# Predicting Mobile Game Success Using Data Analytics

# التنبؤ بنجاح العاب المحمول باستخدام تحليل البيانات

**by**

**KHALED MOHAMMAD ALOMARI**

**A thesis submitted in fulfilment**

**of the requirements for the degree of**

**DOCTOR OF PHILOSOPHY COMPUTER SCIENCE**

**at**

**The British University in Dubai**

**Professor Cornelius Ncube**

**November 2017**

# Predicting Mobile Game Success Using Data Analytics

التنبؤ بنجاح العاب المحمول باستخدام تحليل البيانات

**By**

**KHALED MOHAMMAD ALOMARI**

**A thesis submitted to the Faculty of Engineering & Information Technology in fulfilment of the requirements for the degree of**
**DOCTOR OF PHILOSOPHY COMPUTER SCIENCE**

**at**

**The British University in Dubai**

**November 2017**

**Thesis Supervisor**
**Professor Cornelius Ncube**

**Approved for award:**

| | |
|---|---|
| ______________________________ | __________________________ |
| Name<br>Designation | Name<br>Designation |
| ____________________________ | __________________________ |
| Name<br>Designation | Name<br>Designation |

Date: _____________

# DECLARATION

I warrant that the content of this research is the direct result of my own work and that any use made in it of published or unpublished copyright material falls within the limits permitted by international copyright conventions.

I understand that a copy of my research will be deposited in the University Library for permanent retention.

I hereby agree that the material mentioned above for which I am author and copyright holder may be copied and distributed by The British University in Dubai for the purposes of research, private study or education and that The British University in Dubai may recover from purchasers the costs incurred in such copying and distribution, where appropriate.

I understand that The British University in Dubai may make a digital copy available in the institutional repository.

I understand that I may apply to the University to retain the right to withhold or to restrict access to my thesis for a period which shall not normally exceed four calendar years from the congregation at which the degree is conferred, the length of the period to be specified in the application, together with the precise reasons for making that application.

Signature of the student

# COPYRIGHT AND INFORMATION TO USERS

# ABSTRACT (in English)

Since the advent of arcade games and the development of the Wireless Application Protocol (WAP) at the close of the millennium, the mobile game app industry has exploded; and subsequently has transformed the ideologies of mobile technology and software developers to forward thinking within the dimension of innovative mobile game development. After the first decade of the new millennium has passed, and even though billions of dollars in revenue have been realized from mobile game apps, there is still a gap in literature with regard to mobile game user behavior and methodologies for predicting the likely success of mobile game apps during the development phase. Game features and ARM strategies are analyzed and discussed as primary drivers of mobile game app success.

This study addresses these challenges through data driven research of the mobile gaming application market, mobile gaming application features, user acquisition and retention trends, and monetization strategies using the CRISP-DM model for data mining in order to prove a successful method for predictions of mobile game application success. The attainment of the prediction of one mobile game app from a sample of 50 was accomplished by running a batch prediction for the game features dataset, and a separate batch prediction for the user behavior dataset. The lists were then integrated, a final list of games which appeared in both lists was generated for further comparison. According to the prediction model results for the dual datasets, the most successful mobile game app from the 50 game sample was *Game of War-Fire Age*; the most successful genre was Puzzles, and the most successful developer was EA Sports. Where success is described based on the best match with the results of the study. The most successful game predictions were extracted and compared to the predominating user behaviors for further analysis and implications. Significant outcomes for the comparisons included the predominance of the Social Networking features, Offers, and IAP 90% to 100% of the sample. A model of mobile game app success prediction based upon the game features values that are created proposed.



*Key Terms*: mobile game application, monetization, location-based mobile apps, predictive analysis, CRISP-DM

# ABSTRACT (in Arabic)


منذ ظهور ألعاب الورق وتطوير بروتوكول التطبيقات اللاسلكية (WAP) في نهاية الألفية، انفجرت صناعة تطبيقات لعبة المحمول. بعد ذلك بدأت أيديولوجيات تكنولوجيا الهاتف النقال ومطوري البرمجيات إلى التفكير في تطوير العاب المحمول المبتكرة. بعد مرور العقد الأول من الألفية الجديدة، وجني عشرات مليارات الدولارات التي تحققت من عائدات العاب المحمول، لا تزال هناك فجوة في الأدب فيما يتعلق سلوكيات لعبة المحمول والمنهجيات الكمية للتنبؤ بنجاح لعبة المحمول. وقد تم تحليل مزايت الالعاب واستراتيجيات (ARM) باعتبارها المحركات الرئيسية للنجاح لعبة المحمول.

وتتناول هذه الدراسة هذه التحديات من خلال البحوث القائمة على البيانات من سوق تطبيقات الألعاب المحمولة وميزات الألعاب المحمولة واستراتيجيات اكتساب المستخدم والمحافظة عليه وتحقيق الدخل باستخدام نموذج (CRISP-DM) لاستخراج البيانات من أجل إثبات طريقة ناجحة للتنبؤ بنجاح لعبة المحمول. وذلك خلال التحقيق في عينة من 50 لعبة محمول باستخدام تشغيل التنبؤ دفعة لمجموعة لكل من ميزات اللعبة وسلوك المستخدم كلا على حدا. ثم دمج القوائم، تم إنشاء قائمة نهائية من الألعاب التي ظهرت في كل من القوائم لمزيد من المقارنة. وفقا لنتائج نموذج التنبؤ لمجموعات البيانات المزدوجة كان الأكثر نجاحا من العينة لعبة "*Game of War-Fire Age*" ؛ أنجح فئة كان الألغاز "Puzzles"، وكان المطور الأكثر نجاحا "EA Sports ". وتم استخراج أكثر التنبؤات نجاحا للعبة ومقارنتها مع سلوكيات المستخدم السائدة لمزيد من التحليل والآثار. وشملت بعض النتائج الهامة للمقارنات غلبة ميزات الشبكات الاجتماعية، والعروض، بنسبة 90٪ إلى 100٪ من العينة. واقتراح نموذج للتنبؤ بنجاح لعبة المحمول على أساس قيم ميزات اللعبة التي تم دراستها.

## ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor Prof. Cornelius Ncube for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my advisor, I would like to thank the second supervisor of my thesis: Prof. Khaled Shaalan, for the insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

I thank my colleagues at the British University in Dubai for the stimulating discussions, for the sleepless nights we were working together before deadlines. Also, I thank my colleagues at Abu Dhabi University for the stimulating discussions and support.

Last but not the least, I would like to thank my family: my parents and to my brothers and sister for supporting me spiritually throughout writing this thesis and my life in general. And especially to my wife for understanding my situation in this stage. Also, I thank my children and apologize to them for not giving them enough time during my studies.

Thanks for all your encouragement!

# Table of Contents

## LIST OF TABLES

## LIST OF FIGURES

# LIST OF ABBREVIATIONS

**Table 1. List of Abbreviations**

| Abbreviation | Definition |
|---|---|
| API | Application Program Interface |
| ARG | Alternate Reality Game |
| ARM | Acquisition, Retention and Monetization Funnel |
| ARPPU | Average Revenue per Paying Customer |
| ARPU | Average Revenue Per User |
| CAC | Customer Acquisition Cost |
| CPA | Cost per action |
| CPC | Cost per click |
| CPI | Cost per install |
| CPM | Cost per 1000 impressions |
| CRISP-DM | Cross Industry Standard Process for Data Mining |
| DGBL | Digital Game-Based Learning |
| DMN | Decision Model Notation |
| F2P | Free-to-Play |
| FPS | First Person Shooter game genre |
| GBL | Game-based Learning |
| GPS | Global Positioning System |
| HUD | Heads-up Display |
| iOS | Information Operating System |
| J2ME | Java 2 Platform MicroEdition |
| KPI | Key Performance Indicators |
| LCD | Liquid Crystal Display |
| LTV | Customer Lifetime Value |
| NPC | Non-player characteristic |
| OpenGL ES | Embedded Accelerated 3D Graphic Standard |
| PDA | Personal Digital Assistant |
| P2P | Pay-to-Play |
| ROI | Return on Investment |
| U&G | Uses and Gratification |
| UI/UX | User Interface/User Experience |
| WAP | Wireless Application Protocol |
| DAU | Daily Active Users |
| MAU | Monthly Active Users |

## Publication

- Alomari K.M., Soomro T.R., Shaalan K. (2016) Mobile Gaming Trends and Revenue Models. In: Fujita H., Ali M., Selamat A., Sasaki J., Kurematsu M. (eds) Trends in Applied Knowledge-Based Systems and Data Science. IEA/AIE 2016. Lecture Notes in Computer Science, vol 9799. Springer, Cham. **Related (Chapter 3 and 4)**

- Alomari K.M., ElSherif H.M., Shaalan K. (2017) Arabic Tweets Sentimental Analysis Using Machine Learning. In: Benferhat S., Tabia K., Ali M. (eds) Advances in Artificial Intelligence: From Theory to Practice. IEA/AIE 2017. Lecture Notes in Computer Science, vol 10350. Springer, Cham. **Unrelated**

- ElSherif H.M., Alomari K.M., Al Haddad A. S., Alkatheeri A.O. (2016). Mobile government services satisfaction and usage analysis: UAE government smart services case study, International Journal of Computer Science and Mobile Computing, 5(3), 291-302. **Unrelated**

# Chapter 1. INTRODUCTION

The billion-dollar mobile gaming app market began with the miniaturization of traditional video games into functionalities for handheld devices and the advent of "time waster" games (Mayra, 2015). The union of arcade gaming, television gaming consoles, digital computing and modern art produced the likes of *Snake*, *Tetris*, and a diversity of Wireless Application Protocol (WAP) games (LoWood, 2009; Wright, 2016; MADAB, 2016). Since the onset of the new markets for e-sports, and other mobile gaming application genres, the rate of increase in mobile gaming software development, digital native downloading and usage trends, and mobile game app revenues has been exponential (Mayra, 2015; Alomari, Soomro & Shaalan, 2016). Research of these increases has been analyzed in some disciplines according to data mining approaches to statistical analytics in order to create snapshots of user data leveraging, automation, and relevant insights[1]. However, as a relatively new market, a limited number of guides, innovative software development models, and analytics for effective acquisition, retention and monetization strategies pertaining to the mobile gaming app market have been published. This study will address these challenges through data driven research of the mobile gaming application market, mobile gaming application features, user acquisition and retention trends, and monetization strategies using the CRISP-DM model for data mining in order to prove a successful method for predictions of mobile game application success.

[1] See Firebase Analytics and Google Analytics as discussed by Mobile App Developers' Advisory Board in 2016 publication.

## 1.1 Problem Statement

The mobile gaming industry offers multifaceted challenges to mobile game developers. Making sound decisions regarding the features that will impact a game's success is very tricky (Feijoo, et al., 2012). These are decisions that developers often have to make before starting the development process. In most cases, developers start by conceptualizing viable game concepts. To validate the acquisition and adoption processes of their games the only option left to them is a launch to market strategy that will test on the how the game will perform. Despite the high risk that is incurred by such a methodology game design features are often skewed. In some situations, developers end up adding many unnecessary features that could potentially compromise the computational performance of the games. The end result is that users find themselves in situations where they download game application where they only end up using only one of the many features provided by the developers. The situation is manifested in mobile applications that end up getting numerous downloads followed by a very high number of uninstalls.

Getting the right user acquisition strategy is usually a challenge. This is a factor that has a direct impact on the revenue potential of a game. Developers are often torn apart between which feature would be more lucrative to make the users stick around to the point where they can start monetizing the games.

The bigger debate is usually whether to go the free to play way or the paid pricing model. More often game developers make blind decisions made on intuition or defined business models that have no relationship to the features of the game. The impact is that acquisition becomes a cumbersome task. Some game developers often choose to spend a lot of money in creating a buzz that could potently bring in a viral loop solving the customer acquisition problem. However, the next step of retention is not straight forward, if the features are not spot on many games often end up being trashed or having very discouraging usage statistics.

On high-level mobile trends are very unpredictable. This is because there are so many consideration factors that cannot simply be determined at first sight. User interface design, sound and multimedia integration, target audience, character modelling and many other components that determine the game success are not very straight forward. This challenge manifests itself through mobile games that consistently receive bad reviews on application stores.

The game market often targets a very vast market. The challenge is that there are many target audiences with different perspectives, cultures, and ideologies. The target markets of many games often span across geographical boundaries. Some games are known to achieve global success while others do fairly well in specific regions and very poorly in others. Theming, as well as internationalization of games, has always been a tricky subject. Although some games normally target specific market segments defined by either age brackets or economic settings the success background usually have a common ground.

It is evident that winners of the mobile game market continue to win while looser are consistently losing in numbers. There is no clear definition of market success as this could be a variable of number of downloads, daily active usage, purchases and revenue growth. All these factors have a significant impact on evaluating the success of games. Quality and standards are important factors that should drive every industry. It is key for all mobile game stakeholders to understand the guidelines to offer quality products.

Statistics are key in every financial market. Before sizable investments can be put into mobile game ventures, it is important for the investment return potential to be determined. The games market makes it very difficult for investors to quantify beforehand the valuation of mobile games. It is also very difficult to determine the amount of investment that can be channeled into various aspects of the applications lifecycle. Determining the right customer acquisition cost and the relevant programming costs to ensure customer retention is very important (Feijoo, et al., 2012). Predicting game success is, therefore, a very important aspect in ensuring business continuity that is lacking in the present day market.

### 1.2 Significance of the Study

The application plays a critical role in the overall mobile device user experience (Google, 2015). Still, the disciplines and models for mobile game app design, development, commercialization and analysis are in infancy stages. A proposed set of protocols and guidelines for mobile game apps is lacking. Several researchers are in consensus that a gap exists in the literature in regard to mobile game app development and for analytic data models that accurately measure the potential success or failure of the mobile game apps in the commercial markets (Bohmer et al 2011; Mayra, 2015). Despite the revenue boom the mobile development game industry it is faced with numerous challenges. The field is a relatively new, and most stakeholders are experimenting with various approaches that in some cases end up being disastrous. There is limited publication of guidelines, best practices and success pathways. This study will show how a diversity of factors and social and cognitive theories combine to impact the lifecycle and the success or failure of the mobile game app. Key contributions to the field of computer science also include a proposed analytical model for mobile game apps which will highlight deficiencies in feature provision and enhancement functionalities. Further, the study will propose computation strategies that will enhance the performance of the application on the mobile device.

The mobile game app market is extremely competitive, with high demands for success and sustainability. The information, knowledge and method presented in this study will serve to encourage newcomers to the field of mobile game app development that have become discouraged by failed projects. The study may be used as an important resource in the development of educational curriculums that accommodate protocols which lead to the development of productive game development syllabi that can be implemented by learning and training institutions. In addition, this research may contribute to the development of new fields of research that will profoundly impact the digital gaming industry through an analysis of current market trends and forecasting of future market behaviors.

Like any other venture, this is a serious business endeavor. Developers are faced with a great dilemma choosing the right monetization strategy. The consequences of some choices are significant wastage of games that had massive potential.

Data is key to making every business decision. Blind application designs have cost companies a significant amount of money. The industry already boasts of successful achievers that have continuously won the market share in terms of user acquisition as well as revenue generation. It is clear that there is an imbalance with cases of some seemingly well-designed games getting no usage at all. The key to stabilizing the industry is providing clear models that can be relied upon in setting standards.

Device manufacturers are also innovatively coming up with features that will better enhance customer experiences and make mobile platforms more useful (pwc, 2012). The study will provide a basis for analyzing feature provisions and enhancements as functionalities that are currently lacking. The research will also give a basis for determination of appropriate computation abilities that can help enhance mobile devices and even potentially serve as groundbreaking paths to improving the computational abilities of mobile devices.

Newcomers in the field are also often discouraged by failures at early stages. It is very difficult for them to determine the right way of making money out of their venture. Good applications are monetized wrongly causing outright failure. The situation can only be corrected by determination of the success pathway.

Traditionally there has not been any right or wrong way of designing mobile games. This decision is often left to the creativity of the developers. Also, the target market is often wide with varying market structures defined by geographical and economic boundaries. Applications often have the ability to span the global market, in this case, if the business models of applications that have managed to achieve remarkable market penetrations are adopted their revenues are definitely bound to increase.

Desktop devices and other gaming consoles are also slowly being faced out as the adoption of mobile devices grows stronger (Bonnington, 2016). If mobile games are not well tapped and the potential fully utilized the gaming industry could be facing a huge drawback that could potentially affect the overall computer industry.

The largest companies in the tech industry have ventured into the mobile industry. Google is heavily involved in the mobile industry through the Android operating system. Microsoft which is a computer technology giant has its share fair of the market through the windows phone platform while Apple supports the iOS platform. Revenue from mobile gaming is, therefore, an important part of the income that drives the industry as it helps support top technology companies that have invested heavily in the mobile technology.

### 1.3 Conceptual Framework

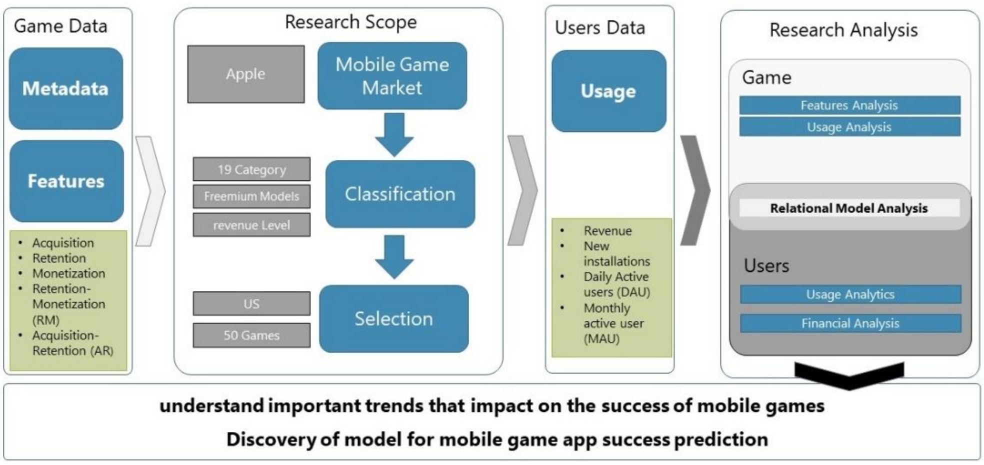



**Figure 1. Conceptual Framework**

Active Users are a primary aspect of the data extracted. Figures regarding active usage provide information on feature usage; determine the impression for given game has on users; and serves as a detailed data model for the determination of the most successful of games based upon performance metrics on a superficial level. The metadata will help to form a basis of analysis, and a number of data fields that are not directly connected to game attributes will be collected.

### 1.4 Research Aims and Objectives

This research presents a multi-faceted approach to statistical analysis and data analytics techniques to determine various mobile gaming application features which significantly and measurably impact the success of a game. More specifically, the aims and the objectives for this study are as follows:

- To identify current data analytics models that are used to analyze mobile game data.
- To perform detailed statistical analysis of existing games to understand features that lead to failure or success of mobile app games.
- To understand mobile gaming trends that shape the growth of the game market.
- To investigate features and game attributes as well as monetization structures that impact the success levels of mobile app games.
- To develop, test and validate a prediction model that can be used to evaluate the market value of games.
- To develop standard evaluation procedures for valuing mobile game apps.

### 1.5 Research Questions

The research questions that the outcomes of this study will answer are as follows:

- **$RQ_1$**: What are the current data analytics models that are used to analyze mobile game app data?
- **$RQ_2$**: What are the main features of mobile game apps that lead to failure or success of the games?
- **$RQ_3$**: What are the mobile gaming trends that shape the growth of the games market?
- **$RQ_4$**: What are best practices to develop standard evaluation procedures for valuing mobile game apps?
- **$RQ_5$**: How can a prediction model be used to evaluate the market value of mobile game apps to be developed, tested, and validated?

### 1.6 Organization of Dissertation

**The reminder on the thesis is organized as follows:**

**Chapter 2** provides a review of literature that encompasses computational aspects of computer science, prior studies of mobile game application success indicators, factors, and associated metrics; fundamental theory in regard to predictive information, prior research of data mining and processing methods for big data, game success prediction models; mobile game app classification and evaluation; user behavior prediction, and an overview of the CRISP-DM model that will be used in the methodology.

**Chapter 3** presents the methodology for the quantitative analysis of mobile gaming app data, to include descriptions and examples of the game categories and features, the framework model used to perform the dual analysis using the BigML statistical analysis tools. A data mining application is developed for fetching and filtering data. This application is a web based data extraction and analysis service. A description of the data, the process of data management and analysis, and the visuals for the quantitative analysis are presented.

**Chapter 4** presents the data analysis for genres and developers, the game features and user behavior, and associations generated by the prediction model. Decision trees for the dual datasets are generated, and a revenue comparison the 50 titles is conducted.

**Chapter 5** presents the results of the data analysis and a discussion of the results. This section will also discuss challenges to data driven research of the mobile gaming application market, relevant user acquisition and retention trends, and monetization strategies that have been deemed as successful using the CRISP-DM model for mining.

**Chapter 6** presents a model of mobile game app success prediction based upon the market values that are created is proposed. An actionable model for mobile game app success prediction is presented.

**Chapter 7** presents the conclusions drawn from the outcomes of the study and this section will also will recommend directions for future research.

# Chapter 2. REVIEW OF LITERATURE

The scope of Theoretical Computer Science (TCS) extends to include the rapidly evolving subdivision of computational science which encompasses computational process models, the algorithm, and associated application oriented analyses. Bharathi et al (2016) presented that a gap exists in the literature in regard to the existence of gaming design features and the degree to which the application features impact mobile game app success or failure. Several studies have been published in regard to time-series extrapolation and practice algorithms as models for future prediction, but not necessarily for the interests of mobile game applications. Sifa et al (2015) contributed that due to the limited number of publications for prediction analysis and qualitative research of commercial games, the evidence-based models and theory from closely associated domains have been used to postulate mobile game prediction theory. In this light, the theme or focus for this review centers upon the tenets of predictive information and data mining, software development theory, predictive models as tools to produce metrics that may be used as success indicators, and market valuations. Most specifically, this review will present fundamental theory in regard to prior research of data mining and analytical methods for big data, game success prediction models; mobile game app classification and evaluation; user behavior prediction, and an overview of the CRISP-DM model that will be used in the method.

## 2.1 Background

The mobile gaming platforms hold potential to generate millions in revenue through commercialization, marketing agility, and sufficient backing for the mobile app brand (Clark, 2012; MADAB, 2016). In a space of three months, *Space Ape Games* evolved from zero downloads to 5 million downloads, generating millions in revenue (MADAB, 2016). A significant driver of the continual success and growth of the mobile gaming application market has been the evolution of data art as a discipline of electronic media, location-based applications, and the progressive development of tools for software and mobile device hardware engineering (Grugier, 2016).

### 2.1.1 The Evolution of the Mobile Game Market

The market for mobile games began during the age of PDAs and bulky cellular phones of the 1990s, during time when mobile communication technologies advanced from merely telephony utilities to include gaming applications (Strain, 2015). The mobile app was designed as a high level interaction between system frameworks and custom code (Apple, 2017a). In 1994, the *Tetris* game was featured on the Hagenuk MT-2000 mobile phone (Beckett, 2016). The arcade game *Snake*, which was developed by Taneli Armanto in 1982, was transformed into a mobile game app in 1997 by Nokia (Osborn, 2014). 1997 was a golden year for mobile game apps, in that the WAP forum was formed, and released the first browser for the mobile device (Wright, 2016). The commercialization of mobile games can be traced back to the introduction of the Japanese NTT DoCoMo iMode platform in 1999 (Beckett, 2016). Further, the success of *Tetris* and *Snake* was followed by the entrance of Gameloft into the market, which produced its first round of mobile games in 1999.

The DoCoMo iMode standards of Compact HTML simplified the process for developers to produce new content (Tercek, 2007). The potential to generate profits solicited inputs from both professional and amateur developers. By the new millennium, the global market race for mobile game app development was well underway. Wright (2016) described the new millennium mobile game market as young and fragile.

However, decreasing costs for mobile communication devices, memory, and batteries along with an infinite number of potential platforms for mobile game app content increased the popularity, demand, and potential profits for mobile gaming. As the distribution costs approached 0, the number of games in the app stores continued to skyrocket (Moreira et al 2014). JAMDAT mobile entered the market at the open of the new millennium and acquired a considerable share of Verizon Wireless' game deck (Vankka, 2014).

At the 2001 Java One Conference, the J2ME language was introduced, along with languages such as Qualcomm's Binary Runtime Environment for Wireless (BREW); all of which significantly increased the quality of the mobile games (Osborn, 2014). The quality of mobile game apps and the costs of mobile game development were also significantly increased by the competitive market pressures of 2004, which resulted from the carrier game service barriers to new product entry (Vankka, 2014). Subsequently, mobile game titles began to be named after sports and entertainment industries characters, such as *Lord of the Rings*, *Tiger Woods Golf*, and *Madden NFL*. In 2003, Nokia pioneered the market once again with the launch of the controversial N-Gage, an odd shaped smartphone and mobile game system that was designed to compete with the Game Boy series (Beckett, 2016).

Over a few years, the market for mobile games, as well as the community of mobile app developers, became fragmented by a diversity of tools, environments, and languages (Clark, 2012). Network infrastructures matured to some degree and the design of mobile device hardware had begun to undergo rapid changes. The mobile game app industry realized a significant boost from monetization in strong markets, and market analysts began to forecast substantial future market growth (Newzoo, 2016). In 2006, Motorola introduced the first series of Razrs, and sold more than 60 million of the mobile devices (Osborn, 2014). A 2011 study showed that a sample of 4,000 android users accessed a mobile application as much as 50 times in one day and spend more than one hour using the application (Mayra, 2015).

The history of the mobile game app reflects contributions from a diversity of developers and mobile communication technology producers. Nonetheless, the final predominating catalyst for the mobile gaming industry came from Apple, when Steve Jobs presented the iPhone as an innovative canvas and technologically adaptable platform for mobile gaming software (Beckett, 2016). The smartphone platform for mobile gaming produced intense competition for the pc and gaming console markets, which also reflected tremendous growth. On one end of the spectrum, the content for mobile gaming took the form of entertainment, educational or learning tools and medical applications while on the other end of the spectrum it took the form of content that was controversial, violent, and adult rated designed purely for entertainment. Alomari et al (2016) found that in 2014, the Apple App Store contained over 850,000 apps and the Google Play Store contained more than 700,000 apps.

### 2.1.2 The Evolution of Data Art and Mobile Game Apps

The digital native demographic, which is the dominant demographic of mobile game app users has continuously expected more, and is willing to pay for more for graphic intensive and data intensive features and customization options for visual applications across all entertainment portal platforms. Therefore, gaming software developers are being continuously challenged to create dynamic digital screen visualizations, as well as a diversity of high functionalities that will impress, and thus, engage the digital native (Clark, 2012). Grugier (2016) argued that data art enables a transformation of abstract, “cold” data into profound, metaphoric beauty within the UI/UX designs which delight the user. In this light, the mobile games for Android, iOS, and Windows Phone with “eye candy graphics” and energizing audios have been the preferred choice among digital native gamers who are willing to pay to play (Wood, 2014).

The intensity of the global market competition for winning mobile game apps has surrounded the creation of data intensive features with high-end visualizations within mobile game applications. Up to 2002, the mobile game app design remained restricted to one dimension and black and white graphics. Then, since 2002, the mobile devices from global producers such as Samsung were released with larger, higher resolution, LCD screen displays, full color and ring tones (Cooper, 2013). All the new design features served as strong precursors to multimedia. In 2003, the mobile game app designs evolved to 2D, color isometric views (Vankka, 2014). To some degree, the functionalities within the mobile app also evolved from simplistic movements to more complex commands and functions. By 2004, the platform capabilities for the visual presentation of mobile game apps had advanced to color, 3D graphics; and by 2005, the visuals consisted of color, 3D, and OpenGL ES. The mobile app market competition has been further compounded by the introduction of the touch screen feature in mobile devices, which created yet another technological revolution for mobile game developers.

The evolution of data art and the functionality of mobile apps can be visualized through a juxtaposition of the *Snake II* game which was embedded into the Nokia 6610. *Snake II* is characterized by simplistic monochrome graphics and black font texts. The user had the options of changing the *Snake II* background color or pausing the game; however, the snake could only move forward in continuous motion. A retro version of *Snake II*, or *Snake 2k,* is available in the Apple App Store for $0.99 and requires 7.5 MB to download on iOS 7.0 or later (iTunes, 2016). The *Asphalt 8: Airbone* mobile gaming app features licensed music and 47 automobiles, all complete with barrel roll, stunt, and 360º jump functionalities. Further, the *Asphalt 8: Airbone* may also be downloaded for $0.99 on Windows Phone and iOS, and for free on Android; however, it requires 819 MB of space for download, as well as additional space required for the new install (Rubino, 2013).

### 2.1.3 The Evolution of Locative Art and Mobile Game Apps

At the beginning of the new millennium, location-based mobile games emerged along with the movement for “locative art”, which capitalized on the increasing ubiquitousness of GPS tracking and location-based media in public spaces (Leorke, 2014, p. 133). The development of the location-based mobile game app features played a critical role in mobile game app user content data acquisition and geographic data retrieval that was not directly volunteered by the user (Winter et al 2011). The objective of the technology art movement was the integration of digital, interactive art and locative media that had been previously limited to consumer and military applications to within the scope of functionalities for the mobile communication device. Further, the dimension of the virtual environment for mobile games that had been traditionally restricted in scope with regard to space and time were expanded in dimension to include the cityscape, local neighborhoods, urban events and performances in interactive play.

During the locative art movement, at the close of the 20th century through the open of the 21st century, the privacy and virtualization of the public space increased in relevance. Pioneering locative artists, to include Cardiff (1999); Lozano-Hemmer (1999); and HuJeBek (2002) introduced the concepts of interactive art images, texts, and sound to the urban space[2]. Media consumption patterns evolved concurrently with spread of WiFi, GPS capabilities, smart devices, handheld gaming, and digital videos (Leorke, 2014). In turn, mobile media technology also expanded beyond the physical demographic to include the virtual, urban public space as location-based mobile game apps enabled location sharing as well as communication between users.

As the location-based mobile game apps began to generate massive amounts of user data, market analysts began to use simple data collection and human knowledge acquisition approaches to semantic interpretations and management of large sets of data. The Red Robot series for *Life is Crime* and *Life is Magic* mobile game apps, released in 2012 and 2013 respectively, incorporated virtual elements of location-based city landscapes, neighborhood gangs and crime, and fantasy. In this light, the simple data collection would encompass the “citizen position” as a sensor; while the human knowledge acquisition approach would define the citizen position in the context of knowledge of the location (Winter et al 2011, p. 1). Alternate Reality Games (ARGs), such as *The Dark Knight* ARG, emerged as real world storylines with virtual experiences which form user interactions through the website, telecommunications, and text messages. From this point, the mobile game app developers needed a systematic plan for collecting market data that would provide valuable insight in regard to premium user demographics.

---

[2] See J. Cadliffe (1999) *Trace*; R. Lozano-Hemmer (1999) *Vectorial Elevation*; and HuJebek (2002) *dot.walk.*

## 2.2 Mobile Game Apps - Acquisition, Monetization and Retention (ARM)

The contemporary market for mobile game apps have progressively transitioned from youth to youth and adult; and from premium game monetization strategies to freemium models. In 2013, the developers of *Candy Crush* announced over half a billion downloads and installations to mobile devices and Facebook in the first year of its release (Madigan, 2015). Released in 2012 by King.com, Candy Crush is structured as a 'match-three puzzle' social game which outpaced *Angry Birds* in sales. Nonetheless, the commercial reality for the average mobile game app developer is a confrontation with the challenges of game balance and pricing (Fields, 2014). In turn, mobile game apps for Android have historically been challenged with payment issues and low averages for revenue per user in juxtaposition to games made for iOS and Windows Phone platforms (Vidyarthi, 2012).

The primary stakeholders of the mobile gaming app industry include the consumer, app developers, App Stores Operators, investors, and marketers. Researchers and market analysts have sought to gain illustrative insights of the mobile application markets through extensive investigations of the perspectives of the stakeholders. In order to analyze the factors or features which drive optimal monetization of mobile apps, Kontagent introduced the Acquisition, Monetization and Retention (ARM) funnel[3]. The ARM model has been used across industries to create data driven business infrastructure designs which dictate how to effectively allocate the business resources and investments. Kotangent President, Josh Williams asserted that in both the initial stages of mobile app development and for apps that have been released and accumulated

[3] See Kontangent President, Josh Williams' comments in N. Vidyarthi 2012. *How Lessons Learned in Social Will Give You a Head Start in Mobile.*

a large user base, mobile game app marketers and developers may benefit significantly from insights of social app strategies to maximize ROIs and the customer lifetime value (LTV) (Vidyarthi, 2012).

### 2.2.1 Mobile App Acquisition Objective

Acquisition endorses analytics which use perpetual flows of data in order to target the user demographics with the most potential, and to identify the most cost effective methods of user engagement through advertising and pricing. The key performance indicators (KPIs) for user acquisition include Customer Acquisition Costs (CACs); Average Revenue per User (ARPU); Retention based upon Channels and Campaigns; and the Acquisition Conversion Funnel (Vidyarthi, 2012). *Gamevil* President, Kyu Lee contributed that prior to the iPhone era, the strength of game developer publications relied upon the device coverage and distribution channels; but that after the iPhone, the strengths relied upon the quality of the game titles (Osborn, 2014). Nonetheless, a large degree of the mobile game app success lies in the successful marketing of the product, as well as in the product design and scope.

Efficient marketing strategies draw new users from pools that have been targeted for frequent purchases and downloads, and long periods of app use. Madigan (2015) supported that gaming is becoming increasingly social and many games provide shared experiences in order to stretch the dimensions of gaming to fit popular culture. A 2014 study showed that users spent approximately 30 hours each month using mobile apps; and approximately 46% of the apps are either games or games related (Nielsen, 2014). Another study showed that the average user has approximately 36 mobile apps installed on their mobile device (Google, 2015). Potential opportunities of the acquisition component in the capacity to reach massive audiences at low cost; the identification of differentiations between Organic, Marketing, and Viral sources; and the more intense game design (Vidyarthi, 2012).

The ecosystem for mobile apps consists of the marketers, publishers, and ad networks, all of which create revenues within the mobile and the app site (Instal, 2014). The marketers consist of the game developers and the companies that promote brands associated with the mobile device, telecommunications, or mobile applications through banners and links to the publisher's website. The ideal methods of promotion must promote the app and also provide a pathway to measure the performance of the app on the market (Instal, 2014). The mobile game app is typically promoted through either discovery apps or ad networks.

The discovery app is used to promote 'daily deals' and to search for the premium apps within an 'app addicted community' across a diversity of sources (Instal, 2014). The developer may achieve high ratings among retailers through the daily deals and associated burst campaigns. The mobile ad network may be used to promote the mobile game app through traffic sales based upon cost-per-click (CPC), impression (CPM), action (CPA), or install (CPI) based commissions (Instal, 2014). The cost-per-impression (CPM) is dependent upon CTR, which identifies the number of users who view the ad, click it, and the number of users who have already clicked it, downloaded it, and installed it. Of the 3 models, the CPI is highly effective and presents the least amount of risk; as the traffic typically comes from discovery apps and the costs are limited to only authentic installs, as opposed to impressions or clicks.

### 2.2.2 Mobile App Retention Objective

The Retention component of the ARM is used to extend the period of consumer engagement and to identify the end of the relationship between the consumer and the product. The KPIs for user retention include Day1, Day 7 and Day 30 Retention; Average Session Lengths; the Customer Event Funnel; the K factor; and the User Lifetime Event Distribution. Moreira et al (2014, p. 4) described the retention objective as a focus upon game design that produce mobile game apps that are “sticky” or “addictive” through game mechanics and gamification dynamics. Further, games with social features which endorse friendship between users tend to exhibit higher percentages of user retention.

The download and installation data for a particular game are not sufficient as sole indicators of the mobile game success. Approximately 1 in every 4 applications that are installed on the mobile device are never used, and approximately 38% of the downloaded app are abandoned or uninstalled immediately after the download[4]. An understanding of the differentiations in user behavior based upon game genres support the development of successful ARM strategies for mobile game apps (Laughlin, 2012). Figure 2 illustrates a loyalty matrix for mobile games based upon a four quadrant Cartesian coordinate system for classifying retention per uses per week for major gaming genres:

---

[4] See Google 2015 study.

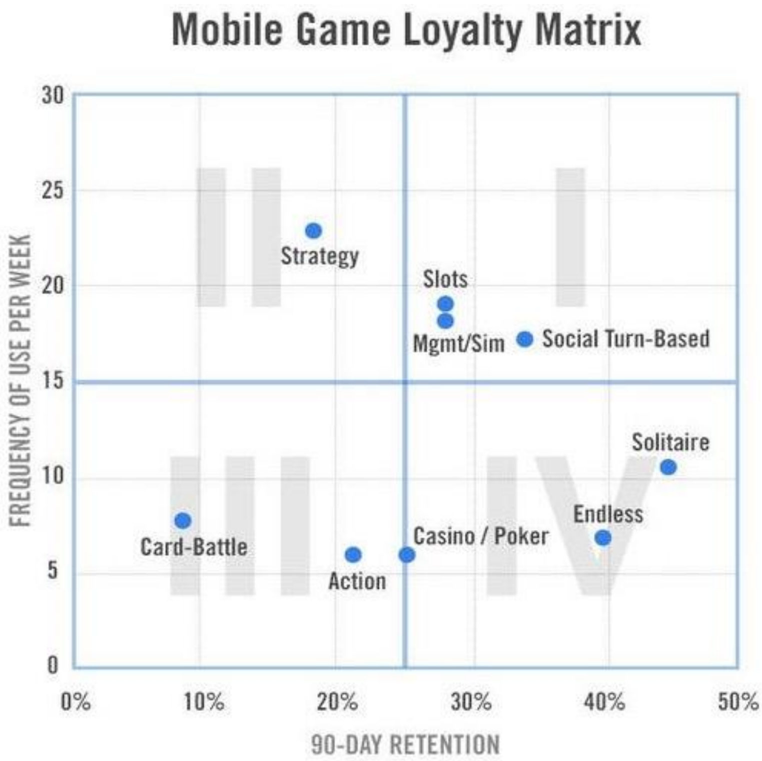



**Figure 2. Mobile Game Matrix**[5]

The matrix categorizes the 9 leading freemium game genres based upon the user frequency against a 90 day retention rate.

The primary reasons for the abandonment of the app are loss of interest, diminished need, diminished usefulness, and the discovery of a better app. At the other end of the spectrum, the leading reasons why the user may re-engage with app usage upon receipt of a coupon or discount toward purchase; promises for bonus or 'exclusive' material; friends, family or associate use; and notifications of upgrades or new features, respectively.

During the first years of the new millennium, global mobile game app analysts such as Ovum, Data monitor, ARC Group and Informa projected that pc gamers would transition to mobile game apps, which would be reflected in the global market revenues

---

[5] Extracted from D. Laughlin. 2012. *The gamification of mobile games*. The matrix was designed based upon a sample of 300 participants and free game titles.

(Tercek, 2007). Between 2005 and 2006, the audience for mobile games, the number of downloads, and the revenues did increase, but not at the anticipated rates. Future market growth evolved from the emergence of the gaming ecosystem and global publishers. The sophistication of ad services for mobile game apps increased to become concurrent with the economy for mobile apps (Laughlin, 2012). Potential opportunities of the retention component of the ARM model include maximized control of in-app user experiences; analytics solutions that enable data filtering based upon the version, location, and type of device; and the free-to-play model proliferations (Vidyarthi, 2012). The developer may increase online presence for the title using app discovery and retain the user through promotions of value-added.

### 2.2.3 Mobile App Monetization Objective

The key objective of the monetization strategy is to solicit long term, continuous payments from the user and to maximize the user LTV. At the open of the market, mobile game app operators introduced “a diversity of billing mechanisms” which directly influenced the game design, presentation, and performance” (Tercek, 2007, p. 15). Modern mobile game apps typically encompass the use of both virtual currencies and real world currencies; particularly those based upon gambling mechanisms. Overall, users expect the mobile app to be free and are willing to pay approximately $1.50 for mobile app games.

The key players of the mobile gaming app value chain, to include the talent, content licensors, outsource platform operators, and portal operators, such as Jamba and Zingy, play a significant role in the gaming ecosystem (Tercek, 2007). The KPIs for monetization include the ARPU; Purchase Conversion Event Funnel; Percentage of Paying Customers; and the Average Revenue per Paying User (ARPPU) (Vidyarthi, 2012). Further, the mobile game app may be monetized through paid downloads, in-app advertisements, and through in-app purchases. The mobile ad network model is ideal for mobile app monetization, in that the approach saves time, provides performance metrics in real time, and provides support throughout the lifecycle of the app (Instal, 2014). The publisher may benefit from the CPI model for monetization and by partnering in contracts with high fill rates, rapid, accurate payment systems, and sufficient support.

The potential benefits of the monetization component include optimization of the mobile app mechanics; tracking of revenue channels; and messaging strategies that encourage the user to stay connected (Vidyarthi, 2012). “Social gaming” has been used as a monetization strategy by game developers, as sociability, or the capacity to facilitate social interactions, has been confirmed as an effective approach to revenue generation (Cheng, 2014). Further, Fields (2014, p. 179) supported that the more sizable the degree to which the game permits the user to “deviate away from standard in-game performance curves, the higher the cost of the game should be”. In 2013, the owners of Supercell reported $2.4 million per day in revenue from approximately 9 million players per day for *Hay Day* and *Clash of the Titans* (Strauss, 2013). However, Moreira et al (2014) asserted that monetization approaches for Pay-to-Play games do not necessarily transfer to Free-to-Play games.

### 2.3 The Psychology of Mobile Game App Success

The mobile game app market is, in itself, a game of hit-or-miss, within which the developer spends significant amounts of time and effort in the development of games that may or may not be accepted as worthy of the required time or money by the intended audiences. James (2016) described the mobile gaming industry as a “gigantic, unregulated beast” which utilizes psychology and gambling methods to create addictive experiences, irrespective of any artistic merit. The complexity of the mobile gaming market is further compounded by the consideration that the trends of what is popular in gaming, with the exception of “free and fun”, change constantly. Here, the developer must design an activity that is stimulating based upon “universal truths” in regard to the human cognitive processes (Madigan, 2016).

Griffiths (1995, p. 14) described the digital gamer addict in respect to the salient component of addiction, in which a “particular activity progressively becomes the singly most important activity in the individual’s life, thinking, emotions, and behavior” where each moment is spent “engaged in the activity or thinking of the next time he or she will be”. The most addictive mobile gaming apps for iOS and Android include *Pokemon Go; Candy Crush Saga*, *Clash of Clans*, *Minecraft, and Angry Birds* (Life’d, 2017). The prediction of the success or failure of such games is typically approached by methods of data analysis which use market variables, such as the user rating, number of downloads and installs, and generated revenues for a specified period of time along with descriptive variables for each mobile game. However, the development of successful games is also mutually inclusive of a higher level of understanding in regard to “cyberpsychology”, or the “psychology behind the gamer motivations” for game selection and mobile media purchasing behaviors (Madigan, 2015, p. xvii; Young, 1998). Some speculate that mobile game apps that allow the gamer to ‘reach the next level’ and to collect in-game objects are more prone to develop addictive gaming behaviors (James, 2016). Razi (2015) contributed that gamers with low frustration tolerance levels are more impulsive and are willing to use more money to progress or to win a game.

A significant attribute of successful mobile game design lies in knowledge of the psychology of the gamer as a consumer and in a behavioral sense, and a grasp of the right measures of anonymity or secret identity, deindividualization, non-player character

(NPCs), and decreased real world accountability (Purslow, 2016). Some gamers enjoy gaming genres that emphasize the anonymity of the players, as the gamer may escape to a virtual environment that is familiar, yet outside of the reality of his or her every day. The psychologically based escape from the real world environment and the barrier to access to those around the gamer play a large role in the addictive effect of the virtual world and futuristic games (Jenkins, 2004). The perception of anonymity has been associated with disinhibition, in which the individual feels less inhibited in action and dialogue and dispositional and situational boundaries are perceived as reduced.

Deindividuation has been described as the psychological state in which the user places his or her personal identity within that of a crowd, as opposed to self-expression as an individual identity (Madigan, 2015). The NPCs within the game are controlled by artificial intelligence, and react to the gamer movements based upon predefined codes (Purslow, 2016). The gamer may play games where only one human player is against one or many computer-controlled opponents for long periods of time. Such concepts are critically important within the context of whether mobile game apps are designed for single or group player strategies, and whether the gameplay environment will be a simulation based upon futuristic or a real world social environments.

### 2.3.1 Gamer Psyche and Addiction

A key objective of the mobile game app developer is design a game that will engage a user demographic and that will retain significant percentages of the user’s time, attention, and finances for long periods of time. Cybersociologist's have acknowledged new sources of psychological preoccupation with digital media, technology, and connection to the Internet[6]. Rubin (1994) supported that media use has often functioned as substitutes to personal human interaction by individuals who are immobile, apprehensive and dissatisfied. An addiction to a particular mobile game or genre of mobile games may stem from a ‘technological addiction’, an ‘Internet addiction’, pathological gambling disorder, a social gaming addiction, a general compulsive behavior disorder, or a combination of such addictions (Young, 1998). Moreover, traditional Uses and Gratification (U&G) theorists have supported that the selection of media, as well as the use of the use of the media, is purposeful, goal directed, and driven; that the digital media competes with other vehicles for socializing or communication in order to become selected, attended to, and used to gratify the individual needs or wants (Katz, Blumler & Gurevitch, 1974).

Behavioral theories of self-control have been used to explain addictive gaming behaviors such as the addictive consumption of avatars and avatar props. In this sense, Griffiths (1995) described technological addiction as a behavioral or non-chemical form of addiction to ‘human-machine interaction’. Further, the technological addiction is mutually inclusive of activities which may also be defined as other types of addiction. Lee & Shin (2004) found that avatars, or animated, digital representations of the gamer, have been extremely successful components of the digital gaming market due to psychological benefits of an artificial sense of well-being within the virtual environment and anonymous personalization capabilities. The avatar cyberproducts also fared well within the demographic of individuals addicted to shopping. Lee & Shin (2004) proposed that avatar addictive consumption may be an effective stress reliever for the user.

---

[6] See Lee & Shin (2014) cyberspace study.

An area of growing concern for consumer and behavioral psychologists has been the increase in addictive behaviors toward social genres of mobile gaming. The rise of information sharing, social media and the integration of applications such as FaceBook and *Candy Crush Saga* have changed the face of the market for social mobile gaming apps. Kaye (2016) supported that the gaming market would benefit from a focus on adaptation to cognitive gaming community psychology, social identities, and social norms; and more specifically, upon expanding the gamer's virtual identity. Cheng (2014) explored the association between psychological states, such as leisure boredom, loneliness, gratification perception, and self-control and mobile social gaming addiction. The study encompassed data collection from a sample of 419 participants in China to determine the purchase motives and gratification derived from playing *Candy Crush Saga*. Self-control and loneliness ranked as the highest predictors of addiction; and leisure boredom ranked as the greatest driver of the intensity of mobile social game app use.

An increased use of mobile gaming apps may be attributed to behavioral or situational phenomena other than addiction, such as scenarios in which the individual spends an increased amount of time in commute, or situations where no other options for entertainment or time-wasting are available. Therefore, researchers have also sought to define the differentiation between habitual gamers and gaming addicts. International Gaming Research Unit Director, Mark Griffiths (2013) presented that the boundary between healthy play and a pathological addiction to mobile gaming is defined by the entrance of financial costs as well as the time drainage which produces deficits in academic performance or family interactions. However, others may perceive random anti-social, wreck less, and socially devastating behaviors as ultimate defining outputs of gaming addiction.

### 2.3.2 Gamer Psyche and Violence

On the other end of the spectrum, mobile games that successfully connect with the gamer psyche could have negative affects upon the individual behavior in the real world. For some, the ultimate gaming experience connects with feelings of undeniable defeat, rage, conflict, rejection and depravity; which in turn, contribute to the cultivation of trash talking, bullying, and other forms of undesirable, abusive real world behaviors (Purslow, 2016). American political activist, Ralph Nader described the violent content in digital games as “unprecedented” and coined digital game developers and publishers as “electronic child molesters” (Kerr, 2013). Mobile game apps for 2015 that were ranked as the most “over-the-top-uber-violent” gaming apps for iOS and Android included *Postal*, *Carmageddon*, *Dead Trigger 2*, *Surgeon Simulator*, and *The Walking Dead* (Barraclough, 2015)[7]. The youth demographic may acquire such titles as many parents assume that all digital gaming media is intended for children; and therefore, the parents ignore the violence and maturity ratings when making purchases for their children (Jenkins, 2004).

Researchers have disagreed in regard to whether digital games such as *Mortal Kombat*, *Splatterhouse*, and *Grand Theft Auto* have the capacity to ‘blur’ the social boundaries between good and evil and to invoke catastrophic effects through subsequent violent acts by gamers (Jaslow, 2013). Bartholow (2011) confirmed that violent digital games with apocalyptic themes desensitize the gamer to the violent graphics over the short term; and that games such as *Grand Theft Auto*, *Killzone* and *Hitman* caused the participants to act more aggressively than the participants who played non-violent games. The study in which 70 participants played either violent or non-violent digital

---

[7] Based upon personal picks tallied by Chris Barraclough for Recombu.

games for 25 minutes, and after which measurements were made of the participants' brain activities which supported that the violent content and interaction enhanced the neural desensitization to violence. Bartholow (2011) also presented that the neural desensitization also served as a predictor to increased aggressive behaviors after repeated exposure to the game violence. However, Massachusetts Institute of Technology (MIT) Director, Henry Jenkins (2004) argued that games are "merely games" and that the youth are well aware of the difference between game rules and rules that apply in the real world. Further, Jenkins (2004) supported that the youth who transfers the tragedy in digital game content to real world tragedies are "severely emotionally disturbed".

The outcomes of associated real world legal cases have also conflicted as to the degree to which the high violent content in games causes or contributes to dangerousness or insanity (Lewis & Jolly, 2012). Violent digital gaming was named as the catalyst for the 2011 tragedies, where Adam Lanza targeted his mother and 26 other victims at the Connecticut Sandy Hook Elementary School; while Anders Breivik murdered 77 individuals by bombs and shootings in Norway (Bates & Pow, 2013; Jaslow, 2013). A "trove" of violent video games were found in the basement of the Lanza home, and further research of 20-year-old Adam Lanza's life revealed he had completed 83,000 "kills" with 22,000 headshots during violent gameplay prior to his rampage of at the elementary school (Bates & Pow, 2013). Thirty-three year old Norwegian Anders Breivik had spent approximately 16 hours a day playing *World of Warcraft, Call of Duty,* and *Modern Warfare* and conceded that he desired to kill even more than the 69 victims at a youth political camp on Utoya Island and the 8 victims through a bombing of Oslo government offices. Breivik also conceded that he had planned “to behead Prime Minister Bruntland while reading a text from his mobile phone” (Lewis & Jolly, 2012).

Nonetheless, market demand has dictated that digital games with highly violent content continue to flow in the markets. *Red Faction II* was removed from the German “List of Media Harmful To Young People”, also referred to as THE INDEX, after 15 years of exhaustive efforts to ban the shooter game (Akio, 2017). The first person shooter (FPS) game, available on video game console and mobile app download, was cited by the Entertainment Software Rating Board as inappropriate due to content associated with Armageddon, blood and gore, graphic language, and extreme violence (Space Squirrel, 2011). The German BPjM formed the list in order to filter games with content that enthralls the gamer through the promotion of drug and alcohol abuse, suicide, socialism, discrimination, racism, and extreme violence which endanger the development of self-reliant and socially responsible personalities” (Akio, 2017).

### 2.4 Mobile Game Apps and the Digital Native Demographic

A large percentage of the modern day user demographic for mobile gaming applications is comprised of digital natives of both genders who reside in an urban center (Shuler, 2009; Jenkins, 2004). Problematic attachments to the mobile device and mobile app use have been confirmed as more prevalent in the younger generation, and in the areas of gaming and text messaging[8]. However, Madigan (2015) argued that the traditional stereotypes of the gamer as youth, social rejects, and basement recluses are

[8] See Griffith, 2013.

outdated, and that gaming is becoming increasingly social with games which provide shared experiences in order to stretch the dimensions of gaming to fit popular culture. The “apps culture” is a terminology that also has emerged in order to describe the millennials, or digital natives, who have grown up in advanced technology hardware and software environments and represent a large part of the universal gamer demographic (Purcell, Entner & Henderson, 2010).

The connection between the digital native and the future potential of the mobile gaming market is expressed as the assumed innate technological capabilities of the digital native and characteristics which highlight multitasking, a desire to remain connected, and a demand for advanced technological innovations for both learning and entertainment. Sad & Enber (2017, p. 180) presented that the life of the digital native is “full of products and services that are reminiscent of the digital tsunami”, with “high exposure to interactive multimedia”; and therefore are more conducive to the technology shift than their faculty, who came from a disturbing archaic era of entertainment and education”. Sad & Goktas (2013) also found that within a sample of university students, mobile applications, communication tools, and the Internet ranked as the preferred functions for learning.

The birth of the mobile game app industry was during a time when the society had not yet become adept at using mobile communication device hardware, and when concepts of network infrastructures, Bluetooth capabilities, social media, and location-based applications were in embryotic stages (Strain, 2015). The digital natives, millennials, or the tech savvy generation of texting, Facebook, smart connectivity, and Hot Spots, came after the open of the 21st century; and their parents and grandparents had little knowledge of products born in the digital age. A 2014 study showed that although most smartphone owners were between the ages of 25 and 44 years of age; the highest use of mobile phone apps came from young adults between 18 and 24 years of age (Nielsen, 2014)[9]. For the digital native, the transition from gaming on Xbox and PlayStation to gaming on iPhones and Samsung Galaxies seemed natural; and for some, the transition to mobile gaming apps has been a necessity in order to 'stay connected'.

### 2.4.1 Mobile Game Apps for Mobile Learning

Mobile communication technologies have become a part of the trajectory of national policies which promote educational advancement in order to prepare the digital native for the future workforce (Shuler, 2009). M-Learning, Mobile Learning, or 'handheld learning' has emerged as a subdivision of e-learning and distance education as a learning model design that is based upon the mobile device platform, with 24 hour access across a wide ranged geographic demographic (Mehdipour & Zerehkafi, 2013). The increase in digital mobile integrations that highlight software, smartphones, tablets, and wearables as educational tools has been reflected significantly within the scope of Augmented Reality and technology centric learning models (Sad & Ebner, 2017). Mehdipour & Zerehkafi (2013) differentiated between the pedagogies for e-learning and the virtual learning environment which is accessed from portable devices as the change from classrooms to anywhere, anytime learning; an increase in audio, graphics, and animations; instruction delivered in multi-languages; and the instant delivery of SMS or email.

The potential of digital media as educational tools to support learning, has been emphasized globally, as national departments of education, researchers and developers strive to develop educational mobile apps, or mobile game based apps, or mGBL, which

[9] Based upon the 2014 Nielsen study of mobile phone usage and apps by 5,000 users 18 and older with iOS and Android devices.

are developmentally appropriate; sustain the student interest in learning; commercialize educational mobile apps in the global market place; and to close the digital divide between students from different socioeconomic backgrounds (Chiong & Shuler, 2010). Popular educational mobile game apps for preschoolers and elementary school aged students that are available for download on iOS and Android include *Super Why!*, *Monkey Preschool Lunchbox*, *Endless Alphabet*, *GazziliScience*, and *Rosetta Stone Kids Lingo Letter Sounds* (Schiola, 2016). Cloud and tablet computing has increased in adoption by educational institutions as an avenue to access Skype, Google Apps for Education, and Dropbox services. Sad & Ebner (2017) estimated that the time-to-adoption for educational apps is between 2 and 3 years. The immaturity of design and usability requirements for iOS and Android mobile apps has been a quality control concern for educational institutions that have been confronted with the infiltration of e-learning and mobile learning platforms. To date, the Apple App Store has instituted a premature set of guidelines for the design and content for mobile app developer inputs; while the Android market has not (Chiong & Shuler, 2010).

A recent study of the knowledge, perceptions and ability of young children to use smartphones, iPads and iPods showed that the children have access to smart devices; have a particular interest in iPads and iPods; are overall, generally adept users; and may learn from mobile apps (Chiong & Shuler, 2010)[10]. However, the outcomes showed the access to smart devices was limited, that the children perceived the mobile device to be a resource to play games, and that the attention span or length of interest in the apps was ‘fleeting’. Concerns have also been expressed in regard to the “disruptive

---

[10] Based upon the Usability Study: Sesame Workshop in which a sample of 114 children between 4 and 7 years of age were interviewed.

track record” of mobile communication device usage within the schools for social purposes (Shuler, 2009, p. 3). Further, studies have shown that mobile learning is challenged by small screens and keys; a short battery life; the potential for physical health problems; potential for distractions; the potential for unethical behaviors; information privacy issues; digital divide in regard to costs and accessibility; and no existing model for mobile learning theory (Shuler, 2009; Mehdipour & Zerehkafi, 2013).

The scope of the technological capabilities of the digital learner and the future of educational mobile apps extends well beyond the pre-k and elementary school students to include all levels k-12, university students, corporate training and education, military classrooms and scientific/research based professional curriculums. Nonetheless, the impending transition to M-learning and the use of educational mobile apps is expected to prove more difficult for the older generations, as the “app culture” has caused a digital divide between the adult learners and the digital native learner. Purcell, Entner & Henderson (2010) found that more than 80% of a sample of adults own a mobile communication device and more than 20% of adults live in homes with at least one mobile phone, but no landline telephone. However, the outcomes of a Pew Research study indicated that although most adults have smartphones, a large percentage of them are not familiar with how to actually use the apps. Only 30% to 35% of the sample used the smartphone to play games, music, or videos. A significant number of the oldest adults in the sample did not use any of the apps on their phone at all; while approximately 11% of the sample asserted that they were not even sure if their smartphone was equipped with apps.

Summarily, Sad & Ebner (2017) supported that as mobile technologies have become pervasive in everyday living; the technologies should also be integrated into educational platforms for seamless learning. Shuler (2009, p. 2) noted that the "ubiquity of adolescent engagement with media" is a critical success factor for education, despite parental and medical professional concerns for the risks of continuous screen viewing to health and development. Further, the potential for advanced mobile technologies has remained untapped, particularly for students in rural, low-income families (Shuler, 2009). Studies show the market for youth apps have greater revenue potential, as consumers are more willing to pay more for educational mobile game apps, and children's book and movie downloads, which average approximately $6 per app (NPD, 2013)[11].

### 2.4.2 Serious Games Mobile Game App Genre

Game-based learning (GBL) is considered a subdivision of serious games which also solicits positive learning outcomes for both adults and children (Susi, Johannesson & Backlund, 2007). In turn, digital game-based learning (DGBL) is a form of GBL that utilizes digital game platforms. The "Serious Games" genre encompasses digital games that are developed to solicit positive social and psychological effects through models outside of the scope of entertainment (Susi, Johannesson & Backlund, 2007).  Most game titles that fall in the serious game category are applications associated with the education, the military, healthcare, and public and private agencies and enable the user to experience situations that would be impossible in the real world due to constraints of time and costs.

[11] Based upon a 2013 survey of 2,248 female parent members of the NPD online panel with children between 2 and 14 years of age. See NPD 2013.

## 2.5 Predictive Information

Information theory characterizes the potential and limits of prediction algorithms and unified extrapolation analyses into a generalized sense of predictability (Bialek et al 2001). Such analyses were adopted from evidence-based economic research and applied as a forecasting tool for consumer behavior and retail market products. Theories of event forecasting and the predictability of human behavior can be traced to prior studies by Shannon (1951), who associated prediction with entropy, and later created a human "mind reading" machine[12]; to Laurent & Thompson (1988), who explored correlations between visual information and human action; and Schultz, Dayan & Montague (1997) who proposed that future event prediction as an inherent trait of adaptive organisms[13].

Laurent & Thompson (1988) described predictive information as an opportunity, or an open door to prospective control. Shannon Vucevic & Yaddow (2012, p. 33) presented that statistical regression "is not limited to the production of predictions; but the production of predictions that are accompanied by significant levels of confidence associated with the prediction" which may enhance intuitive visual comprehension and stimulate creative business insights where this study was agreed with (Shannon Vucevic & Yaddow 2012) and extract the significant levels of confidence associate before developed the prediction model. Merhav & Fedar (1998) addressed arguments that the future and the past must not necessarily be related; and asserted that evidence supports the existence such a relationship, the knowledge of which drives the development of universal prediction models. Diebold & Kilian (2001) contributed that the extent of

[12] See C.E. Shannon (1951) *Prediction and entropy of printed English;* and C. E. Shannon (1993) *The Mind Reading Machine*.

[13] Also see Hale & Saxe (2013) *Theory of the Mind: A Neural Prediction Problem*, in which predictive coding is discussed in support of the theory that the human neural systems make futuristic predictions based upon information received, particularly visualized information.

predictability of a time series is dependent upon the amount of data which is reflected from the past and that may be applied to future values for the series.

Bialek, Nemenman & Tishby (2001) described predictive information $I_{pred}(T)$ as mutual information that is linked between the past and future periods of a time series. Further, the behaviors with regard to sizable observation times T:$I_{pred}(T)$ limitations may be finite, grow in a logarithmic fashion, or grow in the form of a fractional power law. Further, Bialek, Nemenman & Tishby (2001, p. 2416) presented that predictions reflect differentiated averages or proportional perspectives of dense or concentration distributions; while information theory quantifies the "concentration aspect of the distribution in the absence of any assertion in regard to the attractiveness of the averages". Further, observations of the past $x_{past}$ retain only a small percentage of information that is relevant to any prediction, which is expressed as a law of diminishing returns[14]:

**Equation 1**

$$\lim_{\tau \to \infty} \frac{\text{Predictive information}}{\text{Total Information}} = \frac{I_{\text{pred}}(T)}{S(T)} \to 0$$

Entropy $S(T)$ is used to measure all data at time T; with linear growth limits over time.

[14] See Bialek et al 2015. The law of diminishing returns is discussed in regard to the collection of big data for observation times T. Here, the percentage of the data that is useful for the prediction problem based upon assumption of a future which extends toward infinity.

## 2.6 Gamification Theory

Gamification is an application of psychological and scientific tenets to the gaming environment in order to stimulate specific behaviors from the user and to build relationships between the members of targeted market communities. The user may be motivated by the promotion of elements such as autonomy, value and competence in the game design and experience. Successful gamification strategies continuously engage the user and generate repetitive subscriptions and game use. Walsh (2016) contributed that the individual is by nature, competitive, pursuant to being a part of a crowd or community, and a seeker of achievement and validation; universal attributes that are exploited by gamification. Therefore, the most successful mobile game apps immediately name or introduce the user, attach a profile that may be used for interactions; empower the user with frequent game tips and updates; and implement the accumulation of points.  Further, Walsh (2016) pointed out that gamers that are happy with the user experience are likely to adverse the title to others, or create a buzz in their communities.

The vastness of the software development market and the intensity of market competition drove industry leaders back to the drawing board in order to find significant links between information systems, human behavioral trends, and software engineering (Amir & Ralph, 2014). The outcomes have been a defining and pursuit of gamification, or a union of technological systems, immersive dynamics, and entertainment for profit. Law, Kasirun & Gan (2011, p. 353) presented that the "game layer" is added to the mobile application to create fun and to influence. Jiang (2011, p.4) argued that gamification may stifle individual creativity, as the user must overcome functional fixedness"[15].

Bharathi et al (2016, p. 361) defined gamification as "a modern paradigm which uses game thinking and game mechanics to drive behavioral changes" through motivations such as "reward offerings, challenges, and level increases". Law et al (2011) agreed, pointing out that app users are more likely to pay attention to activities which involve a reward, the accumulation of points, or an increase in status than those which do not. However, Vucevic & Yaddow (2012) submitted that the concept of quality is not represented in any mobile application feature or attribute, but rather the relationship between the product and the stakeholders.

Gamification emerged suddenly, along with other software development contexts which were born into the digital era of methodologies, games, and information technology culture. Dragona (2013) described gamification as a "trend, strategy, or buzzword which may be associated with a diversity of activities where badges, leaderboards, and progress bars challenge the individual to increase the level of performance. Werbach & Hunter (2012) presented that the mobile game app features that are relevant to gamification consist of Components, Mechanics, and Dynamics. The mobile app feature components represent instantiations of the dynamics or mechanics. The feature mechanics define the general processes which drive the action and engage the user; while the feature dynamics present the overall aspects of the gamified systems.

Bharathi et al (2016) quantified the gamification features of 60 Google Play Store mobile game apps that recurred most frequently in successful gaming

---

[15] Based upon study outcomes by psychologists Duncker (1930) and Glucksberg (1960) which measured the impact of incentivism.

applications. The highest ranking games in the sample were *Clash of Clans* and *Candy Crush Saga*; while the lowest ranked were *Craps Trainer Pro* and *Flick Golf*. Game features which would potentially increase the probability of game success were identified through the Sequential Minimal Optimization (SMO) algorithm. The study considered the user feedback in regard to the user experience based upon 24 game design features which were extracted from prior studies and classified under either Mechanics and Components:


**Figure 3. Methodology for Game Design Feature Quantification**

Both successful and unsuccessful titles were selected in order to define the game feature combinations which significantly impact the game success or failure (Bharathi et al 2016).

The provision of point accumulation during game play was found to be the most impactful game feature, followed by the incorporation of avatars, challenges and virtual goods. Moreover, the user feedback feature ranked 15th out of the 24 game designing features in regard to the impactfulness on game success predictive power.

Customer-oriented gamers may typically subscribe to mobile game apps which are simple to comprehend and require a minimal amount of user transactions in one setting, as opposed to knowledge-oriented gamers who seek complexity through multiple functions (Vucevic & Yaddow, 2012). Based upon such diversity of cognitive and intellectual gamer demographics, conceptualizations of gamification in the game design construct vary, as some view the application of gamification theory to digital game design detracted from the true meaning or nature of game play. Further, game mechanics have continually progressed toward marketing strategies and revenue generation which amplify the psychological needs of the most frequent users. The target market is identified, along with the motivation for buying the product or using the service and how predictions of consumer behavior can be made; differentiations are made between market segments, along with estimations of how much the patronage of the target market is worth.

Bogost (2011) argued that the context of gamification lies outside of the actual game design and toward the exploitation of the user; and serves to reduce the gamer experience to a psychological round of stimulus-to-response. Kennerly (2003) argued that disinformation and best-case scenarios distort game design theory; therefore, data mining serves to substitute theory with facts, adding specialization to game development. Law et al (2011) argued that although gamification potentially improves user and mobile application relationships, research of such dynamics is lacking. Still, the success of gamification in educational disciplines as a tool to increase student engagement and solicit faster performance within the learning environment has made the method highly attractive to mobile game app developers.

Gamification theory is used for many types of mobile applications, to include apps outside the game playing context such as the *Waze* GPS navigation app. The most appropriate gamification method for mobile game app development is dependent upon the target demographic and the desired outcomes from the plan. Consumers who pay for mobile game apps, or “whales”, have been described as “lucrative yet elusive” (Deloitte, 2016). Approximately 1 in every 650 mobile game app subscribers represent half of the in-app spending on free-to-play game apps. Within the context of engagement and personal investment, such statistics are parallel to arguments against the long-term sustainability of gamification objectives. Bogost (2011) concluded that the essence of gamification is exploitation by capturing the true intentions of the gamer in order to capitalize upon cultural moments which will generate revenues until new trends begin to emerge.

**2.7 Mobile Game App Design and Features**

The capacities of mobile devices have evolved to accommodate rich, stand-alone and distributed client-server apps which retrieve data from web gateways (Holzer & Ondrus, 2009). The same evolution of the mobile device has instigated a global demand for entertainment and organizing applications which extend the capacity well beyond telecommunications, and has created a vast market for innovative mobile application development. However, the mobile game app developer is faced with intense competition amidst game theme and design categories that are grossly over-saturated, a constant flow of new entrants with instantaneous replicas, and the challenge of both free and premium platforms as the basis for revenue generation (Unhelkar & Murugesan, 2010). In this light, the use of predictive power toward the development of successful mobile game applications makes contribution as a solution to challenges of target marketing and the transitioning the free-to-play gamer to the premium player.

Bharathi et al (2016) described the game design feature as a required building block or shared feature of mobile games. The mobile application development process consists of virtual environment, core operating system, and rich operating system solutions. Vucevic & Yaddow (2012) asserted that in the event that the processes used by the software developer for the application development project are ill-defined and unorganized, the quality of the product can be neither predicted or repeated. Wooldridge & Schneider (2010) supported that the winning mobile gaming app must have the right combination of icons, graphics, audio, fonts and action. However, in consideration of the limited availability of resources, the number of attempts to create a successful mobile app design with popular features is limited.

Mobile applications executed on the mobile terminal carry requirements for simultaneous user interaction and services that are location dependent; which creates additional complexity challenges for the developer (Aleksy, Butter & Schader, 2008). A significant challenge to location-based apps is the need to determine the player' s location prior to fully offering access to the application; and the continual location tracking required in order to adapt thereafter. Unhelkar & Murugesan (2010) supported that the evolving mobile application paradigm centers upon location independence and enhancements to user satisfaction through improvements to user personalization and accessibility capacities. The context aware response must generate information based upon the user's location, personal attributes, and the time of day. The mobile terminal restrictions encompass the computational aspects; power supply; input-output capabilities; and communication. Despite reductions in high priced bandwidth, the mobile application requires management of higher periods of latency.

Aalto (2015) supported that the mobile game developer may construct successful game apps more efficiently and faster through a combination of stage gate models and agile development approaches. Figure 4 illustrates a diagram of the Waterfall Classic Software Development Lifecycle:


**Figure 4. Waterfall Classic Software Development Lifecycle Model**

The model begins with the establishment of business requirements, escalates to high level design, and ends with coding and unit testing. During each stage of the cycle, the steps are validated by system, integration, and component tests (Vucevic & Yaddow, 2012).

Based upon the Waterfall model, the mobile application is created using the development tools from the software development kit (SDK), after which the mobile app is published on an application portal for download. Within the cathedral model, approximately half of the users selected either Apple, RIM or Microsoft as the proprietary path (Holzer & Ondrus, 2009). The bazaar model consists of alternative open source technologies such as Google, Nokia or Linux-based platforms. Moreover, the application portal is created as decentralized or centralized.

The operability of the mobile game application must be adapted to smaller screens, screen resolutions, and variations in network connectivity and speeds. Doolittle et al (2012) presented that the need to support multiple operating systems, the landscape for mobile application development is complex, as the application possesses independent hardware, device specific security and capability features which impact the application performance and behaviors. The standards associated with the mobile app development include the IEEE 802.1x network protocols from 2G to 4G; infrared or Bluetooth localized communication; RFID, Wireless VoIP; and WiMax (Unhelkar &

Murugesan, 2010). A diversity of communication protocols are supported by the generic communication framework. Therefore, the communication and network capabilities for mobile apps significantly impact the quality of service. The use of multimedia content with structured data also poses a challenge for mobile app design in the areas of secured client-side content storage, content sourcing and representation, and content mining.

Doolittle (2012) established a 2 step mobile application development process in which a decision matrix is used to measure the suitability of a proposed app, followed by a determination of mobile app delivery. The decision matrix is used to estimate the potential market value of the proposed mobile app based upon the degree that the user experience is optimized and the use case for the app content as well as how well the app can meet security and connectivity standard requirements. The application is deployed through the web browser; through a combination of a web browser and a native HTML5 code; or business application options such as the Mobile Enterprise Application Platform (MEAP) for virtual deployments. From a marketing perspective, the context of the app must be in strong demand based upon current trends.

Aleksy et al (2008) proposed an architecture for mobile application development that focuses upon simplification for applications that are context sensitive. The components of the proposed architecture consist of context and component managers; a Service Discovery Directory; an adaptable User Interface Framework; and a Generic Communication Framework. In the event that the mobile device display format cannot support the graphical representations, the application illustrations are replaced by text. The architecture accommodates the development requirements and terminal restrictions, generic and security components, service oriented software architecture, and position ontologies. In addition, the architecture also addresses the client environment resource restrictions, current context data processing, and the user preferences.

Moreira et al (2014) associated successful game app features with the game statistics for number of downloads and revenue generation and conducted an analysis of the relationships between 37 mobile game app features and success based upon gross revenues and the number of downloads[16]. A total of 34 mobile games were analyzed using CRISP-DM methodology for the top grossing mobile game apps and the apps with the highest number of downloads from the Google Play Store top 100. Here, the methodology consisted of a decision tree model that was trained to identify mobile game patterns, associations, and correlations within the data set and a linear regression model for mapping features and performance. In regard to the download rankings, the Mobage, Achievement, and IAP features were found to positively affect the success of the game. Further, the Upgrade Item, Invite Friends, and Upgrade Status produced negative effects on the game success. Moreira et al (2014) found no correlation between the game app features and the position of the title on the charts. In regard to revenue performance, 9 features were associated with success.

### 2.7.1 Incentives

A general consensus of game theory as well as behavior a lists has been that incentives contribute to psychological and social retention (Lewis, 2004; Jiang, 2011; Law et al 2011; Bharathi et al 2016). Recall that Bharathi et al (2016) defined gamification in terms of game thinking and mechanics which drive behavioral change through motivations, to include reward offerings. Also, Law et al (2011) supported that users are more prone to engage in activities which involve a reward, the accumulation

---

[16] The 37 game features were selected based upon studied which measured the impact of narrative; relationships between the parameters of level design and the player skill or experience; the impact of game aesthetics; and the dynamics of gaming rules.

of points, or an increase in status in juxtaposition to those which do not. Lewis (2004) argued that incentives stimulate the user to shift from a myopic decision making mode to more dynamic, long term decision making processes. Contrastingly, Jiang (2011) theorized that misapplied disincentives promotes cheating, and compromises the integrity of the game.

GameRefinery analyzed the top 170 mobile game features across a platform of top 200+ revenue grossing applications and 700+ games that did not make the top 200 list in order to establish rank for extensive game features and potential commercial improvements (Kanerva, 2017)[17]. According to the study, approximately 40% of the top 100 games incentivized the login with gifts; while approximately 19% of the games that did not make the top 100 incentivized logins. Progressive daily rewards calendars and daily quests were also analyzed as highly monetizing game features. The gacha element as a component of daily rewards, through which the player has a random chance of reward with minimal effort. Approximately 36% of the top 100 games included gacha in the daily rewards; while approximately 21% of the games which did not make the top 100 included the gacha element feature. Moreover, guild mechanics were found to significantly increase player collaboration as well as social responsibility.

### 2.7.2 Pseudorandom Generation

A significant amount of debate has arisen in regard to the optimal approach to mobile game design; as to whether random () or procedural features are more successful than pre-made level generation (Walker, 2014; Davis, 2012). Pseudorandom generation

[17] The study also focused upon the recent revenue and retention statistics for Nintendo's *Fire Emblem: Heroes* in United States and Japanese markets.

is popular in gambling and war game genres and has been used in game design for the issue of unit stats as opposed to cloning levels, map and sprite generation as opposed to hand-drawings, and artificial intelligence for decision making, as opposed to fixed decisions. *Pokemon* utilizes the random generator for encounters; *Spelunky, Hoplite, Terraria* and *Pixel Dungeon* are controlled by random levels and randomly generated environments; and several slot machine games use pseudorandom generators to control the outputs.

Walker (2014) argued that random events have traditionally been calibrated to merely present a mirage of randomness. The potential for boredom has been a primary argument against pre-made level game designs. However, game developers have avoided the randomness approach due to the problem of uneven distributions of numbers in the return, irrespective of the number of if: else conditions. Davis (2012) conceded that the use of pseudo random number generators such as the Random Class in C# may be used to rapidly approximate randomness and that implementation problems may be remedied by use of a “Shuffle Bag” approach. The Shuffle Bag approach encompasses the selection of a range of values with the optimal distribution; combining the values; and then selecting all values, one by one, using basic code.

### 2.7.3 Brand IP

Julkunen (2016) studied the higher success rate of mobile game apps that have been marketed as global licensed brands, such as *Simpsons: Tapped Out, Walking Dead: Road to Survival, Kim Kardashian: Hollywood*, and *Fallout: Shelter* to determine the extent to which the brand IP drives the commercial mobile game app success. The method of measure for the impact of the brand IP was statistical modeling with controlling for more than 150 additional features, in order to prove that the brand IP effect was quantifiable. The popularity of the game apps was established based upon Google trend data. The results reflected a steady increase in the significance of the brand IP of approximately 11% after reaching a Brand Index threshold = 1. Further, Julkunen (2016) found that the popularity of the brand played a critical role in the outcomes and that celebrity based mobile games for celebrities such as Kim Kardashian and Britney Spears were projected to lose at least 50% of revenues without the celebrity name. Thus, the brand IP has a tremendous effect upon the global market success of the mobile game app and are effective solutions for increasing marketing competition and driving stagnant revenues.

### 2.7.4 Usability and Functionality

Mobile application development transitions the technical skill requirements for software engineers and artists toward software development approaches which must balance cross-platform usability and functionality. Doolittle et al (2012) asserted that mobile application development requires an understanding the significance of simplicity and usability; a selection of the most appropriate method for application deployment and the user base for mobile operating systems; comprehension of mobile security and established governance. Also recall that Sad & Ebner (2017) pointed to imature game design and usability requirements for mobile apps as quality control issues in the development of mobile educational applications due to an increase in e-learning and mobile learning platforms.

All mobile applications are subjected to usability and functionality testing. Successful mobile game apps will exhibit unique qualities of user-friendliness, simple game objectives that may be easily learned, and logical timing and scope in regard to task completions and features (Mifsud, 2016). Usability testing establishes what the user will see and experience during game play; while function testing measures the robustness of the capacity to receive inputs and execute seamlessly. Such execution on the mobile device is challenged by a small screen and limitations of the device's processing capacities (Mifsud, 2016). For mobile game apps, usability and the user experience is more of a driving factor than functionality. The significance of usability and functionality testing increases along with the expansion of the user expectations.

The mobile game app may be tested for usability by laboratory-based usability tests with real users; or remote usability tests, which may use simulations. The usability and functionality test data is used to form metrics in regard to task timing, completion rates, errors, user satisfaction ratings and success rates. A usability study concluded that the 10 primary usability features of the success mobile game app will include[18]:

1. Simple input buttons
2. Effective on-screen real estate development
3. Feature feedback
4. Clear Icons
5. Intelligent Interruption Settings
6. Instantaneous Session Starts
7. Intuitive sound settings
8. Simple tutorials
9. Rewards and Goals
10. Asynchronous Multiplayer Options

Attempts to convert successful console games to mobile game apps have been hindered by the challenge of usability and functionality limitations of the mobile device. The conversion of *Grand Theft III* to a mobile app required several on screen push buttons, and the player no longer had the vibration of the joystick and the tactical feedback.

---

[18] See G. Redwood. 2012. Mobile Gaming: A Usability Study.

### 2.8 Data Mining

Data mining, and more specifically, frequent pattern mining for predictive analysis has become a critical tool for high-level business[19]. Kennerly (2003) presented that the data mining process is initiated with "accurate, empirical data" that may be used by the game designer to identify "victims of deficiencies in game balance" and to construct "better game theory". The primary objective of the most common data mining approaches is to generate predictions from pattern discovery for some horizon of interest (Han et al 2001; Poggio & Smale, 2003; Mannila, Toivonen & Verkamo, 1997; Pei et al 2001).

The data mining process begins with the collection of raw data, data processing and cleaning, and exploratory analysis of the data, and the application of models and algorithms. The validity of the data and assumptions are verified. Data mining models may be either descriptive or predictive and may be used to solve problems by description, prediction, classification, segmentation, and a dependency analysis (Chapman et al 2000).

Krauss et al (2008) proposed a data mining methodology which integrates automatic sentiment and social network analysis in order to predict movie success as exemplified by winning an Oscar. The methodology was based upon the "wisdom of the crowds" philosophy for predictions of future trends, which supports that large groups of general populations are more reliable for market trend predictions than the individual expert (Surowiecki 2004). Krauss et al (2008) mined the online discussion posts in

---

[19] Kennerly (2003) presented *Better Game Design through Data Mining*, in which data mining was confirmed as a viable method of production cost reduction, significant increases in user renewal rates, and balanced economy. Kennerly (2003) also supported that, in reference to data mining, mobile game design may not depend upon player feedback.

regard to movies within the IMDb.com online community between 2005 and 2006; and movie information in Box Office Mojo. The first half of the research consisted of an analysis of the posts based upon variables of positivity index, time, and intensity, or frequency. The second half consisted of a success measurement based upon the release date and show times. The probability of success was estimated for the 25 movies for the time period. The results showed that a significant correlation exists between community discussions about the movies and the movie success, and that a high communication intensity rating was a positive indicator of box office success. Krauss et al (2008) also concluded that the level of passivity in the discussion posts was sufficient as a predictor of future movie success. A major difference in the market revenues between movies and mobile game apps is the tendency of movie producers to create only 2 or 3 sequels to box office hits; as opposed to game developers such as Madden NFL, Mario Brothers, and FIFA have produced as many as 10 sequel titles for a popular video console game (Deloitte, 2016).

A fundamental, underlying assumption of data mining approaches is that more knowledge may be obtained from the perspective of the masses, than from one expert, professional, or academician. Through advancements in information technology, the "wisdom of the crowd" may be captured through an analysis of the perspectives and behaviors of large volumes of individuals (Surowiecki, 2004). To generate a definition for the "wise" crowd, Surowiecki (2004) differentiated between analyses that require collective intelligence as either for problems of cognition, coordination, or cooperation. Cognition problems are accompanied by present or future definitive solutions. Coordination problems seek to find optimal coordination models for behavior. Cooperation problems encompass the solicitation of cooperation from disinterested groups.

A great magnitude of data in regard to user preferences and behaviors may be mined from social media networks, consumer profile marketers, and other sources of predictive information that is application oriented. Poggio & Smale (2003, p. 537) presented that the roots of data mining lie in discipline of supervised learning, or “systems trained by examples, or input-output pairs”. Mannila et al (1997) confirmed that event sequences which describe user actions and behaviors could be captured in multiple domains; and proposed an algorithm to discover the frequency of such episodes to be used to generate sequencing rules.

### 2.8.1 Data Mining Algorithms

The identification of frequent patterns is problematic for data mining processes. Agrawal & Srikant (1994) addressed the need to visualize consumer consumption patterns and explored the potential of such research through Association Rule Mining of large sets of market-basket data. Utilizing Apriori and the AprioriTID algorithms based upon the anti-monotone Apriori heuristic, captured data could be analyzed for frequencies by the firm as predictive analyses of consumer preferences and behaviors. Frequent pattern mining by candidate generation-and test or pattern growth became an integral part of data mining associations; sequential, max and frequent closed patterns; clustering; and classification tasks (Pei et al 2001). Han et al (2004) proposed the more cost-efficient, frequent-pattern tree (FP-tree) method of data mining of complete sets of frequency patterns that could be used for performance analytics of large databases.

A diversity of algorithms for data mining have been proposed with cost and time reduction objectives. Pei et al (2001) introduced the H-mine algorithm and H-struct data structure as a rapid, high performance, space saver method of data mining that was also compatible with Han et al’s (2004) FP-trees. Zou et al (2002) presented a pattern decomposition algorithm that would reduce the dataset size in order to increase the efficiency of frequency patterns for larger datasets and argued that the pattern reduction approach outperforms Apriori-based algorithms and FP-trees which are constructed using candidate-set generation approaches.

Advancements in computational power have since significantly increased the capacity for data mining methods and the use of variables from the datasets that may be used for forecasting. Kennerly (2003) argued that the process of raw data collection may be complicated by the use of temporal cycles of time such as seasons, holidays, weekdays, and the time of day; and that the most instructive cycle for data mining is the weekly cycle. Clark & McCracken (2012) pointed out that within the data mining process, the search must be considered during the process of validation in regard to the significance of the outcomes for multiple testing. Traditional bootstrap methods for multiple test problems typically assume the baseline model to be non-nested with a minimum of one alternative model. Moreover, the more modern, simpler baseline models also contain versions of competing models.

The accumulation of information has increased exponentially during the Information and Digital Ages, as more sophisticated databases, data warehouse, and ubiquitous methods of data extraction and sharing have dominated all disciplines of virtual business analytics. Taylor (2017a) presented that decision making is the primary objective of analytic approaches to solving business problems through the development of decision models that will exploit the metrics for success. Moreover, Geisser (1993) supported that data mining approaches would produce more beneficial results by the inclusion of decisions and inferences within the predictive or observable analytical framework.

The most fundamental use of such massive accumulations of data is for prediction (Friedman, 1997). Geisser (1993) described prediction as the most primitive and the most common form of statistical inference that has evolved from parametric estimation approaches to testing. Shmueli (2010) described statistical modeling as a powerful methodology for the development and testing of theories utilizing description, prediction, or causal explanations. Moreover, Geisser (1993) contributed that within the scope of statistical predictions and the degree to which structure may be infused into the problem may be used to define the inferential model, to include the Bayesian, fiducialist or likelihood models; and that the Bayesian method has the capacity to generate prediction probability distributions.

Descriptive models objectively capture data structures parsimoniously. The descriptive modeling approach is less formal and is characterized by a minimal dependency upon a causal theory with no objective for prediction. However, statistical models that are characterized by extensive explanatory power typically possess inherent predictive capabilities (Shmueli, 2010). Empirical methodologies for problem solving in data mining projects may encompass the use Clementine, MineSet, or the Decision Tree (Chapman et al 2000). The predictive model for data mining formulates logical or mathematical estimations for unknown targeted values. Further, the predictive model may be used to manage risk, reduce fraud, and maximize the value added to the consumer (Taylor, 2017a).

### 2.9 Prediction Models

Reliable outputs from predictive analyses of big data are of tremendous value to market analysts as forecasting tools. Statistical models are typically used to build relationships and to test them in order to defend theories of causal relationships that are used as a basis for prediction. Hee (1966) outlined the importance of both qualitative and quantitative testing of the validity of forecasting model formulas and values. Vucevic & Yaddow (2012) supported that mobile game design approaches that are unorganized and that are not clearly defined prevent the prediction and repetition of the quality of the product.

The need to test the validity of theoretical scientific proclamations has been well served by predictive modeling and testing methods which identify existing causal mechanisms and discover differentiations in construct operalizations (Schmueli, 2010). The predictive model may reveal potential improvements to existing models and bridge the gap between theoretical assumptions and application. Classical models for probability and prediction were developed prior to the information technology revolution and pre-programmed algorithms. Further, the quantification of predictability within the model is commonly achieved by a statistical evaluation of distinct features through both explanatory, descriptive, and predictive approaches.

Schmueli (2010) defined the predictive model as a tool for applications of data mining in order to make new and future predictions. The approach for making predictions may be Bayesian, frequentists, parametric, or statistical models. Statistical predictions within dynamic systems fit the Bayesian model; when time and the relative system state are not referenced, the distribution selection is generated by equilibrium distribution. Model-to-data fit may be measured by a prior and posterior predictive checks or from a hybrid of checks designed for the hierarchical model (Gelman, Hwang & Vehtari, 2014). The most common prediction model is the multiple linear prediction model which utilizes predictor and response variables.

The primary objective of logistic regression is the modeling of some probability. More specifically, predictive modeling is used to forecast unknown values based upon other values or attributes that are known. The learning algorithm is used to generate a dependable rule for the prediction of probable outputs for future data (Friedman, 1997). In the case of non-stochastic prediction, the objective is the prediction of output (*Y*) for any new observations with input values (*X*), and observations to time (*t*) are the basis for forecasting future values at a time $t + k$, where $k > 0$ (Geisser, 1993).

Information criteria, the measure used to determine predictive accuracy, may be defined as the deviance, or log predictive density (Gelman, Hwang & Vehtari, 2014). Shmueli (2010) differentiated between predictive modeling and explanatory modeling in that the scientific objectives for data mining are for predictions or for explanation, respectively. Diebold & Kilian (2001) proposed a method of predictability measurement which utilized the ratio of short-term and long term expected loss that could be used with general loss functions, uni- and multivariate information data sets, and with difference and covariance stationary processes. The predictability measure is expressed as:

**Equation 2**

$$\mathrm{P}(\mathrm{L}, \Omega, \mathrm{j}, \mathrm{K}) = 1 - \frac{\mathrm{E}(\mathrm{L}(\mathrm{et} + \mathrm{j}, \mathrm{t}))}{\mathrm{E}(\mathrm{L}(\mathrm{et} + \mathrm{k}, \mathrm{t}))}$$

where $E_{(}L(e_{t+j,t})_{)}$ depicts the most desirable short term forecast; $E_{(}L(e_{t+k,t})_{)}$ depicts the most desirable long-term forecast; E(.) depicts the mathematical expectation conditional for the set; Ω represents the uni-or multivariate information set, and $k = \infty$[20].

[20] Short and long term based upon Granger & Newbold (1986) predictability assessment where j=1.

Benefits of the approach included validation for covariance and difference stationary time series in which $k < \infty$; the allowance for general loss functions; the uni- and multivariate information data set allowance; and the multiple options for *j* and *k* assignments for predictability studies across a diversity of horizons. The study also presented an application of the parametric method to other series of varying sizes based upon a fitting of autoregressive models.

The predictive analysis has traditionally been used as a metric for performance in economic disciplines, and as a tool for forecasting future market trends for the purposes of both government and commerce. Diebold & Moriano (1995, p. 253) supported that the predictive performance and the adequacy of the model are “linked inextricably”; and that predictive failure is “an implication of the inadequacy of the model”. Thus, the value of the predictive model may be determined by measures of the accuracy its predictions, which is the primary criteria and an approach that is also appropriate for juxtapositions of accuracy between models (Gelman, Hwang & Vehtari, 2014). In turn, a juxtaposition of the predictability of multiple data series requires “a common numeraire” (Diebold & Kilian, 2001). Common measures of predictive accuracy include the in-sample predictive accuracy measure, the adjusted in-sample predictive accuracy measure, and cross validation.

## 2.10 Mobile Game App Success Prediction

The development of a mobile game app that will generate millions of dollars upon release is neither probable nor is it ever achieved by most developers (Filho, Moreira & Ramalho, 2014). Despite the trademark digital culture of the millineals, identifying the trends within the mobile game app ecosystem expands the scope of merely providing a form of entertainment. However, for those who do formulate a mobile-first, augmented reality strategy for mobile game design that appeals to the digital native to the degree that they are willing to spend, and continue to spend, a successful development could represent a lifetime of work. *Pokemon Go* reflected such potential in 2016, with a record breaking 500 million downloads and as much as $10 million a day in revenues (Mason, 2016).

Predictions of how the new mobile app will fare on the market may be determined, to some extent, by an analysis of prior market trends and the models used to develop historically successful games. Taylor (2017a) argued that the successful analytic approach will consume more time in gaining insight of the business problem than sorting through large sets of data and will “reduce the white space between analytic and business success. However, the mobile app market structure and value chain are affected by a diversity of factors which are shaped by market leaders such as the Nokia Symbian OS, the iOS, the Windows CE OS, Google Android, the LiMo Linux Mobile platform, and the Blackberry OS (Holzer & Ondrus, 2009).

Mobile game app success is, in itself, defined differently according to the perspectives of different stakeholders. In order to be deemed successful from an economic or accounting perspective , the Life Time Value (LTV) or Customer Lifetime Value (CLV) of the product must be higher for the game than the User Acquisition Cost (UAC) (Sifa et al 2015). From the perspective of the mobile device producers, Unhelkar & Murugesan (2010) pointed out that the mobile device is more personal to the user than the desktop computer or game console; therefore, sensitivity to user personalization is a critical part of mobile application development and design. From the perspective of the software developer, Law et al (2011) asserted that mobile app user retention is dependent upon the relevance and utility of the application features; and that the frequency of use is determined by considerations of leverage, stickiness, and feedback. Collectively, Gualtieri (2011) submitted that the mobile application contexts of immediacy, location, device, locomotion, and intimacy are the most critical to the user experience.

Wilcox and Voskoglou (2015) argued that the mobile app economy revenues have begun to polarize; and that the premium mobile game app segment is owned by iOs, the Android market consumes most other segments, and the Windows and browser platforms are left to fight for any remaining market share. Deloitte (2016) conducted a study of mobile game app success, and predicted that the average revenues per mobile game app will vary mobile games due to the size of installation base, barriers to entry, and the scope of the business models. As the mobile device base increases, the mobile game app revenue will increase; however, the success of the mobile game apps will be distributed across a small number of developers. The required capital for mobile game app developers is expected to maintain the current market stratification. A study conducted by Vision Mobile showed that out of a total of 8,000 developers, approximately 17% asserted that they had generated no revenues from their apps; approximately 18% asserted that they had earned less than $100 per month; and approximately 50% asserted that they earned less than $1,000 in revenues each month (Wilcox & Voskoglou, 2015)[21]. Further, Deloitte (2016) predicted that due to differences in business dynamics, the mobile game app user population will continue to co-exist with, rather than to dominate game console and PC game markets.

[21] The data for the 8th Edition State of the Nation report was collected by a survey of 8,000 mobile game app developers. The report consisted of the results of an analysis of current trends in IoT and mobile developments, tools and models, and enterprise apps versus consumer apps. See Vision Mobile, 2015.

Sifa et al (2015) supported that the capacity to predict the demographic for mobile game app purchases prior to soft launch serves to optimize marketing and customer relationship management customization. Tuckerman (2014) proposed a method for mobile game application success prediction which encompasses the extraction of application features from the Google Play Store for analysis. The data is used to train 3 models for the prediction: A generalized linear model for classification as success or fail; a naive Bayes text classification model as a descriptor; and linear regression for average application rating prediction. The data for over 1.3 million mobile applications were collected by using a web crawler on Amazon Web Services (AWS) Elastic Compute Cloud (EC2). The features which were extracted included the developer rating, Android Application Package (APK) size, average user and star ratings, number of application installations, and most recent updates. Rather than use revenue as the success metric, the number of application installs and the average user ratings were used to find the probability of success based upon description:

**Equation 3**

$$\text{success } x = 1\{x_{score} >= 4.5\} .1\{x_{installs} >= 5x10^4\}$$

where application x has an average rating $x_{score}$ and $x_{installs}$. Further, a penalty was imposed for high price and lengthy descriptions in the GLM model. The predictive power of Tuckerman's (2014) method was limited; however, the outputs exhibited consistency in regard to photos and sharing as the most recurring attributes for the successful applications.

The fundamentals of mobile app success center upon the impact of the game design and features upon the user's decisions. Sifa et al (2015) used 2 models to predict mobile game app purchase decisions from a dataset of 100,000 gamers. The classification model and regression model were to predict purchase occurrence and the number of purchases, respectively. The classification algorithms consisted of Decision

Trees and Support Vector Machines. The regression was performed with Poisson trees for 1,3 and 7 day observations. Premium players were generated synthetically using Synthetic Minority Over-Sampling Technique - Nominal Continuous (SMOTE-NC). Purchases were defined as active spending on in-game items with real currency, to include purchases of in-game currency. The outcomes of the study supported that prior purchase behaviors for a gamer demographic, in-game interaction, and features which are activity related have a strong influence upon the players' future purchasing behaviors.

Filho et al (2014) extended upon the prior research of mobile game success based upon the features of 34 games by expanding the sample data to include 100 games. Thrity-seven features were analyzed for a total of 100 mobile games; 60 titles were extracted from the Top 100 Google Play mobile game apps, and 40 titles from either the Top 400 or the Top 500 Google Play game lists. The methodology consisted of a combination of data discrimination, classification and CRISP-DM used to compare gross revenues and the number of downloads for each title. A decision tree and linear regression model were constructed to extract knowledge of the predictive and class attributes. The Receiver Operating Characteristic (ROC) curve illustrated the binary classifier performance. The IAP Potential, which included game purchases made that were made with real currency, showed a negative correlation with the game revenues. The ability to share game content with others through Facebook and gambling using hard currency produced positive correlations with game success.

Alomari et al (2016) explored the potential of 31 mobile game app features and successful mobile game development environments based upon the gross revenues. The sample data consisted of 50 Apple App Store games for iPhone which generated substantial revenues in the United States. The methodology consisted of classification of data by the ARM funnel, and an analysis by the CRISP-DM with SPSS Modeler. Freemium games were the primary focus with attributes to include the Title, New Installs, Revenues, Developer name, Daily Active Users, and Game category. The data was normalized by the formula:

**Equation 4**

$$\text{value} = \frac{(xi - Min(range\ xi))}{(Max(range\ xi) - \mathrm{Min}(\mathrm{range\ x}i))}$$

to achieve values of 0 or 1. The model was constructed by a decision tree, regression, and a performance evaluation. As with Filho et al (2014), the ROC curve was used as the binary classifier for Facebook, Leaderboard, Event Offers, Skill Tree, Invite Friends and Request Friend Help, Time Skips, Soft, Customizable Currency, Reference Line, and Unlock Content features. The highest grossing game category was casino games, or gambling, followed by casual and strategy games. King had the highest frequency of top grossing games; however, Supercell generated more revenue with less games. The Invite Friends feature produced the highest significance ranking, followed by the Skill Tree and Leaderboard, respectively. The ARM classification presented the significance of the Facebook, Event Offers, Request Friend Help, Time Skips, Soft, Customizable Currency, and Unlock Content features as essentially equally distributed. However, the linear regression shows a more distinct differentiation between scores. The outcomes identified 10 features as significant in the achievement of mobile game app success with an outstanding relationship between the Daily Active User and revenue.

## 2.11 Superior Predictive Ability

The quality of the predictive model significantly impacts the validity and usefulness of the outcomes of the analysis. Tests for the superior predictive ability (SPA) are commonly used for comparisons of models for which one or more do not have the nested benchmark (Aalto, 2015). SPA provides the best forecasting model for performance and may serve as a benchmark against available alternative models. Clarke & McCracken (2012) presented a bootstrap approach which permits serial correlation and conditional heteroskedasticity within the predictive errors. Diebold & Moriano (1995) presented an equal predictive ability (EPA) test which supported the superiority of the benchmark over alternative models. Hansen (2005) presented a test for SPA that was ranked as superior to the data snooping reality check (RC). The methodology consisted of a modification of the RC by the studentized test statistic in order to address the erratic forecasts and to invoke a null distribution that is sample dependent. Optimal sample performance was observed by the models with Phillips Curve structures.

Aalto (2015) presented a tests for quasi likelihood ratio predictability and analyzed the out-of-sample equal predictive capacity with several models in which a parsimonious benchmark model is nested by others. In similar works, Hubrich & West (2010) proposed a direct extension of the pairwise model Wald-type tests juxtaposed in West (1996) and Diebold & Mariano (1995) to a statistic that is chi-squared or the use the largest of the t-statitistics for the MSPE pairwise differences. Clark & McCracken (2012) proposed a bootstrap method in order to estimate the critical values which are asymptotically valid in two tests of forecast and two tests for MSPE. Aalto (2015) presented an alternative to the approaches of Hubrich & West (2010) and Clark & McCracken (2012). In the case where all models nested the benchmark based upon adjusted MSPEs, the properties that were exhibited were reasonable. Upon incorporation of the bootstrap critical values, the sizes decreased. Aalto (2015) also submitted that the t-statistic limiting distribution is nonstandard and dependent upon the predictor characteristics.

### 2.12 CRISP-DM for Mobile App Success Prediction

Methodologies for data mining and logical data organization such as SEMMA and CRISP-DM assist the researcher in the analysis of big data. The data mining models may be used for predictive analysis, as well as other applications such as population clustering and likelihood estimations. Bartlmae (2000) presented that CRISP-DM is a hierarchical process model that provides a blueprint for the steps required for the Knowledge Discovery in Databases (KDD) process model structure and to meet the minimum product or service quality standards. The concept for the CRISP-DM approach was developed in 1996, and is common used by data miners and business process analysts as an analytical solution for knowledge discovery that may be used to make intelligent decisions in regard to business problems (Chapman et al 2000). In 1997, the CRISP-DM Special Interest Group (SIG) was launched as central point for information and concept sharing through a consortium workshop in Amsterdam[22]. In the years to follow, the CRISP-DM standard process model was revised and tested at Mercedes-Benz and OHRA (Chapman et al 2000). By 2000, CRISP-DM 1.0 was released and applied to a diversity of projects by Daimler-Chrysler. The hierarchical process model was presented in the following parts: an introduction, the reference

[22] Special Interest Group includes Pete Chapman and Randy Kerber (NCR); Julian Clinton, Thomas Khabaza, and Colin Shearer (SPSS); and Rudiger Wirth and Thomas Reinartz (Daimler Chrysler).

model, the user guide, reports, and an appendix with a glossary and data mining problem type characterizations.

The data mining context for CRISP-DM represents the mapping that occurs between the generic and specialized groups. The context is presented in the four dimensions of the application domain, the type of data mining problem, the technical aspects and the dimension of tool and technique. The Application encompasses Response Modeling and Churn Prediction; the Data Mining Problem may be Segmentation, Prediction, Dependency Analysis, Description and Summarization, Classification, or Concept Description Chapman et al 2000). The Technical Aspect may be outliers or missing values; while Tool and Technique may be the use of Clementine, Decision Trees, or MineSet.

The CRISP-DM framing activities begin with the first phase of Business Understanding and end with an Evaluation prior to deployment. The model methodology consists of four levels: Phases, Generic Tasks; Specialized Tasks; and Process Instances. Figure 5 shows a breakdown of the four levels of the CRISP-DM methodology:

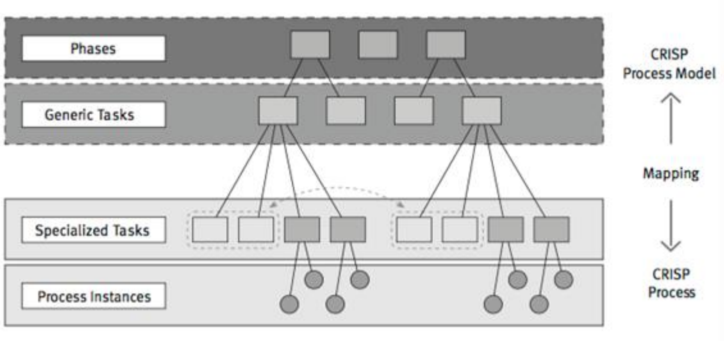



**Figure 5. CRISP-DM Methodology Breakdown**

The Phases and Generic Tasks are divided from the Specialized Tasks and Process Instances by Mapping. The Mapping between generic and specialized levels of the model is driven by the data mining contexts of application domain, the problem type, technical aspects and tool and technique. The Mapping component is a depiction of the mapping between levels for the present and for the future.

The first level divides the data mining processes into the phases of the CRISP Process Model. In the first phase of the process, the objectives are identified and defined within a software development language that may serve as a simplistic representation of the purpose for the project. In the second level, the phases are further divided into generic tasks for all potential data mining scenarios of the CRISP Process (Chapman et al 2000). The third level, Specialized Tasks, defines the actions that are to be executed in specified scenarios. Level four, Process Instances, keeps a record of decisions, actions, and actual outcomes from the data mining and is organized based upon tasks that have been defined at higher levels (Chapman et al 2000).

The CRISP-DM reference model depicts the 6 phase life cycle of the data mining project, which may be classified as subjective, rather than sequential. The life cycle begins with Business Understanding, followed by Data Understanding, Data Preparation, Modeling, Evaluation, and on to Deployment. Figure 6 illustrates the phases of the reference model:

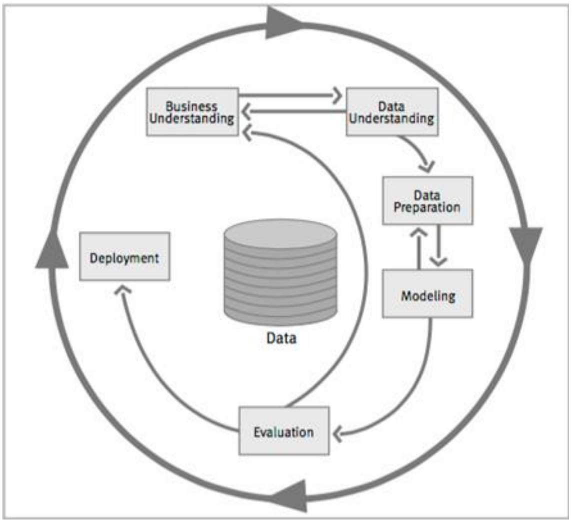



**Figure 6. CRISP-DM Reference Model (Chapman et al 2000)**

The outer circle represents the data mining cyclical nature. The Business Understanding component encompasses the level of clarity to the problem that is to be addressed in order to establish a focus for the project. Within the Business Understanding phase, the business objectives are established, the business problem is assessed, the data mining goals are determined, and a project plan emerges.

Initial data collection begins in the Data Understanding phase, within which the analyst becomes familiar with the data, identifies potential problems, verifies data quality, and forms the first insights associated with the data (Chapman et al 2000). The Data Preparation phase consists of the selection, cleaning, construction, integration, and formatting of the data (Chapman et al 2000). The Modeling phase consists of the selection of a modeling method, the generation of a test design, the construction of the model, followed by the model assessment. The final phase prior to deployment, the Evaluation phase, consists of an evaluation of the outcomes, a review process, and assessment of what should be the next steps. Figure 8 illustrates the generic tasks and the outputs for the reference model:

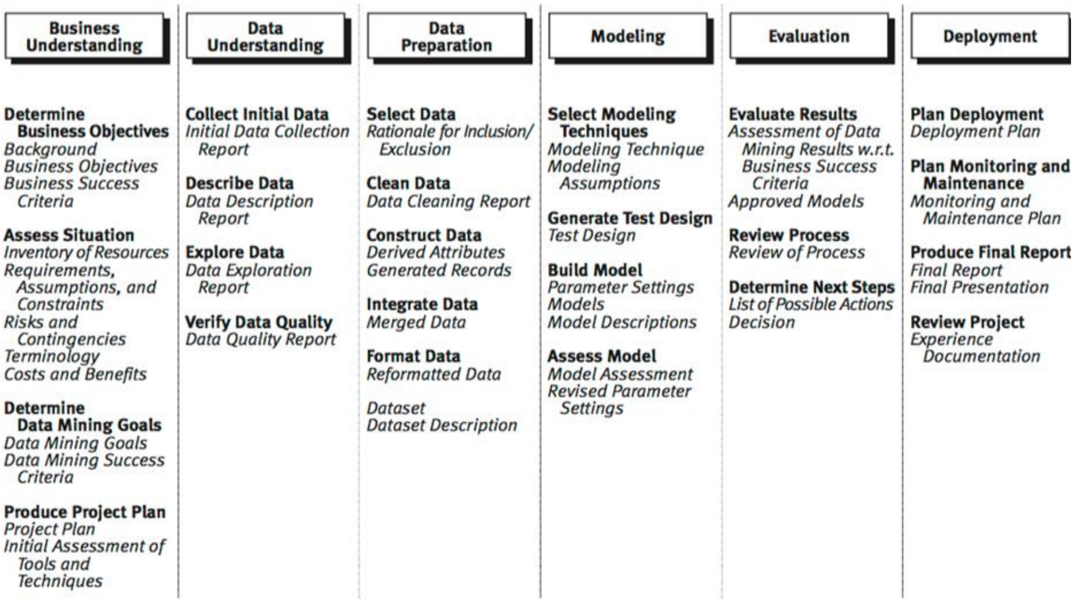



**Figure 7. CRISP-DM Generic Tasks and Outputs (Chapman et al 2000)**

The generic tasks are presented in bold print and the outputs are presented in italic. The outputs include an assessment of the resources available to conduct the analysis, to include the data, human resources, and relevant software tools. The requirements, assumptions, and potential constraints are defined, and a cost-benefit analysis is constructed to exhibit the ratio of project costs to the monetary benefits of success.

The CRISP-DM model is a cyclical process of analytical activity that encompasses non-linear movements between the phases in order to expand the original business problem and integrate the model into an existing operation. The primary states of implementation are analysis and corresponding actions. Information technology plays a significant role in the CRISP-DM processes, along with the collection of the most relevant data in the most appropriate format. Data preparation is significant as gaps are filled, errors are detected, sources are merged. Moreover, the data must be profiled in order to identify potential problems and cleansed prior to processing. The datasets may then be joined and variables are created for the model. Several modeling iterations may be required prior to running the simulations. The source data, data quality issues, methods, and transformations should be identified and documented as well as the outputs obtained by the model. The evaluation phase is significant in that it objectively reviews the process and tests the model performance. In the event that the model is useful as a real world application, it may be operationalized for other projects.

### 2.12.1 Challenges of Using CRISP-DM

The popularity of CRISP-DM as an effective approach to data mining and analysis is not without criticisms by some. The majority of the notations which point out deficiencies in the model encompass the tendency of the researcher or analyst to skip steps or to process some steps half-heartedly, which compromises the timeliness and validity of the results. Gruber (2017) asserted that CRISP-DM is limited in the areas of format, business understanding and deployment; and proposed decision modeling as a solution for data science projects. Taylor (2017b) noted that challenges of using the CRISP-DM model include deficiencies in clarity in understanding the business problem and in evaluations of the analytic outcomes in a business context. Other persistent issues with the application of CRISP-DM that have been noted include mindless reworks, iteration failures, and blind hand-offs to the Information Technology:

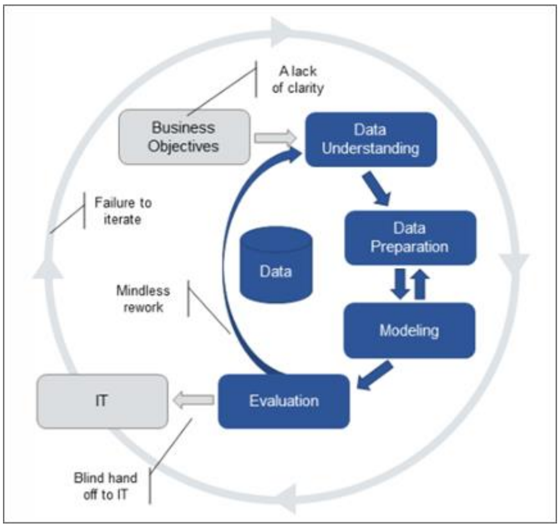



**Figure 8. CRISP-DM Challenges (Taylor, 2017b)**

When the project utilizes shortcuts, the CRISP-DM approach becomes corrupted[23]. The lack of clarity results from adopting business objectives which minimize the costs of the project. Reworks result from predictive model testing based upon ambiguous business objectives. Iteration failure is described as a form of neglect to aging predictive models through deficiencies on monitoring, maintenance, and deficiencies in long-term valuations.

As a solution, Taylor (2017b) proposed a more defined focus upon the decision making process during Business Understanding and the associated changes to the environment based upon the fundamentals of decision modeling.

Bartlmae (2000) generated a knowledge framework from extracted data using CRISP-DM under a KDD process model to solve the business problem. The study was based upon assumptions of case-based reasoning (CBR) theory, by which solutions to new business problems may be derived from insights from similar business problem solutions. An experience management methodology included an application of the CRISP-DM mapping to KDD tasks. An experience factory was used to reuse the knowledge gained from previous projects to improve the model processes. The outcomes of the study indicated that due to the repetitive nature of CRISP-DM

[23] Diagram published by J. Taylor. 2017. *Four Problems in Using Crisp-DM and How to Fix Them.* KD Nuggets.

applications, experience may be useful in projects conducted in succession as the prior knowledge may be integrated with the case-based reasoning (CBR)-based experience factory designed for KDD.

### 2.12.2 CRISP-DM and Decision Modeling

A clear vision of the decision requirements for the analytic project may be developed by creating a model of the decision requirements early on. The Decision Model and Notation (DMN) standard increases clarity towards the business problem, how the analytic will be used, and the approach to a solution with the highest probability of success (Taylor, 2017a). The DMN consists of Knowledge Sources and Business Knowledge Models; and the DMN standards are compatible with the Business Process Model Notation (BPMN) standards for the completion of process model tasks.

The decision model represents an iterative cycle of decision making that completes when all of the project requirements have been met. The decision model may serve to clarify the business understanding phase, to include considerations for how the outcomes of the predictive analysis will be used (Gruber, 2017). The context of the deployment is explored extensively and the decision model illustrates how the outcomes of the predictive analysis will add impact and value to the business case. Therefore, the implementation of decision modeling at the onset of the project, improves the outcomes of using CRISP-DM.

## 2.13 Global Mobile Game App Market Revenues

The history of revenues for the mobile game apps reflect several factors that include the introduction of WAP and JavaME and disruptive information technologies; price fluctuations for mobile devices; the iPhone revolution; information communication technology services and accessories; and variations in the trends of public interests. In 1997, the Nokia 6110 included *Snake*, *Memory* and *Logic* games; a currency converter, calendar and pager; and a battery life of approximately 3.3 hours. Nokia sold approximately 40 million Nokia 6110 phones in 1998, at what would be approximately €1000, or $1239 today (Cooper, 2013)[24]. In 2000, Christopher Kassulke founded *HandyGames*, which evolved into an enterprise that specializes in freemium games (Osborn, 2014).

Still, small developers struggled to enter the market that had been dominated by the pioneering operators. Approximately 60% of mobile game app developers were reported to earn an average of $500 per month (Dogtiev, 2015). Further, mobile gaming porting required great expense across platforms, and the method of configuration tracking, testing, and certification processes are complex (Clark, 2012). More sophisticated approaches to analyzing the mobile app market began to surface. Models such as CRISP-DM[25] were used to process the market data for mobile gaming titles that appeared in the rankings charts which were deemed as successful based upon correlations between the popularity of the features and the market performance. Acquisition, monetization and retention (ARM) models placed a focus upon the relationships between revenue, features, developer, users, and research.

By 2004, market analysts were optimistic about the potential growth of mobile gaming applications, despite the challenges of the new ecosystem and the market volatility. Revenue projections from 2003 to 2006 ranged from $3.6 billion to $18.5 billion (Tercek, 2007). From 2006, the highlights of the market consisted of the trial and

[24] Based upon current conversion rates for the British pound to the US Dollar.
[25] See Chapman 2000.

error of Nokia; the growth of Digital Chocolate through millions in investments; and the bankruptcy of Gizmondo (Osborn, 2014). The iPhone era, which began in 2007, provided the ubiquitous, commercializable designs and futuristic canvas that the mobile game app market desperately needed to recapture the public interest and to realize strong monetization (Osborn, 2014, p. 7). Between 2008 and 2009, the mentality of major mobile communication device manufacturers and mobile game app developers experienced overhauls, during which time innovation and differentiation became real solutions for both developers and marketers.

Dogtiev (2015) presented revenue figures for the global app market from 2011 to 2015 and predictions for 2016 and 2017 annual revenues. The 2011 revenues for all types of mobile apps was reported as $8.32 billion; up to $18.6 billion in 2012; $26.7 billion in 2013; $35 billion in 2014; and $45.4 billion in 2015[26]. The most significant increase was between the periods of 2011 and 2012; as the annual revenue increased by $10.24 billion. A Popcap survey reported that in 2012, approximately 44% of the population played some type of mobile game, 33% of which played the games on smartphones compared to 18% who played the game on a console (Osborn, 2014). Also significant, is that the Android market held 75% of the global market for smartphones in 2012 (Dogtiev, 2015). By 2015, the mobile game apps accounted for approximately 85% of the global market revenues for all types of mobile apps, with annual revenues of $35 billion (Takahashi, 2016)[27].

---

[26] The figures were based upon a 2017d Statista report which included analytics of data from Gartner TechCrunch.

[27] Based upon an App Annie Market Research study on the growth of the mobile app market, which included games as a category.

The total market growth for mobile apps was predicted to extend from $45.4 billion in 2015 to $76.5 billion by 2017; while the projected global market share of iTunes will be approximately 20% by 2020. An App Annie report projected that annual revenues for mobile game app in 2020 will be approximately $74.6 billion (Takahashi, 2016). Leorke (2014) supported that the incorporation of freemium versions of *Shadow Cities* and *Life is Magic* location-based games improved the commercial success of the games and increased the sustainability of their revenue streams. Global revenues for iOS and for Google Play have increased annually, with exceptional performances in the Commonwealth of Independent States (CIS), Russia and the Baltics between 2015 and 2016[28] (App Annie, 2016).

### 2.13.1 Free-to-Play v Pay-to-Play

The high upfront costs of the mobile game apps in the App stores, combined with the rise of Netflix and Spotify, and the consumer reluctance to release private financial and personal information drove the game developers to the concept of freemium games (Osborn, 2014). Leorke (2014, p. 143) argues that the modern digital games market is characterized by a “work as play ethos” which has been assimilated by the integration of freemium games. The market had struggled to solicit consistent revenues for $9.99 downloads. In 2009, Apple introduced the In-App purchase. *Candy Crush Saga*, *CSR Racing* and many other mobile games app titles that had been introduced in 2012 began to monetize. Gameloft released *Despicable Me 2: Minion Rush* as a freemium game, while EA Sports released *Plants vs Zombies 2*, *The Simpsons: Tapped Out!*, and *Fifa* on mobile. By 2013, the freemium pricing strategy

[28] The data was extracted for iOS and Google Play games for 4 publishers: *Murka* in the Ukraine, the Russian *Pixonic* and *Playrix*; and the Belarus *Awem*. See App Annie September 29, 2016.

predominated almost all of the mobile app categories, include mobile app games (Richter, 2013).

The advent of free-to-play (F2P) games brought a new method of revenue generation for mobile game app developers (Vankka, 2014). Revenue streams are realized through the sale of the virtual components in the game, while the game is made available to the user free-of-charge.  The platforms for the development of mobile games are primarily based upon Free-to-Play (F2P), Freemium; Paynium; and Pay-to-Play models which dictate the cost of the app (Alomari et al 2016). In 2016, Super Mario Run opened at $9.99 and generated $14 million in the first 3 days on the market (Thompson, 2016). Nonetheless, the majority of mobile games are developed based upon the Free-to-Play model. However, many freemium games, such as *Shadow Cities*, have been presented as freemium games, but realistically were designed to allow only laborious interactions unless the user makes purchases to obtain access to objects that enhance the game or increase the level of the characters (Leorke, 2009). The financial commitment that such game designs solicited also were purposed to increase user loyalty.

### 2.13.2 The Mobile Game App Lifecycle

Data artist Frick (2016) argues that in order to gain a full understanding of something, it must be measured. Grugier (2016) argues that the demand for studies of complex, diversified data retrieval, storage, and analysis by evaluators, analysts and artists is an ever increasing trend for an ever emerging economic sector. Within the economic sector for mobile game applications, the lifecycle of the mobile game app is finite and decreasing (App Annie, 2015). Figure 9 shows the five primary fixed events of a session mobile game app lifecycle from the Start event to Destroy:

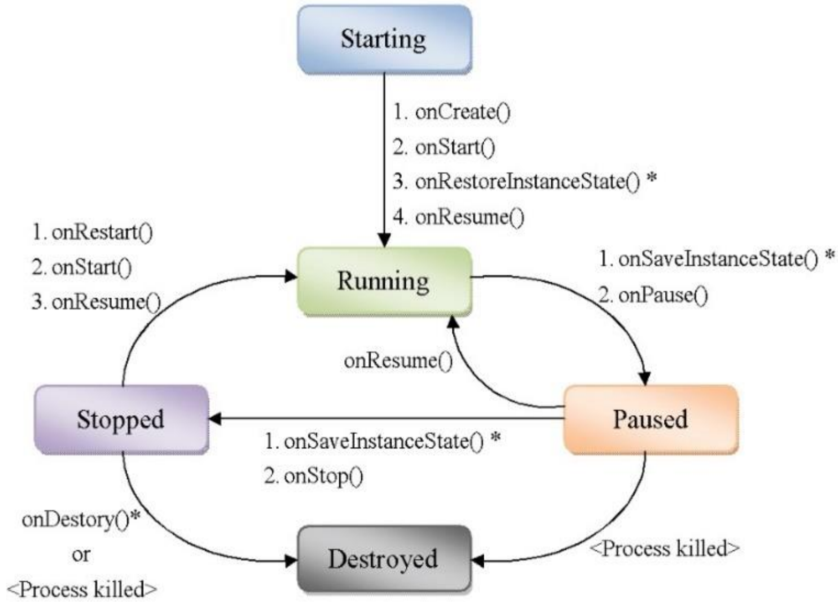



**Figure 9. Mobile Game App Life Cycle Events (Game Science, 2010)**

The lifecycle of the app begins at onCreate () and onStart() and is sustained through onResume (). The lifecycle of the app ends when the uninstall process has been completed by onDestry ()* during the Destroyed event.

The mobile app lifecycle begins with the user download, installation and the first use. From here, the number of times the application is opened and the duration of time spent using the application are classified as *Starting* and *Running* events (Game Science, 2010). The life of the mobile game app is also reflected in its value position within the mainstream demographic (Osborn, 2014). Mobile gaming porting is expensive across platforms, and the method of configuration tracking, testing, and certification processes are complex (Clark, 2012). Further, the input mechanisms and screen dimensions vary across mobile device manufacturers. The life of the game is already impacted by these challenges upon release; at which time, factors such as the visual presentation and the levels of difficulty come into play.

The leaders of the mobile game app market are the Apple App Store and the Google Play Store (Balan, 2016). The industry has been sustained through the continual development of mobile games that entice the user and provide high levels of entertainment according to the unique, socioeconomically-based demands of the user demographic. As the mobile app market competition has increased, game developers have needed data analytic methods which illustrated user behaviors, to include time and reactions to features and levels of difficulty (Fields, 2014). Bohmer et al (2011) developed a method of data collection called *appazaar*, which provides forecasting statistics for mobile game app with an internal *AppSensor* data collection component[29]. The mobile game app lifecycle is defined according to the observation of five events: the number of user installs, opening, updates, closing, and uninstalls. The length of the opening stage, during which the game is installed and being used, is a critical metric for forecasting the mobile app lifecycle.

The penetration of mobile app platform market is now diverse and heavily dependent upon the user demographic and location (Clark, 2012). Fields (2014, p. 179) presented that the sale of games that are not necessarily better by measure, but are differentiated in some manner, have the same potential to “delight” the user. The lifecycle of the game app is also affected by value added content and the extent to which the developer pursues user retention and loyalty. Data art creates aesthetic forms and artistic presentations from the data’s “digital nature” which is generated from big data in the form of statistics, graphics, simulations, and worksheets (Grugier, 2016). The algorithm may then be used to capture data from digital data streams and to quantitatively generate potential selections of content based upon prior user behavioral patterns. For instances in which new user continually subscribe to the mobile app or the time spent using the device increases, the lifecycle of the mobile game app may be extended. App Annie (2015) reported that the lifecycle of the role playing genre of mobile game apps and one tap games tend to mature more rapidly than other genres; while sports racing and gambling genre lifecycles are more consistent and longer

[29] The Mobile Game App Lifecycle is based upon the *AppSensor* definition presented in the Bohmer et al 2011 study of Mobile Application Usage.

lasting[30]. In addition, the first version of the mobile game tends to fare better and longer on the market than the follow-up versions.

### 2.14 The Mobile Game App Publisher Market

The success of mobile game apps in the new millennium and projections of future industry growth have attracted the attention of global investors as the average time Americans spend playing mobile games has exceeded the time spent streaming from Youtube and Netflix. Studies have shown that the publisher has a significant impact upon the gamer decision to download and install mobile game apps (Koetsier, 2017). The mobile game app publisher distributes the mobile game app content and solicits downloads and sales using a diversity of approaches. Common marketing strategies include media buys, public relations schemes, cross promotion, OEM features, Email, and SMS marketing (Mason, 2013). In return, the publisher receives a percentage of any revenues generated by the game. The amount of control the game app developer has over the product and the amount retained from sales may vary significantly based upon the terms of the contract.

The mobile game app industry has become a cash cow for many publishers, who retain a significant percentage of any mobile game app sales. Ballard (2017) addressed investment strategies to capitalize upon the growing mobile game app market, which has become a primary driver of industry growth for gaming. The rate of growth for the mobile game app segment indicates that in the future, mobile game apps may gain a larger market share than the PC gaming and video console sectors. Nonetheless, the

[30] The statistics for the metrics included in the report were derived based upon the App Annie Intelligence methodology for a sample of 4,346 unified mobile game apps from 2010 for iOS and from 2012 for Google Play.

potential for investments in mobile game app publishers appears to be drab, as publically traded Glu Mobile and Zynga have exhibited both profit and growth challenges. Activision Blizzard acquired King Digital, the developer of *Bubble Witch* and *Candy Crush*, for close to $6 billion in 2016, and proceeded to released mobile versions of *Skylanders* and *Call of Duty* (Ballard, 2017). Thus, investments in trending industry game app publishers such as Electronic Arts (EA), Activision Blizzard, and Take-Two Interactive Software which produce games for all platforms have been risky endeavors based upon prediction of mobile game app industry success.

### 2.15 The Apple Mobile Gaming App Store

The advent of the iPhone positioned Apple as an industry leader and opened up an innovative smartphone platform for the development of more sophisticated mobile applications. Since then, mobile app developers have generated billions in revenue, much of which has been from successful mobile app games for iOS. As with the mobile devices, the Apple App Store has solicited premium profits in juxtaposition to other App vendors through the sale of applications that are higher priced and with high market performance. On New Year's Day 2017, Apple mobile device owners spent a record breaking $240 million on digital products (Bradshaw, 2017). Irrespective of the number of downloads, the Apple App Store games generate more revenue than Google downloads. Mobile app developers such as NimbleBit and Zach Gage have produced successful gaming apps such as *Pocket Planes*, *Ridiculous Fishing*, and *Spell Tower* on the iOS platform for years.

Mobile app developers pursue the iOS platform for paid game formats, and capitalize upon the low level entry and specialized touch control designs of the iPhone and the iPad. The Metal API component of the iOS platform and the iOS 10 SDK empower the mobile game app developer to create innovative categories for game titles and features. The iOS 10 also expands the color capabilities for the AVFoundation, Core Graphics, Metal, and Core Image frameworks without compromises to the game app performance (Apple, 2017b). The Heads-up display (HUD) provides a graphic overlay of the user statistics, such as scoring, the number of remaining attempts, current gamer attributes, and power ups. Further, the Bluetooth integration allows for the iPhone to act as a controller to play iPad games that support both iOS platforms. The control type for iOS  game designs include the simulated action button and control stick; flick gesture for tossing, pushing and pulling during gameplay; gyroscope and accelerometer for precise control; and the microphone game controller screen interface elements (Banga, 2011).

Apple and Google continue to dominate the mobile app platforms, along with industry leaders TenCent, Sony and Activision Blizzard. Pioneers of the industry, to include Sega, Disney, and Warner Bros who struggled to keep up in 2016 with revenues that kept them out of the top 10. Toward the close of 2016, Apple was rated number 5 and Google was rated number 8 in a list if the largest gaming producers according to revenue (Ballard, 2017). Apple generated over $4 billion in gaming revenues; while Google generated close to $3 billion[31]. Despite the aggressive pace of the Google Play Store downloads in the past year, Apple is earning much more from its apps. IHS Markit analyst, Jack Kent presented that Apple's mobile game app revenues have topped Google Play revenues as a result of the success of Apple in China, where the Google Android store has been blocked (Bradshaw, 2017). Others speculate that the

[31] Based upon the results of Newzoo Global Games Market Report. See Newzoo  (2017). *Top 25 Companies by Game Revenues.*

Google App Store subscribers are just not as willing to spend as the Apple App Store subscribers.

Efforts to compensate for the growing risk associated with market investments into the iPhone have focused upon driving up online revenues through Apple Music, iCloud, and the Apple App Store. Leading Apple mobile game app titles for 2016 included *Clash Royale*, which grossed over $2 million in daily revenues; *Pokemon Go,* which grossed approximately $1.6 million in daily revenues; and *Candy Crush Saga*, which grossed approximately $1.3 million in daily revenues in the United States[32]. Despite the market challenges of the iPhone, the Apple App Store revenues increased by 40% in 2016, much of which has been accredited to the success of *Super Mario Run* and *Pokemon Go* (Bradshaw, 2017). Market analysts have also noted that the success of Apple's virtual products is due to increased spending by existing product and service Apple users. IOS Mobile app developer profits also increased significantly in 2016. New subscription app pricing trade agreements allowed for the Apple mobile game app developer to retain 85% of the app sales in cases where the consumer subscribes to the app for more than one year (Bradshaw, 2017).

## 2.16 Chapter Summary

This review of literature has addressed several theories associated with gamification, mobile game app development, acquisition and monetization models, and common variables used for mobile game app success prediction. The fundamentals of predictive analysis, data mining, and an overview of the CRISP-DM model are

[32] Data drawn from Statista report for the top grossing mobile game apps for iPhone. See Statista, 2017d.

presented along with challenges associated with the implementation of the model, and how mapping between the levels of the model is driven by the contexts of the application domain, problem and technical aspects for present and future. lastly, decision modeling and the Apple App Store rankings against market competition and a snapshot of the mobile game app publisher industry were discussed. In Chapter 3, the methodology is presented, to include the data selection process and the variables to be used in the model through used the first step of the CRISP-DM model is to identify and clarify the business objective. The business objective for this study is to gain understanding of mobile game app features which achieve remarkable market performance through the extraction of data from a large dataset for analysis. This section will describe sample data and outline the method of data analysis for 50 mobile game apps that were extracted from a large pool of the Apple App Store mobile game titles in order to develop a prediction model that will identify the most successful mobile game features and the most significant user views in order to predict which game titles will be most successful in the market.

# Chapter 3. METHODOLOGY

The scope of mobile game app success is outlined by assessments of the most popular and highest revenue generating titles based upon comparisons of variables such as game category or genre, price, downloads, developer history, and the target markets. The process of game categorization is complex, and encompasses considerations for both game categories and genres which may combine, overlap, and conflict in definition (Grace, 2005; Askelof, 2013). Prior studies have modeled mobile game app success based upon specific game features and ARM techniques (Askelof, 2013; Moreira et al 2014; Filho et al 2014; Alomari et al 2016).

The methodology for this study is based upon the assumptions and structures of the ARM Funnel and CRISP-DM models for data analysis in an approach that is twofold, to include the utilization of two prediction models for game features and for user perspective (see fig. 10). The dual method is designed to utilize statistics in a unique approach to presenting the correlations between the game features and user behavior. The initial data is collected from large datasets of freemium iOS mobile game titles in mobileaction Top Charts; and is prepared, modeled and evaluated based upon the CRISP-DM general tasks[33]. The game app data is extracted, segmented and decision trees are created in order to create the game success prediction model. The user behavior data is mined in order to identify relationships between features such as cost and user retention and classified in order to construct the user behavior prediction

[33] See Mobileaction. 2017. *Top Charts for iPhone*. ASO Intelligence. The ASO Intelligence provides information for all types of mobile app titles in categories of Top Charts, New Apps, Biggest Movers, Biggest Losers, and Publisher Leaderboard. Mobile app data also includes estimated daily downloads, estimated daily generated revenues, category ranking and rating trends, audience geography, and visibility score histories.

model. The final output equations for the two datasets will be used to construct an algorithm or API for future game app performance prediction. The most prominent game and user features may then be combined to create an overall measure for mobile game success prediction. The framework for the methodology is presented in Figure 10:

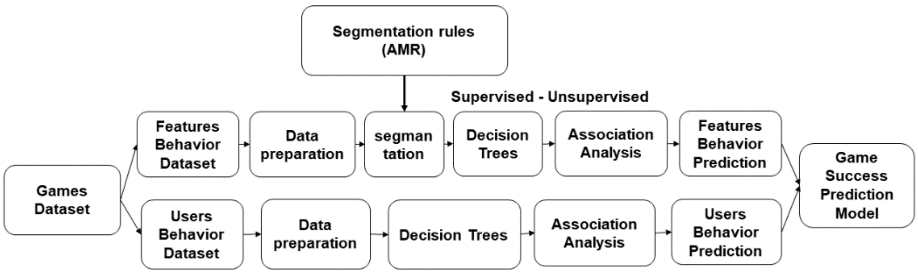



**Figure 10. Twofold Data Analysis Methodology**

The data for the Games dataset and User Behavior variables are presented in Appendices A, B and D. The data analysis for the mobile game features is presented in section 4.2, followed by an analysis of the user behavior variables in section 4.3, within the scope of the Game Genre, Features, and Developers dataset analysis.

### 3.1 Dataset Preparation, Configuration and Analysis

The research will use a total of 50 Apple App Store mobile game app titles which were selected from the listing of 500 top game apps based upon revenues. The method of selection was to extract 50 games from the highest, lowest, and average revenue performances from 500 top games. The final 50 title sample is comprised representations from 19 game categories and the Daily Active User, Average Revenue, Average Downloads, and Monthly Average Users for user behaviors. User perspectives which are constructed as two separate datasets extracted from mobileaction.com for game statistics between February 2016 and January 2017 (see Appendix A). A total of 29 game features were used for the analysis along with the number of downloads, revenues, and daily average usage statistics generated for each game. The outputs are then used to predict game success and user behavior; and the relationships between outputs for the two prediction models is discussed. The first data set consists of the game data drawn from the mobileaction freemium game database of United States-based freemium games from the Apple App Store. The United States market was selected as the focus of this study is the iOS freemium games in the Apple App Store, which is an American company. analysis.  The second dataset is comprised of user behaviors for the selected iOS platform games.

The preparation, configuration, and analysis of the game features is performed in the order of data preparation, segmentation, decision trees, and the final prediction model. The game features and user behavior datasets are prepared, mined for text, features extracted and classified based upon CRISP-DM, antecedent and consequent assessments are made, and lastly, the proposed prediction model is presented. The dual data analysis integrates the data mining and extractions for the features and user behavior variables to perform algorithmic analyses. Analytical results are retrieved from Top Down Induction Decision Trees (TDIDT algorithms) as illustrations of the most significant data relationships and distinctions between classes are produced.

The data was scraped from the mobileaction.com, cleaned and trained based upon the steps of the CRISP-DM model, using the BigML. In order to process the data using BigML supervised machine learning tools, the data which was stored on excel files was converted to CSV files in Google spreadsheets. The uploaded CSV files were uploaded as sources which were used to create datasets that could be selected for the different types of analytics. The *features* and *user behavior* datasets were classified based upon the type, count, number of missing values, errors, and a histogram that underscores the values.

The data is filtered to exclude any fields or entries that could damage the results. In BigML, the filter may be performed by a manual configuration using the

*FilterDataset* feature. Here, the upper and lower bound limits of the filter are entered manually as 'between' values *x* and *y*. The goal that is set for the dataset is the prediction, which is displayed in the BigML *Objective Field*. The predictive model may then be generated manually or by the *1-Click Model* feature. The decision trees that were generated in BigML are presented. The results of the data analysis and proposed model are then presented as decision trees by default for further analysis and discussion.

The prediction models for the game features and user behaviors were presented as decision trees in fig. 17 and fig. 18. The evaluation of the model is achieved through an 80% training dataset and 20% test dataset. An evaluation is run with predictions from the test set inputs. The evaluation is presented in BigML as a benchmark between the model, baselines, and the random predictions. The model performance in comparison to the mean is computed using the R-squared method. The mean absolute percentage error may be decreased by the addition of more features and user behaviors or by simply adding more titles.

### 3.2 Game Genres

The game genres for the 50 app sample were selected from the 19 category listings for the Apple App Store[34]. The selection of game categories for the mobile game apps may be approached according to classification theory for mobile games which dates back to the emergence of the mobile game app era. Wolf (2001) provided a model of 42 game categories based upon the level of interactivity and game play. Grace (2005) differentiated between game categories and game genres in that the game category provides a description of the type of game play; while the game genre describes the narrative context. Askelof (2013) differentiated between game title classifications based upon national culture, citing differences in platforms, genres, and devices between Western and

[34] See 2016. *iTunes Preview*. Apple.

Japanese mobile markets. Lee et al (2014) presented a facet foci method of game genre classification based upon a diversity of facets to include Gameplay, Style, Theme, Audience, Target, Spatial Setting, Mood or Affect.

The 19 categories reflect the genres that are listed in the Apple App Store. Many of the mobile game titles selected for this study fall within the scope of more than one game category and genre, or could be classified as ‘either, or’. Table 2 shows the category listings and the number of game titles from the sample that are represented for each genre:

**Table 2. Sample Data Mobile Game App Genre Classification**

| | Categories | # | Sample Game App Titles |
|---|---|---|---|
| **1** | Action | 6 | 8 Ball Pool; Mortal Kombat X; King Rabbit; Cally’s Cave 3; Dynasty Warriors Unleashed; Rogue Runner |
| **2** | Adventure | 7 | Temple Run 2; The Sims: Free Play; The Walking Dead: Road to Survival; Plants & Zombies 2; Kim Kardashian: Hollywood; Fire Emblem Heroes; Two Dots |
| **3** | Arcade | 4 | Super Mario Run; Pet Rescue Saga; Toy Blast; Genies & Gems |
| **4** | Card | 2 | World Series of Poker; Zynga Poker: Texas Holdem; |
| **5** | Casino | 3 | Double Down Casino; Slotomania Slots; Big Fish Casino |
| **6** | Dice | 1 | Boggle with Friends |
| **7** | Family | 1 | Farmville: Tropic Escape |
| **8** | Kids | 4 | Angry Birds Blast; Crossy Road; Minecraft Pocket; Beneath the Lighthouse |
| **9** | Music | 1 | Piano Tiles 2 |
| **10** | Puzzle | 5 | Candy Crush Saga; Cookie Jam; Panda Pop; Bubble Witch 3 Saga; Gummy Drop |
| **11** | Racing | 3 | CSR Racing 2; Sonic All Star Racing; Asphalt 8: Airborne |
| **12** | Role Playing | 2 | Game of War - Fire Age; Summoners of War |
| **13** | Simulation | 2 | The Simpsons: Tapped Out; Pokemon Go |
| **14** | Sports | 3 | Madden NFL Mobile; MLB Tap Sports Baseball 2017; Super Stickman Golf 3 |
| **15** | Strategy | 5 | Boom Beach; Clash Royale; Clash of Clans; Mobile Strike; Clash of Kings |
| **16** | Word | 1 | Word Cookies |
| | Total | 50 | |

The classifications for the mobile title genre were derived from the categories provided by mobileaction, Think Gaming, and the Apple App Store. The majority of the Adventure, Strategy, and Action games are interchangeable between genres.

However, for the purposes of the modeling and regression analysis, each title is listed under one category based upon the classification from the Apple App Store.

### 3.2.1 Action

The Action category consists of games which characterized by action intensity as the focus and which require reflex response skills to excel (Grace, 2005). The Action games primarily include Shooter games, Stealth games, and Sports. The action may be in the form of fighting, fair competition, adventure. Popular freemium Action mobile game apps for iOS include *Mortal Kombat X*; *Vector 2*, and *Injustice: Gods Among Us*.

### 3.2.2 Adventure

The Adventure category consists of games which require problem solving and exploration based upon virtual storylines or storylines which are replicas of real world events (Grace, 2005). The skills required for Adventure games include reasoning and creativity and depend upon player performativity (Lee et al., 2014). Popular freemium Adventure mobile game apps for iOS include *The Walking Dead: Road to Survival*, *Fire Emblem Heroes*, and *Plants & Zombies*.

### 3.2.3 Arcade

Arcade mobile game apps are designed according to traditional arcade and video console games with similar engagement strategies as RPGs, Family and Trivia games (Apsalar, 2013). The objective may be for points, rewards, higher levels, or additional playtime; however, the Arcade games do not successfully promote in-app purchases. Traditional Arcade games that are available in mobile game format include *PAC-MAN Lite*, *Genies & Gems*, and *Super Mario Run.*

### 3.2.4 Board

The Board games category consists of apps that require a virtual board design and often replicate traditional board games. Board games available as mobile game apps in the Apple App Store include *4 in a Row*, *Monopoly*, and *Checkers*.

### 3.2.5 Card

The Card game category of games are based upon traditional card games such as *Solitaire*, *Poker*, *Spades* and *Hearts*, or may consist of games which require the players to draw from a deck of cards during gameplay. Many Card games also include gambling features, or are based upon Casino card games and offer virtual or real currency rewards. Card games may be played in single player or multiplayer mode.

### 3.2.6 Casino

The Casino games are characterized by gambling elements found in traditional card games, bingo, casino roulette wheels, and slot machines. The Casino game environments are often designed for play in simulations of real world casinos and gambling events. The games may be played for virtual currency or real currency. Popular Casino mobile game apps include *Texas Holdem Poker*, *Heart of Vegas Slots*, *Slotomania*, and *Double Down Casino*.

### 3.2.7 Dice

Dice games include the rolling of die as a part of the gameplay and often replicate traditional dice games. The Dice games may also have gambling features or be replicas of real world casino dice games to be played for fun or for real money. Popular Apple App Store games which feature dice include *YAHTZEE with Buddies: The Classic Dice Game*, *Backgammon Free*, and *Lucky Roulette*.

### 3.2.8 Educational

The Educational games are based upon learning platforms that promote creativity as well as intuitive thinking and problem solving. Educational games may target the infant market, children, and young adults from kindergarten to college with a wide range of subjects from mathematics to physics. The Apple App Store contains an extensive list of educational mobile game app titles to include *Pictorial*, *LEGO Juniors Create and Cruise*, and *Laugh and Learn Shapes & Colors Music Show for Baby*.

### 3.2.9 Family

The Family mobile games are designed for play by groups and are appropriate for both adults and children. The characteristics of Family games such *Farmville: Tropic Escape* and *Deal or No Deal* include multiplayers, family oriented themes, and may be classified in a diversity of game genres.

### 3.2.10 Kids

The Kids games are designed for children under 13 years of age, are simple to play with engaging graphics and content that is approved for minors. The Kids games may also be educational with musical and color themes, and include titles such as *Kids Doodle*, *Where's My Mickey?* and *Cake Bites Maker*.

### 3.2.11 Music

The Music category consists of mobile games with musical and aerobic themes and strategies that require high engagement and longer daily session times. Popular Music game titles include *DuckTales Remastered*, *Piano Tiles 2*, *Song Pop 2: Guess the Song*, *Guitar Hero Live*, *Just Dance Controller,* and *Beat MP3 2.0 - Rhythm Game*.

### 3.2.12 Puzzle

The Puzzle game is characterized by enigma, problem solving, manipulation and navigation (Wolf, 2001). The Puzzle game titles are often designed for single player mode and may or may not be connected to online servers. Apple App Store Puzzle game titles include *Candy Crush Saga*, *Cookie Jam*, *Mahjong Journey,* and *Sudoku*.

### 3.2.13 Racing

The Driving or Racing games provide the player with a diversity of vehicle types in order to engage in the primary action of competitive racing (Lee et al., 2014). Some driving games, such as *Grand Theft Auto*, are in city settings or on roadways rather than race tracks and include additional activities other than driving. Popular racing games include *CSR Racing*, *Asphalt 8: Airborne*, and *Real Racing 3*.

### 3.2.14 Role Playing

Role Playing games (RPGs) consist of a storyline, rich narratives, technical character management, and fantasy. The objectives of the RPG may require long periods of play (Grace, 2005). Lee et al. (2014) contributed that role playing games are characterized by valuations and changes in level and power which is reflected in the avatar characteristics of each player. Popular RPG games on iOS include *Jade Empire: Special Edition*, *The Warlock of Firetop Mountain*, and *Game of War - Fire Age*.

### 3.2.15 Simulation

Simulations provide a demonstration of the real world to include reenactments of social situations in either virtual or simulated real world settings. The simulations are used in Role Playing games and may have extensive game lives that engage the user for long periods of daily gameplay. Monetization strategies include in-app purchases for avatars, items, and objects that are used to design the game space. Lee et al. (2014) supported that simulation games are designed to mimic the activities of the physical world. Combat and Racing game categories are presented against simulation of real world locations and monuments. Popular Simulation games available in iOS include The Sims: Free Play, *Order Up!! To Go*, and *Tiny Tower*.

### 3.2.16 Sports

The Sports games are characterized by replicas of traditional sports games and may be mobile editions of successful video console and PC game titles, and from real world professional sports leagues. Most Sports game are also classified as Action and

Strategy games and provide the players with multiplayer, player-versus-player, and player-versus computer mode options. Popular Sports titles in the Apple App Store include *Madden NFL Mobile*, *Basketball Stars*, *Flick Homerun Free Version*, and *Super KO Boxing*.

### 3.2.17 Strategy

The Strategy games require logic and problem solving and some may present little or no storyline. The player typically must collect, process, and interpret information through the game interface as interventions to achieve desired outcomes (Lee et al., 2014). Apple App Store Strategy game titles include *Clash Royale*, *World of Tanks Blitz*, and *Star Wars Galaxy of Heroes*. A 2013 study showed that Family, Adventure and Strategy game categories solicit the longest average lengths of daily gameplay sessions, which correlate strongly with in-app purchases[35].

### 3.2.18 Trivia

Trivia based games are knowledge testers which are characterized by question and answer or quick quiz designs based upon a diversity of themes from history to movies and culture. Trivia themes are also often used in educational mobile game apps for target audiences of all ages. Popular Trivia mobile game app titles include *Trivia Crack*, *Logos Quiz*, and *Guess the Movie*.

### 3.2.19 Word

The Word games consist of strategies that involve word formation or guessing and may be mobile editions of traditional Word board games. Word games are also

---

[35] See Big Data Lab study of millions of mobile game apps in Apsalar Analytics. 2013. *Top Mobile Game Categories by In-app & Engagement*.

popular for educational mobile game app designs. Popular Word games available in mobile app formats include *Scrabble, Words with Friends*, *AbbleDabble,* and *Boggle*.

The most recurring mobile game app categories for the sample were Action, Adventure, Puzzles and Strategy games. The less popular genres included Music, Family, and Word games. Many of the mobile game titles may be classified in more than category or genre. Many of the Action games are also Strategy Games. Many of the Board games could also be classified as Arcade games. Table 3 shows the general statistics for the game genres:

**Table 3. Game Genre Statistics**

| 50 Game Sample General Statistics | | | |
|---|---|---|---|
| **Mean** | Standard Deviation | Variance | Population |
| **3.125** | 1.8929 | 3.35 | 3.583 |

The genre statistics included measures of mean, variance between the genres, and the standard deviation for the sample population

The game titles are distributed randomly across the sample. The mean indicates that the most recurring categories have approximately 2 to 4 titles. The standard deviation between categories is 1.8929 with a variance of 3.35, which indicates that there is a fair representation of most of the 19 game categories. Figure 11 illustrates the distribution of the sample across genres:

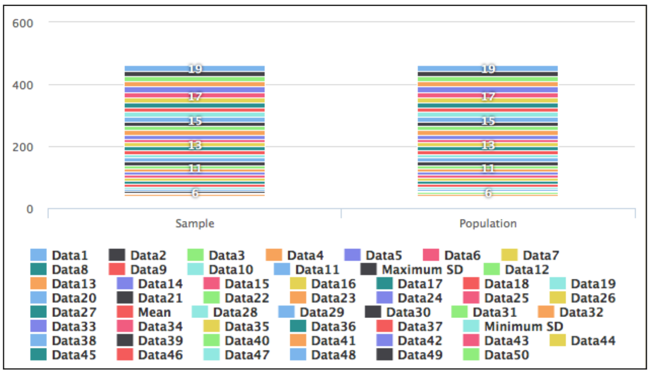



**Figure 11. Sample Game Genre Distribution**

The light blue data strata represents titles which are classified in the Kids category of games. The black data strata represents titles which are classified in the Racing category of games. The yellow data strata represents titles which are classified in the Adventure category of games. The dark green data strata represents titles which are classified in the Puzzle category of games. The red data strata represents titles which are classified as Action.

Only 2 Role Playing games were included in the sample. Of the 19 categories, none of the 50 game titles in the sample contained titles from the Trivia, Board, or Educational categories. However, some of the game titles that could have been classified in more than one category may have been listed in one of these categories, such as Word games, Puzzles, and Family games.

### 3.3 Game Features

The pool of mobile game app features is expanding at a rapid pace, along with expansion of mobile device and software capabilities. The game features are typically designed in order for the developer to realize all of the components of the ARM model in some capacity (Askelöf, 2013; Filho et al., 2014; Moreira et al., 2014). A diversity of game feature combinations have been used to produce successful mobile game apps based upon a focus upon acquisition, retention or monetization, or with a focus upon all of the components in the design of the mobile game app.

A limited number of studies are available that have defined the scope of the game features and how the success of the mobile game app on the market could be associated with such features. Koetsier (2017) found that mobile game app players value game title reviews and ratings; but that the mobile game app description is now the most powerful determinant for whether the user will download the application[36]. In order to differentiate from prior studies of mobile game success prediction, the game

---

[36] Based upon a 2016 survey of 3,005 smartphone users by Tune. See Koetsier (2017). *How mobile users make app install decisions on Google Play and the App Store*. Acquisition and Engagement.

feature variables for this study were selected to represent variables with both strong and weak market outcomes extracted from across several prior studies, to include the research methods and outcomes for Askelöf (2013), Filho et al. (2014) and Moreira et al. (2014). Since these studies were conducted, market trends for freemium games have continued to change as social network features and in-game purchases expand within higher level ARM strategies. However, the fundamental approaches and prediction models remain valid as snapshots of the regional markets at the time the studies were conducted.

The aforementioned studies consist of different combinations of game features used in mobile game apps for different markets. Askelöf (2013) conducted a comparative analysis of 12 social network games for Asian and Western markets, and analyzed the game titles based upon social, mechanics, rewards, time-dependent events, and monetization aspects. Out of 37 game features Moreira et al. (2014) singled out 11 features as significant to download and grossing ranks: Invite Friends; Random Elements; Customizable; Event Offer; Soft Currency Gambling; Item and Status Upgrades; IAP; Achievement; and Mobage. Out of 37 game features, Filho et al. (2014) singled out 12 features as most impactful: Timed Boost; Soft and Hard Currency; IAP Potential; Leaderboard; Time Skips; Hard Currency Gambling; Request Friend Help; Customizable; Consumable; Facebook; and Levels.

Approximately 40 mobile game app features were researched from prior studies; and from the Apple App Store and Google Play Store mobile game app descriptions. On the basis of a selection of features from the pool of game features and categories from the aforementioned studies, 29 game features were selected for this study, and were grouped into the following 6 categories: Social Network and Social Interaction; Offers; Virtual Currency and Purchases; Play Accelerators; Reward Retention and Punish

Absence; and Game Features. Table 4 presents the 29 game features which were selected for this study by category[37]:

**Table 4. Game Features**

| Social Network/Social Interaction | Offers | Virtual Currency/Gambling | Play Alterations | Reward Retention/ Punish Absence | Game Features |
|---|---|---|---|---|---|
| Facebook | Unique Offer | Soft Currency | Time Skip | Gambling Reward Retention | Customizable |
| Invite Friends | Event Offer | Hard Currency | Time Boost | Cumulative Reward | Power-ups |
| Request Friend Help | Daily Offer | IAP | | Non-Cumulative Reward | Skill Tree |
| Line Chat | | | | Achievement | Unlock Content |
| Single Play | | | | | Item Upgrade |
| Versus | | | | | Status Upgrade |
| Competitive Play | | | | | Random Elements |
| Leaderboard | | | | | Levels |
| Cooperative Play | | | | | |

A total of 29 features are analyzed which have been divided into 6 categories.

New game category descriptions have also emerged in the industry, to include lightweight games, Mobile eSports, Alternate Reality, RPG Card, Building Simulation and Brain Puzzles (Hwong, 2016a). The new classifications will provide industry outlooks in future studies of the mobile game app prediction[38].

The selection of game features is challenging for the developer as the pool of game features is ever-increasing and the formula for the most successful combination of features is dependent upon several factors. The game genre or category is a significant

---

[37] Fujita et al (2016) described the grouping of logically related components in order to extract information with value as relative to the choice of information organization. The 29 game features for this study were grouped based upon the functionality relative to the ARM Funnel strategy for mobile game app development and marketing.

[38] See Hwong, C. 2016a Verto Analytics' 2016 Audience Measurement Data study of mobile game app categories and the demographic for players from each category. The sample consisted of American mobile game subscribers 18+.

factor, as the target audience differs in intellectual capacity, skill, interests, and access to currency to make in-app purchases. The player will increase ability as the time spent playing the game increases; therefore, the features and degree of challenge for each game must reflect consideration for such scenarios.

### 3.3.1 Social Network/Social Interaction[39]

The social aspect of the game title has the capacity to strengthen the interpersonal relationships between players, which affects the ARM components in a diversity of ways. The social network games provide avenues to acquire new players through invites, sharing, and chats; to retain players through the ongoing social relationships; and to achieve monetization by in-app purchase offerings which solidify, heighten and prolong the multiplayer gaming experiences (Zichermann & Cunningham, 2011). Behavioral changes are realized within the target market through the achievement awards and other motivational offerings.

#### 3.3.1.1 Facebook

The Facebook App and Game Groups feature allows for the integration of Facebook Groups with the mobile game apps. Freemium games that are often shared on social networks include *Candy Crush Saga*, *Farmville: Tropic Escape*, and *Criminal Case*. The game players may be grouped for sharing through Facebook, which is ideal for games with alliances, VIPs, teams and clans through the API and dialogs. The user may create groups, join, view other members of the group and post. Users may also log in to a mobile game app using a social media application and share gaming

---

[39] Bogost (2014) described four asynchronous gameplay features for social network games: 1. sequential, non-tandem play between mutiplayers; 2. the requirement of a persistent shared gaming environment; 3. breaks in game play for organizing; and 4. asynchronous play which may or may not be the defining element of the game.

achievements on leaderboards. Gold coins may be earned by *Angry Birds Blast* players for logging into Facebook. The Facebook feature was selected for this study due to the popularity of the social application, the degree to which Facebook adds a social element to the gaming experience, and how often the social network is used by developers as a monetization strategy, as well as to acquire and to retain new users.

#### 3.3.1.2 Invite Friends

The Invite Friends feature is a user acquisition strategy for mobile game apps that exploits the capacity of social features to force interactions with other users. Top group play mobile game apps with the Invite Friends feature include *King of Opera, Plunder Pirates*, *Party Doodle*, *Who Can't Draw?*, and *Evil Apples*. The Invite Friends button is embedded in the mobile game app, which enables the user to invite, challenge, or refer friends through Facebook, Email, Twitter, SMS and iMessage. Invites are made during breaks which occur between players are organizing tenets of asynchronous play in which the player actions create multifaceted timelines of the game play and the breaks in game play (Willson & Leaver, 2016). The down time in social network game play is often also the time within which the players socialize or communicate.

Commonly, the player may also gain some benefit or reward for inviting friends, and the friends who accept the invitations may also earn rewards. However, Willson and Leaver (2016) pointed out that player requests for gifts or other items may be beneficial to both parties of the exchange; however, the gifts are not always needed, desired, or used by the recipient and no direct consequent upon game play can be established. The Invite Friends feature was selected due to the large number of mobile game developers who incorporate this feature and as an engagement component of the full suite of SNG features.

#### 3.3.1.3 Request Friend Help

The Request Friend Help feature is used to prompt interactions between users. Players help each other through exchanges of extra lives or items. The degree to which request for help features require persistence of the shared gaming environment has been a subject of debate. The player may request the help of a friend in order to advance to another level in the game; however, the duration of the gaps in game times fluctuates whether the players are engaged in separate or the same game space (Willson & Leaver, 2016). The request occurs during game “downtime”, a period in which the asynchronousness of the social network game is most identifiable. Retention is increased through extended gameplay times; while acquisition rates are increased through the networking nature of the feature. This feature was selected as a measure against the more primary social mobile gaming features such as Facebook and Invite Friends, and to contrast against the Request Friend Help feature.

#### 3.3.1.4 Line Chat

The line chat is described as an “oriental network” with a diversity of supporting mobile games and services (Moreira et al., 2014). Developers have increasingly turned to the line chat feature based upon the success of messaging apps among mobile device users. Like Facebook and Mobage, the line chat is a social feature, which promotes the developer’s acquisition, retention, and monetization strategies. WeChat and the Line messaging app accumulated billions in revenue when introduced to the market, which became a platform for social and multiplayer games. *Madden NFL Mobile* allows up to 32 users to join friends to play or to chat. The line chat feature was selected for this study due to a high association of the chatbot with the ARM monetization component for iOS and Google Play games.

### 3.3.1.5 Single Play

The single play feature is commonly level-based and provides for only one player to participate in a session. The single play feature was selected in order to assess the significance of the social features which provide participation by more than one player per session. Players who prefer playing against the computer to playing against friends subscribe to mobile games with the single player mode feature. Mobile game apps that have single player modes include *Scrabble*, *Clash of Clans*, *Boom Beach*, and *Asphalt 8: Airborne*. This feature was selected in order to gain insight into its effect on user behavior for mobile game apps which are not compatible with online servers; and therefore, cannot offer multiplayer sessions.

### 3.3.1.6 Versus

The versus mode encourages two-player interactions through competition and the use of leaderboards. The Versus mode may be for player-versus-player, multiplayer-versus-multiplayer, or player-versus-computer. Askelöf (2013) supported that the multiplayer features serve as social motivators through cooperation or competition; which in turn, increase the number of daily users for game title. *Minecraft Pocket Edition* provides for two single player versus modes of Survival or Creative. This feature was also selected in order to gain insight into its effect on user behavior for mobile game apps which are not compatible with online servers; and therefore, cannot offer multiplayer sessions.

### 3.3.1.7 Competitive Play

Competitive play games foster social interaction and are common features of mobile game apps (Moreira et al., 2014). In contrast to cooperative game designs, competitive play consists of a struggle between players to obtain the highest achievement during gameplay.  The competitive instincts of the user are exploited through skills and task completion requirements which are tracked and scored. The complexity of the game increasing with each advancement, causing winning or success to become more difficult for the player. The user associates personal identity with the status or rank achieved in the game (Zichermann & Cunningham, 2011).

#### 3.3.1.8 Cooperative Play

Cooperative play is integrated into the game mechanics through cooperative modes and cooperative campaigns which allow the players to use couch, online and split screen play. Multiple players are connected to participate collectively in a game session for optimal social gaming experiences. *Minecraft Pocket Edition*, *Call of Mini: Infinity*, *Crossy Road*, and *Summoners of War* are mobile game apps which offer cooperative play modes. The *Farmville* series of games provide natural cycles in which the players experience nurturing and growth through ongoing engagement. The cooperative play feature was selected as a measure of the success of multiplayer mobile game app platforms for multiple users across WiFi or online as well as how cooperative play contributes to the successfulness of the game in the social networking game markets.

#### 3.3.1.9 Leaderboard

The leaderboard is the score keeping feature in strategic and competitive games[40]. The leaderboard has been associated with Bartle's (1996) Killer player profile and is typically found in social networking games with leaderboards. The scores of all of the players are managed through a central system, which may also display

[40] Ganguly (2016) contributed that the players may be used to market the mobile game app through sharing of Leaderboards, levels that have been bypassed, earned achievements and Avatars, hidden items that have been found, and items that have been purchased in the App Store.

comparisons of the players' scores or performance. The developer is challenged to implement a leaderboard for real time action for scenarios in which millions of players accumulate scores simultaneously. *King Rabbit* provides Game Center leaderboards and achievement lists which the players may follow through Facebook, Twitter, Instagram and Raresloth. *Genies & Gems* allows the players to compete on the leaderboard to save the animals in arcade competitions that are held each week. This feature was selected as a fair assessment of user behaviors in regard to social network games.

### 3.3.2 Offers

In-game offers are significant marketing and monetization strategies which developers use to market the mobile game app, to encourage the user to return to the game, and as monetization strategies by encouraging the user to make in-app purchases. The offers may appear as push button notifications or as advertisements within the game. The object of the offers is to engage the user and to persuade the user to make purchases.

#### 3.3.2.1 Unique Offer

Unique Offers are typically one time offers for items or gifts which must be redeemed within a short span of time. The user is persuaded that the offer is a special opportunity which is not available to all players and must be used within a limited period of time. The user may receive an usually large discount or special free gift by acting within the promotional period. A Unique offer may be for the user to download a pay-to-play game for free for a limited time. This feature was selected due to its value as an indicator of user purchasing behaviors.

#### 3.3.2.2 Event Offer

The Event Offer is presented on holidays or other commemorative dates in order to gain more attention and increase the probability of user response. Event dates such as Christmas or the student Spring Break may be used as the reason for an offer that is actually also presented on other dates. *The Walking Dead: Road to Survival* includes live events which give the player opportunities to acquire valuable resources and new teammates. The event offer feature was also selected due to its value as an indicator of user purchasing behaviors and as a monetization strategy for freemium games.

#### 3.3.2.3 Daily Offer

The daily offer is typically a small discount offer that is presented to the user each day in order to transition the freemium game user to the paynium model (Moreira et al., 2014). The objective is to remain connected with the user through communication each day, and to encourage the player to return to the game. The Star Bonus was added as a *Clash of Clans* daily offer in 2016 as a retention strategy. The Star Bonus must be completed within 24 hours, after which a new Star Bonus is generated. In the event that the player receives 5 Star Bonuses or more, additional loot is added to the player's treasury. *Pokemon Go* offers Stardust and XP as Daily Bonuses for the first Pokemon catch and PokeStop spin of each day. For players who earn the Daily Bonus for 7 consecutive days, or a Pokemon streak, a larger bonus is rewarded. The daily offer feature was selected for the study due to the highly traceability of the user activities that is enabled. The degree to which the user views and accepts the offers provides data for analysis of user behaviors.

### 3.3.3 Virtual Currency/Purchases

The player may purchase currency to use within the game space in order to enhance the game, purchase gifts for friends, or to earn currency and rewards in gambling-based platforms. The primary types of types of currency for mobile game apps are soft and hard currencies which may be purchased with real currency in the app store.

#### 3.3.3.1 Soft Currency

The soft currency model for free-to-play games is earned by gameplay that may be earned at a more rapid rate. The gamer may start the games without paying any real money upfront. As the game progresses, offerings for hard or premium currency are made to improve the game experience by added functionalities, the addition of items, or the unlocking of content. The players will typically use the currency provided within the gambling games to start, and once the currency runs out, the player must purchase more using real currency to keep playing. *Slotomania* and *Big Fish Casino* developers have realized hundreds of millions of dollars in consistent revenues by tapping into the global casino market with freemium Social Casino mobile game apps. *Hearthstone* offers a currency which the player earns by completing the daily quest; and the currency may then be used as real currency. This feature was selected due to prior findings that the use of one soft currency within the freemium game model promotes repeated trips to the in-game app store and may also contribute significantly to conversion optimization strategies. Therefore, soft currency can serve as a viable indicator of user engagement and purchasing behaviors.

#### 3.3.3.2 Hard Currency

The hard currency model for free-to-play games is designed to provide currency by in-app purchases. Moreira et al. (2014) found that the hard currency provides linear gains throughout gameplay and is more controllable by the developer. Hard currency may typically be used to accelerate the progression of the game. Some mobile game apps such as *Farmville*, *Cityville*, Clash *of Clans* and *Snoopy Candy Town* may offer both soft and hard currency. Similar to the monetization strategy for *Clash of Clans*, *Snoopy Candy Town* uses Beagle Bucks for hard currency, and gold for soft currency with varying conversion rates. This feature was selected as a measure of user engagement, user purchasing behaviors, and a contrast of purchasing behaviors between soft and hard currency within the same mobile game app.

#### 3.3.3.3 In-app Purchases

The IAP feature targets gaming whales who are more proned to spend continuously on virtual products and privileges inside the mobile game app. When users purchase gifts for their friends, the gifts motivate the recipient to also login to the gaming environment, and for some to also purchase gifts in return. A 2013 study showed that Trivia, Strategy, and Adventure mobile game categories draw the highest percentage of in-app purchases, followed by Role Playing and family game categories. The items that may be purchased inside the app may range in price from a few cents for avatar clothing; to $100.00 for a nanotech sniper rifle offered in *Contract Killer* or the MIL-3000 Assault Rifle offered in *Deer Hunter: Reloaded;* to a premium price of $600 for an Apathy Bear Gun offered by *Gun Bros*. IAPs appeared in iOS freemium mobile game apps in 2009; and in 2011, Google introduced IAP functionality for Android mobile game apps. The in-app purchase feature was selected as a strong indicator of user behavior and the effectiveness of the developer's monetization strategy for the mobile game title.

### 3.3.4 Play Alterations

The rate of speed during gameplay depends upon the user's device, the speed of the Internet or WiFi connection, and the speed characteristics that are built into the game mechanics by the developer. Play alterations give the player more control over the speed and progression of the game. The duration of the game may be extended, the speed of the game may be accelerated, and certain events may be skipped with play alteration features. *CSR Racing 2* inlcudes speed alteration features such as quick speed jumps by tapping. The play alteration features are significant for mobile game app platforms, as the user may play the game while on the go and will need to pause the game, or otherwise control the progression of the game frequently.

#### 3.3.4.1 Time Skip

The Time Skip feature allows the player to race forward in the progression of the game by speeding up the game's timed mechanics. The play may be accelerated to bypass session limits by instant completions, object finishing, and energy purchases (Askelöf, 2013). Many of the tasks and actions in *The Sims: Free Play* take long periods of time to execute; therefore, the LifeStyle Points may be used to speed up time consuming actions. Commander skills can be used to significantly reduce times for research, building and training using Skill Points in *Mobile Strike*. The time skip feature was selected for this study as an assessment measure of the player demographic or psychology as well as purchase behaviors.

#### 3.3.4.2 Time Boost

The Time Boost feature gives the player an advantage in the game with less physical effort and typically may require specific actions or an in-app purchase. The player is engaged for specific increments of time and may retain players with short attention spans who would otherwise quit the game because of a lack of control or because of the slow pace (Moreira et al., 2014). The game session may also be expanded if the user purchases additional content. *Kim Kardashian: Hollywood* requires constant "energy boosts" which may be earned after five minutes of play, clicking objects, or by reaching higher levels. The boosts significantly speed up tasks such as the 1-hour photo shoot. NBA Digital introduced *NBA InPlay* for mobile devices in 2016 with real time play and “turbo” boosts that can increase points within a specific time frame. The turbo boosts were designed to increase the player engagement and to increase the multiplier effect during market campaigns. The time boost feature was also selected as an assessment measure of the player demographic or psychology, and purchase behaviors (NBA, 2016).

### 3.3.5 Reward Retention/Punish Absence

The psychology of the freemium game model is designed to reward the user for continuous play and for inviting others to join the game. Zichermann and Cunningham (2011) supported that the primary metrics of engagement are a focus upon frequency, recency, virality, duration and ratings. Online communities are now formed through which the users login in multiple times each day to review posts, leaderboards, and offers, as well as to socialize in the forums. Based upon the nature of the game, some metrics may be more significant to the developer than others. Filho et al. (2014) described the punishing absence feature as a game mechanic which subtracts value from the player in some way for failing to consistently return to gameplay. The reward and punishment strategies solicit desired user behaviors in terms of duration, recency, and frequency. When the user is logged out of a game, a notification of pending rewards is sent out to encourage the user to engage. The user may also be punished for absences from the game by losing points or forfeiting eligibility for certain rewards.

#### 3.3.5.1 Gambling Reward Retention

Long periods of gameplay are ideal for Casino game models, as the longer the user engages, typically the more money will be spent. Gambling reward retention strategies reward the user randomly for regular revisits to the game based upon pre-established time intervals (Moreira et al 2014). The reward may be cumulative or non-cumulative. The absence of the reward punishes the user for absence. The gambling reward retention feature was selected in order to measure the impact of gambling content on the game success and to identify purchase behavior patterns.

### 3.3.5.2 Cumulative Reward Retention

The cumulative reward is used to draw the player back to the game each day by reward provisions which increase each day, if the user continues to return to the game. The cumulative reward may be associated with the collection-based ecosystem as well as the gambling platform (Zichermann & Cunningham, 2011). The daily login bonus is a cumulative reward that is used based upon psychological theory of positive reinforcement. *Clash Royale* offers rewards for returning players every four hours by issuing push notifications of a waiting gift chest. The cumulative reward was selected for this study due to the significant potential to retain players according to the ARM Funnel.

### 3.3.5.3 Non-cumulative Reward Retention

Similar to the cumulative reward, the non-cumulative reward is issued each time the player returns to the game; however, the reward amount is the same and does not increase with each new engagement. The rewards may still accumulate each time the user logs on to the game. Casino games and apps with gambling elements such as *Slotomania* and *Big Fish Casino* offer the players both cumulative and noncumulative rewards for returning to the game each day. Badges are common rewards used to engage the user and may rewarded for creating accounts and posting comments, as well as for completing tasks during gameplay. Bonuses are also common types of non-cumulative rewards which may be awarded as one-time credit for the completion of specific activities. This feature was selected in order to observe the contrast in impact between cumulative and non-cumulative rewards on mobile game success.

#### 3.3.5.4 Achievements

The achievement feature provides a bonus reward to the player when a task is completed during gameplay. Moreira et al (2014) associated the achievement feature with retention through an extension of the game lifecycle. *Candy Crush Saga* offers achievements for making candy and stripe/wrap combinations and matching color bombs; and displays the achievements in the gamer profile. *Clash Royale* provides achievements for joining clans, helping friends, reaching a road to glory arena, collecting cards, and for watching a TV Royale match. *Pokemon Go* rewards the player with several achievements on each level, to include the Jogger, Backpacker, Depot Agent, Ruin Maniac, and Pikachu Fan. The achievement feature was selected for this study in order to weigh the degree to which the title remains successful on the market as a result of an infinite number of offerings for levels and potential achievements.

### 3.3.6 In-Game Features

In-game features provide the players with options to enhance gameplay in freemium game apps through in-app purchases. Developers use a diversity of approaches to differentiate their games from similar games by adding creative features with unconventional names and functions. The selection of in-game features can make the difference in how much revenue the developer can realize for each game. Nizan (2014) supported that items which give the player an advantage propels a significant percentage of mobile game app revenues. Zichermann and Cunningham (2011) contributed that virtual objects are prerequisites to the development of a collection-based virtual ecosystem and in order to add value to the objects, scarcity must be present.

Freemium game developers are especially dependent upon the selection of successful combinations of in-game features, as the features are designed to encourage in-app purchases. Dependent upon the nature of game play, the developer can offer a diversity of game features that continuously generate revenues by in-app purchases ranging from $0.99 to hundreds of dollars. Due to the structure of the freemium model, the mobile game app in-game features must be purchased in order for the game to make any substantial amount of money. Simulation and Role-Playing games generate revenues by in-app purchases of customizations, avatars, and other items that expand the capacity to play the game. Casino games generate revenues by in-app purchases of soft and hard currency for game access and gambling.

#### 3.3.6.1 Customizable

The customizable feature allows the player to change the appearance of the characters and the gaming environment which differentiate the user's game presentation from the presentation of other users. *The Sims: Free Play* offers extensive customization features from the characters' gender, ethnicity, and facial features to custom home floor plans and furnishings. The *Pokemon Go* training avatars may be customized in the profile screen. Moreira et al. (2014) supported that the customizable feature does not add significant value to the mobile game app. However, this feature was selected in order to measure whether the trend has changed, or if the customizable feature is weak as a contribution to the mobile game app success on the market.

#### 3.3.6.2 Power-up

Power-ups provide the player with increasing capabilities through the completion of a task or through in-app purchases. Power-ups occur in *Angry Birds Blast* when the player clears balloons or makes matches for large items such as the Laser Gun, Rocket or Bomb and can be combined to create more power to remove tiles. *Pokemon Go* includes IAPs for power-ups, which are used to power-up the Pokemon and evolve with high CPs. The player may use Stardust and Candy received from catching and storing

Pokemon to power-up, which increases the CP and pushes up the white bar for the Pokemon Level. The power-up feature was selected for this study due to its significance as a component of a majority of the game categories; particularly for Strategy and Adventure game genres.

#### 3.3.6.3 Item upgrade

The item upgrade provides the player with the capacity to increase the level of power or the significance of equipment items such as weapons, armor, and jewelry. The item upgrade may be earned by the completion of tasks, through in-app purchases, or as gifts from other players. Typically, in social games that are highly competitive, it is highly probable that the players will continue to purchase item upgrades to retain their status in the game. Further, the complexity of the game may be enhanced by enabling the users to trade item (Zichermann & Cunningham, 2011). Some items may be upgraded several times, while others may be limited to one upgrade per session. *The Walking Dead: Road to Survival* team upgrade feature allows the player to customize the Survivors with items in order to gain skills, weapons, and power. *Plants vs Zombies: Garden Warfare 2* allows the player to upgrade Ammo, Super Meters, and Ice as well as intangible items such as Player Health, Speed, Damage and Toxic Aura. The item upgrade feature was selected for this study as valid indicator of user purchasing behaviors and the popularity of Action, adventure, and Strategy games.

#### 3.3.6.4 Status Upgrade

The status upgrade feature allows the player to upgrade the status of characters within the game. The upgrade may extend the life of the character or increase the number of capabilities (Moreira et al 2014). *Fire Emblem Heroes* allows the player to upgrade the status of the anime warriors to higher star ratings by summoning and advancing the character growth through the “unlock potential” feature. The characters may achieve up to 5 stars, and the player may upgrade the status of more than one character using Great Badges and Hero Feathers from the Training Tower or the Sixth Stratum. *Mortal Kombat X* allows the player to upgrade the team fighters through accumulated gameplay experience, teaching new attack modes, and skill improvements with the artifacts. In-app purchase requirements for the status upgrade increase the revenue generation potential for the mobile game app developer. The status upgrade feature was selected for the sample as it increases player engagement and retention; and therefore, may be directly associated with the ARM model.

#### 3.3.6.5 Skill Tree

The skill tree displays potential skills or capabilities in branches which become accessible to the player after the completion of prerequisites. Skill trees are commonly used in Massively Multiplayer Online Role Playing (MMORPG) and Multiplayer Online Battle Arena (MOBA) games from the Role Playing category. *Game of War: Fire Age* features a Hero Skill Tree with top prizes such as Troop Attack and Troop Training, Construction, and Research; and progresses to other skills on either side of the tree. The skills are purchased from the Hero Skill Tree using Skill Points. The commanders in *Mobile Strike* are assigned a skill tree which are unlocked as the commander advances to each level. The left side of the tree consists of trap and combat training; while the right side of the tree consists of construction speeds, research and resources. The Skill Tree was selected as a feature for this study as an indicator of user engagement and as a representative of the Role Playing games genre.

#### 3.3.6.6 Unlock Content

Mobile game developers use the unlock content feature to differentiate between new users and experienced gamers; for user retention by keeping the player engaged to achieve game objectives; and to create a sense of unveiling in the game experience. The unlock content features provides access to items, characters and locations that are otherwise unusable by the player. The player may unlock content as a reward for completed tasks or by making in-app purchases. *Mortal Kombat X* allows the player to unlock costumes and to unlock special content by connecting the mobile version with the console game. *Temple Run 2* allows the player to unlock characters using coins earned during gameplay. The unlock content feature was selected as a measure of user engagement and as a user retention strategy, as the introduction of new features and levels keep the user interested in the game and extend the life of the application.

#### 3.3.6.7 Versus

The Versus feature encourages two-player competitive play and is common in Sports and Trivia categories. Social relationships are established as players meet new friends to compete against and through sharing of performance and achievements on the leaderboards. *Madden NFL Mobile* features. Asynchronous Head-to-Head play in which the players are each granted a single drive possession. *Mobile Strike* include a State versus State Kill event during which time the player uses skills such as Troop Defense, Troop Health and equipment attack. The Versus feature was selected for this study as it contributes insight into user preferences and as a monetization strategy for competitive players.

#### 3.3.6.8 Levels

The mobile game app may feature a series of levels in which the complexity of the game increases along with the size of the rewards or achievements. The number of levels that may reached in one game increase steadily in order to retain long term players and to maintain an environment of never-ending play. Typically, the player acquires larger rewards and the characters gain capacity as the game level increases. *The*

*Sims Free Play* provides the player with the opportunity to earn new items after leveling up. *Slotomania* offers 7 Status levels, which the player earns by leveling up with Status Points for each purchase. By the close of 2016, *Candy Crush Saga* had introduced 2,000 levels in order to engage players who had become familiar with the game and reached the highest levels (Kelion, 2016). The levels feature was selected for this study in order to measure the impact of levels on long term success as a game-as-a-service and the impact on the success of Strategy games.

**3.3.6.9 Random Elements**

Random elements are game components or elements that are generated and appear within the game unsystematically to create suspense. The elements are erratic and independent of each other in order to reduce the number of identifiable patterns and the predictability of the game outcomes. Nizan (2014) supported that random elements add the potential to "get lucky" to mobile games, and is critical to game success by adding uncertainty and uniqueness, drama, game balance, and reductions in pay-to-win environments. Random elements also have the capacity to increase engagement by impulsive users who will likely respond to impromptu notifications, instructions, and offers. The developer may add random elements to the game by using built-in functions which insert random values or items with different probabilities into the game mechanics. Random elements may include greetings, attacks, playthroughs, and gift offerings. The *Pokemon Go* moves during evolution are randomized and may materialize in a large number of evolved movesets. As an alternative to randomness, the developer may use hidden information as a game mechanic for multiplayer games.

### 3.4 User Behavior

For this study, the user behavior will be assessed based upon the number of new installs, revenue generation, the number of daily and monthly active users, and recent visibility scores for the mobile app (see Appendix D). Player types are based upon goals which are complementary, aspiring to either the game content through action or interaction, or to game control through the players or the game environment. The assessments may then be used in the ARM Funnel model for acquisition, retention, and monetization based upon the user behavior.

Mobile game app game mechanics are coded in order to capitalize upon specific user demographics and behavioral tendencies (Zichermann & Cunningham, 2011). The mobile game app user behavior may be assessed through variables such as daily and monthly game revenues, the number of daily and monthly downloads, or the number of daily active users. The whales, or users who are willing and able to spend the most also require the longest conversion times from free gaming to spending real currency. Therefore, a further understanding of the user behavior may be gained through an evaluation of the types of mobile game app players and their motivations for playing certain types of games. Bartle (1996) categorized the gamer into a taxonomy of four player types: killers, achievers, explorers, and socializers. The objective of the Achievers is progress or mastery; the Explorers become engaged with the story, and become immersed in the game; Socializers yearn for contact with other people and participate in social network games and sharing. Other studies have also considered psychological theories for user behavior based upon the socioeconomic or demographic variables.

### 3.4.1 New Installs

The number of new installs provides an estimation of the mobile game popularity by the number of times it is accessed by new users in one day. The developer may increase the number of daily installs by using mobile app install campaigns to drive up downloads and by offerings that will reward users for game referrals and invites. Koetsier (2017) supported that mobile gamers select apps to install based upon screenshots, publisher, game reviews and ratings, the mobile game app description, and the app video or trailer[41]. The amount of information that is provided to the Apple App Store or Google Play about the game has a profound effect on the number of new installs. Conversion optimizations are more likely to occur when the game has differentiating description components than an engaging title and icon (Koetsier, 2017).

### 3.4.2 Revenue Generation

The freemium game model has been successful in the capacity to generate revenue by in-app purchases once the game has been installed. Developers of freemium games depend upon monetization strategies that will solicit ongoing in-app sales once the game has been downloaded.  Ganguly (2016) supported that revenue generation is achieved in Strategy games by early development of monetization strategies in the initial development phase, investments in Actionable Analytics in real time, and a focus upon Daily User Engagement for user retention and for virality campaigns. The market for mobile game app players who will spend time and real currency is highly

---

[41] According to the Tune study, approximately half of all mobile app new installs are based upon the user perceived need. The user may install based upon the recommendation of a friend; the appearance of the advertisement; special feature by Apple or Google; a top result from a search; according to a specific task; or based upon previous purchases from the same developer.

competitive, as the number of mobile game apps that are introduced each month is increasing rapidly.

### 3.4.3 Daily Active Users (DAUs)

The Daily Active Users (DAUs) is a metric that represents the number of users who actually use the mobile game app each day, rather than only the number of users that install the application or the number of sessions. Many times the user may download an application, but fail to install it. Others may install the game and delete it after only a few sessions. Offers and rewards, effective marketing campaigns, and competitive application models serve as retention strategies which help to increase the daily use of mobile games. Social networking features also increase the probability that the user will engage in game play each day.

### 3.4.4 Monthly Active Users (MAUs)

The monthly active users (MAUs) represents the number of user who have used the mobile app within a specified month, or 30 day period. The monthly active users is useful to compare with the daily active users in order to generate additional user behavior patterns for analysis. If a mobile game app is used 40,000 times by 20,000 players within a window of 30 days, the MAU would be 20,000. Metrics such as the MAU are useful to mobile app developers in estimating the costs of user acquisition.

### 3.4.5 Visibility Score

The visibility score represents the sum total of points allotted to the URL address by the search engines for the mobile app. Vogel (2014) supported that the visibility score is a critical online performance metric for business applications which also serves as an indicator of the application’s competitiveness. The visibility score is also associated with the position of the game title on the search engine results page (SERP) in terms of traffic generation and the potential revenues from user query rankings.

### 3.5 Chapter Summary

This methodology has addressed the first step of the CRISP-DM model is to identify and clarify the business objective, understanding of mobile game app features and user's behavior which achieves remarkable market performance, and present the dataset for each of the game features and user behavior where used in Chapter 4. In Chapter 4, the Data Analysis is presented, to include the data analysis was based upon a dual approach that integrates data mining for the mobile game app genres and features along with an analysis of the user behavior variables. The game title genres and developers are first analyzed in regard to frequency, rankings, revenues, and the sample anomalies. The game features are then analyzed in regard to the data distributions and field importance, sample anomalies, and model predictions. lastly, the user behaviors are also analyzed in regard to the data distributions and field importance, sample anomalies, and model predictions.

# Chapter 4. Data Analysis

The research performs data mining and data algorithmic analyses on the Information Mining Engineering process typically utilizes process models such as CRISP-DM, or KDD to complete information base management tasks. The analysis for this study is carried out in BigML, using the machine learning algorithms for regression analysis and prediction. WhizzML is used to implement the algorithms to simplify the analytic processes for large datasets. In addition to the general analysis, the WhizzML may also compute the prediction objective gradient as associated with the probabilities generated for each class.

The model ensemble is used to learn a number of models across different categories of the sample data, which mitigates the potential to overfit the data using a single model. The dual analysis of mobile game features and the user behavior variables was conducted in order to generate a feasible model for mobile game app success prediction. The mean absolute error assessing the accuracy of the predictions based upon the formula:

**Equation 5**

$$M - \frac{100}{n} \sum_{t=1}^{n} \frac{|A_t - F_t|}{|\quad A_t \quad|}$$

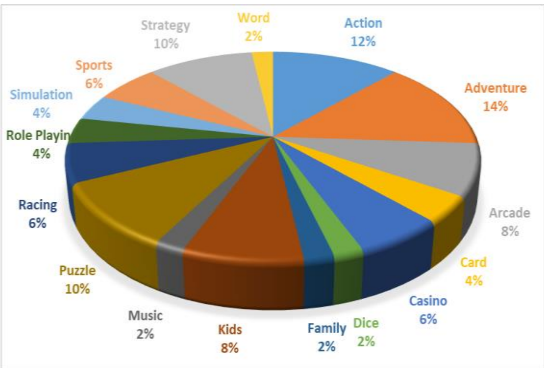
where $A_t$ represent the real value and $F_t$ represents the predicted value. The degree of certainty that accompanies the outcomes is predicated upon the expected error and confidence levels. The confidence and expected error are based upon the class distributions for each node in the decision tree and the number of instances in the sample population. The Wilson score interval may be used to obtain a 95% confidence level for the model predictions.

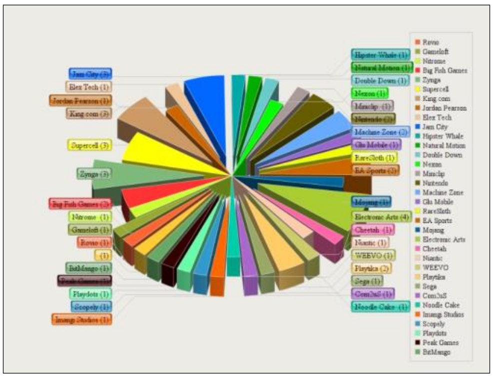
## 4.1 Games Dataset

The mobile game features and user behavior datasets were analyzed using the BigML Machine Learning platform for predictive power and pattern generation using custom algorithms[42]. Section 4.2.1 will present the outcomes for the game dataset genres and developers. Section 4.2.2 will present the outcomes for the game features analyses. Section 4.2.3 will present the outcomes for the user behavior variables.

### 4.1.1 Mobile Game Genres and Developers

The sample of game titles consisted of a combination of 19 mobile game app genres. Figures 12a and 12b illustrate the leading mobile game app developers and genres for the sample:



12a 12b

**Figure 12. 12a Game Genres and 12b Game Developers**

The leading game developers for the sample were Electronic Arts, Supercell, and King.com. The leading genres were Action, Adventure, Puzzles, and Strategy.

#### 4.1.1.1 Mobile Game Genre

Of the 19 genres, the average recurrence of one genre was 2 to 4 genres. The most recurring genre was Adventure, which was represented 7 times or approximately 14%, followed by 6 Action or approximately 12%, and 5 Strategy or approximately

[42] See BigML. 2017b.

10%. Based upon the research, the Adventure, Action and Strategy genres were highly correlated due to the fact that many of the game titles could be categorized interchangeably between the 3 genres. The least recurring genres were Trivia and Board both with 0 representation, and Word and Dice both with a representation of 1 or approximately 2%. The median genre for the sample was for Puzzles, which was represented 5 times or about 10% of the sample. The revenues for the 3 leading genres were calculated and compared in order to further establish valid rankings for the sample in Table 5:

**Table 5 . Comparison of Revenue Statistics for Genres**

| Adventure | Revenues | Action | Revenues | Strategy | Revenues |
|---|---|---|---|---|---|
| **Temple Run 2** | $2,528.00 | **8 Ball Pool** | $67,428.00 | **Boom Beach** | $322,786.00 |
| **The Sims: Free Play** | $28,798.00 | **Mortal Kombat X** | $16,757.00 | **Clash Royale** | $1,685,714.00 |
| **The Walking Dead: Road to Survival** | $37,219.00 | **King Rabbit** | $227,025.00 | **Clash of Clans** | $1,480,799.00 |
| **Plants & Zombies 2** | $5,600.00 | **Cally's Cave 3** | $13,744.00 | **Mobile Strike** | $384,772.00 |
| **Kim Kardashian: Hollywood** | $97,467.00 | **Dynasty Warriors Unleashed** | $97,467.00 | **Clash of Kings** | $29,189.00 |
| **Fire Emblem Heroes** | $29,630.00 | **Rogue Runner** | $7,387.00 | | |
| **Two Dots** | $26,213.00 | | | | |
| **Totals** | $227,455.00 | | $497,236 | | $3,903,260 |
| **Mean** | $32,493.57 | | $51,839.29 | | $780,652.00 |
| **Median** | $28,798.00 | | $42,092.50 | | $384,772 |
| **Standard Deviation** | $31,428.17 | | $84,413.11 | | $748,398.61 |

The order for the genre rankings changes based upon revenue statistics. Although the Adventure genre is the largest, the Strategy games have generated the most revenue and the highest mean for revenue. However, the Strategy games revenue also has the largest standard deviation. The median revenue for the Action games, $42092.50 is larger than the median for the Adventure games of $28798, and for Strategy games. The highest representation for genres based upon revenues is the Strategy games.

### 4.1.1.2 Mobile Game Sample Genre Anomalies

A total of 10 anomalies were detected for 50 instances and a forest size of 128.

Figure 13 shows the leading anomalies for the sample genres:

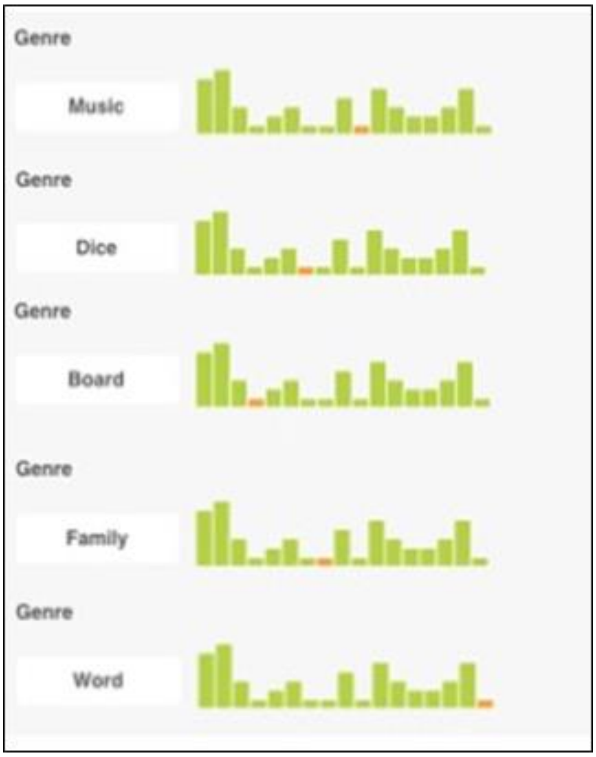



**Figure 13. Leading Sample Anomalies**

The unsupervised Anomaly Detection is a predictive model which identifies instances that are not a part of the regular data patterns. The anomaly detection is used in the data cleansing to highlight outliers. The model produced anomalies for a total of 50 of instances with the largest anomalous characteristic with anomaly scores.

The Anomaly Scores were generated for 10 combinations of the 50 instances in the Decision Forest. The scores were generated for each instance (input_data) for a score between 0 and 1 for each object with a field id and value. A Boolean type anomaly_status property enables the anomaly detector in the BigML HTTP status code. The closer to 1, the more distinct the value in juxtaposition to the other values for the dataset.

Of a total of 10 anomalies, the 5 leading anomalies were for the Dice, Word, Board, Music, and Family respectively (see fig.13). Listed in descending order, the predictive outcomes precluded that Dice reflected an Anomaly Score of 66.42%; the Word genre produced an Anomaly Score of 64.80%; Board Score 62.71%; Music 61.69%; and Family 61.37%.

### 4.1.1.3 Game Features Decision Tree

The mobile game genre data was used to generate a 128 instance decision tree forest. The algorithm for the compacted tree is based upon Reingold, Tilford & Buchheim. Figure 14 illustrates the decision tree and distributions for the 19 game app genres from the sample:

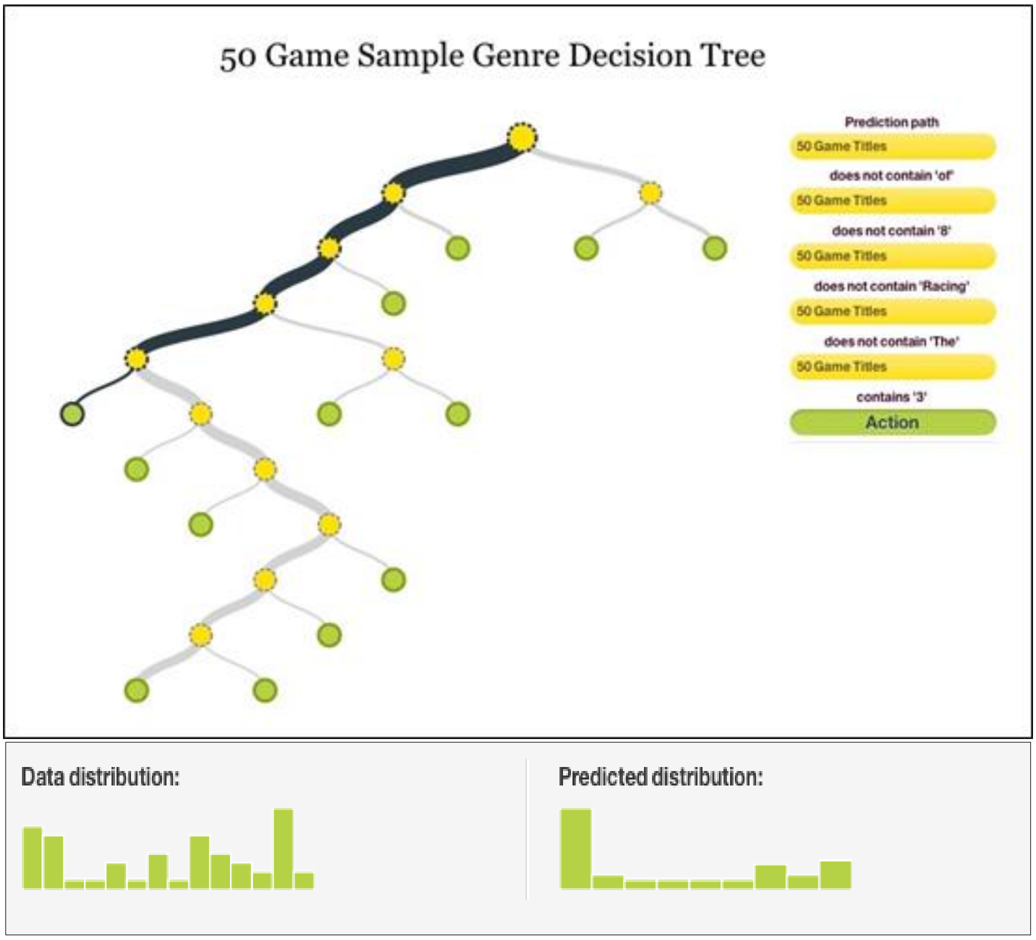



**Figure 14. Sample Genre Decision Tree**

For 50 instances, the Strategy genre produced a confidence level of 9.77%; from 39 instances, the Action genre produced a confidence level of 8.98%. From 25 instances, the Kids genre produced a confidence level of 6.40%. From 1 instance, the Casino genre produced a confidence level of 20.65%. From 7 instances, the Strategy genre produced a confidence level of 64.57%. The data distribution consisted of 9 instances of Strategy, 7 instances of Action, 6 instances of Adventure, and 6 instances of Racing. The Prediction Distribution consisted of 22 instances of Action, 8 instances of Strategy, and 7 instances of Racing[43].

[43] Also see M. Reingold J.Tilford 2001. Tidier drawings of trees. 223-228, 1981. Buchheim, Jünger, and Leipert (2006). Drawing rooted trees in linear time.

#### 4.1.1.4 Mobile Game Developers Revenue Comparison

A total of 34 developers were represented in the sample. The most recurring developers were Electronic Arts for 4, Supercell for 3, King.com for 3, Zynga for 3, and Jam City for 3, or approximately 0.9%.  The least recurring developers, those who occured only 1 time in the sample, were Elex Tech, Jordan Pearson, Nitrome, Gameloft, Rovio, Sega, WEEVO, Noodle Cake, Com2uS, Cheetah, Mojang, Raresloth, Glu Mobile, Nintendo, Miniclip, Nexon, Hipster Whale, Imangi Studios, Scopely, Playdots, Peak Games, and BitMango, all represented once in the sample. The median developers for the sample were Big Fish Games, Playtika, EA Sports, and Machine Zone, all of whom were represented 2 times in the sample.

### 4.1.2 Mobile Game Features

The analysis of the 29 sample game features consisted of statistical measurements for the leading game features based upon frequency, distribution across the sample, associations, and the outputs for the decision tree.

#### 4.1.2.1 Leading Game Features by Field Importance

The most impactful game features are identified in BigML by distributions, pruning and field importance rankings for the instances in the decision tree. Table 6 presents the data and prediction distributions for the sample game features:

Table 6. Data Distributions and Field Importance

| Data Distribution and Field Importance | 0 /no Feature | 1 / Yes Features |
|---|---|---|
| **Data Distribution** | 91.60%; 98 Instances | 80.40%; 402 Instances |
| **Prediction Distribution** | 91.60%; 98 Instances | 80.40%; 402 Instances |
| **Field Importance** | | |
| **Unique Offer** | 26.55% | |
| **Hard Currency** | 13.50% | |
| **Levels** | 12.01% | |
| **Time Boost** | 8.79% | |
| **Line Chat** | 6.11% | |
| **Item Upgrade** | 5.81% | |
| **Status Upgrade** | 5.67% | |
| **Skill Tree** | 4.54%% | |
| **Soft currency** | 4.19%% | |
| **Leaderboard** | 3.07%% | |
| **Gambling** | 1.83% | |
| **Cooperative Play** | 1.57% | |
| **Daily Offer** | 1.41%% | |
| **Request Friend Help** | 1.28% | |
| **Customizable** | 1.16% | |
| **Unlock Content** | 1.01%% | |
| **Non-cumulative** | 0.59% | |
| **Power-ups** | 0.53% | |
| **Competitive Play** | 0.39% | |

The data distribution and prediction distribution were both 91.60% or 98 instances for 0 or no feature and 80.40% or 402 instances for 1 or yes the game has this feature. The field importance provides an indicator of the game features that are the most significant per prediction and according to the indicators from the Boosted Tree. The highest field importance score was assigned to Unique Offer at 26.55%. The lowest field importance score was assigned to Competitive Play at 0.39%. The Field Importance is calculated based upon Breiman's Gini importance.

The field importance represents the relative contribution of the fields to the objective field. For the data training set, the objective field was chosen as the objective field. The higher the field importance of the feature, the higher the influence the feature will have on the predictions. The field importance is calculated in BigML through prediction error estimations for each field in order to decrease the splits in the decision trees (BigML, 2017b). The decision tree model is trained through recursive partitioning of the dataset, by splitting with nodes which have at least 2 instances. The splitting continues until the program reaches a defined stopping criteria.

#### 4.1.2.2 Game Feature Associations

The outputs for the Features Associations included the Antecedent, Consequent, Support, Confidence, Leverage and Lift. The rules for the association computations dictate that high Support features have high frequencies and high Confidence is an indicators of high predictive power. In absolute context, the support percentage is dependent upon factors such as the analysis subject-domain and the primary objectives of the research. The Associations tool in the BigML Analytics suite is a means for the identification of significant correlations between the values of the dataset. The rules or conditions may be configured and the data filtered based upon what outcomes are needed to achieve the aims of the research. The Association tool is cloud-based platform for Unsupervised Learning.

The Features Association analysis for the 50 mobile game app sample produced a total of 139 associations from the 29 game features dataset. Table 7 shows the outputs for 11 associations (see Appendix C for the complete list of game feature associations):

**Table 7. Mobile Game Feature Associations**

| Antecedent | Consequent | Coverage | Support | Confidence | Leverage | Lift |
|---|---|---|---|---|---|---|
| **Time skips = 0** | Timed boost = 0 | 34.69% | 34.69% | 100.00% | 21.95% | 2.7222 |
| **Timed boost = 0** | Time skips = 0 | 36.74% | 34.69% | 94.44% | 21.95% | 2.7222 |
| **Facebook = 1** | Leaderboard = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Invite friends = 1** | Facebook = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Leaderboard = 1** | Invite friends = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Invite friends = 1** | Leaderboard = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Facebook = 1** | Invite friends = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Leaderboard = 1** | Facebook = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Facebook = 1** | IAP <= 2 | 38.78% | 34.69% | 89.47% | 18.87% | 2.1921 |

The Life and Leverage outputs measure the associations between probability of a given rule occurrence (support(A→C) and the projected probability for features that are independent (coverage(A)*coverage(C)) from other features (BigML, 2017a). The Lift computation is a ratio of support and coverage as (Support(A→C)/(Coverage(A)*(Coverage(C)). The Leverage computation is the difference as (Support(A→C)/(Coverage(A)-(Coverage(C)). The Lift Associations are strong for features that appear less frequently; and the Leverage associations prioritize the features which occur with the highest frequency/support within the dataset. The Antecedent and the Consequent occur in sets 13.33% to 21.95% more frequently than as statistically independent components..

The Associations Discovery analysis is based upon the Association rules which produce measures of Support, Lift, Coverage, and Leverage based upon values for each variable as well as the variables of the dataset (BigML, 2017a). The method is underpinned by the filtered-top-$k^2$ method that was developed by Webb, and which is now a common tool for data mining, mortality rate analysis, and enhanced machine learning.

The associations which occur the most frequently have the highest support values. The Support may be expressed as the proportion of dataset instances which contain an itemset:

**Equation 6**

$$Support(item\ from\ dataset) = \frac{|instance\ \in\ D\ itemset\ \subseteq\ instance|}{N}$$

Coverage dictates the number of times the association rule may be applied. The Confidence level is an indicator of how high the predictive power is for a specific association based upon the analysis subject-domain. In respect to the mobile game apps, the Associations feature reflects the probability that one in-app purchase will be accommodated by another.

The Leverage reflects the difference between the rule probability and the anticipated probability in cases where the items are measured independently. Lift measures the number of times the antecedent and consequent occur as a pair more often than if the measure was statistically independent. A high Lift with low Support would indicate that a strong relationship exists between the features, but within the dataset, the frequency is low:

**Equation 7**

$$Lift(\mathrm{A} \rightarrow \mathrm{C}) = \frac{Support(\mathrm{A} \rightarrow \mathrm{C})}{Support(A)\ x\ Support(C)}$$

In order to perform the calculations, the game features data was normalized, or structured in a horizontal data layout as a binary representation of the game features data; where 0 represents an absence of the game feature, and 1 indicates that the game feature is present. The numerical values are discretized into categories fields, or classes. The higher value rules are prioritized using Leverage and the complementary items are avoided.

The game feature Associations showed simultaneous antecedent and consequent occurrences in a range between 7.50% and 21.95%. The lowest occurrence was between Gambling Reward and Daily Offer, which yielded 91.8370% coverage and 91.8370% support, 7.4970% leverage and 1.0089 lift, with 100% confidence.

The highest occurrence was between Time Skips and Time Boost, with 34.6940% coverage, 34.6940% support, 21.9490 leverage and 2.7222 lift with 100% confidence. High occurrences were found for the association between Facebook and Leaderboard, Facebook and Invite Friends, and Leaderboard and Invite Friends all with 38.7760% coverage, 36.7350% support, 21.6990% leverage and 2.4432 lift with 94.7370% confidence. Also, high occurrences were found for IAP and Versus, IAP and Line Chat, Achievements and Skill Tree, and between Gambling and Levels (see Appendix C).

The Time Skip feature appeared as an antecedent or a consequent in 29 instances of the game feature dataset and was associated primarily with the Time Boost and Cumulative game features. The Facebook feature appeared as an antecedent or a consequent in 86 instances of the game feature dataset and was frequently associated with the Invite Friends, Leaderboard, In-App Purchases, and Time Boost game features.

The Gambling Reward feature appeared as an antecedent or a consequent in 36 instances of the game feature dataset and was primarily associated with the In-App Purchase, Levels, Rewards and Daily Offers. The Versus feature appeared as an antecedent or a consequent in 24 instances of the dataset and was primarily associated with Invite Friends, Non-Cumulative, and Line Chat. Both the Leaderboard and Line Chat features appeared as an antecedent or a consequent in 21 instances of the game feature dataset. The Leaderboard was primarily associated with the Invite Friends, Time Boost, and Time Skip features. The Line Chat feature was primarily associated with In-App Purchases, Versus, Competitive Play and Invite Friends features**.**

The Cooperative Play feature appeared as an antecedent or a consequent in 15 instances of the game feature dataset, and was primarily associated with Cumulative and Time Skips. The Skill Tree feature appeared as an antecedent or a consequent in 12 instances of the game feature dataset, and was primarily associated with Status and Item upgrades and Cumulative Rewards.

The In-App purchase appeared as an antecedent or a consequent in 12 instances of the game feature dataset, and was primarily associated with Gambling and Facebook. Competitive Play also appeared in 12 instances and the primary associations were the Line Chat and Versus features.

The associations were further categorized into direct relationships based upon the number of recurrences with fields within the same groups. The majority of the game feature dataset clustered into relationships based upon the type of game feature. Figure 15 illustrates the relationship categorizations that were generated by Associations Discovery:

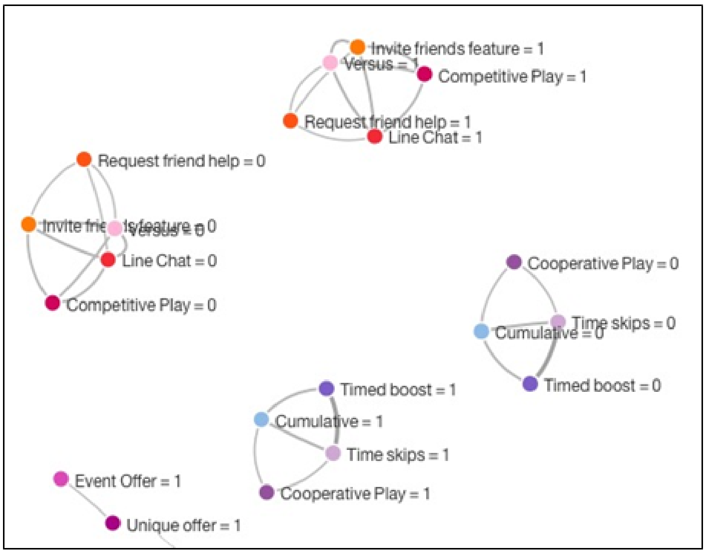



**Figure 15. Game Feature Dataset Relationships**

The Association Discovery generated 139 associations were generated for the mobile game features, with sample dataset settings of 8.21% leverage, 8.1630% minimum support, and 58.8240% confidence. The Offers generated clustered relationships, while other game features created clustered relationships between Time Skips, Time Boost, Cooperative Play, and Cumulative Play. The Social Network features generated a cluster relationship with Invite Friends, Line Chat, and Request Friend Help along with the Competitive Play game feature.

### 4.1.2.3 Sample Game Feature Decision Tree

The permanent and incidental parameters for the decision tree may be customized to replace some of the default values for the internal algorithms. Nodes with less than 1% of the total number of instances are pruned by either Smart, Statistical, or No Statistical Pruning functions. The sums of the split error reductions for the fields of the decision tree are normalized so that the total is equal 1 (BigML, 2017b). The decision tree confidence levels that are estimated by the BigML model are derived from considerations for the terminal node purity and the number of instances within the node.

A high purity assures that the base probability estimates are reliable. The game features data was used to generate a 128 instance decision tree forest in BigML, which produces outcomes that are more definitive than the outcomes of the decision tree.

The pessimistic confidence level of the model predictions reflects the certainty of predictions of each class at the node level as opposed to a completely random guess. A positive prediction represents approximately 80% probability, depending upon the sample (BigML, 2017c). The lower bound Wilson score interval is used to balance the prediction against elements of uncertainty within a smaller number of instances:

**Equation 8**

$$(\mathrm{w}-, \mathrm{w}+) \equiv \frac{\mathrm{p} + z2_{\alpha/2}}{2n} \pm \sqrt[\frac{z\alpha}{2}]{\mathrm{P}(1-\mathrm{P})/\mathrm{n} + \mathrm{z}^2{}_{\alpha/2}/4\mathrm{n}^2/\{1 + \frac{z^2{}_{\alpha/2}}{n}\}}$$

where probability P of the population of size *n* is generated with the highest confidence relative to normal distribution. The number of instances at the node and z = 1.96 is calculated as quantile 1-α/2 of the normal distribution for an error of α = 5%. With the exception of p = 0.5, the Wilson interval is defined as asymmetric. Further, when using the lower bound Wilson score interval, the variability between the predictions and the true state will be lower bound, or less probable than the true prediction. The node expected average squared error is calculated as:

**Equation 9**

$$e^{-} = \frac{\sum_i (vi - x)^2}{n}$$

where $x$ depicts the output node, $v_i \ldots v_n$ depicts the remaining values in each node, and $n$ depicts the number of instances. Figure 16 illustrates the decision tree for the 29 mobile game app features from the sample with the turquoise Levels node as the root:

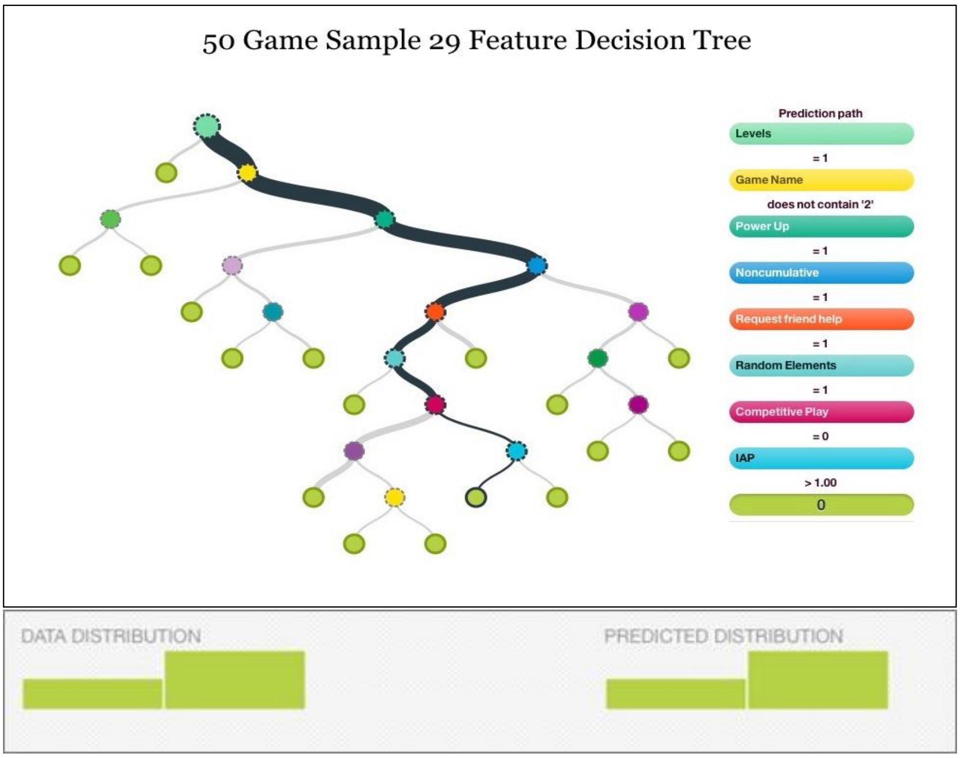



**Figure 16. Game Feature Decision Tree**

The green end nodes represent the final prediction for each prediction path. The decision tree produced a confidence level of 60.66% for Unlock Content, 66.49% for Noncumulative, and 25.05% for Daily Offer. The confidence level for all of the instances in the sample was 51.51%. The statistics for the game features decision tree will be discussed in Chapter 5.

The decision tree *stree* streaming algorithm is used to support the dataset, multi-machine distributions or multi-core parallelism and the Anytime algorithm. The *stree* algorithm grows the decision tree with each addition iteration as the node threshold sets the boundaries. The *mtree* algorithm grows the decision tree according to each split.

Sampling may be achieved randomly or deterministically. For this study, the samples were analyzed using both methods. Due to the relatively small sample size of the 50 game titles, the replacement option was not used; therefore, one instance in the dataset could not be selected more than once. The rows of the dataset are shuffled deterministically, linearly and randomly. When the data is shuffled randomly, the shuffle will produce a different combination for each time the dataset is trained. The deterministic shuffle shuffles the rows of the dataset in the same manner each time the dataset is trained.

The pessimistic decision tree prediction model that is generated in BigML follows the confidence and error calculations for classical prediction models (BigML, 2017c). The 1-Click model utilizes default values for configuration, or the configurations may be made manually to create a model from a dataset or a cluster. Data sampling options for the model included adjustments for a 100% sampling rate and the default sampling range. The 100% sampling rate produces a frequency of the instances that are extracted to include all of the instances in the sample. The dataset subset of instances is comprised in the sampling range.

In the final processes, the decision trees are compacted; however, the data may still be filtered. The histogram is provided to illustrate the instance count and the accompanying values so that the predicted value distribution, the potential range for the objective variables, and the skew of the error to either side of the prediction may be determined. Further, individual datasets may be created from one branch of the tree for more detailed analytics.

### 4.1.3 Game Features Prediction Model Evaluation

The Accuracy measures represents the number of predictions that are correct over the total number of instances in the evaluation for the classification model. The Recall score indicates the amount of positives that are actually being discovered within the classification model. The Precision metric reflects how closely the positives were identified with accuracy. A high precision indicates all of the positives were identified correctly, while a low Precision score indicates that some of the positives are false. The F-Measure represents the balanced harmonic mean of the Recall and the Precision scores with equal weighting. The PHi Coefficient is a metric for the True Negatives. Figures 17 and 18 illustrate the Actual versus Predicted and Predicted versus Actual Confusion Matrices, for the Game Features dataset:

TP FN FP TN

| ACTUAL VS. PREDICTED | 0 | 1 | ACTUAL | RECALL | F | Phi |
|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 33.33% | 0.40 | 0.24 |
| 1 | 1 | 7 | 8 | 87.50% | 0.82 | 0.24 |
| PREDICTED | 2 | 9 | 11 | 60.42% AVG. RECALL | 0.61 AVG. F | 0.24 AVG. Phi |
| PRECISION | 50.00% | 77.78% | 63.89% AVG. PRECISION | 72.73% ACCURACY | | |

**Figure 17. Confusion Matrix Actual vs Prediction**

The Recall for the Actual versus Predicted is 33.33% for Binary 0, and 87.50% for Binary 1, with an Average Recall of 60.42%. And the Precision is 50.00% for Binary 0, and 77.78% for Binary 1, with an Average Precision of 63.89%. and Average Accuracy are 72.73%. The Average F score is 0.61 and Average PHi is 0.24. Key: TP = True Positive; FN = False Negative; FP = False Positive; TN = True Negative.

TP FN FP TN

| PREDICTED VS. ACTUAL | 0 | 1 | PREDICTED | PRECISION | F | Phi |
|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 2 | 50.00% | 0.40 | 0.24 |
| 1 | 2 | 7 | 9 | 77.78% | 0.82 | 0.24 |
| ACTUAL | 3 | 8 | 11 | 63.89% AVG. PRECISION | 0.61 AVG. F | 0.24 AVG. Phi |
| RECALL | 33.33% | 87.50% | 60.42% AVG. RECALL | 72.73% ACCURACY | | |

**Figure 18. Confusion Matrix Prediction vs Actual**

The Precision for the Predicted versus Actual is 50.00% for Binary 0 and 77.78% for Binary 1 with an Average Precision of 63.89%. And the Recall is 33.33% for Binary 0, and 87.50% for Binary 1, with an Average Recall of 60.42%. And Average Accuracy are 72.73%. The Average F score is 0.61 and Average PHi is 0.24.

Histograms with the instances in the decision tree are provided along with values to create estimations of the potential ranges for objective variables and the degree of skewness. The data is partitioned in BigML in order to maximize the gains from each type of value with minimal mean squared error[2]. The partitions then form hierarchies with related predicates. For the categories, a candidate split score is generated for each category, with predicate fieldX==“some_category”, and binary numeric splits with predicate fieldX >= 42 (BigML, 2017c). The training set split was 80%/20% for the games feature dataset using random and deterministic sampling and ordering. The recall model, mode and random values for all classes of the game features model were 66.67%, 100%, and 33.33%, respectively.

### 4.2 Users Behavior Dataset

The variables that were used to measure the user behavior are the Daily Active User, Average Revenue, Average Downloads, and Monthly Average Users.

#### 4.2.1 User Behavior Statistics

A total of 50 instances and a forest size of 128. Table 8 shows the user behavior variables Statistics:

**Table 8. Game User Behavior Statistics**

| User Variable | Daily Active User | New Installs | Revenue |
|---|---|---|---|
| **Minimum** | 9277 | 4078.00 | 1286 |
| **Mean** | 982625.22 | 43529.02 | 185816.84 |
| **Median** | 251559 | 32144.00 | 57677 |
| **Maximum** | 9442801 | 276957.00 | 1685714 |
| **Standard Deviation** | 1706254.75 | 47312.76 | 363605.84 |
| **Kurtosis** | 11.44 | 11.34 | 8.10 |
| **Skewness** | 3.15 | 3.17 | 2.99 |

The kurtosis and skewness metrics provide extended characterizations for the game features of the dataset beyond the variability and location. The kurtosis provides distinction as either light-tailed or heavy-tailed as relative to the distribution and is formulated for univariate data as kurtosis = $\frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^4/N}{s^4}$ . The high kurtosis and heavier tail may be an indicator of an outlier; low kurtosis will reflect a light tail; while the uniform distribution would be the most unconventional arrangement of the data. The skewness measures the degree of symmetry where the Fisher-Pearson univariate data skewness is formulated g1 = $\frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^3/N}{s^3}$ (NIST, 2017).

#### 4.2.2 Sample User Behavior Decision Tree

Figure 19 illustrates the decision tree for the sample user behaviors:

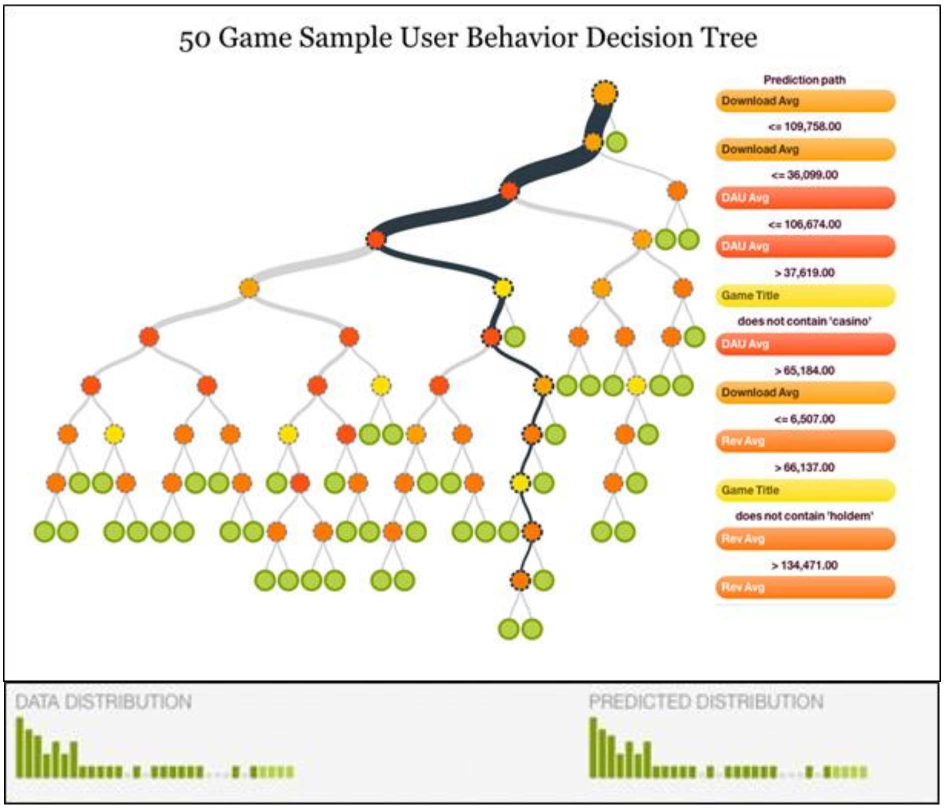



**Figure 19. 50 Game Sample User Behavior Decision Tree**

The user behavior decision tree generated prediction paths for the Daily Average User, Monthly Average User, Downloads, and Revenue.

### 4.2.3 Mobile Game App Game User Behavior Associations

A total of 34 associations were generated for the user behavior dataset, with 11.69% leverage, 15.3850% minimum support, and 72.7270% confidence. The values used for the Association Discovery rules are the same as the rules for the games feature dataset associations. Table 9 shows the user behavior association values:

**Table 9. Associations for the User Behavior Dataset**

| Antecedent | Consequent | Coverage | Support | Confidence | Leverage | Lift |
|---|---|---|---|---|---|---|
| **39038 < MAU Avg <= 90738** | 13232 < DAU Avg <= 28991 | 21.15 % | 21.15 % | 100.00% | 16.68 % | 4.7273 |
| **13232 < DAU Avg <= 28991** | 39038 < MAU Avg <= 90738 | 21.15 % | 21.15 % | 100.00% | 16.68 % | 4.7273 |
| **DAU Avg <= 13232** | MAU Avg <= 39038 | 19.23 % | 19.23 % | 100.00% | 15.53 % | 5.2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **MAU Avg <= 39038** | DAU Avg <= 13232 | 19.23% | 19.23% | 100.00% | 15.53% | 5.2 |
| **DAU Avg > 164638** | Download Avg > 12031 | 19.23% | 19.23% | 100.00% | 15.53% | 5.2 |
| **Download Avg > 12031** | DAU Avg > 164638 | 19.23% | 19.23% | 100.00% | 15.53% | 5.2 |
| **Download Avg > 12031** | MAU Avg > 556357 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **MAU Avg > 556357** | Download Avg > 12031 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **61096 < DAU Avg <= 164638** | 208588 < MAU Avg <= 556357 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **4929 < Download Avg <= 12031** | 61096 < DAU Avg <= 164638 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **MAU Avg > 556357** | DAU Avg > 164638 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **208588 < MAU Avg <= 556357** | 4929 < Download Avg <= 12031 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **DAU Avg > 164638** | MAU Avg > 556357 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **208588 < MAU Avg <= 556357** | 61096 < DAU Avg <= 164638 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **61096 < DAU Avg <= 164638** | 4929 < Download Avg <= 12031 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **4929 < Download Avg <= 12031** | 208588 < MAU Avg <= 556357 | 19.23% | 17.31% | 90.00% | 13.61% | 4.68 |
| **Download Avg <= 1103** | DAU Avg <= 13232 | 21.15% | 17.31% | 81.82% | 13.24% | 4.2546 |
| **DAU Avg <= 13232** | Download Avg <= 1103 | 19.23% | 17.31% | 90.00% | 13.24% | 4.2546 |
| **Download Avg <= 1103** | MAU Avg <= 39038 | 21.15% | 17.31% | 81.82% | 13.24% | 4.2546 |
| **MAU Avg <= 39038** | Download Avg <= 1103 | 19.23% | 17.31% | 90.00% | 13.24% | 4.2546 |
| **28991 < DAU Avg <= 61096** | 90738 < MAU Avg <= 208588 | 17.31% | 15.39% | 88.89% | 12.39% | 5.1358 |
| **90738 < MAU Avg <= 208588** | 28991 < DAU Avg <= 61096 | 17.31% | 15.39% | 88.89% | 12.39% | 5.1358 |
| **90738 < MAU Avg <= 208588** | 2593 < Download Avg <= 4929 | 17.31% | 15.39% | 88.89% | 12.06% | 4.6222 |
| **2593 < Download Avg <= 4929** | 28991 < DAU Avg <= 61096 | 19.23% | 15.39% | 80.00% | 12.06% | 4.6222 |
| **2593 < Download Avg <= 4929** | 90738 < MAU Avg <= 208588 | 19.23% | 15.39% | 80.00% | 12.06% | 4.6222 |
| **28991 < DAU Avg <= 61096** | 2593 < Download Avg <= 4929 | 17.31% | 15.39% | 88.89% | 12.06% | 4.6222 |
| **39038 < MAU Avg <= 90738** | 1103 < Download Avg <= 2593 | 21.15% | 15.39% | 72.73% | 11.72% | 4.202 |
| **1103 < Download Avg <= 2593** | 39038 < MAU Avg <= 90738 | 17.31% | 15.39% | 88.89% | 11.72% | 4.202 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **13232 < DAU Avg <= 28991** | 1103 < Download Avg <= 2593 | 21.15 % | 15.39 % | 72.73 % | 11.72 % | 4.202 |
| **1103 < Download Avg <= 2593** | 13232 < DAU Avg <= 28991 | 17.31 % | 15.39 % | 88.89 % | 11.72 % | 4.202 |
| **Rev Avg > 274819** | DAU Avg > 164638 | 19.23 % | 15.39 % | 80.00 % | 11.69 % | 4.16 |
| **DAU Avg > 164638** | Rev Avg > 274819 | 19.23 % | 15.39 % | 80.00 % | 11.69 % | 4.16 |
| **Download Avg > 12031** | Rev Avg > 274819 | 19.23 % | 15.39 % | 80.00 % | 11.69 % | 4.16 |
| **Rev Avg > 274819** | Download Avg > 12031 | 19.23 % | 15.39 % | 80.00 % | 11.69 % | 4.16 |

The frequency range for the Antecedent and Consequence occurring simultaneously rather than statistically independent was between 16.68% and 11.69%. for the user variables[44]

The Daily Active User behavior appeared as an antecedent or a consequent in 18 instances of the user behavior dataset and was associated primarily with the Monthly Average Users and Downloads variables. The Monthly Average Users behavior appeared as an antecedent or a consequent in 17 instances of the user behavior dataset and was primarily associated with the Daily Average Users and the Downloads user behavior variables.

The Average Downloads User Behavior appeared as an antecedent or a consequent in 22 instances of the user behavior dataset and was primarily associated with the Daily Average User, Monthly Average User and Revenue behaviors. The Average Revenue User Behavior appeared as an antecedent or a consequent in 3 instances of the game feature dataset and was primarily associated with the Daily Average Revenue and Download behaviors. Figure 20 shows the categorization of the 34 associations that were generated for the user behavior dataset:

[44] See Wallis, S. Binomial confidence intervals and contingency tests: mathematical fundamentals and the evaluation of alternative methods. *Journal of Qualitative Linguistics*, 20(3), 2013.

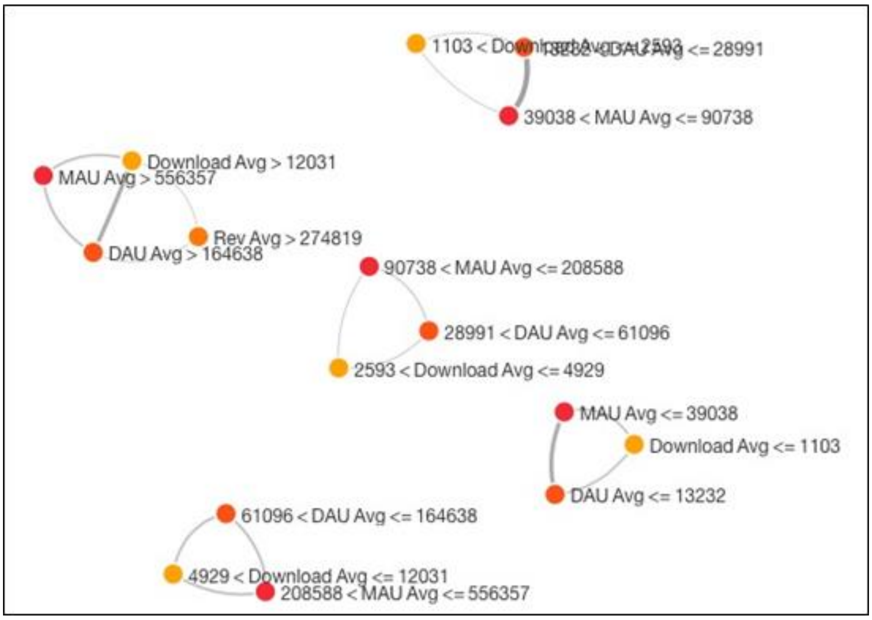



**Figure 20. User Behavior Clusters**

The User Behavior Anomalies were grouped into 5 clusters which reflect the most frequent pairings between DAU, MAU, Downloads and Revenue. The MAU Average <= 556357 appeared in 2 relationship clusters, as did the Download Average > 12031. The Revenue Average appeared in only 1 cluster as > 274819.

#### 4.2.3.1 User Behavior Data Distribution and Field Importance

The training set split used to create the user behavior prediction models was also 80%/20% by both random and deterministic sampling and ordering. The Data Distribution and Prediction Distribution values for the user behavior variables are listed in Appendix E. The Field Importance for user behaviors dataset were Daily Average User 94.35%, Revenue Average 4.01%, Average Downloads 1.10%, and Game Title 0.54%. A comparison of the Field Importance value, root node, and predictions is discussed further in Chapter 5.

### 4.2.4 Mobile Game Title Revenue Comparison

The features of the mobile game app are designed to generate substantial returns based upon some or all of the ARM model components (Moreira et al 2014; Askelof, 2013; Filho et al 2014). In turn, the purchasing behaviors and the total revenues generated serve as indicators of the game app success. The revenues for the developers and revenues for selected game titles were included in the analysis in order to make assessments in regard to monetization, the success or failure of the in-app purchase structures, and how ARM monetization correlates to the game revenue. Table 10 shows the revenues and titles that are represented by each developer:

**Table 10. Game Title Revenues: 2014 to 2016**

| Game Title | # of Titles | 2014 Revenues | 2015 Revenues | 2016 Revenues | Totals |
|---|---|---|---|---|---|
| **Electronic Arts** | 4 | $3.58 Billion | $4.53 Billion | $4.37 Billion | $12.48 Billion |
| **King.com** | 3 | $2.26 Billion | $2.0 Billion | $1.59 Billion | $5.85 Billion |
| **Zynga** | 3 | $690 Million | $764.72 Million | $741.42 Million | $2.196 Billion |
| **Jam City** | 3 | $3.40 B | $3.3 Billion | $3.13 Billion | $9.83 Billion |
| **Big Fish Games** | 2 | $13.9 Million | $113.7 Million | $116.6 Million | $244.2 Million |
| **EA Sports** | 2 | $3.575 Million | $4.515 Million | $4.963 Million | $13.053 Million |
| **Playtika** | 2 | $549 Million | $725 Million | $900 Million | $13 Billion |
| **Machine Zone** | 2 | $831 Million | $1.1 Billion | $5.7 Billion | $7.6 Billion |
| **Gameloft** | 1 | $302.38 Million | $279.37 Million | $80.25 Million | $5.32 Billion |
| **Nintendo** | 1 | $5,72 Billion. | $5 Billion | $4.19 Billion | $14.91 Billion |
| **Glu Mobile** | 1 | $76.2 Million | $61.0 Million | $43.6 Million | $180.8 Million |
| **Rovio** | 1 | $172.62 Million | $154.95 Million | $207.51 Million | $535.08 Million |

Electronic Arts had the largest number of instances in the sample; however, Nintendo realized the most total revenue. The slowest financial performance was by EA Sports and Glu Mobile. On a year to year basis, the highest performer for 2014 and 2015 was Nintendo. The highest performer for 2016 was Machine Zone. The most consistent earnings were from King.com, Jam City, Nintendo, and Machine Zone.

The data reflects the performance of the sample group based upon revenues for 3 consecutive years[45]. Electronic Arts leads the industry with a total of $12.48 billion in

[45] For sources for developer revenue data see Market Watch. 2017. Statista. 2017b, Statista 2017c, Sega Sammy Holdings Annual Report. 2016. Casino Daily News. 2016. Nasdaq. 2017. and Tsipori. 2016.

annual revenues for 2014 to 2016. Playtika leads the median group of developers with $13 billion for the period. The highest earner for the sample of developers is Electronic Arts with annual earnings between $3.58 billion and $37 billion for the period of 2014 to 2016. The second highest earner is Zynga, with revenues between $600 and $750 Million for the period[46]. The highest ranking developers by leading game apps are Supercell for Clash Royale, King.com for Candy Crush, Machine Zone for Mobile Strike, Niantic for Pokemon Go, and Miniclip for 8 Ball Pool (ABC57, 2017)[47]. The lowest performers: EA Sports $13.053 million and Glu Mobile $180.8 MillionThe user behavior regression model yielded values a mean absolute error of 235,173, 418,635, and 487,191.77 random. The mean squared error 324,216,523,223, 697,001,505,710.92, and 495,488,591,938. The R-squared was 0.53, 0. mean, and 0.29.

[46] Statista. 2017a. *Annual revenue  generated by Zynga from 2008 to 2016.*

[47] Rankings based upon Juniper Research Annual Report: *Future Games Market: Emerging Opportunities & Pivotal Publisher Analysis 2017 - 2021*.

# Chapter 5. RESULTS AND DISCUSSION

The prediction model generator by BigML may be used to predict the probability of success or failure of data classifications that have been fed from a data training set. The prediction model has been useful in the development of the inclusion criteria and for covariate adjustments in random trials. Binary predictors support assessments using ROC curves to summarize the model accuracy. For this study, the statistical model is generated based upon predictor variables for a 50 game title dataset. Steyerberg et al (2010) supported that the calibration and discrimination aspects of the prediction model performance are significant, and that other performance measures, such as for reclassification, may be used to increase the level of insight of the value-added of novel predictors to a particular model. Wallis (2013) supported that normal approximations to binomial distributions provide a platform for statistical testing and methodologies, to include the generation of reliable confidence levels, goodness of fit tests, contingencies and model and line fitting. Further, the evaluation of predictive model performance and accuracy may be used beyond the training dataset toward comparisons of the predictive models themselves, and the algorithms used in the model to produce the predictions.

This section provides the results of the statistical analysis for the Game Feature and the User Behavior datasets and a discussion in regard to issues that were encountered with the datasets and the outcomes that provided the most insight toward the machine learning predictions. Section 5.2 provides a summary of the outcomes for the game features dataset. Section 5.3 provides a summary of the outcomes for the user behavior dataset. Section 6.1 and 6.2 presents a discussion of the outcomes for the game features and user behavior analyses, to include the associations and predictions. Section

6.3 presents the results for both datasets and the predictions are discussed in regard to rankings and market value, and a juxtaposition of game genres and developers for the highest and lowest performers.

### 5.1 Decision Trees

The BigML decision trees are color coded and arranged according to specific logic. The color of the nodes is determined by the input fields that are associated with them. The root node is the first node in the decision tree and depicts the field which may best be used to perform the best split based upon the objective field. Further, the root node is connected to any number of child nodes by the branches. The root node is not necessarily the most significant node, although it has been determined by the algorithm to be the optimal point to split the data. The significance of the nodes is also depicted in the degree of separation from other nodes in the training sets. The final green node contains the prediction for each feature that appears in a subset of the tree.

The game feature and user behavior datasets were assessed for the number of instances, missing values for each field, and errors that occur from unidentified data and token formats. The Confidence Level for the decision tree model is estimated based upon the number of instances for each node and the purity of the terminal node (BigML, 2017b). A 50% confidence level is equivalent to tossing a coin for heads or tails. Anything more or less than 50% adds tor detracts to the reliability of the confidence level. Thus, the base estimation of an accurate probability is represented in the node purity. The formula used in BigML to estimate the probability, the Wilson score Interval, is based upon the balance between the percentage of the class that is being predicted, or the Support, and the degree of uncertainty that is associated with the number of instances in the dataset.

### 5.1.1 Game Features

The analysis of the game features produced a decision tree from the training data with predictions and accompanying purity, or levels of confidence for 15 of 29 features. As the last field in the dataset, Unlock Content was selected as the objective field by default for the game features. The game features are divided into 2 classes: 0 or 1 for not present or present, respectively. For the game features dataset, the root node is the Levels feature based upon the objective field of Unlock Content. Following the Levels root node for the game features model, the split produced Power Up and Item Upgrades as the first level of child nodes for all game titles. The Item Upgrades node did not produce any child nodes. However, the Power Ups node produced Time Skips and Noncumulative Rewards as child nodes. The Time Skips node produced a Customizable child node. The remainder of the child nodes for the tree are produced from Noncumulative Rewards.

The 2 child nodes for Noncumulative Rewards were Request Friend Help and Daily Offers. Daily Offers produced recursive child nodes for Skill Tree and Unique Offer. The remainder of the child nodes produced in the tree were from Request Friend Help.  Request Friend Help produced Random Elements, which produced Competitive Play. The Competitive Play node produced nodes for Cooperative Play and In-App Purchases. Cooperative Play produced all game titles. The predictions ended at the child nodes for the Item Upgrade, Customizable, Unique Offers, and In-App Purchases features. Figure 21 shows the Confidence Level and percentage of data for the red node, Request Friend Help:

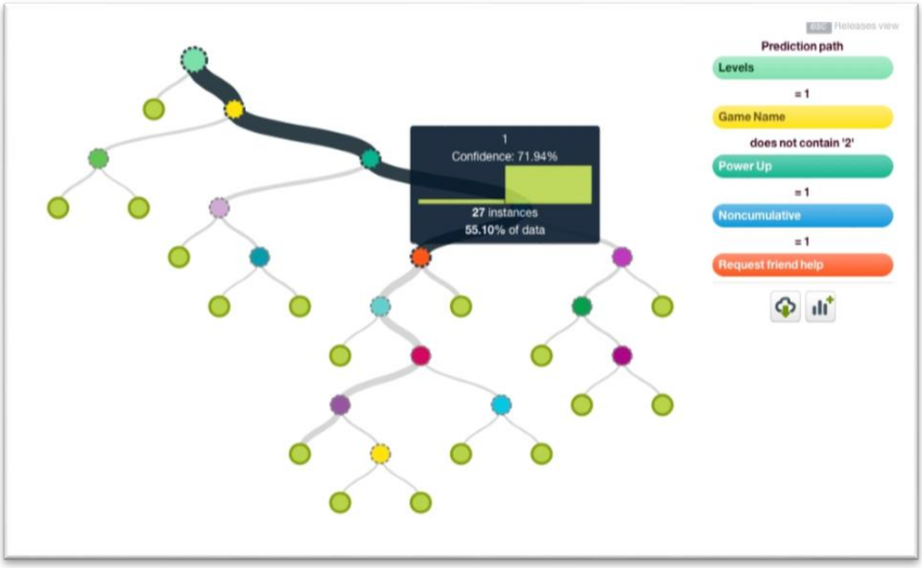



**Figure 21. Confidence and percentage for Request Friend Help**

Across 33 instances of Request Friend Help, the red node separated the classes of features with a confidence level is 71.94% and the red node covered approximately 55% of the data. The histogram depicts the underlying data distributions.

Here, the churn for the mobile game app consumer over approximately 30 days is represented in the percentage of data. The game features that would produce the highest probability for churn over a 30-day period are Levels with a confidence of 51.31%, IAP with a confidence level of 9.45%, Power Up with a confidence level of 60.66%, Noncumulative Rewards with a confidence level of 66.49% confidence, Request Friend Help with 71.94% confidence, and Random Elements with 60.78% confidence, and Competitive Play with 66.65% confidence.

The Game Name is depicted by the yellow node and appears at the beginning and the end of the decision tree as an assessment of words that could be significant due to frequency or absence. From the assessment of the recurrence of words in the game titles, the mobile games with the words ‘Blast’, ’Clash’, ‘Poker’, and ‘War’ also held significance in the assessment of the text fields for *Game Name*. The binary fields for the game features are treated as categories in BigML. Although features such as Levels and IAP can be found in 100% of the dataset, varying user behaviors and revenue generation affects the degree to which the features affect success.

The analysis also provided an assessment of the frequency of specific words within the game titles. A total of 14 game features were eliminated from the success prediction based upon features. Table 11 shows the 15 features which held significance in the decision tree prediction model for the game features dataset:

**Table 11. Game Feature Decision Tree Outcomes**

| Results for Game Features Dataset | | |
|---|---|---|
| **Game Feature** | Confidence Level | % of Data |
| **Levels** | 51.31% | 100% |
| **Unlock Content** | 34.24% | 66% |
| **Item Upgrade** | 30.06% | 74% |
| **Unique Offer** | 20.77% | 82% |
| **Power Up** | 60.66% | 83.67% |
| **Random Elements** | 60.78% | 86% |
| **Time Skips** | 25.05% | 66% |
| **Customizable** | 30.06% | 93.49% |
| **Request Friend Help** | 71.94% | 55.10% |
| **Noncumulative** | 66.49% | 66.39% |
| **Daily Offer** | 25.05% | 89.29% |
| **Cooperative Play** | 70.18% | 76% |
| **Skill Tree** | 23.07% | 46% |
| **Competitive Play** | 66.65% | 76% |
| **IAP** | 9.45% | 100% |

The decision tree identified 15 of the 29 features with varying degrees of confidence across the percentages of the sample

Lastly a SunBurst visualization is generated in BigML to depict the game title categorizations from a different view or perspective. Figure 22 presents a SunBurst visualization that was generated in BigML for the game features dataset:

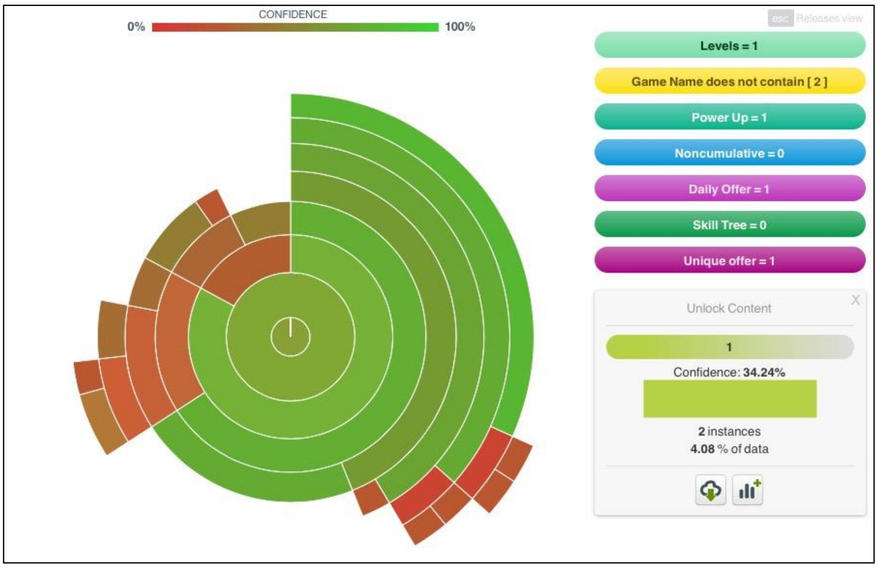



**Figure 22. SunBurst Visualization for Game Features**

The confidence level for the game features are lowest on the starburst color map at red and highest at green. The least predictive areas in the map are depicted by the brown areas. The game features are also depicted in a decision tree hierarchy by color, and also the subsets that were created.

The root node, Levels, is at the center of the SunBurst. The layers of the SunBurst are an indicator of the complexity of the prediction paths. The length of each ring arch is a depiction of the training set percentage that was considered to produce the child node. The arc lengths are also indicators of the amount of Support for the child node; as smaller arcs define less Support. A partial summary of the prediction rules for the game features is displayed in part on the right side of Figure 21. The full listing for the rules, or Rule Summaries for the game features and user behavior datasets are listed in Appendices F and G, respectively.

The 15 selected features were distributed across the game feature categories as follows:

- Social Network - Request Friend Help, Competitive Play, Cooperative Play
- Offers - Unique and Daily Offers
- Currency/Gambling - IAP
- Play Alterations - Time Skip
- Reward Retention/Punish Absence - Noncumulative Reward
- Game Feature - Customizable, Skill Tree, Random Elements,

Power Ups, Unlock Content, Item Upgrade, Levels

The highest game feature category representation was for Social Network and Game Features. Offers produced the second highest performance. Due to being present in all of the game titles, the In-App Purchases scored the lowest in significance to game app success. Correspondingly, all of the Game Features were included in the decision tree with the exception of Status Upgrade, compared to 3 Social Network features.

The BigML provides several color coded maps, trees, scatterplots and other diagrams for the dataset which provide a diversity of perspectives of the data by category and frequency. Fukita et al (2016) supported that data mining methods which are based upon intelligent systems have the capacity to retrieve analytical results through TDIDT algorithms, Bayesian networks, and Self-Organizing Maps (SOMs). Symbolic illustrations of the data are obtained through the TDIDT algorithm which may be used as evidence of the distinctions between the data classes. The Self-Organizing map may be used to form data clusters which extend the generalization of the categories to groups that reflect deeper representations of new data. The discriminative attributes and behavioral patterns are identified in big data by Bayesian networks. The 15 features in the decision tree were clustered as follows:

- Cluster 1: Levels, Item Upgrade
- Cluster 2: Power Up, Time Skip, Customizable
- Cluster 3: Noncumulative, Request Friend Help, Random Elements, Competitive Play, Cooperative Play, In-app purchases
- Cluster 4: Daily Offer, Skill Tree, and Unique Offer

The 14 features that were not represented in the outcomes for the decision tree were Facebook, Leaderboard, Soft and Hard Currency, Invite Friends, Line Chat, Single

Play, Achievement, Versus, Cumulative, Status Upgrade, Gambling Reward, Time Boost, and Event Offer.

## 5.2 User Behavior

The decision tree that was created for the user behavior dataset produced outcomes in a different form from the game features as the user behavior data was not binary. For the user behavior dataset, the root node is Average Downloads based upon the objective field of Monthly Average Users (MAU). Figure 23 illustrates the Download Average split into the dark orange node for Daily Average Users and the medium orange node for Average Revenue:

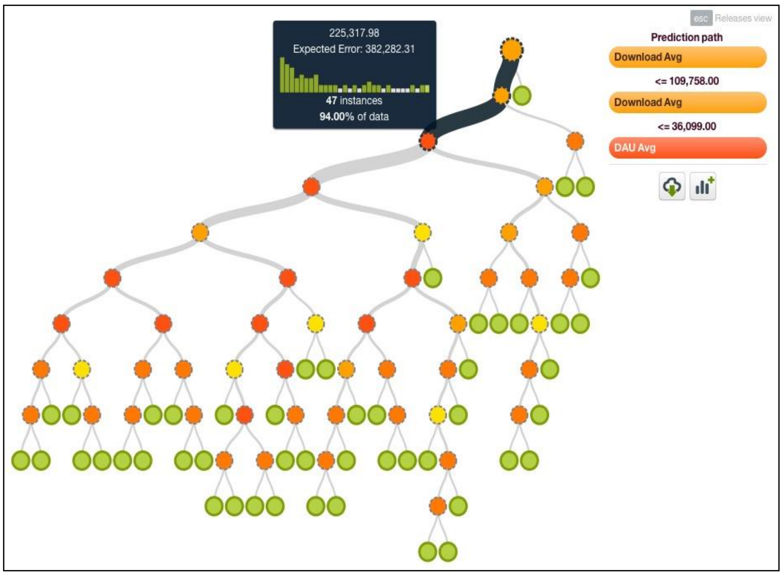



**Figure 23. Download Average Split to DAU and Revenue Average**

The Daily Average User produces 225,317.98 with 382,282.31 expected error and 95% confidence bound. The Average Revenue produces 2,055,121.50 with 3,674,415.52 expected error and 95% confidence bound.

The prediction path for user behaviors is furthered through the Daily Average Users node, which splits back to Download Averages in 2 instances, with a third path that is initiated based upon *Game Name*. Unlike the game feature prediction model, a prediction is given for the root node of Download Average, which also repeats as the first step down the prediction path. The first split of the Download Average node is between Daily Average Users Average Revenue. Similar to the Item Upgrade node in the game features model, the Average Revenue node does not produce any child nodes. The largest child nodes are produced by the Download Average feature. However, all of the final predictions were made based upon the Revenue Average.

In order to develop a complete representation of all of the user variables, particularly the data for Average Revenues, the model was adjusted to change the objective filed to Average Revenue. When the objective field was changed from Monthly Average Users to Average Revenue, the root node changed Average Downloads to Daily Active Users. Also significant was a sizable increase in the importance of the game titles, as the Game Name drew 2 closed ended predictions between the daily and monthly active users:

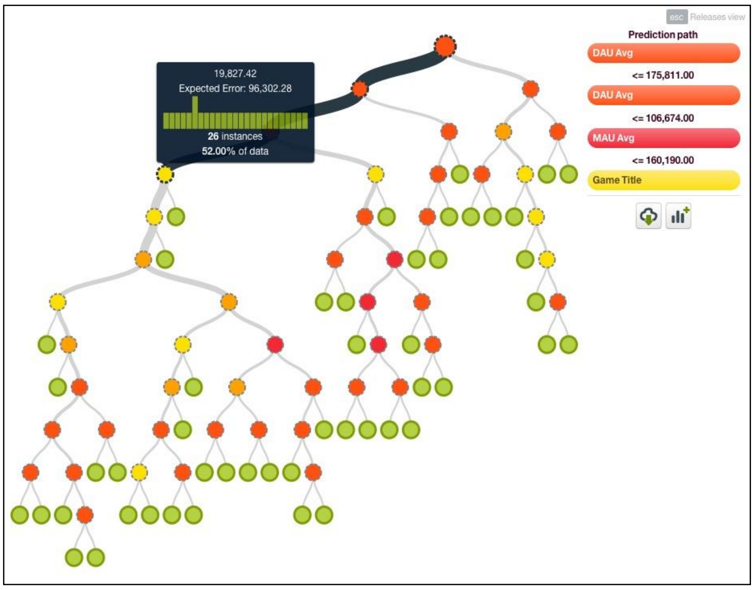



**Figure 24. User Behavior Decision Tree with Revenue Objective Field**

The changed objective increased the number of Game Name nodes along with their prediction levels. The first Game Title for 26 instances produced 19,827.42 with an expected error of 96,302.28 and 95% confidence bound. The second node produced 10,166.04 with an expected error of 42,696.64 with a 95% confidence bound.

Steyerberg et al (2010) supported that the prediction model performance may be measured by a diversity of approaches and using a diversity of metrics to include concordance statistic for the area under the ROC curve, goodness-of-fit, and the Brier Score. The probability that the user will select the mobile game app from the list of results is also assessed by the popularity rating or by the visibility rank, which may improve or decline from year to year. The popularity ratings are based upon surveys of users from a diversity of group demographics in addition to considerations based upon financial assessments. Further, the visibility score is an integration of the algorithm, Organic and Paid results, Search Behavior Intelligence, and Search for Info versus Search to Buy metrics. In both versions of the model, the Daily Average User prevailed in the prediction as the most significant user behavior indicator of mobile game app success. Overall, in comparison to the game features decision tree, the user behavior model produced a larger number of nodes and leaves, which also is a reflection of the different types of input data. The Item Upgrades feature and Revenue Averages user behavior indicator produced distinct patterns and predictions in the analysis.

# Chapter 6.  The Final Prediction Outcomes

The attainment of the prediction of one mobile game app from a sample of 50 was accomplished by running a batch prediction for the game features dataset, and a separate batch prediction for the user behavior dataset. The lists were then integrated, a final list of games which appeared in both lists was generated for further comparison. The batch prediction in BigML computes predictions for each instance in the dataset in one request based upon a Boolean argument for all_fields. However, the option to name a category that best describes the prediction was not used. The confidence for the model prediction was also included as an additional Boolean argument. A prediction confidence is returned as positive for the instances in the classification with when the prediction is above the threshold that was established for the model (BigML, 2017c). In the event that a positive is not returned, the classification defaults back to the majority classification. For the game features dataset, the classification was based upon whether the value for each feature was 0 or 1. Further, the class value is replaced by the least frequent category in the training dataset. According to the prediction model results for the dual datasets, the most successful mobile game app from the 50 game sample was *Game of War-Fire Age*; the most successful genre was Puzzles, and the most successful developer was EA Sports.

### 6.1 Game Feature Batch Prediction

The batch prediction is created asynchronously in BigML by using the model, logistic regression, or topic model_id and a dataset_id (BigML, 2017c). The status object properties for the batch prediction consist of the code, elapsed time, status messaging, and a progress float between 0 and 1. The predictions for both the game features and the user behaviors may be adjusted through the parameter settings. The mobile game app predictions for this study were created using the model id and dataset_id.

The batch prediction for the game features dataset yielded a total of 25 or 50% of the dataset: *Asphalt 8: Airborne*, *Boggle with Friends*, *Bubble Witch Saga*, *Cally's Cave 3*, *Cookie Jam*, *Crossy Roads*, *Dynasty Warriors Unleashed*, *Farmville: Tropic Escape*, *Fire Emblem Heroes*, *Game of War - Fire Age*, *Gummy Drop!*, *Madden NFL Mobile*, *MLB Tap Sports Baseball 2017*, *Mortal Kombat X*, *Panda Pop*, *Piano Tiles 2*, *Pokemon Go*, *Slotomania*, *Sonic All Star Racing*, *Summoners of War*, *Super Stickman Golf 3*, *The Simpsons: Tapped Out*, *The Sims Free Play*, *Toy Blast*, *Word Cookies*. The predictions trained set processes and output preview for the game features predictions in the BigML API are depicted in Figure 25:


**Figure 25. Game Feature Batch Predictions**

The batch prediction produced outcomes based upon the modeling analysis of the game features for 50 games.

## 6.2 User Behavior Batch Prediction

The most successful game predictions were also extracted and compared to the predominating user behaviors for further analysis and to develop a visualization of the future market implications. The predictions trained set processes and output preview for the user behavior predictions in the BigML API are depicted in Figure 26:

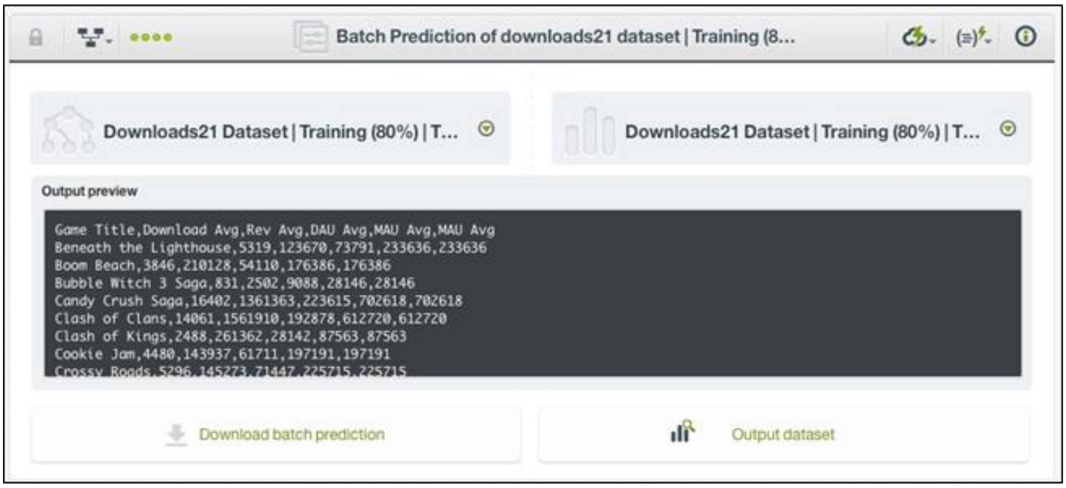



**Figure 26. User Behavior Batch Prediction**

The batch prediction produced 24 outcomes based upon the modeling analysis of the user behavior for 50 games.

A total of 24 titles, or 49% of the sample, were produced with the highest probability of success based upon the user behavior data: *Beneath the Lighthouse*, *Boom Beach*, *Bubble Witch Saga 3*, *Candy Crush Saga*, *Clash of Clans*, *Clash of Kings*, *Cookie Jam*, *Crossy Roads*, *Double Down Casino*, *Fire Emblem Heroes*, *Game of War - Fire Age*, *Genies and Gems*, *Kim Kardashian: Hollywood*, *King Rabbit*, *Madden NFL Mobile*, *MLB Tap Sports Baseball 2017*, *Mobile Strike*, *Mortal Kombat X*, *Panda Pop*, *Rogue Runner*, *Temple Run 2*, *The Walking Dead: Road to Survival*, *Two Dots*, and *Zynga Poker: Texas Holdem*. The 24 games that were produced in the batch prediction based upon the user behavior variables are depicted in Table 12:

**Table 12. User Behavior Success**

| Batch Predictions based upon User Behavior Variables | | | | |
|---|---|---|---|---|
| **Game Title** | Download Avg | Rev Avg | DAU Avg | MAU Avg |
| **Beneath the Lighthouse** | 5319 | 123670 | 73791 | 233636 |
| **Boom Beach** | 3846 | 210128 | 54110 | 176386 |
| **Bubble Witch Saga 3** | 831 | 2502 | 9088 | 28146 |
| **Candy Crush Saga** | 16402 | 1361363 | 223615 | 702618 |
| **Clash of Clans** | 14061 | 1561910 | 192878 | 612720 |
| **Clash of Kings** | 2488 | 261362 | 28142 | 87563 |
| **Cookie Jam** | 4480 | 143937 | 61711 | 197191 |
| **Crossy Roads** | 5296 | 145273 | 71447 | 225715 |
| **Double Down Casino** | 2211 | 11121 | 25423 | 74568 |
| **Fire Emblem Heroes** | 2133 | 123250 | 24859 | 76796 |
| **Game of War - Fire Age** | 13133 | 2282495 | 177019 | 552528 |
| **Genies & Gems** | 2288 | 11404 | 27114 | 85012 |
| **Kim Kardashian: Hollywood** | 8516 | 702465 | 116639 | 369597 |
| **King Rabbit** | 1542 | 12213 | 16092 | 44105 |
| **Madden NFL Mobile** | 12443 | 458318 | 174604 | 560186 |
| **MLB Tap Sports Baseball 2017** | 661 | 2392 | 3803 | 8254 |
| **Mobile Strike** | 21468 | 2291221 | 287039 | 899418 |
| **Mortal Kombat X** | 555 | 2360 | 1294 | 3117 |
| **Panda Pop** | 11620 | 133333 | 154672 | 480518 |
| **Rogue Runner** | 886 | 1731 | 8180 | 19519 |
| **Temple Run 2** | 2900 | 967 | 21980 | 56062 |
| **The Walking Dead: Road to Survival** | 2158 | 12349 | 22734 | 71451 |
| **Two Dots** | 2509 | 8001 | 24139 | 68949 |
| **Zynga Poker: Texas Holdem** | 474 | 1922 | 448 | 804 |

The user behavior dataset produced a sample with significant similarities and differences from the game feature prediction outcomes. A comparison of the user behavior averages for sample presented significant variations between the values in regard to downloads, revenue, and average daily and monthly use. Zynga Poker: Texas Holdem generated only 474 average downloads, compared to 21,468 for Mobile Strike.

Two of the Casino games, Double Down Casino and Zynga Poker: Texas Holdem made the user variables batch prediction, but not the game feature batch prediction. The Big Fish Casino game made the game features batch prediction, but not the user behavior batch prediction.

### 6.3 Prediction from Integration of Dual Datasets

The dual analysis produced a list of the most successful games based upon features, and a list based upon the most successful user behavior indications. By selecting only the titles which appeared on both lists, the model produced a total of 9 titles from the sample with the highest probability of success: *Bubble Witch Saga 3*, *Cookie Jam*, *Crossy Roads*, *Fire Emblem Heroes*, *Game of War - Fire Age*, *Madden NFL Mobile*, *MLB Tap Sports Baseball 2017*, *Mortal Kombat X*, and *Panda Pop*. From the 9 title sample, the most successful game was generated by a comparison of user behavior outcomes for the final sample. Table 13 shows the comparisons for the game genre, developer, revenue, downloads and average uses:

**Table 13. Comparison of 9 leading mobile game apps user behavior data**

| Game Title | Genre | Developer | Down Avg | DAU Avg | MAU Avg | Rev Avg |
|---|---|---|---|---|---|---|
| **Bubble Witch Saga 3** | Puzzle | King.com | 831 | 9068 | 28146 | 2502 |
| **Cookie Jam** | Puzzle | Jam City | 4480 | 61711 | 197191 | 143937 |
| **Crossy Roads** | Kids | Hipster Whale | 5296 | 71447 | 225715 | 145273 |
| **Fire Emblem Heroes** | Adventure | Nintendo | 2133 | 24859 | 76796 | 123250 |
| **Game of War - Fire Age** | Role Playing | Machine Zone | 13133 | 177019 | 552528 | 2282495 |
| **Madden NFL Mobile** | Sports | EA Sports | 12443 | 174604 | 560186 | 458318 |
| **MLB Tap Sports Baseball 2017** | Sports | EA Sports | 661 | 3803 | 8254 | 2392 |
| **Mortal Kombat X** | Action | Electronic Arts | 555 | 1294 | 3117 | 2360 |
| **Panda Pop** | Puzzle | Jam City | 11620 | 154672 | 480518 | 133333 |

Significant outcomes for the comparisons included the appearance of the Puzzle genre in 33.33% of the sample. Also significant was the amount of variation between the number of downloads, revenue, and usage between the highest and lowest ranking game titles.

The highest performing game features from Chapter 4 were compared with the batch prediction sample of 9 mobile game apps:

- Social Network - Request Friend Help, Competitive Play, Cooperative Play
- Offers - Unique and Daily Offers
- Currency/Gambling - IAP
- Play Alterations - Time Skip
- Reward Retention/Punish Absence - Noncumulative Reward
- Game Feature - Customizable, Skill Tree, Random Elements, Power Ups, Unlock Content, Item Upgrade, Levels

Table 14 shows the comparison of game features for the final 9 games:

**Table 14. Comparison of 9 leading mobile game apps game feature data**

| Game Title | FB | II | RFH | LC | SP | VS | CMM | COOO | SC | GA | HC | CUM | NC | AC | IAP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bubble Witch Saga 3 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| Cookie jam | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| Crossy Roads | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Fire Emblem Heroes | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| Game of War - Fire Age | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Madden NFL Mobile | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| MLB Tap Sports Baseball 2017 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| Mortal Kombat | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Panda Pop | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| | | | | | | | | | | | | | | | |
| **Game Title** | UOO | EO | DO | TS | TB | CUS | Ran | STree | PUP | LVL | Stat | Item | IAP | Unlock | |
| Bubble Witch 3 Saga | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | |
| Cookie Jam | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | |
| Crossy Roads | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | |
| Fire Emblem Heroes | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | |
| Game of War - Fire Age | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | |
| Madden NFL Mobile | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | |
| MLB Tap Sports Baseball 2017 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | |
| Mortal Kombat | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | |
| Panda Pop | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | |

Significant outcomes for the comparisons included the predominance of the Social Networking features, Offers, and IAP 90% to 100% of the sample.

### 6.4 Most Successful Mobile Game Genre Prediction

Puzzle games have been ranked as the most popular mobile game genre based upon user behaviors. Barus, Tobing & Pratiwi (2015) used the Unity Test Tool methodology to test a Puzzle mobile game app and found that the game was free of defects and bugs. Hwong (2016b) confirmed that the Puzzle mobile game genre is the most popular based upon a study by Verto analytics[48]. More specifically, Brain Puzzles were ranked as the genre players spent the most time playing per month, with a stickiness of 40%. Brain Puzzles were followed by Matching Puzzles, Alternate Reality and Action/Strategy as the most highly engaging genres. Further, the Puzzle games attracted players from an age demographic of between 35 and 64.

Sonders (2016) also found that Puzzle games are the highest ranking mobile game app genre based upon time spent and engagement[49]. Each active user in the sample was found to average 105 minutes per month in Puzzle game play in the United States. The players were also found to play Puzzle games approximately 5.7 times more often each month than players of educational games with approximately 44% higher engagement than Arcade game players.

The genres and developers of the mobile game apps which appeared in both batch predictions were compared to the preliminary findings for the sample genres and developers in Chapter 4. Based upon the batch predictions, the leading genre was

---

[48] Verto Analytics study consisted of a sample of American gamers 18 years of age and older. The study measured the amount of time spent playing, Net reach and Stickiness of the games by genre.

[49] Based upon Survey Monkey Intelligence Study conducted in July 2016.

Puzzles, followed by Sports. The Role Playing, Action and Adventure genres were all included in the final prediction, and are also considered interchangeable as genres. The genre predictions are consistent with the findings for genres in section 4.2.1, in which the leading genres were Action, Adventure, Puzzles, and Strategy. The Sports genre was not included in the genre outcomes of section 4.2.1; however, the Sports genre could also be interchanged with Action and Strategy.

### 6.5 Most Successful Mobile Game Developer

The most successful developer was EA Sports, as the only developer with more than 1 title in the final 9 game sample. Headquartered in Redwood City, California, Electronic Arts was founded in 1982 and has since evolved into "global interactive entertainment software" corporation (Electronic Arts, 2017). The entertainment software is produced for the video console, computer, mobile phone, and tablet markets. The most profitable, recent titles from Electronic Arts include *FIFA*, *Madden NFL Mobile*, *NBA Live Mobile*, *Dragon Age*, *The Sims*, *The Secret Life of Pets Unleashed*, *Bejeweled Stars*, *Need for Speed No Limits*, *The Sims: FreePlay*, and *Plants vs. Zombies*. In the past few years, Electronics Art has begun to focus upon mobile game apps within 3 of the company's 4 major franchises for sports (Rogers, 2016). The company's net revenues for 2016 exceeded $4 billion (Electronic Arts, 2017). The company realized a substantial growth in digital revenues due to the growth of the mobile gaming revenues and the number of downloads of full game versions.

The leading developers found in section 4.2.1 were Electronic Arts, Supercell, and King.com. Based upon the batch predictions, the leading developer was EA Sports, followed by 7 developers which did include Electronic Arts and King.com. *Mortal*

*Kombat X* may also be credited to the same developer to comprise 33.33% of the 9 game sample.

### 6.6 Most Successful Mobile Game App Prediction

The objective of this study was to produce a prediction and a model for the prediction of success mobile game apps based upon a dual approach to analyzing the game features and user behavior. The most successful mobile game app from the 50 game sample, based upon the dual batch predictions for game features and user behaviors, is *Game of War - Fire Age*, which was first released in 2013 by Machine Zone:


**Figure 27. Game of War - Fire Age**

The multiplayer fantasy game achieved its first global number 1 rankings in October 2015 based upon Apple App Store Revenues (App Annie, 2015). The title gained popularity in the European, United States, and Japanese markets and remained at the top of the charts as number 1 or number 2 for the duration of October 2015. Much of the success of Game of War has been the ad-based monetization strategies and IAP promotions that appeal to whales. Features other than those selected for the analysis include language translation and messaging.

The *Game of War - Fire Age* app was selected based upon final rankings for performance from the final 9 game sample group:

- Highest Download Average - *Game of War Fire Age*
- Highest DAU Avg - *Game of War- Fire Age*
- Highest MAU Avg - *Game of War Fire Age*
- Highest Revenue Average - *Game of War -Fire Age*

The outcomes are based upon data extracted from mobileaction.com, thinkgaming.com, and App Annie for mobile game apps from 2013 to 2016.

*Game of War - Fire Age* was developed by Machine Zone and released in July of 2013 as a relaunch of the 2010 *Game of War* (Bridgman, 2015). The objective of the game is to build a city and develop the resources to attack the cities of other players. The initial progressive pace of the game is moderate; however, the momentum increasingly depends upon the player's investment in time and IAP. Initial investments of $5 escalate to IAPs for packs of gold that typically range from $16.99, to $50, and on to $100 to speed up the game. A Belgian teenager reportedly spent approximately $46,000 on *Game of War - Fire Age* IAPs for virtual gold, over a few months, using his grandfather's credit card (Mendoza, 2014).

The time investment increases rapidly for the players, as virtual gold, VIP and Power points accumulate. Taking breaks from the game without experiencing substantial losses requires the purchase of a Peace Shield for 24 hours to 30 days. Research that initially may require minutes or hours begins to require days to months to complete. With no endgame and over 160 statistics to improve upon, the benefits earned in *Game of War - Fire Age* are temporary, as the troops and resources that have been purchased can be lost (Hill & Croghan, 2015). The title is available for both iOS and Android, and requires approximately 250 MB of space to download. The successful mobile game may be categorized as Action or Strategy. Major market competitors include *Mobile Strike*, *Clash of Clans*, *Boom Beach*, *Clash Royale* and *Clash of Kings*.

Machine Zone invested large amounts of capital in the advertising for *Game of War - Fire Age*, and the game was a top grossing mobile game app in 2014 and 2015. In 2015, Machine Zone spent $40 million on an epic-themed, 4-month, global advertising campaign through television commercials, Facebook and Twitter, and Youtube action trailers which featured the face and bust of Kate Upton, and later featured Mariah Carey (Bridgman, 2015; App Annie, 2015):


**Figure 28. Game of War Kate Upton**

A large part of the success of the *Game of War* ad campaigns has been attributed to the selection of Kate Upton as the first model for the high end advertisement strategies. Kate Upton was later replaced by Mariah Carey as the game's supermodel.

The $40 million investment brought a return of more than $1 million per day in revenues in the United States, close to 41,000 in daily installs, and 2.9 million daily active users (Think Gaming, 2017a). In 2017, Game of War - Fire Age has realized almost $400,000 in daily revenue, ranks number 5 on the Top Grossing Games chart and number 43 on the chart for Top Free Games (Think Gaming, 2017b).

From a social networking perspective, the Game of War app provides global, simultaneous, multiplayer combat with the potential for alliances in real time.


**Figure 29. Game of War Facebook Referrals**

The Facebook referral boosts may be obtained on Stronghold levels 5 to 21 and may range from a Rare Gem Chest to 600,000 units of Dark Energy.

The players earn rewards for Facebook referral boosts, which must accumulate to increase the Stronghold level. The player uses the Invite Your Friends feature to connect with others who join the game. The player receives the reward after the friend downloads the game, completes the tutorial, and connects the game to their Facebook page. The number of Facebook rewards per player are limited to 10 to 20. Along with Tumblr, Instagram, and Twitter, *Game of War* may also be downloaded, shared, and discussed in Facebook. The Facebook likes average 250,000 to 300,000 per week (GOW Likes, 2017). Line chat is enabled with friends and alliances within the game through Alliance Chat, Chat with the Entire World, and 1-on-1 Chat.

*Game of War - Fire Age* was the second highest grossing game in the Apple App Store in 2015, despite some negative reviews by consumers (Johnston, 2015). The substantial reinvestment of daily revenues into the advertisement campaign contributed to the highly competitive rankings and revenues for *Game of War* in comparison to mobile game titles such as *Candy Crush Saga* and *Clash of Clans*. The high revenues and rankings for *Game of War* have remained consistent for the past 3 years.

The *Game of War* players purchase virtual gold to use in the game, which allows them to skip through the stages of the game faster. However, the game has a high burn-out rate; therefore, to replace the players that are lost each year, the advertisements are used to draw large numbers of new players. In 2015, Game of war - Fire Age was valued at $3 billion dollars (Bridgman, 2015). Studies show that mobile game players spend an average of $87 (ARPPU) each year on IAP for free-to-play games; however, the average annual spending on *Game of War - Fire Age* for 2015 was $550 (Grubb, 2016)[50]:

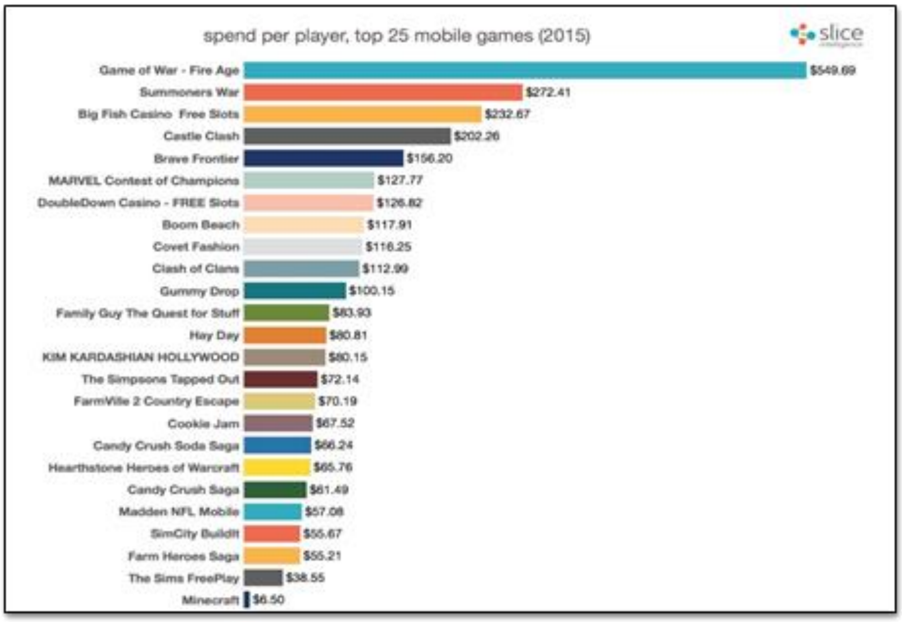



**Figure 30. Per Player Spending 2015**

In 2015, the average spending for *Game of War - Fire Age* was $549.69, which was twice that for *Summoners of War*. Several games from the 50 game sample for this study were included in the analytics by Slice Intelligence as a part of the top 25 revenue generators. Neither No. 2 *Summoners of War* or No. 3 *Big Fish Casino* were a part of the final 9 sample prediction.

Game of War - Fire Age IAPs far exceeded the IAPs for competitors. The majority of the average $550 was spent on virtual gold. In addition, Grubb (2016)

---

[50] Average revenue data extracted from a study by Slice Intelligence published by T. Stanton in 2016. Hardly pocket change: mobile gamers spend an average of $87 on in-app purchases.

supported that *Game of War* has a disproportionately higher number of whale players than other games.

### 6.7 Proposed Actionable Model Validation

The success prediction model accurately selects the mobile game app with the highest probability of success from a representative sample of games from several developers based upon a dual analysis. The pioneering underlying functionalities of the *Game of War - Fire Age* app model enable global simultaneous interactions between millions of players during real time, continuous play. Further, *Game of War - Fire Age* has generated the most consistent history of top grossing games since its release in 2013 (Think Gaming, 2017a). This section will present the code designed for a robust model to predict mobile game app success in Python. The anomaly detection was applied to the game feature and user behavior datasets during the data cleaning process in the development mode. The actionable model was then created by retraining the dataset. The model provides an interpretation of the data and the associations between features and also user behaviors (see Appendix H) :

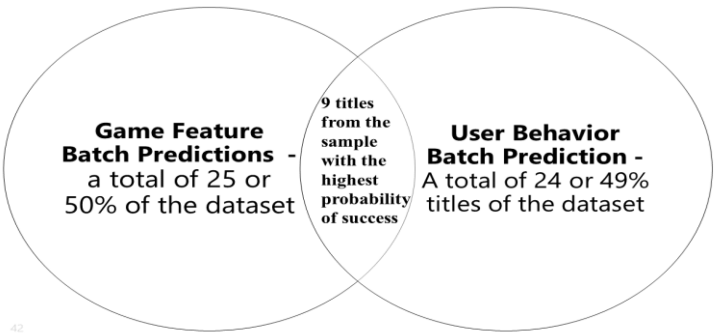



**Figure 31. Proposed Actionable Model**

The dual analysis produced a list of the most successful games based upon features, and a list based upon the most successful user behavior indications

The actionable model predicts the success of the mobile game app from a sample based upon the most outstanding performance of the input variables. The variables for the model may be substituted, and the sample size may be increased. Local predictions may be made for any sample of mobile game apps by downloading the model to a computer; or by caching the model on a server and serving with predetermined API bindings provided by BigML for python, node.js, java, and c-sharp. The local prediction may also be generated by embedding the actionable model into the webpage and by using java script bindings.

# Chapter 7. CONCLUSION

This study presented a dual methodology which produced success prediction outcomes that are more reliable, in that the model provides for more sources of input for similar analyses. Based upon the outcomes of the dual analysis, the major factors that drive mobile game app success include the extent of advertisement investments and approaches; the presence of violent and sexual content; the opportunity to gamble; the opportunity to compete, and the illusions of fantasy. The successful free-to-play game will require the player to invest substantial amounts of time, as well as currency. Further, irrespective of the genre, the game has to have an addictive element, as opposed to ‘just for fun’.

Data analytics is currently mature enough to depend upon in analyzing and discovering critical elements that are hidden behind the figures of mobile game performance. This study, therefore, gives a promising potential of being able to discover a new field of research that can be a pathway to redefine the usual game market. A combination of data analytics and predictive algorithms is a technique that brings a new angle to the way we look at mobile applications and should be greatly considered as a field worth venturing in.

Pioneering mobile game apps were significantly limited in visual and function capacities, which limited the degree to which mobile game apps could compete with video game consoles and pc games (LoWood, 2009; Wright, 2016). The digitalization of the gaming space has transitioned significantly (Rogers, 2017). As the technological capabilities of smartphones rapidly increases, the market for mobile game apps may be expected to continue to expand with the addition of innovative features and offerings. The game plays have become more complex with larger source codes and incorporate resources that are more complex, to include 2D and 3D graphics, sound, and artificial intelligence (Barus, Tobing & Pratiwi, 2015).

The increased potential for mobile game app functionality has attracted software developers who specialize in high functionality applications, such as sports and other action games which require several types of input from the user. In turn, the number of sports, war, and simulated living game apps continuously increase.

The value of this study is very evident. The end results are very beneficial and can serve to improve the levels of revenues from mobile devices opening up great potential for more research and improvement of the computational abilities of mobile phones. In addition, experimental approaches to demonstrating important game features give a less risky approach that will save developers a lot of money and time investment compared to the risks of prototyping tests in the normal game development iteration cycle.

The importance of growth in any industry cannot be underestimated. Many newcomers find it very difficult to successfully pick up and succeed in the game industry which is dominated by well-established industry leaders. Recommendations of this study will offer a soft landing spot for students as well as other individuals who join the mobile game development market.

Academic improvement and curriculum development is key in every field of research. Such advancements can only be made by forecasting and accommodating trends that will remain relevant and withstand market changes. This research is perfectly suited to providing a knowledge base that will greatly improve the information technology industry.

## 7.1 Contributions

The amount of literature on predictive mobile game performance analysis based on application features is very scanty. The research would focus on discovering new ways of predicting game revenue potential based on appropriate data analytics techniques.

Today most research works have focused on analyzing individual game features in relation to the overall experience served to the users as a factor in determining the acquisition and retention levels of mobile games (Átila, et al., 2014). The research would bring a new approach to the analysis of the problem domain by additionally seeking to investigate how monetization strategies affect the gaming experience then trickle down to demystifying the impact of to individual game features by also accommodating a top to bottom approach on top of the conventional bottom to top perspective.

The primary foundation of the research will be to combine statistical methods that have not been combined before with analytic techniques that will make it easier to understand the statistical relationships between data sets. Unlike the usual approach of analyzing game features, the study will marshal up two data prediction techniques and define the relationships between them to make unique conclusive findings that can be used by other researchers.

Predictive analysis, especially in the field of mobile game development, is often a complex study domain that requires analysis of many independent factors (Thomas & Kielman, 2016). This study will focus on simplifying the problem domain by making it easy for people to understand important trends that impact on the success of mobile games. Over the cause of the study factors will be graded based on weight metrics that will serve as basis of filtering out important aspects that ultimately determine the success of a mobile game.

The research further envisions to start a new field of research that through analysis of current trends and forecasting of future market behaviors that will shape the gaming industry. The main focus will revolve around formalizing the important aspects of current statistics to provide standards that can then be used for further analysis.

The study would also be an important resource in the development of curriculums that accommodate standards which lead to development of productive game development syllabus that can be employed by training institutions.

## 7.2 Limitations

The study has a scope of getting data from official mobile application stores. This might not necessarily reflect on the actual performance of games in various regions. There are a number of applications that are also published in other platforms such as the Vodafone application store. Some regions also have their own mobile stores especially for Android applications which may have a significant number of downloads that may be overlooked (Monica, 2015). Strategic evaluation can also be applied to cover up for the market shares dominated by other methods of distribution.

To get a full overview of the actual application performance statistics, a wide variety of data parameters are critical for the data analytics process. Some of these data values are important but very difficult to obtain. A significant aspect that can be used to analyze an application’s feature is the number of active users. This information is not freely available on application stores and may require workarounds to obtain. Also, getting data such as specific feature accesses is also not very straight forward. The number of uninstalls might not be openly accessed for data extraction and evaluation.

Mobile games are also subject to various environmental factors that affect the performance of a game. Economic aspects of a given time duration, as well as the prominence of specific mobile devices that support a specific device well, can be a contributing factor that is not associated with any analytics data. This could potentially affect the accuracy of the prediction model.

In many regions, game performances are also affected by a viral effect that span from a variety of sources which cannot be clearly defined. This could potentially be media mentions, or social media memes or attractive catch phrases whose occurrences cannot be predetermined. The research might will not be able to incorporate the viral effect factor accurately in the development of data models.

Some statistical data and application statistics such as revenue details are only available in specific regions. In some situations, device manufacturers especially for devices that run on open source platforms such as Android pre-bundle games with their devices. Such applications cannot be easily analyzed as they do not conform to the usual acquisition, retention, and monetization cycle.

In as much as the research will classify data based on time conscious models, many applications have varying trends of picking adoption. Some take a relatively longer duration but gradually become stable while others hit peak success at very early stages. The evaluation of these durations may require deep learning algorithms to process and understand. Detailed analysis of all the possible cases might, therefore, be challenged by computational abilities for data processing.

Finally, 29 features with sample of only 50 may lead to significant overfitting. Where overfitting is a significant practical difficulty for decision tree models and many other predictive models. Overfitting happens when the learning algorithm continues to develop hypotheses that reduce training set error at the cost of an increased test set error.

### 7.3 ARM, Advertising and Whales

The ARM model focuses upon the relationships between revenue, features, developer, users, and future research. In this study, it has been concluded that the advertising investments by Machine Zone had a significant effect on the final rankings of the game from 2014 to 2016. Further, the advertisements may be associated with the large number of whales who continuously invest substantial amounts of revenue into *Game of War - Fire Age*. Tapjoy (2016) identified 5 personas for players to subscribe to free-to-play games: new players, minnows, dolphins, whales, risk-of-churners, and offer takers. The lower level spenders are classified as minnows, the mid-level spenders are classified as dolphins, and the largest scaleof spending is by whales. Offer takers respond to rewarded advertisements. Risk-of-churners invest no revenue during gameplay and do not respond to the rewarded advertisements. The monetization strategy which appeals to whales through differentiated offerings and opportunities to increase value at a discount has been a successful business model for game developers. As much as 70% of the IAP revenue and 60% of the total revenues from free-to-play games comes from the whale spending (Tapjoy, 2016). Further, the whales have exhibited significantly longer life spans as gamers in comparison to the other personas. *Game of*

*War* players play longer, buy more virtual gold, and are highly likely to connect with other whales.

### 7.4 Social Impact of War Games

The success of mobile games with war scenarios, overly aggressive and violent themes, and psychological stimulants of aggression comes with social costs. Grossman & DeGaetano (2014) addressed theories that electronic games which are based upon war and violent content, and supported that the games simulate the methods used by military training to eliminate the soldier's aversion to act of killing. Although the violence is a significant part of the appeal of war games for many players, the continuous exposure sensationalized presentations of murder and torture serve as primers for the player to conceptualize killing as an acceptable act. Psychological research of links between violent media and aggressive behaviors have shown that the relationship is significant, particularly as an avenue for imitation of the violent acts and strategies.

- Pleasure and reward due to death and human suffering have been associated with stimulates for ongoing engagement with war games (Grossman & DeGaetano (2014).

- Grossman & DeGaetano (2014, p. 168) categorized games that are harmful as:

- Plots driven by rapid scenes filled with gratuitous violence that is delivered frequently, graphically and with salient technical effects.

- Sadistic, revengeful, and torturous methods, inhumane practices in the context of trivialization, raucous fun, humor, or glibness.

- Explicit depictions of violence achieved by special effects, background music, lighting, and glamorous, heroic presentations which encourage imitation.

- Depictions of social power due to weapons, domination, and violence; explicit violent details that are unnecessary to understand the central context of the message.

The studies also supported that the higher the intensity of the violence, the higher the intensity of the negative effects on the player. Although some games are rated as M, or Mature, due to the maturity of the content, a considerable number of players who choose war games are under the age of 18 (Mendoza, 2014). Recall that Jenkins (2004) pointed out that the youth demographic acquires mature game titles with violence and sexual content as many parents assume that all digital gaming media is intended for children; and therefore, the violence and maturity ratings are ignored during purchases for their children.

### 7.5 New Horizons in Mobile Game App Development

Market analysts predict that by the year 2020, the mobile game app industry will reach approximately $200 billion in revenues (Golmack, 2017). Much of the revenue is expected to come from the hybrid monetization model that includes in-app advetising and purchases. Brands will also lay a significant part in the future of mobile game app development, as large companies have established reputations and more capital to invest in advertising and market research. In turn, small developers have turned to crowdfunding as a monetization model in order to get their games to the market.

Facebook Instant and iMessage games over messaging will spur more growth, particulalry for social networking and competitive play genres. The 2D and 3D games are an emerging trend for the future profits for mobile game app developers and publishers, which will spur growth for virtual reality play with friends and fantasy genres. In contrast to the gore and sexually explicit content in mature war games, media market leaders such as the Cartoon Network have begun to establish mobile apps for children such as *Super Slime Blitz*, *Teeny Titans*, and *Rockstars of Oooo* (Cartoon Network, 2017). Walt Disney has also become a player in the mobile game app market, targeting young children with games such as *Marvel Future Fight, Frozen Free Fall: Icy Shot,* and *Moana: Rhythm Run* (Disney, 2017).

## 7.6 Implications for Future Research

The reliability of the success prediction outcomes that were presented in this study, and that are presented by the proposed model may be further enhanced by the removal of variables that are found in more than 90% of the sample population; by increasing the number of mobile game apps in the sample, and by adding new variables to the model, a more efficient scoring system is enabled that includes the developers. Features such as Facebook and IAPs were represented in 100% of the population; and therefore, had no identifiable impact in contrast to the remaining features. Also, future studies which include comparisons of the performance of paid games and free-to-play games may also add valuable insight.

# Reference


Aalto, O. 2015. *Mobile Game Product Development Models.* Business Information Technology, Kajaani University of Applied Sciences.

Agrawal, R. and Srikant, R. 1994. Fast Algorithms for Mining Association Rules in Large Databases. *VLDB Proceedings of the 20th International Conference on Very Large Data Bases*, pp. 487–499.

Akio, K. 2017. *'Red Faction II' Latest News: Game Removed From Germany's List of Media Harmful To Young People; THQ Nordic Did It Again!* [Online]. Available at: http://www.mobilenapps.com/articles/42785/20170228/red-faction-ii-latest-news-game-removed-from-germanys-list-of-media-harmful-to-young-people-thq-nordic-did-it-again.htm [Accessed: 03 Mar. 2017].

Aleksy, M. et al. 2008. Architecture for the Development of Context-Sensitive Mobile Applications. *Mobile Information Systems* 4(2), pp. 105–117.

Alomari, K. M. et al. 2016. Mobile Gaming Trends and Revenue Models. In: Fujita, H. et al. eds. *Trends in Applied Knowledge-Based Systems and Data Science: 29th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2016, Morioka, Japan, August 2-4, 2016, Proceedings*. Cham: Springer International Publishing, pp. 671–683.

Amir, B. and Ralph, P. 2014. Proposing a theory of gamification effectiveness. In: Jalote, P. et al. eds. *36th International Conference on Software Engineering (ICSE Companion 2014)*. New York, USA: ACM Press, pp. 626–627.

Apsalar. 2013. *Top Mobile Game Categories by In-app & Engagement* [Online]. Available at: https://apsalar.com/2013/02/top-mobile-game-categories-by-in-app-purchase-engagement/ [Accessed: 10 Apr. 2017].

Annie, A. 2015. *Mobile Games: Now You Can Predict the Future* [Online]. Available at: https://www.appannie.com/insights/app-adoption-cycle/ [Accessed: 20 Feb. 2016]

Annie, A. 2016. *Russia, the CIS and the Baltic region feature some up-and-coming mobile mega-powers. See which local game publishers are leading* [Online]. Available at: https://www.appannie.com/insights/mobile-gaming/russia-cis-baltic-region-spotlight-outpacing-global-gaming-market/ [Accessed: 05 Jan. 2017].

Apple. 2017a. *App Programming Guide for iOS : The App Life Cycle* [Online]. Available at: https://developer.apple.com/library/content/documentation/iPhone/Conceptual/iPhoneOSProgrammingGuide/TheAppLifeCycle/TheAppLifeCycle.html [Accessed: 22 Feb. 2017].

Apple. 2017b. *What's New in iOS: iOS 10.0* [Online]. Available at: https://developer.apple.com/library/content/releasenotes/General/WhatsNewIniOS/Articles/iOS10.html [Accessed: 20 Mar. 2017].

Askelöf, P. 2013. *Monetization of Social Network Games in Japan and the West.* Production Management, Lund University.

Balan, P. K. 2016. *What makes a mobile game successful.* Computer science, TEXAS A&M UNIVERSITY.

Ballard, J. 2017. *The Best Way to Invest in Mobile Gaming* [Online]. Available at: https://www.fool.com/investing/2017/03/24/the-best-way-to-invest-in-mobile-gaming.aspx [Accessed: 27 Mar. 2017].

Banga, C. 2011. *IPhone user interface cookbook: A concise dissection of Apple's iOS user interface design principles*. Birmingham: Packet Pub.

Barraclough, C. 2015. *The most violent, bloody, gory mobile games for iOS and Android* [Online]. Available at: https://recombu.com/mobile/article/the-most-violent-bloody-gory-mobile-games-for-ios-and-android [Accessed: 06 Feb. 2017].

Bartle, R. 1996. Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD research* 1(1), p. 19.

Bartle, R. A. 2004. *Designing virtual worlds*. Indianapolis, Ind.: New Riders Pub.

Bartlmae, K. 2000. An experience factory for knowledge discovery in databases: Representing KDD-experience. In: Khosrowpour, M. ed. *Challenges of Information Technology Management in the 21st Century*. Anchorage, Alaska, USA: IGI Global, pp. 1064–1065.

Bates, D. and Pow, H. 2013. *Lanza's descent to madness and murder: Sandy Hook shooter notched up 83,000 online kills including 22,000 'head shots' using violent games to train himself for his massacre* [Online]. Available at: http://www.dailymail.co.uk/news/article-2516427/Sandy-Hook-shooter-Adam-Lanza-83k-online-kills-massacre.html [Accessed: 16 Feb. 2017].

Beckett, S. 2016. *Mobile Gaming Apps: Past, Present, And Future* [Online]. Available at: https://www.casino.org/blog/mobile-gaming-apps-past-present-and-future/ [Accessed: 07 Feb. 2017].

Bharathi, A. K. B. G. et al. 2016. Knowledge discovery of game design features by mining user-generated feedback. *Computers in Human Behavior* 60, pp. 361–371.

Bialek, W. et al. 2001. Predictability, complexity, and learning. *Neural computation* 13(11), pp. 2409–2463.

Bialek, W. et al. 2015. Predictive information in a sensory population. *Proceedings of the National Academy of Sciences of the United States of America* 112(22), pp. 6908–6913.

BigML. 2017a. *Anomaly Detection and More!* [Online]. Available at: https://blog.bigml.com/2014/09/26/bigml-late-summer-release-anomaly-detection-and-more/ [Accessed: 15 Apr. 2017].

BigML. 2017b. *Building Your Models. Support* [Online]. Available at: https://support.bigml.com/hc/en-us/articles/206616219-How-is-the-confidence-being-estimated-in-a-decision-tree-model- [Accessed: 20 May. 2017].

BigML. 2017c. *Classification and Regression with the BigML Dashboard* [Online]. Available at: https://static.bigml.com/pdf/BigML_Classification_and_Regression.pdf [Accessed: 21 May. 2017].

Blumler, J. G. and Katz, E. eds. 1974. *The Uses of Mass Communication.* Beverly Hills: SAGE Publications.

Bogost, I. 2011. *Persuasive Games: Exploitationware* [Online]. Available at: http://www.gamasutra.com/view/feature/134735/persuasive_games_exploitationware.php [Accessed: 22 Mar. 2017].

Böhmer, M. et al. 2011. Falling asleep with Angry Birds, Facebook and Kindle. In: Bylund, M. et al. eds. *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. New York, NY: ACM, pp. 47–56.

Bonnington, C. 2015. *In Less Than Two Years, a Smartphone Could Be Your Only Computer* [Online]. Available at: https://www.wired.com/2015/02/smartphone-only-computer/ [Accessed: 27 Aug. 2016].

Bradshaw, T. 2017. Apple App Store sales leap 40% in 2016: Financial Times [Online]. Available at: https://www.ft.com/content/61d88902-d333-11e6-b06b-680c49b4b4c0 [Accessed: 28 Mar. 2016].

Brodie, M. 2016. *The Mobile Revolution: Reconstructing Familiar Interactions* [Online]. Available at: https://uxmag.com/articles/the-mobile-revolution-reconstructing-familiar-interactions [Accessed: 08 Aug. 2016].

Buchheim, C. et al. 2006. Drawing rooted trees in linear time. *Software: Practice and Experience* 36(6), pp. 651–665.

Chaffey, D. 2017. *Mobile Marketing Statistics compilation* [Online]. Available at: http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/ [Accessed: 20 Mar. 2017].

Chapman, P. et al. 2000. *CRISP-DM 1.0 Step-by-step data mining guide*. Spss.

Chen, C. 2014. *Are you addicted to Candy Crush Saga? An exploratory study of linking psychological factors to mobile social game addiction.* Science, The Chinese University of Hong Kong.

Chen, C. and Leung, L. 2016. Are you addicted to Candy Crush Saga?: An exploratory study linking psychological factors to mobile social game addiction. *Telematics and Informatics* 33(4), pp. 1155–1166.

Chen, Y. et al. eds. 2017. *Smart Graphics: 13th International Symposium, SG 2015, Chengdu, China, August 26-28, 2015, Revised Selected Papers.* Cham: Springer International Publishing.

Chiong, C. and Shuler, C. 2010. Learning: is there an app for that? Investigations of young children's usage and learning with mobile devices and apps. *The Joan Ganz Cooney Center at Sesame Workshop*.

Clark, J. F. 2012. *History of Mobile Applications: MAS 490: Theory and Practice of Mobile Applications* [Online]. Available at: http://www.uky.edu/~jclark/mas490apps/History%20of%20Mobile%20Apps.pdf [Accessed: 05 Feb. 2017].

Clark, T. E. and McCracken, M. W. 2012. Reality Checks and Comparisons of Nested Predictive Models. *Journal of Business & Economic Statistics* 30(1), pp. 53–66.

Clover, J. 2016. *Apple's Best of the App Store in 2016:' Prisma' and 'Clash Royale' Win Top Honors* [Online]. Available at: https://www.macrumors.com/2016/12/06/apple-best-of-app-store-2016/ [Accessed: 10 Apr. 2017].

Cooper, m. 2013. *A History of Mobile Phones* [Online]. Available at: http://www.stelladoradus.com/history-mobile-phones/ [Accessed: 24 Feb. 2017].

Cui, M. and Zhu, L. 2017. Usability Evaluation Methods of User Interface Based on Mobile Games Using Fuzzy Methods. In: Chen, Y. et al. eds. *Smart Graphics: 13th International Symposium, SG 2015, Chengdu, China, August 26-28, 2015, Revised Selected Papers*. Cham: Springer International Publishing, pp. 124–131.

Daily, C. N. 2016. *Big Fish Games Acquisition Drives Record Revenue Growth for Churchill Downs Incorporated* [Online]. Available at: http://www.casinonewsdaily.com/2016/02/25/big-fish-games-acquisition-drives-record-revenue-growth-churchill-downs-incorporated/ [Accessed:

Davis, J. 2012. *Shuffle Bags: Making Random() Feel More Random* [Online]. Available at: https://gamedevelopment.tutsplus.com/tutorials/shuffle-bags-making-random-feel-more-random--gamedev-1249 [Accessed: 24 Mar. 2016].

Day, N. 2016. *Monetizing Mobile Gaming* [Online]. Available at: https://techcrunch.com/2016/02/28/mobile-gaming-trends/ [Accessed: 20 Aug. 2016].

Deloitte. 2016. *Mobile games: leading, but less lucrative* [Online]. Available at: https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Technology-Media-Telecommunications/gx-tmt-prediction-mobile-games.pdf [Accessed: 25 Mar. 2017].

Diebold, F. X. and Kilian, L. 2001. Measuring predictability: Theory and macroeconomic applications. *Journal of Applied Econometrics* 16(6), pp. 657–669.

Diebold, F. X. and Mariano, R. S. 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* 13(3), pp. 253–263.

Dobosz, K. and Ptak, J. 2016. How to Control a Mobile Game. In: Miesenberger, K. et al. eds. *Computers Helping People with Special Needs: 15th International Conference, ICCHP 2016, Linz, Austria, July 13-15, 2016, Proceedings, Part I*. Cham: Springer International Publishing, pp. 523–529.

Dogtiev, A. 2015. *App Revenue Statistics 2015* [Online]. Available at: http://www.businessofapps.com/app-revenue-statistics/ [Accessed: 21 Jan. 2017].

Doolittle, J. et al. 2012. *White Paper : Building a mobile application development framework*. USA: Intel.

Dragona, D. 2013. *Counter-gamification: Emerging forms of resistance in social networking platforms*. Gamification Lab at Centre for Digital Cultures.

Engelhardt, C. R. et al. 2011. This is your brain on violent video games: Neural desensitization to violence predicts increased aggression following violent video game exposure. *Journal of Experimental Social Psychology* 47(5), pp. 1033–1036.

Feijoo, C. et al. 2012. Mobile gaming: Industry challenges and policy implications. *Telecommunications Policy* 36(3), pp. 212–221.

Fields, T. 2014. *Mobile & Social Game Design: Monetization Methods and Mechanics, Second Edition*. Taylor & Francis.

Filho, V. V. et al. 2014. Deepening the Understanding of Mobile Game.*SBGames*. Los Alamitos, CA: Conference Publishing Services, IEEE Computer Society, pp. 183–192.

Frick, L. 2016. *Your data : from creepy to irresistible* [Online]. Available at: https://storage.googleapis.com/think-v2-emea/v2/f6421_Google%20Madab%20-%20ebook%20v05%20content%20with%20titles%20(2).pdf [Accessed: 22 Feb. 2017].

Friedman, J. H. 1997. On Bias, Variance, 0/1—Loss, and the Curse-of-Dimensionality. *Data Mining and Knowledge Discovery* 1(1), pp. 55–77.

Game Science 2010. *Mobile Game App Lifecycle* [Online]. Available at: http://nkeegamedev.blogspot.ae/2010_05_01_archive.html [Accessed: 28 Dec. 2016].

Gaming, T. 2017. *Top Grossing iPhone - Games* [Online]. Available at: https://thinkgaming.com/app-sales-data/ [Accessed: 12 Apr. 2017].

Ganguin, S. and Hoblitz, A. 2012. Mobile Worlds: Mobile Gaming and Learning? In: Göbel, S. et al. eds. *E-Learning and Games for Training, Education, Health and Sports: 7th International Conference, Edutainment 2012 and 3rd International Conference, GameDays 2012, Darmstadt, Germany, September 18-20, 2012. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 111–120.

Ganguly, A. 2016. *3 Trusted Ways to Drive Revenue in Strategy Games* [Online]. Available at: http://blogs.shephertz.com/2016/07/13/trusted-ways-drive-revenue-strategy-games/?utm_campaign=QuoraAnswers&utm_medium=questions&utm_source=Quora&utm_term=quoraans [Accessed: 08 Apr. 2017].

Gao, S. et al. 2014. The Adoption of Mobile Games in China: An Empirical Study. In: Liu, K. ed. *Service science and knowledge innovation*. Vol. 426. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 368–377.

Geisser, S. 1993. *Predictive inference: An introduction*. New York: Chapman & Hall.

Gelman, A. et al. 2014. Understanding predictive information criteria for Bayesian models. *Statistics and Computing* 24(6), pp. 997–1016.

Göbel, S. ed. 2015. *Serious Games: First Joint International Conference, JCSG 2015, Huddersfield, UK, June 3-4, 2015 Proceedings / Stefan Göbel [and 5 others] (Eds.).* Cham: Springer International Publishing.

Google. 2015. *Mobile App marketing Insights: How Consumers Really Find and Use Your Apps* [Online]. Available at: https://think.storage.googleapis.com/docs/mobile-app-marketing-insights.pdf [Accessed: 22 Jan. 2017].

Grace, L. 2005. *Game type and game genre* [Online]. Available at: http://www.lgrace.com/documents/Game_types_and_genres.pdf [Accessed: 09 Apr. 2017].

Granger, C. W. J. and Newbold, P. 1986. *Forecasting economic time series*. 2nd ed. ed. Orlando: Academic Press.

Grannell, C. and Saldana, G. 2017. *The best free iPhone games on the planet* [Online]. Available at: http://www.techradar.com/news/phone-and-communications/mobile-phones/60-best-free-iphone-games-on-the-planet-669893 [Accessed: 13 Apr. 2017].

Griffiths, M. 1995. Technological addictions.*Clinical Psychology Forum*. p. 14.

Griffiths, M. D. 2013. Adolescent mobile phone addiction: a cause for concern? *Education and Health* 31, pp. 76–78.

Gruber, M. 2017. *Decision Modeling and CRISP-DM for Modern Data Science Projects* [Online]. Available at: http://www.decisionmanagementsolutions.com/crisp-dm-and-decision-modeling-for-modern-data-science-projects/ [Accessed: 26 Mar. 2017].

Grugier, M. 2016. *The digital age of data art* [Online]. Available at: https://techcrunch.com/2016/05/08/the-digital-age-of-data-art/ [Accessed: 22 Jan. 2017].

Gualtieri, M. 2011. *Forrester's Mobile App Design Context: Location, Locomotion, Immediacy, Intimacy, And Device* [Online]. Available at: http://blogs.forrester.com/mike_gualtieri/11-04-13-forresters_mobile_app_design_context_location_locomotion_immediacy_intimacy_and_device [Accessed: 20 Mar. 2017].

Hamer, C. ed. 2007. *Creating Mobile Games: Using Java ME Platform to Put the Fun into Your Mobile Device and Cell Phone.* 1 ed. Berkeley, CA: Apress.

Han, J. et al. 2004. Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach. *Data Mining and Knowledge Discovery* 8(1), pp. 53–87.

Hansen, P. R. 2005. A Test for Superior Predictive Ability. *Journal of Business & Economic Statistics* 23(4), pp. 365–380.

Hee, O. 1966. Contributed Papers: Marketing, Prices, and Consumption: Tests for Predictability of Statistical Models. *Journal of Farm Economics* 48(5), p. 1479.

Hesselman, C. and Giannelli, C. eds. 2009. *Mobile Wireless Middleware, Operating Systems, and Applications - Workshops: Mobilware 2009 Workshops, Berlin, Germany, April 2009, Revised Selected Papers.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Hirsch, D. and Madria, S. 2013. Data Replication in Cooperative Mobile Ad-Hoc Networks. *Mobile Networks and Applications* 18(2), pp. 237–252.

Holzer, A. and Ondrus, J. 2009. Trends in Mobile Application Development. In: Hesselman, C. and Giannelli, C. eds. *Mobile Wireless Middleware, Operating Systems, and Applications - Workshops: Mobilware 2009 Workshops, Berlin, Germany, April 2009, Revised Selected Papers*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 55–64.

Huang, A. 2011. *Top 7 Social Game Metrics of 2011: The Rise of the Social Web*.

Hubrich, K. and West, K. D. 2010. Forecast evaluation of small nested model sets. *Journal of Applied Econometrics* 25(4), pp. 574–594.

Hwong, C. 2016a. *The most popular mobile game genres: Who plays what, when?* [Online]. Available at: http://www.vertoanalytics.com/the-most-popular-mobile-game-genres-who-plays-what-when/ [Accessed: May 13. 2017].

Hwong, C. 2016b. *Report: Leveling Up Your Mobile Game: Using Audience Measurement Data to Boost User Acquisition and Engagement*.

Instal. 2014. *Best Practices for Mobile and App Business: Acquire app's users and monetize mobile traffic* [Online]. Available at: https://storage.googleapis.com/instal-dashboard/static/public/pdf/Best_practices_for_mobile_and_app_business.pdf [Accessed: 02 Mar. 2017].

iTunes. 2016. *iTunes Preview: Snake 2K by dsd 164* [Online]. Available at: https://itunes.apple.com/us/app/snake-2k/id527862914?mt=8 [Accessed: 23 Jan. 2017].

James. 2016. *Why Mobile Games Make You Their Slave and You Pay for the Privilege* [Online]. Available at: http://joyofandroid.com/mobile-gaming-addiction/ [Accessed: 18 Mar. 2017].

Jaslow, R. 2013. *Violent video games and mass violence: A complex link* [Online]. Available at: http://www.cbsnews.com/news/violent-video-games-and-mass-violence-a-complex-link/ [Accessed: 07 Mar. 2017].

Jenkins, H. 2004. *Reality Bytes: Eight Myths About Video Games Debunked* [Online]. Available at: http://www.pbs.org/kcts/videogamerevolution/impact/myths.html [Accessed: 12 Feb. 2017].

Jeong, Y.-S. ed. 2014. *Ubiquitous information technologies and applications: CUTE 2013 / Young-Sik Jeong [and three others], editors.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Jiang, K. 2011. *The Dangers of Gamification: Why We Shouldn't Build a Game Layer on Top of the World* [Online]. Available at: https://krystlejiang.files.wordpress.com/2011/07/the-dangers-of-gamification.pdf [Accessed: 25 Mar. 2017].

Jin, D. Y. 2017a. The Emergence of Asian Mobile Games: Definitions, Industries, and Trajectories. In: Jin, D.Y. ed. *Mobile Gaming in Asia: Politics, Culture and Emerging Technologies*. Dordrecht: Springer Netherlands, pp. 3–20.

Jin, D. Y. ed. 2017b. *Mobile Gaming in Asia: Politics, Culture and Emerging Technologies.* Dordrecht: Springer Netherlands.

Julkunen, V.-P. 2016. *The Effect of Licensed IP on Mobile Game Success Measured!* [Online]. Available at: http://www.gamerefinery.com/licensed-ip-mobile-game-effect/ [Accessed: 15 Mar. 2017].

Kanerva, T. 2017. *Fire Emblem: Heroes – GameRefinery's Recipe for Bouncing the Game Back to the Top* [Online]. Available at: http://www.gamerefinery.com/fire-emblem-heroes-gamerefinerys-recipe/ [Accessed: 25 Mar. 2017].

Katz, E. et al. 1974. Utilization of mass communication by the individual. In: Blumler, J.G. and Katz, E. eds. *The Uses of Mass Communication*. Beverly Hills: SAGE Publications, pp. 19–32.

Kaye, L. 2016. Applying Psychology within Games Development: What Can the Gaming Industry Learn from the Discipline? In: Attrill, A. and Fullwood, C. eds. *Applied Cyberpsychology: Practical Applications of Cyberpsychological Theory and Research*. London: Palgrave Macmillan UK, pp. 179–196.

Kelion, L. 2016. *Candy Crush Saga: Life beyond level 2,000* [Online]. Available at: http://www.bbc.com/news/technology-37484114 [Accessed: 07 Apr. 2017].

Kennerly, D. 2003. *Better Game Design Through Data Mining* [Online]. Available at: http://www.gamasutra.com/view/feature/131225/better_game_design_through_data_.php [Accessed: 22 Mar. 2017].

Kerr, D. 2013. *Ralph Nader: Video game firms are 'electronic child molesters'* [Online]. Available at: https://www.cnet.com/news/ralph-nader-video-game-firms-are-electronic-child-molesters/ [Accessed: 21 Jan. 2017].

Khosrowpour, M. ed. 2000. *Challenges of Information Technology Management in the 21st Century.* Anchorage, Alaska, USA: IGI Global.

Koetsier, J. 2017. *How mobile users make app install decisions on Google Play and the App Store* [Online]. Available at: https://www.tune.com/blog/mobile-users-make-app-install-decisions-google-play-app-store/ [Accessed: 08 Apr. 2017].

Koster-Hale, J. and Saxe, R. 2013. Theory of mind: a neural prediction problem. *Neuron* 79(5), pp. 836–848.

Krauss, J. et al. 2008. Predicting Movie Success and Academy Awards through Sentiment and Social Network Analysis. *ECIS 2008 Proceedings* 116.

Lai, A. S. Y. and Leung, S. Y. 2014. Ubiquitous Mobile Game Development Using Arduino on Android Platform. In: Jeong, Y.-S. ed. *Ubiquitous information technologies and applications*. Vol. 280. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 271–276.

Laughlin, D. 2012. *The Gamification of Mobile Games* [Online]. Available at: http://flurrymobile.tumblr.com/post/113380478710/the-gamification-of-mobile-games [Accessed: 16 Jan. 2017].

Laurent, M. and Thomson, J. A. 1988. The Role of Visual Information in Control of a Constrained Locomotor Task. *Journal of Motor Behavior* 20(1), pp. 17–37.

Law, F. L. et al. 2011. Gamification towards sustainable mobile application. In: Harun, M.F. ed. *5th Malaysian Conference in Software Engineering (MySEC), 2011*. Piscataway, NJ: IEEE, pp. 349–353.

Lee, J. H. et al. 2014. Facet Analysis of Video Game Genres.*iConference 2014 Proceedings*. Vol. 57. iSchools, pp. 125–139.

Lee, O. and Shin, M. 2004. Addictive Consumption of Avatars in Cyberspace. *CyberPsychology & Behavior* 7(4), pp. 417–420.

Leorke, D. 2014. Location-based gaming apps and the commercialization of locative media.*Mobility and Locative Media: Mobile Communication in Hybrid Spaces*. Florence: Taylor & Francis, pp. 132-148.

Lewis, M. 2004. The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention. *Journal of Marketing Research* 41(3), pp. 281–292.

Lewis, M. and Jolly, D. 2012. *Norwegian Defends Shooting and Regrets Death Toll Wasn't Higher* [Online]. Available at: http://www.nytimes.com/2012/04/20/world/europe/norwegian-says-attack-was-in-the-works-for-years.html [Accessed: 11 Feb. 2017].

Life'd. 2017. *10 Most Addictive Free Mobile Game Apps for All Ages* [Online]. Available at: http://www.lifed.com/10-most-addictive-free-mobile-game-apps-for-all-ages [Accessed: 05 Feb. 2017].

Liu, K. ed. 2014. *Service science and knowledge innovation: 15th IFIP WG 8.1 International Conference on Informatics and Semiotics in Organisations, ICISO 2014, Shanghai, China, May 23-24, 2014, proceedings / Kecheng Liu, Stephen R. Gulliver, Weizi Li, Changrui Yu (eds.).* Berlin, Heidelberg: Springer Berlin Heidelberg.

Lowood, H. 2009. Videogames in Computer Space: The Complex History of Pong. *IEEE Annals of the History of Computing* 31(3), pp. 5–19.

MADAB 2016. *The Art of Automation: Mobile App Developers' Advisory Board*.

Madigan, J. 2015. *Getting gamers: The psychology of video games and their impact on the people who play them*. Rowman & Littlefield, p. 320.

Mannila, H. et al. 1997. Discovery of Frequent Episodes in Event Sequences. *Data Mining and Knowledge Discovery* 1(3), pp. 259–289.

Mason, J. 2016. *PREDICTIONS FOR SEARCH AND SOCIAL IN 2017* [Online]. Available at: https://www.accuracast.com/articles/digital-marketing-predictions-2017/ [Accessed: 28 Mar. 2017].

Mason, M. 2013. *Understanding a Publishing Contract for Mobile Games* [Online]. Available at: http://developers.magmic.com/understanding-publishing-contract-mobile-games/ [Accessed: 27 Mar. 2017].

Mäyrä, F. 2015. Mobile Games.*The International Encyclopedia of Digital Communication and Society*. John Wiley & Sons, Inc.

Mehdipour, Y. and Zerehkafi, H. 2013. Mobile learning for education: Benefits and challenges. *International Journal of Computational Engineering Research* 3(6), pp. 93–101.

Merhav, N. and Feder, M. 1998. Universal prediction. *IEEE Transactions on Information Theory* 44(6), pp. 2124–2147.

MGA. 2016. *MGA publishes Consultation Feedback on the adoption of Enhanced Automated Reporting Platforms (EARP) to enhance regulatory compliance & reporting* [Online]. Available at: http://www.mga.org.mt/mga-publishes-consultation-feedback-adoption-enhanced-automated-reporting-platforms-earp-enhance-regulatory-compliance-reporting/ [Accessed: 31 Aug. 2016].

MGAIN. 2003. *Mobile entertainment in Europe: Current state of the art. Mobile Entertainment Industry and Culture* [Online]. Available at: http://s3.amazonaws.com/zanran_storage/www.knowledge.hut.fi/ContentPages/1120235.pdf [Accessed: 28 Aug. 2016].

Mifsud, J. 2016. *Usability Testing Of Mobile Applications: A Step-By-Step Guide* [Online]. Available at: http://usabilitygeek.com/usability-testing-mobile-applications/ [Accessed: 22 Mar. 2017].

Miller, C. 2016. *iOS App Store brings in 75% more revenue than Play Store despite difference in downloads* [Online]. Available at: https://9to5mac.com/2016/01/20/app-store-ios-downloads-vs-android-revenue/ [Accessed: 16 Mar. 2017].

Mobileaction. 2017. *Top Charts for iPhone* [Online]. Available at: https://insights.mobileaction.co/appreport/?trackId=1095254858&store=ios [Accessed: 01 Feb. 2017].

Monica, G. 2015. *What You Need to Know About App Distribution in East Asia* [Online]. Available at: http://www.oneskyapp.com/blog/asia-app-distribution/ [Accessed: 31 Aug. 2016].

Moreira, Á. V. et al. 2014. Understanding mobile game success: a study of features related to acquisition, retention and monetization. *SBC Journal on Interactive Systems* 5(2), pp. 2-13.

Naib, M. and Chhabra, A. 2014. Predicting Primary Tumors using Multiclass Classifier Approach of Data Mining. *International Journal of Computer Applications* 96(8), pp. 9–13.

NarayanaSwamy, M. and Hanumanthappa, M. 2012. Predicting Academic Success from Student Enrolment Data using Decision Tree Technique. *International Journal of Applied Information Systems* 4(3), pp. 1–6.

Nasdaq. 2017. *EA Company Financials* [Online]. Available at: http://www.nasdaq.com/symbol/ea/financials?query=income-statement [Accessed: 22 Apr. 2017].

NBA. 2016. *NBA InPlay Autotrader Challenge Official Rules* [Online]. Available at: http://www.nba.com/inplay/rules/ [Accessed: 02 Apr. 2017].

Newzoo. 2016a. *2016 Global Games Market Report: An Overview of Trends & Insights* [Online]. Available at: http://resources.newzoo.com/global-games-market-report [Accessed: 14 Jan. 2017].

Newzoo. 2016b. *THE GLOBAL GAMES MARKET REACHES $99.6 BILLION IN 2016, MOBILE GENERATING 37%* [Online]. Available at: https://newzoo.com/insights/articles/global-games-market-reaches-99-6-billion-2016-mobile-generating-37/ [Accessed: 31 Aug. 2016].

Newzoo. 2017. *TOP 25 COMPANIES BY GAME REVENUES* [Online]. Available at: https://newzoo.com/insights/rankings/top-25-companies-game-revenues/ [Accessed: 01 Apr. 2017].

Nielsen. 2014. *SMARTPHONES: SO MANY APPS, SO MUCH TIME* [Online]. Available at: http://www.nielsen.com/us/en/insights/news/2014/smartphones-so-many-apps--so-much-time.html [Accessed: 22 Jan. 2017].

NIST. 2017. *Measures of Skewness and Kurtosis* [Online]. Available at: http://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm [Accessed: 08 Apr. 2017].

Nizan, Y. 2014. *3 Reasons to Add a Luck Element in Your Games* [Online]. Available at: http://blog.soom.la/2014/04/3-reasons-to-add-a-luck-element-in-your-games.html [Accessed: 05 Apr. 2017].

NPD. 2013. *Over Half of Children Now Using Smart Devices* [Online]. New York: The NPD Group, Inc. Available at: https://www.npd.com/wps/portal/npd/us/news/press-releases/over-half-of-children-now-using-smart-devices/ [Accessed: 13 Feb. 2017].

Osborn, G. 2014. *A HISTORY OF MOBILE GAMING* [Online]. Available at: http://www.proelios.com/wp-content/uploads/2013/12/A-History-of-Mobile-Gaming.pdf [Accessed: 05 Jan. 2017].

Osipov, I. V. et al. 2015. Monetization as a Motivator for the Freemium Educational Platform Growth.

Pei, J. et al. 2001. H-mine: hyper-structure mining of frequent patterns in large databases. In: Cercone, N. et al. eds. *2001 IEEE international conference on data mining*. IEEE Comput. Soc, pp. 441–448.

Poggio, T. and Smale, S. 2003. The Mathematics of Learning: Dealing with Data. *Notices of the American Mathematical Society* 50(5), pp. 537–544.

Prato, G. d. et al. 2014. The Road to E-Services: Online and Mobile Games as Enablers. In: Baek, Y. et al. eds. *Trends and Applications of Serious Gaming and Social Media*. Singapore: Springer Singapore, pp. 15–23.

Prensky, M. 2001. *Digital Natives, Digital Immigrants*. MCB University Press.

Price, D. 2017. *83 best free iPhone games | Best free games for iPhone 2017* [Online]. Available at: http://www.macworld.co.uk/feature/iosapps/83-best-free-iphone-games-ios-2017-3507064/ [Accessed: 06 Apr. 2017].

Purcell, K. et al. 2010. *The Rise of Apps Culture* [Online]. Pew Research Center. Available at: http://pewinternet.org/Reports/2010/The-Rise-of-Apps-Culture.aspx [Accessed: 12 Mar 2017].

Purslow, M. 2016. *Why are gamers awful to each other? The psychology of online abuse* [Online]. Available at: https://www.pcgamesn.com/league-of-legends/why-are-gamers-awful-to-each-other-the-psychology-of-online-abuse [Accessed: 05 Mar. 2017].

PwC. 2012. *How to drive innovation and business growth: Leveraging emerging technology for sustainable growth* [Online]. Available at: https://www.pwc.com/us/en/supply-chain-management/assets/pwc-oracle-innovation-white-paper.pdf [Accessed: 30 Aug. 2016].

Razi, S. 2015. *Level of Frustration Discomfort Tolerance and Social Interaction Anxiety in Online Gaming Addicts* [Online]. Available at: http://joyofandroid.com/mobile-gaming-addiction/ [Accessed: 01 Jan. 2017].

Redwood, G. 2012. *Mobile Gaming: A Usability Study* [Online]. Available at: http://www.simpleusability.com/beinspired/2012/01/mobile-gaming-usability/ [Accessed: 29 Mar. 2017].

Reed, R. and Sims, S. 2014. *Property Development: Chapter 4 Development Finance*. 6th ed. UK: Routledge.

Reingold, E. M. and Tilford, J. S. 1981. Tidier Drawings of Trees. *IEEE Transactions on Software Engineering* SE-7(2), pp. 223–228.

Richter, F. 2013. *Freemium is the No.1 Pricing Strategy in Most App Categories* [Online]. Available at: https://www.statista.com/chart/1733/app-monetization-strategies/ [Accessed: 01 Mar. 2017].

Rogers, A. 2016. *How Did Electronic Arts Perform in Fiscal 2Q17?* [Online]. Available at: http://marketrealist.com/2016/11/mobile-gaming-key-driver-electronic-arts/ [Accessed: 11 May. 2017].

Rubin, A. M. 1994. Media uses and effects: A uses-and-gratifications perspective. In: Bryant, J. and Zillmann, D. eds. *Media Effects: Advances in Theory and Research.* Hillsdale, NJ. Lawrence Erlbaum Associates Inc, pp. 417-436

Rubino, D. 2013. *Gameloft's Asphalt 8: Airborne arrives for Windows Phone 8 as a non-Xbox racer* [Online]. Available at: http://www.windowscentral.com/gamelofts-asphalt-8-airborne-arrives-windows-phone-8 [Accessed: 08 Feb. 2017].

Şad, S. N. and Ebner, M. 2017. *Digital Tools for Seamless Learning*. Information Science Reference.

Şad, S. N. and Göktaş, Ö. 2014. Preservice teachers' perceptions about using mobile phones and laptops in education as mobile learning tools. *British Journal of Educational Technology* 45(4), pp. 606–618.

Schiola, E. 2016. *15 Android and iOS games to keep your kids entertained (and quiet)* [Online]. Available at: http://www.digitaltrends.com/mobile/best-apps-for-kids/ [Accessed: 27 Feb. 2017].

Schultz, W. et al. 1997. A Neural Substrate of Prediction and Reward. *Science* 275(5306), pp. 1593–1599.

Shannon, C. E. 1951. Prediction and Entropy of Printed English. *Bell System Technical Journal* 30(1), pp. 50–64.

Shannon, C. E. 1993. The mind reading machine. In: Shannon, C.E. et al. eds. *Shannon's Collected Papers*. New York, N.Y.: Institute of Electrical and Electronics E#ngineers, pp. 688–689.

Shmueli, G. 2010. To Explain or to Predict? *Statistical Science* 25(3), pp. 289–310.

Shoukry, L. et al. 2015. Blitzmerker: Learning Idioms with a Mobile Game. In: Göbel, S. ed. *Serious Games*. Vol. 9090. Cham: Springer International Publishing, pp. 77–88.

Shuler, C. 2009. *Pockets of Potential: Using Mobile Technologies to Promote Children's Learning*. New York: The Joan Ganz Cooney Center at Sesame Workshop.

Sifa, R. et al. 2015. Predicting Purchase Decisions in Mobile Free-to-Play Games. *The Eleventh AAAI Conference on Artificia Intelligence and Interactive Digital Entertainment (AIIDE-15).*

Squirrel, S. 2011. Red Faction: Armageddon's blood, gore, strong language and violence detailed.

Statista. 2016. *Statistics and facts on Mobile Gaming* [Online]. Available at: https://www.statista.com/topics/1906/mobile-gaming/ [Accessed: 28 Aug. 2016].

Statista. 2017a. *Annual revenue generated by Zynga from 2008 to 2016 (in million U.S. dollars)* [Online]. Available at: https://www.statista.com/statistics/273567/zyngas-annual-revenue/ [Accessed: 15 May. 2017].

Statista. 2017b. *Annual revenue of King.com from 2010 to 2016 (in million U.S. dollars)* [Online]. Available at: https://www.statista.com/statistics/288974/king-annual-revenue/ [Accessed: 03 Apr. 2017].

Statista. 2017c. *Net sales of Sega Sammy Holdings from fiscal 2010 to 2016* [Online]. Available at: https://www.statista.com/statistics/550553/sega-sammy-holdings-net-sales/ [Accessed: 28 Mar. 2017].

Statista. 2017d. *Top grossing iPhone mobile gaming apps in the United States as of January 2017, ranked by daily revenue (in U.S. dollars)* [Online]. Available at: https://www.statista.com/statistics/263988/top-grossing-mobile-ios-gaming-apps-ranked-by-daily-revenue/ [Accessed: 30 Mar. 2017].

Steyerberg, E. W. et al. 2010. Assessing the Performance of Prediction Models. *Epidemiology* 21(1), pp. 128–138.

Strain, M. 2015. *1983 to today: a history of mobile apps*. *The Guardian*. 13 Feb. 2015. [Online]. Available at: https://www.theguardian.com/media-network/2015/feb/13/history-mobile-apps-future-interactive-timeline [Accessed: 07 Jan. 2017].

Strauss, K. 2013. *The $2.4 Million-Per-Day Company: Supercell* [Online]. Available at: https://www.forbes.com/sites/karstenstrauss/2013/04/18/the-2-4-million-per-day-company-supercell/#6af34e786fc1 [Accessed: 20 Mar. 2017].

Surowiecki, J. 2004. *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economics, societies, and nations / James Surowiecki*. 1 ed. New York: Doubleday.

Susi, T. et al. 2007. Serious Games – An Overview: Technical Report No. HSIKITR07001.Sweden.

Takahashi, D. 2016. *Mobile games hit $34.8B in 2015, taking 85% of all app revenues* [Online]. Available at: http://venturebeat.com/2016/02/10/mobile-games-hit-34-8b-in-2015-taking-85-of-all-app-revenues/ [Accessed: 05 Feb. 2017].

Taylor, J. 2017a. *Bringing Business Clarity To CRISP-DM* [Online]. Available at: http://www.kdnuggets.com/2017/01/business-clarity-crisp-dm.html [Accessed: 31 Mar. 2017].

Taylor, J. 2017b. *Four Problems in Using CRISP-DM and How To Fix Them* [Online]. Available at: http://www.kdnuggets.com/2017/01/four-problems-crisp-dm-fix.html [Accessed: 31 Mar. 2017].

Tercek, R. 2007. *1997 to 2007: The First Decade of Mobile Games Part Two* [Online]. San Francisco: Available at: http://www.roberttercek.com/resources/Decade+of+Mobile+Games+PART+TWO+GDCM07+Tercek.pdf [Accessed: 02 Feb. 2017].

Thomas, J. and Kielman, J. 2009. Challenges for Visual Analytics. *Information Visualization* 8(4), pp. 309–314.

Thompson, E. 2016. *Super Mario Run Generates $14M in First Three Days* [Online]. Available at: https://www.appannie.com/insights/mobile-gaming/super-mario-run-generates-14m-first-three-days/ [Accessed: 02 Feb. 2017].

Tsipori, T. 2016. *Caesars receives $4b bids for Playtika - report* [Online]. Available at: http://www.globes.co.il/en/article-caesars-seeking-to-sell-playtika-report-1001124141 [Accessed: 20 Apr. 2017].

Tuckerman, C. J. 2014. PREDICTING MOBILE APPLICATION SUCCESS. *Stanford University Project*.

Unhelkar, B. and Murugesan, S. 2010. The Enterprise Mobile Applications Development Framework. *IT Professional* 12(3), pp. 33–39.

Vankka, E. 2014. *Free-to-play Games: Professionals' Perceptions.* University of Tampere.

Vidyarthi, N. 2012. *How Lessons Learned in Social Will Give You a Head Start in Mobile* [Online]. Available at: http://www.adweek.com/digital/how-lessons-learned-in-social-will-give-you-a-head-start-in-mobile/ [Accessed: 03 Feb. 2017].

Vogel, L. 2014. *What is a Visibility Score and why is it important?* [Online]. Available at: https://www.linkedin.com/pulse/20140921211734-236895113-what-is-a-visibility-score-and-why-is-it-important [Accessed: 13 Apr. 2017].

Vucevic, D. and Yaddow, W. 2012. *Testing the Data Warehouse Practicum: Assuring Data Content, Data Structures and Quality*. Trafford Publishing.

Walker, C. 2014. *Developing Games with Random Number Generators* [Online]. Available at: http://www.gamasutra.com/blogs/CharlotteWalker/20141021/228292/Developing_Games_with_Random_Number_Generators.php [Accessed: 22 Mar. 2017].

Wallis, S. 2013. Binomial Confidence Intervals and Contingency Tests: Mathematical Fundamentals and the Evaluation of Alternative Methods. *Journal of Quantitative Linguistics* 20(3), pp. 178–208.

Walsh, S. 2016. *Gamification as a Strategy for Your Mobile App* [Online]. Available at: http://www.business2community.com/mobile-apps/gamification-strategy-mobile-app-01507503#OzyZ8b0mrXWz72Lq.97 [Accessed: 30 Mar. 2017].

Watch, M. 2017. *Electronic Arts Inc: Annual Financials for Elecrtonic Arts* [Online]. Available at: http://www.marketwatch.com/investing/stock/ea/financials [Accessed: 28 Apr. 2017].

Werbach, K. and Hunter, D. 2012. *For the Win: How Game Thinking Can Revolutionize Your Business*. New York: Wharton Digital Press, p. 149.

West, K. 1996. Asymptotic Inference about Predictive Ability. *Econometrica* 64(5), pp. 1067–1084.

Wilcox, M. and Voskoglou, C. 2015. *Developer Economics: The State of the Developer Nation Q1 2015*. London: VisionMobile.

Willson, M. A. and Leaver, T. 2016. *Social, casual and mobile games: The changing gaming landscape*. Bloomsbury Academic, p. 320.

Winter, S. et al. 2011. Location-Based Mobile Games for Spatial Knowledge Acquisition. *Cognitive Engineering for Mobile GIS* 780, pp. 1-8.

Wolf, M. J. P. 2001. *Genre and the Video Game*. Austin: University of Texas Press.

Wood, C. 2014. *Five of the best: Mobile games with eye candy graphics* [Online]. Available at: http://newatlas.com/best-graphics-mobile-games-2014/33500/ [Accessed: 03 Feb. 2017].

Wooldridge, D. and Schneider, M. 2010. *The business of iPhone app development: Making and marketing apps that succeed*. 1 ed. New York: Apress, p. 408.

Wright, C. 2016. *A Brief History of Mobile Games: In the beginning, there was Snake* [Online]. Pocket gamer. Available at: http://www.pocketgamer.biz/feature/10619/a-brief-history-of-mobile-games-in-the-beginning-there-was-snake/ [Accessed: 11 Apr 2017].

Young, K. S. 1998. Internet Addiction: The Emergence of a New Clinical Disorder. *CyberPsychology & Behavior* 1(3), pp. 237–244.

Zichermann, G. and Cunningham, C. 2011. *Gamification by design: Implementing game mechanics in web and mobile apps*. Sebastopol: O'Reilly Media, p. 182

Zou, Q. et al. 2002. A Pattern Decomposition Algorithm for Data Mining of Frequent Patterns. *Knowledge and Information Systems* 4(4), pp. 466–482.

# APPENDIX A: 50 Game Genre and Developer Dataset

**Table 15. 50 Game Genre and Developer Dataset**

| Game Titles | Genre | Develop | Game Titles | Genre | Develope |
|---|---|---|---|---|---|
| Angry Birds Blast | Kids | Rovio | Madden NFL | Sports | EA Sports |
| Asphalt 8: Airborne | Racing | Gameloft | Minecraft Pocket | Kids | Mojang |
| Beneath the | Kids | Nitrome | MLB Tap Sports | Sports | EA Sports |
| Big Fish Casino | Casino | Big Fish | Mobile Strike | Strategy | Machine |
| Boggle with Friends | Dice | Zynga | Mortal Kombat X | Action | Electronic |
| Boom Beach | Strategy | Supercell | Panda Pop | Puzzle | Jam City |
| Bubble Witch 3 Saga | Puzzle | King.com | Pet Rescue Saga | Arcade | King.com |
| Candy Crush Saga | Puzzle | King.com | Piano Tiles 2 | Music | Cheetah |
| Cally's Cave 3 | Action | Jordan | Plants vs | Adventur | Electronic |
| Clash Royale | Strategy | Supercell | Pokemon Go | Simulatio | Niantic |
| Clash of Clans | Strategy | Supercell | Rogue Runner | Action | WEEVO |
| Clash of Kings | Strategy | Elex Tech | Slotomania | Casino | Playtika |
| Cookie Jam | Puzzle | Jam City | Sonic All Stars | Racing | Sega |
| Crossy Roads | Kids | Hipster | Summoners of | Role | Com2uS |
| CSR Racing 2 | Racing | Natural | Super Mario Run | Arcade | Nintendo |
| Double Down Casino | Casino | Double | Super Stickman | Sports | Noodleca |
| Dynasty Warriors | Action | Nexon | Temple Run 2 | Adventur | Imangi |
| 8 Ball Pool | Action | Miniclip | The Simpsons: | Simulatio | Electronic |
| Farmville: Tropic | Family | Zynga | The Sims: Free | Adventur | Electronic |
| Fire Emblem Heroes | Adventur | Nintendo | The Walking | Adventur | Scopely |
| Game of War - Fire | Role | Machine | Two Dots | Adventur | Playdots |
| Genies & Gems | Board | Jam City | Toy Blast | Arcade | Peak |
| Gummy Drop | Puzzle | Big Fish | Word Cookies | Word | BitMango |
| Kim Kardashian: | Adventur | Glu | World Series of | Card | Playtika |
| King Rabbit | Action | RareSlot | Zynga Poker: | Card | Zynga |

# APPENDIX B: 50 Game Features Dataset

**Table 16. 50 game features Dataset**

| Game Title | FB | II | RFH | LC | SP | VS | CMM | COOO | SC | GA | HC | CUM | NC | AC | IAP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Angry Birds Blast | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Asphalt 8: Airborne | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Beneath the Lighthouse | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| Big Fish Casino | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| Boom Beach | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Boggle with Friends** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| **Bubble Witch 3 Saga** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| **Candy Crush Saga** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| **Cally's Cave 3** | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Clash of Clans** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Clash of Kings** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Clash Royale** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Cookie Jam** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| **Crossy Roads** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **CSR Racing 2** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| **Double Down Casino** | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| **Dynasty Warriors Unleashed** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **8 Ball Pool** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| **Farmville: Tropic Escape** | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Fire Emblem Heroes** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| **Game of War - Fire Age** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Genies and Gems** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Gummy Drop!** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| **Kim Kardashian: Hollywood** | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| **King Rabbit** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| **Madden NFL Mobile** | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Minecraft: Pocket Edition** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **MLB Tap Sports Baseball 2017** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| **Mobile Strike** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Mortal Kombat** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Panda Pop** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| **Pet Rescue Saga** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| **Piano Tiles 2** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Planets & Zombies 2** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Pokemon Go** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Rogue Runner** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Slotomania** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Sonic All Star Racing** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| **Summoners War** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Super Mario Run** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Super Stickman Golf 3** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| **Temple Run 2** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| **The Simpsons(tm): Tapped Out** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **The Sims: Free Play** | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **The Walking Dead: Road to Survival** | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Two Dots** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Toy Blast** | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Word Cookies** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **World Series of Poker: Free Texas Holdem** | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Zynga Poker: Texas Holdem** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

| Game Title | UOOOOO | EO | DO | TS | TB | CUS | Ran | STree | PUP | LVL | Stat | Item | IAP | Unlock |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Angry Birds Blast** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Asphalt 8: Airborne** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| **Beneath the Lighthouse** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **Big Fish Casino** | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Boom Beach** | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| **Boggle with Friends** | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| **Bubble Witch 3 Saga** | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| **Candy Crush Saga** | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| **Cally's Cave 3** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Clash of Clans** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| **Clash of Kings** | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| **Clash Royale** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| **Cookie Jam** | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Crossy Roads** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| **CSR Racing 2** | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| **DoubleDown Casino** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| **Dynasty Warriors Unleashed** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **8 Ball Pool** | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| **Farmville: Tropic Escape** | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| **Fire Emblem Heroes** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| **Game of War - Fire Age** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| **Genies and Gems** | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| **Gummy Drop!** | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| **Kim Kardashian: Hollywood** | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| **King Rabbit** | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| **Madden NFL Mobile** | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Minecraft: Pocket Edition** | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| **MLB Tap Sports Baseball 2017** | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Mobile Strike** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| **Mortal Kombat** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Panda Pop** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| **Pet Rescue Saga** | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Piano Tiles 2** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| **Planets & Zombies 2** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Pokemon Go** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Rogue Runner** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Slotomania** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| **Sonic All Star Racing** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Summoners War** | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Super Mario Run** | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Super Stickman Golf 3** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| **Temple Run 2** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| **The Simpsons(tm): Tapped Out** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **The Sims: Free Play** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **The Walking Dead: Road to Survival** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| **Two Dots** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| **Toy Blast** | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Word Cookies** | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **World Series of Poker: Free Texas Holdem** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Zynga Poker: Texas Holdem** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

# APPENDIX C: Game Feature Associations

**Table 17. Game Feature Associations**

| Antecedent | Consequent | Coverage | Support | Confidence | Leverage | Lift |
|---|---|---|---|---|---|---|
| **Cumulative = 1** | Timed boost = 1 | 65.31% | 57.14% | 87.50% | 15.83% | 4.0833 |
| **Competitive Play = 0** | Line Chat = 0 | 24.49% | 20.41% | 83.33% | 15.41% | 4.0833 |
| **IAP <= 2** | non-cumulative = 0 | 40.82% | 30.61% | 75.00% | 18.12% | 4.0833 |
| **non-cumulative = 0** | IAP <= 2 | 30.61% | 30.61% | 100.00% | 18.12% | 4.0833 |
| **Cumulative = 0** | Timed boost = 0 | 28.57% | 28.57% | 100.00% | 18.08% | 4.0833 |
| **Timed boost = 0** | Cumulative = 0 | 36.74% | 28.57% | 77.78% | 18.08% | 4.0833 |
| **Customizable = 0** | Skill tree = 0 | 30.61% | 28.57% | 93.33% | 17.95% | 4.0833 |
| **Time skips = 0** | Customizable = 0 | 34.69% | 28.57% | 82.35% | 17.95% | 4.0833 |
| **Cumulative = 1** | Timed boost = 1 | 65.31% | 57.14% | 87.50% | 15.83% | 4.0833 |
| **Competitive Play = 0** | Line Chat = 0 | 24.49% | 20.41% | 83.33% | 15.41% | 4.0833 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **IAP <= 2** | non-cumulative = 0 | 40.82% | 30.61% | 75.00% | 18.12% | 4.0833 |
| **non-cumulative = 0** | IAP <= 2 | 30.61% | 30.61% | 100.00% | 18.12% | 4.0833 |
| **Cumulative = 0** | Timed boost = 0 | 28.57% | 28.57% | 100.00% | 18.08% | 4.0833 |
| **Timed boost = 0** | Cumulative = 0 | 36.74% | 28.57% | 77.78% | 18.08% | 4.0833 |
| **Customizable = 0** | Skill tree = 0 | 30.61% | 28.57% | 93.33% | 17.95% | 4.0833 |
| **Time skips = 0** | Status upgrade = 0 | 34.69% | 28.57% | 82.35% | 17.95% | 4.0833 |
| **Achievement = 0** | Time skips = 0 | 30.61% | 28.57% | 93.33% | 17.95% | 4.0833 |
| **Achievement = 0** | Status upgrade = 0 | 30.61% | 28.57% | 93.33% | 19.20% | 3.0489 |
| **Customizable = 0** | Time skips = 0 | 30.61% | 28.57% | 93.33% | 17.95% | 3.0489 |
| **Status upgrade = 0** | Achievement = 0 | 30.61% | 28.57% | 93.33% | 19.20% | 3.0489 |
| **Status upgrade = 0** | Customizable = 0 | 30.61% | 28.57% | 93.33% | 19.20% | 3.0489 |
| **Customizable = 0** | Status upgrade = 0 | 30.61% | 28.57% | 93.33% | 19.20% | 3.0489 |
| **Achievement = 0** | Customizable = 0 | 30.61% | 28.57% | 93.33% | 19.20% | 3.0489 |
| **Unique offer = 0** | Event Offer = 0 | 18.37% | 18.37% | 100.00% | 12.75% | 3.2667 |
| **Event Offer = 0** | Unique offer = 0 | 30.61% | 18.37% | 60.00% | 12.75% | 3.2667 |
| **Unique offer = 0** | Event Offer = 0 | 18.37% | 18.37% | 100.00% | 12.75% | 3.2667 |
| **Timed boost = 0** | Time skips = 0 | 36.74% | 34.69% | 94.44% | 21.95% | 2.7222 |
| **Facebook = 1** | Leaderboard = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Invite friends feature = 1** | Facebook = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Leaderboard = 1** | Invite friends feature = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Invite friends feature = 1** | Leaderboard = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Facebook = 1** | Invite friends feature = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **Leaderboard = 1** | Facebook = 1 | 38.78% | 36.74% | 94.74% | 21.70% | 2.4432 |
| **IAP <= 2** | Versus = 0 | 40.82% | 36.74% | 90.00% | 20.91% | 2.3211 |
| **Versus = 0** | IAP <= 2 | 38.78% | 36.74% | 94.74% | 20.91% | 2.3211 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Leaderboard = 1** | Timed boost = 0 | 38.78% | 34.69% | 89.47% | 20.45% | 2.4357 |
| **Timed boost = 0** | Time skips = 0 | 36.74% | 34.69% | 94.44% | 21.95% | 2.7222 |
| **Timed boost = 0** | Leaderboard = 1 | 36.74% | 34.69% | 94.44% | 20.45% | 2.4357 |
| **Skill tree = 0** | Status upgrade = 0 | 34.69% | 30.61% | 88.24% | 19.99% | 2.8824 |
| **Achievement = 0** | Skill tree = 0 | 30.61% | 30.61% | 100.00% | 19.99% | 2.8824 |
| **Skill tree = 0** | Achievement = 0 | 34.69% | 30.61% | 88.24% | 19.99% | 2.8824 |
| **Status upgrade = 0** | Skill tree = 0 | 30.61% | 30.61% | 100.00% | 19.99% | 2.8824 |
| **Invite friends feature = 1** | Request friend help = 1 | 38.78% | 32.65% | 84.21% | 19.99% | 2.5789 |
| **Request friend help = 1** | Invite friends feature = 1 | 32.65% | 32.65% | 100.00% | 19.99% | 2.5789 |
| **Customizable = 0** | Timed boost = 0 | 30.61% | 30.61% | 100.00% | 19.37% | 2.7222 |
| **Timed boost = 0** | Customizable = 0 | 36.74% | 30.61% | 83.33% | 19.37% | 2.7222 |
| **IAP <= 2** | Line chat = 0 | 40.82% | 32.65% | 80.00% | 19.33% | 2.45 |
| **Line chat = 0** | IAP <= 2 | 32.65% | 32.65% | 100.00% | 19.33% | 2.45 |
| **Levels = 1** | Gambling = 0 | 32.65% | 30.61% | 93.75% | 19.28% | 2.7022 |
| **Gambling = 0** | Levels = 1 | 34.69% | 30.61% | 88.24% | 19.28% | 2.7022 |
| **Versus = 0** | non-cumulative = 0 | 38.78% | 30.61% | 78.95% | 18.74% | 2.6902 |
| **non-cumulative = 0** | Versus = 0 | 30.61% | 30.61% | 100.00% | 18.74% | 2.6902 |
| **Cumulative = 0** | Time skips = 0 | 34.69% | 28.57% | 82.35% | 16.54% | 2.4178 |
| **Timed boost = 0** | Leaderboard = 1 | 36.74% | 34.69% | 94.44% | 20.45% | 2.4357 |
| **Skill tree = 0** | Status upgrade = 0 | 34.69% | 30.61% | 88.24% | 19.99% | 2.8824 |
| **Achievement = 0** | Skill tree = 0 | 30.61% | 30.61% | 100.00% | 19.99% | 2.8824 |
| **Skill tree = 0** | Achievement = 0 | 34.69% | 30.61% | 88.24% | 19.99% | 2.8824 |
| **Status upgrade = 0** | Skill tree = 0 | 30.61% | 30.61% | 100.00% | 19.99% | 2.8824 |
| **Invite friends feature = 1** | Request friend help = 1 | 38.78% | 32.65% | 84.21% | 19.99% | 2.5789 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Request friend help = 1** | Invite friends feature = 1 | 32.65% | 32.65% | 100.00% | 19.99% | 2.5789 |
| **Customizable = 0** | Timed boost = 0 | 30.61% | 30.61% | 100.00% | 19.37% | 2.7222 |
| **Timed boost = 0** | Customizable = 0 | 36.74% | 30.61% | 83.33% | 19.37% | 2.7222 |
| **IAP <= 2** | Line chat = 0 | 40.82% | 32.65% | 80.00% | 19.33% | 2.45 |
| **Line chat = 0** | IAP <= 2 | 32.65% | 32.65% | 100.00% | 19.33% | 2.45 |
| **Levels = 1** | Gambling = 0 | 32.65% | 30.61% | 93.75% | 19.28% | 2.7022 |
| **Gambling = 0** | Levels = 1 | 34.69% | 30.61% | 88.24% | 19.28% | 2.7022 |
| **Versus = 0** | non-cumulative = 0 | 38.78% | 30.61% | 78.95% | 18.74% | 2.6902 |
| **non-cumulative = 0** | Versus = 0 | 30.61% | 30.61% | 100.00% | 18.74% | 2.6902 |
| **Time skips = 1** | Timed boost = 1 | 65.31% | 63.27% | 96.88% | 21.95% | 2.4178 |
| **Cumulative = 0** | Time skips = 0 | 34.69% | 28.57% | 82.35% | 16.54% | 2.4178 |
| **Time skips = 0** | Cumulative = 0 | 34.69% | 28.57% | 82.35% | 16.54% | 2.4178 |
| **Time skips = 1** | Cumulative = 1 | 65.31% | 59.18% | 90.63% | 16.54% | 2.4178 |
| **Cumulative = 1** | Time skips = 1 | 65.31% | 59.18% | 90.63% | 16.54% | 2.4178 |
| **Invite friends feature = 0** | Versus = 0 | 20.41% | 20.41% | 100.00% | 16.24% | 2.4178 |
| **Line Chat = 0** | Invite friends feature = 0 | 20.41% | 20.41% | 100.00% | 16.24% | 2.4178 |
| **Invite friends feature = 0** | Line Chat = 0 | 20.41% | 20.41% | 100.00% | 16.24% | 2.4178 |
| **Time skips = 0** | Leaderboard = 1 | 34.69% | 32.65% | 94.12% | 19.20% | 2.4272 |
| **Leaderboard = 1** | Time skips = 0 | 38.78% | 32.65% | 84.21% | 19.20% | 2.4272 |
| **Invite friends feature = 1** | IAP <= 2 | 38.78% | 34.69% | 89.47% | 18.87% | 2.1921 |
| **IAP <= 2** | Leaderboard = 1 | 40.82% | 34.69% | 85.00% | 18.87% | 2.1921 |
| **IAP <= 2** | Facebook = 1 | 40.82% | 34.69% | 85.00% | 18.87% | 2.1921 |
| **Facebook = 1** | IAP <= 2 | 38.78% | 34.69% | 89.47% | 18.87% | 2.1921 |
| **Gambling = 0** | IAP <= 2 | 34.69% | 32.65% | 94.12% | 18.49% | 2.1778 |
| **IAP <= 2** | Gambling = 0 | 40.82% | 32.65% | 80.00% | 18.49% | 2.1778 |
| **Timed boost = 0** | Facebook = 1 | 36.74% | 32.65% | 88.89% | 18.41% | 2.1778 |
| **Timed boost = 0** | Invite friends feature = 1 | 36.74% | 32.65% | 88.89% | 18.41% | 2.1778 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Invite friends feature = 1** | Timed boost = 0 | 38.78% | 32.65% | 84.21% | 18.41% | 2.1778 |
| **Facebook = 1** | Timed boost = 0 | 38.78% | 32.65% | 84.21% | 18.41% | 2.1778 |
| **Line Chat = 1** | Versus = 1 | 79.59% | 79.59% | 100.00% | 16.24% | 2.1778 |
| **Versus = 1** | Line Chat = 1 | 79.59% | 79.59% | 100.00% | 16.24% | 2.1717 |
| **Line Chat = 1** | Invite friends feature = 1 | 79.59% | 79.59% | 100.00% | 16.24% | 2.1717 |
| **Invite friends feature = 1** | Versus = 1 | 79.59% | 79.59% | 100.00% | 16.24% | 2.1717 |
| **Versus = 1** | Invite friends feature = 1 | 79.59% | 79.59% | 100.00% | 16.24% | 2.1717 |
| **Invite friends feature = 1** | Line Chat = 1 | 79.59% | 79.59% | 100.00% | 16.24% | 2.1717 |
| **Timed boost = 0** | Cumulative = 0 | 36.74% | 28.57% | 77.78% | 15.83% | 2.1717 |
| **Cumulative = 0** | Timed boost = 0 | 34.69% | 28.57% | 82.35% | 15.83% | 2.1717 |
| **Timed boost = 1** | Cumulative = 1 | 63.27% | 57.14% | 90.32% | 15.83% | 2.1717 |
| **Line Chat = 0** | Request friend help = 0 | 20.41% | 20.41% | 100.00% | 13.33% | 2.8824 |
| **Invite friends feature = 0** | Request friend help = 0 | 20.41% | 20.41% | 100.00% | 13.33% | 2.8824 |
| **Versus = 0** | Request friend help = 0 | 20.41% | 20.41% | 100.00% | 13.33% | 2.8824 |
| **Request friend help = 0** | Versus = 0 | 34.69% | 20.41% | 58.82% | 13.33% | 2.8824 |
| **Cumulative = 0** | Cooperative Play = 0 | 34.69% | 22.45% | 64.71% | 13.95% | 2.6422 |
| **Time skips = 0** | Leaderboard = 1 | 34.69% | 32.65% | 94.12% | 19.20% | 2.4272 |
| **Leaderboard = 1** | Time skips = 0 | 38.78% | 32.65% | 84.21% | 19.20% | 2.4272 |
| **Invite friends feature = 1** | IAP <= 2 | 38.78% | 34.69% | 89.47% | 18.87% | 2.1921 |
| **Cooperative Play = 0** | Cumulative = 0 | 24.49% | 22.45% | 91.67% | 13.95% | 2.6422 |
| **Cooperative Play = 0** | Time skips = 0 | 24.49% | 22.45% | 91.67% | 13.95% | 2.6422 |
| **Cooperative Play = 1** | Cumulative = 1 | 75.51% | 63.27% | 83.78% | 13.95% | 1.2829 |
| **Time skips = 1** | Cooperative Play = 1 | 65.31% | 63.27% | 96.88% | 13.95% | 1.2829 |
| **Cumulative = 1** | Cooperative Play = 1 | 65.31% | 63.27% | 96.88% | 13.95% | 1.2829 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Cooperative Play = 1** | Time skips = 1 | 75.51% | 63.27% | 83.78% | 13.95% | 1.2829 |
| **Line Chat = 0** | Request friend help = 0 | 20.41% | 20.41% | 100.00% | 13.33% | 2.8824 |
| **Invite friends feature = 0** | Request friend help = 0 | 20.41% | 20.41% | 100.00% | 13.33% | 2.8824 |
| **Request friend help = 0** | Invite friends feature = 0 | 34.69% | 20.41% | 58.82% | 13.33% | 2.8824 |
| **Request friend help = 0** | Line Chat = 0 | 34.69% | 20.41% | 58.82% | 13.33% | 2.8824 |
| **Versus = 0** | Request friend help = 0 | 20.41% | 20.41% | 100.00% | 13.33% | 2.8824 |
| **Versus = 1** | Request friend help = 1 | 79.59% | 65.31% | 82.05% | 13.33% | 1.2564 |
| **Line Chat = 1** | Request friend help = 1 | 79.59% | 65.31% | 82.05% | 13.33% | 1.2564 |
| **Invite friends feature = 1** | Request friend help = 1 | 79.59% | 65.31% | 82.05% | 13.33% | 1.2564 |
| **Request friend help = 1** | Line Chat = 1 | 65.31% | 65.31% | 100.00% | 13.33% | 1.2564 |
| **Request friend help = 1** | Versus = 1 | 65.31% | 65.31% | 100.00% | 13.33% | 1.2564 |
| **Request friend help = 1** | Invite friends feature = 1 | 65.31% | 65.31% | 100.00% | 13.33% | 1.2564 |
| **Unique offer = 0** | Event Offer = 0 | 18.37% | 18.37% | 100.00% | 12.75% | 3.2667 |
| **Event Offer = 0** | Unique offer = 0 | 30.61% | 18.37% | 60.00% | 12.75% | 3.2667 |
| **Unique offer = 1** | Event Offer = 1 | 81.63% | 69.39% | 85.00% | 12.75% | 1.225 |
| **Versus = 1** | Competitive Play = 1 | 79.59% | 75.51% | 94.87% | 15.41% | 1.2564 |
| **Competitive Play = 1** | Versus = 1 | 75.51% | 75.51% | 100.00% | 15.41% | 1.2564 |
| **Competitive Play = 1** | Line Chat = 1 | 75.51% | 75.51% | 100.00% | 15.41% | 1.2564 |
| **Line Chat = 1** | Competitive Play = 1 | 79.59% | 75.51% | 94.87% | 15.41% | 1.2564 |
| **Cooperative Play = 1** | Cumulative = 1 | 75.51% | 63.27% | 83.78% | 13.95% | 1.2829 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Time skips = 1** | Cooperative Play = 1 | 65.31% | 63.27% | 96.88% | 13.95% | 1.2829 |
| **Cumulative = 1** | Cooperative Play = 1 | 65.31% | 63.27% | 96.88% | 13.95% | 1.2829 |
| **Cooperative Play = 1** | Time skips = 1 | 75.51% | 63.27% | 83.78% | 13.95% | 1.2829 |
| **Event Offer = 1** | Unique offer = 1 | 69.39% | 69.39% | 100.00% | 12.75% | 1.225 |
| **Gambling Reward = 0** | Daily Offer = 0 | 8.16% | 8.16% | 100.00% | 7.50% | 1.225 |
| **Daily Offer = 0** | Gambling Reward = 0 | 8.16% | 8.16% | 100.00% | 7.50% | 1.225 |
| **Daily Offer = 1** | Gambling Reward = 1 | 91.84% | 91.84% | 100.00% | 7.50% | 1.0889 |
| **Gambling Reward = 1** | Daily Offer = 1 | 91.84% | 91.84% | 100.00% | 7.50% | 1.0889 |

# APPENDIX D: User Behavior Variables

**Table 18. User Behavior Variables**

| Game Title | New Ins. | Rev | DAU | Vis Score | IAP | Game Title | New Inst. | Rev. | DAU | Vis Score | IAP |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Angry Birds Blast** | 19,686 | $27,921.00 | 251152 | 42 | Yes | Madden NFL Mobil | 42,837 | $108,399.00 | 1,669,588 | 100 | Yes |
| **Asphalt 8: Airborne** | 23,240 | $1,286.00 | 110,301 | 100 | Yes | Minecraft Pocket | 4,078 | $28,853.00 | 22775 | 19 | Yes |
| **Beneath the** | 47,123 | $1,489.00 | | 9 | Yes | MLB Tap Sports Baseball 2017 | 175,424 | $86,841.00 | 50,000 | 100 | Yes |
| **Big Fish Casino** | 35,646 | $270,790.00 | 2,520,336 | 0 | Yes | Mobile Strike | 64,108 | $384,772.00 | 340494 | 82 | Yes |
| **Boggle with Friends** | 43,746 | $322,786.00 | 917,255 | 51 | Yes | Mortal Kombat X | 7,372 | $16,757.00 | 71984 | 100 | Yes |
| **Boom Beach** | 43,746 | $322,786.00 | 917,255 | 28 | Yes | Panda Pop | 39,931 | $36,716.00 | 830,056 | 99 | Yes |
| **Bubble Witch 3 Saga** | 37,487 | $64,163.00 | 1,807,670 | 60 | Yes | Pet Rescue Saga | 31,688 | $92,382.00 | 1,353,772 | 60 | Yes |
| **Candy Crush Saga** | 148,556 | $1,005,806.00 | 9,442,801 | 100 | Yes | Piano Tiles 2 | 57,222 | $4,148.00 | 45,900 | 100 | Yes |
| **Cally's Cave 3** | 13,744 | $19,041.00 | 18567 | 2 | Yes | Plants & Zombies Garden | 118,453 | $5,600.00 | 85131 | 63 | Yes |
| **Clash Royale** | 98,018 | $1,685,714.00 | 5,534,911 | 100 | Yes | Pokemon Go | 29,101 | $166,637.00 | 41,307 | 100 | Yes |
| **Clash of Clans** | 28,789 | $1,480,799.00 | 251559 | 100 | Yes | Rogue Runner | 9,459 | $7,387.00 | 9785 | 0 | Yes |
| **Clash of Kings** | 9,362 | $29,189.00 | 87722 | 44 | Yes | Slotomania | 15,906 | $145,728.00 | 14577 | 59 | Yes |
| **Cookie Jam** | 40,847 | $87,561.00 | 1,480,102 | 62 | Yes | Sonic All Stars Racing | 45,670 | $133,190.00 | 35414 | 22 | Yes |
| **Crossy Roads** | 28,537 | $2,464.00 | 700,000 | 5 | Yes | Summoners of War | 31,989 | $51,191.00 | 569,773 | 40 | Yes |
| **CSR Racing 2** | 41,762 | $29,930.00 | 1,062,688 | 100 | Yes | Super Mario Run | 39,639 | $19,514.00 | 41233 | 100 | Yes |
| **Double Down Casino** | 33,214 | $164,191.00 | 988,702 | 1 | Yes | Super Stick Golf 3 | 17,876 | $306,300.00 | 21014 | 50 | Yes |
| **Dynasty Warriors** | 20,423 | $17,754.00 | 9,277 | 47 | Yes | Temple Run 2 | 30,049 | $2,528.00 | 19055 | 100 | Yes |
| **8 Ball Pool** | 60,126 | $67,428.00 | 1,398,319 | 100 | Yes | The Simpsons: Tapped Out | 23,278 | $102,810.00 | 1,845,065 | 9 | Yes |
| **Farmville: Tropic** | 6,658 | $17,433.00 | 324518 | 49 | Yes | The Sims: Free Play | 24,493 | $28,798.00 | 136672 | 0 | Yes |
| **Fire Emblem Heroes** | 12,037 | $29,630.00 | 11,087 | 0 | Yes | The Walking Dead: Road to Survival | 9,349 | $37,219.00 | 11726 | 52 | Yes |
| **Game of War - Fire** | 43,440 | $1,178,504.00 | 2,554,285 | 62 | Yes | Two Dots | 22,816 | $26,213.00 | 79662 | 35 | Yes |
| **Genies & Gems** | 11,776 | $26,932.00 | 81199 | 20 | Yes | Toy Blast | 276,957 | $83,007.00 | 4,358,095 | 18 | Yes |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Gummy Drop** | 35,038 | $47,369.00 | 615,037 | 65 | Yes | Word Cookies | 57,887 | $74,700.00 | 851,936 | 100 | Yes |
| **Kim Kardashian:** | 37,030 | $97,467.00 | 3,479,869 | 7 | Yes | World Series of Poker: Free Texas Holdem | 25,414 | $28,853.00 | 22775 | 19 | Yes |
| **King Rabbit** | 32,299 | $227,025.00 | 1,006,235 | 91 | Yes | Zynga Poker: Texas Holdem | 23,125 | $86,841.00 | 50,000 | 100 | Yes |

# APPENDIX E: User Behavior Data and Prediction Distributions

**Table 19. User Behavior Data and Prediction Distributions**

| Data Distribution | % | Instances |
|---|---|---|
| **111767** | 4.00% | 2 |
| **147807** | 4.00% | 2 |
| **155461** | 2.00% | 1 |
| **1575** | 6.00% | 3 |
| **164920** | 2.00% | 1 |
| **1706349** | 2.00% | 1 |
| **176386** | 2.00% | 1 |
| **198326** | 2.00% | 2 |
| **20033.5** | 4.00% | 1 |
| **20033.5** | 4.00% | 2 |
| **215844** | 2.00% | 1 |
| **225541** | 4.00% | 1 |
| **233636** | 2.00% | 1 |
| **2403894** | 2.00% | 2 |
| **250021** | 2.00% | 1 |
| **30132** | 4.00% | 1 |
| **308505** | 2.00% | 1 |
| **369597** | 2.00% | 4 |
| **378740** | 2.00% | 1 |
| **43492.25** | 8.00% | 1 |
| **480518** | 2.00% | 1 |
| **552528** | 2.00% | 1 |
| **560186** | 2.00% | 1 |
| **56062** | 2.00% | 1 |
| **571652** | 2.00% | 1 |
| **612720** | 2.00% | 1 |
| **6793230** | 2.00% | 1 |
| **704797** | 4.00% | 2 |
| **72941** | 8.00% | 4 |
| **86383** | 6.00% | 3 |
| **899418** | 2.00% | 1 |
| **977503** | 2.00% | 1 |
| **9916** | 4.00% | 2 |

| Predicted Distribution | % | Instances |
|---|---|---|
| 109998 | 2.00% | 1 |
| 113536 | 2.00% | 1 |
| 11578 | 2.00% | 1 |
| 147141 | 2.00% | 1 |
| 148474 | 2.00% | 1 |
| 155461 | 2.00% | 1 |
| 164920 | 2.00% | 1 |
| 1706349 | 2.00% | 1 |
| 19519 | 2.00% | 1 |
| 197191 | 2.00% | 1 |
| 199462 | 2.00% | 1 |
| 20548 | 2.00% | 1 |
| 215844 | 2.00% | 1 |
| 225367 | 2.00% | 1 |
| 225715 | 2.00% | 1 |
| 233636 | 2.00% | 1 |
| 2403894 | 2.00% | 1 |
| 250021 | 2.00% | 1 |
| 28146 | 2.00% | 1 |
| 308505 | 2.00% | 1 |
| 3117 | 2.00% | 1 |
| 32118 | 2.00% | 1 |
| 369597 | 2.00% | 1 |
| 37182 | 2.00% | 1 |
| 378740 | 2.00% | 1 |
| 44059 | 2.00% | 1 |
| 44105 | 2.00% | 1 |
| 480518 | 2.00% | 1 |
| 48623 | 2.00% | 1 |
| 552528 | 2.00% | 1 |
| 560186 | 2.00% | 1 |
| 56062 | 2.00% | 1 |
| 571652 | 2.00% | 1 |
| 612720 | 2.00% | 1 |
| 6793230 | 2.00% | 1 |

| | | |
|---|---|---|
| 68949 | 2.00% | 1 |
| 702618 | 2.00% | 1 |
| 706976 | 2.00% | 1 |
| 71451 | 2.00% | 1 |
| 74568 | 2.00% | 1 |
| 76796 | 2.00% | 1 |
| 804 | 4.00% | 2 |
| 8254 | 2.00% | 1 |
| 85012 | 2.00% | 1 |
| 86574 | 2.00% | 1 |
| 87563 | 2.00% | 1 |
| 899418 | 2.00% | 1 |
| 977503 | 2.00% | 1 |

# APPENDIX F. User Behavior Model Rules Summary

**Table 20. User Behavior Model Rules Summary**

**Rules summary:**

109998: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and 31941 < DAU Avg <= 37619 and Game Title contains super [Confidence: 1863690.82%]

113536: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and 31941 < DAU Avg <= 37619 and Game Title does not contain super [Confidence: 1863690.82%]

11578: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 3834 < DAU Avg <= 8634 and Game Title contains run [Confidence: 2494801.19%]

147141: (data 0.00% / prediction 2.00%) 3766 < Download Avg <= 36099 and 37619 < DAU Avg <= 53386 and Game Title does not contain casino and Rev Avg > 6227 [Confidence: 702176.33%]

148474: (data 0.00% / prediction 2.00%) 3766 < Download Avg <= 36099 and 37619 < DAU Avg <= 53386 and Game Title does not contain casino and 2612 < Rev Avg <= 6227 [Confidence: 702176.33%]

155461: (data 2.00% / prediction 2.00%) 3766 < Download Avg <= 36099 and 37619 < DAU Avg <= 53386 and Game Title does not contain casino and Rev Avg <= 2612 [Confidence: 2271135.53%]

164920: (data 2.00% / prediction 2.00%) Download Avg <= 3766 and 37619 < DAU Avg <= 53386 and Game Title does not contain casino [Confidence: 3312928.05%]

1706349: (data 2.00% / prediction 2.00%) 36099 < Download Avg <= 109758 and Rev Avg <= 1240626 [Confidence: 367441552.06%]

176386: (data 2.00% / prediction 2.00%) Download Avg <= 36099 and 53386 < DAU Avg <= 65184 and Game Title does not contain casino and Rev Avg > 177032 [Confidence: 6463717.85%]

19519: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 3834 < DAU Avg <= 8634 and Game Title does not contain run and Rev Avg <= 1837 [Confidence: 542040.09%]

197191: (data 0.00% / prediction 2.00%) Download Avg <= 36099 and 53386 < DAU Avg <= 65184 and Game Title does not contain casino and 75846 < Rev Avg <= 177032 [Confidence: 1196280.91%]

199462: (data 0.00% / prediction 2.00%) Download Avg <= 36099 and 53386 < DAU Avg <= 65184 and Game Title does not contain casino and Rev Avg <= 75846 [Confidence: 1196280.91%]

20548: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 3834 < DAU Avg <= 8634 and Game Title does not contain run and Rev Avg > 1837 [Confidence: 542040.09%]

215844: (data 2.00% / prediction 2.00%) Download Avg <= 6507 and 65184 < DAU Avg <= 106674 and Game Title contains holdem and does not contain casino and Rev Avg > 66137 [Confidence: 2961394.59%]

225367: (data 0.00% / prediction 2.00%) Download Avg <= 6507 and 65184 < DAU Avg <= 106674 and Game Title does not contain casino or holdem and Rev Avg > 224471 [Confidence: 183313.85%]

225715: (data 0.00% / prediction 2.00%) Download Avg <= 6507 and 65184 < DAU Avg <= 106674 and Game Title does not contain casino or holdem and 134471 < Rev Avg <= 224471 [Confidence: 183313.85%]

233636: (data 2.00% / prediction 2.00%) Download Avg <= 6507 and 65184 < DAU Avg <= 106674 and Game Title does not contain casino or holdem and 66137 < Rev Avg <= 134471 [Confidence: 2376925.80%]

2403894: (data 2.00% / prediction 2.00%) 36099 < Download Avg <= 109758 and Rev Avg > 1240626 [Confidence: 367441552.06%]

250021: (data 2.00% / prediction 2.00%) Download Avg <= 6507 and 65184 < DAU Avg <= 106674 and Game Title does not contain casino and Rev Avg <= 66137 [Confidence: 4484687.06%]

28146: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 8634 < DAU Avg <= 13263 and 2076 < Rev Avg <= 5423 [Confidence: 2092306.37%]

308505: (data 2.00% / prediction 2.00%) 6507 < Download Avg <= 36099 and 65184 < DAU Avg <= 106674 and Game Title does not contain casino [Confidence: 10697446.49%]

3117: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and DAU Avg <= 3834 and 2141 < Rev Avg <= 2376 [Confidence: 678693.41%]

32118: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 8634 < DAU Avg <= 13263 and Rev Avg <= 2076 [Confidence: 2092306.37%]

369597: (data 2.00% / prediction 2.00%) Download Avg <= 10203 and DAU Avg > 106674 and Rev Avg > 405138 [Confidence: 4816202.70%]

37182: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 8634 < DAU Avg <= 13263 and Rev Avg > 5423 [Confidence: 2301756.68%]

378740: (data 2.00% / prediction 2.00%) Download Avg <= 10203 and DAU Avg > 106674 and Rev Avg <= 405138 [Confidence: 4816202.70%]

44059: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 13263 < DAU Avg <= 37619 and Rev Avg <= 5795 [Confidence: 1332496.91%]

44105: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 13263 < DAU Avg <= 37619 and Rev Avg > 10689 [Confidence: 2379919.48%]

480518: (data 2.00% / prediction 2.00%) 10203 < Download Avg <= 15493 and DAU Avg > 106674 and Rev Avg <= 208828 [Confidence: 16806842.44%]

48623: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and 13263 < DAU Avg <= 37619 and 5795 < Rev Avg <= 10689 [Confidence: 2379919.48%]

552528: (data 2.00% / prediction 2.00%) 10203 < Download Avg <= 15493 and DAU Avg > 106674 and Rev Avg > 1370406 and Game Title does not contain clash [Confidence: 4033958.25%]

560186: (data 2.00% / prediction 2.00%) 10203 < Download Avg <= 15493 and DAU Avg > 106674 and 371320 < Rev Avg <= 1370406 and Game Title does not contain clash [Confidence: 4033958.25%]

56062: (data 2.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and DAU Avg <= 26268 and Game Title contains run [Confidence: 2846645.32%]

571652: (data 2.00% / prediction 2.00%) 10203 < Download Avg <= 15493 and DAU Avg > 106674 and 208828 < Rev Avg <= 371320 and Game Title does not contain clash [Confidence: 4891685.00%]

612720: (data 2.00% / prediction 2.00%) 10203 < Download Avg <= 15493 and DAU Avg > 106674 and Rev Avg > 208828 and Game Title contains clash [Confidence: 10906438.20%]

6793230: (data 2.00% / prediction 2.00%) Download Avg > 109758 [Confidence: 154446714.66%]

68949: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and DAU Avg <= 24499 and Game Title does not contain run and Rev Avg <= 10175 [Confidence: 1317963.38%]

702618: (data 0.00% / prediction 2.00%) 15493 < Download Avg <= 36099 and DAU Avg > 106674 and 768819 < Rev Avg <= 1826292 [Confidence: 2295637.25%]

706976: (data 0.00% / prediction 2.00%) 15493 < Download Avg <= 36099 and DAU Avg > 106674 and Rev Avg <= 768819 [Confidence: 2295637.25%]

71451: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and DAU Avg <= 24499 and Game Title does not contain run and Rev Avg > 10175 [Confidence: 1317963.38%]

74568: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and 24499 < DAU Avg <= 26268 and Game Title does not contain run and Rev Avg <= 67185 [Confidence: 1173630.06%]

76796: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and 24499 < DAU Avg <= 26268 and Game Title does not contain run and Rev Avg > 67185 [Confidence: 1173630.06%]

804: (data 0.00% / prediction 4.00%) Download Avg <= 1926 and DAU Avg <= 3834 and Rev Avg <= 2141 [Confidence: 479908.71%]

8254: (data 0.00% / prediction 2.00%) Download Avg <= 1926 and DAU Avg <= 3834 and Rev Avg > 2376 [Confidence: 1429103.34%]

85012: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and 26268 < DAU Avg <= 27257 [Confidence: 653673.14%]

86574: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and 27257 < DAU Avg <= 31941 and Rev Avg <= 136312 [Confidence: 520969.54%]

87563: (data 0.00% / prediction 2.00%) 1926 < Download Avg <= 36099 and 27257 < DAU Avg <= 31941 and Rev Avg > 136312 [Confidence: 520969.54%]

899418: (data 2.00% / prediction 2.00%) 15493 < Download Avg <= 36099 and DAU Avg > 106674 and Rev Avg > 1826292 [Confidence: 57117520.05%]

977503: (data 2.00% / prediction 2.00%) Download Avg <= 36099 and 37619 < DAU Avg <= 106674 and Game Title contains casino [Confidence: 45352795.19%]

# APPENDIX G. Game Features Model Rules Summary

**Table 21 Game Features Model Rules Summary**

0: (data 34.69% / prediction 34.69%)

· 23.53%: Levels = 0 [Confidence: 51.01%]

· 17.65%: Levels = 1 and Game Name does not contain 2 and Power Up = 0 and Time skips = 1 [Confidence: 43.85%]

· 17.65%: Levels = 1 and Game Name contains 2 and Item Upgrade = 1 [Confidence: 43.85%]

· 11.76%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 0 and Daily Offer = 1 and Skill Tree = 1 [Confidence: 34.24%]

· 5.88%: Levels = 1 and Game Name does not contain 2 and Power Up = 0 and Time skips = 0 and Custom = 0 [Confidence: 20.65%]

· 5.88%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 0 and Daily Offer = 1 and Skill Tree = 0 and Unique offer = 0 [Confidence: 20.65%]

· 5.88%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 1 and Request friend help = 1 and Random Elements = 0 [Confidence: 20.65%]

· 5.88%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 1 and Request friend help = 1 and Random Elements = 1 and Competitive Play = 0 and IAP > 1 [Confidence: 20.65%]

· 5.88%: Levels = 1 and Game Name does not contain 2 or Blast and Power Up = 1 and Noncumulative = 1 and Request friend help = 1 and Random Elements = 1 and Competitive Play = 1 and Cooperative Play = 0 [Confidence: 20.65%]

1: (data 65.31% / prediction 65.31%) Levels = 1
· 40.63%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 1 and Request friend help = 1 and Random Elements = 1 and Competitive Play = 1 and Cooperative Play = 1 [Confidence: 77.19%]

· 28.13%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 1 and Request friend help = 0 [Confidence: 70.08%]

· 9.38%: Levels = 1 and Game Name does not contain 2 and Power Up = 0 and Time skips = 0 and Custom = 1 [Confidence: 43.85%]

· 6.25%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 0 and Daily Offer = 0 [Confidence: 34.24%]

· 6.25%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 0 and Daily Offer = 1 and Skill Tree = 0 and Unique offer = 1 [Confidence: 34.24%]

· 3.13%: Levels = 1 and Game Name does not contain 2 and Power Up = 1 and Noncumulative = 1 and Request friend help = 1 and Random Elements = 1 and Competitive Play = 0 and IAP <= 1 [Confidence: 20.65%]

· 3.13%: Levels = 1 and Game Name contains Blast and does not contain 2 and Power Up = 1 and Noncumulative = 1 and Request friend help = 1 and Random Elements = 1 and Competitive Play = 1 and Cooperative Play = 0 [Confidence: 20.65%] · 3.13%: Levels = 1 and Game Name contains 2 and Item Upgrade = 0 [Confidence: 20.65%]

# APPENDIX H. The actionable model

**Table 22 The actionable model Code**

```
def predict_unlock_content(data={}):
    """ Predictor for Unlock Content from model/58fa79c37e0a8d6c300037bc

        Predictive model by BigML - Machine Learning Made Easy
    """

    import re

    tm_tokens = 'tokens_only'
    tm_full_term = 'full_terms_only'
    tm_all = 'all'

    def term_matches(text, field_name, term):
        """ Counts the number of occurences of term and its variants in text

        """
        forms_list = term_forms[field_name].get(term, [term])
        options = term_analysis[field_name]
        token_mode = options.get('token_mode', tm_tokens)
        case_sensitive = options.get('case_sensitive', False)
        first_term = forms_list[0]
        if token_mode == tm_full_term:
            return full_term_match(text, first_term, case_sensitive)
        else:
            # In token_mode='all' we will match full terms using equals and
            # tokens using contains
            if token_mode == tm_all and len(forms_list) == 1:
                pattern = re.compile(r'^.+\b.+$', re.U)
                if re.match(pattern, first_term):
                    return full_term_match(text, first_term, case_sensitive)
            return term_matches_tokens(text, forms_list, case_sensitive)


    def full_term_match(text, full_term, case_sensitive):
        """Counts the match for full terms according to the case_sensitive
              option

        """
        if not case_sensitive:
```

```
        text = text.lower()
        full_term = full_term.lower()
    return 1 if text == full_term else 0

def get_tokens_flags(case_sensitive):
    """Returns flags for regular expression matching depending on text
          analysis options
    """
    flags = re.U
    if not case_sensitive:
        flags = (re.I | flags)
    return flags

def term_matches_tokens(text, forms_list, case_sensitive):
    """ Counts the number of occurrences of the words in forms_list in
           the text

    """
    flags = get_tokens_flags(case_sensitive)
    expression = ur'(\b|_)%s(\b|_)' % '(\\b|_)|(\\b|_)'.join(forms_list)
    pattern = re.compile(expression, flags=flags)
    matches = re.findall(pattern, text)
    return len(matches)

term_analysis = {
    "game_name": {
        "token_mode": 'all',
        "case_sensitive": True,
    },
}
term_forms = {
    "game_name": {
    },
}
if (data.get('levels') is None):
    return u'1'
if (data['levels'] == '1'):
    if (data.get('game_name') is None):
        return u'1'
    if (term_matches(data['game_name'], "game_name", "2") > 0):
        if (data.get('item_upgrade') is None):
            return u'0'
        if (data['item_upgrade'] == '1'):
            return u'0'
        if (data['item_upgrade'] == '0'):
            return u'1'
    if (term_matches(data['game_name'], "game_name", "2") <= 0):
        if (data.get('power_up') is None):
            return u'1'
        if (data['power_up'] == '1'):
            if (data.get('noncumulative') is None):
                return u'1'
            if (data['noncumulative'] == '1'):
                if (data.get('request_friend_help') is None):
                    return u'1'
                if (data['request_friend_help'] == '1'):
                    if (data.get('random_elements') is None):
                        return u'1'
                    if (data['random_elements'] == '1'):
                        if (data.get('competitive_play') is None):
                            return u'1'
                        if (data['competitive_play'] == '1'):
                            if (data.get('cooperative_play') is None):
                                return u'1'
                            if (data['cooperative_play'] == '1'):
                                return u'1'
                            if (data['cooperative_play'] == '0'):
```

```
                                    if (term_matches(data['game_name'],
"game_name", "Blast") > 0):
                                        return u'1'
                                    if (term_matches(data['game_name'],
"game_name", "Blast") <= 0):
                                        return u'0'
                            if (data['competitive_play'] == '0'):
                                if (data.get('iap') is None):
                                    return u'1'
                                if (data['iap'] > 1):
                                    return u'0'
                                if (data['iap'] <= 1):
                                    return u'1'
                        if (data['random_elements'] == '0'):
                            return u'0'
                    if (data['request_friend_help'] == '0'):
                        return u'1'
                if (data['noncumulative'] == '0'):
                    if (data.get('daily_offer') is None):
                        return u'1'
                    if (data['daily_offer'] == '1'):
                        if (data.get('skill_tree') is None):
                            return u'0'
                        if (data['skill_tree'] == '0'):
                            if (data.get('unique_offer') is None):
                                return u'1'
                            if (data['unique_offer'] == '1'):
                                return u'1'
                            if (data['unique_offer'] == '0'):
                                return u'0'
                        if (data['skill_tree'] == '1'):
                            return u'0'
                    if (data['daily_offer'] == '0'):
                        return u'1'
            if (data['power_up'] == '0'):
                if (data.get('time_skips') is None):
                    return u'0'
                if (data['time_skips'] == '1'):
                    return u'0'
                if (data['time_skips'] == '0'):
                    if (data.get('custom') is None):
                        return u'1'
                    if (data['custom'] == '1'):
                        return u'1'
                    if (data['custom'] == '0'):
                        return u'0'
    if (data['levels'] == '0'):
        return u'0'
```