ABSTRACT

• Digital gaming has a history of more than 50 years. The industry started in the late 1960’s when game titles such as Pong, Centipede and Odyssey were introduced to consumer markets.
• It is reported that in 2011, the total revenue of the industry amounted to about 17 billion USD.
• The aim is to analyze the user reviews collected from online websites such as metacritic.com and gamespot.com in order to understand the sentiment of the user reviews.
• This study will benefit the game developers to identify the interests and expectations of users from a game and also the users who think of buying a game after reading the reviews on different websites.
• For this study, we have collected 10,000 user reviews mostly from metacritic.com and some from gamespot.com.
• We used SAS Enterprise Miner nodes to generate a quick summary of terms and their relationship with each other.
• We used Text Rule Builder node to build model to explain rules that are used to classify reviews as positive and negative.

METHODOLOGY

• It is an essential factor for any business who designs console games to collect and analyze user feedback and make necessary changes for the organization to be able to design what their customers exactly need.
• The process flow below was developed in SAS Enterprise Miner and shows all the nodes used viz., Text Parsing, Text Filter and Text cluster.
• The data from all the user reviews was imported to the enterprise miner using File Import node.

APPROACH

Text Filtering:

• After importing the data, we used Text Parsing node to parse the data i.e., to clean and modify any unstructured data in the imported data.
• The Text Filter node was attached to the Text Parsing node to remove the low frequency terms and also to perform a spell check.
• The import synonym option in the Text Filter node was used to group all the similar words as synonyms and also we can keep or remove the terms in the interactive filter viewer.

Text Clustering:

• The Text cluster node in Enterprise Miner was used to group all the terms which are closely associated with each other into different clusters.
• There are total 10 clusters formulated with related descriptive terms in each cluster which were associated but different from the terms in another cluster.
• Cluster ID 8, which had terms like game, play fun, and multiplayer had the highest frequency percentage of 26%. The frequent occurrence of this cluster indicates that most reviews contained terms from this cluster and also from Cluster ID 7 with frequency percentage of 22%.
• This showed that the players were concerned about good graphics, story, fun time and a good gameplay.
The concept link diagram is developed in Enterprise Miner to determine the strength of association between a dependent term and other terms which were most frequently used with it. The width of each link denotes the strength of association between those two terms and wider the link, stronger is the association. From the figure, it was observed that terms like awesome, gameplay, well, design and story were strongly associated with the term great which only tells us that a game with good design, gameplay and story was perceived to be a good game in users’ view.

**RESULTS**

**Predictive Model Using Text Rule Builder:**

- We used Text Rule builder node inside the SAS Enterprise Miner to build a model to classify reviews into positive and negative
- Text Rule builder node is a Boolean rule based categorizer that automatically generates an ordered set of rules for describing and predicting a target variable.
- For example some of the rules obtained from Text Rule Builder say that, if a review has a term repetitive and it is not a great game then it is negative review with a precision of 94.34%
- And also, rule #8 says if a review has a term great and does not have terms such as money, disappointment, worst then it is a positive review with a precision of 91.4%

**REFERENCES**

- Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS® by Goutam Chakraborty, Murali Pagolu, Satish Garla

**LIMITATIONS AND FUTURE SCOPE**

- The study has several imitations which provides scope for further research and exploration
- We couldn’t include many reviews because they were not in English. Additional linguistic research is needed.
- Advanced analysis could have been done with proper domain expertise.
- Deeper analysis on sentiments of people based on different consoles such as XBOX, PLAYSTATION, PC is what we hope to achieve in the future.

**CONCLUSION**

- Game reviews give an insight of what people expect from a game, this can be used by the developers to come up with games that can satisfy and reach the expectations of the people.
- A Score node can be used to test new reviews. They can be classified into positive and negative reviews with the help of the text rule builder.
- We can also get an insight of what the users want from a game, like graphics and gameplay seem to be most important here, and if the game is repetitive they feel it is boring, this can be seen from the concept link of the term repetitive.
- It can also be seen from the concept link of the term repetitive that fighting games and shooting games are most repetitive and if so, can be really boring.