

# Exploring Gamer Sentiments on Steam using NLP

Christian Anthony C. Stewart  
*Department of Computer, Information  
Sciences and Mathematics  
University of San Carlos  
Cebu City, Philippines  
christianstewart5111@gmail.com*

**Abstract**—Emotion theory proposes that human beings have a limited number of emotions [1]. This paper aims to uncover the underlying sentiments expressed by gamers through the application of Natural Language Processing (NLP) techniques. By examining the intricate language patterns present in user reviews, the objective is to utilize TF-IDF (Term Frequency-Inverse Document Frequency) values to construct a predictive model. This model is designed to predict whether a game will be recommended by a player based on the textual content of the review. Moreover, the study utilizes TextBlob for sentiment analysis to explore potential correlations between the sentiment polarity of reviews and the likelihood of game recommendation. Leveraging a comprehensive training dataset, the model will undergo rigorous training to capture the subtleties of gamer feedback. Subsequently, the model's performance will be evaluated using a separate dataset to ensure its efficacy in generating predictions for new and unseen reviews. Embark on this analytical journey to uncover the motivations behind gamers' recommendations—or lack thereof—for their preferred games.

**Index Terms**—Natural Language Processing, Sentiment Analysis

## I. INTRODUCTION

Understanding the sentiments of gamers is crucial in deciphering their preferences and behaviors within the gaming community. Emotion theory suggests that human beings experience a limited range of emotions, which serve as fundamental drivers behind their actions and interactions. Within the gaming landscape, these emotions play a significant role in shaping user experiences and influencing gaming choices. By delving into the realm of Natural Language Processing (NLP), researchers can unlock valuable insights hidden within user reviews and discussions on gaming platforms like Steam.

The ability to understand human emotions from written text, out of the various modalities, has been a significant challenge in the field of natural language processing. Despite such predicament, sentiment analysis is critical in capturing the nuances of human emotion especially in text and researchers rely on machine learning to develop models to tackle such [2]. In the field of NLP, the challenge lies in extracting and analyzing sentiments from the vast trove of user-generated content on gaming platforms.

By harnessing the power of machine learning algorithms and linguistic analysis, researchers can gain deeper insights into the motivations and preferences driving gamers' recommendations and critiques. Through this exploration of gamer sentiments on Steam using NLP analysis, the study

aims to unveil the underlying emotions shaping the gaming community and pave the way for enhanced user experiences and game development strategies.

The core of the study is aimed to develop a predictive sentiment model for analyzing user reviews of games on the Steam platform.

Specifically, the objectives were the following:

- 1) Preprocess a Steam Reviews Dataset [3].
- 2) Develop a predictive model for sentiment analysis.
- 3) Evaluate the performance of the model.
- 4) Provide a sanity check of predicted results.

## II. REVIEW ON RELATED LITERATURE

### A. Gaming Industry

The gaming industry has grown exponentially over the past few decades, becoming a major segment of the entertainment sector. Alongside this growth, the importance of understanding player sentiment has become critical for game developers and marketers. This growth has resulted in the accumulation of a substantial amount of valuable information, such as reviews and ratings pertaining to a wide range of products and services [4].

Given the industry's vast reach, understanding consumer feedback through analysis is crucial for appreciating market requirements, evaluating game performance, and understanding the impact of games on players.

### B. Sentiment Prediction in Games

Sentiment prediction, also known as sentiment analysis, is the process of determining the emotional tone behind a body of text. In the context of games, it involves analyzing player reviews, social media posts, in-game chat logs, and other forms of user-generated content to gauge player sentiment.

Various techniques are employed for sentiment analysis in the gaming industry, ranging from traditional machine learning approaches to more advanced deep learning models. Some ways to measure the accuracy is by applying and comparing different ML models to predict if a game is recommended or not on Steam based on different features of user reviews, as well as by exploring potential correlations between different groupings of features and their prediction outcomes [5].

### C. Binary Classifier Model: TD-IDF and Logistic Regression

A Baseline Binary Classifier Model using TF-IDF (Term Frequency-Inverse Document Frequency) and Logistic Regression is a foundational approach in natural language processing (NLP) for text classification tasks [6]. This can be a good starting point in sentimental analysis for games.

In the context of sentiment analysis or text classification, the model serves as a benchmark or starting point for more sophisticated models. It involves preprocessing the text data, such as tokenization and removing stop words, followed by transforming the text into numerical vectors using TF-IDF [7]. Then, Logistic Regression is applied to these TF-IDF vectors to train a model that can predict the sentiment or classify the text into predefined categories [4].

While simple, this baseline model provides a straightforward yet effective approach for text classification tasks. It is easy to implement and interpret, making it a suitable choice for initial exploratory analysis or as a reference point for evaluating more complex models.

## III. PROCEDURE AND RESULTS

### A. A Peek at the Data

The training dataset contains a game title column, a year column, a user review column that contains the textual gamer feedback and a user suggestion (recommendation) column. User suggestions are marked as "1" for recommended and "0" otherwise in the "user suggestion" column.

TABLE I  
FIRST 5 ROWS OF STEAM GAME REVIEW DATASET

	title	developer	publisher		tags	overview
0	Spooky's Jump Scare Mansion	Lag Studios	Lag Studios	[Horror, 'Free to Play', 'Cute', 'First-Pers...	Can you survive 1000 rooms of cute terror? Or...	
1	Sakura Clicker	Winged Cloud	Winged Cloud	[Nudity, 'Anime', 'Free to Play', 'Mature'...	The latest entry in the Sakura series is more...	
2	WARMODE	WARTEAM	WARTEAM	[Early Access, 'Free to Play', 'FPS', 'Multi...	Free to play shooter about the confrontation o...	
3	Fractured Space	Edge Case Games Ltd	Edge Case Games Ltd	[Space, 'Multiplayer', 'Free to Play', 'PvP'...	Take the helm of a gigantic capital ship and g...	
4	Counter-Strike: Global Offensive	Valve	Hidden Path Entertainment	Valve	[FPS, 'Multiplayer', 'Shooter', 'Action', 'T...	Counter-Strike: Global Offensive (CS: GO) exa...

We strip the unnecessary information from the dataset as part of the preprocessing and it would look like how it is in Table. II.

TABLE II  
DATA INSTANCE FROM STEAM REVIEWS DATASET

ID	Title	Year	User Review	Suggestion
1	Spooky's...	2016.0	I'm scared and ...	1
2	Spooky's...	2016.0	Best game, better...	1
3	Spooky's...	2016.0	A little iffy on the...	1
4	Spooky's...	2015.0	Great game, fun and...	1
5	Spooky's...	2015.0	Not many games have...	1

Here, we group the reviews by title and investigate the difference in recommendation numbers (recommended or not) to get an intuitive feel for the most and least popular games in the dataset. The resulting percentage is similar to a common gaming measure of how many gamers "like the game", as on Google, etc.

From the Fig. 1, we can obtain the 5 Best and Worst titles based on user suggestions. This is visible in Table III

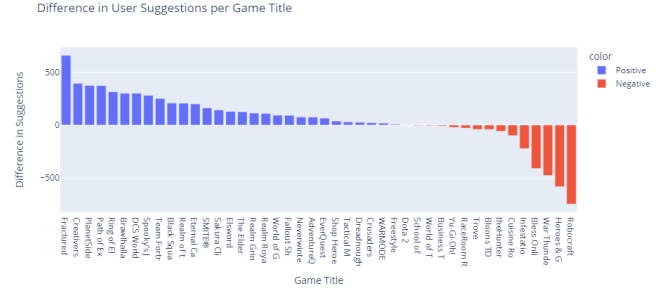


Fig. 1. Difference in User Suggestions per Game Title

and Table IV. In the tables, user suggestions are marked as "1" for recommended and "0" otherwise in the "user suggestion" column.

TABLE III  
5 BEST TITLES

Title	0	1	Difference	Percentage	color
Fractured Space	30	688	658	95.82	Positive
Creativerse	49	443	394	90.04	Positive
PlanetSide 2	49	423	374	89.62	Positive
Path of Exile	43	415	372	90.61	Positive
Ring of Elysium	52	367	315	87.59	Positive

From the game overview dataset, there are about 64 unique titles which is a quite limited number but it is a good start for this term paper.

TABLE IV  
5 WORST TITLES

Title	0	1	Difference	Percentage	color
Infestation	350	129	-221	26.93	Negative
Bless Online	561	151	-410	21.21	Negative
War Thunder	598	122	-476	16.94	Negative
Heroes Generals	663	82	-581	11.01	Negative
Robocraft	794	48	-746	5.70	Negative

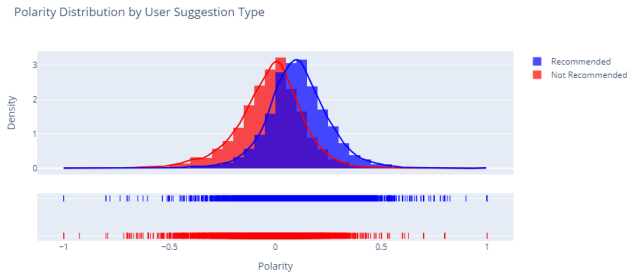
A good way to visualize the most common words in a corpus of text is to generate a word cloud. A word cloud is a visual representation of text data where the size of each word indicates its frequency or importance in the document. Fig. 3 is a word cloud that is generated from words across all reviews which displays the top 10 most frequent words, and then generates and plots a word cloud with the 100 most common words set against a white background.

### B. Result

The baseline model produces an accuracy score of 82.7% in predicting recommendations based on the textual content of user reviews. The Fig. 2 shows a visualization of the predicted labels for the test set and Table.V is the breakdown of the evaluation metrics.



processing (NLP) tasks, such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more. TextBlob is built on top of NLTK (Natural Language Toolkit) and Pattern, two popular NLP libraries in Python [8].

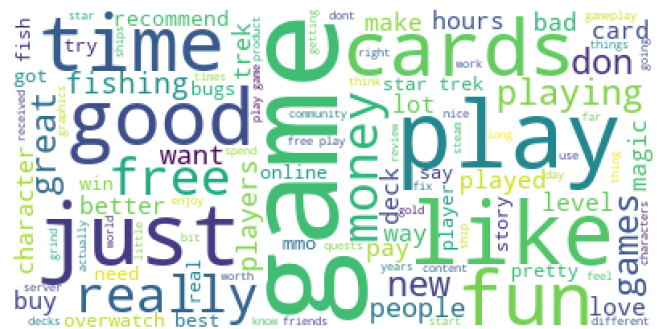


Here, a histogram of the sentiment polarity of reviews is plotted with kernel density estimation (KDE) for sets of recommended and not recommended reviews to visualize the correlation between positive sentiment in recommended reviews and neutral/negative sentiment in non-recommended reviews. Sentiment polarity is categorized as follows: negative ( $< 0$ ), neutral ( $= 0$ ) and positive ( $> 0$ ).

TABLE VII  
POLARITY TABLE OF USER SUGGESTION

Table VII shows this relationship. Positive sentiment polarity reviews yielding suggestions (similar to an accuracy score) are at 78.74% which means more than half of positive reviews align with recommendations for the game while 55.3% of Negative (also neutral reviews without suggestion so skewed in terms of proportion) polarity reviews that did not yield a suggestion. This means that about half of the negative/neutral reviews align with non-recommendations for the games of the entire dataset.

In this section, a different Steam user review dataset is used to see how the developed model performs on further unseen data. The dataset is recorded in 'test.csv' and



A word cloud is generated for the test set (as before) as means of comparison with the training set and for an informative glance at the common terms in this dataset as seen in Fig. 6.

Using the TD-IDF model established previously for the training set. The transformed data is then passed through the optimised logistic regression model to predict whether or not a gamer would suggest the game to others. (1 or 0, respectively, in the dataset). This can be seen in Table VIII.

### B. Sentiment Analysis

Table IX shows this relationship. Positive sentiment polarity reviews yielding suggestions (similar to an accuracy score) are at 80.52% predicted by the model. This is an improvement over the test dataset while 57.05% of Negative (also neutral reviews without suggestion so skewed in terms

Polarity Distribution by User Suggestion Prediction Type

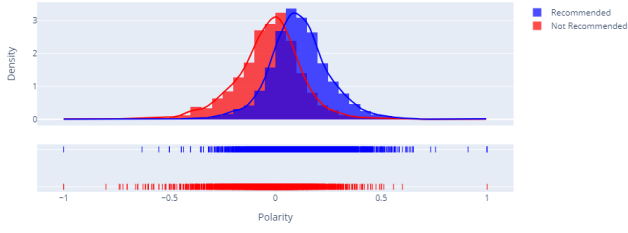


Fig. 7. Polarity Distribution of User Suggestion Prediction

of proportion) polarity reviews that did not yield a suggestion. This means that about half of the negative/neutral predicted suggestions align with non-recommendations for the games of the training dataset which is also an improvement from the test dataset.

TABLE IX  
POLARITY TABLE OF USER SUGGESTION

Sentiment Polarity	Reviews	Total User Suggestion	Percentage
Positive	3812	4734	80.52%
Negative/Neutral	1889	3311	57.05%

### C. Sanity Check of Predicted Results

To ensure the validity of the predicted outcomes without access to the actual user suggestion data for the test dataset, a further verification step is performed by retrieving recommendation data directly from Steam's user review API [9], providing a baseline understanding of anticipated recommendation levels for the games under consideration.

By extracting the Steam game IDs for the games in the dataset, we then use this information to obtain review data for a specific date range (in this case, relevant to the dataset before December 2018). We then calculate the total number of positive and negative recommendations, and compute the ratio of positive to negative recommendations. The extracted positive suggestion rate can then be compared to the predicted suggestion rate of 58.84% achieved using the model.

The ratio of positive to negative recommendations provides an insight into the proportion of positive recommendations expected for a set of reviews for a given game, much like a sample compared to a population (this being the population and the dataset, a sample). In the standard mean calculation above, the spread of user reviews per game is not taken into account. This could lead to skewed results for a larger number of games that are either very positively or negatively rated.

To fix this issue, the calculated positive suggestion ratio should be adjusted to include the proportion of games that make up the dataset; the Up/Down ratios for each game should be weighted depending on the number of reviews for

TABLE X  
COMPARISON OF BASELINE USER REVIEW AND PREDICTED RESULTS

	Game Title	Game ID	Up Total	Down Total	Ratio Up/Down	Weighted Ratio Up/Down	Multiplied Ratios
0	Counter-Strike: Global Offensive	730	2344547	309459	0.883399	0.050841	0.044913
1	World of Warships	552990	6809	2417	0.738023	0.033576	0.024780
2	Star Trek Online	9900	7142	2114	0.771608	0.054286	0.041887
3	Paladins®	444090	147387	26313	0.848515	0.055794	0.047342
4	Shadowverse CCG	453480	4979	1708	0.744579	0.041648	0.031010
5	Tree of Savior (English Ver.)	372000	8845	5516	0.619904	0.048614	0.029942
6	VEGA Conflict	339600	1361	1232	0.524875	0.011287	0.005924
7	Minion Masters	489520	5538	1052	0.840364	0.041888	0.035201
8	The Lord of the Rings Online™	212500	6385	1427	0.817332	0.044600	0.036453
9	Fishing Planet	380600	9582	2966	0.763628	0.047270	0.036097
10	Crush Crush	459620	6137	801	0.884549	0.036613	0.032386
11	Dungeon Defenders II	236110	9691	3415	0.739432	0.050919	0.037651
12	Governor of Poker 3	436150	1953	718	0.731187	0.008089	0.005915
13	Digimon Masters Online	537180	2541	1211	0.677239	0.026096	0.017673
14	Shakes and Fidget	438040	5683	879	0.866047	0.014748	0.012773
15	Champions Online	9880	2019	796	0.717229	0.034145	0.024490
16	Magic Duels	316010	1214	204	0.856135	0.095032	0.081360
17	Aura Kingdom	268420	102	29	0.778626	0.031358	0.024416
18	HY21	433850	103386	82468	0.556275	0.029387	0.016347
19	GUNS UPI	446150	0	0	NaN	NaN	NaN

that game in the dataset as seen in Table X. These weighted ratios can then be used to calculate a weighted average of expected positive recommendations which better represents the test dataset in question.

The overall, weighted suggestion/recommendation rate for the games in the test set is thus: 58.66%. This value is much more in line with the rate calculated from our predictive model: 58.84%.

## VI. CONCLUSION

Indeed, the resulting weighted positive suggestion rate is calculated to be 58.66%, which is much closer to the value (58.84%) based on the predictions of the model. Understanding the sentiments of gamers is paramount in deciphering their preferences and behaviors within the gaming community. Emotion theory suggests that human beings experience a limited range of emotions, which serve as fundamental drivers behind their actions and interactions. These emotions play a significant role in shaping user experiences and influencing gaming choices within the dynamic landscape of gaming platforms like Steam.

Despite the challenges inherent in extracting sentiments from written text, sentiment analysis is critical in capturing the nuances of human emotion, especially in text-based interactions. Through the utilization of natural language processing (NLP) techniques, researchers can delve into the vast trove of user-generated content on gaming platforms to unearth valuable insights hidden within user reviews and discussions.

The core objective of the study is to develop a predictive sentiment model for analyzing user reviews of games on the Steam platform. This involves preprocessing the data, developing machine learning models, and evaluating their performance using standard evaluation metrics. The study also aims to establish a correlation between sentiment polarity in reviews and the likelihood of a game being recommended by the user. Clearly, the objectives of the

paper was met.

Through the rigorous analysis and interpretation of the findings, the study contributes to the advancement of natural language processing techniques in understanding and predicting user sentiments in the gaming domain. Ultimately, this research serves as a foundation for future studies aimed at enhancing user experiences and informing game development strategies in the ever-evolving landscape of digital gaming.

#### REFERENCES

- [1] C. D. Wilson-Mendenhall, L. F. Barrett, and L. W. Barsalou, "Neural evidence that human emotions share core affective properties," *Psychological Science*, vol. 24(6), pp. 947-956, 2013.
- [2] M. Z. Ashgar et.al., "A deep neural network model for the detection and classification of emotions from textual content," *Complexity*, pp. 1-12, 2022.
- [3] Steam Game Review Dataset. (2020, December 24). Kaggle. <https://www.kaggle.com/datasets/arashnic/game-review-dataset>
- [4] Jaiswal, P., Setia, H., Raghuwanshi, P., and Randhawa, P. (2023). A Natural Language Processing Model for Predicting Five-Star Ratings of Video Games on Short-Text Reviews. *Engineering Proceedings*. <https://doi.org/10.3390/engproc2023059058>
- [5] Zhang, A. (2022). Sentiment without Sentiment Analysis: Using the Recommendation Outcome of Steam Game Reviews as Sentiment Predictor. Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. <https://openprairie.sdstate.edu/etd2/490/>
- [6] Karabiber, F. (n.d.). Binary classification. <https://www.learnatasci.com/glossary/binary-classification/>
- [7] Zuo, Z. (2018). Sentiment Analysis of Steam Review Datasets using Naive Bayes and Decision Tree Classifier. University of Illinois. <https://www.ideals.illinois.edu/handle/2142/100126>
- [8] PYPI · The Python Package Index. (n.d.). PyPI. <https://pypi.org/>
- [9] Steam Store. (n.d.). <https://store.steampowered.com/>