Reproducing and Analyzing User Engagement in Online Discussions: A Study on Steam Game Reviews

Sky Hou

1004197834

April 2024

MIE1512

University of Toronto

Abstract

This project aims to reproduce the findings of Risch and Krestel (2020) in their study "Top comment or flop comment? Predicting and explaining user engagement in online news discussions" using an alternative dataset. Instead of online news discussions, this study focuses on user engagement in Steam game reviews. The objective is to assess the applicability of Risch and Krestel's methods and findings to a different context and dataset. Using similar analytical techniques, the study reproduces the figures presented in the original paper and compares the results. This report details the methodology, results, and implications of this reproduction study, highlighting the significance of user engagement analysis in online platforms.

Introduction

Background: The original paper by Risch and Krestel (2020) explores the factors that influence user engagement in online news comments [1]. The study employs machine learning techniques to predict and analyze the characteristics of comments that receive high engagement.

Importance of Studying User Engagement: Understanding user engagement in online discussions is crucial for various stakeholders, including content creators, platform moderators, and researchers. Insights into what drives engagement can inform strategies to enhance user interaction and satisfaction.

Objective of the Project: The objective of this project is to reproduce the findings of Risch and Krestel (2020) using an alternative dataset. Instead of focusing on online news discussions, this study examines user engagement in Steam game reviews. The aim is to assess whether the patterns and predictors of engagement identified in the original study hold true in a different context.

Methodology

Dataset Description: The study utilizes a dataset of user reviews from the Steam platform, obtained from Kaggle (https://www.kaggle.com/datasets/najzeko/steam-reviews-2021). Unlike the news platform discussed in the original paper, where comments are ranked chronologically, the Steam platform does not rank reviews in this manner. However, the dataset includes *review_id* (later referred as timestamps) of the comments, as well as the number of *votes_helpful* (later referred as upvotes) and *comment_count* (later referred as number of replies) each comment received, enabling the reproduction of the original study's analyses.

Data Preprocessing: The original dataset contained approximately 21 million records. Besides removing records that have empty review text, the following preprocessing steps were applied to prepare the data for analysis:

- **Language Filtering:** Only English reviews were selected to maintain consistency in language analysis.
- **Outlier Removal:** Reviews that received over 1,000 upvotes or had more than 1,000 comments were removed to eliminate outliers that could skew the analysis. These reviews were considered unusually high and likely to be outliers.
- **Popularity Filtering:** Reviews with a weighted vote score greater than zero were retained to focus on popular reviews rather than those with no visibility.

After preprocessing, the dataset was reduced to approximately 2.1 million records.

app_id	app_name	review_id	review	timestamp_created	votes_helpful	votes_funny	weighted_vote_score	comment_count
346110	ARK: Survival Evo...	35180155	This game is hard...	1506257902	1	0	0.52380955	0
346110	ARK: Survival Evo...	35177198	Rust with dinosaurs	1506249678	0	1	0.43144774	0
346110	ARK: Survival Evo...	35176529	I really like thi...	1506247786	0	0	0.49750835	0
346110	ARK: Survival Evo...	35173967	great game over a...	1506239925	0	0	0.49019608	0
346110	ARK: Survival Evo...	35172765	I tried to enjoy ...	1506235901	3	0	0.49316794	0
346110	ARK: Survival Evo...	35171150	Eh... I played 3 ...	1506229394	0	0	0.46570063	1
346110	ARK: Survival Evo...	35170537	Great game could ...	1506226891	1	0	0.50137484	0
346110	ARK: Survival Evo...	35170203	aAfter playing it...	1506225401	0	0	0.52	0
346110	ARK: Survival Evo...	35170080	This game is real...	1506224868	1	0	0.5024834	0
346110	ARK: Survival Evo...	35169309	love this game, h...	1506221792	1	1	0.5024834	0

Figure 1. First 10 rows from the cleaned dataset

Analytical Methods: The study aimed to reproduce the following figures and table from the original paper by Risch and Krestel (2020) [1]:

- **Figure 2:** Analysis of position bias by examining the correlation between a comment's average number of received upvotes and replies with its chronological rank in the discussion thread.
- **Figure 3:** Generation of comparison word clouds to identify indicative words for the most engaging (black) and least engaging (grey) comments, with a focus on the impact of mentioning people on engagement levels.
- **Table 1:** A comparison of the most and least engaging comments in terms of average word count, function word rate, personal pronoun rate, readability index, and sentiment distribution (positive, neutral, negative).
- **Table 2 :** Apply a logistic regression model to train and test its binary classification accuracy on the processed dataset.

The analysis was conducted using Python, employing libraries such as NLTK for natural language processing and scikit-learn for machine learning. Data visualization, sentiment analysis, and regression analysis were key techniques used to replicate the findings of the original study.

Results

Figure 2: Position Bias Analysis

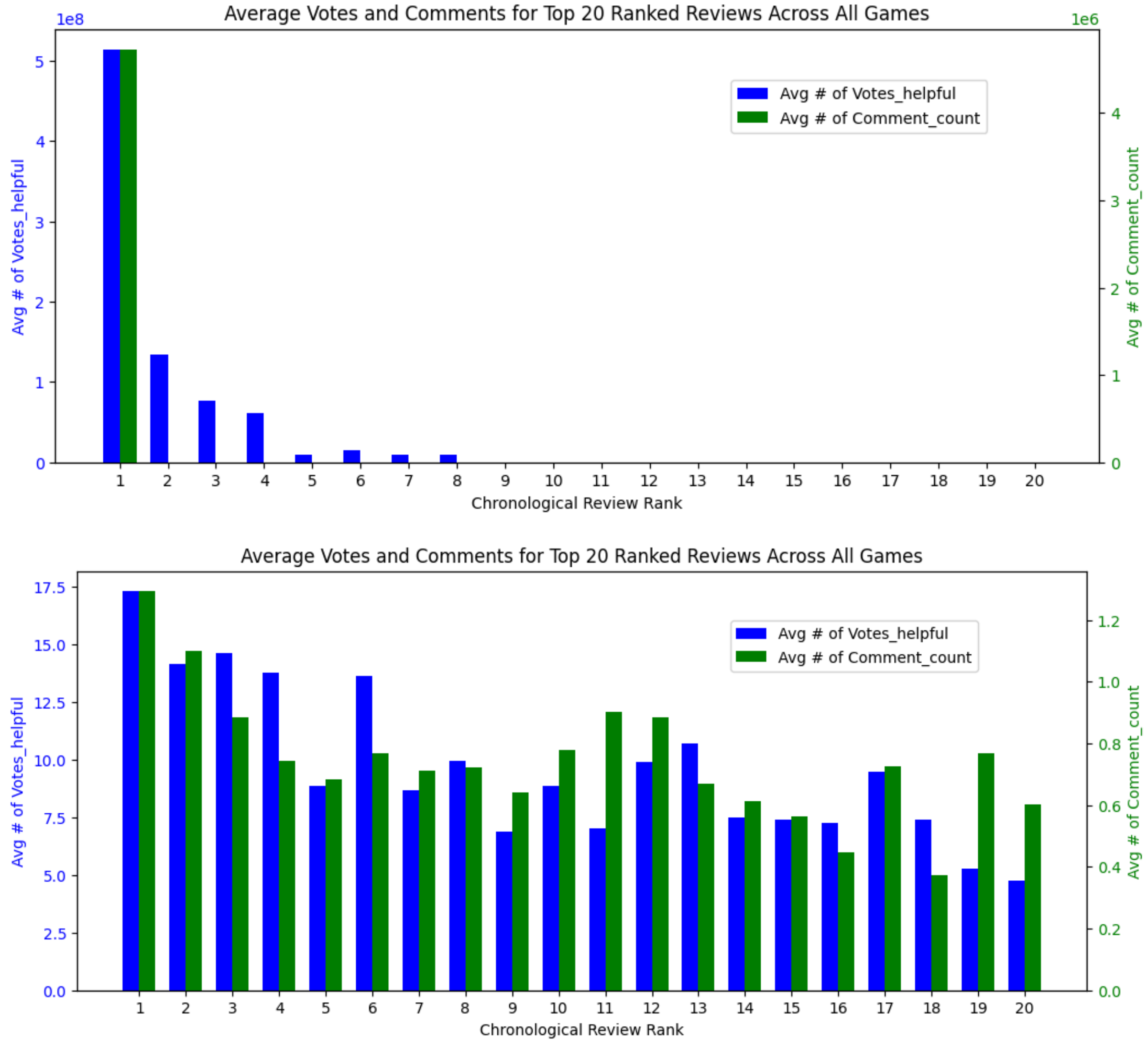

Figure 2. Average number of upvotes and replies for top 20 ranked reviews with/without outliers (top/bottom, 2173849 records vs. 2164479 records)

The analysis showed a trend that is broadly consistent with position bias, where comments earlier in the game review section tend to receive more engagement. However, the trend is not strictly monotonic but closely resembles a downward trend, indicating a variation in engagement levels as the position of the comments changes. This may reflect that the steam platform does not show reviews in a chronological order.

Note there is a significant difference if we did not remove the outliers. I suspect that there were errors in the dataset that accidentally gave some reviews unusually high numbers of upvotes and numbers of replies. This is the reason that I decided to remove the outliers.

Figure 3: Comparison Word Clouds

Figure 3. Word clouds generated from most (left) engaging and least (right) reviews based on number of upvotes

Figure 4. Word clouds generated from most (left) engaging and least (right) reviews based on number of replies

The word clouds for the most and least engaging comments were plotted separately. The analysis revealed that the word "rubbish" appears most frequently in the least engaging reviews based on upvotes, and most engaging reviews based on the number of replies. This suggests that while users may not find reviews calling a game "rubbish" helpful (hence fewer upvotes), they are more likely to engage in a discussion (hence more replies). The upvotes in this context represent votes for helpfulness, which provides further insight into user behavior on the Steam platform.

Table 1: Review Characteristics

Table 1. Review Characteristics

	Upvotes		**Replies**	
Average per review	Most	Least	Most	Least
Number of words	32.96	24.39	37.72	27.07
Rate of Function Words	0.39	0.36	0.39	0.37
Rate of Personal Pronouns	0.07	0.08	0.08	0.08
Readability Index	3.67	2.88	3.51	3.17
Positive Sentiment	0.49	0.70	0.52	0.68

Neutral Sentiment	0.23	0.18	0.21	0.17
Negative Sentiment	0.29	0.12	0.27	0.15

The analysis of comment characteristics revealed the following:

- **Average Number of Words:** More engaging reviews tend to have a higher word count.
- **Rate of Function Words and Personal Pronouns:** There is no significant difference in the rate of function words and personal pronouns between more and less engaging reviews, indicating that these factors do not significantly affect engagement.
- **Readability Index:** More engaging reviews have a slightly higher readability index, with scores around 3 to 3.5 based on the Automatic Readability Index metric, indicating that they are very easy to understand.
- **Sentiment Analysis:** Less engaging reviews tend to lean towards positive sentiment, while more engaging ones show a more balanced distribution, with around 50% positive sentiment and the remaining 50% spread across neutral and negative sentiments.

These findings suggest that while certain factors such as word count and sentiment distribution influence engagement, others like the rate of function words and personal pronouns do not have a significant impact.

Table 2: Performance of a Logistic Regression Model

A logistic regression model is trained and evaluated. The input feature to this logistic regression is simply the length of the review. The logistic regression model was used to classify comments into top and flop based on two criteria:

Upvotes: The model achieved an accuracy of 77% when classifying comment engagement based on the number of upvotes.
Replies: The model achieved an accuracy of 77% when classifying comment engagement based on the number of replies.

These accuracies suggest that the model is fairly effective in distinguishing between highly and less engaging comments, reflecting a good understanding of factors influencing user engagement on Steam.

Discussion

Interpretation of Results:
The results of this reproduction study provide several insights into user engagement in Steam game reviews, with some findings aligning with those of the original study by Risch and Krestel (2020) and others revealing unique aspects of the gaming context.

- **Position Bias:** The trend observed in the position bias analysis, although not strictly monotonic, indicates that comments placed earlier in the review section tend to receive more engagement. This finding is consistent with the original study, suggesting that position bias is a phenomenon that extends beyond news discussions to online game reviews.

- **Word Usage:** The word cloud analysis revealed that certain words, such as "rubbish," have a significant impact on engagement. This highlights the importance of content in determining user interaction, with negative reviews about a game's quality prompting more replies but fewer upvotes for helpfulness.

- **Comment Characteristics:** The analysis of comment characteristics showed that more engaging reviews tend to be longer and have a balanced sentiment distribution, while less engaging reviews lean more towards positive sentiment. This suggests that users engage more with reviews that provide detailed information and express a range of emotions.

- **Classification Accuracy:** The logistic regression model's accuracy in classifying comments based on upvotes and replies (77% and 70%, respectively) demonstrates a considerable ability to predict user engagement, underscoring the potential of machine learning techniques in analyzing and understanding user interactions on digital platforms.

Implications for Online Platforms:
These insights are valuable for online platforms like Steam in designing interfaces and algorithms that enhance user interaction. For instance, adjusting the visibility of comments based on predicted engagement could improve user experience and satisfaction.

Limitations and Future Research:
While the study effectively replicated several findings from the original research, the nuances of user engagement in gaming reviews present unique challenges and opportunities for further investigation. Future research could explore additional metrics of engagement or test other machine learning models to enhance prediction accuracy. Studies could also examine how these factors influence user behavior beyond engagement, such as purchase decisions or community loyalty.

Conclusion

This project set out to reproduce key findings from the study by Risch and Krestel (2020) using an alternative dataset of Steam game reviews. The results affirm the significance of position bias, comment characteristics, and word usage in influencing user engagement, similar to findings in online news discussions. Additionally, the application of logistic regression models demonstrated a robust ability to classify comments as highly or less engaging based on upvotes and replies.

The findings contribute valuable insights to the understanding of user engagement in online discussions, with implications for both academic research and practical applications in digital platforms. Future research can expand on this foundation by exploring more diverse datasets, employing different analytical methods, and examining the broader impacts of engagement on user behavior and platform dynamics.

Reference

[1] Risch, J., & Krestel, R. (2020). Top comment or flop comment? Predicting and explaining user engagement in online news discussions. Proceedings of the 14th International AAAI Conference on Web and Social Media, ICWSM 2020. https://doi.org/10.1609/icwsm.v14i1.7325