Exploring User Sentiments and Themes: A Comprehensive Analysis of Software Reviews

Reuben Jooste
University of Pretoria
Computer Science Department
u21457060@tuks.co.za

ABSTRACT

This paper aims to enhance customer experience for software products by utilizing data mining techniques. In this paper we demonstrate how using sentiment analysis and topic modelling can help stakeholders derive meaningful insights from customer reviews. Furthermore, we classify sentiments as positive, negative, or neutral using a pre-trained model such as TextBlob. Additionally, we employ the Latent Dirichlet Allocation (LDA) for analysing the key themes in the reviews such as easy of use or product quality. By analysing statistical evaluations and visualisations such as coherence scores and distribution plots, we explore sentiment trends as well as how sentiments and topics vary across different ratings. Finally, we present our insights obtained from our results which stakeholders can use to enhance product offerings, increase customer satisfaction, and make data-driven decisions.

1. INTRODUCTION

Over recent years, natural language processing (NLP) has become a popular research field, specifically using data mining techniques to extract insights from customer reviews. This can also be referred to as *sentiment analysis* which involves analysing people's opinions, attitudes, and/or emotions towards certain topics. These topics can include movies, services, and software products (e.g., mobile apps, games, etc.). By extracting insights from the customer reviews, businesses and organisations can gain a better understanding on how customers feel about their products. Furthermore, businesses and organisations can use these insights to facilitate better decision-making and increase customer satifaction. In this study we are only concerned with the top 10 most popular products from a software reviews dataset. We will first analyse the sentiment towards the popular products and then identify the key topics discussed in the reviews for each of these products. Finally, we combine topic modelling with sentiment analysis to analyse the sentiment towards the key topics for each product. In doing this, we hope to answer the following research questions:

1. How can insights derived from sentiment analysis and topic modeling be used to enhance customer experience and improve product offerings?
2. How effective are sentiment analysis and topic modeling techniques, such as TextBlob and LDA, in analyzing customer feedback for software products?
3. Do customers show more positive, negative, or neutral sentiment towards software products?
4. What are the key themes that customers frequently discuss in software product reviews?

2. RELATED WORK

Previous work involved using sentiment analysis or topic modelling to extract insights from customer reviews. Sentiment Analysis is the process of analysing opinions from text to help identify the attitude of a person towards a certain topic. On the other hand, topic modelling is the process of analysing text to identify the key themes discussed, thereby grouping the words in the text in different categories or topics.

Different techniques have been explored to improve the accuracy of sentiment analysis tasks. These techniques include deep-learning approaches using models such as BERT or RoBERT as well as ensemble learning approaches, or simple supervised learning methods such as k-nearest neighbour, naive-bayes, and random forest [5; 7; 9; 12]. Despite, the impressive performance shown by leveraging deep learning approaches, sentiment analysis is still a difficulty task due to language models struggling with understanding sarcastic texts [13].

Topic modelling has also received a lot of attention over the years with applications in various domains. Researchers have explored using topic modelling to extract key themes for understanding customer perspectives on project portfolio management software [3], identifying reasons behind software product returns [2], and have also shown to be useful to enhance software solutions [10]. However, identifying the key themes only provide high-level insights to managers and stakeholders about customer satisfaction towards their products. Studies have shown that models such as the TextBlob and VADER can have a great impact for analysing sentiment and using it in conjunction with topic modelling can help provide deeper insights to trending products [1]. More recent work has shown that using TextBlob for sentiment analysis together with Latent Dirichlet Allocation (LDA) for topic modelling can greatly impact e-commerce businesses by allowing them to improve their understanding on customer satisfaction towards their products [4].

Despite the recent advancements, there are many challenges associated with sentiment analysis. Among these challenges are cross-domain sentiment classification which is often difficult to implement due to variability in aspect information, and the inclusion of irrelevant features often leads to incorrect classifications. Therefore, including only relevant features is a crucial part in data preprocessing [11].

In this paper we aim to address issues related to ambiguity and informality found in texts that can lead to misclassified sentiment [13]. Our approach involves carefully analysing the dataset and preprocessing it to ensure that contextually relevant stop words are preserved. Additionally, we carefully apply feature selection and feature engineering to ensure only the most revelant features are used. Furthermore, a pre-trained model, TextBlob, is used to classify a subset of the data as positive, negative, or neutral and label the data for sentiment analysis. Based on the labelled data we will analyse the sentiment trends by product ratings, time, and purchase verification. For topic modelling, we will apply similar preprocessing techniques as the sentiment analysis approach, with an emphasis on text cleaning. This will remove any special characters and irrelevant content. Using the LDA method, we will extract the key topics from all the reviews. Based on the extracted topics, we can then examine the distribution across different product ratings as well as track how certain topics got more popular or less popular over time.

We anticipate providing stakeholders with visualized insights into sentiment trends by product ratings, time, and purchase verification status, as well as identifying high-frequency topics and their variations across customer segments. These insights will enable more data-driven decision-making for improved customer satisfaction and product development strategies.

3. PROBLEM STATEMENT

Businesses such as Amazon try to improve their products and services daily and sometimes it's challenging to understand how they can improve them. By using data mining techniques such as topic modelling and sentiment analysis we can derive meaningful insights from customer's reviews on Amazon's software products. It is quite interesting how something as simple as understanding customer feedback can lead to an increase in customer experience. We believe that analysing the reviews can reveal sentiment trends, trending topics/products, and what customers like and dislike, allowing the business to make changes to their services to better fit customer needs. In doing so it can lead to an increase in customer satisfaction and potentially increased revenue for the company.

4. METHODOLOGY

This section presents a detailed explanation for the methodology used for implementing our research. We first discuss our exploratory data analysis (EDA) approach and present our initial insights gained from it. Next, we discuss the data preprocessing steps taken to ensure our data is in a consistent format such that we can use it for conducting sentiment analysis and topic modelling. Finally, we discuss the implementation of using the TextBlob model to analyse the sentiment of the reviews and the LDA method used for identifying the key themes within the reviews. Figure 1 shows an illustration of our methodology.

4.1 Data Analysis and Preprocessing

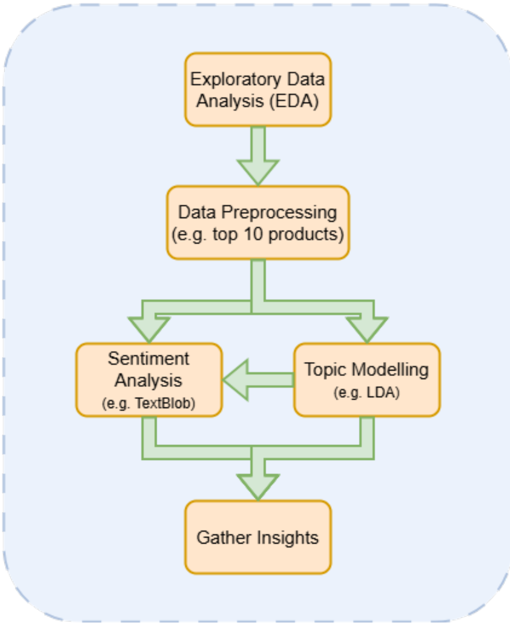

Figure 1: Methodology

Our methodology started by first analysing the datasets. We used the publicly available Software reviews dataset created by Amazon [6]. The dataset can be downloaded by clicking the 'review' link next to the 'Software' category. The dataset contains approximately 4.8 million records, with approximately 2.6 million unique users' reviews in JSON format. Our dataset contains a few irrelevant fields such as the 'images' and 'helpful_vote'. We applied feature selection methods to ensure only revelant features are used in our study. Feature engineering was also used to add an additional 'category' attribute by mapping the 'parent_asin' attribute to the revelant metadata ID. This enabled us to analyse the products by grouping them by category name.

Descriptive statistics was also used to gain a better understanding of the attribute values. The mean rating given by customers are 3.94/5 which means most customers gave positive ratings for the products. The minimum rating is 1/5 with the maximum being 5/5. This indicates that there are some customers who disliked the product completely. However, based on the Interquartile range (IQR) of 3/5 (25%) to 5/5 (75%) the number of negative ratings given are in the minority. In terms of the 'timestamp' attribute, we analysed that reviews between 1996 and 2023 are included in our dataset. However, the IQR falls between 2014 and 2018, indicating that customers made a lot more reviews within this time period. Furthermore, the mean timestamp is 2016 which could indicate that a lot of new products or updates were released causing customers to leave more reviews in 2016.

In terms of erroneous records, the dataset included 6134 missing values for the 'category' field with 51027 duplicate records. Due to the dataset being so large, we dropped all missing values and duplicate records as these only made up $\approx$ 1% of the data and will have little affect on our analysis. We then also plotted a few visualisations to help analyse the average rating for each of the top 10 products and the number of reviews received over the years per product. These visualisations are shown in figures 2 and 3 (due to the size of the visualisations, we could only show two products' activity per year). Finally, we also filtered our dataset to only include the top 10 products since we want to analyse the sentiment and popular topics towards the most popular products.

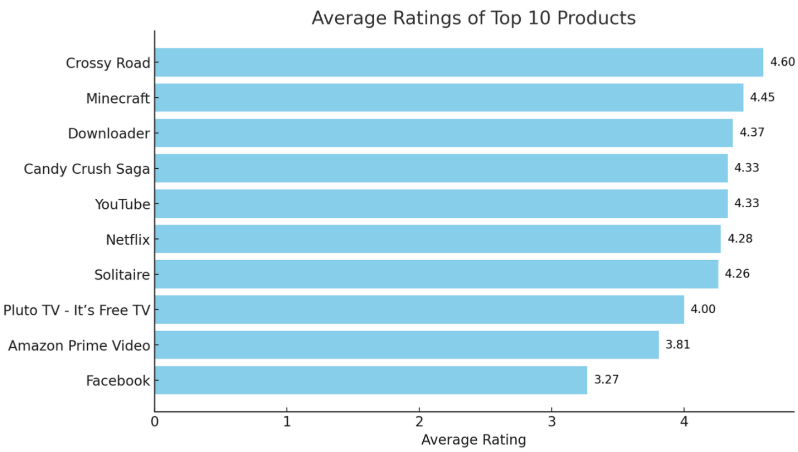

Figure 2: Average rating per top 10 product

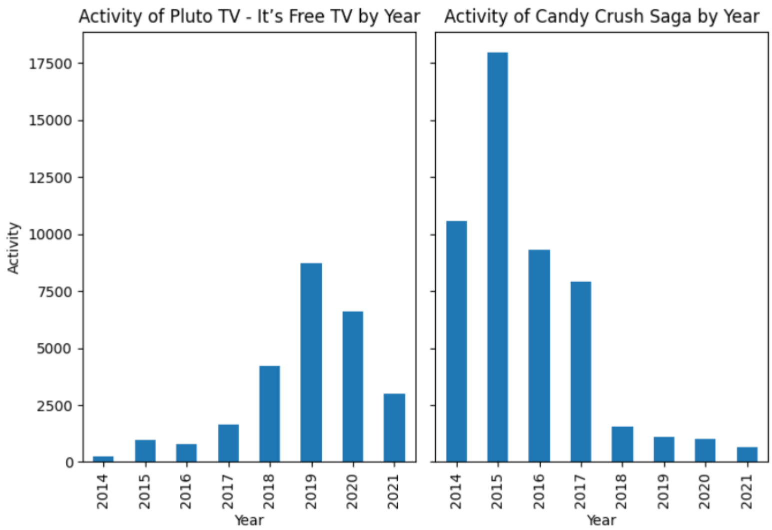

Figure 3: Activity per top 10 product (only two products shown)

4.2 Sentiment Analysis and Topic Modelling

After analysing our data and ensuring it is free of any erroneous data we implement the sentiment analysis and topic modelling techniques. The pre-trained, TextBlob, model is used for analysing the sentiment of the top 10 most popular products. The popularity of a product is based on the number of reviews the product has received. Before using the TextBlob model, we first preprocess the text to tokenise it and ensure all irrelevant stop words are removed. The tokenised text is then passed into the model to perform sentiment analysis. The sentiments are grouped into three categories namely positive, neutral, and negative which help us in labelling the customer reviews accordingly. The classified sentiments are then further used to extract insights for each separate product. The labelled data is then used to plot visualisations for each of the top 10 products and we present our findings based on the sentiment towards each product.

After performing sentiment analysis, we use the Latent Dirichlet Allocation (LDA) algorithm to extract the key themes discussed in the software product reviews and we present our initial findings from analysing the key themes discussed. Additionally, we apply sentiment analysis on these topics to analyse the sentiment towards key topics such as easy-of-use. Visualisations such as word clouds are also used to help visualise the most frequent words used in customer reviews.

Finally, we present all our findings obtained from performing sentiment analysis and topic modelling. These insights will enable stakeholders to make data-driven decisions to increase customer experience and make targeted product offerings based on the customer feedback.

5. RESULTS AND DISCUSSION

In our exploratory data analysis we already derived some initial insights on the top 10 products. We analysed the frequency of reviews per year for each of the top 10 products as well as each product's average rating. In this section we present our findings on the sentiment towards each product as well as the key topics identified per product's reviews.

We first used the NLTK and TextBlob libraries to perform the sentiment analysis and added an additional attribute to our dataset to store the sentiment label (e.g. positive, neutral, negative). The resulting dataset was used to create visualisations of the sentiment distribution for each of the top 10 products. Figure 4 helps visualise the results.

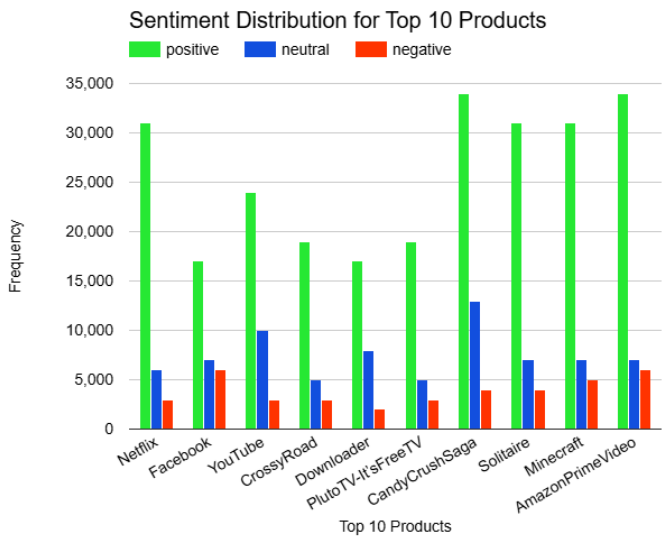

Figure 4: Top 10 Product sentiment distribution

We clearly analyse that all the products have a majority of positive reviews than neutral or negative reviews. This indicates that customers enjoy the features of the products. Despite the customer satisfaction being high, we still need to consider why some customers gave negative reviews for the products such that stakeholders could further improve customer satisfaction. However, the reason behind the sentiment is still unclear and this is why topic modelling is required to help identify the key themes discussed in the reviews. We'll look at the results for topic modelling later but first we analyse the customer sentiment further by analysing the sentiment trends for two products product. From our analysis earlier we visualised the number of reviews for each product over the years, specifically showing the visualisations for the "Pluto TV - It's Free TV" and "Candy Crush Saga" products. Figure 5 helps visualise the sentiment of these products over the years.

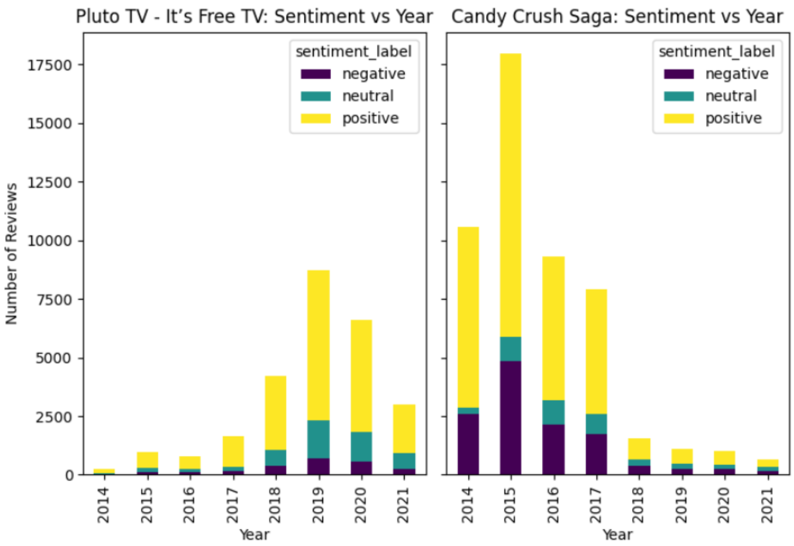

Figure 5: Sentiment Activity per top 10 product (only two products shown)

From the graph we can now analyse that the large spike in activity was mainly due to a large number of positive reviews from customers which shows that the features or updates made in the years 2018-2020 and 2014-2017 for the respective products, satisfied the customers' needs. However, the developers of these products should also be concerned with the significant drop in product reviews. Although, this could indicate that users were satisfied with the current state of the application and did not need to make reviews, it could also indicate that users stopped using the application, creating a negative impact on the company who developed it. Therefore, developers need to focus on implementing features that closely aligns with the users' needs and should listen more closely to their feedback.

Let us analyse the potential reasons behind the sentiment shown towards these products. Due to space limitations we only present the results for two products. If the reader would like to see all the results, please visit our GitHub repository [8]. Figure 6 helps visualise the most common words used per identified topic, specifically for the "Candy Crush Saga" product. From the word clouds we analyse that a lot of positive words such as "fun", "love", or "great" is used which indicates that customer do enjoy playing the game. Furthermore, if we analyse the output shown in figure 7, we analyse that the majority of the topics show positive sentiment with some indicating negative sentiment. Additionally, the topics

Figure 6: Topic Modelling for Product: Candy Crush Saga

derived are the following:

- Topic 0 - *Addiction and Positive Experience* - Users seem to enjoy the gameplay and find it very engaging since they use words such as 'addictive' or 'addicting'
- Topic 1 - *Challenging and Fun Gameplay* - This topic emphasizes the challenging aspects of the game, with words like "challenging," "hard," and "frustrating." However, it also includes positive terms like "fun," "love," and "entertaining," suggesting that users find the challenges enjoyable.
- Topic 2 - *Engaging and Mind-Stimulating* - The topic highlights on the cognitive aspect of the game, with words like "mind," "brain," and "challenges." Users seem to appreciate how the game keeps their minds active and engaged.
- Topic 3 - *Frustration with In-App Purchases and Level Design* - Centers around negative experiences, particularly related to in-app purchases and level design. Words like "money," "pay," and "wait" indicate frustration with the game's monetization strategy.
- Topic 4 - *Positive Sentiments and Gratitude* - Overall positive sentiment towards the game, with words like "love," "awesome," and "happy." Users express gratitude for the game's availability and their enjoyment of playing it

From the results obtained we observe that customers do enjoy playing the Cany Crush Saga application due to it being challenging. However, developers need to ensure that the challenges are not so difficult that it would lead to frustration. The users also do not enjoy the game's monetization strategy so developers should explore alternative monetization strategies or offer more generous rewards to reduce negative sentiment. Analyzing user feedback on level design can help identify areas for improvement, such as reducing excessive difficulty or monotony. The developers could also try and implement better social interaction features such as leaderboards to enhance the overall gaming experience.

Topic #0: great game addicting good time fun play playing way love pass ok stop day addicted enjoy hours loves kill work

Topic #1: fun addictive game challenging play lots like easy really love lot great frustrating recommend playing entertaining highly relaxing time hard

Topic #2: game fun levels keeps love challenging level challenge time play mind like brain really kids think makes great challenges good

Topic #3: game lives dont level like play just wait money time fun levels pay buy playing wish free want facebook friends

Topic #4: love game kindle candy crush awesome play finally glad best app playing phone happy im saga available addicted thank addictive

Figure 7: Topic Modelling Text for Candy Crush Saga application

Despite presenting only the results for the sentiment analysis for a few products and the topic modelling results for the "Candy Crush Saga" application, the other products showed similar results. Developers should focus on maintaining the positive sentiment shown towards their products but should also focus on addressing issues raised by any customer. In doing this, they can greatly enhance the customer satisfaction and get more users to engage with their products and this could potentially lead to an increase in revenue, especially if the customers do not mind spending purchasing in-app purchases.

6. CONCLUSION

In this paper we showed how sentiment analysis can be used to extract valuable insights from customer reviews on software products. Additionally, we showed that using topic modelling can help identify the reasons behind positive, neutral, or negative sentiment. These insights are valuable to stakeholders and developers as it highlights areas where improvements can be made to enhance customer satisfaction. Future work could explore using deep learning techniques for classifying sentiment of software product reviews. Furthermore, our research mainly focused on analysing software product reviews and future work can expand on our research by using a similar approach to ours to analyse reviews for products or services in various other domains.

7. REFERENCES

[1] O. Abiola, A. Abayomi-Alli, O. A. Tale, S. Misra, and O. Abayomi-Alli. Sentiment analysis of covid-19 tweets from selected hashtags in nigeria using vader and text blob analyser. *Journal of Electrical Systems and Information Technology*, 10(1):5, 2023.

[2] N. Ali, S. Hwang, and J.-E. Hong. Your opinions let us know: mining social network sites to evolve software product lines. *KSII Transactions on Internet and Information Systems (TIIS)*, 13(8):4191–4211, 2019.

[3] M. Daradkeh. Understanding the factors affecting the adoption of project portfolio management software through topic modeling of online software reviews. *International Journal of Information Technology Project Management (IJITPM)*, 10(3):91–114, 2019.

[4] N. Darraz, I. Karabila, A. El-Ansari, N. Alami, and M. El Mallahi. Detecting trending products through moving average and sentiment analysis. *Multimedia Tools and Applications*, pages 1–22, 2024.

[5] A. Feizollah, S. Ainin, N. B. Anuar, N. A. B. Abdullah, and M. Hazim. Halal products on twitter: Data extraction and sentiment analysis using stack of deep learning algorithms. *IEEE Access*, 7:83354–83362, 2019.

[6] Y. Hou, J. Li, Z. He, A. Yan, X. Chen, and J. McAuley. Bridging language and items for retrieval and recommendation. *arXiv preprint arXiv:2403.03952*, 2024.

[7] R. S. Jagdale, V. S. Shirsat, and S. N. Deshmukh. Sentiment analysis on product reviews using machine learning techniques. In *Cognitive Informatics and Soft Computing: Proceeding of CISC 2017*, pages 639–647. Springer, 2019.

[8] JsteReubsSoftware. Cos781-project, 2024.

[9] Z. Li, Y. Fan, B. Jiang, T. Lei, and W. Liu. A survey on sentiment analysis and opinion mining for social multimedia. *Multimedia Tools and Applications*, 78:6939–6967, 2019.

[10] A. Mor, C. Orsenigo, M. Soto Gomez, and C. Vercellis. Shaping the causes of product returns: topic modeling on online customer reviews. *Electronic Commerce Research*, pages 1–35, 2024.

[11] N. Raghunathan and K. Saravanakumar. Challenges and issues in sentiment analysis: A comprehensive survey. *IEEE Access*, 11:69626–69642, 2023.

[12] D. Suhartono, K. Purwandari, N. H. Jeremy, S. Philip, P. Arisaputra, and I. H. Parmonangan. Deep neural networks and weighted word embeddings for sentiment analysis of drug product reviews. *Procedia Computer Science*, 216:664–671, 2023.

[13] K. L. Tan, C. P. Lee, and K. M. Lim. A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences*, 13(7):4550, 2023.