sciendo

The Business Value of Text Analysis and Topic Modeling: Evaluating Sentiment Shifts in Mobile Game Reviews after Updates

Bogdan MIREA*
Bucharest University of Economic Studies, Bucharest, Romania
**Corresponding author, mireabogdan17@stud.ase.ro*

Giani-Ionel GRĂDINARU
Bucharest University of Economic Studies, Bucharest, Romania
Institute of National Economy-Romanian Academy, Bucharest, Romania
gianigradinaru@gmail.com

Abstract. *This article explores the potential of web scraping and text analysis as tools to extract business value from user reviews on Google Play. By collecting and analysing app reviews, the study investigates customer sentiment and perceptions before and after the release of specific app features. The proposed methodology leverages natural language processing (NLP) techniques to identify key trends and insights, providing actionable feedback for developers and stakeholders. This approach demonstrates how businesses can use real-time user feedback to assess feature performance, improve user satisfaction, and inform strategic decisions. A particular method used is splitting the data set in two subsets based on a specific date, trying to unveil potential shifts in users' perceptions of the app before and after a major update. Bringing new perspectives to business value, a topic modeling analysis was used on the two subsets, observing changes in the main point of discussion the users had, signalling if changes brought to the app had influenced users and their opinions as a result. LDA topic modeling method was used along with descriptive analysis of textual data.*

Keywords: text analysis, topic modeling, web scrapping, user feedback analysis, natural language processing (NLP).

Introduction

Text mining is an emerging technique aiming to summarize vast amounts of data and extract the main information by recognizing the most frequent used words and correlating topics based on the sentiment of the main corpus. Additions to text mining would include natural language processing, topic extraction and visual analysis methods, all with a focus on the text generating techniques that are of broad and current interest in all industries, used to address various business problems and innovate the way industries function.

Analysing users' feedback can be a challenge in industries that have a numerous user base, so using text analysis and web scrapping methods comes in as a solution to capture the overall perspective and even summarize main topics using topic modeling analysis. As a growing industry, the gaming companies are facing a task that can be addressed with this type of method and ensure developing and creating apps that are in sync with their clients' core ideals. Conducting user-centered research can set a company to develop updates that can positively influence game success as well as retention and revenue. Monitoring feedback to improve user experience is a must for game developers given the growth of users along with the growth of users' feedback even though the large amount of unstructured user reviews can become a challenge.

DOI: 10.2478/picbe-2025-0082

The text mining techniques are a suitable and relevant approach to this problem, as web scrapping can create a usable data source by organizing text from different sources in a way that can be furthermore used in statistical programs where insights can be generated. Using this as a main tool would reduce the need to manually monitor the feedback from users and cut not only costs but also time, offering a time efficient solution and a high-speed framework.

Positive, neutral and negative themes will be analysed based on the score the users gave when they left the review. Shifts in sentiment can be a relevant measure that reveals an important conclusion about whether the business decisions around the update solved the most important problems signaled and improved the user experience. Even though herding behaviour can influence users' opinions, the study has a technical approach and is not evaluating the psychological aspects of sentiment analysis, identifying general feedback and the most expressed problems being the focus of this investigation.

The research questions are around whether the game updates impact overall user sentiment, whether certain types of updates correlate with a negative or positive sentiment change and whether negative feedback is related to a certain problem that can be addressed. In a first step, a descriptive analysis was performed to present the differences between the feedback before and after the update of the game application, considering distinct categories (positive, negative, neutral), then a topic modeling analysis with LDA was implemented, using hyperparameters tuning to capture the best model for our dataset. Python libraries and Jupyter Lab were used for implementing this methodology.

Using a similar framework, app developers can make data-driven decisions using the reviews from their applications, taking into considerations main problems signaled by the end user and shaping business decisions around the client, assuring an increase of satisfaction and a potential revenue addition. This approach can be implemented as an automated and recurrent step in a business strategy, exhibiting the main interests of users, by integrating Python text analysis methods into business intelligence front-end dashboards that usually display various financial KPIs.

Literature review

User feedback can be categorized as a new and emerging procedure of social listening, offering the chance to create a direct connection to the end user, eliminating the need for formal feedback gathering and allowing a more colloquial expression, easing and facilitating the whole process. These types of insights can be important to anticipate any type of business inefficiency and offers a way to correct it (Tian). According to a study by MarketsandMarkets, the NLP market dimension is expected to rise from USD 10.2 billion in 2019 to USD 26.4 billion in 2024, with a Compound Annual Growth Rate (CAGR) of 21.0% during the forecast period (von Aulock). This growth is driven by the increasing use of NLP in data-driven decision-making and the growing demand for enhanced customer experience. In the past years, the rapid expansion and diversification of use of mobile applications has led to an increased need of understanding user feedback to enhance app quality with regards to users' needs. Analysing user opinions, particularly in the gaming industry, can bring to light valuable insights into user experiences and expectations. This literature review investigates the role of text mining and Natural Language Processing (NLP) methods in evaluating user feedback, prioritizing their business value in strengthening and improving mobile gaming applications post-update. NLP covers all the techniques used for analysing natural language, Sentiment Analysis Theory being a subdomain where categorizing unstructured information express in text and identifying patters is the focus. Sentiment analysis refers to determining polarity (neutral, negative or positive) and intensity of opinions (Bing). For game reviews analysis, this theory showcases different techniques like lexicon-based classifiers or supervised machine learning models like BERT. These types of approaches offer a way to use sentiment theory and identify main expressions related to game features (‘crash’, ‘server issues’ or ‘fun play’).

Importance of User Reviews in App Development

User reviews serve as an important feedback mechanism, offering developers direct insights into user experiences, challenges, and desires. Studies have proven a significant relationship between analysing online consumer reviews and financial and strategic performance, bringing into focus the necessity for developers to observe and evaluate this feedback following a carefully planned and systematically implemented methodology. For instance, research indicates that carefully selecting, integrating, and presenting essential components that contribute to the overall depth of analysis from consumer-generated opinions and detailed evaluations into app updates is associated with an observable increase in the average user ratings, indicating greater approval of the latest app enhancements and feature improvements (Shen).

Text Mining Techniques for Analysing User Feedback

Given the unstructured form of textual reviews, text mining is systematically integrated into global frameworks and practices to extract actionable insights. Latent Dirichlet Allocation (LDA) is recognized as a standard approach topic modeling technique that identifies and synthesizes latent themes, providing a deeper understanding of the central theme within extensive datasets containing textual information. For example, an analysis of user reviews for mobile role-playing games (RPGs) utilized LDA to uncover prevalent topics, guiding developers in enhancing game design and user engagement (Youm). Another approach was comparing different topic modeling techniques (NMF, LDA and BERTopic) on game reviews concluded that BERTopic was the best algorithm at identifying topics after the performance evaluation and was also offering a novel way of interpreting the extracted topics (Tong). The platforms used for playing games have a strong influence on users’ interests, changing their feelings and with that the main topics, mobile gaming users have a special interest in character customization options (Sim).

The integration of sentiment analysis enhances topic modeling outcomes by evaluating the emotional state of user feedback. By classifying reviews as negative, positive and neutral, developers can analyse and quantify user sentiment towards defined aspects of feature development and updates. This combined and complementary strategy of topic modeling and sentiment analysis enhances the feasibility of a complete and well-informed perspective of user feedback, supporting specific improvements in the application functionality and end-user satisfaction.

Impact of App Updates on User Sentiment

The dynamic industry of mobile applications requires periodical enhancements, which can drastically drive and modify perceptions of users. Research exploring the relationship between app release patches and user reviews revealed that developers often take into consideration user feedback when organizing and implementing updates. This strategic fit not only ensures user needs and expectations are met but also promotes user satisfaction and favourable perception with respect to newly released versions of the application (Wan).

Moreover, studies have identified various roles that user feedback can play in app enhancements, such as providing developers with system errors reports, propose functional upgrades or usability insights Understanding these roles through text mining allows developers to prioritize updates that resonate with user needs, thereby enhancing user satisfaction and app performance (Liu). A methodical and practical approach is TOUR (TOpic and sentiment analysis of User Reviews), an automated tool designed to dynamically analyse user feedback over multiple app releases. Unlike previous studies that summarized user sentiment towards app features, TOUR provides a straightforward solution for highlighting and synthesizing app, revealing sentiment change, and assigning higher importance to critical user reviews for developers. The tool implements an online topic modeling method and sentiment prediction analysis, allowing developers to tune hyperparameters for tailored results. TOUR improves developer decision-making by reducing noise from high volumes of reviews and portraying insights interactively. In an evaluation survey involving 15 developers, all participants affirmed the practical utility of TOUR's insights, demonstrating its effectiveness in assisting developers with feature improvements and issue resolution. This study highlights the importance of automated text analysis methodologies in enhancing application update strategies and developing and refining them in a manner that prioritizes user needs, ensuring a seamless and satisfying experience (Yang).

Sentiment analysis, made possible through AI algorithms and implemented through Natural Language Processing (NLP) methods, has become a powerful tool for extracting clear and implementable insights from customer opinions, social media content or other digital sources and platforms. By using sentiment analysis models, textual data can be processed and divided into positive, neutral, or negative sentiments, offering businesses a chance to systematically interpret user perceptions and respond in a proper manner to the most revealed problems. Using this approach, frequently revealed themes - whether they focus on highlighting product strengths, exposing recurring issues faced by users, or tracking emerging patterns in consumer behavior - are identified and translated into measurable and impactful marketing strategies that drive brand positioning and audience engagement (Kauffmann). The insights derived from sentiment analysis have been shown to facilitate the optimization of marketing campaigns and the development of products. By identifying the most valued functionalities and addressing operational inefficiencies, sentiment analysis has been demonstrated to contribute to enhanced user experience, improved customer satisfaction, and driven strategic business decisions.

Methodology

In this study, web scraping technique was utilized for data collection to extract user reviews from the Google Play Store, leveraging google_play_scraper package in Python. Regarding the vast number of user-generated reviews available, automated extraction was a necessity to gather and analyse data for sentiment and topic modeling in an efficient mode. The google_play_scraper package offers a structured mode to access app-related data, from reviews, ratings, and metadata, by taking advantage of Google Play's API endpoints rather than parsing HTML manually. The data collection process implies installing the package, extracting reviews for a specified application based on language, country, sorting order, and count, and converting the extracted JSON data into a structured pandas DataFrame for further processing. The ethical views of web scraping were also reviewed, assuring compliance with Google Play's terms of service and excluding personally identifiable information (PII) to maintain users' right to privacy. This method has been found to offer several advantages, including efficiency, scalability, and accuracy, which makes it a reliable approach for extracting user sentiment data. By leveraging this automated technique, a substantial volume of mobile game reviews was obtained, which served as the basis for subsequent text mining analyses. These analyses were aimed at understanding shifts in user sentiment following game updates.

Chosen application for research purpose

Riot Games' mobile adaptation of its popular PC MOBA, League of Legends, known as League of Legends: Wild Rift, has been met with considerable success since its release in 2020. The game has amassed a large player base and surpassed $500 million in revenue by 2022, indicating its viability as a subject for study in user review analysis, amassing 2.3 million reviews left by users from all around the world. Designed for fast-paced 5v5 battles with optimised controls for mobile and console, League of Legends: Wild Rift follows a free-to-play model, generating revenue through microtransactions for cosmetic items like champion skins (Rousseau). Riot Games employs a variety of strategies to maintain player engagement, including the regular implementation of update patches, the introduction of new champions, balance adjustments, gameplay improvements, and seasonal content. These updates are designed to ensure that the game remains competitive and aligned with player expectations. Given its active player community and frequent updates, Wild Rift offers significant opportunities for the study of player feedback trends, sentiment shifts, and the impact of game changes, making it a valuable subject for analysis in the field of user review studies.

As first observable insights from the scrapped data, out of the total number of 287155 reviews left by US resident, English speaking users, the company replied to 36053 (12,6%), usually redirecting the user to directly contact support team, leaving the 87.4% of users without a formal reply. The sentiment column was created using the score given with the review, from 1 to 5, grouping 1,2 scores as negative sentiment, 3 as neutral and 4,5 as positive sentiment.

Data cleaning and text analysis

Text data cleaning constitutes a fundamental preprocessing step in Natural Language Processing (NLP), with the objective of ensuring that raw text is structured and free from inconsistencies prior to analysis. A primary step in the text cleaning process involves the conversion of all text to lowercase, a procedure that assists in maintaining uniformity and preventing the unnecessary creation of distinctions between words that differ only in capitalization. Furthermore, the removal of text enclosed in square brackets serves to eliminate superfluous metadata, such as inline citations or editorial notes, that do not contribute to the core meaning of the text. Links and URLs are also discarded since they provide little value in most linguistic analyses and can introduce noise into the dataset. Another important step is the removal of punctuation, which helps in standardising text and improving tokenization for downstream tasks. Similarly, words containing numbers are often eliminated, as they may not hold significant linguistic meaning unless specifically required for the analysis. Beyond basic cleaning, more advanced techniques are applied to refine textual data. One approach is the handling of contractions, where shortened word forms are expanded to their full versions to preserve meaning, such as converting "don't" to "do not". Stopword removal further improves the quality of text by eliminating commonly used words like "the", "is", or "and" which, while essential in natural speech, contribute little to meaningful analysis in many NLP tasks or worse, they give false results. Another crucial step is lemmatization, where terms are simplified to base form based on their meaning and context, ensuring that variations of a word, such as "helpers" and "helping," are all recognized as their base form, "help". Through these steps, text data is properly refined, limiting noise and redundancy, which significantly strengthens the accuracy of machine learning models and improves insights derived from text analysis techniques such as sentiment analysis and topic modeling.

To ensure a comprehensive analysis of user feedback trends, the study examines Google Play reviews within defined temporal windows surrounding major updates to League of Legends:

Wild Rift. Specifically, three key update patches are considered: Patch 4 (released on 12 January 2023), Patch 5 (released on 18 January 2024), and Patch 6 (scheduled for 9 January 2025). The analysis focuses on a six-month period before and after Patch 5, capturing late-stage issues from Patch 4 and early-stage feedback from Patch 5. The selected review period extends from 18th July 2023 to 17th January 2024, representing the pre-update phase, and from 18th January 2024 to 18th July 2024, summing the post-update phase. The present study employs this approach with the objective of identifying whether the updates introduced in Patch 5 led to significant shifts in sentiment and the dominant themes within user feedback. This will offer insights into the effectiveness of Riot Games' update strategy in addressing player's main concerns.

Comparative analyses were then conducted on the splitter data sets.

Table 1. Scores summary

Reviews	**Before-Update**					**After-Update**				
Sentiment	Negative (0)		Neutral (1)	Positive (2)		Negative (0)		Neutral (1)	Positive (2)	
Score	1	2	3	4	5	1	2	3	4	5
No. reviews	6292	1127	996	781	4644	5547	870	769	548	2525
[%]	45%	8%	7%	6%	34%	54%	8%	7%	5%	25%

Source: Authors' own research.

The two categories have similar scores, with a modest decline in positive sentiment indicated by an increase in negative reviews from 52% to 62% following the implementation of the new game patch. This denotes a shift of 10% from positive to negative sentiments after the update, while the neutral stance persists at 7% of the reviews. Comparatively, the before-update base consists of a total of 13840 reviews, while the after-update sums a total of 10259 reviews for the same six months window.

Figure 1. Rating distributions for before (left) and after (right) app update

Source: Authors' own research.

The distribution of length is also following a similar path but has slight differences, with the most frequent review length between 0-19 characters for 4485 reviews in the before-update data, while the after-update data shows a reduction of shorter reviews, with only 2722 reviews that have lengths between 0-19 characters.

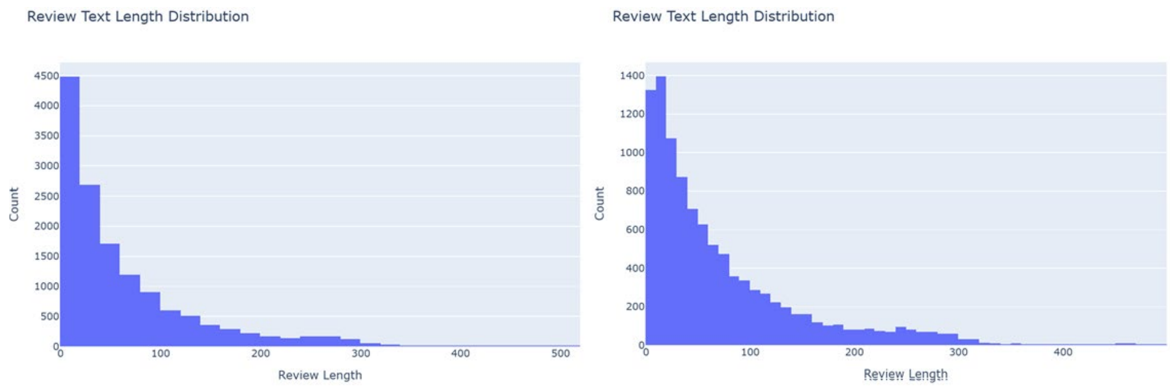

Figure 2. Lengths distributions of reviews before (left) and after (right) app update
Source: Authors' own research.

For both before and after-update sets, the positive reviews have similar frequently used words like 'good moba' or 'great' but some of them raised connection issues after the update ('unable connect server', 'fix log issue') therefor a potential problem happening regarding the server connection after the app was updated. Users have a positive sentiment about the game overall but could migrate to another cluster if the problems they encounter are not addressed.

The main difference in negative reviews is recorded where before-update reviews have words signaling matchmaking problems ('match', 'crash') while the after-update are raising concerns regarding server connections ('unable connect server', 'say fail load') and may also talk about the changes not conforming with their wishes for the game play overall.

Neutral raters have the most potential of leaning to one of the two extremes, therefore special attention should be paid to their requests or complaints. In this category frequent phrases like 'stick verify files', 'would often crash' or 'fix matchmaking' suggest a clear solution that could attract a more positive view of the app. Players would likely shift to a positive sentiment if the matchmaking procedure were reworked, or if the server issues would be addressed.

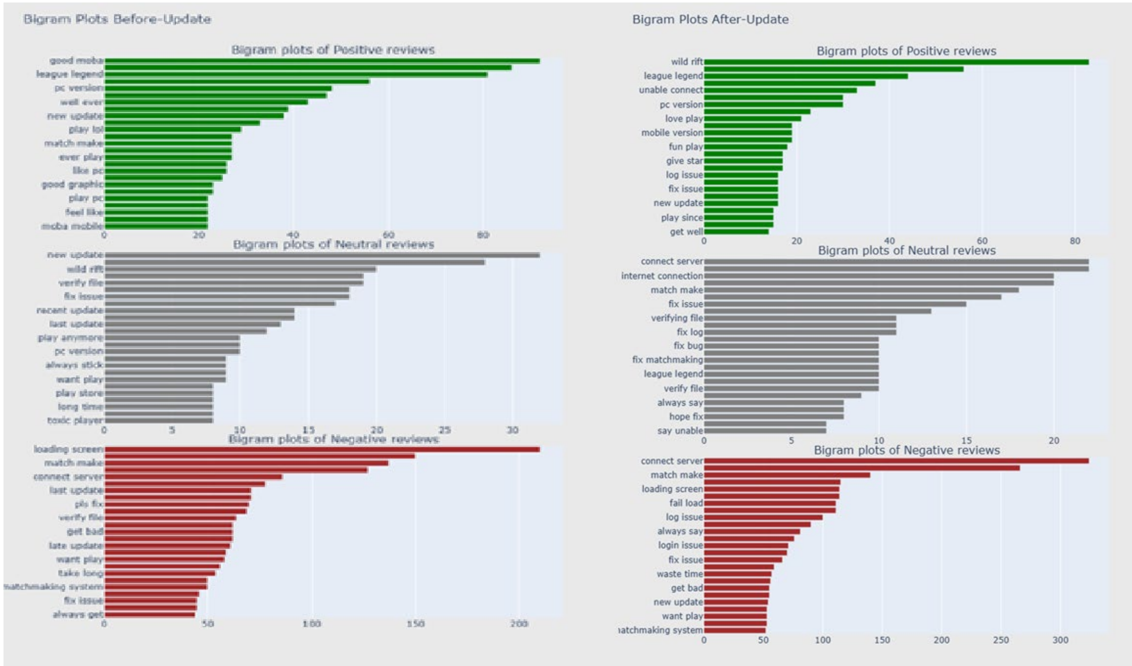

Figure 3. Bigram plots by sentiment of review before (left) and after (right) app update

Source: Authors' own research.

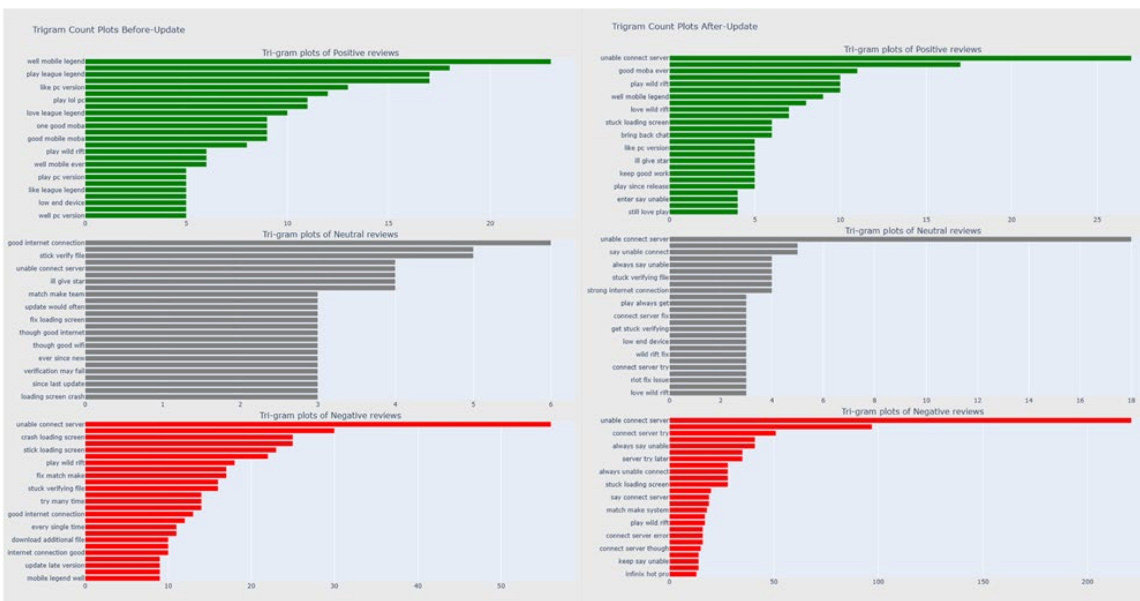

Figure 4. Trigram plots by sentiment of review before (left) and after (right) app update

Source: Authors' own research.

Another way of analysing textual data is through wordclouds, a visual representation of text data where the frequency of words is represented by their dimension. The most common words appear larger, while rarer words are smaller. Word clouds find widespread use in Natural Language Processing (NLP) and text analytics to quickly find trending topics, keywords, or sentiment trends in vast datasets. For our target users that can move to a different cluster easier, the neutral sentiment users, seems to keep similar requests as the most frequent words are 'update', 'play', 'fix' in the before-update category, only ‘update’ being replace with ‘log’ in the after-update category as the largest three words, suggesting connection issues and reinforcing the importance of continuous progress in updating the app and making it more suitable to user’s needs.

Figure 5. Wordcloud graphics of neutral sentiment reviews before (left) and after (right) app update
Source: Authors’ own research.

Topic modeling analysis
As a prerequisite to topic modeling methods, the tokenization process used was based on whitespace tokenization, where text is split into individual words using spaces as delimiters. This method uses the .split() function, which separates each review into a list of words. Following tokenization, a dictionary of unique words is created using the corpora.Dictionary() function, assigning an ID to each unique word in the dataset. The text is then transformed into a Bag-of-Words (BoW) representation, where each document becomes a list of tuples indicating the word ID and the frequency in the text. This numerical representation facilitates text analysis, including topic modeling. However, the process is based on whitespace separation, words with punctuation marks remaining unchanged, a condition we already satisfied by preparing the text at the beginning of our analysis.

Topic modeling is a way of using machine learning to automatically identify topics in a big collection of text. These topics are made up of words that tend to appear together in the dataset. This method is useful for grouping similar documents, summarising large amounts of text and extracting key information from big datasets. By uncovering hidden themes, topic modeling helps make sense of complex textual data without needing to manually categorise it.

Latent Dirichlet Allocation (LDA)
As an unsupervised machine learning algorithm that does not need labeled data or categories, LDA works on probabilistic model to reveal main topics contained by documents. As this algorithm works on the assumption that the documents are a mixture of topics and the topics are a mixture of words, the reviews corpus used is assumed to have different percentages of topics about matchmaking, graphics or bugs, while the topics may contain words like ‘play’, ‘system’, ‘team’ relating to these topics.

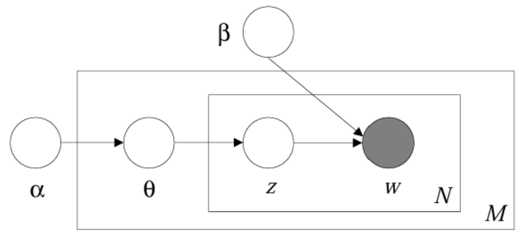

Figure 6. Graphical model representation of LDA

Source: https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf

In this model, the hyperparameters α and β must be set manually. They represent the per-topic distribution (α) and the per-topic word distribution (β), respectively. In document M, the topic is denoted by the letter Z, and the specific word is denoted by the letter W. The presence of specific words (W) in the documents can be attributed to their status as observable variables, with all others being latent. The matrix θ is made up of rows that represent documents and columns that represent topics. The element θ*(i, j)* shows the probability of document *i* containing topic *j*.

To ensure the best model was chosen, hyperparameters tuning was performed on both data sets. Model was run on a number of topics ranging in (2, 5, 7, 10), α in (‘symmetric’,0.3, 0.5, 0.7) and β in (‘auto’, 0.3, 0.5, 0.7), both datasets settling at a similar maximum coherence score.

Table 2. Hyperparameters tuning results

Before-updates model

n	alpha	beta	Coherence score
7	0.3	0.5	0.684

After-updates model

n	alpha	beta	Coherence score
5	0.7	0.7	0.664

Source: Authors’ own research.

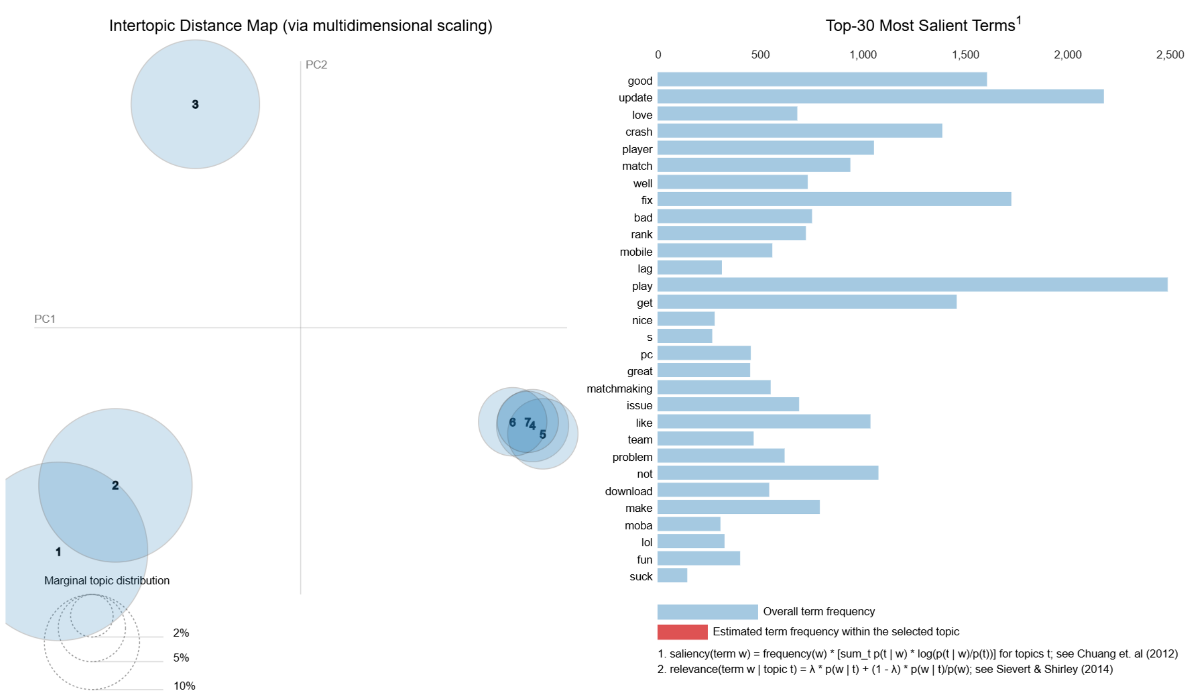

Figure 7. Results of topic modeling for before-updates dataset using pyLDAvis

Source: Authors' own research.

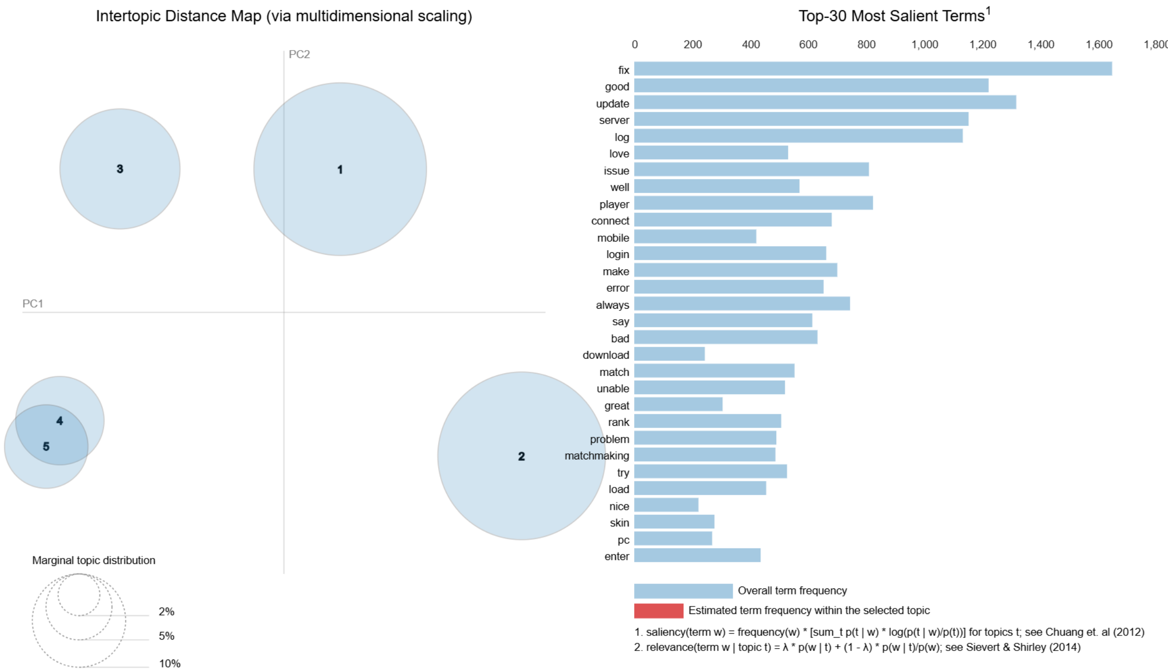

Figure 8. Results of topic modeling for after-updates dataset using pyLDAvis

Source: Authors' own research.

Applying the models resulted in distinct themes within the dataset, each represented by a set of highly associated words. The extracted topics reveal key discussion areas.

Table 3. Grouped terms and derived topics

	Before-updates model			After-updates model	
Topic	Grouped terms	Derived topics	Topic	Grouped terms	Derived topics
1	(update, fix, crash, play, not, issue).	User complaints regarding updates, crashes and unresolved issues	1	(play, matchmaking, system, rank, experience, bad)	Gameplay experience, matchmaking quality and player behavior
2	(player, match, teammate, matchmaking, troll, afk)	Gameplay experience, matchmaking quality and player behavior	2	(fix, update, server, log, issue, connect)	User complaints regarding updates, crashes and unresolved issues
3	(good, play, well, fun, graphic, overall, skin)	Players' enjoyment, gameplay quality, and appreciation of graphics	3	(good, love, graphic, great, new, skin)	Players' enjoyment, gameplay quality, and appreciation of graphics

Source: Authors' own research.

For the first model, containing before-updates reviews, the first and most important topic seems to be about the issues players have, the second topic relates to the gameplay experience while the third one suggests positive feedback and appreciation. The dataset that contains after-updates reviews seems to shift the first two topics importance, making the gameplay experience topic the most important, suggesting a shift in users' focus from the server problems and connection issues to the quality of the gameplay and the satisfaction they get from being part of the players community. This also validates the major game update the company deployed with Patch 5.0, as the topic of complaints is no longer the most important one.

Results and discussions

This study proves that text analysis and topic modeling can discover significant information within many user reviews. We applied Natural Language Processing (NLP) methods to clean the text so that we could examine user emotions, shared topics, and emerging topics up close, grouping users' reviews in three different groups and discovering their main topics of interest.

The method of splitting the database into reviews from two six-month periods before and after a major app update allowed us to research the shifts in users' opinions from different perspectives, revealing a set of different problems reported, with users that had a negative sentiment complaining about matchmaking problems (match', 'crash') before the update and server connections after the update. The neutral category has as main concerns a 'fix of matchmaking', suggesting an improvement on how the game pairing process group players in the game based on their past results or skills and loading issues, and by fixing this type of problems game developers have the potential to shift neutral users toward a positive sentiment.

By using Latent Dirichlet Allocation (LDA) to discover significant topics of the dataset, we identified main topics such as user complaints regarding technical issues or players' experience and behaviour. In regards with the splitting dataset, we identified a switch of the important topics after the update, the topic regarding technical issues becoming less important while the topic about the gameplay experience became more critical. The results of this analysis provide informative recommendations to the business owners and game developers, bringing into focus the need to update the whole game experience while trying to also resolve technical issues rather than only focusing on technical difficulties. It informs them about user concerns to a greater degree and enables them to develop future updates accordingly. This framework can be applied to other domains other than gaming to examine user reviews, businesses that offer a diverse set of services could benefit even more from this type of topic modeling. A core downside of using this method in the gaming industry is that the main topics will not be highly varied, therefore grouping the feedback will result in close related topics and will reduce the value this technique can bring.

Conclusion

Topic modeling and text analysis are great ways of finding hidden patterns, themes and insights in large amounts of text. By using techniques like Latent Dirichlet Allocation (LDA) and text analysis, researchers can easily process and interpret unstructured text data. These methods let you unveil important topics, spot trends and understand sentiment dynamics, which can be valuable in a wide range of industries. But there are still challenges, like how models can be interpreted, adjusting the parameters to obtain a performant model and dealing with noisy data. We can expect future improvements in natural language processing (NLP) and machine learning to make these techniques even more accurate and useful for making decisions based on unstructured data.

Companies can use this type of approach as indirect market research where the clients are not directly asked what features they like or what improvements they would make but rather use clients' feedback in a comparative manner after a change and identifying shifts in clients' perception regarding that change. As stated in this research, all categories of user sentiment have common themes, usually relating to technical problems, but specific problems can be identified based on sentiment or even main topics of discussion can change after major changes and that can bring businesses a valuable idea on next development iterations.

Future work can examine novel advancements such as the usage of sentiment analysis, deep learning techniques or other methods of topic modeling like Non-negative Matrix Factorization (NMF), Latent Semantic Analysis (LSA) or BERTopic (Transformer-Based Topic Modeling). Other perspectives of user feedback can be analysed as well, looking into seasonality of text reviews and the sentiment or comparing users' opinions based on social-demographics variables or even applying the same methods for different languages.

References

Bing, L. (2012). Sentiment analysis and opinion mining. Morgan & Claypool Publishers.

Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3, 993-1022. Retrieved from https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf

Kauffmann, E., & Peral, J. (2018, August 5). Managing marketing decision-making with sentiment analysis: An evaluation of the main product features using text data mining. Sustainability. doi:https://doi.org/10.3390/su11154235

Rousseau, J. (2022, July 22). League of Legends: Wild Rift has amassed $500m in global revenue. Retrieved from https://www.gamesindustry.biz/to-date-league-of-legends-wild-rift-has-amassed-usd500m-in-global-revenue

Shen, Q., & Han, S. (2023, April 23). User review analysis of dating apps based on text mining. PLOS ONE. doi:https://doi.org/10.1371/journal.pone.0283896

Tian, T., & Zichen, H. (2024, May 7). Enhancing organizational performance: Harnessing AI and NLP for user feedback analysis in product development. Journal of Academy of Business and Economics, 145-159. doi:10.18374/JABE-24-1.11

Liu, T., & Wan, C. (2023, September). RoseMatcher: Identifying the impact of user reviews on app updates. *Information and Software Technology*. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0950584923001155

Tong, X., & Willcock, I. (2024). Unraveling player's insights: A comparative analysis of topic modeling techniques on game reviews and video game developers' perspectives. IEEE Transactions on Games, 17(1), 167-180.

von Aulock, I. (2024, April 8). Sentiment analysis: A comprehensive, data-backed guide for 2024. Retrieved from https://penfriend.ai/blog/sentiment-analysis

Wan, C., & Liu, T. (2021, October 17). The role of user reviews in app updates: A preliminary investigation on app release notes. 28th Asia-Pacific Software Engineering Conference (APSEC). IEEE. Retrieved from https://ieeexplore.ieee.org/document/9712100

Yang, T., Gao, C., & Zang, J. (2021, March 26). TOUR: Dynamic topic and sentiment analysis of user reviews for assisting app release. Retrieved from

https://soarsmu.github.io/papers/2021/Dynamic%20Topic%20and%20Sentiment%20Analysis%20of%20User%20Reviews.pdf
Youm, D., & Kim, J. (2022). Text mining approach to improve mobile role playing games using users' reviews. Applied Sciences, 12(12), 6243. doi:10.3390/app12126243
Sim, Y., & Choi, T.-S. (2025). Analyzing cross-platform gaming experiences using topic modeling. Entertainment Computing.

Acknowledgments

This paper was co-financed by The Bucharest University of Economic Studies during the PhD program and was funded by the EU's NextGenerationEU instrument through the National Recovery and Resilience Plan of Romania - Pillar PNRR-III-C9-2022 – I8, managed by the Ministry of Research, Innovation and Digitalization, within the project entitled „CauseFinder: Causality in the Era of Big Data and AI and its applications to innovation management", contract no. 760049/23.05.2023, code CF 268/29.11.2022