E-commerce Product Review Text Summarization Using Natural Language Toolkit

V. B. Mhaske[1], D.S. Jadhav[2]
[1] *PVG's, College of Science and Commerce, Pune 411009*
[2]*Yashavantrao Chavan Institute of Science, Karmaveer Bhaurao Patil, University Satara-415001*

***Abstract*— In recent years, the volume of textual data has grown rapidly as people increasingly rely on written communication to share their ideas and opinions, especially through e-commerce product reviews. Most online shopping platforms enable customers to rate products and submit written feedback. These reviews play a crucial role by offering organizations valuable information about product performance while also helping prospective buyers develop opinions and make well-informed purchasing decisions.**

Conventional product evaluation methods mainly depend on numerical ratings to judge quality. However, this study introduces an alternative approach that focuses on analyzing review texts through text summarization techniques using the Natural Language Toolkit (NLTK). The primary aim of this research is to derive meaningful insights from large collections of customer reviews to support better decision-making. Text summarization can function as an effective decision-support mechanism, assisting consumers in determining whether to purchase a product and enabling companies to gain deeper, more actionable feedback about their products.

***Index Terms*—Comparison, NLTK, Reviews, Similarity score, Text summarization.**

I. INTRODUCTION

In the contemporary digital environment, online shopping has become highly convenient, enabling consumers to purchase a wide variety of products with minimal effort. The integration of advanced technologies such as artificial intelligence and data analytics has transformed e-commerce platforms by facilitating personalized product recommendations based on individual preferences and behavioral patterns. Moreover, secure payment mechanisms and efficient logistics systems have significantly improved the reliability, speed, and overall user experience of online transactions. Through continuous technological advancement, secure financial systems, and the strategic utilization of social media, e-commerce has emerged as a major driver of economic growth and digital development [1].

E-shopping, as an important segment of e-commerce, allows customers to acquire goods and services through online interfaces. Prior research on supply chain strategies in e-shopping has proposed analytical models to determine optimal order quantities and maximize inventory-related profits under varying demand scenarios. These studies reported that supplier- or manufacturer-led supply strategies often outperform approaches that rely exclusively on e-shopping platforms [2]. Furthermore, several models have been developed to minimize operational costs, including expenses related to distribution, inventory holding, backorders, and defective products [3]. Understanding evolving consumer behavior is also essential for the sustained success of e-commerce systems. Consequently, future research has emphasized the need for broader methodological perspectives and the examination of diverse business models across multiple contexts to establish more flexible and comprehensive frameworks [4].

Consumer preferences further influence the structure and effectiveness of online business strategies. Studies indicate that businesses must cater to diverse market segments by offering high-quality product varieties, balancing affordability and quality, and enhancing software design to improve user satisfaction. Rapid adaptation to shifting consumer demands is essential for maintaining competitiveness in the dynamic digital marketplace [5]. Additionally, the effectiveness of different e-commerce models varies across industrial clusters, highlighting the importance of customizing strategies to suit specific economic and technological environments [6].

Despite the advantages of e-commerce, several challenges persist, including fake reviews, irregular review frequency, click fraud, and code-mixing in recommendation systems. These issues not only affect consumer trust but also open important directions for future research [7]. In this context, sentiment analysis has gained prominence as a strategic tool, offering valuable insights that help organizations improve product quality, customer engagement, and analytical accuracy [8].

In the current technological landscape, consumers are exposed to an overwhelming number of choices. Although online platforms provide abundant user-generated reviews, the excessive volume of textual feedback makes it increasingly difficult for individuals to read, analyze, and interpret all available opinions. Additionally, modern consumers tend to spend less time reading lengthy content, further complicating the decision-making process [9]. As a result, there is a growing need for effective mechanisms that can assist users in quickly understanding collective consumer opinions.

Purchasing decisions often require evaluating numerous, diverse viewpoints and identifying features that align with personal priorities. This process is complex, unstructured, and cognitively demanding. Therefore, this research focuses on the role of review summarization in supporting consumer decision-making. The central objective is to generate review summaries tailored to consumer priorities, enabling users to grasp essential product attributes efficiently [10]. Consumers are particularly concerned with specific product features and performance indicators, and summarized reviews provide concise representations of general consumer sentiment [11].

Furthermore, review summarization facilitates the extraction of recurring themes and significant keywords from large volumes of feedback. By identifying commonly occurring terms and expressions such as “durable,” “fast,” or “reliable,” consumers can better understand dominant purchasing factors and satisfaction drivers [12].

This study aims to analyse consumer behaviour through online product reviews by examining the influence of review summaries on purchasing decisions and preference identification [13]. Given the exponential growth of online textual data, there is a critical need for advanced Natural Language Processing (NLP) techniques capable of ensuring accurate, scalable, and automated evaluation of consumer opinions [14]. By identifying what consumers value most, businesses can design and refine products that align more closely with customer expectations, thereby enhancing market relevance and competitive advantage.

II. DATASET

The proposed techniques are implemented on a real-world dataset. For this study, review data for the Redmi 13C 5G smartphone were collected from Amazon. A total of 100 customer reviews were randomly selected and used for analysis.

III. PROCEDURE

In most review analysis systems, sentiment analysis is commonly employed. However, in this study, we focus on generating keyword-oriented review summaries rather than only predicting sentiment. In the first stage, significant keywords are extracted to capture the core aspects of customer feedback. Users are then allowed to select keywords based on their individual priorities or manually enter their own terms. The selected keyword becomes the central element for subsequent analysis.

Next, reviews and sentences containing the chosen keyword are identified and examined with respect to their contextual meaning. This process enables a deeper understanding of consumer concerns and interests related to specific product features. Based on this filtered content, a concise summary is generated to explain why consumers are discussing that particular keyword. The resulting summary helps potential buyers make more informed purchasing decisions according to their personal preferences. Additionally, it provides companies with valuable insights into which product aspects attract the most attention and what improvements may be required to enhance user experience.

Before describing the detailed methodology, the following key terms are introduced.

IV. DATA ANALYSIS TECHNIQUES

A. Figures and Tables

In text mining, preprocessing is crucial to prepare textual data for analysis.

- Stopwords Removal:

Stopwords are common words (such as "the", "is", "and", etc.) that occur frequently in text but carry little meaningful information. Example: Removing stopwords from the sentence "The quick brown fox jumps over the lazy dog" would result in "quick brown fox jumps lazy dog."

- Porter Stemmer:

A stemming algorithm that reduces words to their root or base form by removing suffixes. Helps in normalizing words so that variations of the same word are treated as identical. Example: "Running", "runs", and "ran" all stem to "run" when processed with a Porter Stemmer.

- Removing Punctuation:

Strips punctuation marks (such as “!@#$%^&*():;'",' etc.) from the text. Helps in reducing the complexity of text and focuses on the meaningful content. Example: Removing punctuation from the sentence "Hello, world!" would result in "Hello world".

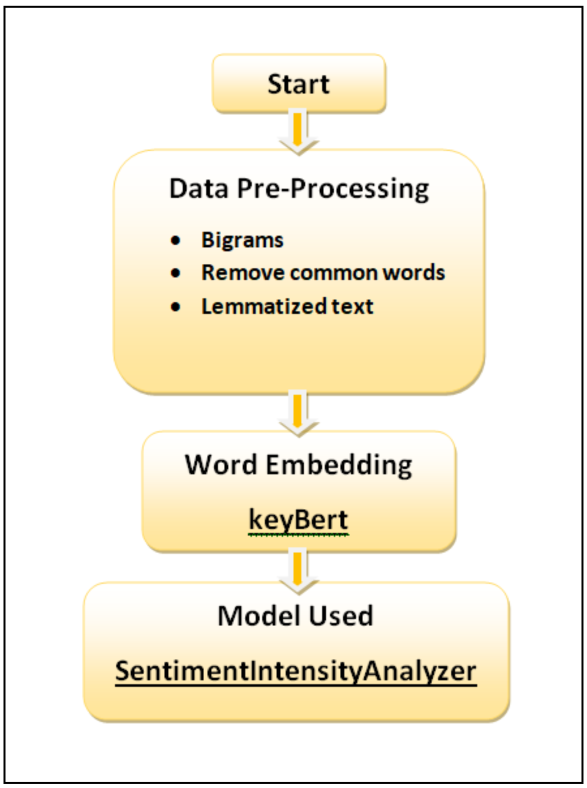

Fig 1: Steps involved in review analysis

B. Word Embeddings

Word embeddings are numerical representations of words in a continuous vector space, where words with similar meanings are represented by vectors that are close to each other in this space. The SemtenceTransformer library is used for transforming sentences or text into fixed-length vector representations, known as embeddings, that capture semantic meaning. These embeddings are useful for various tasks like clustering, similarity search, classification, and information retrieval. There are various pretrained models, in this particularly we use all-MiniLM-L6-v2. It’s great for tasks like finding relevant documents, grouping similar texts, or categorizing information quickly, even on devices with limited power.

C. KeyBERT

KeyBERT is a Python library designed for keyword extraction using BERT (Bidirectional Encoder Representations from Transformers). It identifies the most relevant words or phrases in a piece of text without requiring extensive training or labeled data. Uses BERT embeddings to understand the context, ensuring more accurate and relevant keyword selection.

D. keyphrase_ngram_range

The keyphrase_ngram_range parameter in KeyBERT specifies the size (length) of the phrases (n-grams)

A unigram is a single word (e.g., "machine").

A bigram is a two-word phrase (e.g., "machine learning").

A trigram is a three-word phrase (e.g., "machine learning models").

E. sia.polarity_score

polarity_scores method from the SentimentIntensityAnalyzer (SIA) in the VADER (Valence Aware Dictionary and sEntiment Reasoner) library. It provides a quick and straightforward way to assess sentiment in text without the need for extensive preprocessing or training. VADER is specifically designed for social media, reviews, and informal text. It understands context, for example “very good" vs. "good".

F. Summarizer

Summarizer is used in the Hugging Face Transformers library to create a text summarization pipeline. pipeline is a simple interface provided by Hugging Face to perform various natural language processing (NLP) tasks such as summarization, sentiment analysis, translation, and more. "summarization” Specifies the task you want to perform, in this case, generating a summary of a given text. If you don't specify a model, the pipeline automatically uses a default model fine-tuned for summarization, such as BART or T5.

G. Transformers

It allows a model to give important words according to content. It basically assigns weight while encoding and decoding process. Transformer-based models provide a balance between model size, computational efficiency, and performance, making them versatile options for various NLP applications.

V. HIERARCHICAL FRAMEWORK FOR RESEARCH DESIGN

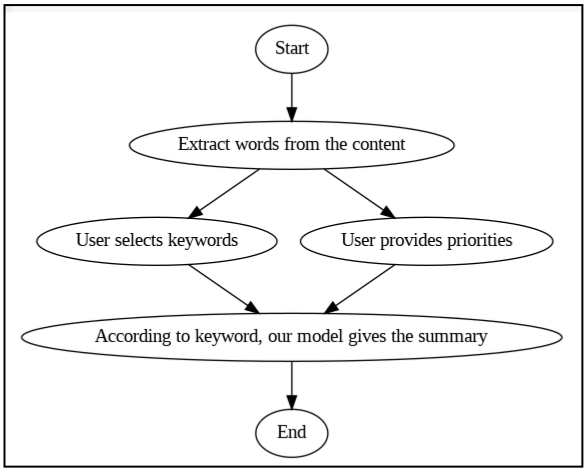

Fig 2: Flow of the process

VI. RESULT

We used reviews available on amazon for the Redmi 13C 5G. Here we are giving some output before final output.

Step 1) Extract Keywords

Extracted Keywords (bigrams, Excluding Numbers, 'redmi', and with Lemmatization):

phone lag: 0.4877
phone lacking: 0.4861
performance phone: 0.4757
quality phone: 0.4605
phone quality: 0.4556
phone budget: 0.4485
new phone: 0.4482
budget phone: 0.4423
recommend phone: 0.4413
poor phone: 0.4407
best phone: 0.4349
average phone: 0.4312
bad phone: 0.4300
problem phone: 0.4191

Essentially, these scores help in identifying the most relevant phrases that describe issues, features, or qualities of a product (in this case, a phone).

Step 2) either user select keyword from above or give keyword

A) User select keyword from above

Suppose user select keyword related "performance phone"

Model Output:

Your max_length is set to 50, but your input length is only 39. Since this is a summarization task, where outputs shorter than the input are typically wanted, you might consider decreasing max_length manually, e.g. summarizer('...', max_length=19)

Positive Comments related to 'performance phone' Summary:

No positive comments found related to performance phone

Negative Comments related to 'performance phone' Summary:

Phone in general is good but the PROBLEM is the phone goes silent of it's own (no ring when someone calls) Overall, I'm not satisfied with its performance.

System gives a summary related to the word "performance phone" with positive and negative sentiment. System makes a summary of it so that the user can easily make decisions according to priority. Sometimes customers make comments negative or positive according to their choice. So it is important to know, particular feature is negative or positive by which mean.

B) User give keyword

Let's consider a user interested in "heating" quality but the machine does not give that keyword. In this case user can give the keyword.

Model Output:

Your max_length is set to 50, but your input_length is only 35. Since this is a summarization task, where outputs shorter than the input are typically wanted, you might consider decreasing max_length manually, e.g. summarizer('...', max_length=17)

Positive Comments related to Heating Summary:

Redmi 13 C 5g phone is value for money, as its battery goes long, performs well as per expectations . No heating issue observed so far, and it could be better if its size is reduced to be pocketed comfortably .

Negative Comments related to Heating Summary:

The phone is Lacking and heating problem. Performance is also slow and battery is very long lasting .To much heating is a main problem with the phone .

If user don't have time to read above comment machine make summary of it from that user easily take decision according to priority.

Your max_length is set to 50, but your input length is only 39. Since this is a summarization task, where outputs shorter than the input are typically wanted, you might consider decreasing max_length manually, e.g. summarizer('...', max_length=19)

Positive Comments related to 'performance phone' Summary:

No positive comments found related to performance phone

Negative Comments related to 'performance phone' Summary:

Phone in general is good but the PROBLEM is the phone goes silent of it's own (no ring when someone calls) Overall, I'm not satisfied with its performance.

VI. CONCLUSION

In many cases, sentiment analysis is used for making decisions about reviews. It gives a number of positive and negative sentiment numbers. However, a sarcastic approach often leads to biased results. The method proposed here is a more reliable alternative. Method simplifies review analysis to help consumers decide what to buy. According to consumer priority we can find out the overview of product with respect to that selected word

Text mining techniques offer valuable insights for both consumers and manufacturers in analyzing consumer reviews. By identifying prevalent issues discussed by consumers, these techniques aid decision-making processes. Consumers benefit from informed purchasing decisions, while manufacturers gain valuable feedback on areas for improvement in their products.

REFERENCES

[1] S. Amin, K. Kansana, and J. Majid, "A review paper on E-commerce," TIMS 2016-International Conference, Gwalior, Feb. 2016.

[2] D. S. Jadhav and V. H. Bajaj, "A study on different supply chain strategy for e-shopping agencies," Int. J. Sci. Res., vol. 4, no. 3, pp. 372–374, 2015.

[7] S. Elzeheiry, W. A. Gab-Allah, N. Mekky, and M. Elmogy, "Sentiment analysis for e-commerce product reviews: Current trends and future directions," 2023.

[8] A. Daza, N. D. G. Rueda, M. S. A. Sánchez, W. F. R. Espíritu, and M. E. C. Quiñones, "Sentiment analysis on e-commerce product reviews using machine learning and deep learning algorithms: A bibliometric analysis and systematic literature review, challenges and future works," Int. J. Inf. Manag. Data Insights, vol. 4, no. 2, p. 100267, 2024.

[9] D. Aggarwal, D. Sharma, and A. B. Saxena, "Enhancing the online shopping experience of consumers through artificial intelligence," 2024.

[10] S. Gensler, F. Völckner, M. Egger, K. Fischbach, and D. Schoder, "Listen to your customers: Insights into brand image using online consumer-generated product reviews," Int. J. Electron. Commerce, vol. 20, no. 1, pp. 112–141, 2015.

[11] A. R. Raheem, P. A. R. M. A. R. Vishnu, and A. M. Ahmed, "Impact of product packaging on consumer's buying behavior," Eur. J. Sci. Res., vol. 122, no. 2, pp. 125–134, 2014.

[12] M. A. Rahman, M. A. Islam, B. H. Esha, N. Sultana, and S. Chakravorty, "Consumer buying behavior towards online shopping: An empirical study on Dhaka city, Bangladesh,"

Cogent Bus. Manag., vol. 5, no. 1, p. 1514940, 2018.

[13] T. Chen, P. Samaranayake, X. Cen, M. Qi, and Y. C. Lan, "The impact of online reviews on consumers' purchasing decisions: Evidence from an eye-tracking study," Frontiers Psychol., vol. 13, p. 865702, 2022.

[14] B. Gayval and V. Mhaske, "Evaluation of descriptive probability approach, cosi pretrained mo," J. Sci. Res., vol. 67, no. 2, 2023.

[15] F. Cuconasu, G. Trappolini, F. Siciliano, S. Filice, C. Campagnano, Y. Maarek, and F. Silvestri, "The power of noise: Redefining retrieval for RAG systems," in Proc. 47th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2024, pp. 719–729.

[16] S. Simon, A. Mailach, J. Dorn, and N. Siegmund, "A methodology for evaluating RAG systems: A case study on configuration dependency validation," arXiv preprint arXiv:2410.08801, 2024.

[17] S. Bag, A. Gupta, R. Kaushik, and C. Jain, "RAG beyond text: Enhancing image retrieval in RAG systems," in Proc. Int. Conf. Elect., Comput. Energy Technol. (ICECET), 2024, pp. 1–6.

[18] M. Arslan, H. Ghanem, S. Munawar, and C. Cruz, "A survey on RAG with LLMs," Procedia Comput. Sci., vol. 246, pp. 3781–3790, 2024.