Analisis Sentimen Ulasan Pengguna Aplikasi Keuangan Di Google Play Store Untuk UMKM (Studi Kasus: Bukuwarung, Pencinta Uang, Kledo)

Sentiment Analysis Of User Reviews Of Financial Application In Google Playstore For MSME (Case Study : Bukuwarung, Moneylover, Kledo)

Rega Primantara[1], Indira Rachmawadi[2]
[1] International ICT Business, School of Economics and Business, Telkom University, Indonesia, regaprimantara@student.telkomuniversity.ac.id
[2] International ICT Business, School of Economics and Business, Telkom University, Indonesia, indirarachmawati@telkomuniversity.ac.id

Abstrak
MSME (Usaha Mikro, Kecil, dan Menengah) sangat penting bagi pertumbuhan Indonesia, namun akses internet terbatas dan infrastruktur yang terbatas menghambat digitalisasi. Integrasi solusi digital menjadi sangat penting. Mengoptimalkan pengalaman pelanggan, memanfaatkan big data, analisis sentimen, dan pemodelan topik muncul sebagai strategi krusial. Memanfaatkan atribut big data memberikan keunggulan kompetitif. Analisis sentimen menghasilkan wawasan dari konten yang dihasilkan pengguna, sementara pemodelan topik mengungkap tema dengan efisien. Konsep-konsep ini membentuk ulang pengambilan keputusan dan hubungan pelanggan, menjadi dasar untuk berkembang di era digital. Penelitian ini menyasar ulasan pengguna pada aplikasi BukuWarung, MoneyLover, dan Kledo, menggunakan analisis sentimen untuk merekomendasikan aplikasi yang paling sesuai untuk MSME. Proses ini memberikan wawasan tentang pengalaman pelanggan, membantu MSME dalam pemilihan aplikasi.

Kata Kunci-UMKM, manajemen keuangan, big data, analisis sentimen, pemodelan topik, solusi digital

---

Abstract
MSMEs are vital for Indonesia's growth, but limited internet access and infrastructure hinder digitization. Integrating digital solutions becomes crucial. Optimizing customer experience, harnessing big data, sentiment analysis, and topic modeling emerge as pivotal strategies. Leveraging big data's attributes offers a competitive edge. Sentiment analysis extracts insights from user-generated content, while topic modeling uncovers themes efficiently. These concepts reshape decision-making and customer relations, becoming cornerstones for thriving in the digital age. Research targets user reviews on BukuWarung, MoneyLover, and Kledo applications, using sentiment analysis to recommend the most fitting app for MSMEs. The process yields insights into customer experiences, aiding MSMEs in application selection.

Keywords-MSMEs, financial management, big data, sentiment analysis, topic modeling, digital solutions

---

I. INTRODUCTION

According to recent data from We Are Social (2023), only 23% of the population in Indonesia remains unconnected to the Internet. This figure is based on the total population of 278.69 million, indicating that approximately 214.59 million people in Indonesia use the Internet.

The primary reasons people use the Internet in Indonesia, as revealed by We Are Social (2023), include searching for information, staying connected with friends and family, staying updated with news and events, watching videos and entertainment, conducting research, and managing finances and savings. Notably, financial management and savings hold a significant share of 33.7% in terms of Internet usage.

Mobile phones are the most commonly used devices to access the Internet in Indonesia, with a substantial percentage of 92.3%. Among mobile phone users, the Android operating system dominates with 92.39%, compared to iOS with 7.39%. This prevalence of Android users makes the Google Play Store a highly competitive marketplace

for application developers. As of Q3 2022, the Google Play Store hosted approximately 3.5 million applications, illustrating the intense competition (CNBC Indonesia, 2022).

Furthermore, the multitude of available applications on the Google Play Store caters to users across various sectors and activities. For example, micro, small, and medium enterprises (MSMEs) utilize these applications to enhance productivity and solve business-related challenges. Indonesia is a significant player in the ASEAN region when it comes to MSMEs. In 2021, it had the highest number of MSMEs, with around 65.46 million units, far surpassing neighboring countries. Indonesian MSMEs have shown substantial contributions, absorbing 97% of the workforce, contributing 60.3% to GDP, and contributing 14.4% to national exports. However, there is room for improvement, as Indonesian MSMEs lag behind in terms of export contribution compared to some neighboring countries like Myanmar.

The Indonesian government is actively promoting the digitization of MSMEs as part of the Go-Digital program, aiming to digitize 30 million MSMEs by 2024 to boost their competitiveness on a global scale.

Despite the potential benefits of digitization and financial management applications for MSMEs, there are challenges. Data from KataData (2021) reveals the top reasons why MSMEs in Indonesia choose to close their businesses, including a lack of production costs, decreased demand, government regulations, difficulties in accessing financial resources, and challenges in obtaining raw materials.

To address some of these challenges, financial management applications on the Google Play Store offer solutions, particularly in improving financial access and aiding in cost management. Three applications, BukuWarung, Money Lover, and Kledo, were chosen for further research based on their user downloads and reviews. Analyzing user reviews of these applications can provide valuable insights into customer needs and product quality.

In light of the significant impact and use of the Internet and its potential benefits for both large businesses and MSMEs, further research is warranted. This research, titled "Sentiment Analysis of User Reviews of Financial Applications for MSMEs," aims to explore how these applications assist MSMEs in managing their financial statements and better understanding customer needs.

A. Research questions:
1. What is the customer (MSMEs) sentiment towards financial statement management service provider (BukuWarung, MoneyLover, and Kledo)?
2. What is the customer (MSMEs) topic modelling towards financial statement management service provider (BukuWarung, MoneyLover, and Kledo)?
3. Which application, between BukuWarung, MoneyLover, and Kledo, is the most preferred by customer?

B. Research purposes:
1. To understand the perception of consumers towards BukuWarung, MoneyLover and Kledo application based on text reviews on Google Play Store using Naïve Bayes Algorithm.
2. To understand the perception of consumers towards BukuWarung, MoneyLover and Kledo application based on text reviews on Google Play Store using Topic Modeling
3. To determine which application, between BukuWarung, MoneyLover, and Kledo has the most positive sentiment shown on Google Play Store reviews.

C. Research benefits:

The study enriches management knowledge by combining vast data understanding. It offers insights for technology-based management, aiding MSMEs and contributing to the writer's education. Practical benefits include input for improving services, systems, and customer experience for users and acting as a reference for technology-based management. Recommendations for financial application choices and benchmarks for MSMEs against e-commerce are also anticipated.

II. LITERATURE REVIEW

A. Customer Experience

Customer Experience encompasses cognitive, emotional, physical, sensorial, and social aspects of user interaction, both directly and indirectly with businesses (Bilgihan et al., 2016). Online customer experience refers to consumers' mental perceptions of their interactions with a company's online value proposition. Four dimensions, outlined by Bilgihan (2016), enhance online customer experience in e-commerce:

1. Ease of use and usefulness: Accessibility and efficiency in accessing application features, affecting user satisfaction.

2. Hedonic and ethical features: Features influencing purchase decisions and user satisfaction, including application interface aesthetics.
3. Enjoyment: The user's happiness and satisfaction while using the application.
4. Social interactions: Company-customer interactions through support panels, live chats, or social media, resolving user needs and complaints.

Zippia (2022) divides customer experience into three components:

1. Discover: Relevant customer contact strategies.
2. Engagement: Customer interaction with products and services.
3. Delivery: Timely and consistent service provision.

Enhancing customer experience yields benefits such as influencing purchasing decisions, repeat purchases, brand engagement, and word-of-mouth propagation (Bilgihan, 2016).

B. Big Data

Big Data's five 'V's'—Volume, Velocity, Variation, Veracity, and Value—define its essence. It encompasses substantial data growth, quick data analysis, diverse data types, data reliability, and strategic insights (Akter et al., 2016). Schroeck et al. (2012) elaborate on big data's characteristics: volume, variety, speed, and truth. Big data's application benefits include improving competitive performance through decision-making.

C. Web Scraping

According to Flores et al. (2020) web scraping is a technique used to obtain data or information from a website automatically and then save it to a system file or database to be used as analysis material. This data can be in the form of text, links, video, audio, or documents. The web scraping process from the internet can be divided into two sequential steps, namely:

1. Acquiring web resources.
2. Extracting the desired information from the data obtained.

D. Text Pre-processing

Arsi et al. (2021) said that text pre-processing is the initial stage in processing data before entering the classification process which aims to transform the data into a more structured and simple one. Text pre-processing is carried out to reduce noise in text which can then help improve classification performance and speed up the classification process. Some of the processes carried out in text pre-processing are as follows:

1. Case folding, this process is carried out to change all the letters in a sentence to uppercase or lowercase.
2. Cleaning, this process is carried out to clean non-alphabetic characters such as numbers, symbols and emoticons from text.
3. Stemming, this process is carried out to get the basic word by removing affixes, both prefixes and suffixes.
4. Stopword removal, this process is carried out to delete words with low information content. Some examples of stop words are the words 'from', 'in', 'to', 'the', 'with', 'and', 'or'.
5. Tokenization, this process is carried out to break sentences into tokens/parts.

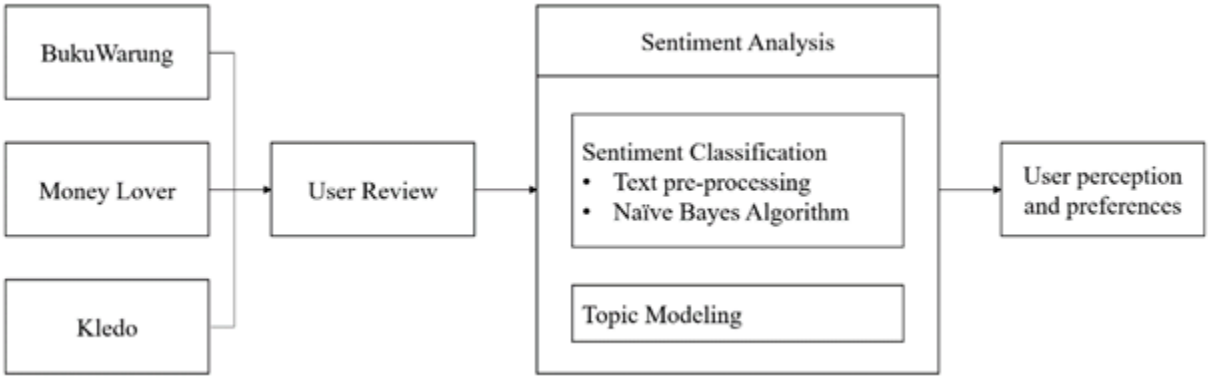
E. Sentiment Analysis

Sentiment analysis gauges authors' sentiment toward various aspects or polarity contexts of text (Pamadja et al., 2013). It extracts subjective information to determine positive, negative, or neutral opinions (Alamsyah et al., 2019). This method aids in studying emotions and opinions in unstructured data, like opinion text. Sentiment analysis employs tokenization, word segmentation, part-of-speech tagging, and other stages to process unstructured data. Naïve Bayes Classifier (NBC) is suitable for Customer Experience analysis in e-commerce (Sari et al., 2018; Al-rubaiee et al., 2018).

F. Topic Modelling

Topic Modelling identifies major topics in customer reviews, aiding the recognition of customer perceptions (Alamsyah et al., 2019). This algorithm groups text or documents based on similarities, revealing latent patterns (Jacobi et al., 2016). Latent Dirichlet Allocation (LDA), a widely-used method, significantly impacts Natural Language Processing and statistical machine learning, making it a valuable tool for Topic Modelling.

G. Theoritical Framework



H. Research Method

A. Google Playstore Data Retrieval

Text mining, as per R. Feldman (2016), extracts knowledge from patterns within text documents through analysis tools. Google Collab software is utilized to scrape user review text data from applications—BukuWarung, MoneyLover, and Kledo.

B. Data Pre-processing

Data is cleaned by preprocessing .csv-formatted data from Google Playstore review scraping. Text pre-processing involves four stages, per Uysal and Gunal (2014):

1. Lowercase Conversion:
   Transforming all letters to lowercase before classification.

2. Tokenization:
   Segregating text into words, phrases, or significant parts.

3. Stopword Removal
   Eliminating unimportant words like conjunctions and prepositions.

4. Stemming
   Returning words to their original form.

C. Sentiment Analysis

Sentiment analysis, according to Zhang (2018), studies opinions, attitudes, and emotions towards entities using the Naïve Bayes method. Hussein (2018) highlights Naïve Bayes' accuracy in text classification, making it a common choice.

D. Analysis

Sentiment analysis results are analyzed based on object positivity levels, recommending the most valuable option for MSMEs.

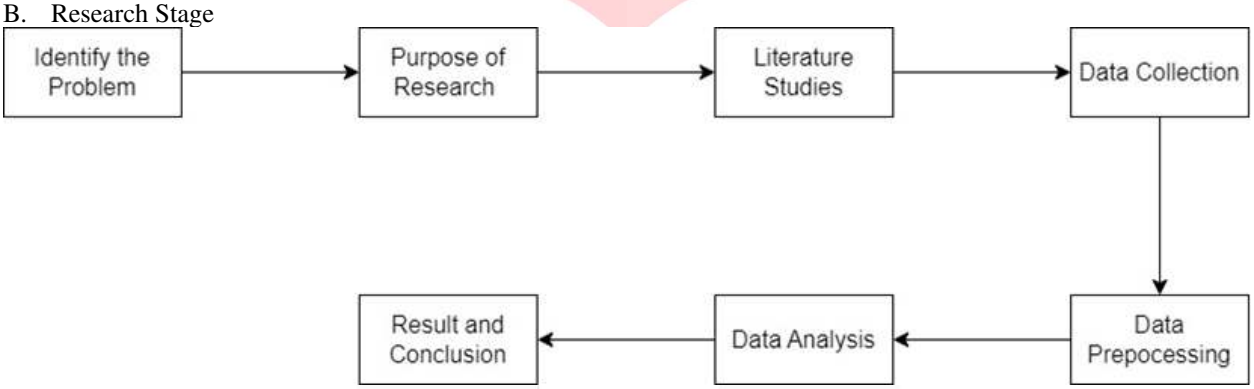
III. RESEACH METHODS

A. Research Characteristic

No	Research Characteristic	Type
1	Based on Research Methodology	Mixed
2	Based on Purpose	Descriptive
3	Based on Units of Analysis	Individual
4	Based on Research Time	Cross Section
5	Based on The Involvement of Researchers	Minimum

According to Sugiyono (2016) research methods can be interpreted as research methods that link between qualitative and quantitative research methods so that the results obtained will be more valid, objective and comprehensive. According to Siyoto and Sodik (2015) descriptive research is related to observing phenomena in more detail and comparing them with other phenomena. According to Sekaran and Bougie (2016) unit analysis refers to the level of aggregation of data collected during the data analysis phase, research analysis is divided into five unit analyses, namely individuals, couples, groups, organizations and culture. According to Indrawati (2015) Cross-Sectional research is a study in which data collection is carried out over a period, which then the data is processed, analysed and produces a conclusion. The involvement of researchers was divided into three levels of intervention, namely minimal, moderate and excessive (Sekaran and Bougie;2016).

B. Research Stage



C. Population and Sample

1. Population

The population used by the author in conducting this research is user reviews on the application BukuWarung, MoneyLover and Kledo contained in the Google Playstore that use Indonesian in their reviews.

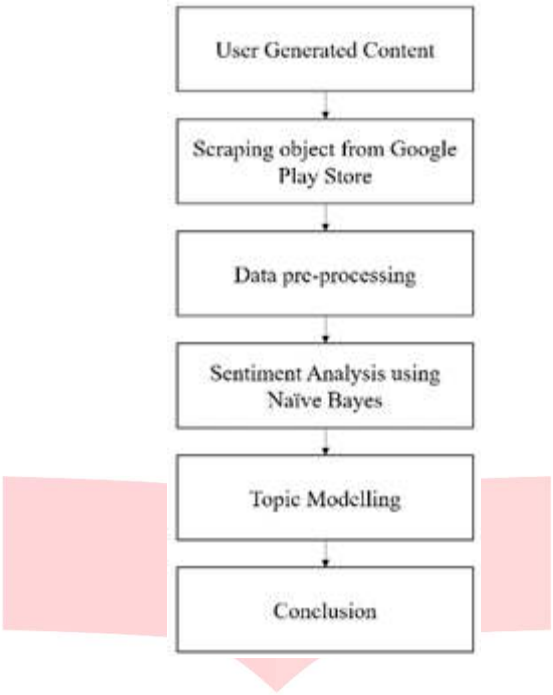
2. Sample

Sample are members of selected members to be considered in the research either to be observed, available a lot, or little about the research being conducted (Indrawati, 2015). Some of the criteria used are reviews that use Indonesian with a research period from the beginning of this application released until May 30, 2023 on the latest application version of BukuWarung, MoneyLover and Kledo.

3. Data Collection and Data Sources

In conducting the study, researchers used secondary data. Secondary Data is data collected by others for use in research (Sugiyono;2016). In this study, researchers used secondary data from User Generated Content (UGC) on Google Playstore. The source of data used by the author is user reviews on the application BukuWarung, Money Lover and Kledo. The result of scraping data will be saved in the format .tsv and will be converted in the format .csv for subsequent processing.

4. Data Analyse Technique



IV. RESULT AND DISCUSSION

A. Data Characteristics**

The study utilizes data from Google PlayStore reviews of BukuWarung, MoneyLover, and Kledo applications, collected through Python-based scraping. A total of 139,141 reviews were obtained.

B. Research Results

1. Data Pre-processing

Collected data undergoes cleaning and preprocessing stages to make it machine-readable. Uysal and Gunal (2014) in his research states that text pre-processing consists of 4 years, namely:

a. lowercase conversion

The process of converting all letter characters to their smallest form before going through the classification process

b. tokenization

Text segmentation process in which the separation of text into words, or phrases

c. Stop words removal

Stop words are words without dependence on a specified topic, for example, conjunctions, articles, prepositions, and so on. Stop words removal is used to remove words that are not so important in the text.

d. Stemming

Stemming is a way to turn words in text back into their original form.

2. Topic Modelling

Topic modelling, identifying latent patterns in text, is used to group texts based on similarity. A Word Cloud visually represents frequent words in reviews, with larger size indicating higher frequency. Google Collaboratory is employed for topic modelling analysis.

C. Discussion of Research Results

1. Discussion of Topic Modelling Result

Word Cloud visualizations illustrate common words in reviews for each app. BukuWarung's prominent words include "Buku Warung," "Sangat Membantu," "Aplikasi," "Sangat Bagus," indicating positive sentiment. MoneyLover's Word Cloud shows "Sangat Membantu" prominently along with "Tidak Bisa," suggesting a mix of positive and negative feedback. Kledo's Word Cloud displays positive terms like "Bagus," "Aplikasi," "Bisa."

Topic Modelling Result

Apps	Topwords	Topic Extracted from Keywords
BukuWarung	Buku Warung, sangat, membantu, usaha, saya, aplikasi, sangat, bagus, mudah, dipahami	Ease and excellence of BukuWarung in helping business users BukuWarung application.
MoneyLover	Sangat, membantu, mengatur, keuangan, fitur, banyak, mantap	The use of the money lover application shows the advantages of Money Lover in terms of features.
Kledo	Aplikasi, bagus, mudah, mantap, lengkap, usaha, pemula	The majority of Kledo users are of the opinion that this application is easy to use and complete for beginner businesses.

2. Discussion of Sentiment Analysis Result

Subject	BukuWarung	MoneyLover	Kledo
Positive	71,947%	66,04%	60,909%
Negative	28,052%	33,96%	39,091%

Sentiment analysis reveals the distribution of positive and negative sentiments for each app. BukuWarung has 71.947% positive and 28.052% negative sentiment. MoneyLover has 66.04% positive and 33.96% negative sentiment. Kledo exhibits 60.909% positive and 39.091% negative sentiment from customers.

V. CONCLUSION AND SUGGESTION

A. Conclusion

Based on the results of research and discussions that have been carried out regarding the Sentiment analysis of Buku Warung, Money Lover and Kledo with data collection period start from the application launch until May 30th 2023 can be taken a conclusion that Financial Application that have first rank in term of Positive sentiment is Buku Warung with 71.947% Positive Sentiment from 126.953 Reviews, In the second place there is Money Lover have 66,04% Positive sentiment from 11.968 reviews and in the last place there is Kledo with 60,909% Positive sentiment from 220 reviews. This show that Buku Warung is more recommend to use especially for MSMEs compare to two other application. That according to topic extracted from keywords BukuWarung make user think that Buku Warung Application is Ease and excellence in helping business users BukuWarung application.

B. Suggestion

1. Theoretical Aspect
   For future research, considering the following suggestions based on the current study's results is recommended:
   a. Utilize a broader time range for review data retrieval to provide a more comprehensive view of customer experiences over different periods.
   b. Include data from platforms like Facebook, Instagram, or news portals, enhancing the preprocessing phase to ensure accurate machine processing and analysis across various aspects.
   c. Conduct a combined analysis with star ratings to compare and correlate results from sentiment analysis and star ratings, offering a more comprehensive understanding of user reviews.

2. Practical Aspect
   In a practical context, the following suggestions are provided based on the research outcomes:
   a. Businesses can employ sentiment analysis and topic modelling to enhance online customer experiences by evaluating user feedback. This analysis can help companies identify their online value propositions, strengths, weaknesses, and uniqueness in terms of customer experiences.
   b. Companies can leverage User Generated Content from Google Play Store user reviews to analyze consumer perceptions and evaluate their performance. Unlike platforms like Twitter and Facebook, Google Play Store reviews offer reliable and valuable insights with minimal spam content.

REFERENCES

AL-Rubaiee, H., Alomar, K., Qiu, R., & Li, D. (2018). Tuning of Customer Relationship Management (CRM) via Customer Experience Management (CEM) using sentiment analysis on aspects level. International Journal of Advanced Computer Science and Applications, 9(5), 300–312. https://doi.org/10.14569/IJACSA.2018.090540

Alamsyah, A., & Bernatapi, E. A. (2019). Evolving Customer Experience Management in Internet Service Provider Company using Text Analytics. Proceeding - 2019 International Conference on ICT for Smart Society: Innovation and Transformation Toward Smart Region, ICISS 2019, November 2019. https://doi.org/10.1109/ICISS48059.2019.8969828

Alamsyah, A., Rizkika, W., Nugroho, D. D. A., Renaldi, F., & Saadah, S. (2018). Dynamic large scale data on Twitter using sentiment analysis and topic modeling case study: Uber. 2018 6th International Conference on Information and Communication Technology, ICoICT 2018, 0(c), 254–258. https://doi.org/10.1109/ICoICT.2018.8528776

Bilgihan, A., Kandampully, J., & Zhang, T. (Christina). (2016). Towards a unified customer experience in online shopping environments: Antecedents and outcomes. International Journal of Quality and Service Sciences, 8(1), 102–119. https://doi.org/10.1108/IJQSS-07-2015-0054

Daulay, E. D. P., & Asror, I. (2020). Sentimen Analisis pada Ulasan Google Play Store Menggunakan Metode Naïve Bayes. E-Proceeding of Engineering , 7(2),8400.https://openlibrary.telkomuniversity.ac.id/pustaka/161809/sentimen-analisis-pada-ulasan-google-play-store-menggunakan-metode-naive-bayes.html

Imbug, N., Nabila, S., Ambad, A., & Bujang, I. (2018). The Influence of Customer Experience on Customer Loyalty in Telecommunication Industry. September. https://doi.org/10.6007/IJARBSS/v8-i3/3909