



# An Evaluation of User Satisfaction with the Ruangguru EdTech App Using a Lexicon-Based Sentiment Analysis Approach

Rahmat Febrian Putra<sup>1\*</sup>, Wahyu Supriyatin<sup>2</sup>

<sup>1,2</sup>Department of Information Systems, Faculty of Computer Science and Information Technology, Gunadarma University, Indonesia

## ABSTRACT

The development of educational technology (EdTech) in Indonesia has established Ruangguru as one of the largest online learning platforms. Users of Ruangguru's EdTech platform can provide feedback regarding their satisfaction and complaints directly via the Google Play Store. Manually evaluating thousands of reviews would be inefficient; therefore, an automated approach is required to understand user perceptions. This study aims to evaluate the level of user satisfaction with the Ruangguru app by classifying user reviews into positive, negative or neutral sentiment groups. The results of this classification can be used to identify aspects of the service that need to be improved in the Ruangguru app. This study employs a Lexicon-Based Sentiment Analysis approach. Data was collected via user review scraping on the Google Play Store. The research stages included data preprocessing (case folding, data cleaning, normalisation, stopword removal and stemming), sentiment scoring and labelling using an Indonesian lexicon, and sentiment distribution analysis. The results of the study, conducted on 1,000 user reviews, showed that approximately 75.2% fell into the positive sentiment category, 22.1% into the neutral sentiment category, and 2.7% into the negative sentiment category. Users were predominantly positive, expressing satisfaction with the Ruangguru learning app. This study demonstrates that the Lexicon approach is effective in classifying the sentiment of user reviews. The evaluation results can serve as a reference for Ruangguru's developers to improve features and enhance service quality in the future. Further research could be conducted using a larger dataset, a longer data collection period, reviews in various languages, and the exploration of other sentiment analysis approaches such as Machine Learning (Naive Bayes, SVM, Random Forest) or Deep Learning (CNN, BERT, RNN).

## ARTICLE INFO

### Article History:

Submitted/Received 15 Jan 2026

First Revised 20 Feb 2026

Accepted 09 Mar 2026

Publication Date 30 May 2026

### Keyword:

Analysis, Lexicon, Ruangguru, Sentiment, Web Scraping.

## 1. INTRODUCTION

Advances in information and communication technology have brought about significant changes in various aspects of life, one of which is the world of education (Juniarti & Sulastika, 2025) (Palandi, Sriyuliawati, & Aziz, 2025) (Ariyansyah, 2025). The integration of information and communication technology into the learning process has given rise to the concept of learning transformation through a new approach that is effective, efficient, engaging, and enriches students' learning experiences (Juniarti & Sulastika, 2025) (Wahyudi & Jatun, 2024). One form of learning transformation is the use of e-learning as a flexible and adaptive learning tool (Ariyansyah, 2025). In the digital age, the use of e-learning has become an effective solution for the learning process and provides broad access to information (Simangunsong, 2024). E-learning, as a form of ICT application in the learning process, offers a range of conveniences and flexibility for learners (Palandi, Sriyuliawati, & Aziz, 2025). One of the e-learning platforms that dominates the Educational Technology (EdTech) sector in Indonesia is Ruangguru.

Ruangguru is a technology-based application designed to help students learn through educational videos, practice exercises and online tutors (Permana, et al., 2025) (Rahmadani & Setiawati, 2019). Ruangguru is regarded as the largest educational technology company in Indonesia (Zuhair, Islamy, Saputri, & Rika, 2025). Ruangguru supports students' independent and integrated learning processes (Permana, et al., 2025). The success of a public service application depends on the level of user satisfaction. Increasing the utilisation of digital applications such as Ruangguru requires evaluation to determine user satisfaction levels and assess users' perceptions of service quality (Rizki, Khabib, Rahmayuna, & Utomo, 2026). On the Google Play Store platform, users can provide reviews describing their experience of using the application. User reviews constitute User-Generated Content (UGC) for evaluating the quality of information systems. Evaluation is required not only from the technical perspective of the application but also to ensure that users perceive the benefits of the service (Rizki, Khabib, Rahmayuna, & Utomo, 2026). Manual evaluation of thousands of user reviews inefficient and prone to problems.

Automated evaluation is required to assess user satisfaction with the Ruangguru app by applying Natural Language Processing (NLP) techniques in the form of sentiment analysis (Rizki, Khabib, Rahmayuna, & Utomo, 2026). Sentiment analysis, or opinion mining, is a branch that examines evaluations, opinions, attitudes, judgements, sentiments and emotions expressed by the public (Ratnaswari, Wibowo, & Kartika, 2025) (Darman, 2023). Sentiment analysis is a process of automatically searching for, understanding, extracting, and processing data to obtain sentiment information contained in public opinion regarding the assessment of trustworthy applications on social media (Darman, 2023) (Amaliah & Nuryana, 2022). The sentiment analysis approach to be used is the Lexicon-Based Approach. The Lexicon-Based Approach detects sentiment by utilising a list of words/phrases (lexicon) where each word/phrase is associated with a specific sentiment (Wikarsa, Angdresey, & Kapantow, 2022). The Lexicon-Based Approach is a process or approach in sentiment analysis used to select important words in a document by using a dictionary or lexicon containing a list of words with associated sentiment values (Anam, et al., 2023) (Putra & Setiawan, 2024).

This study aims to evaluate the level of satisfaction among users of the Ruangguru app by classifying user reviews into positive, negative or neutral sentiment groups. The results of this classification and evaluation can be used to identify aspects of the service that need to be

improved within the Ruangguru app. A lexicon-based sentiment analysis approach was employed by collecting data through user review scraping on the Google Play Store. This was followed by data pre-processing, sentiment scoring and sentiment distribution analysis.

Sentiment analysis research related to user reviews on the Google Play Store was conducted by Syah, Nurdiyansyah and Rahman (Syah, Nurdiyansyah, & Rahman, 2024). The study was conducted to analyse the sentiment of user reviews on the Shopee, Tokopedia, Lazada and Blibli apps on the Google Play Store using a lexicon-based approach and the Random Forest algorithm. The results of the study showed that Lazada performed best with an accuracy of 88.33% compared to Blibli, Shopee and Tokopedia.

Wijaya and Panjaitan conducted a sentiment analysis of Instagram app reviews on the Google Play Store (Wijaya & Panjaitan, 2024). The study was carried out by developing a Naive Bayes sentiment analysis method and combining it with a lexicon-based approach to obtain positive or negative sentiment results. The research was conducted to address the discrepancy between star ratings and the content of user reviews. The research was divided into three datasets. The results showed that differences in accuracy between the three datasets were due to the influence of sample size, data composition diversity, temporal considerations and potential sampling bias.

An evaluation of user perceptions of securities applications was carried out by Thoib, Candra, Sururi, Nugraha and Kholifah. The research was conducted using a lexicon-based sentiment analysis approach and evaluated classification performance using the Random Forest algorithm (Thoib, Candra, Sururi, Nugraha, & Kholifah, 2025). The data analysed comprised 130,905 user reviews from ten popular securities apps on the Google Play Store. The analysis results showed that the Ajaib app received the highest and most representative positive sentiment score, with the largest number of reviews compared to other apps.

The distinction between previous research on evaluating user satisfaction with apps via the Google Play Store and the current study lies in the approach adopted, which focuses exclusively on lexicon-based sentiment analysis. The current study involves identifying user reviews of the Ruangguru app on the Google Play Store, specifically those written in Indonesian. This allows us to determine whether the sentiment identified in relation to the Ruangguru app is positive, negative or neutral..

## 2. METHODS

This study employs a lexicon-based sentiment analysis approach. The lexicon-based method is an approach used to extract sentiment from a text (Darman, 2023). The lexicon works by taking text as input and generating a numerical sentiment value from the words in the sentence (Tanjung, Iqbal, & Sitorus, 2025). The lexicon works by creating a dictionary of opinion words (lexicon), and the words in the dictionary are used to identify the sentiment of a sentence (Darman, 2023). Lexicon-based systems are capable of identifying and assessing sentiment in text by matching words found in the text with a sentiment dictionary (Anam, et al., 2023).

This study will evaluate user reviews of the Ruangguru app on the Google Play Store platform. The user reviews to be analysed are all Indonesian-language reviews, with a dataset

comprising 1,000 reviews from the period 2022 to 2024. The workflow of the lexicon-based sentiment analysis approach in evaluating user satisfaction with the Ruangguru app, as shown in Figure 1, is as follows :

1. Data Collection

Collecting user reviews from the Google Play Store relating to the Ruangguru app. Data collection was carried out using web scraping techniques with the Python programming language and the Google Play Scraper library.

2. Data Preprocessing

A research stage involving the cleaning and preparation of review data (raw text data) to enable better processing and analysis (Haq & Mulyani, 2025). Data preprocessing includes the stages of case folding, text cleaning, normalisation, filtering (stopword removal) and stemming. The preprocessing process is carried out using the Python programming language.

- a. Case Folding Stage

A stage aimed at converting all text from uppercase to lowercase to ensure greater consistency (Tanjung, Iqbal, & Sitorus, 2025).

- b. Text Cleaning Stage

A stage in which unnecessary characters are removed from the review text (Darman, 2023). Special characters such as numbers, hashtags, commas, emoticons, full stops, and excessive spaces are removed from the user review data. Cleaning is performed to reduce random errors (noise) in the data (Haq & Mulyani, 2025).

- c. Normalisation Stage

The stage of converting non-standard words into standard words in accordance with the KBBI (Kamus Besar Bahasa Indonesia) (Haq & Mulyani, 2025).

- d. Filtering Stage (Stopword Removal)

This stage involves removing common words that appear in large numbers but have no significant meaning (stopwords) (Darman, 2023) (Tanjung, Iqbal, & Sitorus, 2025). The removal of stopwords in text mining is used to ensure the focus remains solely on important words (Darman, 2023).

- e. Stemming Stage

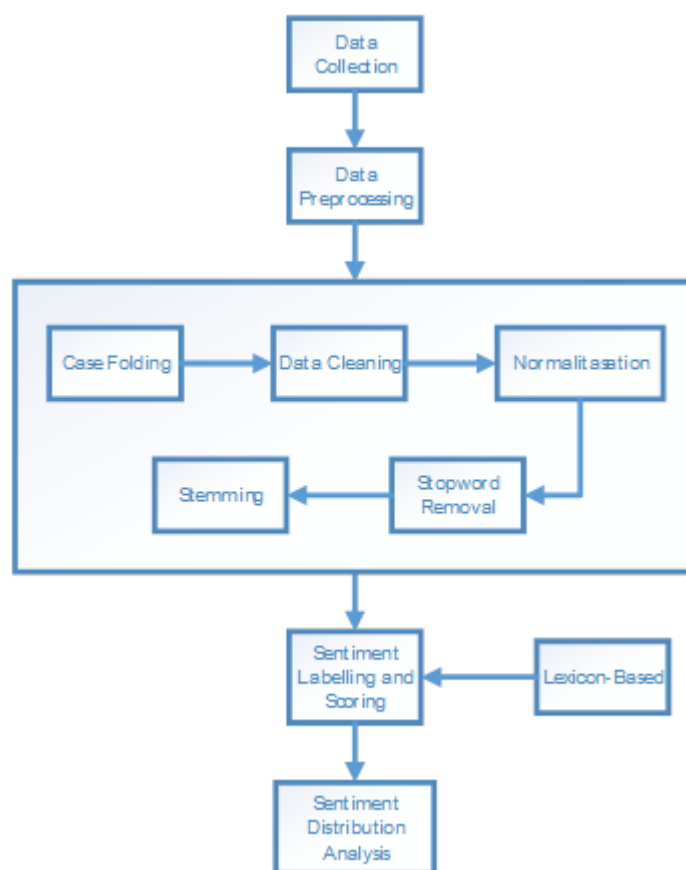
A text pre-processing technique to identify or convert words into their base forms in accordance with the correct structure of the Indonesian language (using the Sastrawi Stemmer) (Tanjung, Iqbal, & Sitorus, 2025) (Haq & Mulyani, 2025). During the stemming stage, all affixes are removed from derived words, including prefixes, infixes, suffixes and combined affixes (Darman, 2023) (Haq & Mulyani, 2025).

3. Sentiment Labelling and Scoring

Performing a lexicon-based process using an Indonesian lexicon from the collected reviews.

4. Sentiment Distribution Analysis

The final stage of sentiment analysis involves presenting the results in the form of visual distributions via graphs and tables. The results of the sentiment distribution analysis are presented as positive, negative, or neutral sentiment classifications. The analysis will identify sentiment trends and the factors influencing user reviews.



**Figure 1.** Research Flowchart

### 3. RESULTS AND DISCUSSION

The evaluation of user satisfaction with the Ruangguru app was conducted using a lexicon-based sentiment analysis approach. The evaluation was carried out using user reviews of the Ruangguru app on the Google Play Store. The Python programming language was used as a tool to analyse user reviews. The results of the analysis, in the form of positive, negative or neutral sentiment, can be used by developers for future improvements. The steps taken in this research are as follows:

#### 1. Data Collection Stage

The initial stage of the research involved collecting user review data for the Ruangguru app via the Google Play Store platform. The review data extracted from the Google Play Store platform consists of review text, star ratings, review timestamps, and the app version used. The technique employed to collect user reviews utilised web scraping with the assistance of the `google-play-scraper` library, which is based on the Python programming language. Web scraping was chosen as it enables the direct extraction of review data from the Google Play Store, including all associated metadata.

The initial step involves setting up the data environment in the Google Colab code editor, with the installation of modules or feature packages for extraction. Figure 2 shows the result of the installation process for the Python libraries: google-play-scraper, pandas, matplotlib, nltk and textblob, indicating that the Google Colab environment has been successfully set up for the analysis.

```

Collecting google-play-scraper
  Downloading google_play_scraper-1.2.7-py3-none-any.whl.metadata (59 kB)
    -----59.2/59.2 kB 1.7 MB/s eta 0:00:00
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (2.2.2)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (3.8.0)
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.9.1)
Requirement already satisfied: textblob in /usr/local/lib/python3.10/dist-packages (0.17.1)
Requirement already satisfied: numpy<=1.22.4 in /usr/local/lib/python3.10/dist-packages (from pandas) (1.26.4)
Requirement already satisfied: python-dateutil<=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz<=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas) (2024.2)
Requirement already satisfied: tzdata<=2022.7 in /usr/local/lib/python3.10/dist-packages (from pandas) (2024.2)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.3.1)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (4.55.3)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.4.7)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (24.2)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (11.0.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (3.2.0)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk) (1.4.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk) (2024.9.11)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.66.6)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil<=2.8.2>pandas) (1.17.0)
Downloading google_play_scraper-1.2.7-py3-none-any.whl (28 kB)
Installing collected packages: google-play-scraper
Successfully installed google-play-scraper-1.2.7
    
```

Figure 2. Results of the Python library installation

The review data to be used consists of 1,000 reviews in Indonesian, covering the period from 2022 to 2024. The review data obtained is sorted by highest rating. The review data is converted into a DataFrame using the pandas library and saved in a CSV file. Figure 3 shows the Ruangguru review DataFrame that has been generated (data scraping).

	user_name	rating	review_text	date
0	Indonesia Maju	5	Aplikasi yang bukan main² isinya. Walaupun han...	2024-12-08 13:23:26
1	Nadini Fairuz shafir	5	Sangat mengedukasi terimakasih telah membantu ...	2024-11-05 09:55:25
2	Regina	5	Kalo secara pribadi, aku kalo belajar bosan ba...	2024-12-01 13:52:36
3	Dyah Aulia	5	Fitur nya keren keren semua,banyak mata pelaja...	2024-10-14 14:36:00
4	Fake Smile	5	Buat developer aku ad saran buat update selanj...	2024-11-15 12:48:20
...	...	...	...	...
995	Pengguna Google	5	Alhamdulillah setelah satu tahun kemarin aku L...	2019-07-18 12:36:04
996	Pengguna Google	5	Terimakasih ruangguru, karena sudah membantu s...	2019-08-01 14:06:51
997	Ira Kusuma Dewi	5	Ruang guru dan skill academy adalah platform t...	2020-09-12 11:57:38
998	Affiah Zahra	5	Aplikasinya keren banget, semua fasilitasnya s...	2020-09-30 05:18:59
999	Sayazia .perikecil	5	Assalamualaikum Halo.. Apk Ruangguru ini bnr' ...	2022-10-19 12:13:29

1000 rows x 4 columns

Figure 3. Ruangguru Google Play Store Review DataFrame

## 2. Data Preprocessing Steps

Once the user review data has been collected, the next step is data preprocessing to clean and prepare the data so that it is ready for analysis. This stage is carried out to simplify the text and focus on words that are meaningful in the context of sentiment analysis. The stages in data preprocessing are:

### a. Case Folding Stage

This stage is carried out to convert all letters in the text to lowercase. Figure 4 shows the result of the case folding process applied to the DataFrame in Figure 3.

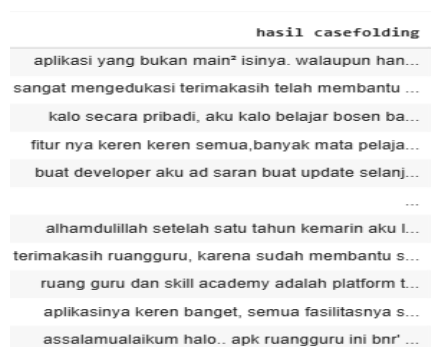
	user_name	rating	review_text	date	hasil casefolding
0	Indonesia Maju	5	Aplikasi yang bukan main² isinya. Walaupun han...	2024-12-08 13:23:26	aplikasi yang bukan main² isinya. walaupun han...
1	Nadini Fairuz shafir	5	Sangat mengedukasi terimakasih telah membantu ...	2024-11-05 09:55:25	sangat mengedukasi terimakasih telah membantu ...
2	Regina	5	Kalo secara pribadi, aku kalo belajar bosan ba...	2024-12-01 13:52:36	kalo secara pribadi, aku kalo belajar bosan ba...
3	Dyah Aulia	5	Fitur nya keren keren semua,banyak mata pelaja...	2024-10-14 14:36:00	fitur nya keren keren semua,banyak mata pelaja...
4	Fake Smile	5	Buat developer aku ad saran buat update selanj...	2024-11-15 12:48:20	buat developer aku ad saran buat update selanj...
...	...	...	...	...	...
995	Pengguna Google	5	Alhamdulillah setelah satu tahun kemarin aku L...	2019-07-18 12:36:04	alhamdulillah setelah satu tahun kemarin aku L...
996	Pengguna Google	5	Terimakasih ruangguru, karena sudah membantu s...	2019-08-01 14:06:51	terimakasih ruangguru, karena sudah membantu s...
997	Ira Kusuma Dewi	5	Ruang guru dan skill academy adalah platform t...	2020-09-12 11:57:38	ruang guru dan skill academy adalah platform t...
998	Affiah Zahra	5	Aplikasinya keren banget, semua fasilitasnya s...	2020-09-30 05:18:59	aplikasinya keren banget, semua fasilitasnya s...
999	Sayazia .perikecil	5	Assalamualaikum Halo.. Apk Ruangguru ini bnr' ...	2022-10-19 12:13:29	assalamualaikum halo.. apk ruangguru ini bnr' ...

1000 rows x 5 columns

Figure 4. Case Folding Results

b. Data Cleaning Stage

This stage involves removing elements that are irrelevant to sentiment analysis, such as punctuation marks, numbers and special characters. Figure 5 shows the case-folded data before cleaning, whilst Figure 6 shows the cleaned data from Figure 5.

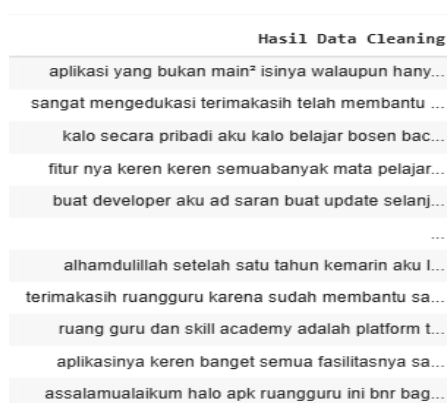


hasil casefolding

aplikasi yang bukan main² isinya. walaupun han...  
sangat mengedukasi terimakasih telah membantu ...  
kalo secara pribadi, aku kalo belajar bosan ba...  
fitur nya keren keren semua,banyak mata pelaja...  
buat developer aku ad saran buat update selanj...  
...  
alhamdulillah setelah satu tahun kemarin aku l...  
terimakasih ruangguru, karena sudah membantu s...  
ruang guru dan skill academy adalah platform t...  
aplikasinya keren banget, semua fasilitasnya s...  
assalamualaikum halo.. apk ruangguru ini bnr' ...

**Figure 5.** Case Folding Data

Figure 5 shows text that still contains several special characters that are irrelevant to sentiment analysis; therefore, data cleaning is required, resulting in the output shown in Figure 6



Hasil Data Cleaning

aplikasi yang bukan main² isinya walaupun hany...  
sangat mengedukasi terimakasih telah membantu ...  
kalo secara pribadi aku kalo belajar bosan bac...  
fitur nya keren keren semuabanyak mata pelajar...  
buat developer aku ad saran buat update selanj...  
...  
alhamdulillah setelah satu tahun kemarin aku l...  
terimakasih ruangguru karena sudah membantu sa...  
ruang guru dan skill academy adalah platform t...  
aplikasinya keren banget semua fasilitasnya sa...  
assalamualaikum halo apk ruangguru ini bnr bag...

**Figure 6.** Data Cleaning Results

c. Normalisation Stage

Following text cleaning, data normalisation is carried out. The normalisation stage focuses on simplifying word forms, so that the resulting words are simpler and more structured. The steps involved in data normalisation are the removal of stopwords and stemming.

d. Filtering Stage (Removal of Stopwords)

The stopword removal stage involves removing words that frequently appear in the text but do not provide information relevant to sentiment analysis, such as “and”, “or”, “which”, “at” and “to”. Stopwords are removed so that the focus remains solely on meaningful words within the text. Stopwords are removed using the nltk library to eliminate irrelevant sentences and affixes that do not alter the sentiment of the review. Figure 7 shows the result of stopword removal from Figure 6.

Data stopwords
aplikasi main <sup>2</sup> isinya aplikasi bermanfaat butuh les ga mengurangi jam sibuk diajari video singkat dirangkum lengkap welldone rg
mengedukasi terimakasih membantu proses belajar belajar nyaman praktis manapun jadwal padat ya min vidio nya fitur teks nya biar malem gk pakai suara hehehe udah bagus
kalo pribadi kalo belajar bosen baca buku nonton sih nanya roboguru kalo mapel lengkap yaaa isi mapel kadang susah nyari pembelajaran lengkap
fitur nya keren keren semuabanyak mata pelajarannya tingkatannyacocok banget fitur ujian tesdulu memahami mata pelajaran ruang guru memahami mata pelajaran memahami ulangi ruang guru the best

**Figure 7.** Results of Stopword Removal

e. Stemming Stage

The stage in which words are converted to their base forms, such as the words “belajarnya” or “belajarkan” being converted to the base form “belajar”. The stemming process is carried out using the Sastrawi Stemmer dictionary library on the text data in the DataFrame. The Sastrawi library is designed specifically for Indonesian language reviews. Stemming ensures that the resulting text is homogeneous and can be used in sentiment analysis. Figure 8 shows the results of stemming the review data from Figure 7. Figure 8 represents the final stage of pre-processing in normalisation, where there are no longer any words or sentences with affixes and everything is in its base form.

Data Hasil Stemming
aplikasi main isi aplikasi manfaat butuh les ga kurang jam sibuk ajar video singkat rangkum lengkap welldone rg
edukasi terimakasih bantu proses ajar ajar nyaman praktis mana jadwal padat ya min vidio nya fitur teks nya biar malem gk pakai suara hehehe udah bagus
kalo pribadi kalo ajar bosen baca buku nonton sih nanya roboguru kalo mapel lengkap yaaa isi mapel kadang susah nyari ajar lengkap
fitur nya keren keren semuabanyak mata ajar tingkatannyacocok banget fitur uji tesdulu paham mata ajar ruang guru paham mata ajar paham ulang ruang guru the best
developer ad saran update gimana kalo live teachingnya sistem rewind gitu kak kalo tinggal kembalikan slide yg kak gak ribet nunggu nunggu riwayat sesi

**Figure 8.** Stemming Results

3. Sentiment Labeling and Scoring Steps

This step focuses on the use of a lexicon-based approach in sentiment analysis. Sentiment analysis is performed using a dictionary of words labeled with sentiment categories. Words in the text are assigned scores for classification. Analysed text are compared against the dictionary, and sentiment scores are calculated based on the number of words found to be either positive or negative. Table 1 is an example of an Indonesian sentiment-labelled dictionary and its sentiment scores.

**Table 1.** Example of a Sentiment-Labelled Dictionary

Word	Sentiment Score
like	1
dislike	-1

fun	1
bad	-1
disappointed	-1
satisfied	1
satisfying	1
negative	-1
positive	1
exciting	1
eror	-1
error	-1
errorr	-1
errorrr	-1
errorrrr	-1
please	-1
thank God	1
help	1
expensive	-1
beneficial	1

If a word is not found in the dictionary, its default value is zero. The sentiment score is calculated from the stemmed text. If the total sentiment score is greater than zero ( $Stotal > 0$ ), the text falls into the positive sentiment category. If the total sentiment score is less than zero ( $Stotal < 0$ ), the text falls into the negative sentiment category. If the total sentiment score is equal to zero ( $Stotal = 0$ ), the text falls into the neutral sentiment category. Figure 9 shows the sentiment scores and sentiment labels from the analysed stemming data text.

	Data Hasil Stemming	sentiment_score	sentiment
0	aplikasi main isi aplikasi manfaat butuh les g...	1	Positive
1	edukasi terimakasih bantu proses ajar ajar nya...	2	Positive
2	kalo pribadi kalo ajar bosan baca buku nonton ...	0	Neutral
3	fitur nya keren keren semuabanyak mata ajar ti...	2	Positive
4	developer ad saran update gimana kalo live tea...	0	Neutral
...	...	...	...
995	alhamdulillah kemarin langgan ruangguru digita...	0	Neutral
996	terimakasih ruangguru bantu gak suka ajar male...	3	Positive
997	ruang guru skill academy platform baik berkual...	1	Positive
998	aplikasi keren banget fasilitas bantu ajar jel...	2	Positive
999	assalamualaikum halo apk ruangguru bnr bagus s...	2	Positive

**Figure 9.** Sentiment Scores and Sentiment Labels

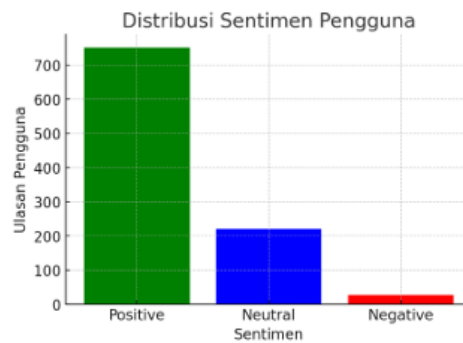
#### 4. Stages of Sentiment Distribution Analysis

The final stage involves analysing the sentiment distribution of user reviews within the dataset to obtain sentiment scores and labels. Of the 1,000 reviews analysed, scores  $> 1$  were classified as positive, scores  $< 1$  as negative, and scores  $= 0$  as neutral. The visualisation of the sentiment distribution analysis of user reviews for the Ruangguru app was carried out using tables and graphs. Table 2 shows the results of the sentiment distribution analysis of a total of 1,000 user reviews from the Google Play Store.

**Table 2.** Sentiment Distribution Analysis Results

Variable	Count
Positive	752
Neutral	221
Negative	27

Table 2 shows that there were 752 positive sentences (75.2%), 221 neutral sentences (22.1%) and 27 negative sentences (2.7%). Figure 10 shows the three sentiment categories classified from the sentiment analysis results in Table 2.



**Figure 10.** Results of Review Data Labelling

#### 4. CONCLUSION

The evaluation of user satisfaction with the Ruangguru app using a lexicon-based sentiment analysis approach can be summarised as yielding three sentiment distributions: positive, neutral and negative. The evaluation of 1,000 user reviews yielded 752 positive sentiment reviews (75.2%), 221 neutral sentiment reviews (22.1%) and 27 negative sentiment reviews (2.7%). The higher positive sentiment score indicates that positive reviews of the Ruangguru app are more dominant among users. The Lexicon-Based approach used in this study is simpler and more efficient, thus not requiring complex model training or large training datasets. However, the Lexicon-Based approach has limitations, such as difficulty in handling words with multiple meanings or complex contexts. Further development related to this study could involve increasing the number of reviews used and combining the Lexicon-Based approach with other methods such as Naive Bayes, Random Forest or SVM. This would yield a more complex distribution of analysis results.

## 5. REFERENCES

- Amaliah, F., & Nuryana, I. K. (2022). Comparison of the Accuracy of Lexicon-Based Methods and the Naive Bayes Classifier in Sentiment Analysis Public Opinion of Investment Apps on Twitter. *JINACS (Journal of Informatics and Computer Science)*, 384-393.
- Anam, M. K., Fitri, T. A., Agustin, Lusiana, Firdaus, M. B., & Nurhuda, A. T. (2023). Sentiment Analysis for Online Learning using The Lexicon-Based Method and The Support Vector Machine Algorithm. *ILKOM Jurnal Ilmiah*, 290-302.
- Ariyansyah. (2025). Using E-Learning as a Strategy for Optimizing Learning at the Elementary School Level. *Jurnal Pemikiran dan Penelitian Pendidikan Dasar*, 249-258.
- Darman, R. (2023). Sentiment Analysis Of Twitter Responses To The Requirements Of The Social Security Administrator (Bpjs) At The Land Office. *Jurnal Widya Bhumi*, 113-136.
- Haq, L. N., & Mulyani, A. (2025). Sentiment Analysis of Jakone Mobile Application on Google Play Store Using Support Vector Machine Method. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 1-9.
- Juniarti, A., & Sulastika, R. (2025). Transforming Learning with Digital Technology. *Journal of Innovation in Teaching and Instructional Media*, 594-708.
- Palandi, E. H., Sriyuliawati, F., & Aziz, A. (2025). The Role of Technology in Developing Interactive and Effective E-Learning Systems for Education. *Journal Scientific of Mandalika (JSM)*, 1987-1997.
- Permana, M. W., Rahmah, N. A., Wisesa, N. S., Efendi, N. S., Khulliyatunnisa, N., & Setiawan, B. (2025). Implementation of the Ruangguru Application in Indonesian Language Lessons on Student Learning Outcomes at MTS Persis 1 Bandung. *LEARNING : Jurnal Inovasi Penelitian Pendidikan dan Pembelajaran*, 106-119.
- Putra, M. B., & Setiawan, E. (2024). Lexicon-Based Method for Sentiment Analysis of Twitter Users on ISP Performance (Case Study: Indihome, Biznet, MyRepublic). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 12079-12087.
- Rahmadani, N. S., & Setiawati, M. (2019). The "Ruang Guru" Online Education Application as a Increasing Millennial Generation Learning Interest in Responding to The Development of The Industrial Revolution 4.0. *Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 241-246.
- Ratnaswari, S., Wibowo, N. C., & Kartika, D. S. (2025). Sentiment Analysis Using Lexicon-Based and Support Vector Machine Methods for the Indonesian President and Vice President for the 2024–2029 Period. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, 362-368.
- Rizki, D. S., Khabib, M. S., Rahmayuna, N., & Utomo, V. G. (2026). Sentiment Classification of Google Play Store Public Service App User Reviews Using NLP and ML. *Jurnal Tekno Kompak*, 51-64.
- Simangunsong, W. N. (2024). Utilizing E-Learning for Flexible Learning and Easy Access to Information Anywhere (Case Study of Informatics Engineering, Asahan University). *Jurnal Penelitian Multidisplin Bangsa*, 492-499.

- Syah, A., Nurdiyansyah, F., & Rahman, A. Y. (2024). Sentiment Analysis Of Shopee, Tokopedia, Lazada, And Blibli Apps Using Lexicon And Random Forest . JITET (Jurnal Informatika dan Teknik Elektro Terapan), 3576-3587.
- Tanjung, M. R., Iqbal, M., & Sitorus, Z. (2025). Sentiment Analysis of Google Reviews on the Quality of Education at STIE Al-Washliyah Sibolga Using the Lexicon Method and the Naive Bayes Algorithm. *Jatilima: Jurnal Multimedia Dan Teknologi Informasi*, 400-412.
- Thoib, I., Candra, B. P., Sururi, N., Nugraha, D. S., & Kholifah, B. (2025). Evaluation of Lexicon-Based Sentiment Labeling Methods on Securities App Reviews on Google Play Store. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 450-457.
- Wahyudi, N. G., & Jatun. (2024). Technology Integration in Education: Challenges and Opportunities for Digital Learning in Elementary Schools. *Indonesian Research Journal on Education*, 444-451.
- Wijaya, N., & Panjaitan, E. S. (2024). Sentiment Analysis of Instagram App Reviews on Google Play Store: Multinomial Naive Bayes and Lexicon-Based Approaches. *Building of Informatics, Technology and Science (BITS)*, 921-929.
- Wikarsa, L., Angdresey, A., & Kapantow, J. D. (2022). Implementation of the Naive Bayes Method and the Lexicon-Based Approach to Classify Netizen Sentiment in Indonesian-Language Tweets. *Jurnal Ilmiah Realtech*, 15-24.
- Zuhair, A. D., Islamy, A. D., Saputri, A. S., & Rika, A. R. (2025). Competitive Strategy of the Ruangguru Application Platform in the Indonesian Edtech Industry. *Jurnal Transformasi Pembelajaran dan Inovasi*, 204-217.