# Anjis Sapto Nugroho

## Comparing RNN and LSTM Algorithms with FastText Embeddings for Sentiment Analysis on Educational Platforms

Quick Submit
Quick Submit
Universitas 17 Agustus 1945 Semarang

### Document Details

Submission ID
trn:oid:::1:3120530707

Submission Date
Dec 21, 2024, 3:59 PM GMT+7

Download Date
Dec 21, 2024, 4:26 PM GMT+7

File Name
Draft_Jurnal_Publikasi_Sinkron_in_English_-_Revisions_Final.docx

File Size
3.2 MB

12 Pages

6,841 Words

40,906 Characters

# 19% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

## Filtered from the Report

- Bibliography

## Match Groups

**50** Not Cited or Quoted 17%
Matches with neither in-text citation nor quotation marks

**19** Missing Quotations 3%
Matches that are still very similar to source material

**0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation

**0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

## Top Sources

17% Internet sources
16% Publications
12% Submitted works (Student Papers)

## Integrity Flags

**0 Integrity Flags for Review**

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

## Match Groups

- **50** Not Cited or Quoted 17%
  Matches with neither in-text citation nor quotation marks
- **19** Missing Quotations 3%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

## Top Sources

17% Internet sources
16% Publications
12% Submitted works (Student Papers)

## Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

| | | |
|---|---|---|
| 1 | Student papers | |
| Universitas Prima Indonesia | | 8% |
| 2 | Student papers | |
| University of Sunderland | | 1% |
| 3 | Internet | |
| jurnal.polgan.ac.id | | 1% |
| 4 | Internet | |
| ebin.pub | | 1% |
| 5 | Internet | |
| www.ojs.uma.ac.id | | 0% |
| 6 | Student papers | |
| University of Wales, Bangor | | 0% |
| 7 | Publication | |
| Zauyik Nana Ruslana, Rudi Setyo Prihatin, Sulistiyowati Sulistiyowati, Kristiawan ... | | 0% |
| 8 | Internet | |
| doctorpenguin.com | | 0% |
| 9 | Internet | |
| www.degruyter.com | | 0% |
| 10 | Internet | |
| www.mdpi.com | | 0% |

11 Internet
journal.itts.ac.id 0%

12 Student papers
Malta College of Arts,Science and Technology 0%

13 Internet
ejournal.uksw.edu 0%

14 Internet
emrbi.org 0%

15 Internet
journal.rescollacomm.com 0%

16 Internet
jutif.if.unsoed.ac.id 0%

17 Student papers
Sim University 0%

18 Student papers
University of Hertfordshire 0%

19 Internet
scie-journal.com 0%

20 Internet
www.springerprofessional.de 0%

21 Publication
Chrisdion Andrew Ramaputra, Mohammad Hamim Zajuli Al Faroby, Berlian Rahm... 0%

22 Publication
Pethuru Raj, N. Gayathri, G. Jaspher Willsie Kathrine. "Artificial Intelligence for Pr... 0%

23 Internet
admin.calitatea.ro 0%

24 Publication
Imam Fathur Rahman, Anisa Nur Hasanah, Nono Heryana. "ANALISIS SENTIMEN ... 0%

25 Internet
aclanthology.org 0%

26 Internet
github.com 0%

27 Internet
pdffox.com 0%

28 Internet
www.polgan.ac.id 0%

29 Publication
Sayyid Muh. Raziq Olajuwon, Kusrini Kusrini, Kusnawi Kusnawi. "Analyzing Public ... 0%

30 Internet
elibrary.mec.edu.om 0%

31 Internet
www.researchgate.net 0%

32 Publication
"Information, Communication and Computing Technology", Springer Science and... 0%

33 Publication
Elly Indrayuni, Achmad Nurhadi. "Sentiment Analysis About COVID-19 Booster Va... 0%

34 Publication
Fatima De Gloria Da Conceicao, Isho muddin, Dewi Nurwantari. "Effective and Pro... 0%

35 Publication
H.L. Gururaj, Francesco Flammini, S. Srividhya, M.L. Chayadevi, Sheba Selvam. "Co... 0%

36 Publication
I Gede Totok Suryawan, Made Sudarma, I Ketut Gede Darma Putra, Anak Agung K... 0%

37 Publication
Nilaa Raghunathan, Saravanakumar Kandasamy. "Challenges and Issues in Senti... 0%

38 Publication
Noureen, Sharin Hazlin Huspi Huspi, Zafar Ali. "Sentiment Analysis on Roman Urd... 0%

39 Publication

Syeda Nida Hassan, Mudassir Khalil, Humayun Salahuddin, Rizwan Ali Naqvi, Dae... 0%

40 Student papers

Universitas Dian Nuswantoro 0%

41 Publication

V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challeng... 0%

42 Internet

agris.fao.org 0%

43 Internet

assets-eu.researchsquare.com 0%

44 Internet

documents.mx 0%

45 Internet

journal.ump.edu.my 0%

46 Internet

jurnal.uns.ac.id 0%

47 Internet

www.iieta.org 0%

48 Publication

Joice. C Sheeba, M. Selvi. "Pedagogical Revelations and Emerging Trends", CRC Pr... 0%

49 Publication

Kousik Barik, Sanjay Misra, Luis Fernandez-Sanz. "Adversarial attack detection fra... 0%

50 Publication

Wiga Maulana Baihaqi, Arif Munandar. "Sentiment Analysis of Student Comment ... 0%

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

# Comparing RNN and LSTM Algorithms with FastText Embeddings for Sentiment Analysis on Educational Platforms

Anjis Sapto Nugroho[1)*], Kristiawan Nugroho[2)]
[1,2,3,4)] Department of Information Technology, Faculty of Information and Industrial Technology, Universitas Stikubank, Semarang, Indonesia
[1)] anjissapto0003@mhs.unisbank.ac.id, [2)] kristiawan@edu.unisbank.ac.id

**Submitted** :xxxxx | **Accepted** : xxxx | **Published** : xxxxx

**Abstract:** As of 2024, the Merdeka Mengajar Platform has been used by more than 3.5 million teachers across Indonesia. This number represents an increase of more than 3.85% compared to the previous academic year, which was 3.37 million. However, the utilization of this application has not yet reached the expected target number of users, so an analysis is needed to identify the factors causing this. This research uses Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) to perform sentiment analysis on reviews of the Merdeka Mengajar platform. RNN and LSTM are chosen for their advantages in handling sequential data, particularly in text processing for sentiment analysis. This research aims to address the challenges in understanding the positive or negative sentiments of users on the platform. The research methodology includes important stages such as data cleaning, preprocessing, and transforming text into numerical vectors using FastText embedding. Next, RNN and LSTM models are applied to predict sentiment based on patterns in the text data. The research results show that the LSTM model is capable of capturing long-term relationships in sequential data with an expected accuracy of 93.58%. Meanwhile, the RNN model yields a lower accuracy of 91.70%. The LSTM model is more effective in classifying sentiment with high accuracy, especially in text data with complex temporal contexts. This research contributes to understanding user perceptions and feedback regarding the Merdeka Mengajar platform, which is expected to provide insights for platform developers to enhance service quality.

**Keywords:** Sentiment analysis, Merdeka Mengajar, RNN, LSTM, Accuracy.

## INTRODUCTION

Advances in information and communication technologies have had a significant impact on various sectors, including education (Maru et al., 2021). In Indonesia, the Ministry of Education, Culture, Research, and Technology (Kemendikbudristek) launched the Merdeka Mengajar (PMM) platform designed to improve the quality of teaching provided by teachers in Indonesia as part of efforts to support the implementation of the Merdeka program (Ndari et al., 2023). Since its inception in 2022 to 2024, this platform has been used by more than 3.5 million teachers in Indonesia, an increase of 3.85% from the previous year (Iqmatul Pratiwi, Desi Rahmawati, 2024). However, the number is still far from the expected goal of all teachers in Indonesia. The use of this application has not been optimal due to various challenges, such as lack of access to remote areas and technical difficulties faced by teachers (Ketaren et al., 2022). This shows a gap between the platform's goals and the users' needs. Ultimately, this situation raises questions about how well this platform meets the users' needs, which can be revealed through sentiment analysis of user reviews.

Given this, sentiment analysis of user reviews becomes a crucial tool for assessing user experiences in order to comprehend user opinions and remarks regarding technological services, like the Merdeka Mengajar platform (Gomez-Adorno et al., 2024), and to provide data-driven suggestions for enhancement. This analysis can categorize user comments as either positive or negative (Munna & Zuliarso, 2024) using natural language processing (NLP) techniques (David & Renjith, 2021). This helps developers enhance the quality of service and gives insight into the factors influencing teachers' adoption of the platform. Complex sequential data processing is a challenge for traditional machine learning techniques like Naive Bayes and Support Vector Machine (Merinda

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Lestandy et al., 2021). These constraints are addressed by the capacity of recurrent neural networks (RNNs) (Kurniasari & Setyanto, 2020) and long short-term memory (LSTM) as deep learning algorithms (Khan et al., 2022) to record temporal patterns and associations between words. In a variety of sentiment analysis tasks, LSTM has proven to be more effective than RNNs due to its sophisticated memory mechanism, particularly when dealing with complex sequential data and massive amounts of data (Atikah et al., 2022).

Since most research focuses on international platforms or commercial applications, there is still little literature on the use of RNN and LSTM algorithms in the sentiment analysis of educational applications in Indonesia, despite their widespread use (Zuraiyah et al., 2023). In order to close this gap, this study uses RNN and LSTM algorithms based on FastText embeddings, which are able to capture word semantic associations, especially in the Indonesian language (Al-jumaili, 2024). Through the use of FastText embeddings, which are rarely investigated in the context of education, this study seeks to increase the precision of sentiment prediction in reviews, especially for Merdeka Mengajar Platform users.

The research problem addressed in this study is how to improve the accuracy of sentiment analysis (Liu et al., 2021) in PMM user reviews to achieve better results and how these analysis results can be used to understand user perceptions and improve the quality of platform services (Maulana et al., 2020). To answer this question, the research takes a comprehensive approach that includes data cleaning, preprocessing, and the use of FastText input to represent text as digital vectors (Shumaly et al., 2021). RNN and LSTM models were then applied to predict sentiment based on patterns in text mining data (Tan et al., 2022). The primary target audience for this research includes academics interested in developing NLP and deep learning methods, as well as practitioners in the field of educational technology, especially educational application developers. For academics, this research contributes to understanding the advantages and limitations of RNN and LSTM algorithms and includes the development of NLP and deep learning sciences. For platform developers, the results of this research can be used as a guide to improve the quality of service and user experience in educational applications in Indonesia.

## LITERATURE REVIEW

A prior study on sentiment analysis of COVID-19 tweets using the RNN method was carried out by (Nemes & Kiss, 2021), and it focused on reviews of the virus on Twitter. In addition to describing the usage of an embedding layer to represent words as numerical vectors that may be processed by deep learning models, this study describes preprocessing procedures such as punctuation removal, case normalization, tokenization, and stemming. RNN classification performance is assessed in this study, with a 90% accuracy rate. The findings of the investigation demonstrate that this method can precisely spot sentiment trends and offer profound understanding of how society reacts to the pandemic via digital channels. RNN classification performance is assessed in this study, with a 90% accuracy rate. The findings of the investigation demonstrate that this method can precisely spot sentiment trends and offer profound understanding of how society reacts to the pandemic via digital channels.

The efficacy of the Long Short-Term Memory (LSTM) algorithm and Naive Bayes in sentiment analysis was examined in the study by (Isnain et al., 2022), specifically in light of public responses to the "New Normal" policy during the COVID-19 pandemic. With 1,590 postings from Twitter, this study used a balanced sample that focused on both positive and negative sentiments. With an accuracy, precision, and recall of 83.33% as opposed to 82% for Naive Bayes, the results demonstrate that LSTM performed better than Naive Bayes and was also faster. The significance of sentiment analysis for comprehending public opinion and directing policy decisions is emphasized by this study.

In order to solve the problem of imbalanced data in sentiment classification, the following study uses the Recurrent Neural Network (RNN) method to conduct a sentiment analysis of user evaluations of the Shopee Indonesia application (Utami, 2022). The Synthetic Minority Oversampling Technique (SMOTE) and Tomek Links were used in conjunction for data preparation in this study, which improved performance measures such as accuracy (80%) and F1 score (88.1%). These results highlight how crucial it is to control unbalanced data in order to enhance the caliber of sentiment categorization outcomes.

According to a study that investigates the application of deep learning techniques, Word2Vec and Long Short-Term Memory (LSTM) are being used in an effort to analyze the sentiment of movie reviews. It talks about the difficulties in deciphering lengthy text documents and shows how well LSTM handles these kinds of sequences. Using a dataset of 25,000 movie reviews, this study tests several word vector dimensions using the Skip-Gram approach. The best accuracy of 88.17% is obtained with a word vector size of 100, while the CBOW method achieves an accuracy of 87.68% at a dimension of 200. The significance of word vector dimensions and the possibility of additional parameter adjustment to increase accuracy are highlighted in this work (Widayat, 2021).

Another study by (Subowo et al., 2022) improves sentiment analysis of comments on online shopping applications that accept installment payments by using Bi-LSTM and the Corpus Text approach for autolabeling. The Bi-LSTM model for sentiment classification and FastText embedding for word representation are combined

*name of corresponding author

SinkrOn JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

in this study to analyze platform data. The study's results show that the Bi-LSTM model approach with an autolabeled corpus text may achieve an accuracy of 81% when compared to traditional methods. This provides more accurate data on how people feel about the online installment shopping application.

The aforementioned research indicates that sentiment analysis with RNN and LSTM algorithms can produce findings with a high degree of accuracy and enhance the end product. The bulk of earlier research, however, relied on Word2Vec embeddings (Cahyani & Patasik, 2021) or TF-IDF (Abubakar & Umar, 2022), which have limitations in capturing semantic relationships between words (Dang, N. C., Moreno-García, M. N., & De la Prieta, 2022). As a result, there are still a number of gaps that need to be filled. Furthermore, the majority of earlier research has concentrated on international platforms or the social media and commercial sectors (Nadia Ristya Dewi et al., 2022), which means it has little bearing on the Indonesian educational context. Through the analysis of user sentiment ratings on the Merdeka Mengajar Platform (PMM), this study attempts to close the gap by evaluating the performance of RNN and LSTM algorithms based on FastText embeddings (Ariyus & Manongga, 2024). FastText embeddings were selected over TF-IDF or Word2Vec because of their superior capacity to capture the morphology of the language by handling out-of-vocabulary (OOV) terms, which are frequently seen in informal Indonesian writing (Darussalam & Arief, 2021).

One of the things that sets this study apart from a number of previous studies that have been described is the data employed. By employing K-Fold Cross Validation techniques (C. Kalaiarasan, 2024) and the Confusion Matrix (Hasnain et al., 2020), this study seeks to increase the model's accuracy (Liu et al., 2021) and stability (Azizah et al., 2022) in predicting both positive and negative sentiments, which are frequently disregarded in earlier research (Sharma et al., 2021). These two features set this research apart from earlier investigations.

## METHOD

As illustrated in Figure 1, this study employs the Sampling, Explore, Modify, Model, and Assess (SEMMA) process. Predictive models for big datasets have been successfully developed using the SEMMA method (Suwitono & Kaunang, 2022). The SEMMA approach is simple to comprehend and can be used as a guide for data mining projects. A sample is a procedure used to gather important information and data. The stage of exploration is where data sets pertaining to the concepts that will be developed are sought after. The process of grouping variables and altering data is called "modify." The technique of modeling data to forecast the intended result is called model. Access is the review of data for analysis. Based on the model's performance, the classification model developed in the earlier step will be evaluated.

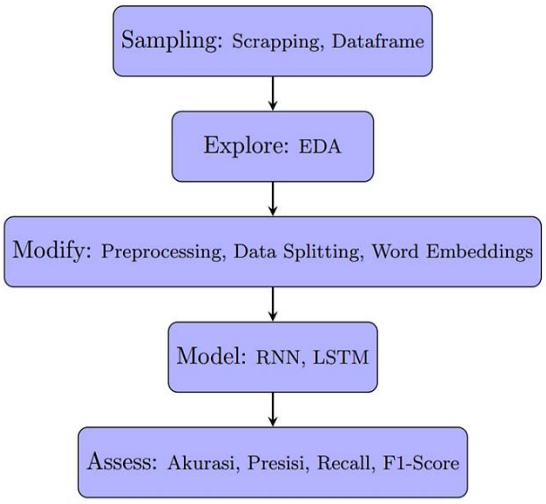


Fig 1. SEMMA Method

Examining the model's output in terms of recall, accuracy, and precision allows for evaluation. The accuracy value shows how accurate the model is, with accurate predictions that match the available data. The loss value, on the other hand, shows how many errors there are in each cycle. The quality of the model produced increases with decreasing loss. The percentage of data that is projected to be positive is known as precision. The percentage of true positive data is called recall. With 0 being the poorest value and 1 representing the greatest, the F1 score is the average of precision and recall (Merinda Lestandy et al., 2021).

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

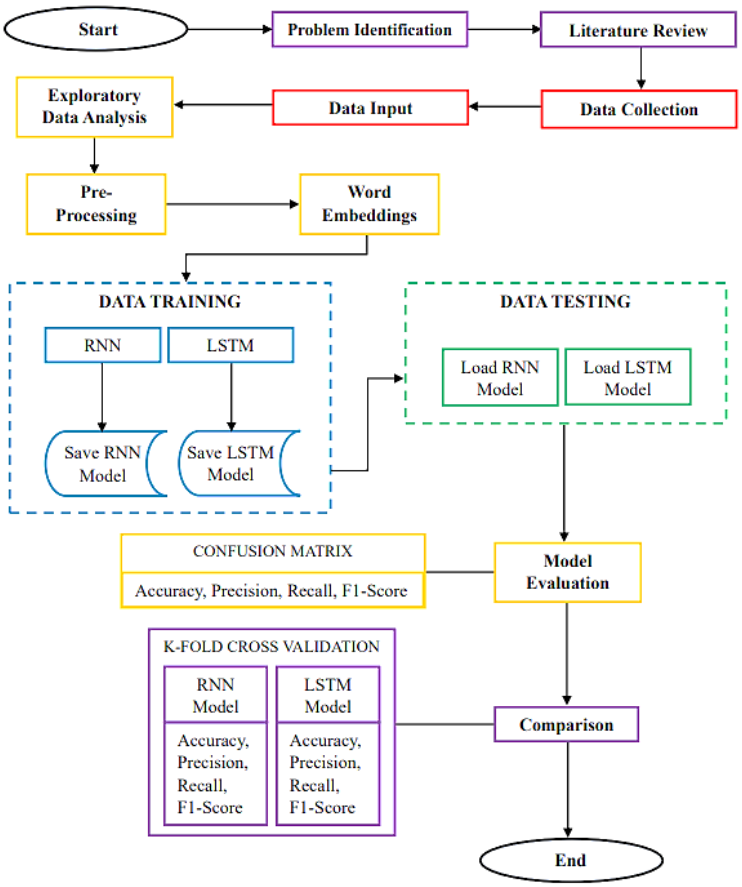


Fig 2. Research Workflow Diagram

The first step in this research is data collection, which involves obtaining review data from the Play Store. Next comes exploratory data analysis, which looks at the frequency distribution and data characteristics. Finally, data pre-processing, which involves cleaning and transforming the data. The next phase is word embeddings, which use FastText embeddings (Al-jumaili, 2024) to turn text data into numerical vectors that can be used as input for the classification methods of the RNN and LSTM models. As seen in Figure 2, the K-Fold Cross Validation model (Azizah et al., 2022) will employ the outputs of each base model as input after being provided with a Confusion Matrix (Sharma et al., 2021).

**Data Collection**

The Python Scraper package was used to gather review data for the Merdeka Mengajar platform on the Google Play Store. Once the library has been installed via the package manager, import it. The primary key for data retrieval is the unique ID assigned to each Play Store application. After determining the language characteristics, country, sorting, amount of reviews, and filters, the ID used in this study is id.belajar.app. There are 17,848 review data in total, according to the findings of utilizing Google Colab's Python scraper package to scraping reviews from the Merdeka Mengajar Platform.

The research began in September 2024, and data was gathered from January 2022 to September 2024. The JSON structure of the acquired review data includes details such as ReviewId, userName, userImage, content, score, thumbsUpCount, reviewCreatedVersion, at, replyContent, repliedAt, and appVersion. Only the score and content data—which are user reviews—are used in this study. The score value ranges from 1 to 5, with a score of 4 or higher denoting positive sentiment and a score of 3 or lower denoting negative sentiment. The score and content will then be saved in a data frame.

**Data Labeling**

The information used in this study is limited to user-reviewed content and scores. A number of three or lower indicates negative sentiment, while a score of four or above indicates good sentiment. The scores range from 1 to 5. The score and content will then be saved in a CSV data frame. The score' column, which is thought to portray sentiment, will then be converted to the sentiment column with 'Positive' or 'Negative' categories according to the threshold value of 3. With 'content' as the feature (X) and sentiment as the aim (Y), the data is then separated into

*name of corresponding author

SinkrOn JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

training and test data, with a training data percentage of 80% and a test data proportion of 20%. The code below shows how the Python programming language is used to complete the data labeling process at this stage using the pandas module in Google Colab.

```
# Creating sentiment score categories (score: 1-3 negative, 4-5 positive)
df['sentiment'] = pd.cut(df['score'], bins=[0, 3, 5], labels=['negative', 'positive'], include_lowest=True)
# If score > 3 (or any other threshold), consider it positive, otherwise negative
df['sentiment'] = df['score'].apply(lambda score: 'Positive' if score > 3 else 'Negative')
# Splitting training and testing data
X = df['content']
y = df['sentiment']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

**Pre-processing**

Prior to being ready for sentiment categorization on Google Play Store app evaluations, the pre-processing phase is crucial. The data frame must first be prepared for processing by eliminating any unnecessary columns, leaving only the content and score columns.

Table 1. Data frame

| Content | Score |
|---|---|
| Jelek selalu tak bener suka error' | 1 |
| APLIKASI PMM MEMBUAT BEBAN GURU MAKIN BERAT | 1 |
| Kesulitan upload bukti karya | 2 |
| Hp iphone tdk bisa download | 1 |
| Saya tidak bisa me reset e kin yg telah disetujui. | 2 |

The image's flowchart illustrates how Play Store review data is processed (Figure 3), beginning with dataset retrieval and moving on to data cleaning to eliminate unnecessary components. After that, the data is tokenized, normalized, stopwords eliminated, and returned to its original form (stemming) (Nadia Ristya Dewi et al., 2022). The Sastrawi Python library is utilized in this study (Munna & Zuliarso, 2024). The Sastrawi stemming algorithm breaks words down into their most basic forms by applying the morphological principles of the Indonesian language. The data is prepared for use in additional analysis once all the procedures have been finished.

The process of checking review data for missing values and determining how to manage them is known as data cleaning. At this point, outliers are eliminated, duplicate data is also removed, and letters, punctuation, special symbols like "!?#&*," and other meaningless symbols are removed. In order for the model or algorithm to comprehend the meaning of each separated word, tokenization then completes the process of dividing the review text into tokens or distinct words. All words will be transformed into lowercase and put into the same format during the normalization phase. Additionally, words with the same meaning will be combined into a single form.

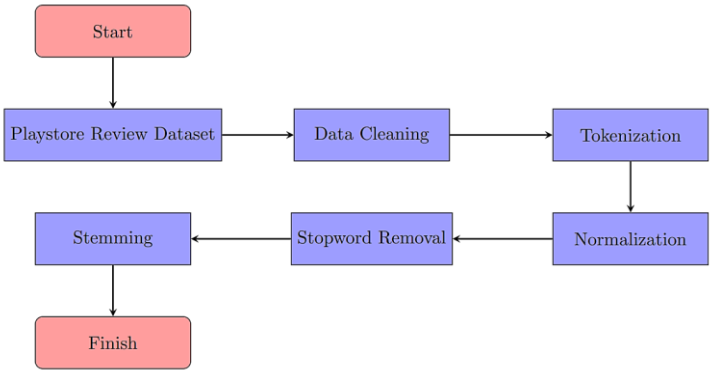


Fig 3. Preprocessing Stages

Stopword elimination is the process of eliminating popular Indonesian terms like "dan," "atau," "yang," "ini," and so forth that have no significance. The process of breaking words down into their most basic forms is called stemming. The Sastrawi Python library is utilized in this study. The Sastrawi stemming algorithm breaks words down into their most basic forms using the morphological laws of the Indonesian language. Since the terms "meminjam" and "pinjam" change to "pinjam" and "mengangsur" to "angsur," respectively, words that have the same meaning but distinct forms can be regarded as interchangeable. Table 2 displays the results of the data preprocessing.

*name of corresponding author

SinkrOn JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 2. Preprocessing

| Preprocessing | Review | Result |
|---|---|---|
| Data Cleaning | Teruntuk para PENELAAH Aksi Nyata, Dengan tegas saya sampaikan, Banyak aksi nyata yg sudah sesuai panduan tetapi di buat perbaikan, sementara saya menemukan ssalah satu akun yg aksi nyatanya cuma 4 slide, tetapi di Validasi. Ini sangat merugikan kami yg membuat aksi nyata dengan nyata beraksi bukan copas lalu upload. | Teruntuk para PENELAAH Aksi Nyata Dengan tegas saya sampaikan Banyak aksi nyata yg sudah sesuai panduan tetapi di buat perbaikan sementara saya menemukan ssalah satu akun yg aksi nyatanya cuma slide tetapi di Validasi Ini sangat merugikan kami yg membuat aksi nyata dengan nyata beraksi bukan copas lalu upload |
| Tokenization | Teruntuk para PENELAAH Aksi Nyata Dengan tegas saya sampaikan Banyak aksi nyata yg sudah sesuai panduan tetapi di buat perbaikan sementara saya menemukan ssalah satu akun yg aksi nyatanya cuma slide tetapi di Validasi Ini sangat merugikan kami yg membuat aksi nyata dengan nyata beraksi bukan copas lalu upload | ["Teruntuk", "para", "PENELAAH", "Aksi", "Nyata", "Dengan", "tegas", "saya", "sampaikan", "Banyak", "aksi", "nyata", "yg", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "ssalah", "satu", "akun", "yg", "aksi", "nyatanya", "cuma", "slide", "tetapi", "di", "Validasi", "Ini", "sangat", "merugikan", "kami", "yg", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copas", "lalu", "upload"] |
| Lowercasing | ["Teruntuk", "para", "PENELAAH", "Aksi", "Nyata", "Dengan", "tegas", "saya", "sampaikan", "Banyak", "aksi", "nyata", "yg", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "ssalah", "satu", "akun", "yg", "aksi", "nyatanya", "cuma", "slide", "tetapi", "di", "Validasi", "Ini", "sangat", "merugikan", "kami", "yg", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copas", "lalu", "upload"] | ["teruntuk", "para", "penelaah", "aksi", "nyata", "dengan", "tegas", "saya", "sampaikan", "banyak", "aksi", "nyata", "yg", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "ssalah", "satu", "akun", "yg", "aksi", "nyatanya", "cuma", "slide", "tetapi", "di", "validasi", "ini", "sangat", "merugikan", "kami", "yg", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copas", "lalu", "upload"] |
| Normalization | ["teruntuk", "para", "penelaah", "aksi", "nyata", "dengan", "tegas", "saya", "sampaikan", "banyak", "aksi", "nyata", "yg", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "ssalah", "satu", "akun", "yg", "aksi", "nyatanya", "cuma", "slide", "tetapi", "di", "validasi", "ini", "sangat", "merugikan", "kami", "yg", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copas", "lalu", "upload"] | ["untuk", "para", "penelaah", "aksi", "nyata", "dengan", "tegas", "saya", "sampaikan", "banyak", "aksi", "nyata", "yang", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "salah", "satu", "akun", "yang", "aksi", "nyatanya", "hanya", "slide", "tetapi", "di", "validasi", "ini", "sangat", "merugikan", "kami", "yang", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copy", "lalu", "unggah"] |
| Stopword Removal | ["untuk", "para", "penelaah", "aksi", "nyata", "dengan", "tegas", "saya", "sampaikan", "banyak", "aksi", "nyata", "yang", "sudah", "sesuai", "panduan", "tetapi", "di", "buat", "perbaikan", "sementara", "saya", "menemukan", "salah", "satu", "akun", "yang", "aksi", "nyatanya", "hanya", "slide", "tetapi", "di", "validasi", "ini", "sangat", "merugikan", "kami", "yang", "membuat", "aksi", "nyata", "dengan", "nyata", "beraksi", "bukan", "copy", "lalu", "unggah"] | ["penelaah", "aksi", "nyata", "tegas", "sampaikan", "banyak", "aksi", "nyata", "sesuai", "panduan", "buat", "perbaikan", "menemukan", "salah", "akun", "aksi", "nyatanya", "hanya", "slide", "validasi", "merugikan", "membuat", "aksi", "nyata", "beraksi", "bukan", "copy", "unggah"] |
| Stemming | ["penelaah", "aksi", "nyata", "tegas", "sampaikan", "banyak", "aksi", "nyata", "sesuai", "panduan", "buat", "perbaikan", "menemukan", "salah", "akun", "aksi", "nyatanya", "hanya", "slide", "validasi", "merugikan", "membuat", "aksi", "nyata", "beraksi", "bukan", "copy", "unggah"] | ["penelaah", "aksi", "nyata", "tegas", "sampai", "banyak", "aksi", "nyata", "sesuai", "panduan", "buat", "perbaiki", "temu", "salah", "akun", "aksi", "nyata", "hanya", "slide", "validasi", "rugi", "buat", "aksi", "nyata", "beraksi", "bukan", "copy", "unggah"] |

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

**Word Embeddings**

Word embeddings are carried out using a corpus from FastText Embedding, which handles suitably vast and diverse data by dynamically learning embeddings during model training. Within the embedding matrix, tokenized word inputs are transformed into numerical vectors using the model's FastText embeddings function (Darussalam & Arief, 2021). X_train and X_test are converted by the embedding layer into fixed-dimensional word vector representations (128 for each word) that maximize training and capture the semantic relationships or meaning of words. Depending on the particular data that is given, embedding will learn representations that are pertinent. Facebook AI Research created the word embedding technique known as FastText. By learning to anticipate the likelihood of a word given a particular context, its context, or vice versa, FastText expresses words as vectors.

The FastText method breaks words down into character n-grams, which are smaller word units. FastText is especially helpful for languages with rich morphology or when working with words that are not in the lexicon since it can capture morphological and syntactic information by taking these subword units into account. During training, FastText builds a lexicon of words and their subword units. After that, it learns to give words and subwords vector representations according to the patterns of co-occurrence in a particular text corpus. Semantic linkages between words and their sub-word components can be captured by the word vectors that are produced. (Ariyus & Manongga, 2024). At the Word Embeddings stage, this research utilizes FastText Embeddings with the following corpus:

[['mantap'], ['sangat', 'membantu'], ['mantap'], ['bagus'], ['sangat', 'baik', 'dan', 'memuaskan', 'dan', 'yg', 'utama', 'sangat', 'membantu', 'dalam', 'mengajar'], ['sangat', 'mempermudah'], ['mantab'], ['bagus'], ['sangat', 'membantu', 'sebagai', 'bahan', 'referensi'], ['membantu', 'guru', 'hebat']]

Using FastText embeddings, which were utilized in this study, word embedding makes text analysis more efficient by representing words as high-dimensional vectors that capture semantic links between words and language morphology.

**Modeling**

Classification is the next action to take. In data analysis, classification becomes a crucial step that helps organize data into classes or categories according to particular traits or attributes. Classification aids in locating pertinent patterns or traits in a dataset, which enables us to predict the class of fresh data using a trained model in the context of data mining and machine learning (Darojah et al., 2023). This approach is highly helpful in a number of domains, including recommendation systems for classifying products of user interest, medical diagnostics for classifying patients according to disorders, and sentiment analysis for classifying positive or negative evaluations.

At this point, a sentiment analysis model is developed to assess user reviews of the Merdeka Mengajar platform. Because of its sequential character and architecture tailored to sequence and list-shaped data, the Recurrent Neural Network (RNN) was selected as the classification model (Zuraiyah et al., 2023). RNN processes sequential input and uses an internal loop method to retain information from earlier steps. The ideal model is also developed in this procedure using LSTM (Amrustian et al., 2022). An instrument for examining pre-existing evaluation texts is LSTM. The input gate, which chooses which values to update; the forget gate, which must decide which information to use or not; and the output gate, which chooses the context to be generated, are the first three stages in the LSTM computation process. By controlling the information flow using unique gates, LSTM enables the model to retain and recall data for extended periods of time.

**Evaluation**

Following the completion of classification modeling in the preceding step, an assessment of the machine learning classification modeling's performance is carried out in this step. The purpose of this step is to evaluate the performance and efficacy of the two machine learning classification models. The confusion matrix is the method employed at this point. The overall results of both correct and wrong classifications are summarized using the confusion matrix. The combination of True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN) data can be used to view the confusion matrix. Using the formulas in Table 3, the confusion matrix yields four measurement results: accuracy, precision, recall, and F1-score (Hasnain et al., 2020)

Table 3. Structure of the Confusion Matrix

| | Positive Prediction | Negative Prediction |
|---|---|---|
| Positive Actual | True Positive (TP) | False Negative (FN) |
| Negative Actual | False Positive (FP) | True Negative (TN) |

The K-Fold Cross Validation approach is used in the subsequent evaluation step. With a value of k=5, this method yields five results that are shown for the metrics of accuracy, precision, recall, and F1-score obtained from the K-Fold Cross Validation procedure. The data is split into five parts (folds) for K-Fold Cross Validation, with one portion serving as the test data and the other four as the training data for each iteration. In 2022, Azizah et al. produced five sets of assessment metrics for each tested model as a result of this process being done five times (Azizah et al., 2022). To get a more reliable and precise performance estimate, this is done multiple times.

*name of corresponding author

SinkrOn JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

## RESULT

The Merdeka Mengajar Platform application's review score frequency distribution is displayed in Table 4, where a score of 5 with a frequency of more than 14,000 reviews predominates, indicating extremely high customer satisfaction. Conversely, a score of 4 indicates a moderate level, whereas low scores like 1, 2, and 3 have far lower frequencies.

Table 4. Distribution Review

| Scores | Frequency | Percentage (%) |
|---|---|---|
| 1 | 712 | 3,9892 |
| 2 | 207 | 1,1598 |
| 3 | 497 | 2,7846 |
| 4 | 1503 | 8,4211 |
| 5 | 14929 | 83,6452 |
| Total | 17,848 | 100,00 |

The distribution of sentiment classes is then calculated. The data shown in Table 4 indicates that the positive sentiment class is 16432 (92.1%), whereas the negative sentiment class is 1416 (7.9%). Reviews that are positive point to a positive and fulfilling Merdeka Mengajar user experience. Negative reviews reveal issues or discontent that customers have, like trouble utilizing the program, missing functionality, or bothersome bugs.

Next, Figure 4's temporal graph displays a rising trend in good ratings, but some peaks also show spikes in unfavorable reviews, such as when a particular app update is made available. Terms like "bug" and "difficult access," which characterize particular problems encountered by customers, are revealed in negative evaluations. Significantly positive emotion is also seen in the review sentiment distribution. Richer insights into user experience patterns can be obtained by presenting this information in greater detail during the debate.

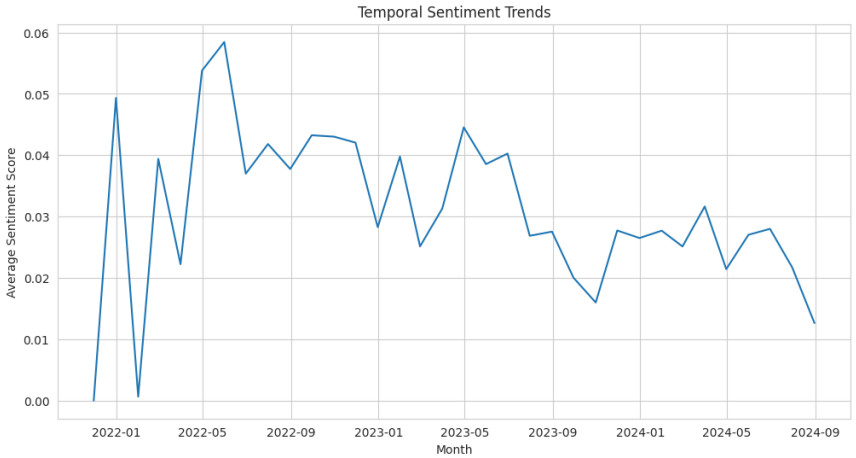


Fig 4. Temporal Sentiment Trend

Figure 5 plots the data index (data row) on the horizontal axis (X) against the ThumbsUpCount data (number of likes) on the vertical axis (Y). Many of the data's values are low (around 0), while several indices have notable peaks (high values). The greatest peak, which shows that certain data has a significantly higher number of likes than others, is located close to the 17500 index.

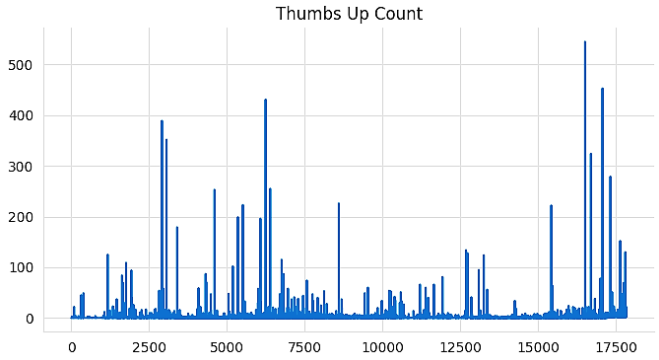


Fig 5. Plot of Thumb Count Data

*name of corresponding author

SinkrOn JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The data from FastText Embeddings, which is already in the form of numerical vectors, is used as input for the algorithm model with an 80% training data percentage and a 20% test data percentage at the confusion matrix comparison stage between the RNN and LSTM algorithms. Each RNN and LSTM model is subjected to sentiment analysis during a 20-epoch period. A confusion matrix table with a colored heatmap that contrasts the expected and actual values is then used to display the results.

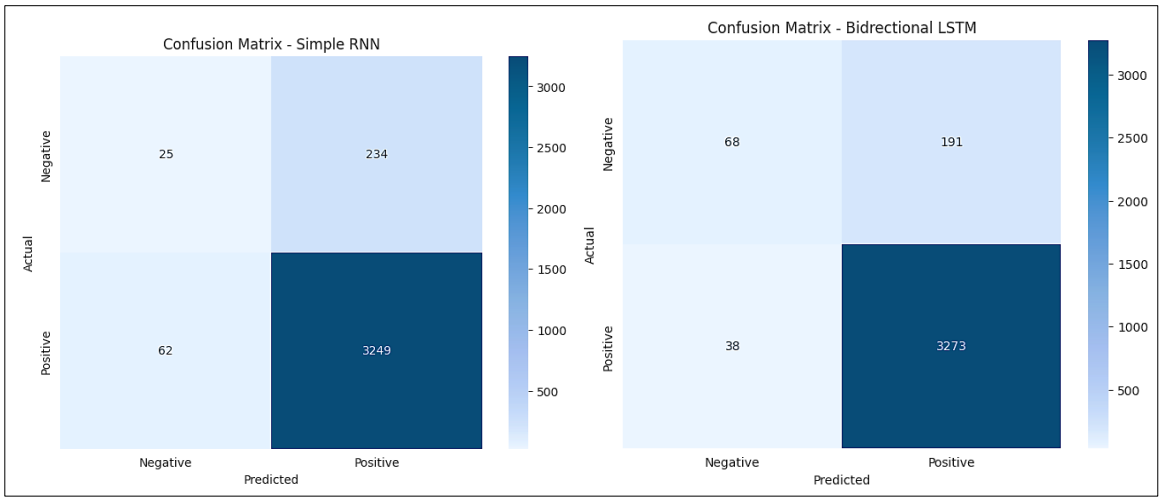


Fig 6. Confusion Matrix – Comparison of RNN and LSTM

The test results of Table 5 show that the RNN model achieved an accuracy of 91.70%, precision of 88.59%, recall of 91.70%, and an F1-Score of 89.75%. Meanwhile, the LSTM algorithm achieved an accuracy of 93.58%, precision of 92.28%, recall of 93.58%, and an F1-Score of 92.23%. These results support the hypothesis that LSTM-based algorithms are superior in capturing complex patterns compared to RNN for sequential data of user reviews on educational platforms.

Table 5. Comparison of RNN and LSTM Accuracy Result

| Algorithm | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| RNN | **91.70%** | 88.59% | 91.70% | 89.75% |
| LSTM | **93.58%** | 92.28% | 93.58%, | 92.23%. |

The results of the K-Fold Cross Validation cross-validation process are shown in Table 6, where the average accuracy, precision, recall, and F1 score of the RNN model were 92.23%, 89.06%, 92.23%, and 89.12%, respectively, when using the RNN and LSTM models. The model performed better with the LSTM, achieving an average accuracy of 92.70%, precision of 91.03%, recall of 92.70%, and F1-Score of 90.68%.

Table 6. Comparison of RNN and LSTM Accuracy Average

| Algorithm | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| RNN | **92,23%** | 89,06% | 92,23% | 89,12% |
| LSTM | **92,70%** | 91,03% | 92,70% | 90,68% |

Based on these findings, LSTM continues to outperform RNN across all evaluation metrics. The duration of each step in the training process of both models also shows that LSTM takes a little longer than RNN, which makes sense given that the LSTM architecture is more intricate due to the forget gate mechanism and cell state in processing sequential data. Because the K-Fold Cross Validation technique can lessen bias in data splitting, model evaluation becomes more stable.

The findings of this study are compared with those of earlier sentiment analysis research employing RNN and LSTM algorithms with varied embedding strategies and settings in Table 7. In comparison to earlier studies like Subowo et al. (2022) with an accuracy of 81% in the context of e-commerce and Widayat (2021) with 88.17% in movie reviews, the results of this study showed the highest accuracy of 93.58% using LSTM with FastText embeddings in the context of educational platforms. FastText's proficiency with out-of-vocabulary (OOV) words and the intricacy of the Indonesian language, which has not been thoroughly examined in prior research, allowed LSTM to outperform other models in this study.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 7. *Comparison with Previous Research.*

| Research | Algorithm | Embeddings | Context | Dataset | Best Accuracy |
|---|---|---|---|---|---|
| Nemes & Kiss (2021) | RNN | Word2Vec | COVID-19 on Twitter | 500 | 90% |
| Widayat (2021) | LSTM | Word2Vec | Movie Review | 25,000 | 88.17% |
| Isnain et al. (2022) | LSTM & Naive Bayes | TF-IDF | Kebijakan *New Normal* di Twitter | 15,000 | 83.33% (LSTM) |
| Utami (2022) | RNN | TF-IDF | Shopee application (Online Shop) | 800 | 80% |
| Subowo et al. (2022) | LSTM | FastText | E-commerce | 1000 | 81% |
| This Research | RNN & LSTM | FastText | Educational Platforms | 17,848 | **91.70%** (RNN), **93.58%** (LSTM) |

Thus, this study demonstrates how well LSTM and FastText work to increase model accuracy and make considerable progress in sentiment analysis, particularly in the area of educational platforms.

## DISCUSSIONS

The research results show that the Long Short-Term Memory (LSTM) algorithm has a significant advantage over the Recurrent Neural Network (RNN) in sentiment analysis of user reviews on the Merdeka Mengajar platform. With an accuracy of 93.58%, precision of 92.28%, recall of 93.58%, and an F1 score of 92.23%, LSTM proves to be more effective in capturing complex temporal relationships through the forget gate and cell state mechanisms. On the other hand, although RNN achieved a high accuracy of 91.70%, this architecture has limitations in handling long-term context. This difference is consistent with the findings of Isnain et al. (2022), which show that LSTM is superior in capturing language morphology and semantic context compared to RNN. However, LSTM requires longer training time, which poses a challenge for real-time applications.

Cross-validation with the K-Fold technique confirms that LSTM is more stable compared to RNN, with an average accuracy improvement of 1.47% during validation. These findings demonstrate the reliability of LSTM in data generalization. Sentiment distribution analysis reveals that the majority of user reviews are positive (92.1%), reflecting a good experience. However, negative reviews (7.9%) noted complaints such as application bugs and feature limitations, which support Ketaren et al.'s (2022) findings on technical challenges in educational platforms.

The application of FastText embeddings demonstrates the effectiveness of this technique in capturing semantic relationships between words, especially in the complex Indonesian language. FastText-based representations significantly contribute to the improvement of model accuracy, particularly in handling out-of-vocabulary words. This supports the research of Ariyus & Manongga (2024), although further studies on the influence of embedding parameters are needed for datasets with local context. This research has a main strength in the use of FastText embeddings and reliable cross-validation, but it also has limitations. The imbalance in sentiment distribution can lead to bias, so oversampling techniques or additional evaluation metrics such as the Matthews Correlation Coefficient need to be considered. Additionally, a dataset that only includes Google Play Store reviews may not fully reflect the experiences of teachers in various regions.

These results have important implications for application developers, such as fixing bugs, enhancing features, and optimizing application accessibility to support users in areas with network limitations. Future studies are recommended to explore more complex model architectures, such as transformer-based models, and to expand the dataset to include field surveys in order to provide more comprehensive insights.

## CONCLUSION

This research addresses how deep learning algorithms such as LSTM and RNN can be applied to improve the accuracy of sentiment analysis on user reviews of the digital education platform Merdeka Mengajar. The advantage of LSTM in capturing temporal relationships through the forget gate and cell state mechanisms provides a significant edge for analyzing complex sequential data. The distribution of mostly positive reviews (92.1%) also reflects a high level of user satisfaction with the platform's features, particularly in the aspect of self-training. The implementation of FastText embeddings in this study proved effective in capturing semantic relationships between words, especially for the complex Indonesian language. This contributes to the improvement of sentiment classification accuracy, especially in handling out-of-vocabulary words often found in informal reviews.

*name of corresponding author

SinkrOn JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.xxx.xxx p-ISSN : 2541-044X

With an accuracy of 93.58%, LSTM outperforms the RNN model, which has an accuracy of 91.70%. However, this study has limitations, including data bias from the uneven sentiment distribution (92.1% positive) and reliance on Google Play Store reviews, which might not accurately reflect user experiences in Indonesia. The 7.9% of negative evaluations include multiple grievances about feature limits, application glitches, and access issues in remote locations. According to this analysis, developers should prioritize testing their applications across a range of devices and network conditions. They should also think about adding features like offline mode to support accessibility in places with inadequate infrastructure. These insights will help developers of educational applications enhance the quality of their services and platform features.

## REFERENCES

Abubakar, H. D., & Umar, M. (2022). Sentiment Classification: Review of Text Vectorization Methods: Bag of Words, Tf-Idf, Word2vec and Doc2vec. *SLU Journal of Science and Technology*, *4*(1&2), 27–33. https://doi.org/10.56471/slujst.v4i.266

Al-jumaili, A. S. A. (2024). *NAMED ENTITY RECOGNITION*. *25*, 5258–5264. https://doi.org/10.12694/scpe.v25i6.3365

Amrustian, M. A., Widayat, W., & Wirawan, A. M. (2022). Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM. *Jurnal Media Informatika Budidarma*, *6*(1), 535. https://doi.org/10.30865/mib.v6i1.3527

Ariyus, D., & Manongga, D. (2024). *Enhancing Sentiment Analysis of Indonesian Tourism Video Content Commentary on TikTok : A FastText and Bi-LSTM Approach*. *14*(6), 18020–18028.

Atikah, L., Purwitasari, D., & Suciati, N. (2022). *DETEKSI KEJADIAN LALU LINTAS PADA TEKS TWITTER DENGAN PENDEKATAN KLASIFIKASI MULTI-LABEL BERBASIS DEEP LEARNING MULTI-LABEL CLASSIFICATION USING DEEP LEARNING APPROACH ON TWITTER*. *9*(1), 87–96. https://doi.org/10.25126/jtiik.202295206

Azizah, R. A., Bachtiar, F., & Adinugroho, S. (2022). Klasifikasi Kinerja Akademik Siswa Menggunakan Neighbor Weighted K-Nearest Neighbor dengan Seleksi Fitur Information Gain. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *9*(3), 605–614. https://doi.org/10.25126/jtiik.2022935751

C. Kalaiarasan, K. P. (2024). "Web Services Performance Prediction with Confusion Matrix and K-Fold Cross Validation to Provide Prior Service Quality Characteristics." *Journal of Electrical Systems*, *20*(2s), 284–292. https://doi.org/10.52783/jes.1139

Cahyani, D. E., & Patasik, I. (2021). Performance comparison of tf-idf and word2vec models for emotion text classification. *Bulletin of Electrical Engineering and Informatics*, *10*(5), 2780–2788. https://doi.org/10.11591/eei.v10i5.3157

Dang, N. C., Moreno-García, M. N., & De la Prieta, 2022. (2022). Sentiment Analysis Based on Deep Learning in E-Commerce. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *13369 LNAI*, 498–507. https://doi.org/10.1007/978-3-031-10986-7_40

Darojah, Z., Susetyoko, R., & Ramadijanti, N. (2023). Strategi Penanganan Imbalance Class Pada Model Klasifikasi Penerima Kartu Indonesia Pintar Kuliah Berbasis Neural Network Menggunakan Kombinasi SMOTE dan ENN. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *10*(2), 457–466. https://doi.org/10.25126/jtiik.20231026480

Darussalam, & Arief, G. (2021). Jurnal Resti. *Resti*, *1*(1), 19–25.

David, M. S., & Renjith, S. (2021). Comparison of word embeddings in text classification based on RNN and CNN. *IOP Conference Series: Materials Science and Engineering*, *1187*(1), 012029. https://doi.org/10.1088/1757-899x/1187/1/012029

Gomez-Adorno, H., Bel-Enguix, G., Sierra, G., Barajas, J. C., & Álvarez, W. (2024). Machine Learning and Deep Learning Sentiment Analysis Models: Case Study on the SENT-COVID Corpus of Tweets in Mexican Spanish. *Informatics*, *11*(2). https://doi.org/10.3390/informatics11020024

Hasnain, M., Pasha, M. F., Ghani, I., Imran, M., Alzahrani, M. Y., & Budiarto, R. (2020). Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking. *IEEE Access*, *8*, 90847–90861. https://doi.org/10.1109/ACCESS.2020.2994222

Iqmatul Pratiwi, Desi Rahmawati, T. T. D. S. (2024). *Lectura : Jurnal Pendidikan*. *15*, 596–610.

Isnain, A. R., Sulistiani, H., Hurohman, B. M., Nurkholis, A., & Styawati, S. (2022). Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, *8*(2), 299. https://doi.org/10.26418/jp.v8i2.54704

Ketaren, A., Rahman, F., Meliala, H. P., Tarigan, N., & Simanjuntak, R. (2022). Monitoring dan Evaluasi

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume xx, Number xx, month Year
DOI : https://doi.org/10.33395/sinkron.xxx.xxx
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Pemanfaatan Platform Merdeka Mengajar pada Satuan Pendidikan Aswinta. *Jurnal Pendidikan Dan Konseling*, *4*(6), 10340–10343. https://doi.org/https://doi.org/10.31004/jpdk.v4i6.10030

Khan, L., Amjad, A., Afaq, K. M., & Chang, H. T. (2022). Deep Sentiment Analysis Using CNN-LSTM Architecture of English and Roman Urdu Text Shared in Social Media. *Applied Sciences (Switzerland)*, *12*(5). https://doi.org/10.3390/app12052694

Kurniasari, L., & Setyanto, A. (2020). Sentiment Analysis using Recurrent Neural Network. *Journal of Physics: Conference Series*, *1471*(1). https://doi.org/10.1088/1742-6596/1471/1/012018

Liu, C., Fang, F., Lin, X., Cai, T., Tan, X., Liu, J., & Lu, X. (2021). Improving sentiment analysis accuracy with emoji embedding. *Journal of Safety Science and Resilience*, *2*(4), 246–252. https://doi.org/10.1016/j.jnlssr.2021.10.003

Maru, M. G., Tamowangkay, F. P., Pelenkahu, N., & Wuntu, C. (2021). Teachers' perception toward the impact of platform used in online learning communication in the eastern Indonesia. *International Journal of Communication and Society*, *4*(1), 59–71. https://doi.org/10.31763/ijcs.v4i1.321

Maulana, R., Rahayuningsih, P. A., Irmayani, W., Saputra, D., & Jayanti, W. E. (2020). Improved Accuracy of Sentiment Analysis Movie Review Using Support Vector Machine Based Information Gain. *Journal of Physics: Conference Series*, *1641*(1). https://doi.org/10.1088/1742-6596/1641/1/012060

Merinda Lestandy, Abdurrahim Abdurrahim, & Lailis Syafa'ah. (2021). Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(4), 802–808. https://doi.org/10.29207/resti.v5i4.3308

Munna, A., & Zuliarso, E. (2024). *Interpretasi model Stacking Ensemble untuk analisis sentimen ulasan aplikasi pinjaman online menggunakan LIME*. *21*(2), 183–196.

Nadia Ristya Dewi, Yulia Puspaningrum, E., & Maulana, H. (2022). Analisis Sentimen Tweet Vaksinasi Covid-19 Menggunakan RNN Dengan Metode TF-IDF Dan Word2Vec. *Jurnal Informatika Dan Sistem Informasi*, *3*(1), 56–65. https://doi.org/10.33005/jifosi.v3i1.449

Ndari, W., Suyatno, Sukirman, & Mahmudah, F. N. (2023). Implementation of the Merdeka Curriculum and Its Challenges. *European Journal of Education and Pedagogy*, *4*(3), 111–116. https://doi.org/10.24018/ejedu.2023.4.3.648

Nemes, L., & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication*, *5*(1), 1–15. https://doi.org/10.1080/24751839.2020.1790793

Sharma, R., Shrivastava, S., Kumar Singh, S., Kumar, A., Saxena, S., & Kumar Singh, R. (2021). Deep-Abppred: Identifying antibacterial peptides in protein sequences using bidirectional LSTM with word2vec. *Briefings in Bioinformatics*, *22*(5), 1–19. https://doi.org/10.1093/bib/bbab065

Shumaly, S., Yazdinejad, M., & Guo, Y. (2021). Persian sentiment analysis of an online store independent of pre-processing using convolutional neural network with fastText embeddings. *PeerJ Computer Science*, *7*, 1–22. https://doi.org/10.7717/peerj-cs.422

Subowo, E., Adi Artanto, F., Putri, I., & Umaedi, W. (2022). Algoritma Bidirectional Long Short Term Memory untuk Analisis Sentimen Berbasis Aspek pada Aplikasi Belanja Online dengan Cicilan. *Jurnal Fasilkom*, *12*(2), 132–140. https://doi.org/10.37859/jf.v12i2.3759

Suwitono, Y. A., & Kaunang, F. J. (2022). Implementasi Algoritma Convolutional Neural Network (CNN) Untuk Klasifikasi Daun Dengan Metode Data Mining SEMMA Menggunakan Keras. *Jurnal Komtika (Komputasi Dan Informatika)*, *6*(2), 109–121. https://doi.org/10.31603/komtika.v6i2.8054

Tan, K. L., Lee, C. P., Anbananthen, K. S. M., & Lim, K. M. (2022). RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network. *IEEE Access*, *10*, 21517–21525. https://doi.org/10.1109/ACCESS.2022.3152828

Utami, H. (2022). Analisis Sentimen dari Aplikasi Shopee Indonesia Menggunakan Metode Recurrent Neural Network. *Indonesian Journal of Applied Statistics*, *5*(1), 31. https://doi.org/10.13057/ijas.v5i1.56825

Widayat, W. (2021). Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning. *Jurnal Media Informatika Budidarma*, *5*(3), 1018. https://doi.org/10.30865/mib.v5i3.3111

Zuraiyah, T. A., Mulyati, M. M., & Harahap, G. H. F. (2023). Perbandingan Metode Naïve Bayes, Support Vector Machine Dan Recurrent Neural Network Pada Analisis Sentimen Ulasan Produk E-Commerce. *Multitek Indonesia*, *17*(1), 27–43. https://doi.org/10.24269/mtkind.v17i1.7092

*name of corresponding author