ISSN 2087-3336 (Print) | 2721-4729 (Online)

TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika
Vol. 13, No. 1, 2026, page. 1-8
http://jurnal.sttmcileungsi.ac.id/index.php/tekno
https://doi.org/10.37373/tekno.v13i1

The implementation of text mining to improve google classroom performance based on user review

Anggita Yekti Pawestri, Annisa Uswatun Khasanah*

*Universitas Islam Indonesia, Gedung K.H. Mas Mansyur, Kampus Terpadu UII, Jl. Kaliurang KM 14,5, Umbulmartani, Ngemplak, Krawitan, Sleman, Daerah Istimewa Yogyakarta, 55584

* Corresponding Author: annisa.uswatun@uii.ac.id

Submitted: 23/4/2025 Revised: 15/6/2025 Accepted: 14/7/2025

ABSTRACT

The internet is developing rapidly, especially in education (e-learning). Google Classroom is an e-learning platform that has been widely used for online learning. Google Classroom gets an average rating of 3.5 out of 5 on Google Play, so it is necessary to research to increase the application's rating based on user reviews with text mining. The data used in this study was 77,454 user reviews from Google Play. Data processing uses the Textblob and Naïve Bayes Classifier methods to determine user sentiment. The sentiment analysis results with a 70/30 split data yield an accuracy of 92.39%. Data with negative sentiments is then processed with the Association Rules method to find words used as problem keywords, which include keywords' assignment', 'upload,' 'submit,' 'file,' 'class,' 'notification,' and 'dark.' The word is then analyzed using a fishbone diagram to find the root cause. The root of the problems includes problems uploading files that are integrated with Google Drive, there is no setting to change the display to 'dark mode,' The user has a good internet connection, but the file upload process is slow, and so on. The recommendations for improvements are synchronizing file integration with Google Drive, changing the display to dark mode according to user preferences, and updating the application server so that the file upload process can be done more quickly.

Keywords: Sentiment analysis; text mining; Naïve Bayes classifier; Google Classroom

1. Introduction

The rapid development of the internet has significantly influenced various aspects of society, including education. Globally, 59.5% of the population uses the internet [1], while in Indonesia, this figure is even higher at 79.5% [2]. This widespread internet penetration has paved the way for the integration of digital tools in education, leading to the emergence of e-learning as a critical component of modern teaching and learning practices. E-learning leverages information technology to facilitate training, deliver learning materials, and enable communication between students and teachers, thereby streamlining the learning process. One of the primary tools in e-learning is the Learning Management System (LMS). As described by Cavus and Alhih [3], an LMS is a technological platform designed to manage educational or training resources, distribute materials online, and monitor learning progress effectively.

With the rapid advancement of digital technology, the use of Learning Management Systems (LMS) in education has become increasingly widespread. Initially accelerated during the Covid-19 pandemic, online learning has now become a standard complement to traditional education. Traditional education systems heavily relied on teachers, textbooks, and rote memorization, often leading to monotonous and disengaging learning experiences for students. Modern technology-enhanced learning, however, complements the educational process by accommodating diverse learning styles and abilities, ultimately enhancing student engagement and improving learning outcomes [4]. The integration of technology, communication, and information in education has had a profound impact [5]. Key trends in learning include (1) a shift from teacher-centered to student-centered approaches, (2) the increasing popularity and expansion of open and distance learning, and (3) the enhanced accessibility to a wider variety of learning resources.

Even in the post-pandemic era, online and hybrid learning models continue to play a vital role in education. Schools and universities leverage LMS platforms not only to provide remote learning opportunities but also to enhance in-class experiences. Many distance learning activities have collaborated with online learning platforms. Google Classroom is one of the most widely used learning platforms. Based on a survey conducted by Arus Survei Indonesia (ASI), 26.1% of people have used the Google Classroom platform during distance learning [6]. Google Classroom occupies the highest percentage as an application for distance learning activities, followed by other platforms such as Ruangguru (17.1%) and Rumah Belajar (15.2%). Platforms like Google Classroom remain popular due to their user-friendly interface and integration with other digital tools, enabling educators to efficiently manage assignments, distribute materials, and foster communication with students. This trend signifies a shift toward a more technology-driven approach in education, ensuring accessibility and innovation in learning.

The increasing number of users of Google Classroom has raised criticism and suggestions from users. Based on the rating level expressed by users on Google Play in 2024, the Google Classroom application received an average rating of 3.5 out of 5 in the Google Play store. The rating is relatively low compared to other Google applications, so improvements need to be made to increase the rating of the Google Classroom application. Before enhancements are made, it is necessary to find the root of the problem that causes the Google Classroom application to have a low rating. Reviews from Google Classroom app users are used as a source of data to explore issues from the app. Text mining is one technique that can be implemented to classify data by finding interesting patterns from large text data sets [7]. According to Antons et al. [8], text mining requires a quantitative approach to analyze text data.

In this research, text mining is used to explore the review data of Google Classroom application users, especially on the Android platform. Reviews that become data used in this research are in English. The language was chosen because it is the majority language used in giving reviews, so that it can represent criticism from users. Data collection will be focused on the Google Play Store platform by scraping data through the AppFollow website. In this research, sentiment classification analysis is carried out to classify reviews with positive and negative sentiments. The Naïve Bayes Classifier method is used in this research because it is considered one of the sentiment analysis methods that are simple, easy to understand, and have a reasonably high accuracy rate [9]. This method can project possibilities that can happen in the future based on experiences that have occurred, which is now known as Bayes' Theorem. In addition, a lexicon-based labeling library in Python, Textblob, will also be used. Textblob is one of the libraries in Python for text data processing and uses NLTK for Natural Language Processing (NLP), which is unlabeled [10]. The sentiment analysis classification is a hybrid method between machine learning and lexicon-based sentiment analysis. Association rule mining will also be used in this research. According to Santoso [11], association rules are used to find rules or combinations of items that occur most frequently in data. This study uses association rules to find combinations of these words. After being classified, the root of the problem can be found using root cause analysis in the form of a fishbone diagram so that it can be used as a reference for developing recommendations for improvements to the Google Classroom application.

2. Method

Text mining has been widely used for sentiment analysis in user reviews. Conducted sentiment analysis on 66,000 MOOC reviews using text mining techniques [12]. The study compared machine learning, ensemble learning, and deep learning methods for sentiment classification in education. Results showed that deep learning architectures achieved the highest predictive accuracy among the methods tested. Conducted a sentiment analysis about online education during the Covid-19 era from Twitter [13]. This research aims to evaluate the effectiveness of e-learning by examining people's sentiments toward it. Integrated text mining and machine learning to analyze student evaluations and uncover how sentiments, emotions, and gender influence teacher recommendation using the proposed model, the EPDM + ML [14]. The model achieved perfect prediction accuracy, highlighting its effectiveness in improving teaching quality through data-driven insights. Conducted study in sentiment analysis for TikTok using Naive Bayes and Support Vector Machine [15].

ISSN 2087-3336 (Print) | 2721-4729 (Online)
DOI 10.37373/tekno.v13i1.1479

This current study applied text mining to determine user sentiment in Google Classroom. The data used in this study consists of 77454 user reviews. This data was obtained using web scraping through the AppFollow site for six months. The data preprocessing stage is carried out to clean the data of unnecessary information and anomalies. This process was carried out using Jupyter Notebook and Python 3.9. The data preprocessing includes several steps:

a. Case folding: Homogenizes all letters so that letters that were initially capitalized become lowercase in Table 1.

Table 1. Case folding step

Before *Case Folding*	After *Case Folding*
This app is mostly like trash ðŸ¤¦â€â™€ï¸ I've been using this app for a while and i changed my acc password, and just like that, i can't log out and whenever I use it, it will tell me something went wrong and wants me to try again… 🤷🤷	*this app is mostly like trash ðŸ¤¦â€â™€ï¸ i've been using this app for a while and i changed my acc password, and just like that, i can't log out and whenever i use it, it will tell me something went wrong and wants me to try again…* 🤷🤷

b. Cleaning: remove punctuation marks to delete unwanted symbols or punctuation marks in text data, remove non-ASCII characters, which are Character Encoding Standard for Electronic Communication. And the last was removing emojis in the text data to Table 2, Table 3, Table 4, Table 5

Table 2. Punctuation removal step

Before *Remove Punctuation*	After *Remove Punctuation*
this app is mostly like trash ðŸ¤¦â€â™€ï¸ i've been using this app for a while and i changed my acc password, and just like that, i can't log out and whenever i use it, it will tell me something went wrong and wants me to try again... 🤷🤷	*this app is mostly like trash ðŸ¤¦â€â™€ï¸ ive been using this app for a while and i changed my acc password and just like that i cant log out and whenever i use it it will tell me something went wrong and wants me to try again* 🤷🤷

Table 3. Non-ASCII character removal step

Before Non-*ASCII Character Removal*	After Non-*ASCII Character Removal*
this app is mostly like trash ðÿ¤¦â€â™€ï¸ ive been using this app for a while and i changed my acc password and just like that i cant log out and whenever i use it it will tell me something went wrong and wants me to try again 🤷🤷	*this app is mostly like trash ive been using this app for a while and i changed my acc password and just like that i cant log out and whenever i use it it will tell me something went wrong and wants me to try again* 🤷🤷

Table 4. Emoji removal step

Before *Emoji Removal*	After *Emoji Removal*
this app is mostly like trash ive been using this app for a while and i changed my acc password and just like that i cant log out and whenever i use it it will tell me something went wrong and wants me to try again 🤷🤷	*this app is mostly like trash ive been using this app for a while and i changed my acc password and just like that i cant log out and whenever i use it it will tell me something went wrong and wants me to try again*

c. Tokenizing**:** separates the words in a sentence, leaving only the needed words.

Table 5. Tokenizing step

Before *Tokenizing*	After *Tokenizing*
this app is mostly like trash ive been using this app for a while and i changed my acc password and just like that i cant log out and whenever i use it it will tell me something went wrong and wants me to try again	*['this', 'app', 'is', 'mostly', 'like', 'trash', 'ive', 'been', 'using', 'this', 'app', 'for', 'a', 'while', 'and', 'i', 'changed', 'my', 'acc', 'password', 'and', 'just', 'like', 'that', 'i', 'cant', 'log', 'out', 'and', 'whenever', 'use' 'it', 'it', 'will', 'tell', 'me', 'something', 'went', 'wrong', 'and', 'wants', 'me', 'to', 'try', 'again']*

d. Stopwords removal**:** remove meaningless words, including the, and, but, and or, and others to Table 6.

Table 6. Stopwords removal step

Before *Tokenizing*	After *Tokenizing*
['this', 'app', 'is', 'mostly', 'like', 'trash', 'ive', 'been', 'using', 'this', 'app', 'for', 'a', 'while', 'and', 'i', 'changed', 'my', 'acc', 'password', 'and', 'just', 'like', 'that', 'i', 'cant', 'log', 'out', 'and', 'whenever', 'use' 'it', 'it', 'will', 'tell', 'me', 'something', 'went', 'wrong', 'and', 'wants', 'me', 'to', 'try', 'again']	*['app', 'mostly', 'like', 'trash', 'using', 'app', 'changed', 'acc', 'password', 'like', 'log', 'whenever', 'use', 'tell', 'something', 'went', 'wrong', 'wants', 'try', 'again']*

After preprocessing, the data label must be given. Textblob, one of the Python libraries that performs sentiment analysis with a lexicon-based approach, provides labels on the data. The sentiment analysis results with Textblob will give the output of subjectivity and polarity values. The polarity value is used as a reference to determine the sentiment of text data into the negative, neutral, or positive category. Labeling criteria on Textblob refer to the polarity score with the following conditions [16]:

a. Polarity score < 0, then the sentiment is negative
b. Polarity score = 0, then neutral sentiment
c. Polarity score > 0, then positive sentiment

Based on the results of sentiment labeling using the Textblob library, the data was divided into negative sentiment (27,233 data), positive sentiment (27,931 data), and neutral sentiment (22,290 data). The visualization of the percentage of sentiment labeling results is shown in Figure 1. It can be known that positive and negative sentiments have similar percentages, which means from the Google Classroom data review the number of negative comments and positive comments were almost the same.

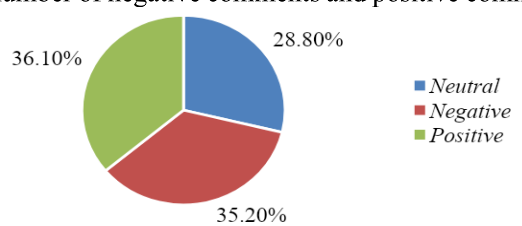

Figure 1. Visualization of sentiment labeling results

3. Result and Discussion

a. Classification Using Naïve Bayes Algorithm

Classification with the Naïve Bayes algorithm is done to determine the accuracy of the classification model in predicting sentiment. The Naïve Bayes Classifier algorithm is a machine learning model that is supervised or needs to be trained. The illustration of the classification model is shown in Figure 2**.**

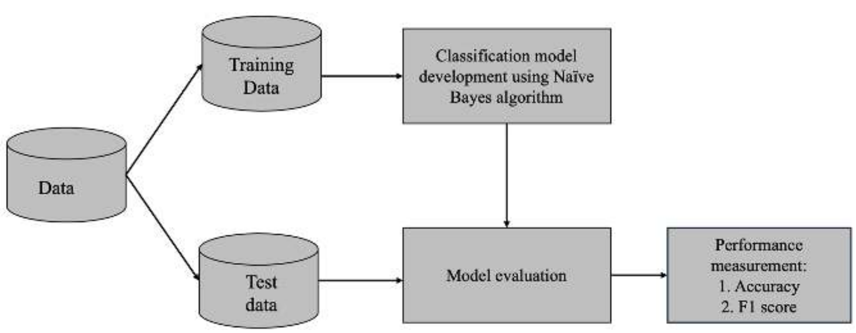

Figure 2. Classification model

Therefore, training and test data are required to apply the classification model. Data splitting uses the 'train_test_split' library from sklearn in Python. Several combinations of training and testing data were generated to get the best accuracy; the combinations are: 50/50, 60/40, 70/30, and 80/20 [17].

Training data was used to build the model, and test data was used to evaluate the model. The performance measurement was performed by calculating the model accuracy and F1 score. The comparative results of classification using the Naïve Bayes classifier with four combinations of split data are shown in Table 7.

Table 7. Comparison of split data size, accuracy, and f1 score

		Accuracy	F1 Score
Split Data	50/50	92.19%	92.06%
	60/40	92.27%	92.3%
	70/30	92.39%	92,.5%
	80/20	92.33%	92.17%

Based on the accuracy rate and F1 score values obtained, it is known that classification with a data splitting size of 70% training data and 30% test data produces the best output with an accuracy value of 92.39%.

b. Association Rules

The data extracted is 27,233 customer reviews with negative sentiments containing criticisms and complaints about the Google Classroom application. In data processing, the stemming process is implemented using the Snowball Stemmer algorithm, which aims to minimize word diversity. The Snowball Stemmer algorithm is a development of the Porter Stemmer algorithm, which is considered to have the best output compared to other stemmer algorithms [18]. Association rules processing uses the FP-Growth algorithm. The FP-Growth algorithm is an efficient algorithm for finding frequent item sets, which are considered to have better performance when compared to the Apriori algorithm [19]. Data mining is done to find the most frequent itemset in the data with FP-Growth by creating an FP-Tree structure) [20].

Association rules data with the FP-Growth algorithm is applied in RapidMiner with min support of 0.001 and min confidence of 0.1. The support value represents the percentage of words in the data, while the confidence value represents the strength of the association between the words formed in Table 8. If the resulting lift ratio value is more than 1, it can be concluded that the relationship or association between the conclusion and premises is significant. The words which are frequently found in data are 'assignment', 'upload', 'submit', 'file', 'class', 'notification', and 'dark'.

Table 8. Frequent words

Words	**Support**
class	0.042
assignment	0.036
upload	0.024
submit	0.019
file	0.018
notification	0.017
dark	0.009

Based on the AR results, it can be concluded that the frequent words are correlated with several other words to determine the specific issues. The assignment was correlated with miss, late, post, notification, etc. Upload was correlated with slow, speed, assignment, etc. Submit was correlated with time, assignment, worst, late, take, miss, etc. File was correlated with attachment, time, take, upload, slow, etc. Class was correlated with students, teacher, online, attend, worst, etc. Notifications ware correlated with received, late, missed, assignment, etc. Dark was correlated with the mode, please, theme.

c. Recommendations for Improvement

Based on the data processing results with association rules, the problems that can be analyzed related to assignments; upload and submit files; class entry system; notifications issues; and displays problems. The root cause is analyzed based on four aspects: product, people, place, and process. The product aspect is related to the product feature, which is the Google Classroom application. The people aspect is associated with the human error problems related to obstacles in using the application. The place aspect refers to difficulties in accessing the Google Classroom application. The last aspect is the process related to the problem users encounter when running the application. After the data has been analyzed, the improvement recommendation is formulated. The recommendations based on four aspects are shown in Table 9, Table 10, Table 11, Table 12.

Table 9. Recommendations for improvement according to product aspects

Product Aspects	
Problems	Recommendations for Improvement
1. Obstacles in uploading files integrated with Google Drive 2. No theme variation for the application interface 3. No setting to change the display to 'dark mode'	1. Improve file integration with Google Drive 2. Add a feature to change the interface theme of the Google Classroom application to make it more attractive to users 3. Settings to change the display to dark mode according to the user's preference

Table 10. Recommendations for improvement according to the people aspect

People Aspects	
Problems	Recommendations for Improvement
Communication between teachers and students is not optimal, leading to miscommunication,	Improve and enhance the communication forum between students and teachers.

Table 11. Recommendations for improvement according to the place aspect

Place Aspects	
Problems	Recommendations for Improvement
1. There is no notification even though the user has an internet connection 2. Users have a good internet connection, but the file uploading process is slow	1. Optimize notifications so users can still receive notifications and do not experience delays 2. Update the Google Classroom server so the file upload process can be done faster

Table 12. Recommendations for improvement according to the process aspect

Process Aspects	
Problems	Recommendations for Improvement
1. The assignment is submitted on time, but the status is late 2. Assignments that have been collected have a missed status on the Dashboard 3. There is an error in the grading feature 4. Problems collecting photo and video file types 5. The result of the uploaded file is not visible 6. Failed to enter the classroom 7. Late notification 8. No notification when there is a new comment	1. Improving the assignment assessment feature 2. Improvements need to be made to the submission process for various types of files 3. Improvements need to be made to the server 4. Updates are needed to resolve bugs.

4. Conclusion

Based on the study that has been conducted, it can be concluded that the best combination of splitting data to build the classification model is 70/30. The accuracy rate using Naïve Bayes Classifier is 92.39%. Based on the processing of association rules with the FP-Growth method, it can be found that the word appears most often when the support value of the word is examined. The words that appear most often in Google Classroom application user reviews are 'class', 'assignment', 'upload', 'submit', 'file', 'notification', and 'dark'. The root of the problem can be formulated using a fishbone diagram based on four aspects: product, people, place, and process. The improvement recommendations are developed to increase the performance of Google Classroom.

Reference

[1] A. Petrosyan, “Number of internet and social media users worldwide as of October 2024,” https://www.statista.com/statistics/617136/digital-population-worldwide/.

[2] Asosiasi Penyelenggara Jasa Internet Indonesia, “APJII Jumlah Pengguna Internet Indonesia Tembus 221 Juta Orang,” https://apjii.or.id/berita/d/apjii-jumlah-pengguna-internet-indonesia-tembus-221-juta-orang#:~:text=Asosiasi%20Penyelenggara%20Jasa%20Internet%20Indonesia%20(APJII)%20mengumumkan%20jumlah%20pengguna%20internet,jiwa%20penduduk%20Indonesia%20tahun%202023.

[3] N. Cavus and M. S. Alhih, “Learning Management Systems Use in Science Education,” *Procedia Soc Behav Sci*, vol. 143, pp. 517–520, Aug. 2014, doi: 10.1016/j.sbspro.2014.07.429. https://doi.org/10.1016/j.sbspro.2014.07.429

[4] D. Ambarwati, U. B. Wibowo, H. Arsyiadanti, and S. Susanti, “Studi literatur: Peran inovasi pendidikan pada pembelajaran berbasis teknologi digital,” *Jurnal Inovasi Teknologi Pendidikan*, vol. 8, no. 2, pp. 173–184, 2022.

[5] R. Dwi Gunawan, A. Sutisna, and E. Fitri Ana, “Literature review: The role of learning management system (LMS) in improving the digital literacy of educators,” *Jurnal Inovasi Teknologi Pendidikan*, vol. 11, no. 2, pp. 116–123, 2024, doi: 10.21831/jitp.v11i2.56326.

[6] I. Kamil and B. Galih, “Survei: Google Classroom Jadi Platform Belajar Paling Sering Digunakan Saat PJJ,” https://nasional.kompas.com/read/2020/10/16/18264341/survei-google-classroom-jadi-platform-belajar-paling-sering-digunakan-saat?page=all#google_vignette.

[7] E. E. Pratama and R. L. Atmi, “A Text Mining Implementation Based on Twitter Data to Analyse Information Regarding Corona Virus in Indonesia,” *Journal of Computers for Society*, vol. 1, no. 1, pp. 91–100, 2020, [Online]. Available: https://ejournal.upi.edu/index.php/JCS

[8] D. Antons, E. Grünwald, P. Cichy, and T. O. Salge, “The application of text mining methods in innovation research: current state, evolution patterns, and development priorities,” in *R and D Management*, Blackwell Publishing Ltd, Jun. 2020, pp. 329–351. doi: 10.1111/radm.12408. https://doi.org/10.1111/radm.12408

[9] P. S. M. Suryani, L. Linawati, and K. O. Saputra, “Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia,” *Majalah Ilmiah Teknologi Elektro*, vol. 18, no. 1, p. 145, May 2019, doi: 10.24843/mite.2019.v18i01.p22. https://doi.org/10.24843/MITE.2019.v18i01.P22

[10] C. Kaur and A. Sharma, “Twitter Sentiment Analysis on Coronavirus using Textblob,” *EasyChair Preprint*, no. 2974, 2020.

[11] M. H. Santoso, “Application of Association Rule Method Using Apriori Algorithm to Find Sales Patterns Case Study of Indomaret Tanjung Anom,” *Brilliance: Research of Artificial Intelligence*, vol. 1, no. 2, pp. 54–66, Dec. 2021, doi: 10.47709/brilliance.v1i2.1228. https://doi.org/10.47709/brilliance.v1i2.1228

[12] A. Onan, “Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach,” *Computer Applications in Engineering Education*, vol. 29, no. 3, pp. 572–589, May 2021, doi: 10.1002/cae.22253. https://doi.org/10.1002/cae.22253

[13] M. Mujahid *et al.*, “Sentiment analysis and topic modeling on tweets about online education during covid-19,” *Applied Sciences (Switzerland)*, vol. 11, no. 18, Sep. 2021, doi: 10.3390/app11188438. https://doi.org/10.3390/app11188438

[14] K. Okoye, A. Arrona-Palacios, C. Camacho-Zuñiga, J. A. G. Achem, J. Escamilla, and S. Hosseini, “Towards teaching analytics: a contextual model for analysis of students’ evaluation of teaching through text mining and machine learning classification,” *Educ Inf Technol (Dordr)*, vol. 27, pp. 3891–3933, 2022. https://doi.org/10.1007/s10639-021-10751-5

[15] Friska Aditia Indriyani, Ahmad Fauzi, and Sutan Faisal, “Analisis sentimen aplikasi tiktok menggunakan algoritma naïve bayes dan support vector machine,” *TEKNOSAINS : Jurnal Sains, Teknologi dan Informatika*, vol. 10, no. 2, pp. 176–184, Jul. 2023, doi: 10.37373/tekno.v10i2.419. https://doi.org/10.37373/tekno.v10i2.419

[16] P. Shah, “Sentiment Analysis using TextBlob,” https://towardsdatascience.com/my-absolute-go-to-for-sentiment-analysis-textblob-3ac3a11d524.

[17] A. Rácz, D. Bajusz, and K. Héberger, “Effect of dataset size and train/test split ratios in qsar/qspr multiclass classification,” *Molecules*, vol. 26, no. 4, Feb. 2021, doi: 10.3390/molecules26041111. https://doi.org/10.3390/molecules26041111

[18] M. Anjali and G. Jivani, “A Comparative Study of Stemming Algorithms,” *Int.J. Comp. Tech. Appl*, vol. 2, no. 6, pp. 1930–1938, 2011, [Online]. Available: www.ijcta.com

[19] M. S. Mythili and A. R. M. Shanavas, “Performance Evaluation of Apriori and FP-Growth Algorithms,” *Int J Comput Appl*, vol. 79, no. 10, pp. 975–8887, 2013. https://doi.org/10.5120/13779-1650

[20] D. Widiastuti and N. Sofi, “Analisis Perbandingan Algoritma Apriori Dan Fp-Growth Pada Transaksi Koperasi,” *UG Journal*, vol. 8, no. 1, 2014.