Published online on the page : https://ejournal.iainbukittinggi.ac.id/index.php/ijokid/

Knowbase : International Journal of Knowledge in Database

| ISSN (Print) 2798-0758 | ISSN (Online) 2797-7501 |

Sentiment Analysis And Topic Modeling on User Reviews of Online Tutoring Applications Using Support Vector Machine and Latent Dirichlet Allocation

Mohammad Rezza Fahlevvi[1,*]

[1]*Institut Pemerintahan Dalam Negeri, Bandung, Indonesia*

Article Information

Article History:
Accepted by the Editor: October 27, 2022
Final Revision: October 26, 2022
Published Online: November 28, 2022

Key Word

Ruangguru
Google Play Store (GPS)
Sentiment Analyst
Topic Modelling
Support Vector Machine
Latent Dirichlet Allocation (LDA)
Confusion Matrix
Rouge

Correspondence

E-mail: rezza@ipdn.ac.id*

A B S T R A C T

Ruangguru is an online non-formal education application in Indonesia. There are several appealing features that encourage students to study online. The app's release on the Google Play Store will assist app developers in receiving feedback through the review feature.Users submit various topics and comments about Ruangguru in the review feature of Ruangguru, making it difficult to manually identify public sentiments and topics of conversation. Opinions submitted by users on the review feature are interesting to research further. This study aims to classify user opinions into positive and negative classes and model topics in both classes. Topic modeling aims to find out the topics that are often discussed in each class. The stages of this study include data collection, data cleaning, data transformation, and data classification with the Support Vector Machine method and the Latent Dirichlet Allocation method for topic modeling. The results of topic modeling with the LDA method in each positive and negative class can be seen from the coherence value. Namely, the higher the coherence value of a topic, the easier the topic is interpreted by humans. The testing process in this study used Confusion Matrix and ROUGE. The results of model performance testing using the Confusion Matrix are shown with accuracy, precision, recall, and f-measure values of 0.9, 0.9, 0.9, and 0.89, respectively. The results of model performance testing using ROUGE resulted in the highest recall, precision, and f-measure of 1, 0.84, and 0.91. The highest coherence value is found in the 20th topic, with a value of 0.318. Using the Support Vector Machine and Latent Dirichlet Allocation algorithms are considered adequate for sentiment analysis and topic modeling for the Ruangguru application.

1. Introduction

The rapid development of technology, especially the Internet, opens opportunities to develop information services in educational institutions [1] —the Internet benefits every area of business, academia (education), government, and organizations. The Indonesian Internet Service Providers Association (APJII) grew internet use from 143.26 million to 171.17 million. Thus, of the total population of Indonesia, 64.8% has been connected to the Internet [2].

The electronic learning medium, as a controller, is operated by the user so that he can control and access needs. The study "Interactions in online learning: Important factors and research evidence from the literature" shows that most factors proved that traditional parameters apply to the environment. Online learning schools are viewed from the subject [3]. Self-study and interactive, providing experiences, sending emails, and commenting on discussion forums [4].

Ruangguru is a non-formal education mobile application in Indonesia founded in 2014 by Belva Devara and Isman. In 2018, the Ruangguru application was added to the company's learning by providing an application-based platform to conduct online training [5]. The presence of exciting features in the Ruangguru application makes students eager to learn and willing to accept lessons. This finding is consistent with research indicating that the features of the Ruangguru online tutoring application can provide application users with satisfaction [6].

The release of the Ruangguru application on the Google Play Store (GPS) will help application developers get feedback from their application users through reviews. Users' sentiments contain information that is invaluable for analysis [7]. Manually extracting information from a set of reviews is difficult due to the large number of reviews and diverse review topics, necessitating the use of a fast and precise data analysis model. Sentiment analysis and topic modeling are two data analysis models that can be used to help find hidden information in a set of reviews [8].

Sentiment analysis is a subset of text mining that aims to classify text documents as opinions in order to generate sentiment information that can have both positive and negative implications [9]. The sentiment analysis method can be used to examine a person's feelings and opinions in response to a specific topic [10], for example, public sentiment toward the Ruangguru review, so that user satisfaction can be determined. Owners and stakeholders of the application can use the sentiment analysis results to help them make decisions about the future development and management of Ruangguru.

According to the research, many researchers use classification to explore text data in order to produce critical information, such as research [11] using Nave Bayes and Latent Dirichlet Allocation for sentiment analysis and modeling of Lombok tourism topics The purpose of this study is to divide tourist opinions into positive and negative categories and to conduct topic modeling in both categories. Topic modeling seeks to identify the topics that are frequently discussed in each class. This study's stages include data collection, data cleaning, data transformation, data classification using the Naive Bayes method, and topic modeling using the Latent Dirichlet Allocation (LDA) method. The accuracy, precision, recall, and specificity of the model performance testing using the Naive Bayes algorithm are 92%, 100%, 83.84%, and 100%, respectively.

Another study by [12] used the Long Short-term Memory (LSTM) method and a pre-trained global vector for word representation in Indonesian obtained from fast text, a model with the highest accuracy value of 71.13% and the lowest accuracy value of 63.48% was produced. The Latent Dirichlet Allocation (LDA) method was used in this study to generate topics discussed by online media, whether positive, negative, or neutral. The topics derived from sentiment analysis results on Indonesian online news headlines can be well interpreted. Topics with insights from a news issue can be obtained from the visualization results presented.

Then [13] conducted sentiment analysis using the Support Vector Machine (SVM) method as a classification method, Term Frequency-Inverse Document Frequency (TF-IDF) as a character weighting method in analyzing sentiment based on user reviews of the PPI Dkemendagri application on the Google Play Store was able to be applied well in carrying out sentiment analysis so that application owners could use it to make decisions in the future. The data used is comment data derived from reviews of the PPIDKemendagri application on the Google Play Store, with as many as 700 data with 85 positive and 615 negative labels. In the evaluation stage of sentiment analysis, the average k-fold yielded 88%, precision 94%, recall100%, f-measure 97%, and accuracy 97%.

In this study, we will use the Support Vector Machine method to conduct sentiment analysis on the Ruangguru user reviews. This method has several advantages, including excellent time series prediction performance. The SVM algorithm is based on statistical learning theory and produces promising results that are superior to other methods. Using the [new] kernel technique, SVM also performs well on high-dimensional data. Furthermore, using the Latent Dirichlet Allocation (LDA) method, this study focuses on identifying the most discussed topics in each class, namely positive and negative. LDA is used to determine several topics on which each class has an opinion. The LDA method has the advantage of accurately extracting topics from a relatively large data set [8].

This study has two goals: first, conduct sentiment analysis on Twitter with the topic of Ruangguru user reviews into two classes, positive and negative, using the Support Vector Machine method, and then measure the model's performance results using a confusion matrix based on four criteria, namely accuracy, precision, recall, and f-measure.The second uses Latent Dirichlet Allocation (LDA) to perform topic modeling on each of the positive and negative classes to identify the main topics frequently discussed in both classes, and then measures the performance results of the LDA model based on coherence values. And use the Rouge method for sentiment testing and topic modeling with three criteria: precision, recall, and f-measure. The findings of this study will be used to help Ruangguru owners and stakeholders make policies and decisions about the development of Ruangguru that are relevant to user needs.

2. Method

This has the advantage of forming evidence more easily ahead of understanding by following an organized explanatory effort of the perception of examples followed by complex evidence [14]. The figure of the reservoir marginalizes the analysis efforts for the KDD system. The following figure is flow of this research.

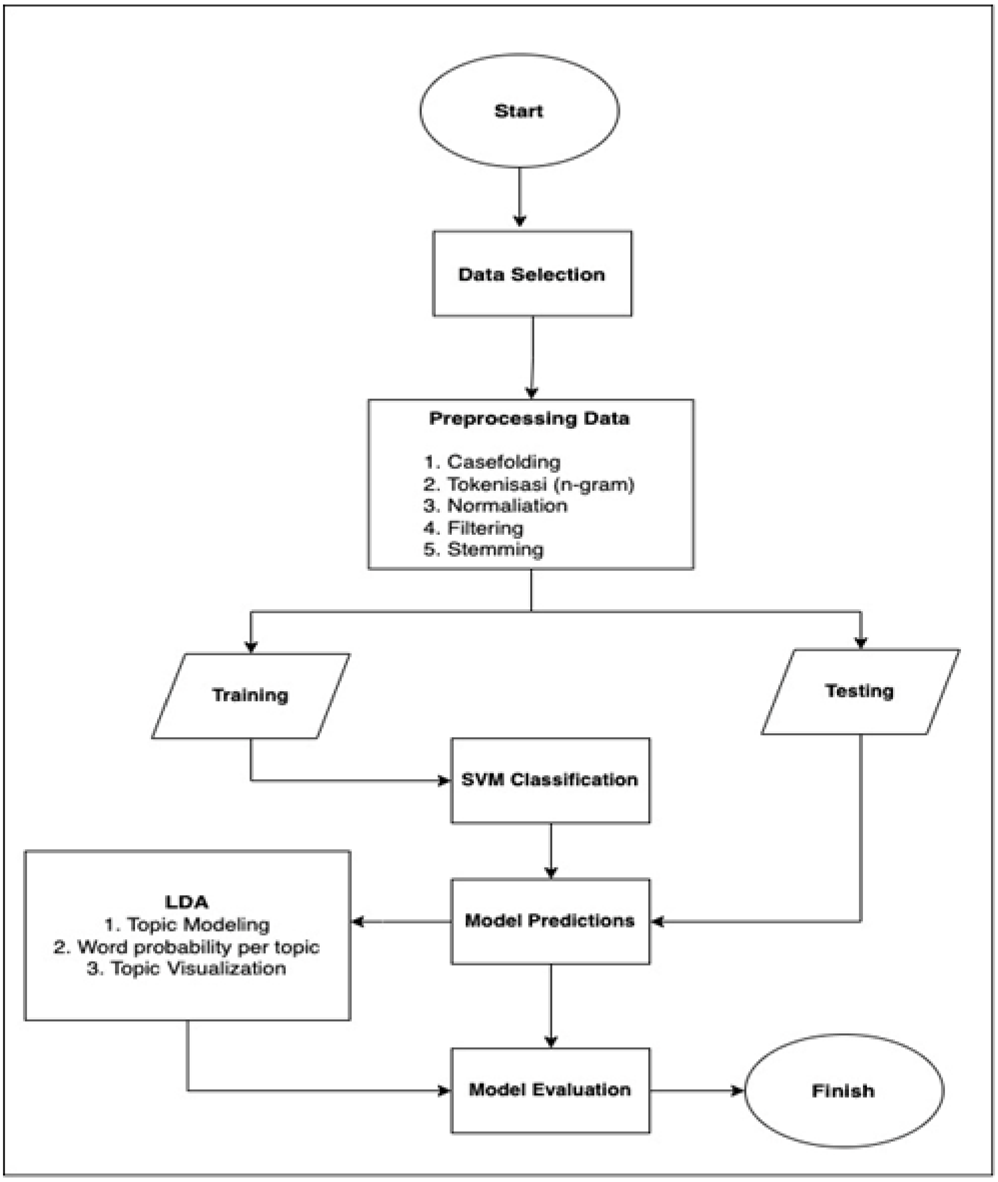

Figure 1. Research Flow

The research flow in Figure 1 is explained as follows:

2.1. Data Selection

The first stage of KDD is data selection, in which data is collected, selected, and labeled. The data collected was 1277 response data from the Google Play Store (GPS) page with the name Ruangguru application using Web Scraping (WS). A technique for extracting data from the World Wide Web (WWW) and storing it in a file system or database for data analysis is known as WS or Web Extraction. The user can run WS manually or through bots or web crawlers.

The WS Internet procedure is divided into two steps: retrieving web resources and extracting the desired information from the captured data. Specifically, the WS program sends an HTTP request to retrieve the website's resources. This request can be formatted as a URL with a GET request or as an HTTP request with a POST request. The requested resource is retrieved from the webpage and returned to the WS program after the request is successfully received and processed by the target website.

Delete HTML, XML, JSON, and multimedia data such as images, audio, and video from web pages. These resources may be available in a variety of formats. WS is comprised of two major modules. One is for writing HTTP requests, like Urllib2 and Selenium, and the other is for parsing and extracting information from raw HTML, like BeautifulSup and Pyquery [14].

2.2. *Data Preprocessing*

Preprocessing is the first step in converting input data into an appropriate format and processing it. This preprocessing transforms the raw data collected into data that can be used later. Various necessary processes during preprocessing include merging, reshaping data, cleaning, merging, minifying, and converting customers. Furthermore, existing preprocessing techniques may be composed of single-process operations or combinations of the aforementioned. The methods available are determined by the preprocessing goals. The preprocessing stage includes the following five functions: (1) Case folding is a preprocessing technique that converts uppercase letters to lowercase letters in order to make them more structured and easier to set up. (2) Tokenization or encryption aims to remove punctuation marks, tags, and emojis from sentences and break them down into distinct words. The encoding stage is customized to your processing requirements. (3) Normalization to correct spelling (typos) or descriptions of non-standard words. (4) Filtering is the process of selecting important words from segmentation results and excluding comments such as except, but, and possibly. (5) Stemming is the process of returning a word to its basic (original) form.

2.3. *Vector Machine Support Classification*

The Support Vector Machine (SVM) classification is the study's fourth stage. SVM is one of many algorithms used for classification and inclusion in the category of supervised learning. SVM's working principle is to find the best hyperplane or boundary line to separate the two classes [13]. Figure 2 depicts the Support Vector Machine (SVM) algorithm's operation. At the data mining stage, the SVM algorithm will be used to perform sentiment classification on the review data.

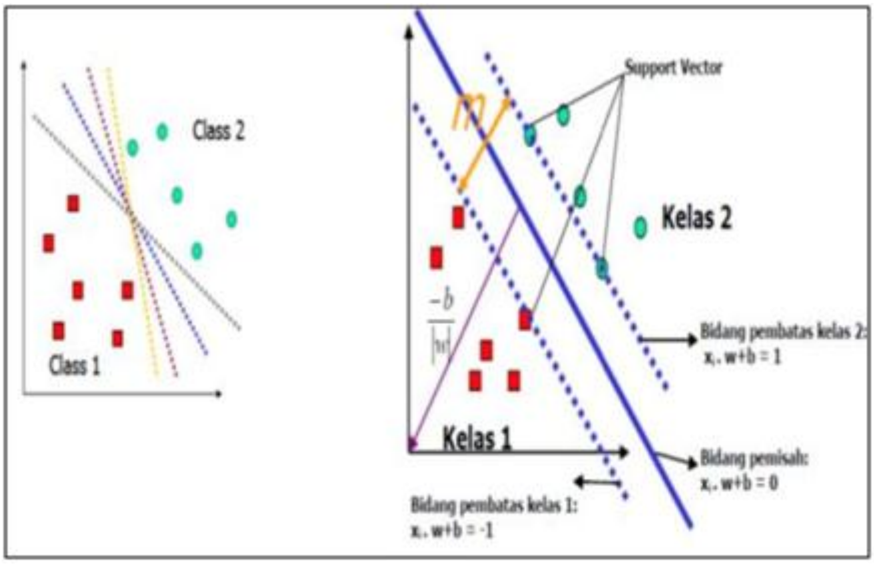

Figure 2. SVM Algorithm's Operation

The hyperplane limit calculation finds the maximum point to get the best hyperplane line when separating data into two classes. Equation (1) can be used to obtain a hyperplane on an SVM for classified data.

$$(w.\,xi) + b = 0 \;.......(1)$$

The data xi, which belongs to class -1, can be formulated in equation (2).

$$(w.\, xi + b) \leq 1,\ yi = -1 \text{ ....(2)}$$

While the data xi, which belongs to the +1 class, can be formulated as in equation (3).

$$(w.\, xi + b) \geq 1,\ yi = 1 \text{ ....(3)}$$

2.4. *Topic Modeling Process*

Machine learning methods are evolving at a rapid pace these days. One of them is a text analysis method. Text analysis is used to infer opinions or arguments from different parties based on a collection of words, such as those found on social media. This method allows us to analyze groups of words in large-sized documents that cannot be read in their entirety. The most straightforward method is to compute the frequency of occurrence of words in the data.The various words in the document are listed one by one, and then it is determined which words appear the most frequently in the data. The word that appears the most often indicates that it is the most frequently discussed in such a collection of texts and can be considered the main topic of conversation. However, because the resulting words are independent and cannot be attributed to other resulting words, one or more of them do not contain complete meaning and information [13].

The Latent Dirichlet Allocation is a popular topic modeling technique (LDA). The LDA's basic principle is that each document is represented as a mix of hidden and unknown topics, with each topic consisting of a word distribution [15]. To determine the interpretation of the main topics frequently discussed by users of the Ruangguru application, each positive and negative sentiment will be processed using the LDA method. Blei depicts the LDA method as a probabilistic model in Figure 3 [16].

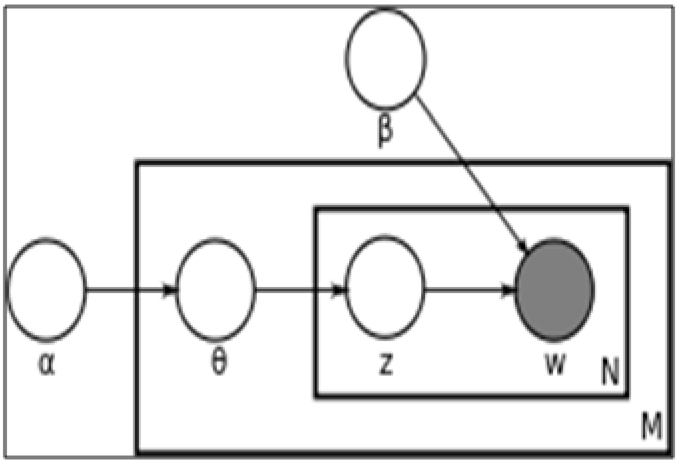

Figure 3. LDA Model Representation

Based on the illustrations in figure 3, the LDA can be explicitly described by using mathematical notations seen in equation 4 [16].

$$p(\beta_{1:k}, \theta 1 : D, z_{1:D}, w_{1:D}) = \prod_{i=1}^{K} p(\beta_i) \prod_{d=1}^{D} p(\theta_i) (\prod_{n=1}^{N} p(z_{d.n}|\theta_d) p(w_{d,n}|\beta_{1:K}, z_{d,n})) \text{ ....(4)}$$

N is the number of words in the document (Nd), Wd,n is the nth word on document d, Zd,n is the nth topic on document d, θd is the number of topics per document identified, βk is the distribution of topics in vocabulary as well as α, ɳ is the parameter of Dirichlet [16].

In general, how the LDA method works in this study is [17]: (1) Create a dictionary and corpus of a collection of positive and negative sentiments. (2) Initialize parameters, namely the number of documents, the number of topics, the number of iterations, the random state, the alpha value, the beta value, and others. (3) Determining the words for a particular topic based on the distribution of Dirichlet. (4) Display word probabilities per topic. (5) Repeat the grooves b through d for all words in the corpus.

2.5. *Testing*

2.5.1. Confusion matrix

The confusion matrix method was used to test the support vector machine algorithm, with measurements of accuracy, precision, recall, and f-measure shown in equations 5, 6, 7, and 8.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \text{.......(5)}$$

$$\text{Precision} = \frac{TP}{TP+FP} \text{.......(6)}$$

$$\text{Recall} = \frac{TP}{TP+FN} \text{.......(7)}$$

$$\text{F-Measure} = \frac{TN}{TN+FP} \text{.......(8)}$$

Information:

TP (True Positive) = Correctly classified positive data.

TN (True Negative) = Correctly classified negative data.

FP (False Positive) = Negative data is classified as positive.

FN (False Negative) = Positive data is classified as negative.

2.5.2. Rouge

ROUGE stands for Recall Oriented Understudy for Gisting Evaluation, and it is a testing procedure that involves calculating the same number of n-grams between a system summary and a manual summary. The ROUGE-1 test was used in this study because it had a high recall significant test for testing commentary labels with topic modeling. ROUGE-1 is supposed to count the number of unigrams of words that exist between a system summary and a manual summary. The work evaluation parameters are Recall, Precision, and F-Measure, which are calculated using equations 9, 10, and 11[18].

$$Rouge\ Recall = \frac{Common\ N-grams\ (Peer,Reference)}{N-grams(Reference)} \text{.......(9)}$$

$$Rouge\ Precision = \frac{Common\ N-grams\ (Peer,Reference)}{N-grams\ (Peer)} \text{.......(10)}$$

$$Rouge\ F\text{-}Measure = \frac{2*Precision*Recall}{Precision+Recall} \text{.......(11)}$$

In equation (9), equation (10), and equation (11), there are three essential components, namely Common N-grams (Peer, Reference) is the number of n-grams in the system summary and manual summary. N-grams(reference) is the number of n-grams in the manual summary. N-grams(peer) is the number of n-grams in the system summary [19].

3. Results and Discussion

3.1. *Data Selection*

Data retrieval for the validation of the Ruangguru application using scraping techniques and the Google Play Scraper library yielded 1277 response data. Review, userName, user image, content, at, content, score, thumbsUpCount, reviewCreateVersion, response content, and RespondAt are the five properties of the data collected. Three annotators then manually labeled the images to be divided into two classes, positive and negative, by writing down the label and comment attributes.The label selection procedure is carried out by selecting the majority label provided by three annotators, as shown in figures 4 and 5, which show the plot bar of the data label. The data labeling performed by the annotators resulted in 790 positive and 487 negative reviews.

	label	komentar
0	Negative	aplikasi ruang guru sangat bagus lebih bagus k...
1	Positive	ruang jadi percaya diri giat ajar ajar langsun...
2	Negative	mau ajar buka video lancar padahal jaring h ag...
3	Positive	wah aplikasi keren tidak sesal deh download ap...
4	Positive	semenjak pakai aplikasi ruang tidak perlu taku...
5	Negative	apps update jadi henti terus vidionya pakai wi...
...	...	...
1270	Positive	sobat semua nama malin april safir kasu bintan...
1271	Positive	anak malas ajar lewat buku tebal tebal tiap ma...
1272	Positive	bimbing ajar online sangat ikut kembang zaman ...
1273	Negative	ngabisin kuota walaupun di download tetap buka...
1274	Positive	kimah kasih ruang nilai ku sekarang naik drast...
1275	Negative	beli biasa erti kalau ruang guru ingin ajar ku...
1276	Positive	suka banget apak bisa ajar mana dan kapan ajar...

Figure 4. Data Selection Results

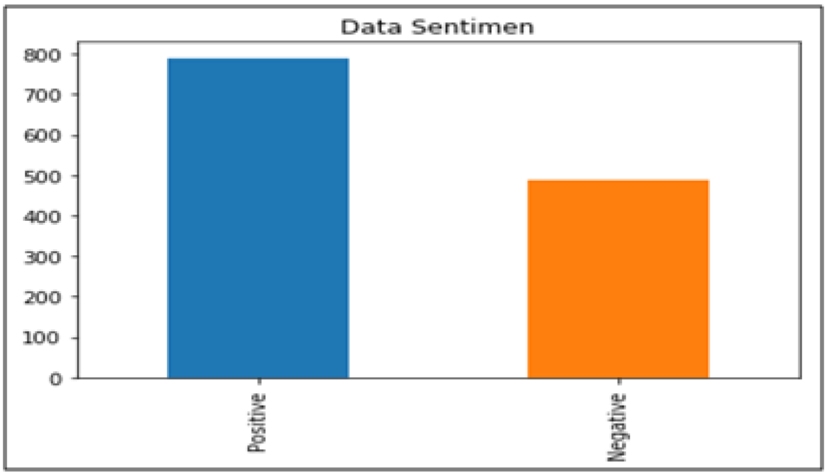

Figure 5. Plot Bar Label Data

3.2. *Preprocessing*

During this phase, the data is adjusted and modified so that it can be processed at a later stage. Table 1 shows an example of the results of the five preprocessing stages: Case Folding, Tokenizing, Normalization, Filtering, and Stemming.

Table 1. Preprocessing Results [18]

Process	Result
Data Collection	The application is good, useful for my work now
Case Folding	The application is good, useful for my work now
Tokenizing	'app', 'good', 'useful', 'for', 'work', 'me', 'now'
Normalization	'app', good, 'useful', 'for', 'work', 'me', 'now'
Filtering	'app', good, 'useful', 'work'
Stemming	'app', good, 'useful', 'work'

3.3. *Support Vector Machine classification*

Two classes determine the classification process. That is, positive and negative with multi-class SVM one vs. one. The dataset is then divided into training sets and test sets using a ratio of 0.1. This means that 90% of the data will be used in the training process and 10% in the testing process. Training data is used to build models using the SVM algorithm. Data testing evaluates the model's performance obtained through the training process. The training begins with calculating the initial value of each class (positive and negative), which is continued by calculating the term n odds on a document.Then proceed to calculate the chances of a document entering a class, and the last stage is to determine the class of the document by choosing the highest probability value. The SVM algorithm has core functions and parameters that can be used during the classification process. You can assign values to the parameters used, and different parameter values can affect your model's performance. Therefore, a testing process is carried out by parameter-adjusting kernel functions to create machine-learning models with optimal performance.

3.4. *Topic Modeling with LDA*

Sentiment result data will be further explored using the Latent Dirichlet Allocation (LDA) method. Topic modeling is carried out for each class of topics in this study to extract information from the collection of opinions of users of the application. The information will be interpreted as a collection of central topics in positive and negative classes. The initial stage of topic modeling is to create a dictionary and corpus for positive and negative classes. The data is the same for the SVM classification of 1277 response data, 790 positive and 487 negative reviews.

An important stage in the process of modeling the topic is the formation of a dictionary and corpus for data on positive classes and negative classes. Furthermore, the process of forming an LDA model is carried out using the help of the gensim library. The first step is to set the value of the parameter to be used such as topic number, random state = 100, update every = 1, chunk size = 100, passes= 10,alpha=auto. This study will take a range of 1 to 20 topics to be tested to find the best group of topics. Topics resulting from topic modeling are relatively easy to interpret. Therefore, a calculation of the coherence of topics is carried out to distinguish which topics are good and bad. Indicators of a topic are said to be good based on the ease with which words are interpreted semantically, while a topic is said to be bad if the words in the topic are difficult to interpret.

The results of calculating the coherence value for the topic are shown in table 2. The higher the coherence value of a topic, the easier it is to interpret its meaning based on the set of words that compose it. In other words, the more often the words in the topic appear simultaneously, the higher the coherence value of the topic. Based on table 2, 20 topics can be interpreted. The results of topic interpretation are used to see the comment trends of Ruangguru application users.

Table 2. The Value of Coherence Topics

Topic	Coherence Values Results
0	'0.017*"terima_kasih" + 0.012*"mudah_paham"+0.010*"tanya" +"0.010*"sangat_bantu"+0.009*"mudah_erti"+0.008*"sekarang" +"0.008*"bisa_buka"+0.008*"download"+0.008*"bagus_banget" + "0.008*"bayar"'
1	'0.042*"sangat_bantu" + 0.025*"terima_kasih" + 0.017*"bantu" +0.011*"nilai"+ 0.009*"di" + 0.009*"baik" + 0.009*"banyak" + 0.008*"lebih" + 0.008*"kasih" + 0.007*"bibel"'
2	'0.052*"sangat_bantu" + 0.018*"terima_kasih" + 0.016*"bantu" + ' '0.015*"orang_tua" + 0.015*"indonesia" + 0.012*"didik" + 0.010*"orang" + "0.010*"sekarang" + 0.009*"soal" + 0.009*"ku"'
3	'0.017*"rata_rata" + 0.016*"mau_tanya" + 0.015*"mudah" + 0.014*"di" + ' '0.014*"sangat_manfaat" + 0.012*"suka" + 0.011*"rata" + 0.010*"sama" + ' '0.010*"bisa_buka" + 0.009*"nilai"'
4	'0.018*"latih_soal" + 0.016*"sangat_bantu" + 0.012*"bagus_banget" + ' '0.011*"buka" + 0.009*"video" + 0.009*"sangat_manfaat" + 0.009*"mana" + "0.008*"soal" + 0.008*"yang" + 0.008*"lebih"'
5	'0.013*"kuota" + 0.013*"mudah_paham" + 0.012*"terima_kasih" + 0.011*"punya" ' '+ 0.010*"sama" + 0.010*"pakai" + 0.010*"video" + 0.010*"makin" + 0.009*"g" "+ 0.009*"bagus"'
6	'0.013*"terima_kasih" + 0.011*"bagus" + 0.010*"tidak" + 0.010*"soal" + ' '0.009*"nilai" + 0.009*"sejak" + 0.008*"dapat" + 0.008*"sangat_bantu" + ' '0.008*"asyik" + 0.007*"video"'
7	'0.035*"sangat_bantu" + 0.028*"terima_kasih" + 0.018*"bagus_banget" + ' '0.013*"bagus" + 0.013*"bantu" + 0.012*"saya" + 0.011*"lancar" + ' '0.010*"video" + 0.009*"buat" + 0.009*"padahal"'
8	'0.022*"terima_kasih" + 0.020*"bisa_buka" + 0.018*"mudah_erti" + ' '0.014*"bagus" + 0.013*"bagus_banget" + 0.012*"nilai" + 0.011*"lebih" + ' '0.011*"bisa" + 0.011*"sekali" + 0.010*"sekarang"'
9	'0.015*"terima_kasih" + 0.012*"kakak" + 0.009*"sangat_bantu" + 0.009*"punya" ' '+ 0.008*"buat" + 0.008*"waktu" + 0.008*"anak" + 0.008*"lebih" + ' '0.008*"paham" + 0.007*"kalau"'
10	'0.020*"kasih_bintang" + 0.018*"sangat_bantu" + 0.013*"soal" + 0.012*"baca" ' '+ 0.009*"bisa_buka" + 0.009*"buka" + 0.008*"bagus_banget" + 0.008*"video" + ' '0.008*"kakak" + 0.008*"bisa"'
11	'0.014*"latih_soal" + 0.013*"terima_kasih" + 0.011*"bisa_buka" + ' '0.010*"sangat_manfaat" + 0.009*"kakak" + 0.009*"wah" + 0.009*"bagus" + "0.009*"tambah" + 0.008*"kalau" + 0.008*"langgan"'
12	'0.027*"terima_kasih" + 0.014*"materi" + 0.011*"bagus_banget" + ' '0.010*"banget" + 0.010*"kasih" + 0.009*"untuk" + 0.008*"teman_teman" + "0.008*"soal" + 0.008*"terus" + 0.008*"tambah"'
13	'0.015*"terima_kasih" + 0.011*"sekali" + 0.011*"tidak" + 0.009*"banget" + "0.009*"video" + 0.009*"nilai" + 0.009*"sangat_bantu" + 0.009*"kelas" + ' '0.008*"ku" + 0.008*"sama"'
14	'0.033*"terima_kasih" + 0.015*"sangat_bantu" + 0.012*"anak_sk" + ' '0.012*"masuk" + 0.012*"kasih" + 0.011*"latih_soal" + 0.010*"banyak" + ' '0.009*"buat" + 0.009*"bagus" + 0.009*"nilai"'
15	'0.033*"sangat_bantu" + 0.021*"mudah_paham" + 0.015*"terima_kasih" + ' '0.010*"bantu" + 0.009*"materi" + 0.009*"ada" + 0.009*"mudah" + ' '0.009*"bagus" + 0.008*"paham" + 0.008*"pakai_kode"'
16	'0.018*"sangat_bantu" + 0.014*"terima_kasih" + 0.011*"baik" + 0.010*"saya" + "0.010*"video" + 0.010*"nilai" + 0.009*"cara" + 0.009*"kakak" + "0.008*"lebih" + 0.008*"tarik"'
17	'0.023*"terima_kasih" + 0.023*"sangat_bantu" + 0.015*"nilai" + 0.013*"anak" "+ 0.012*"bagus_banget" + 0.012*"bagus" + 0.011*"bantu" + 0.010*"sekarang" + "0.009*"paham" + 0.009*"suka"'
18	'0.017*"bagus_banget" + 0.016*"paham_materi" + 0.014*"tank_yu" + '

Topic	Coherence Values Results
	'0.012*"teman_teman" + 0.010*"kalau" + 0.010*"terima_kasih" + ' '0.010*"bank_soal" + 0.009*"sangat_bantu" + 0.009*"lebih" + 0.009*"paham"'
19	'0.013*"sangat_bantu" + 0.011*"anak" + 0.011*"mudah_erti" + ' '0.010*"mudah_paham" + 0.010*"buat" + 0.010*"sekali" + 0.010*"langgan" + "0.009*"orang_tua" + 0.009*"suka" + 0.009*"kasih_bintang"'

Visualization of the modeling of the topic in this study's form of workload is shown in Figure 6. Figure 6 shows a visualization based on the topic distribution shown by the topic indicating that the topic has several of the exact constituent words.

Figure 6. Word Cloud

3.5. *Evaluation*

3.5.1. *Confusion Matrix*

Input values for each parameter used, such as complexity and gamma, are used to run tests. The procedure is the RBF kernel function of the Support Vector Machine (SVM) algorithm and is carried out in accordance with the parameters. The Taguchi method determined the optimal parameter value for each parameter as(1,1.25,1.50,1.75,2) C =andγ= (0.2, 0.4,0.6,0.8). Following the feature selection stage from the dataset, the feature selection results are evaluated using various combinations of parameter values, namely SVM classification with parameterscomplexity(C) andgamma(γ). The optimal parameter values for each parameter include (1,1,25,1,50,1,75,2) C and (0,2,0.4,0.6,0.8). Table 3 displays the evaluation results based oncomplexityparameter (C) andgamma (γ).

The test is performed by providing input values for each parameter, such as complexity and gamma. The process is an RBF kernel function of the Support Vector Machine (SVM) algorithm and is executed according to the parameters. The values of each parameter specified using the Taguchi method as optimal parameter values include (1,1.25,1.50,1.75,2) C = andγ=(0.2, 0.4,0.6,0.8). The feature selection stage of the dataset is carried out, followed by the evaluation stage of the feature selection results using various combinations of parameter values, namely the SVM classification with complexity (C) and gamma (γ) parameters. The values of each parameter set as the optimal parameter value include (1,1,25,1,50,1,75,2) C and (0,2,0.4,0.6,0.8). Table 3 shows the following evaluation results based on complexity parameters (C) and gamma (γ).

Table 3. Best Accuracy Based on Complexity (C) and Gamma (γ) Parameters

Runs	C	γ	*Average K-Fold*	*Accuracy*	*precision*	*recall*	*f-measure*
1	1	0.2	0,88	0,5	0,5	1	0,67
2	1	0.4	0,88	0,5	0,5	1	0,67
3	1	0.6	0,88	0,5	0,5	1	0,67
4	1	0.8	0,88	0,62	0,57	1	0,73
5	1,25	0.2	0,88	0,5	0,5	1	0,67
6	1,25	0.4	0,88	0,5	0,5	1	0,67
7	1,25	0.6	0,88	0,69	0,62	1	0,76
8	1,25	0.8	0,88	0,84	0,5	1	0,67
9	1,50	0.2	0,88	0,5	0,5	1	0,67
10	1,50	0.4	0,88	0,62	0,57	1	0,73
11	1,50	0.6	0,88	0,84	0,76	1	0,86
12	1,50	0.8	0,88	0,97	0,94	1	0,97
13	1,75	0.2	0,88	0,50	0,50	1	0,67
14	1,75	0.4	0,88	0,72	0,64	1	0,78
15	1,75	0.6	0,88	0,94	0,89	1	0,94
16	1,75	0.8	0,88	0,97	0,94	1	0,97
17	2	0.2	0,88	0,50	0,50	1	0,67
18	2	0.4	0,88	0,75	0,67	1	0,80
19	2	0.6	0,88	0,97	0,94	1	0,97
20	2	0.8	0,87	0,97	0,94	1	0,97

3.5.2. Rouge

To determine whether a comment fits into a specific topic, a theme requires topic-related initialization. The researcher created four topics in this study: the first is about application access or service, the second is about the impact or feedback on application use, the third is about application features or modules, and the fourth is about the application learning process. The definition of topic conversion is shown in Table 4.

Table 4. Topic Conversion

Code	**Name**
A	access or service application
B	the impact or feedback of using the application
C	app features or modules
D	application learning process

Figure 7 depicts many data predictions where we evaluate our test data after initializing the topic. The first attributes will be kode_topik, then kode_asli, and finally text. The topic code indicates that a comment belongs to a specific topic.

	kode_topik	kode_asli	text
0	C	15	aplikasi ruang guru sangat bagus lebih bagus k...
1	A	18	ruang jadi percaya diri giat ajar ajar langsun...
2	C	15	mau ajar buka video lancar padahal jaring h ag...
3	C	13	wah aplikasi keren tidak sesal deh download ap...
4	B	1	semenjak pakai aplikasi ruang tidak perlu taku...
5	C	13	apps update jadi henti terus vidionya pakai wi...
...	...	...	...
1273	C	6	ngabisin kuota walaupun di download tetap buka...
1274	B	1	kimah kasih ruang nilai ku sekarang naik drast...
1275	B	17	beli biasa erti kalau ruang guru ingin ajar ku...
1276	B	0	suka banget apak bisa ajar mana dan kapan ajar...

Figure 7. Predict a Lot of Data

Table 5 shows the unit's prediction with the comment, "This apk is fantastic. can study at any time and from any location Even though I am in high school, the lessons are the same. Thank you very much, Ruangguru "Following the prediction process, it was discovered that the comment was included in topic B, namely the impact or feedback on the application's use. The rouge method was evaluated based on the training results, and optimal results were obtained by obtaining values. Precision is 1.0, recall is 0.84, and f-measure is 0.91.The dominance of the approximation feature between sentences indicates the best rouge method score calculation. This feature has a significant impact on topic modeling on user evaluation of the online tutoring application for the Ruangguru topic application.

Table 5. Unit Data Prediction

no	Topic_code	Ori_code	text
126	B	0	suka banget dengan apk ini. bisa belajar di mana saja dan kapan saja. pelajarannya juga sama padahal saya smk. terimaaksih Ruangguru

4. Conclusion

For the topic of Ruangguru user reviews, this study was successful in conducting sentiment analysis using the Support Vector Machine method and topic modeling using the Latent Dirichlet Allocation (LDA) method. Based on the sentiment analysis results, 1277 reviews were obtained, with 790 positive reviews and 487 negative reviews, implying that more users gave positive responses than negative responses. According to the results of model performance testing using the Confusion Matrix, the Support Vector Machine algorithm was able to classify the sentiment of Ruangguru user reviews well, with values of accuracy, precision, recall, and f-measure of 0.9, 0.9, 0.9, and 0.89, respectively.Meanwhile, topic modeling with the LDA algorithm yields the best topic with the highest coherence value, 0.318, in the 20th topic. And the results of ROUGE model performance testing yielded the highest recall, precision, and f-measure values of 1, 0.84, and 0.91, respectively. For sentiment analysis and topic modeling in the Ruangguru application, the Support Vector Machine and Latent Dirichlet Allocation algorithms are deemed adequate.

References

[1] Pujilestari, Yulita. (2020). Dampak Positif Pembelajaran Online Dalam Sistem Pendidikan Indonesia Pasca Pandemi

Covid-19. Adalah: Buletin Hukum & Keadilan, 1(1), 49-56.
[2] Apjii.or.id. (2018). Asosiasi Penyelenggara Jasa Internet Indonesia. Rectrieved Juni 29, 2020. https//apjii.or.od/content/read/104/348/BULETIN-APJIIEDISI-22- Maret-2018.
[3] Yang, D., Lavonen, J. M., & Niemi, H. (2018). Online learning engagement: Critical factors and research evidence from literature. Themes in ELearning, 11(1), 1–18.
[4] Atsani, KH. L. G. M. Z. (2020). Transformasi Media Pembelajaran Pada Masa Pandemi COVID-19. Jurnal Studi Islam, 1(1), 82-93.
[5] Iman, Usman. 2019. Mulai Belajar. Jakarta: Gramedia Pustaka Utama.
[6] Shofi, S. A., Rachmadi, A., & Herlambang, A. D. (2019). Analisis Kebutuhan Pengguna Aplikasi Ruangguru Menggunakan Metode Fuzzy Kano. Jurnal Pengembangan Teknolohi Informasi dan Ilmu Komputer, 3(5), 4307-4315.
[7] R. Ferdiana, F. Jatmiko, D. D. Purwanti, A. S. T. Ayu, and W. F. Dicka, “Dataset Indonesia untuk Analisis Sentimen,” J. Nas. Tek. Elektro dan Teknol. Inf., vol. 8, no. 4, pp. 334–339, 2019.
[8] A. Alamsyah, W. Rizkika, D. D. A. Nugroho, F. Renaldi, and S. Saadah, “Dynamic large scale data on Twitter using sentiment analysis and topic modeling case study: Uber,” in 2018 6th International Conference on Information and Communication Technology, ICoICT 2018, 2018, vol. 0, no. c, pp. 254–258.
[9] R. Ardianto, T. Rivanie, Y. Alkhalifi, F. S. Nugraha, and W. Gata, “Sentiment Analysis on E-Sports For Education Curriculum Using Naive Bayes and Support Vector Machine,” J. Comput. Sci. Inf., vol. 13, no. 2, pp. 109–122, 2020.
[10] M. Cendana and S. D. H. Permana, “Pra-Pemrosesan Teks Pada Grup Whatsapp Untuk Pemodelan Topik,” Junal Mantik Penusa, vol. 3, no. 3, pp. 107–116, 2019.
[11] N. L. P. M. Putu, Ahmad Zuli Amrullah, and Ismarmiaty, “Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation”, J. RESTI (Rekayasa Sist. Teknol. Inf.) , vol. 5, no. 1, pp. 123 - 131, Feb. 2021.
[12] Naury, Chairullah, Dhomas Hatta Fudholi, and Ahmad Fathan Hidayatullah. "Topic modelling pada sentimen terhadap headline berita online berbahasa indonesia menggunakan LDA dan LSTM." Jurnal Media Informatika Budidarma 5.1 (2021): 24-33.
[13] Ramos, S., Soares, J., Cembranel, S. S., Tavares, I., Foroozandeh, Z., Vale, Z., & Fernandes, R. (2021). Data mining techniques for electricity customer characterization. Procedia Computer Science, 186, 475 – 488. https://doi.org/10.1016/j.procs.2021.04.168.
[14] Fahlevvi, M.R., 2022. Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia di Google Playstore Menggunakan Metode Support Vector Machine. Jurnal Teknologi dan Komunikasi Pemerintahan, 4(1), pp.1-13.
[15] M. Cendana and S. D. H. Permana, “Pra-Pemrosesan Teks Pada Grup Whatsapp Untuk Pemodelan Topik,” Junal Mantik Penusa, vol. 3, no. 3, pp. 107–116, 2019.
[16] D. Blei, L. Carin, and D. Dunson, “Probabilistic topic models,” IEEE Signal Process. Mag., vol. 27, no. 6, pp. 55–65, 2010.
[17] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent Dirichlet Allocation,” J. Mach. Learn. Res., vol. 3, pp. 993–1022, 2003.
[18] B. Zhao, "Web Scraping," in Springer International Publishing AG, USA, 2017.
[19] Fahlevvi, M.R., 2019. Pemodelan Topik Pada Portal Berita Online Menggunakan Latent Dirichlet Allocation (LDA). Sleman: Universitas Gadjah Mada.