Sentiment Analysis Study Tour Bus Ban on Twitter Using Support Vector Machine Method

Ony Hizri Kaifa Purba[*], Ilka Zufria

Science and Technology, Computer Science, State Islamic University of North Sumatera, Medan, Indonesia
Email: [1,*]onyhizri12@gmail.com, [2]ilkazufria@uinsu.ac.id,
Correspondence Author Email: onyhizri12@gmail.com
Submitted: 01/08/2024; Accepted: 10/08/2024; Published: 10/08/2024

Abstract−Study tour is an activity outside the classroom that has the purpose of learning about the process of something directly. This activity is usually carried out by the school once a year. This activity is not only a learning tool for students, but also a recreational activity.In this activity, there are many things that need to be prepared, such as transportation, lodging, meals, and so on. This is sometimes troublesome, because not all tourists or business people have the time and willingness to prepare it. Therefore, they need services during their trip. Especially now that it is even semester, where every school usually holds a study tour, as well as a final class farewell. As a response to concerns, some parents may choose to find alternative activities that are considered safer for their children, such as joining activities in the city or at school. Based on this need, it makes opportunities for business people engaged in the tour agency industry. SVM (Support Vector Machine) is a machine learning method that works on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperlane separating two classes in the input space. Simply put, SVM (Support Vector Machine) has the concept of finding the best hyperlane, which serves as the boundary of two classes The results of sentiment classification on Study Tour Buses using the Support Vector Machine algorithm that matches the actual data amount to 176 data out of a total of 240 test data. It is known that of the 1200 data obtained regarding sentiment towards there are 519 reviews that are positive and 681 reviews that are negative.The accuracy value of the Study Tour Bus sentiment classification using the Support Vector Machine (SVM) algorithm obtained is 73%.

Keywords: Sentiment Analysis; Support Vector Machine; Twitter

1. INTRODUCTION

Study tour is an activity outside the classroom that has the purpose of learning about the process of something directly. This activity is usually carried out by the school once a year. This activity is not only a learning tool for students, but also a recreational activity.

In this activity, there are many things that need to be prepared, such as transportation, lodging, meals, and so on. This is sometimes troublesome, because not all tourists or business people have the time and willingness to prepare it. Therefore, they need services during their trip. Based on this need, it makes opportunities for business people engaged in the tour agency industry [1]. Especially now that it is even semester, where every school usually holds a study tour, as well as a final class farewell. As a response to concerns, some parents may choose to find alternative activities that are considered safer for their children, such as joining activities in the city or at school.

This study tour activity triggers cons for parents, especially in the middle to lower class. Because in addition to the cost, the risk of parental concern for children is also very large. For example, a long trip can make children tired or tired for a long time, and what is more dangerous is an accident. The study tour bus accident is a tragedy that not only affects the victims and their families, but also reflects the need for improved transportation safety standards for educational activities.[2]

Sentiment Analysis is the sentiment of the subjective text analyzing, processing, summarizing, and inferential processing. Sentiment analysis is currently divided into the use of machine learning classification and rule-based classification methods, machine learning methods use emotion words as feature classification, and emotion dictionaries can be used to realize the selection of sentiment characteristics quickly and efficiently. In combination with classification in solving the task, the commonly used methods are Naïve Bayes (NB), Support Vector Machine (SVM) and Maximum Entropy (EM) [4]

Machine Learning is the study of how computers can learn or improve performance based on data to automatically recognize data patterns to make intelligent decisions based on data [5].

Text mining is part of data mining, which is the process of gaining knowledge using a set of analytical tools where users interact with a set of documents over time. Like data mining, text mining seeks to extract information from a data source (a set of documents) through the identification and exploration of existing patterns [6]. Text mining is a technique used to handle classification, information extraction and information retrieval. The concept of text mining is used in the classification of textual documents with the aim of classifying documents that match the topic of discussion. The difference between data mining and text mining lies in preprocessing, in data mining preprocessing focuses on indexing and data normalization, while text mining focuses on identification and feature extraction [7].

Classification in this study uses the Support Vector Machine method with feature selection. Feature selection in classification is expected to be more efficient by reducing the amount of data analyzed by identifying features that will then be processed based on the classifier model that has been generated from the training

process in this study the data is divided into two parts, namely training data and testing data. According to Burges, the classification process is divided into two types based on the kernel used, namely SVM-Linear and Non-Linear SVM [8].

At this stage using data in the form of text data from the crawling process and will perform a data preprocessing process because the data still has an unstructured text form that has a lot of noise, so the data needs to be cleaned first. Preprocessing aims to improve the accuracy of the data. If the data used for the knowledge search process is of low quality, the resulting knowledge will be low as well.

SVM (Support Vector Machine) is a machine learning method that works on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperlane separating two classes in the input space. Simply put, SVM (Support Vector Machine) has the concept of finding the best hyperlane, which serves as the boundary of two classes [9].

Twitter is a popular social media platform in Indonesia, used by people to make friends, express feelings, and share information. Data shows that Indonesia is ranked as the first largest twitter user in Asia and sixth in the world. With more than 237,556,363 people in Indonesia, about 2.41% of the total population, recorded as active twitter users [10].

Python is increasingly popular among application developers (desktop, mobile and web), because it is easy to understand and use. Python is a high-level programming language known for its simple, easy-to-understand, and flexible syntax. Developed by Guido van Rossum [11].

Google Colab, which stands for Google Collaboratory, is a web-based cloud platform provided by Google that was launched in March 2017. Google Colab provides a Jupyter Notebook-based development environment that allows users to write and run Python code interactively. Google Colab is designed to provide an interactive, easy-to-use, and collaborative Python development environment. Its main goal is to facilitate development and experimentation in data science, machine learning, and deep learning. It allows users to develop, run, and share Python code interactively [12].

Previous research was conducted by Dian Siti Utami, Adhitia Erfina in 2021 entitled "Sentiment Analysis of Online Loans on Twitter Using the Support Vector Machine (SVM) Algorithm" where the results of the implementation of the Support Vector Machine (SVM) algorithm, in this study the data shows that the level of accuracy for online loans is 62.00% [13]. The accuracy results in the study are considered quite good. Sentiment analysis with the Support Vector Machine (SVM) algorithm successfully classifies community sentiment on Twitter about online loans. The classification results show that negative sentiment dominates over positive sentiment, there are 59% negative reviews and 41% positive reviews. It can be concluded that there are still many people who are hesitant about online loans because the risk is quite high. And the number of victims of online loans makes people afraid to try. on the other hand, some people think positively about online loans because they can help financial needs [14].

Based on this background, researchers want to examine this research with the Support Vector Machine method with the title "Sentiment Analysis Study Tour Bus Ban on Twitter Using the Support Vector Machine Method".

2. RESEARCH METHODOLOGY

2.1 Research Method and Materials

In this study, the authors used a qualitative approach to evaluate the highest level of accuracy and positive or negative sentiment on the study tour bus ban. The research framework contains an explanation of the stages of the process that will be carried out during research activities, the order of this research framework is an overview of the stages that will be carried out in solving the problems discussed.

This research process begins with planning, namely determining the topic to be discussed. The topic of this research is sentiment analysis of the study tour bus ban on twitter social media using the support vector machine method.

SVM (Support Vector Machine) is a machine learning method that works on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperlane separating two classes in the input space. Simply put, SVM (Support Vector Machine) has the concept of finding the best hyperlane, which serves as the boundary of two classes.

The data used in this research comes from public opinion related to the conflict over the study tour bus ban on the Twitter platform, which recently occurred and triggered many opinions related to study tours. The opinion data retrieval process is carried out using the Scraping technique, with the data retrieval stage using Python to retrieve 1200 datasets. The data will then go through a preprocessing process before further analysis.

Data analysis in this research means preparing tweet data as the initial process of data processing. This process is used to describe twitter users' tweets about the study tour bus ban. By displaying data, it can make it easier to research and organize data. The stages of conducting sentiment analysis are carried out using the SVM method, starting from entering training data and test data, which are then processed until the classification process in the preprocessing stage.
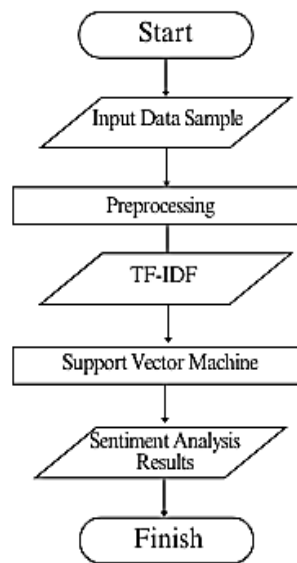
Figure 1. Flowchart of the Research

From the Figure 1, the implementation of this research begins with obtaining a Twitter dataset containing tweets related to the study tour bus ban. Make sure this dataset has been labeled with positive, or negative sentiment. Clean the data from unnecessary special characters, URLs, and punctuation marks. Perform tokenization to break each sentence into words [15]. Remove stop words (common words like "and", "or", "which", etc.). Perform stemming to convert the words into their base form. Use feature extraction methods such as TF-IDF (Term Frequency-Inverse Document Frequency) to convert the text into a numerical representation. Then calculate word weights to obtain opinion classification using the SVM method and classify the resulting dataset with a confusion matrix so that it becomes two classes, namely the negative class and the positive class [16].

3. RESULT AND DISCUSSION

3.1 Data Analysis

In May 2024, there have been 5 cases of student study tour bus accidents that claimed 15 lives. One of them was the accident of SMK Lingga Kencana students in Ciater, Subang on (11/5/2024) at around 18.00 WIB. News about the study tour bus accident has spread widely on various social media platforms such as Facebook, TikTok, Instagram, and Twitter. Social media is a means of information that is easily used by the public.

Previous accidents can lead to greater trauma and fear. As a result, many parents feel less trust in the organization of study tours, including schools or travel agents. Especially now that it is even semester, where every school usually holds a study tour, as well as the final class farewell. As a response to concerns, some parents may choose to find alternative activities that are considered safer for their children, such as participating in activities within the city or at school.

Therefore, researchers want to conduct research on public responses or responses to determine the level of sentiment towards the study tour bus ban with the Support Vector Machine (SVM) method with a total of 1200 data. Which data was taken in May-June 2024.

This research aims to analyze sentiment towards study tour buses on Twitter social media using the Support Vector Machine (SVM) method. The data taken are tweets from May 11, 2024 to May 30, 2024. The total data processed is 1200 sentiments, with 519 sentiments showing support for the bus study tour and 681 sentiments expressing disapproval of the policy. Prior to analysis, the crawled sentiments were labeled using a lexicon dictionary. The sentiment labeling process is done automatically by the system by utilizing the lexicon dictionary imported into the Python development environment.

Next, the sentiment data is cleaned and processed through pre-processing stages, including data cleaning, text normalization, tokenization, stopword removal, and stemming/ lemmatization. SVM algorithm is used to classify the sentiment towards the study tour bus. SVM models were trained using training data and evaluated using testing data. The evaluation metrics used include accuracy to measure the performance of the classification model [16].

The total data processed reached 179 sentiments, with 91 sentiments showing support for the bus study tour and 88 sentiments expressing disapproval of the policy.
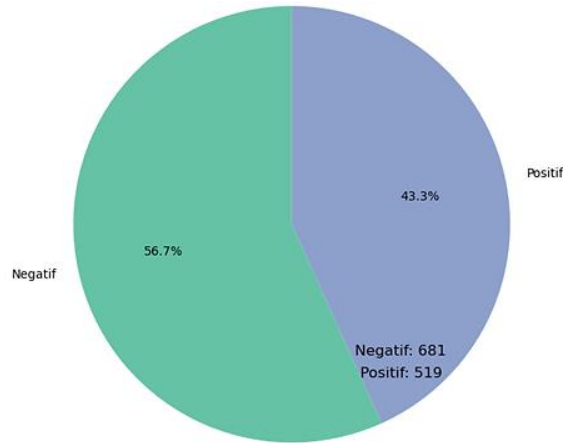
Figure 2. Sentiment Data Distribution

3.2 Text Pre-Processing

The sentiment analysis process of the study tour bus ban on Twitter social media begins with an important stage, namely data preprocessing. The data obtained from scraping Twitter is still in raw condition and contains a lot of noise that needs to be cleaned and structured so that it is ready for in-depth analysis. This preprocessing stage is important to ensure that the data used in the research is of good quality so that the analysis results obtained can be an accurate representation of public opinion regarding the study tour bus ban.

The following flow describes the system stages during the data preprocessing process which includes case folding, case folding, filtering, stemming and tokenizing.

Table 1. Preprocessing

Content	After Pre-Processing
@convomf 15 jam kalo ga salah dulu abis study tour dari bali habis utama	['jam', 'habis', 'study', 'tour', 'bal', 'habis', 'utama ']
@convomf Gue angkatan 21 yg batal pergi rekreasigaragara covid-19 :)	['angkat', 'batal', 'pergi', 'rekreasi', 'garagara', 'covid']
@convomf libur sekolah	['libur', 'sekolah']
@convomf terakhir hendak tamasya dari sekolah 2019 aku bawa 1 5 nder	['hendak', 'tamasya', 'sekolah', 'bawa', 'nder']
@convomf D kota gw ngga ada bali wisata	['kota', 'bal', 'wisata']

In the next stage, TF-IDF weighting. This technique assigns a value to each word, indicating how important it is in representing the meaning of the document. By understanding TF-IDF, we can identify essential key words and dive into the deepest meaning of each text. The following is a sample calculation of TF value and DF value from 5 pieces of training data and 2 test data. The labeling of the training data was done manually by taking into account the sentiment before the preprocessing stage. Where the full sentiment is analyzed using the researcher's reasoning then the class that corresponds to the sentiment is concluded [17].

Furthermore, the TF-IDF value is normalized to equalize the interval of each data, as for the equation used to normalize the data is as follows.

$$TF_{norm}(t,d) = \frac{TF(t,d)}{\sqrt{\sum_i (TF(t,d))^2}}$$

The following are the results of the data normalization calculations carried out :

Table 2. Data Normalization

No.	D1	D2	D3	D4	D5
1	0.206	0	0	0	0
2	0.412	0	0	0	0
3	0.206	0	0	0	0
4	0.206	0	0	0	0
5	0.206	0	0	0	0

3.3 Support Vector Machine Classification

After cleaning and structuring the sentiment data, the next step is to enter the classification stage. In this stage, the Support Vector Machine (SVM) algorithm acts as a tutor that will be trained to distinguish between positive and negative sentiments in the text.

The classification process begins by dividing the data into two parts: training data and testing data. In this research, the ratio used is 8:2, where the training data is the majority used to train the SVM. The training data serves as the subject matter for the SVM, where the algorithm learns the characteristics and patterns of the text that can be categorized as positive or negative sentiment. Meanwhile, the testing data, which is a small subset of the overall data, is used to test the SVM's ability to classify new text that it has never seen before. The purpose of this test is to assess the success rate of the SVM in performing the classification.

In the classification performed, the type of kernel used is a linear kernel because the data entered is linear data. The following is the equation used to calculate the linear kernel value.

Table 3. Sample Representation

	x1	x2	x3	x4	x5
y1	K(x1,y1)	K(x2,y1)	K(x3,y1)	K(x4,y1)	K(x5,y1)
y2	K(x1,y2)	K(x2,y2)	K(x3,y2)	K(x4,y2)	K(x5,y2)
y3	K(x1,y3)	K(x2,y3)	K(x3,y3)	K(x4,y3)	K(x5,y3)
y4	K(x1,y4)	K(x2,y4)	K(x3,y4)	K(x4,y4)	K(x5,y4)
y5	K(x1,y5)	K(x2,y5)	K(x3,y5)	K(x4,y5)	K(x5,y5)

Example calculation for column 1 row 1.

$$K(x,y) = x * y$$

$$K(x,y) = (t1d1 * t1d1 + t1d2 * t1d2 + t1d3 * t1d3 + t1d4 * t1d4 + t1d5 * t1d5)$$

$$K(x,y) = (0.206 * 0.206 + 0 * 0 + 0 * 0 + 0 * 0 + 0 * 0)$$

$$D11 = 0.042$$

The results of the calculation of the linear kernel from the sample data owned as presented in the table below.

Table 4. Linear Kernel Calculation

No.	1	2	3	..	18	19	20
1	0.042	0.085	0.042	..	0	0.034	0
2	0.085	0.169	0.085	..	0	0.068	0
3	0.042	0.085	0.042	..	0	0.034	0
4	0.042	0.085	0.042	..	0	0.034	0
5	0.042	0.085	0.042	..	0	0.034	0

After the kernel value is known, the next step is the calculation of the Hessian matrix. Before calculating the Hessian matrix, several parameters will be determined, including αi, C, γ, λ and maximum iteration. The following is a description of the parameters that will be used in the Hessian matrix calculation stage.

The calculation step of the Hessian matrix starts by initializing the value $\alpha = 0$ and then calculating using the following equation.

$$D_{ij} = y_i y_j (K(x_i, x_j) + \lambda^2$$

Example of calculating the value of the Hessian matrix in column 1 row 1

$$D_{11} = y_i y_j (K(x_i, x_j) + \lambda^2 = 1 * 1(0.042) + 0.5^2 = 0.292$$

The following are the results of the Hessian matrix calculation.

Table 5. Hessian Matrix Calculation Result

No.	1	2	3	..	18	19	20
1	0.292	0.335	0.292	..	0.25	0.284	0.25
2	0.335	0.419	0.335	..	0.25	0.318	0.25
3	0.292	0.335	0.292	..	0.25	0.284	0.25
4	0.292	0.335	0.292	..	0.25	0.284	0.25
5	0.292	0.335	0.292	..	0.25	0.284	0.25

The next step after calculating the hessian matrix value is to calculate sequential training using the following equation.

$$E_i \sum_{j=1}^{n} \alpha_j D_{ij}$$

Repeat the sequential training process until the maximum iteration, while the maximum iteration in this sample is 3, therefore the process is repeated up to 3 iterations to get the αi value needed to find the support vector.

After the calculation process described earlier, a calculation is made to find the support vector of each document. From the latest αi value. The largest value of each class is taken.

Table 6. Support Vector Determination

No.	D1	D2	D3	D4	D5	α	Kelas
1	0.206	0	0	0	0	0.25599	1
2	0.166	0	0	0	0.166	0.252738	1
3	0.206	0	0	0	0	0.25599	1
4	0	0.206	0	0	0	0.25599	1
5	0	0.206	0	0	0	0.25599	1

The next step is to calculate the kernel function of each class using the highest α value in each class. Manual calculation for K(xi, x+) is obtained based on the largest αi value of the positive class which is 0.259225, while the value of K(xi, x-) is obtained based on the largest αi value which is 0.343225. Based on this value, it is seen from the hessian matrix of the 11th and 10th columns. So that the kernel value of each class will be obtained as follows.

$K(x_i, x^+) = \sum \alpha_i y_i D_i$

= (0.2579 * 1 * 0.25) + (0.2555 * 1 * 0.25) + (0.2579 * 1 * 0.25) + (0.25695 * 1 * 0.25) + (0.25695 * 1 * 0.25) + (0.25669 * 1 * 0.25) + (0.25695 * 1 * 0.25) + (0.25695 * 1 * 0.25) + (0.25837 * 1 * 0.25) + (0.34323 * -1 * -0.292) + (0.25923 * 1 * 0.292) + (0.25923 * 1 *0.292) + (0.25923 * 1 * 0.292) + (0.34323 * -1 * -0.292) + (0.34323 * -1 * -0.292) + (0.25599 * 1 * 0.25) + (0.25599 * 1 * 0.25) + (0.25599 * 1 * 0.25) + (0.25274 * 1 * 0.25) + (0.25599 * 1 * 0.25) = 1.425463

$K(x_i, x^-) = \sum \alpha_i y_i D_i$

= (0.2579 * 1 * -0.25) + (0.2555 * 1 * -0.25) + (0.2579 * 1 * -0.25) + (0.25695 * 1 * -0.25) + (0.25695 * 1 * -0.25) + (0.25669 * 1 * -0.25) + (0.25695 * 1 * -0.25) + (0.25695 * 1 * -0.25) + (0.25837 * 1 * -0.25) + (0.34323 * -1 * 0.292) + (0.25923 * 1 * -0.292) + (0.25923 * 1 * -0.292) + (0.25923 * 1 * -0.292) + (0.34323 * -1 * 0.292) + (0.34323 * -1 * 0.292) + (0.25599 * 1 * -0.25) + (0.25599 * 1 * -0.25) + (0.25599 * 1 * -0.25) + (0.25274 * 1 * -0.25) + (0.25599 * 1 * -0.25) = -1.42546

The values of K(xi, x+) and K(xi, x-) have been calculated, the next step is to calculate the bias value using the equation below.

$b = \frac{1}{2}[\sum_{i=1} \alpha_i y_i K(x_i x^+) + \sum_{i=1} \alpha_i y_i K(x_i x^-)]$

If the data produces a positive value, then the data is classified into a positive class. Meanwhile, if the data is negative, then the data is classified into a negative class.

Support vector machine predicts positive or negative classes by testing data and learning the knowledge contained in the training data. In the training data there are positive and negative sentiment classes, SVM will learn the characteristics of the words contained in each class. Here are the positive and negative sentiment wordclouds [21].
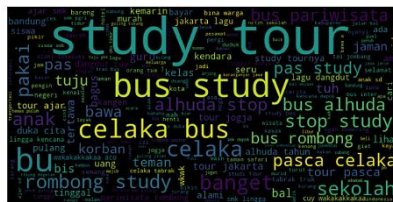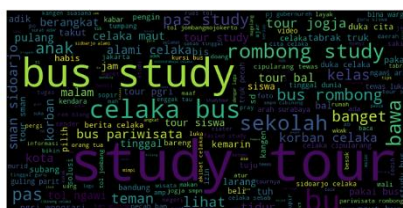

Figure 3. Wordcloud for Sentiment Category "positive"

Figure 4. Wordcloud for Sentiment Category "negative"

3.4 Evaluation of Results

After completing the testing process of the Support Vector Machine algorithm, the results obtained are the labels of the test data generated by the model during the training process. The classification results of the test data, which are the sentiment classes obtained from the program, are compared with the actual class data to determine the accuracy value of the model against the dataset. To see the accuracy value can be seen in the confusion matrix. The following is a confusion matrix of the sentiment analysis results performed using the Support Vector Machine algorithm.
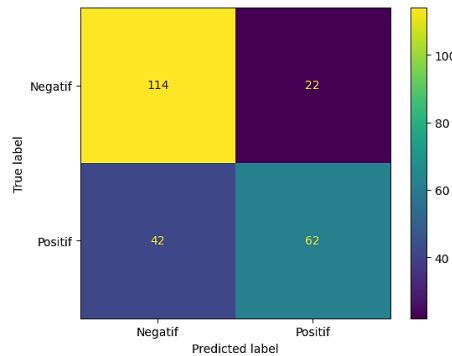

Figure 5. Confusion Matrix

Accuracy is a metric that measures how well a classification model predicts the correct class of data. Then the value of accuracy can be calculated using the equation below.

$$Accuracy = \frac{114+62}{114+22+42+62} \times 100\% = \mathbf{73}\%$$

The purpose of this process is to evaluate the performance of the developed system in analyzing sentiment towards the study tour bus ban policy on Twitter social media. The tools used in this research are Visual Studio Code and Python as the programming language. The classification model used is 8:2. Where 80% of the total data is used as training data and 20% is used as test data. The results of sentiment analysis will be presented in the form of a confusion matrix, where through the confusion matrix and classification report, the accuracy level of the research that has been done will be known. This will provide a better understanding of how public opinion on the study tour bus ban is expressed on social media and how well the system is able to classify the sentiment.

4. CONCLUSION

Study tour is an activity outside the classroom that has the purpose of learning about the process of something directly. This activity is usually carried out by the school once a year. This activity is not only a learning tool for students, but also a recreational activity. In this activity, there are many things that need to be prepared, such as transportation, lodging, meals, and so on. This is sometimes troublesome, because not all tourists or business people have the time and willingness to prepare it. Therefore, they need services during their trip. Based on this need, it makes opportunities for business people engaged in the tour agency industry. This study tour activity triggers cons for parents, especially in the middle to lower class. Because in addition to the cost, the risk of parental concern for children is also very large. For example, a long trip can make children tired or tired for a long time, and what is more dangerous is an accident. The study tour bus accident is a tragedy that not only affects the victims and their families, but also reflects the need for improved transportation safety standards for educational activities. The results of sentiment classification on Study Tour Buses using the Support Vector Machine algorithm that matches the actual data amount to 176 data out of a total of 240 test data. It is known that of the 1200 data obtained regarding sentiment towards there are 519 reviews that are positive and 681 reviews that are negative. The accuracy value of the Study Tour Bus sentiment classification using the Support Vector Machine (SVM) algorithm obtained is 73%.

REFERENCES

[1] B. T. Saputra and D. W. Ariani, "Optimalisasi Media Sosial sebagai Sarana Pemasaran Agen Perjalanan Atmaja Tour Bantul" 2024.
[2] P. E. Saing and J. Y. Sinurat, "ANALISIS PROGRAM STUDY TOUR WISATA SEJARAH SEBAGAI SUMBER BELAJAR PADA SISWA DI SMA NEGERI 2 KOTA JAMBI" 2024.
[3] M. R. F. Septiazi and N. Yuliana, "ANALISIS PENGARUH MEDIA SOSIAL TERHADAP GERAKAN BOIKOT PRODUK ISRAEL DI INDONESIA," *Open Access*, vol. 2, no. 4, 2023.

[4]   A. R. Isnain, A. I. Sakti, D. Alita, and N. S. Marga, "SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM," *JDMSI*, vol. 2, no. 1, p. 31, Feb. 2021, doi: 10.33365/jdmsi.v2i1.1021.

[5]   I. S. Tinendung and I. Zufria, "Pengelompokan Status Stunting Pada Anak Menggunakan Metode K-Means Clustering," *mib*, vol. 7, no. 4, p. 2014, Oct. 2023, doi: 10.30865/mib.v7i4.6908.

[6]   A. Firdaus and W. I. Firdaus, "Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan)," vol. 13, no. 1, 2021.

[7]   M. Furqan, S. Sriani, and S. M. Sari, "Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia," *tc*, vol. 21, no. 1, pp. 51–60, Feb. 2022, doi: 10.33633/tc.v21i1.5446.

[8]   Sonia Wanda Mafriza and Armansyah, "Klasifikasi karir mahasiswa bidang web developer menggunakan algoritma naïve bayes," *infotech*, vol. 4, no. 2, pp. 270–280, Dec. 2023, doi: 10.37373/infotech.v4i2.907.

[9]   R. Kurniawan R and I. Zufria, "Penerapan Text Mining Pada Sistem Penyeleksian Judul Skripsi Menggunakan Algoritma Latent Dirichlet Allocation(LDA)," *ijcs*, vol. 11, no. 3, Dec. 2022, doi: 10.33022/ijcs.v11i3.3120.

[10]   Reski Amalia Salam, Tasrif Akib, and Muhammad Dahlan, "Analisis Wacana Kritis terhadap Sarkasme dalam Twitter Sejak Bulan September-November 2023," *Onoma*, vol. 10, no. 2, pp. 1855–1863, May 2024, doi: 10.30605/onoma.v10i2.3621.

[11]   A. Ma'rif, *Buku Ajar Pemrograman Lanjut Bahasa Pemrograman Python*. Yogyakarta: Universitas Ahmad Dahlan, 2020.

[12]   S. A. Ashari, M. W. A. Saputra, E. Larosa, and B. S. Rijal, "Analisis Sentimen pada Aplikasi Translate Google Menggunakan Metode SVM (Studi Kasus: Komentar Pada Playstore)," *jt*, vol. 21, no. 2, pp. 168–182, Dec. 2023, doi: 10.37031/jt.v21i2.412.

[13]   R. Husna El, R. Wasono, and M. Al Haris, "Analisis Sentimen Pada Twitter Mengenai Netflix Diblokir Telkom Menggunakan Metode Support Vector Machine," *Seminar Nasional Variansi*, pp. 214–222, 2020.

[14]   D. Hardianti, R. R. Pertama, and A. Abdurahman, "KEPASTIAN HUKUM ATAS HAK CIPTA SEBAGAI OBJEK WAKAF BERDASARKAN HUKUM POSITIF DI INDONESIA," vol. 5, 2021.

[15]   Z. Alhaq, A. Mustopa, S. Mulyatun, and J. Santoso Dwi, "Penerapan Metode Support Vector Machine untuk Analisis Sentimen Pengguna Twitter," *Journal of Information System Management*, vol. 3, 2021.

[16]   S. Ailiyya, *Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Tokopedia Menggunakan Metode Support Vector Machine*. Jakarta: UIN Syarif Hidayatullah, 2020.

[17]   E. R. Lidinillah, T. Rohana, and A. R. Juwita, "Analisis sentimen twitter terhadap steam menggunakan algoritma logistic regression dan support vector machine," *tekno*, vol. 10, no. 2, pp. 154–164, Jul. 2023, doi: 10.37373/tekno.v10i2.440.

[18]   M. Ikhsan and R. Kurniawan, "Penerapan Text Mining pada Sistem Rekomendasi Pembimbing Skripsi Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier," *Indonesian Journal of Computer Science*, vol. 12, no. 6, 2023.

[19]   A. H. Lubis, L. P. A. Lubis, and Sriani, "Sentiment analysis on twitter about the death penalty using the support vector machine method," *TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika*, vol. 11, no. 2, pp. 312–321, 2024, doi: 10.37373.

[20]   R. W. Pratiwi, S. F. H, D. Dairoh, D. I. Af'idah, Q. R. A, and A. G. F, "Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector Machine (SVM)," *INISTA*, vol. 4, no. 1, pp. 40–46, Dec. 2021, doi: 10.20895/inista.v4i1.387.

[21]   O. Al Farobey, *Implementasi Metode Support Vector Machine Untuk Mengetahui Respon Masyarakat Terhadap Pemberian Vaksin Sinovac*. Jakarta: UIN Syarif Hidayatullah, 2021.