

## COMPARISON OF SVM AND NAÏVE BAYES CLASSIFIER ALGORITHMS ON STUDENT INTEREST IN JOINING MSIB

Amira Aida Rashifa<sup>1\*</sup>; Hendra Marcos<sup>2</sup>; Pungkas Subarkah<sup>3</sup>; Siti Alvi Sholikhatin<sup>4</sup>

Informatics<sup>1,2,3</sup>

Digital Business<sup>4</sup>

Universitas Amikom Purwokerto, Indonesia<sup>1,2,3,4</sup>

www.amikompurwokerto.ac.id<sup>1,2,3,4</sup>

amiraaidarashifa@gmail.com<sup>1\*</sup>, hendra.marcos@amikompurwokerto.ac.id<sup>2</sup>,

subarkah@amikompurwokerto.ac.id<sup>3</sup>, sitialvi@amikompurwokerto.ac.id<sup>4</sup>

(\*) Corresponding Author

(Responsible for the Quality of Paper Content)



The creation is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License.

**Abstract**— Machine learning (ML) is a branch of artificial intelligence (AI) that deals with the development of systems capable of learning from data to make predictions or decisions without being explicitly programmed. In this study, we conducted an analysis of students' interest in the Internship and Certified Independent Study Program (MSIB) in the context of the Independent Campus Learning policy. The method used is a survey by distributing questionnaires to students of Amikom Purwokerto University in the MSIB batch 5 in year 2023. The results of this study can provide understanding and predictions about students' interest in the MSIB program based on relevant variables, such as study program, semester, cumulative grade point average (GPA), semester credit system (SKS), and previous work experience. The research results indicate that GPA and Study Program greatly influence students' interest in MSIB. The Naïve Bayes algorithm yielded an accuracy of 0.6875 on the training data and 0.25 on the testing data, with a confusion matrix of (0, 1, 0; 0, 1, 2; 0, 0, 0). Meanwhile, the Support Vector Machine (SVM) algorithm yielded an accuracy of 0.4375 on the training data and 0.75 on the testing data, with a confusion matrix of (0, 1; 0, 3). The machine learning model developed in this study is expected to help predict students interest based on new data provided, thus supporting decision-making in optimizing the MSIB program.

**Keywords:** merdeka belajar kampus merdeka, MSIB, naive bayes classifier, support vector machine

**Intisari**— Machine learning (ML) adalah cabang dari kecerdasan buatan yang berkaitan dengan pengembangan sistem yang dapat belajar dari data untuk membuat prediksi atau keputusan tanpa harus diprogram secara eksplisit. Dalam penelitian ini, kami melakukan analisis terhadap minat mahasiswa terhadap program Magang dan Studi Independent Bersertifikat (MSIB) dalam konteks kebijakan Merdeka Belajar Kampus Merdeka. Metode yang digunakan adalah survei dengan menyebarkan kuesioner kepada mahasiswa Universitas Amikom Purwokerto pada MSIB batch 5 Tahun 2023. Hasil penelitian ini dapat memberikan pemahaman dan prediksi tentang minat mahasiswa terhadap program MSIB berdasarkan variabel-variabel yang relevan, seperti program studi, semester studi, indeks prestasi kumulatif (IPK), sistem kredit semester (SKS), dan pengalaman kerja sebelumnya. Hasil penelitian menunjukkan bahwa IPK dan Program Studi sangat mempengaruhi minat mahasiswa terhadap MSIB. Algoritma naive bayes menghasilkan akurasi pada data training sebesar 0.6875 dan pada data testing sebesar 0.25, dengan confusion matrix (0, 1, 0; 0, 1, 2; 0, 0, 0). Sementara itu, algoritma support vector machine (SVM) menghasilkan akurasi pada data training sebesar 0.4375 dan pada data testing sebesar 0.75, dengan confusion matrix (0, 1; 0, 3). Model machine learning yang dikembangkan dalam penelitian ini diharapkan dapat membantu memprediksi minat mahasiswa berdasarkan data baru yang diberikan, sehingga dapat mendukung pengambilan keputusan dalam mengoptimalkan program MSIB.

**Kata Kunci:** merdeka belajar kampus merdeka, MSIB, naive bayes classifier, support vector machine

## INTRODUCTION

The Merdeka Belajar Kampus Merdeka (MBKM) program is an innovative policy issued by the Ministry of Education, Culture, Research, and Technology [1]. One of the key aspects of this policy is the Certified Internship and Independent Study (MSIB) program [2]. MSIB aims to encourage students' interest in mastering various sciences that will be an important provision when they enter the increasingly competitive world of work. Universities that have adopted the MBKM program have succeeded in creating a more flexible and independent learning environment. This allows students to be actively and innovatively involved in the learning process. This program pays close attention to student needs, covering three important aspects, namely attitudes, knowledge, and skills [3]. Participants who take part in the MSIB program can get a maximum value conversion of 20 credits [4]. This is the main attraction of the program, as it gives students the opportunity to gain real-world experience with academic value.

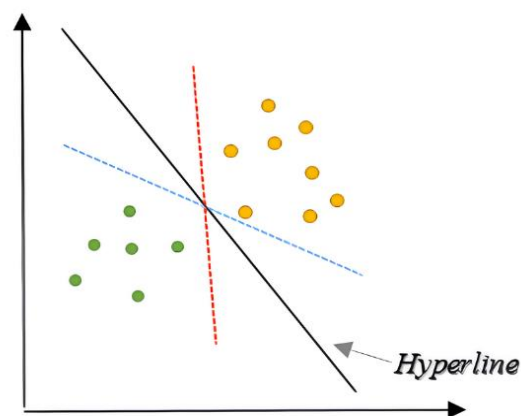
The implementation of the MBKM policy on campuses is able to influence student interest and involvement in the program. The success of this policy largely depends on the extent to which it is able to motivate students to be actively involved, which in turn will help improve their abilities. There is a relevant previous study by Rochana et al. (2021), which noted the determinants that shape student interest and involvement in the independent campus policy. This research provides additional insight into the factors that may influence student interest in the MSIB program [5]. In addition, other studies have compared support vector machine (SVM) and naïve bayes classification (NBC) algorithms in other contexts. For example, research by Sulton Nur Hakim (2021) conducted sentiment analysis of myindihome user perceptions using the support vector machine (SVM) and naïve bayes classification (NBC) methods. The results show that SVM has better performance in classifying myIndiHome review data than the NBC method [6]. Classification method is one of the useful approaches to determine which category is most suitable for a data object, in this case, the category of student interest in the MSIB MBKM program [7]. SVM is an algorithm that focuses on optimal separation between classes in a dataset [8], while NBC is an algorithm that relies on simple probabilities based on Bayes' theorem [9][10]. Both will be used to understand students' interest in the MSIB program.

The research to be conducted aims to analyze and compare student interest in the MSIB program. The research data is taken from the survey results filled in by Amikom Purwokerto University students. This data will be processed into a classification model to determine student interest in the MSIB MBKM program with the help of machine learning. This research will compare two different machine learning algorithms, namely support vector machine (SVM) and naïve bayes classifier (NBC) in analyzing student interests.

## MATERIALS AND METHODS

Naïve Bayes classifier (NBC) is a classification algorithm based on probability and Bayes theorem [11], [12]. In this research, NBC will be used to classify students into interest groups based on datasets and to classify student interests based on probability [13]-[15].

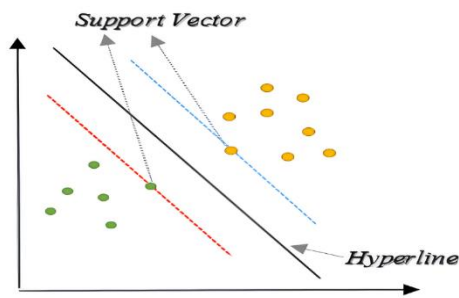
Support Vector Machine (SVM) is an algorithm that focuses on the optimal separation between classes in a dataset [16], [17]. In this study, SVM will be used to create an optimal separation [18] between groups of students based on factors that reflect their interest in the MSIB program.



Source: (Research Result, 2024)

Figure 1. SVM Division Illustration 1

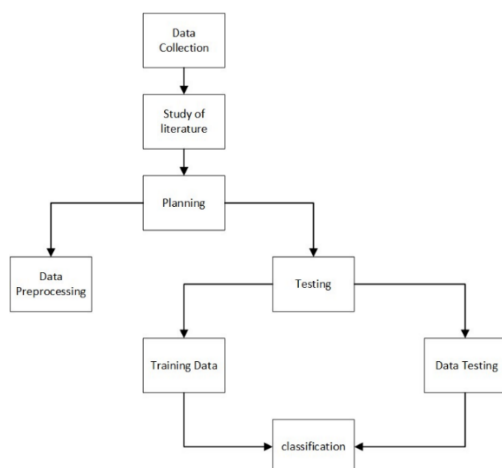
There are three boundary lines shown in Figure 1. The hyperline line located in the center of the figure is the hyperline line, and the other two lines are the separators between the vector points identified by the yellow and green points. For the purpose of finding accuracy, the SVM technique divides the yellow-green vector groups by finding the vector points that are closest to the group and that are close to the hyperline [19]. The vector point closest to the hyperline in Figure 2 should be in a straight direction with respect to the hyperline [20].



Source: (Research Result, 2024)  
Figure 2. Illustration of SVM Division 2

In the process of data processing, there is a stage that is carried out to measure the performance assessment of a system. This stage is known as performance evaluation measure (PEM), and is used in this system to find the identification result of information comparison with confusion matrix. This is achieved by calculating the accuracy, precision, recall and F1-score values. Accuracy is the degree of accuracy between the user's query and the system's answer, recall is the comparison of the system's exact reading with previously executed information, and F1-score is the harmonic mean or average value of accuracy and recall. Accuracy is the result of the data read by the system from all information. Accuracy is the degree of accuracy between user requests and answers from the system [21]. In this system design, google colab is used as the coding environment, and Python is the programming language used. If this environment is used in the machine learning calculation process, there is no need to manually calculate the accuracy, precision, recall and F1-score levels. Python already has the ability to directly calculate these levels automatically[22].

The research flow consists of several stages as shown in Figure 3.



Source: (Research Result, 2024)  
Figure 3. Research flow

Figure 3 is the flow of the research methodology to be carried out, starting with data collection. In this research, the data collection method used is the survey method. The survey was conducted by distributing questionnaires to students of Amikom Purwokerto University to collect data about their interest in MSIB participation. The questionnaire included questions related to motivation, expected benefits, and preferences in participating in MSIB activities. Next is the literature study step, which uses references from previous research by utilizing the support vector machine (SVM) and Naïve Bayes Classifier (NBC) methods [23] for the process of classifying and grouping data. After the survey data is collected, a data pre-processing step is performed to prepare the raw data into a format that is more suitable for analysis with machine learning algorithms[24], [25]. Some pre-processing methods that can be used are as follows:

- Data cleaning: removing irrelevant, duplicate, or incomplete data, as well as troubleshooting missing or invalid data.
- Data normalization: converting data into a standard or normal form to eliminate formatting differences, for example, converting text to all lowercase or replacing synonyms with similar words.
- Categorical data conversion: converting categorical variables into a numerical form that can be understood by the algorithm, such as using one-hot encoding or label encoding.

After the data cleaning process was complete, the data was then categorized according to the labels. This test is conducted by collecting data to identify whether students are interested in MSIB or not. The bound variable used was whether students were interested or not (yes or no), and the independent variables used were course, semester, credits, and work experience. The computational procedures were performed in a computerized manner, which means that codes were entered into the system. This process also includes the performance of the dataset. In testing and training, data training serves as the basis for determining whether the data used is workable. Data testing, on the other hand, refers to using new data to evaluate the accuracy of a system and determine whether it is already functioning according to the required standards. The next step in the identification process is to calculate accuracy, precision, recall and F1-score on the system that has been built [21]. This is done in order to determine the results of the NBC and SVM algorithms, which will allow for comparisons to be made between the two approaches.

**RESULTS AND DISCUSSION**

Table 1 below is the result of use based on the variables that have been planned in the method section.

Table 1. Required data variables

Variable	Variable's description
X1	Study Program
X2	Semester
X3	GPA
X4	SKS
X5	Working Status or Not
Y1	Students' interest level to join MSIB

Source: (Research Result, 2024)

In this research, the main objective is to develop a machine learning model that can help predict student interest in internships and certified independent study (MSIB). MSIB has the potential to provide students with practical experience and additional knowledge outside the traditional academic environment. To achieve this goal, classification algorithms such as naïve bayes classifier and support vector machine (SVM) will be used. These algorithms will learn patterns from training data involving various relevant variables. In this context, the variables taken into account include the student's study program, semester of study, cumulative grade point average (GPA), semester credit system (SKS), and previous work experience.

Number of rows : 120  
 Number of columns : 6  
 Data dimensions : (120, 6)

Source: (Research Result, 2024)

Figure 4. Number of rows, columns, and dimensions of the dataset

Figure 4 is a description of the dataset, where the first row is used to print 120 rows of data, the second row prints the number of columns as many as 6, and the third row prints the dimensions of the data in the form of tuples with dimensions (120,6).

Study Program	SKS	GPA	Semester	Work Experience	MSIB
Informatics	60	3.59	4	Yes	Interested
Informatics	63	3.74	4	No	Interested
Informatics	20	3.83	2	No	Very Interested
Informatics	144	3.73	6	Yes	Very Interested
Other	57	3.15	3	Yes	Not Interested
Informatics	63	3.54	4	No	Fairly Interested
Informatics	86	3.94	4	No	Very Interested
Other	20	3.79	2	No	Interested
Informatics	63	3.90	4	No	Interested
Informatics	83	3.65	4	Yes	Fairly Interested
Informatics	40	3.72	2	Yes	Not Interested
Informatics	60	3.13	4	No	Fairly Interested
Informatics	63	3.56	4	No	Fairly Interested
Informatics	63	3.62	4	Yes	Interested
Information Systems	63	3.47	4	No	Interested
Information Systems	23	3.90	4	No	Very Interested
Information Technology	123	2.75	6	No	Not Interested
Information Technology	63	3.97	4	No	Interested
Informatics	109	3.82	6	No	Very Interested
Informatics	63	3.60	4	No	Interested

Number of data rows: 20

Source: (Research Result, 2024)

Figure 5. Results of dataset testing of students' interest in MSIB

Next Figure 5 is the data frame function 'df'. First, the relevant columns are selected, and then 20 rows of data are randomly displayed with better variation. The following Figure 5 is the result of the model test with one dependent variable and four independent variables based on the dataset.

	Study Program	SKS	GPA	Semester	Work Experience	MSIB
1	1	60	3.59	4	1	2
2	1	63	3.74	4	0	2
3	1	20	3.83	2	0	1
4	1	144	3.73	6	1	1
5	4	57	3.15	3	1	4
6	1	63	3.54	4	0	3
7	1	86	3.94	4	0	1
8	4	20	3.79	2	0	2
9	1	63	3.90	4	0	2
10	1	83	3.65	4	1	3
11	1	40	3.72	2	1	4
12	1	60	3.13	4	0	3
13	1	63	3.56	4	0	3
14	1	63	3.62	4	1	2
15	2	63	3.47	4	0	2
16	2	23	3.90	4	0	1
17	3	123	2.75	6	0	4
18	3	63	3.97	4	0	2
19	1	109	3.82	6	0	1
20	1	63	3.60	4	0	2

Source: (Research Result, 2024)

Figure 6. Data normalization

Figure 6 shows the 'df' data frame that has been transformed in several columns. The 'Study Program' column is transformed into a numerical representation where 'Informatics' is represented as 1, 'Information Systems' as 2, 'Information Technology' as 3, and 'Other' as 4. Then the 'MSIB' column is transformed into a numerical value where 'Very Interested' is represented as 1, 'Interested' as 2, 'Moderately Interested' as 3, 'Not Interested' as 4, and 'Very Not Interested' as 5. Lastly, the column 'Work Experience' is transformed into a boolean value, with 'Yes' represented as 1 and 'No' represented as 0. This transformation helps convert categorical data into numerical form, facilitating data analysis and modeling, and the result is a data frame consisting of numerical data that can be further used for analysis and processing.

Accuracy: 25.00%

Classification Report:

	precision	recall	f1-score	support
1	0.00	0.00	0.00	1
2	0.50	0.33	0.44	3
3	0.00	0.00	0.00	0
accuracy			0.25	4
macro avg	0.17	0.11	0.13	4
weighted avg	0.38	0.25	0.30	4

Source: (Research Result, 2024)

Figure 7. Accuracy and naïve bayes classification report

Figure 7 shows the performance evaluation of a classification model, which may use methods such as Support Vector Machine (SVM) or Naïve



Bayes Classifier to predict the target variable 'MSIB' based on pre-selected independent variables. Let's discuss each of the evaluation metrics presented in the classification report:

- a. Accuracy: 25.00%
- b. The 25% accuracy indicates that the model predicts correctly for only a quarter of the data.
- c. Classification report:
  1. Precision:  
Class 1: 0.00 (No correct predictions for this class)  
Class 2: 0.50 (Half of the class 2 predictions were correct)  
Class 3: 0.00 (No correct predictions for this class)
  2. Recall or Sensitivity:  
Class 1: 0.00 (No correct predictions for this class)  
Class 2: 0.33 (Part of the class 2 data was detected correctly)  
Class 3: 0.00 (No correct prediction for this class)
  3. F1-Score:  
Class 1: 0.00 (No correct prediction for this class)  
Class 2: 0.40 (Balanced combination of precision and recall for class 2)  
Class 3: 0.00 (No correct prediction for this class)
  4. Support:  
Indicates the number of instances (samples) that support each class
  5. Macro Average and Weighted Average:  
Macro avg takes the average of the metric values for each class regardless of how much each class contributes to the total data, while weighted avg gives each class a weight based on the number of samples in each class.

Accuracy: 75.00%					
Classification Report:					
		precision	recall	f1-score	support
	1	0.00	0.00	0.00	1
	2	0.75	1.00	0.86	3
	accuracy			0.75	4
	macro avg	0.38	0.50	0.43	4
	weighted avg	0.56	0.75	0.64	4

Source: (Research Result, 2024)

Figure 8. Accuracy and classification report of support vector machine

Figure 8 explains the model evaluation results, there is an increase in accuracy to 75.00%.

- a. Accuracy: 75.00%
- b. The 75% accuracy indicates that the model predicted correctly in three out of four cases.

- c. Classification report:
  1. Precision:  
Class 1: 0.00 (No correct predictions for this class)  
Class 2: 0.75 (Three-quarters of the class 2 predictions were correct)
  2. Recall or Sensitivity:  
Class 1: 0.00 (No correct prediction for this class)  
Class 2: 1.00 (All class 2 data is correctly detected)
  3. F1-Score:  
Class 1: 0.00 (No correct prediction for this class)  
Class 2: 0.86 (High F1 score indicates a balance between precision and recall for class 2)
  4. Support:  
Indicates the number of instances (samples) that support each class
  5. Macro Average and Weighted Average:  
Macro avg takes the average of the metric values for each class regardless of how much each class contributes to the total data, while Weighted avg gives each class a weight based on the number of samples in each class.

Accuracy on training data: 0.6875  
Accuracy on testing data: 0.25

Source: (Research Result, 2024)

Figure 9. Accuracy of naïve bayes training and testing data

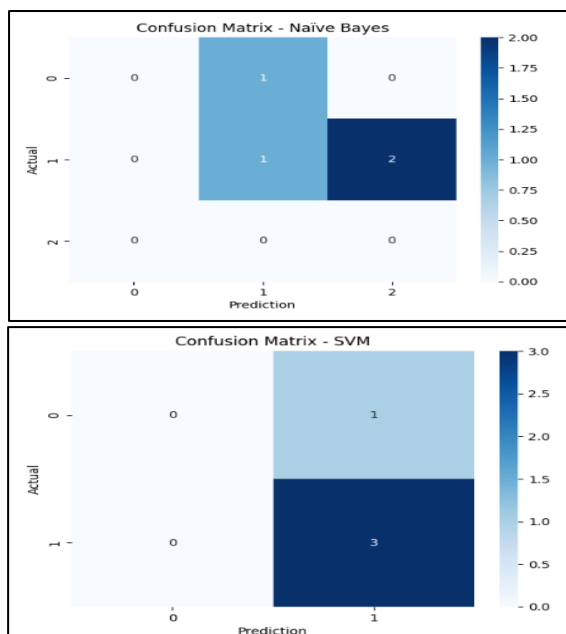
In Figure 9, the results of high accuracy on training data (0.6875) and low accuracy on testing data (0.25) indicate overfitting in the model. Overfitting occurs when the model is overly adapted to specific training data and cannot generalize well to data it has never seen.

Accuracy on training data: 0.4375  
Accuracy on testing data: 0.75

Source: (Research Result, 2024)

Figure 10. Accuracy of support vector machine training and testing data

Figure 10 shows that there is a significant difference between the accuracy on the training and testing data (0.4375 on the training data and 0.75 on the testing data), indicating that the model may be overfitting and the opposite, underfitting. Underfitting occurs when the model is not complex enough to capture the patterns present in the training data.



Source: (Research Result, 2024)

Figure 11. Confusion matrix comparison

From the data results in Figure 11 information about missing values = 0.00%, so the data is very accurate. Strongly affecting student interest in MSIB are GPA and Study Programme. The naïve bayes classifier algorithm produces accuracy on training data = 0.6875 and accuracy on testing data = 0.25, so the confusion matrix = (0, 1, 0; 0, 1, 2; 0, 0, 0). The support vector machine algorithm produces accuracy on training data = 0.4375 and accuracy on testing data = 0.75, so the confusion matrix = (0, 1; 0, 3). A student's programme of study can provide insight into the specific interests a student may have in a particular field. Semester of study can give an indication of a student's level of progress and readiness to engage in internship or independent study activities. GPA and credits can reflect students' academic performance and the extent to which they have engaged themselves in college activities. Previous work experience can also play an important role, as it can provide insights and skills that are relevant in the context of internships and independent studies.

Utilizing these variables, the machine learning model will be trained using historical training data that includes information on student interest in internships and certified independent study. The data will be used to train the model to learn the relationship between relevant variables and student interest. After going through the training process, the trained model will be used to predict students' interest based on the new data provided. Thus, the model can provide assistance in

understanding and predicting students' interest in internships and certified independent studies based on the previously described variables.

## CONCLUSION

Based on the results of research that has been carried out on the data test model above with the dependent variable of interest or not (yes or no) and independent variables such as study programme, semester, credits, work experience, it can be concluded to determine student interest in the MSIB program at Amikom Purwokerto University can be analyzed using machine learning methods, such as support vector machine (SVM) and naïve bayes classifier (NBC). Survey data that has been processed through pre-processing can be used as input to train machine learning models. SVM and NBC models can predict student interest in the MSIB programme with different levels of accuracy. The accuracy of the SVM model is higher (75%) than the NBC model (25%). Model evaluation using methods such as accuracy, precision, recall, F1-score, and confusion matrix can provide information about the model's performance in predicting student interest. Variables such as study programme, semester of study, GPA, credits, and previous work experience can be factoring that influence student interest in the MSIB programme. The machine learning model developed can help better decision-making and campus policy development in optimizing the MSIB program. The factors that strongly influence student interest in MSIB are GPA and Study Programme. The machine learning model developed in this research can provide assistance in predicting student interest based on new data provided, so that it can support decision making in optimizing the MSIB program.

## REFERENCE

- [1] Y. B. Bhakti, M. R. R. Simorangkir, A. Tjalla, and A. Sutisna, "Kendala Implementasi Kebijakan Merdeka Belajar Kampus Merdeka (MbkM) Di Perguruan Tinggi," *Res. Dev. J. Educ.*, vol. 8, no. 2, p. 783, 2022, doi: 10.30998/rdje.v8i2.12865.
- [2] A. N. B. Sulistyanningrum, N. A. Nirwana, D. R. Januar, and N. N. Hilalia, "Performa Kebijakan Kampus Merdeka pada Program Magang dan Studi Independen Bersertifikat," *J. Multidisiplin Madani*, vol. 2, no. 6, pp. 2771-2786, 2022, doi: 10.55927/mudima.v2i6.489.
- [3] B. Bagus and N. Eko, "Evaluasi Program



- Kegiatan Magang dan Studi Independen Bersertifikat (MSIB) Batch 2 Bidang Drafter Mahasiswa Program Studi Pendidikan Teknik Sipil dan Perencanaan Fakultas Teknik Universitas Negeri Yogyakarta," *Jurnal Elektronik Mahasiswa Pendidikan Teknik Sipil (JEPTS)*, vol. 11, no. 1, pp. 79-87, 2023.
- [4] R. Rochana, R. M. Darajatun, and M. A. Ramdhany, "Pengaruh Implementasi Kebijakan Kampus Merdeka terhadap Minat dan Keterlibatan Mahasiswa," *J. Bus. Manag. Educ.*, vol. 6, no. 3, pp. 11-21, 2021, doi: 10.17509/jbme.v6i3.40165.
- [5] S. N. Hakim, "Analisis Sentimen Persepsi Pengguna Myindihome Menggunakan Metode Support Vector Machine (Svm) Dan Naïve Bayes Classifier (NBC)." Universitas Islam Indonesia, 2021.
- [6] Y. A. Singgalen, "Jurnal Media Informatika Budidarma Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Decision Tree (DT)," *Jurnal Media Informatika Budidarma*, vol. 7, no. 3, pp. 1551-1562, 2023, doi: 10.30865/mib.v7i3.6461.
- [7] A. Asrifan, B. Buhari, and I. Manda, "The Views And Energy Of Students Regarding The Implementation Of MBKM At The English Department Of Universitas Muhammadiyah Sidenreng Rappang," *La Ogi English Lang. J.*, vol. 9, no. 2, pp. 71-80, 2023, doi: 10.55678/loj.v9i2.1012.
- [8] M. Asfi and N. Fitrianiingsih, "Implementasi Algoritma Naive Bayes Classifier sebagai Sistem Rekomendasi Pembimbing Skripsi," *J. Nas. Inform. dan Teknol. Jar.*, vol. 5, pp. 45-50, 2020, [Online]. Available: <https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/2536>
- [9] I. Padiku, "Penerapan Metode Naive Bayes Classifier (Nbc) Untuk Klasifikasi Kondisi Internal Program Studi," *J. Tek.*, vol. 19, no. 1, pp. 65-74, 2021, doi: 10.37031/jt.v19i1.118.
- [10] R. A. Ibrahim Alhayali, M. A. Ahmed, Y. M. Mohialden, and A. H. Ali, "Efficient method for breast cancer classification based on ensemble hoeffding tree and naïve Bayes," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 18, no. 2, pp. 1074-1080, 2020, doi: 10.11591/ijeecs.v18.i2.pp1074-1080.
- [11] P. Subarkah, P. Arsi, D. I. S. Saputra, A. Aminuddin, B. Berlilana, and N. Hermanto, "Indonesian Police in the Twittersverse: A Sentiment Analysis Perspectives," *2023 IEEE 7th Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng.*, pp. 76-81, 2023, doi: 10.1109/ICITISEE58992.2023.10405357.
- [12] M. H. Effendy, D. Anggraeni, Y. S. Dewi, and A. F. Hadi, "Classification of Bank Deposit Using Naïve Bayes Classifier (NBC) and K - Nearest Neighbor ( K -NN) ," *Proc. Int. Conf. Math. Geom. Stat. Comput. (IC-MaGeStiC 2021)*, vol. 96, pp. 163-166, 2022, doi: 10.2991/acsr.k.220202.031.
- [13] I. Maulana, W. Apriandari, and A. Pambudi, "Analisis Sentimen Berbasis Aspek Terhadap Ulasan Aplikasi Mypertamina Menggunakan Support Vector Machine," *Idealis: Indonesia Journal Information System*, vol. 6, no. 2, pp. 172-181, 2023. doi: 10.36080/idealis.v6i2.3022.
- [14] P. Subarkah, P. W. Rahayu, I. Darmayanti, and R. Riyanto, "Sentiment Analysis on Reviews of Women'S Tops on Shopee Marketplace Using Naive Bayes Algorithm," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 9, no. 1, pp. 126-133, 2023, doi: 10.33480/jitk.v9i1.4179.
- [15] P. Subarkah, W. Risma, and R. Aditya, "Comparison of correlated algorithm accuracy Naive Bayes Classifier and Naive Bayes Classifier for heart failure classification," *ILKOM Jurnal Ilmiah*, vol. 14, no. 2, pp. 120-125, 2022, doi: 10.33096/ilkom.v14i2.1148.120-125.
- [16] A. Jalu, N. Kisma, P. Arsi, and P. Subarkah, "Sentiment Analysis Regarding Candidate Presidential 2024 Using Support Vector Machine Backpropagation Based," *TAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 8, no. 1, pp. 96-108, 2024, doi: 10.31764/jtam.v8i1.17294.
- [17] S. Muawanah, U. Muzayanah, M. G. R. Pandin, M. D. S. Alam, and J. P. N. Trisnaningtyas, "Stress and Coping Strategies of Madrasah's Teachers on Applying Distance Learning During COVID-19 Pandemic in Indonesia," *Qubahan Acad. J.*, vol. 3, no. 4, pp. 206-218, 2023, doi: 10.48161/Issn.2709-8206.
- [18] O. Rostami and M. Kaveh, "Optimal feature selection for SAR image classification using biogeography-based optimization (BBO), artificial bee colony (ABC) and support vector machine (SVM): a combined approach of optimization and machine learning," *Comput. Geosci.*, vol. 25, no. 3, pp.



- 911-930, 2021, doi: 10.1007/s10596-020-10030-1.
- [19] J. Zhang, D. Shen, D. Chen, D. Ming, D. Ren, and Z. Diao, "ISMSFuse: Multi-modal fusing recognition algorithm for rice bacterial blight disease adaptable in edge computing scenarios," *Comput. Electron. Agric.*, vol. 223, no. March, p. 109089, 2024, doi: 10.1016/j.compag.2024.109089.
- [20] M. Maulita, "Pendekatan Data Mining Untuk Analisa Curah Hujan Menggunakan Metode Regresi Linear Berganda (Studi Kasus: Kabupaten Aceh Utara)," *Idealis: Indonesia Journal Information System*, vol. 6, no. 2, pp. 99-106, 2023, doi: 10.36080/idealis.v6i2.3034.
- [21] M. R. Qisthiano, T. B. Kurniawan, E. S. Negara, and M. Akbar, "Pengembangan Model Untuk Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu dengan Metode Naïve Bayes," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 3, p. 987, 2021, doi: 10.30865/mib.v5i3.3030.
- [22] R. Monat, A. Ouadjaout, and A. Miné, "Static type analysis by abstract interpretation of python programs," *Leibniz Int. Proc. Informatics, LIPIcs*, vol. 166, 2020, doi: 10.4230/LIPIcs.ECOOP.2020.17.
- [23] D. Atika, A. Ari Aldino, S. Informasi, J. Pagar Alam No, L. Ratu, and K. Kedaton, "Term Frequency-Inverse Document Frequency Support Vector Machine Untuk Analisis Sentimen Opini Masyarakat Terhadap Tekanan Mental Pada Media Sosial Twitter," *Jurnal Teknologi dan Sistem Informasi*, vol. , no. 4, pp. 86-97, 2022, doi: 10.33365/jtsi.v3i4.2054.
- [24] P. Subarkah, E. P. Pambudi, S. Oktaviani, and N. Hidayah, "Perbandingan Metode Klasifikasi Data Mining untuk Nasabah Bank Telemarketing," *Matrik: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 1, pp. 139-148, 2020, doi: 10.30812/matrik.v20i1.826.
- [25] A. Andreyestha and A. Subekti, "Analisa Sentiment Pada Ulasan Film Dengan Optimasi Ensemble Learning," *J. Inform.*, vol. 7, no. 1, pp. 15-23, 2020, doi: 10.31311/ji.v7i1.6171.

