Jurnal Teknik Informatika (JUTIF)
Vol. 5, No. 4, August 2024, pp. 241-251

DOI: https://doi.org/10.52436/1.jutif.2024.5.4.1983
p-ISSN: 2723-3863
e-ISSN: 2723-3871

EXPERIMENTAL COMPARISON OF MACHINE LEARNING ALGORITHM PERFORMANCE FOR OPTIMIZING ELECTIVE SUBJECT SELECTION IN PHASE F OF THE MERDEKA CURRICULUM

Dedy Mulyadi[*1]

[1]Informatics, Universitas Pradita, Indonesia
Email: [1]dedy.mulyadi@student.pradita.ac.id

(Article received: April 15, 2024; Revision: May 06, 2024; published: August 01, 2024)

Abstract

In phase f of the Merdeka Curriculum, electives are an important element at the senior high school level. Students are faced with the challenge of choosing four out of twelve elective subjects that are relevant to their talents, interests, further study plans and career goals over a two-year study period. Applying machine learning with the right algorithm is a solution for the effectiveness and efficiency of elective selection. The dataset used comes from the 10th grade report card data, the results of the interest, aptitude, further study, and career choice tests, and the manual selection of electives chosen by students in the previous year. The use of a small data set requires a cross-validation method to improve the generalizability of the model and to optimize the data set, thereby increasing the validity of the results. The test will be conducted using an application that tests five machine learning algorithm models suitable for small datasets, namely Naive Bayes, Decision Tree, Random Forest, Support Vector Machine, and k-Nearest Neighbors. The test focuses on comparing the performance of the five algorithms based on the best accuracy, recall, and confusion matrix and the results obtained Support Vector Machine (SVM) algorithm has the best performance results by achieving the highest accuracy of 57.3770%, the highest recall of 0.574, and the highest true positive (TP) of 0.574. The Support Vector Machine (SVM) algorithm will be a recommendation for further research, namely the development of machine learning for the selection of f-stage elective subjects at Atisa Dipamkara senior high school, to provide relevant guidance to students in making decisions regarding the selection of elective subjects more accurately and according to their respective characteristics.

Keywords: *cross validation, decision tree, KNN, machine learning, merdeka curriculum phase f, naive bayes, random forest, small dataset, SVM*.

1. INTRODUCTION

1.1. Background

In the world of education, the fundamental for the implementation of an education system is the curriculum which is a set of plans and arrangements regarding the objectives, content, and learning materials as well as the methods used as guidelines for organizing learning activities to achieve curriculum objectives. In Indonesia itself, the curriculum has undergone several changes, ranging from the 1947 curriculum, 1994 curriculum, 2006 curriculum, 2013 curriculum, to the ***Merdeka Curriculum*** used in most schools in Indonesia today.

The ***Merdeka Curriculum*** is a curriculum with diverse intra curricular learning where content will be optimized so that learners have enough time to explore concepts and strengthen competencies [1]. In the ***Merdeka Curriculum*** there are have 6 phases, namely: Phase A for grades 1 - 2 elementary school, Phase B for grades 3 - 4 elementary school, Phase C for grades 5 - 6 elementary school, Phase D for grades 7 - 9 junior high school, Phase E for grades 10 senior or vocational high school and Phase F for grades 11 - 12 senior or vocational high school.

In phase f of the ***Merdeka Curriculum*** for grades 11 and 12 of senior high school there are elective subject groups that each student can choose to pursue for 2 years of study. As stated in the Decree of the Minister of Education, Culture, Research and Technology of the Republic of Indonesia number 262/M/2022 concerning guidelines for curriculum implementation in the framework of learning recovery on pages 21-23, the curriculum structure for phase f of the senior high school level is divided into three parts, namely the compulsory subject group, the elective subject and the local content subject group.

For the group of compulsory subjects, which consists of 8 subjects and the local content subjects developed by the school based on its specificity. All these subjects are followed by all students. Meanwhile, in the group of elective subjects, which consists of at most 19 subjects, schools are required to provide at least 7 elective subjects tailored to the learning needs of students and the resources owned by the school, then each student will choose 4 - 5 elective subjects that will be pursued during the 2 years of study in phase f of the ***Merdeka Curriculum***. The selection of elective subjects in phase f of the

Merdeka Curriculum will be the focus of this research.

Atisa Dipamkara senior high school as the third batch of ***sekolah penggerak*** received full guidance from the Banten Provincial Driving Teacher Centre, which is an extension of the Ministry of Education and Culture to implement the ***Merdeka Curriculum*** correctly and precisely, which has been implemented from 2022 to date.

In relation to the focus of this research, Atisa Dipamkara senior high school, based on student's learning needs and the school's resources, provides 12 elective subjects from which students can choose 4 elective subjects. The elective subjects chosen by each student are unique, so students are given full freedom and responsibility to choose the elective subjects they will pursue for 2 years of study. However, choosing 4 of the 12 elective subjects provided by the school is a challenging task for students and their parents. This was also found in other senior high schools with the root cause of students lack of understanding from the relationship between the subjects chosen and the majors in college and related job opportunities. In addition, students also need guidance in understanding the difference between interests and talents and how to identify them in order to choose appropriate subjects [2].

Based on the background of these problems, Atisa Dipamkara senior high school implemented a mentoring program that began with conducting an interest and talent test in collaboration with an external psychological institution that provided results in the form of potential talents, interests, learning styles, personality types based on MBTI as well as further study and career preferences of each student. The next mentoring program is the curation of 10th grade transcripts to understand each student's strengths and weaknesses by subject. The results of these two support programs are then processed by the counselling teacher and vice principal for curriculum to produce recommendations for 4 elective subjects for each student. These recommendations have so far been able to answer the doubts and anxieties of students and parents in choosing elective subjects so that the next process is only a matter of convincing each student's personal self to choose based on the recommendations given and pursue it for the next 2 years of study. However, it turns out that this process takes a long time and requires a lot of accuracy and resources and is still done manually so that it is heavily dependent on the role of humans who process the data, they are counselling teacher and vice principal for curriculum.

1.2. Motivation

In the current technological era, the application of machine learning has wide applicable on various sectors including education [3]. The application of predictive model algorithm in machine learning becomes very useful in providing a prediction based on the given dataset [4]. This research is driven by the need of Atisa Dipamkara Senior High School to develop a machine learning system that can provide recommendations with high accuracy and precision for each student to choose relevant elective subjects in phase f of the ***Merdeka Curriculum***. The main motivation in this research is to facilitate the process of selecting elective subjects and more efficient the working time and resources used and certainly reduce the dependence on the human factor.

1.3. Objectives

The purpose of this research is to find the best machine learning algorithm for small datasets based on the metric evaluation test results that have been set at the beginning of the research. The best machine learning algorithm can later become a foundation for further research to develop a machine learning algorithm that can be used on small datasets, capable of providing recommendations for elective subjects that suit each student's characteristics.

1.4. Contribution

This research will provide a contribution in the form of recommendations for the best machine learning algorithms that can be used in further research to develop machine learning for the selection of ***Merdeka Curriculum*** phase f elective subjects at Atisa Dipamkara senior high school.

1.5. Literature Study

Machine learning algorithms used to design descriptive models with small datasets include Naive Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM) and k-Nearest Neighbours (KNN).

Naive Bayes is an algorithm based on Bayes' theorem, in which the attributes have no relationship or dependency, so that each attribute is mutually independent [5]. In the Naive Bayes machine learning algorithm using collaborative filtering can make personalized recommendations [6]. Classification with the Naive Bayes algorithm provides a practical way because the data taken can be combined [7] . The probability of the Naive Bayes algorithm uses the equation [7]:

$$(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

Where:

- X = Data whose class is unknown
- H = Data hypothesis is class specific
- P(H|X) = Probability of hypothesis H based on condition X (a posteriori probability)
- P(H) = Probability of hypothesis H (prior probability)
- P(X|H) = Probability of X based on the condition in hypothesis H

- P(X) = Probability

The Decision Tree algorithm is used to build a personalized online learning system to help direct learners to appropriate educational resources, including documents, courses, sections, videos and more, based on students' knowledge evaluation [8]. A Decision Tree is a successive model that brings together a series of basic tests in an efficient and cohesive manner where numerical features are compared to a threshold value on each test [9] .

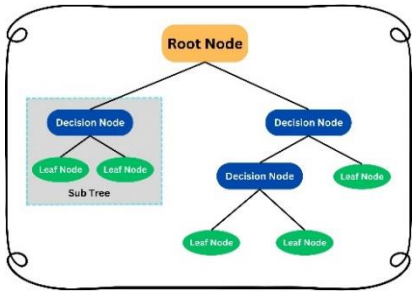

Figure 1. Decision Tree Diagram

The interesting thing about Decision Trees is the use of a tree structure like show on figure 1 that serves to represent the rules formed from the classification results. In a tree, attributes are represented by nodes, and classes are represented by leaves. Each tree has a root, which is the node at the top.

A machine learning model was developed with python and google collab, using Decision Tree and Random Forest algorithms. The model showed promising results with an accuracy rate of 69% for Decision Tree and 70% for Random Forest, demonstrating accuracy in predicting skill groups [10]. Another study on personalized music recommendation using Random Forest algorithm by applying data pre-processing, parameter adjustment, and learning index selection [11]. In medicine, the shift towards personalized medicine in cancer treatment has several challenges, for example, very high-dimensional data, which can lead to the problem of overfitting the classification model. To overcome this, two machine learning methods are used Random Forest (RF) and a hybrid model that combines RF and Support Vector Machine (RF-SVM) [12].

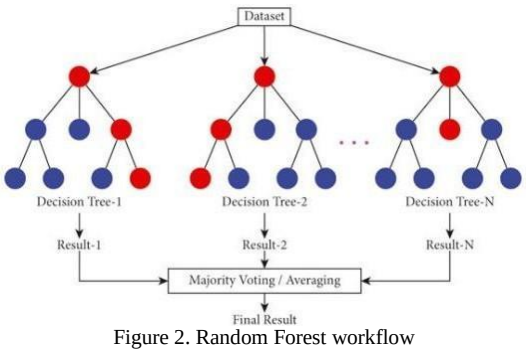

Figure 2. Random Forest workflow

The workings of the Random Forest algorithm like show in figure 2 can be described in the following steps:

1) The algorithm selects a random sample from the provided dataset.
2) Create a Decision Tree for each selected sample. Then the prediction results will be obtained from each Decision Tree that has been made.
3) A voting process is performed for each prediction result. For classification problems, the mode (most frequent value) will be used, while for regression problems, the mean (average value) will be used.
4) The algorithm will select the most voted prediction result as the final prediction.

In one study, the Support Vector Machine (SVM) algorithm produced effective recommendations with superior accuracy and recall compared to collaborative filtering algorithms [13]. Simultaneously, SVM has been successfully applied in stock prediction, giving an accuracy of about 60%-70% for simple SVM, which is further improved by incorporating methods such as Random Forest, genetic algorithm, and gives more accurate results [14]. SVM learning can be easier when using kernel tricks.

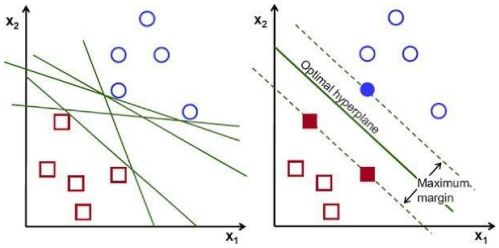

Figure 3. Support Vector Machine model

The Support Vector Machine algorithm show in figure 3 is used to find the best hyperplane in N-dimensional space that clearly classifies the data points. Hyperplane is a function that is used as a separator between one class and another. This function is used to classify in a higher dimensional class space.

K-Nearest Neighbor (KNN) is one of the classification methods in data mining that belongs to supervised learning. The application of the k-nearest neighbor machine learning algorithm to predict crime hotspots by using the distance between data to estimate the results, and the research findings show that this algorithm effectively identifies and understands crime patterns [15]. KNN method is done with a distance approach with the nearest neighbor [16]. In the KNN algorithm, the Euclidean Distance is used as a distance metric to determine the nearest neighbors of the point to be predicted using the formula:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{n}(a_r(x_i) - a_r(x_j))^2} \qquad (2)$$

Where $d\left(x_{i_j} x_j\right)$ is the euclidean distance, x_i and x_j are the two points to be calculated, with x_i being the ith data record and x_j being the jth data record, and a_r is the rth data with i,j being 1,2,3,...n.

The five algorithms above are suitable for testing with small datasets supported by repeated cross validation methods that are proven to increase the validity of the results [17].

2. RESEARCH METHODS

2.1. Data Description

This study uses 122 data of students in grade 10 (phase e) of Atisa Dipamkara senior high school who are currently in grade 11 (phase f) and have taken elective subjects according to the individual characteristics of each student based on the results of manual data processing. The source of the dataset is the report card data from grade 10 (phase e) with 18 columns of subject grades and ratio data type, data on the results of interest, aptitude, further study preferences, career tests consisting of 4 columns of nominal data type, and data on the results of manual processing of elective subject selection consisting of 4 columns of nominal data type. So that the size of the tested dataset is 3.172 data, which includes a small-sized dataset, so it is necessary to use the cross-validation method, which aims to optimize the dataset, generalize the model and validate the research results.

2.2. Algorithm Selection

In this study, five machine learning algorithms that are suitable for small datasets are used, namely the Naive Bayes algorithm (B Derick P et al, 2023) [6], Decision Tree (Kragbi Olivier, Tiemoman and Koffi, 2023) [8], Random Forest (Shu, Shen and Zeng, 2023) [11], Support Vector Machine (SVM) (Kurani et al., 2023) [14] and k-Nearest Neighbors (KNN) (K Vinothkumar et al., 2023) [15] which will be tested and compared for performance.

2.3. Algorithm Performance Evaluation

Weka 3.8.6. was chosen as the data mining tool used in this study to identify and predict relationships between variables in a dataset that can perform experiments with cross validation methods, such as data visualization and quantitative analysis that shows the results of comparing the performance tests of the five algorithms used.

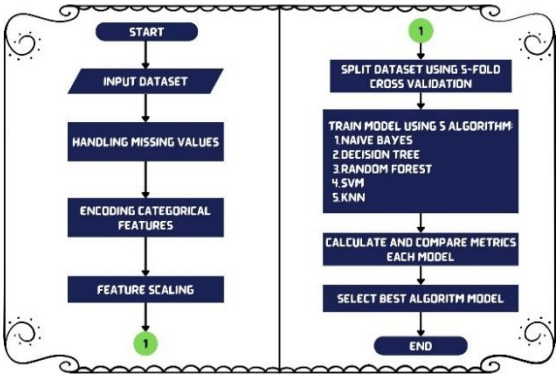

Figure 4. Research Flow

The research flow is explained using the flowchart in figure 4 which can be further described as follows:

1) Dataset input

The Atisa Dipamkara senior high school dataset containing 3.172 data was inputted into Weka.

2) Handling missing values

This step is carried out with the aim of checking and dealing with missing or empty values in the dataset, this is because machine learning models generally cannot handle missing values, therefore, it is necessary to manage missing values so that the model can work properly. The actions performed in this process are delete rows or columns containing missing values or fill the missing values with the mean, median, or mode or use imputation techniques or machine learning models to estimate the missing values.

3) Encoding Categorical Features

This step is to convert categorical variables into numerical form so that they can be incorporated into the machine learning model. This is a must because most machine learning algorithms require numerical input, not categorical variables. The actions performed in this process are one-hot encoding which is creating an additional column for some category and assigning a nominal value to each category and label encoding which is replacing each category with the corresponding nominal value.

4) Feature Scaling

This step is to re-scale the features so that they have similar or normal scales. This step must to do because some machine learning algorithms are very sensitive to scale differences between features. Scaling can help ensure that each feature has a balanced influence on the model. The actions taken in this process are min-max scaling which normalizes the feature values into a certain range, for example [0, 1] or z-score scaling (standardization) which converts the feature values into a standard normal distribution with a mean of zero and a standard deviation of one.

5) Split dataset into 5-fold cross validation

At this stage, the data set is divided into 5 parts, with 1 part of test data and 4 parts of training data, iterated repeatedly using the cross-validation method to improve the data performance so that it can continue to the next process, which is testing with different descriptive model algorithms. Determination of 5 folds is based on general perceptions related to the ideal number of folds, which is 5% of the total object data.

6) Train the model using 5 algorithms

Once the data is ready and cross-validated, the data is trained and tested using a pre-defined evaluation metric for five pre-built machine learning algorithms: Naive Bayes, Decision Tree, Random Forest, Support Vector Machines, and k-Nearest Neighbors.

7) Calculate and compare metrics for each model

After the dataset test using the five algorithms is complete, then an evaluation and comparison of the metric evaluation results of each model is carried out. The test results provided by:

a. Accuracy

Accuracy gives an idea of how well the model can predict overall.

b. Recall

Recall measures how well the model can identify all true positive instances. Recall is defined as the ratio of true positives to the total number of true positives.

c. Confusion Matrix

Confusion matrix is a table that shows the performance of the model by comparing the model predictions with the actual situation. In the confusion matrix there are the following components:

- True Positive (TP): the number of correctly predicted positive instances.
- True Negative (TN): the number of negative instances that were correctly predicted.
- False Positive (FP): the number of negative instances that were wrongly predicted as positive.
- False Negative (FN): the number of positive instances that were wrongly predicted as negative.

8) Select best algorithm model

Table 1. Metric Evaluation

No	Metrics Evaluation	Best Algorithm Model
1.	Accuracy	Highest
2.	Recall	Highest
3.	True Positive (TP)	Highest
4.	False Positive (FP)	Lowest

Confusion Matrix evaluation by referring to the general principle that the best model is the one with the highest TP and or the lowest FP.

Refer to table 1, the best machine learning algorithm ideally has the highest accuracy, recall and TP (true positive) results and or has the lowest FP (false positive). In the event of inconsistencies in the test results of each algorithm model, researchers will focus on the highest accuracy results as a basis for determining the best predictive model algorithm.

3. RESULTS AND DISCUSSION

3.1. Result

First, the dataset with 122 instances consisting of 18 attributes of subject grades, 4 attributes of interest and aptitude test results and 4 attributes of elective subjects formed in numeric data types was confirmed to have no missing values, then enter to Weka version 3.8.6 to perform the initial process of preparing the dataset. The initial process carried out is to change the numeric data type to nominal for 4 attributes of interest and aptitude test results and 4 attributes of elective subjects, and then standardization is carried out with a normal distribution.

Figure 5. Normalization Grade Distribution of Subjects

From figure 5, the visualization of data standardized with normal distribution, it is obtained that some subjects have a well distributed normal data distribution including B. Indonesia, physics, biology, sociology, economics, geography, PJOK, while some subjects appear to be negatively normally distributed including PAB, civics, computer science, music, fine arts and for mathematics subjects with positive normal distribution. In addition, from the data visualization above, it can be sure that 4 data on the results of the interest and aptitude test and 4 data on selected subjects with nominal data types are presented according to the proportion of the amount of data in each class variant formed.

Table 2. Label attributes with nominal data type

Label	MBTI	Advance Study and Career Recommendation	Elective Subjects
1	ENFJ	Administration and Human Resources	Advanced English
2	ENFP	Architecture and Construction	Advanced Chinese
3	ENTJ	Digital, Computer & Information Technology	Biology
4	ENTP	Law and Crime	Digital Business
5	ESFJ	Health Sciences	Visual Communication Design
6	ESFP	Science & Technology	Economics
7	ESTJ	Counseling and Psychology	Physics
8	ESTP	Management and Business	Computer Science
9	INFJ	Manufacturing	Chemistry
10	INFP	Media and Advertising	Advanced Mathematics
11	INTJ	Sports	Sociology
12	INTP	Marketing, Selling and Service	
13	ISFJ	Government and Public Administration	
14	ISFP	Banking & Finance	
15	ISTJ	Hospitality, Tourism and Events	
16	ISTP	Agriculture and Food	
17		Arts and Design	
18		Social Sciences and Humanities	
19		Natural Resources	
20		Transportation, Distribution and Logistics	
21		Education and Training	

Table 2 shows that for each attribute in the nominal data type is denoted by numbers so that the data can be processed by machine learning algorithms, however they are not in standardized form.

After the data is ready to be tested, the tests are performed with five machine learning algorithms those have been prepared, namely the Naive Bayes algorithm, Decision Tree, Random Forest, Support Vector Machine (SVM), and k-Nearest Neighbors (KNN), with parameter settings that produce maximum performance from each algorithm tested. Here we present the test results of the five machine learning algorithms.

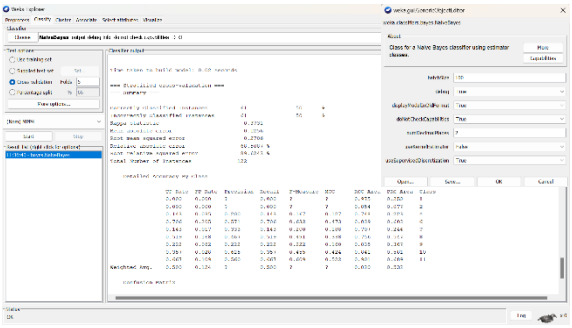

Figure 6. Testing Results with Naive Bayes Algorithm

The first test was conducted using the Naive Bayes algorithm as show in figure 6 by changing several parameters, namely debug, display model in old format, use supervised discretization, which were originally false to true. Changing debug and model format can help in the analysis and debugging process, while the use supervised discretization parameter can improve model accuracy. Discretization is the process of converting continuous variables into discrete variables or categories, which means that the division of data into categories is done by taking to account the class labels. This can help improve model performance by providing a better understanding of the distribution of feature values based on their class labels, which is necessary because the tested dataset, although in the small dataset category, has high complexity with 11 classes. From this test, the results obtained an accuracy value of 50%, recall value of 0.500, TP value of 0.500 and FP value of 0.124.

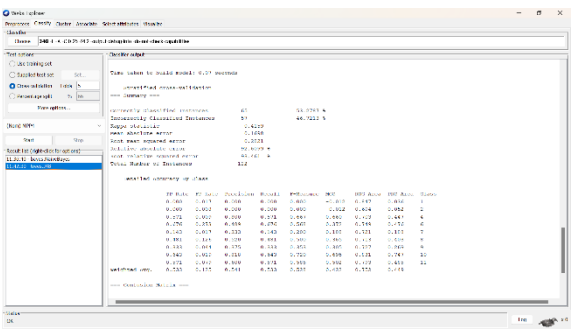

Figure 7. Testing Results with Decision Tree Algorithm

In the second test on the same data set, the Decision Tree algorithm was used. In Weka, the Decision Tree algorithm is known as J48. The name J48 really comes from the name of the previous algorithm called C4.5 that was developed by Ross Quinlan in 1993. Later, in Weka, an improved version of this algorithm was implemented and named J48. Some parameters are adjusted to get the best accuracy, including Debug, do not check capabilities, save instance data, use Laplace from false to true. Setting debug, do not check capabilities, and save instances to true is done for analysis and debugging in the modeling process, while the use Laplace parameter is set to true to be able to help overcome overfitting problems and improve model generalization, especially when the data has large variations such as the currently tested dataset, and can ultimately improve model accuracy. From this test as show in figure 7, the accuracy value is 53.2787%, the recall value is 0.533, the TP value is 0.533 and the FP value is 0.125.

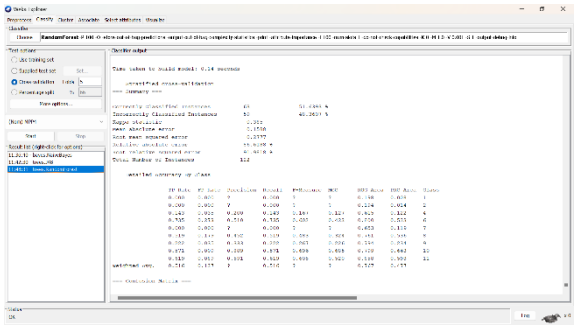

Figure 8. Testing Results with Random Forest Algorithm

The third test was performed using the Random Forest algorithm with adjustments to several parameters, including: compute out of bag, compute attribute importance, debug, do not check capabilities, output out of bag complexity statistics, print classifiers, save out of bag prediction from false to true. The purpose of setting some of these parameters is to provide additional information, enable more in-depth analysis, and assist in evaluating model performance. From this test as show in figure 8, the accuracy value is 51.6393%, recall value is 0.516, TP value is 0.516 and FP value is 0.137.

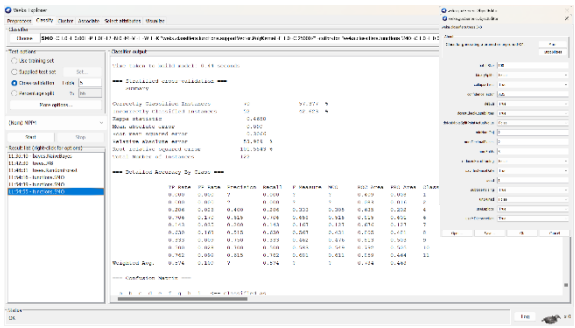

Figure 9. Testing Results with Support Vector Machine (SVM) Algorithm

The fourth test uses the Support Vector Machine (SVM) algorithm known as Sequential Minimal Optimization (SMO) because the implementation of the SVM algorithm in weka uses the sequential minimal optimization technique developed by John Platt in 1998. Some parameters are adjusted including: build calibration models, debug, do not check capabilities from false to true and calibrator parameters are changed from logistic to SMO. The build calibration models parameter is used to build calibration models to improve the prediction probability of the SVM model to approach the true probability. The debug parameter can be used to print additional debug information during the model building process, which helps in troubleshooting and further understanding how the algorithm works, thus aiding in analysis. Do not check capabilities parameter is set so that the dataset bypasses the algorithm's capabilities check because it uses a small dataset where the data has been previously verified and fits the algorithm's needs, so this setting can improve the algorithm's performance. SMO is a calibration model in which calibrator parameters are selected to improve the accuracy and reliability of prediction probabilities. SMO works by selecting a pair of dual variables to optimize, then improving one of those variables while holding the other fixed, and repeating the process until it meets convergence criteria. This test results as show in figure 9 have an accuracy value of 57.377%, a recall value of 0.574, a TP value of 0.574, and a FP value of 0.109.

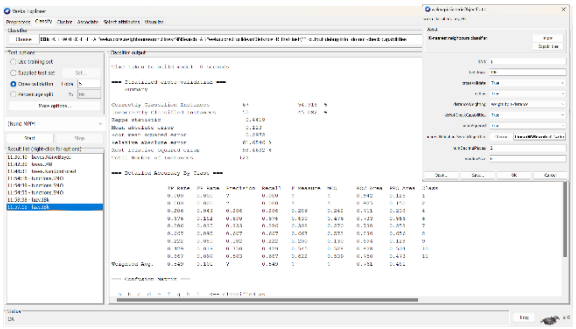

Figure 10. Testing Results with k-Nearest Neighbors (KNN) Algorithm

The final test uses the k-Nearest Neighbors (KNN) algorithm known as IBk, which stands for Instance-Based k. The term instance-based refers to the nature of the KNN algorithm, which uses instances (data points or samples) directly from the data set to make predictions. While k refers to the number of nearest neighbors used to make predictions. Some custom parameters include: cross-validate, debug, do not check capabilities, mean square of false to true and the distance weighting parameter is set using weight by 1-distance. The cross-validate setting aims to perform cross-validation when training the KNN model, which helps to avoid overfitting and provides a better estimate of model performance on unseen data. The mean squared parameter uses the mean squared error (MSE) metric to evaluate the performance of the model, which can provide a better estimate of performance. By setting the distance weighting parameter to weight by 1-distance, the weight for closer neighbors will be larger and the weight for farther neighbors will be smaller. This means that the influence of closer neighbors will be greater in determining the prediction. Overall, this setting can provide additional functionality and the possibility to improve the performance of the KNN model, especially in terms of evaluation and adaptation to data patterns. From the test results as show in figure 10, the accuracy value is 54.918%, the recall value is 0.549, the TP value is 0.549 and the FP value is 0.101.

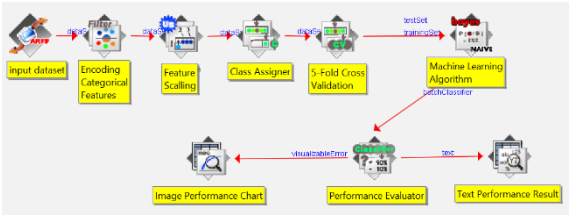

Figure 11. Knowledge Flow Experiment in WEKA

The knowledge flow diagram show in figure 11 explains the work process of the test performed, starting with the dataset input process, then the codification process is performed to label attributes of nominal data type so that they can be processed by the system, after which the standardization process is performed with a certain scale so that all data are distributed in the normal distribution range. The next process is class preparation for 5-way cross validation. After the data is ready, testing is performed by adjusting the parameters using 5 types of algorithms, namely Naive Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM) and k-Nearest Neighbors (KNN). The performance results of each algorithm are then presented in the form of text and graphs. Data on the performance results of testing each algorithm is then used to draw conclusions about the best performing algorithm based on predetermined evaluation metrics.

From the tests carried out on five machine learning algorithms using the Atisa Dipamkara senior high school elective subject in small dataset, the following results are obtained:

Table 3. Performance Comparison of Five Machine Learning Algorithms

No	Algorithm	Accuracy *Highest*	Recall *Highest*	TP *Highest*	FP *Lowest*
1.	Naive Bayes	50.0000 %	0.500	0.500	0.124
2.	Decision Tree	53.2787 %	0.533	0.533	0.125
3.	Random Forest	51.6393 %	0.516	0.516	0.137
4.	**Support Vector Machine (SVM)**	**57.3770 %**	**0.574**	**0.574**	0.109
5.	k-Nearest Neighbors (KNN)	54.9180 %	0.549	0.549	**0.101**

The results of testing as show in table 3, the Atisa Dipamkara senior high school elective subject dataset using five types of machine learning algorithms show variations in the performance of each algorithm. The Support Vector Machine (SVM) algorithm performed best with the highest accuracy of 57.3770 %, also had the highest recall and true positive (TP), both at 0.574. These results show that SVM is capable of classify the dataset accurately and efficiently, and can identify the true positive class well. However, it should be noted that SVM has complexity in parameter tuning and requires longer computation time.

In addition, KNN performs quite well with an accuracy of 54.9180% and a recall of 0.549. KNN has the advantage of easy implementation and does not require much parameter tuning. However, KNN is sensitive to the size and dimension of the data and requires a large memory for its training data. Decision Tree and Random Forest perform quite well with accuracies of 53.2787% and 51.6393%, respectively. However, the resulting models tend to be difficult to interpretation, especially for complex models.

Meanwhile, Naive Bayes achieved an accuracy of 50.0000% and a recall of 0.500. Naive Bayes is known for its simplicity and speed of implementation, but it is often considered too simple and unable to handle dependencies between features in a data set. The results of this test highlight the importance of selecting an algorithm that matches the characteristics of the dataset and the needs of the project. When testing the Atisa Dipamkara senior high school elective subject dataset, SVM emerged as the top choice due to its better performance compared to other algorithms.

3.2. Discussion

The experimental comparison of machine learning algorithms for optimizing the selection of elective subjects in Phase F of the Merdeka Curriculum yielded insightful results, revealing different performances among the tested algorithms. The discussion aims to explore the implications of these results, provide insights into the suitability of each algorithm, and compare them with similar studies cited in the references.

The performance analysis showed that SVM was the best performing algorithm in terms of accuracy, recall, and true positive rate. Despite its computational complexity and parameter tuning requirements, SVM demonstrated superior classification capability on the Atisa Dipamkara Senior High School dataset. This is consistent with the findings of studies (Kurani et al., 2023) [14], which conducted a comparative study between Artificial Neural Networks (ANN) and SVM and confirmed the effectiveness of SVM in various classification tasks.

Each algorithm had distinct advantages and limitations. SVM's robust performance underscores its suitability for datasets with complex decision boundaries, despite its computational overhead. Conversely, Naive Bayes, while simple and fast, struggled to handle feature dependencies, which affected its performance. This is consistent with the findings of Tejawati et al [16], which emphasize the importance of algorithm selection based on dataset characteristics.

Comparison with similar studies provides valuable insights. For example, Shu et al. [11] applied Random Forest in a music recommendation system, demonstrating its versatility beyond classification tasks. In addition, studies such as Prayoga Permana et al. [5] and Charbuty et al. [9] highlighted the effectiveness of decision trees in predictive modeling, albeit with interpretability challenges similar to those observed in this research.

The implications of algorithmic performance extend to educational contexts, particularly in curriculum optimization. The ability to accurately predict elective choices can streamline educational pathways for students, improving their academic journey. By leveraging machine learning algorithms, educators can gain valuable insights into student preferences and tailor curriculum offerings accordingly, in line with the goals of the ***Merdeka Curriculum***.

Future research could explore ensemble techniques or hybrid models to further improve predictive accuracy while addressing interpretability concerns. In addition, investigating the impact of additional features or refining existing feature engineering methods may provide deeper insights into students' decision-making processes.

In conclusion, the experimental comparison of machine learning algorithms provides valuable insights into their applicability in optimizing the selection of electives within the Merdeka Curriculum. While SVM emerged as the top performer in this study, the discussion highlights the nuanced considerations in algorithm selection, dataset characteristics, and implications for educational practice. These findings contribute to the broader discourse on the use of machine learning in educational decision making.

4. CONCLUSION

When making decisions about the performance results of machine learning algorithms on a dataset, it is important to consider the evaluation indicators of the previously set metrics. In this research, the evaluation metrics used are accuracy with the highest value, recall with the highest value, and confusion matrix with respect to the highest true positive (TP) and/or the lowest false positive (FP) values. Accuracy gives an idea of how well the model can predict overall. Recall measures how well the model can identify all true positives. Recall is defined as the ratio of true positives to the total number of instances that are actual positive. Confusion Matrix is a table

that shows the performance of the model by comparing the model's predictions with the actual situation, where true positive (TP) is the number of positive instances that are correctly predicted and false positive (FP) is the number of negative instances that are incorrectly predicted as positive.

From the results of testing 5 different machine learning algorithms on the Atisa Dipamkara senior high school elective subject dataset, the Support Vector Machine (SVM) algorithm has the best performance results by achieving the highest accuracy of 57.3770 %, the highest recall of 0.574, and the highest true positive (TP) of 0.574, compared to other algorithms, which means the best in 3 out of the 4 indicators set from the evaluation metrics.

This result show that SVM has a better ability to classify the Atisa Dipamkara senior high school elective subject dataset accurately and efficiently. This performance is due to SVM's ability to find optimal linear or non-linear separators between different classes in the feature space, especially in complex and non-linear datasets such as the tested dataset which it has 11 classes with 26 attributes consisting of 18 attributes of numeric data type and 8 attributes of nominal data type, because of that the best accuracy result only in the range of 57%.

Although SVMs perform better in terms of accuracy and recall, it should be noted that SVMs also have some drawbacks. One of the main disadvantages of SVM is the complexity of parameter tuning. SVM has several parameters that need to be tuned appropriately, such as calibrator and kernel parameters. Errors in tuning these parameters can lead to poor performance or overfitting. In addition, SVM requires more computation time than other algorithms due to the complex kernel calculations and optimizations performed by the algorithm.

On the other hand, the k-Nearest Neighbors (KNN) algorithm, which ranked second in this test, is easy to implement and does not require much parameter tuning. However, KNN tends to be sensitive to the size and dimension of the data and requires a large amount of memory for its training data. Decision Tree and Random Forest perform reasonably well, but tend to produce models that are difficult to interpret, especially for complex models. Naive Bayes, on the other hand, is easy to implement and fast, but is often considered too simple and cannot handle dependencies between features in the data set.

Thus, when evaluating the best algorithm for the Atisa Dipamkara senior high school elective subject dataset, SVM emerged as the top choice due to its best performance in terms of accuracy, recall, and TP.

5. FUTURE WORK AND SUGGESTION

The results of testing the dataset of selected subjects from Atisa Dipamkara senior high school using five different machine learning algorithms showed that the performance of the algorithms was still not optimal, with the highest accuracy level only around 57%. An algorithm is said to perform well if it has at least 90% accuracy. From the test results, there are several thoughts that can be considered for future research:

a. Multiclass approach: Using this approach can help the algorithm to classify the dataset better, as it can predict the classes of more than two categories that may exist in the dataset, thus improving the prediction accuracy.
b. Deep learning approaches: Deep learning approaches such as neural networks can help in handling complex patterns in the data set. This is useful for extracting more abstract and complex features from the data, which may be difficult for traditional machine learning algorithms to understand.
c. Data simplification: Reducing the number of attributes and classes in a dataset can make the data easier for algorithms to process. This can be done by identifying the most relevant attributes and removing unimportant or redundant attributes.
d. Adding data instances: Increasing the number of data instances can be done by collecting more samples or using data augmentation techniques. This allows the model to look at more instances and learn from more variation in the data.
e. Use other algorithms: Explore the use of other machine learning algorithms that are more effective for this data set, such as the ensemble learning algorithm Gradient Boosting or XGBoost, to improve the predictive performance of the model. It works by combining the strengths of several weak model algorithms. Ensemble learning allows us to generate more robust and accurate models that can be used in various machine learning applications.
f. More comprehensive model validation: Use more comprehensive model validation techniques, such as bootstrapping, to test the reliability and generalizability of the model.
g. Model interpretation: Perform deeper model interpretation analysis to understand what factors influence the model's predictions and how the model makes decisions.
h. Interdisciplinary Collaboration: Collaborate with experts from various fields such as education, statistics, and computer science to obtain different viewpoints and more innovative solutions.

However, the results of this study can be further developed in further research, namely the development of machine learning for the selection of phase f elective subjects at Atisa Dipamkara senior high school using the Support Vector Machine (SVM) algorithm.

REFERENCES

[1] S. Ariga, “Implementasi Kurikulum Merdeka

Pasca Pandemi Covid-19," *EDU Soc. J. PENDIDIKAN, ILMU Sos. DAN Pengabdi. Kpd. Masy.*, vol. 2, no. 2, pp. 662–670, Aug. 2023, doi: 10.56832/edu.v2i2.225.

[2] M. D. R. Adhyatma, "Program Pendampingan Integrasi Bakat dan Minat dalam Memilih Mata Pelajaran sebagai Implementasi Kurikulum Merdeka Tingkat Sekolah Menengah Atas (SMA)," *Bubungan Tinggi J. Pengabdi. Masy.*, vol. 5, no. 4, p. 1651, Dec. 2023, doi: 10.20527/btjpm.v5i4.9572.

[3] L. LÓPEZ-GARCÍA, C. LINO-RAMÍREZ, V. M. ZAMUDIO-RODRÍGUEZ, and J. DEL VALLE- HERNÁNDEZ, "Predictive model for the analysis of academic performance and preventing student dropout using machine learning techniques," *Rev. Educ. Técnica*, pp. 1–5, Nov. 2022, doi: 10.35429/JOTE.2022.16.6.1.5.

[4] J. Rust and S. Autexier, "Causal Inference for Personalized Treatment Effect Estimation for given Machine Learning Models," in *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*, IEEE, Dec. 2022, pp. 1289–1295. doi: 10.1109/ICMLA55696.2022.00206.

[5] A. Prayoga Permana, K. Ainiyah, and K. Fahmi Hayati Holle, "Analisis Perbandingan Algoritma Decision Tree, kNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up," 2021. [Online]. Available: https://www.kaggle.com/manishkc06/startup-success-prediction.

[6] D. P. B, M. K, V. K, and Y. D, "Job and Course Recommendation System using Collaborative Filtering and Naive Bayes algorithms," in *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, IEEE, Jun. 2023, pp. 1–4. doi: 10.1109/ICAECA56562.2023.10200758.

[7] A. H. Wicaksono, A. A. Supianto, S. H. Wijoyo, D. Krisnandi, and A. Heryana, "KLASIFIKASI SISWA SLOW LEARNER UNTUK MENDUKUNG SEKOLAH DALAM MENINGKATKAN PEMAHAMAN SISWA MENGGUNAKAN ALGORITMA NAÏVE BAYES," vol. 9, no. 3, pp. 589–596, 2022, doi: 10.25126/jtiik.202295509.

[8] P. Kragbi Olivier, K. Tiemoman, and K. Koffi, "MODEL OF A PERSONALISED E-LEARNING PROCESS BASED ON A DECISION TREE ALGORITHM," *Int. J. Adv. Res.*, vol. 11, no. 11, pp. 371–382, Nov. 2023, doi: 10.21474/IJAR01/17839.

[9] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

[10] A. B. Wiratman and W. Wella, "Personalized Learning Models Using Decision Tree and Random Forest Algorithms in Telecommunication Company," *JOIV Int. J. Informatics Vis.*, vol. 8, no. 1, Mar. 2024, doi: 10.62527/joiv.8.1.1905.

[11] Z. Shu, Q. Shen, and T. Zeng, "Research and Application of Music Personalized Recommendation System Based on Random Forest Algorithm," in *2023 IEEE 6th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC)*, IEEE, Feb. 2023, pp. 1387–1391. doi: 10.1109/ITNEC56291.2023.10082503.

[12] W. Alsahly, K. Poterlowicz, and H. Aldoghaither, "Using Random Forest and a Hybrid model to Predict Drug Response in Skin Cutaneous Melanoma," in *2023 IEEE International Biomedical Instrumentation and Technology Conference (IBITeC)*, IEEE, Nov. 2023, pp. 46–51. doi: 10.1109/IBITeC59006.2023.10390944.

[13] L. Lin, "Learning information recommendation based on text vector model and support vector machine," *J. Intell. Fuzzy Syst.*, vol. 40, no. 2, pp. 2445–2455, Feb. 2021, doi: 10.3233/JIFS-189239.

[14] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A Comprehensive Comparative Study of Artificial Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting," *Ann. Data Sci.*, vol. 10, no. 1, pp. 183–208, Feb. 2023, doi: 10.1007/s40745-021-00344-x.

[15] V. K, R. K. S, V. R. R, N. Mekala, S. P. Sasirekha, and R. Reshma, "Predicting High-Risk Areas for Crime Hotspot Using Hybrid KNN Machine Learning Framework," in *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, IEEE, Aug. 2023, pp. 848–852. doi: 10.1109/ICIRCA57980.2023.10220738.

[16] A. Tejawati, A. Septiarini, R. Rismawati, and N. Puspitasari, "COMPARISON OF K-NEAREST NEIGHBOR AND NAIVE BAYES METHODS FOR CLASSIFICATION OF NEWS CONTENT," *J. Tek. Inform.*, vol. 4, no. 2, pp. 401–412, Mar. 2023, doi: 10.52436/1.jutif.2023.4.2.676.

[17] C.-C. Chang, C.-H. Chen, J.-G. Hsieh, and J.-

H. Jeng, "Iterated cross validation method for prediction of survival in diffuse large B-cell lymphoma for small size dataset," *Sci. Rep.*, vol. 13, no. 1, p. 1438, Jan. 2023, doi: 10.1038/s41598-023-28394-6.