Journal of Applied Data Sciences
Vol. 4, No. 3, September 2023, pp. 264-275

Ensemble learning techniques to improve the accuracy of predictive model performance in the scholarship selection process

Nurhayati Buslim [1,*], Zulfiandri [2], Kyungoh Lee [3]

[1,2] Department of Informatics, Faculty of Science and Technology, Syarif Hidayatullah State Islamic University Jakarta, Indonesia
[3] Computer Engineering, Sun Moon University Korea, South Korea
[1] nurhayati@uinjkt.ac.id *; [2] zulfiandri@uinjkt.ac@uinjkt.ac.id; [3] leeko@sunmoon.ac.kr
* corresponding author

(Received July 18, 2023; Revised August 30, 2023; Accepted September 10, 2023; Available online September 21, 2023)

Abstract

Scholarship selection with big volumes of college student data in an university undoubtedly required a lot of resources and time. Besides the inefficient factor, there are also human-error occurred in the scholarship selection process. Error and risk can be reduced with ensemble learning approach. The different with another method is that usually research will only choose one algorithm or doing comparison to search the best algorithm. But in ensemble learning, some of algorithms called base learner combined to shape a new more-established model. With ensemble, more accurate result of the scholarship selection produced and also had the most minimum error value. In this research, there are two algorithm used as the base learner which are K-Nearest Neighbor (KNN) and Support Vector Machine (SVM). Experiment showed result of KNN with 91% of accuracy and SVM with at the rate of 75% accuracy. This base learner combined into an ensemble learning model using voting classifier. After last experiment, ensemble learning model succesfully created and produced the scholarship prediction result up to 100% of accuracy on training data. With streamlit, an application has been made which can automatically determine whether a student is accepted or rejected in the scholarship selection process. From the result, ensemble learning model can be used by academics as a Decision Support System (DSS) for determining scholarship recipients. This model can also be used as a development for institutions in the academic field.

Keywords: Ensemble Learning, Linear Regression, Exponential Smoothing, Prediction Model

1. Introduction

Processing large amounts of data, such as student data surely need many resources: human resources, cost, and time [1]. Human-error cases were also found where sometimes the data isn't efficient enough and many error occurred. Human errors in making decisions can be suppressed by using certain technologies, which in this case Machine Learning algorithm approach used to do the prediction process. In the prediction process itself, of course, there will be also error value in process. Because of this there will be some prediction inaccuracies and decreasing performance of the data. To encounter this problem, the ensemble technique can be used as one of method to reduce the error of prediction data [2]. With the ensemble learning mechanism, data performance increased and can get more accurate prediction result [3].

There are large of number scholarship program for each department in an educational institution. Each program has its own mechanism and found that each process taking lot of times due to the large amount of data received from the student. There are requirements for a system that is capable of determining scholarship recipients in order to improve the academic system. This research intend to obtain a scholarship recipient selection model with high performance and high accuracy using Ensemble Learning Technology. The results of this study can be used to create a decision-making system to quickly and accurately process and monitor scholarship recipients in an institution, especially educational institution.

Ensemble learning is a model that combines predictions from two or more algorithm models [4]. The models that contribute to the ensemble, referred to as base learner, may be consist of the same type or different types and also may or may not be trained on the same training data. The Predictions made by ensemble members can be combined

using statistics, or by more sophisticated methods. By combining some algorithms as a base, ensemble has a aim to achieve result with better performane and moreover the accuracy itself [5].

2. Related Works

Explaining Several past studies have made a scholarship prediction using K-Nearest Neighbor (KNN) and Support Vector Machine (SVM). The first related work is “Using KNN Algorithms for Determining the Recipient of Smart Indonesia Scholarship Program” by Purwanto and Dadang Syarif Sihabudin Sahid. The authors made a prediction for a scholarship selection program named Smart Indonesia using KNN and after that analyzed the accuracy performance. The result shows that the model has an accuracy rate of 66.79% which is quite good for datasets with high variance [6].

Wirawati Dewi Ahmad and Azuraliza Abu Bakar compared and analyzed the performance of the K-Means and K-Medoids algorithms in the scholarship selection process in a study titled "Classification Models for Higher Learning Scholarship Award Decisions". In this study, they used five algorithms: Naive Bayes, SVM, J48, Random Tree (RT), and finally ANN (Artificial Neuron Network). They used 10-fold cross-validation and split the dataset 90% into the training dataset and 10% into the test dataset. As a result, it was found that the SVM model has the highest accuracy compared to other models because it shows an accuracy of 86.9% [7].

From these studies, even though the algorithm shows a good accuracy but model performance still can be improved the model by joining these 2 models. There’s research from Ujjwal Pasupulety, Aiman Abdullah Anees, Subham Anmol, and Biju R Mohan who predict stock prices using ensemble learning named “Predicting Stock prices using Ensemble Learning and Sentiment Analysis”. To make the ensemble model they combine 2 models, which are KNN and Random Forest. They found out that the ensemble has better performance on a dataset spanning 5 months [8].

Another research that predicts using ensemble was performed by Amit Kumar Patra, Ratula Ray, Azian Azamimi Abdullah, and Satya Ranjan entitled “Prediction of Parkinson's disease using Ensemble Machine Learning classification from acoustic analysis”. They did research to understand whether patients have Parkinson's Disease or not using some base classifiers such as decision tree, logistic regression, and K-Nearest Neighbor. Their study shows that the ensemble model has the best accuracy compared to the base classifier in most of the cases [9]. In conclusion, various studied proved that ensemble can be used to increase accuracy of the model and also the performance quality.

3. Method

3.1. Support Vector Machine (SVM)

Support Vector Machine (SVM) is one of the algorithms that is widely used to do the classification process that works by finding the hyperplane with the largest margin value [10]. This algorithm was first introduced by Boser, Vapnik, and Guyon in 1992. SVM is supervised learning [11] where before being able to do the classification process this model needs to be taught first through the existing data so later when asked to do the classification it can see a return to previously trained data for decision consideration [12].

3.2. K-Nearest Neighbor (KNN)

K-Nearest Neighbor or better known as KNN is one of the supervised machine learning [13] where the modeling is first done by teaching and training the model with the existing data. In its implementation, KNN is widely used to solve classification and regression problems. KNN is considered as a very simple classification method to implement in classifying [14][15] because it is based on the closest distance to its neighbors or the k variable [16]. The value of the k variable means the number of nearest neighbors that will be used to predict the value of the point in the future. For example, suppose the value of n_neighbors is 3, then this model will use the three closest neighbors to predict the value [17].

3.3. Ensemble Learning

Ensemble Learning is an algorithm to search for the best predictive solution compared to other algorithms because this method uses several learning algorithms to achieve better predictive solutions. In ensemble learning, there are several classifier methods such as adaptive boosting-based [18], voting-based, and bagging-based.

In this study voting method, which is one of the easiest ensemble learning techniques used. The method starts by creating two or more separate models with the same dataset. Then the Voting-based Ensemble model can be used to wrap the previous model and perform aggregation of Predictions created by sub-models and can be weight [19][20]. Streamlit, which is currently a popular tool for deploying machine learning methods also used for application implementation. In Streamlit, HTML code can be used within a python file and there is no need to create another template [21].

3.4. Data Collection and Analysis

To obtain data on UIN students who received scholarships, an online questionnaire was conducted using Google Forms and used. In the data input collection stage through the online survey, a total of 106 data were derived. Respondents are students from various existing majors who have received or are currently receiving both on-campus and off-campus scholarships at UIN Jakarta.

The process of conducting a survey is carried out by first determining the necessary scholarship variables and grouping them into questions. Features included in this study are supporting variables for personal data such as date of birth, campus, faculty, department, single tuition (UKT), and achievement. For the label, it can also be referred to as the output of the model, in this case, a state variable indicating whether the student is eligible for a scholarship. After dividing into features and labels, the data can be divided into training data and test data.

3.5. Research Flow

Before doing the training process, first flow of the research must be knowed. Here's given the flow chart to give a brief explanation of step by step that must be done to complete the research.

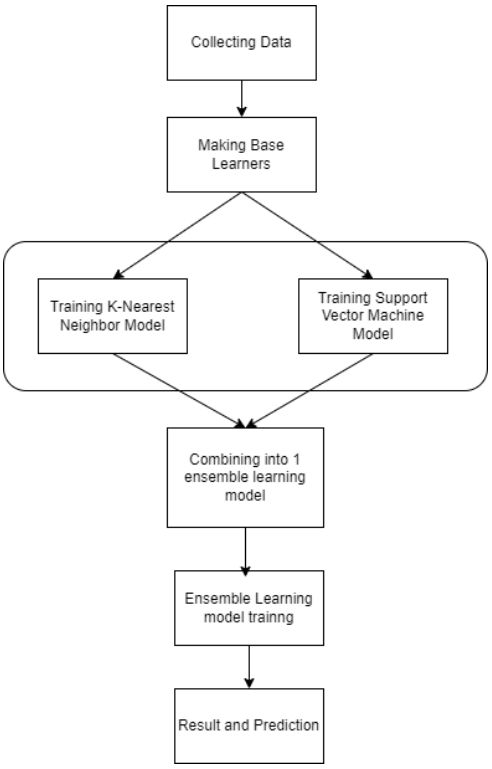


Figure. 1. Flow of the research

From figure 1, in general there are 3 main process: collecting data, training the data, and lastly predict the data using model that have been made. Previous section already explain how data gathered by using questionnaire. From the questionnaire there are some process to clean data: erasing redundancy data, fixing typos, and convert it into a csv file. Training process will be process after all the data preprocessing step finished. After that research began to make the base learner: KNN and SVM. Base learners don't have to reach 90% accuracy because base learner will be combined later, but having a good base learners are preferable. proceed to the next step: combining the base learners into 1 ensemble learning model and then train it to get better performance and accuracy. After ensemble learning model created existing model ready to predict data, whether it is from the data test or input new data.

3.6. UML Diagram

To facilitate the system design process, relationships between actors and systems need to be analyzed. UML diagrams are needed to document and define the system [22]. There are actually thirteen diagrams [23], but in this research two types of diagrams used: structural diagrams, which focus more on the structure of the system, and behavioral diagrams, or called as activity diagram which focus on showing the system flow [24], represent the activities user can do within the system. The following list shows examples of both use case diagrams and activity diagrams.

In use case diagram, there are 3 main component used: use case, actor, and the relation itself [25]. From use case diagrams, there are informations about the behavior of the system and the activities that actors in system can perform. Here are the use case diagram for the ensemble systems:

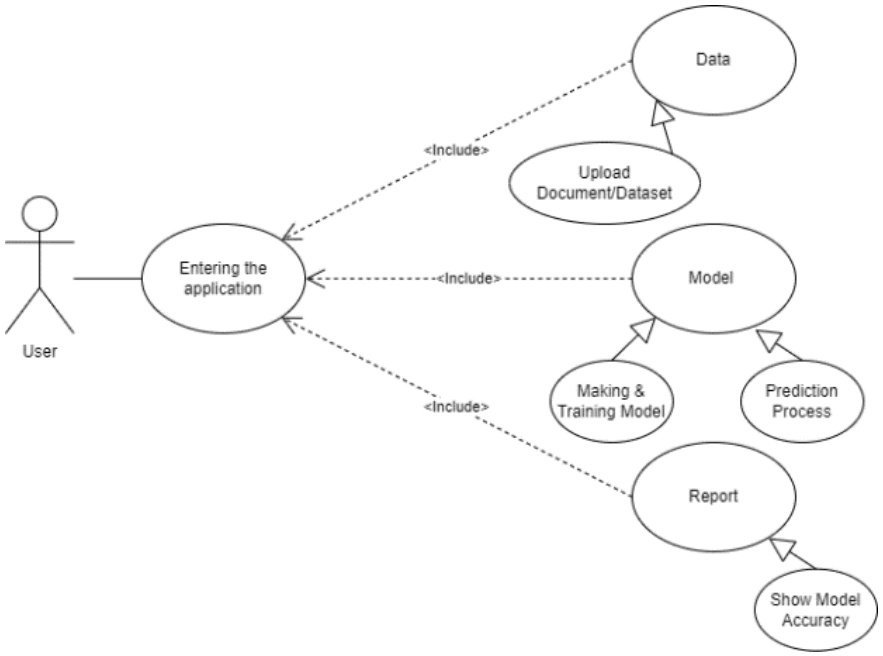


Figure. 2. Use Case Diagram of System

In order to make it easier for user to predict, there are also UI application for prediction. Figure 2 showed by using the application user can simply upload the dataset and then system will automatically generate the prediction result. After entering the application there are 3 main menu user can use: Data, Model, and Report. In data section user can upload the document which in this case the scholarship selection document. In model section out system will train the data and then showing the result, whether the person is accepted or not for the scholarship. In report section, user can see the model performance. If the model isn't reached the desirable accuracy, user can repeat the training process.

There are also activity diagram serves to show the steps and system activities from the beginning to the end of the activity [26]. In this diagram, the activity between the actors and the system can be seen.

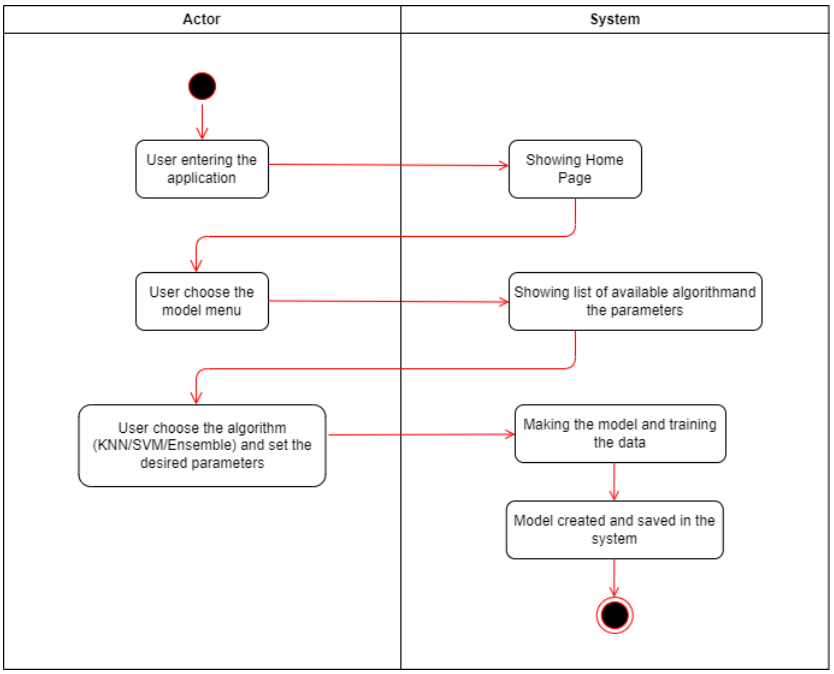


Figure. 3. Activity Diagram – Making Model Process

Based on Figure 2, it can be seen that the process of creating a model is performed by the user. When the user enters the system, the user can go directly to the model menu. After that, the system will display a list of available methods and parameter options. The system then displays a list of available methods and parameter options. The user selects a requirement and the system creates a model directly. The generated model is stored in the system and can be used later to predict the scholarship acceptance process.

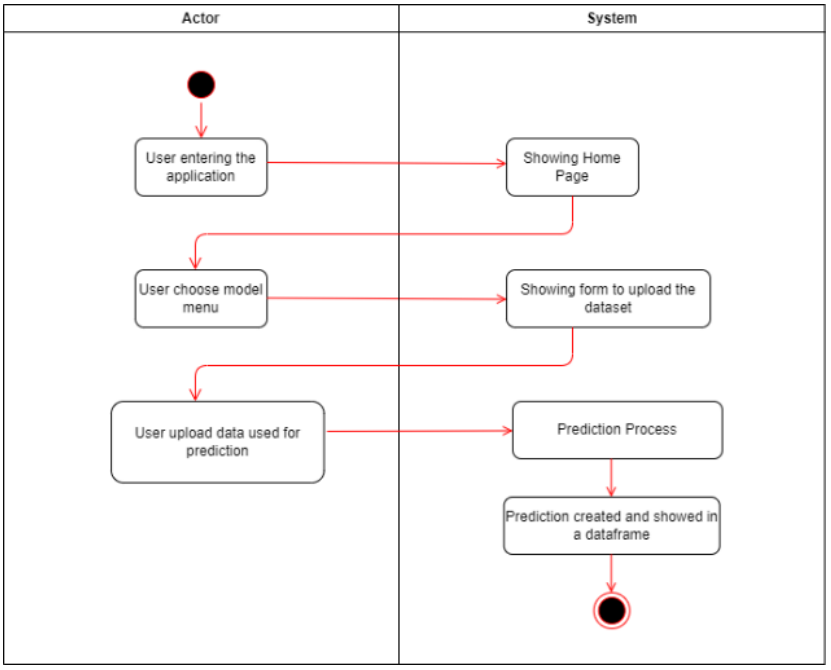


Figure. 4. Activity Diagram – Prediction Process

Figure 3 provides information about the forecasting process. After the user uploads data to train the generated model, the user can upload the data set to make predictions. After that, the system will perform the prediction process and display the prediction results for the student's scholarship receipt displayed in the data frame.

4. System Implementation Model

4.1. Support Vector Machine Model

The next step is the development of an algorithm with a generative model of the SVM itself. Classifier generation is done using a package from "scikit-learn". This study uses default parameters/linearities. In very complex cases, such as predicting scholarship recipients, an additional parameter called C is used to avoid noise [15]. A classifier is trained to generate a model. Model training using the training data. Once the model is formed, the prediction process can be performed with the test data.

4.2. K-Nearest Neighbor Model

In building the KNN algorithm to apply the model used in "scikit-learn" [24], the first step begins with determining the hyperparameters and creating the correct class model. For the kNN algorithm, the "scikit-learn" implementation must choose k values called n_neighbors. For the fabrication of ensemble learning, multiple models needed to run 3 experiments with different n_neighbors values in KNN, choosing 2, 3, 4 k values.

The generated model is stored in the KNN variable and ".fit()" allows the model to learn from the data. The next step is prediction, which predicts whether students will accept their scholarship status. In data preprocessing, predictions are made using previously shared test data. Predictions can be made using the KNN.predict() command, which can be fed into the model variable KNN.

4.3. Ensemble Learning Model

The model used as the basis for ensemble learning is a combination of SVM and KNN algorithms. Next, choosing the ensemble learning method - Voting. First, the voting-based ensemble model is imported using Python's "sklearn" package, and voting is provided through the VotingRegressor and VotingClassifier classes. VotingClassifier is assigned to a variable called "clf" which contains the list of estimators and the type of vote. The estimator is later followed by a list of models in the form of a list containing several models of previously generated SVMs and KNNs.

The original SVM algorithm does not predict probabilities but can be configured to predict a score equal to probability by setting the "probability" argument to "True" in the SVC class. Because soft voting does the sum of predicted probabilities (or scores equal to probabilities) for each class label and produces a final prediction, it predicts the label with the greatest probability [27]. When predicting whether a scholarship student will pass or fail, if the predicted value is greater than 0.5, it is rounded to 1 (eg, if it is less than 0.5, it is rounded to 0 and rejected). The prediction results are stored in an array created using the NumPy package.

5. Experimental Result

Simulations are performed by training the previously partitioned test data using the Support Vector Machine algorithm and the K-Nearest Neighbor algorithm. After three tests, it is finally integrated into an ensemble learning method and use the trained data to create a report model.

Table. 1. SVM Model Test Result (C= 0.4)

Number	Support Vector Machine Model Performance				
		Precision	Recall	F1-Score	Support
1	0	0.86	1.00	0.92	12
2	1	1.00	0.80	0.89	10
3	Accuracy			0.91	22

4	Macro average	0.93	0.90	0.91	22
5	Weighted average	0.92	0.91	0.91	22

From table 1, the accuracy of SVM is 91%. SVM has an overall good performance because all of the parameters which are precision, recall, and, F1-Score also have good score. For example the precision can be up to 100%, same as recall. However, the score especially F1-Score still can be improved. Because of this, SVM can be a good base learner.

Table. 2. KNN Model Test Result (k= 4)

Number	**K-Nearest Neighbor Model Performance**				
		Precision	**Recall**	**F1-Score**	**Support**
1	0	0.71	1.00	0.83	12
2	1	1.00	0.50	0.67	10
3	Accuracy			0.77	22
4	Macro average	0.85	0.75	0.75	22
5	Weighted average	0.84	0.77	0.75	22

From table 2, Performance score of KNN isn't really good. Even thought it has 83% in one of F1-Score but the accuracy is only 77%. The score worsen in recall because it only has 50% of score. If KNN is the only algorithm used in this research, then the prediction result won't be good. Because of this, KNN will be combined with SVM to make a new ensemble model to create better model with better performance in general.

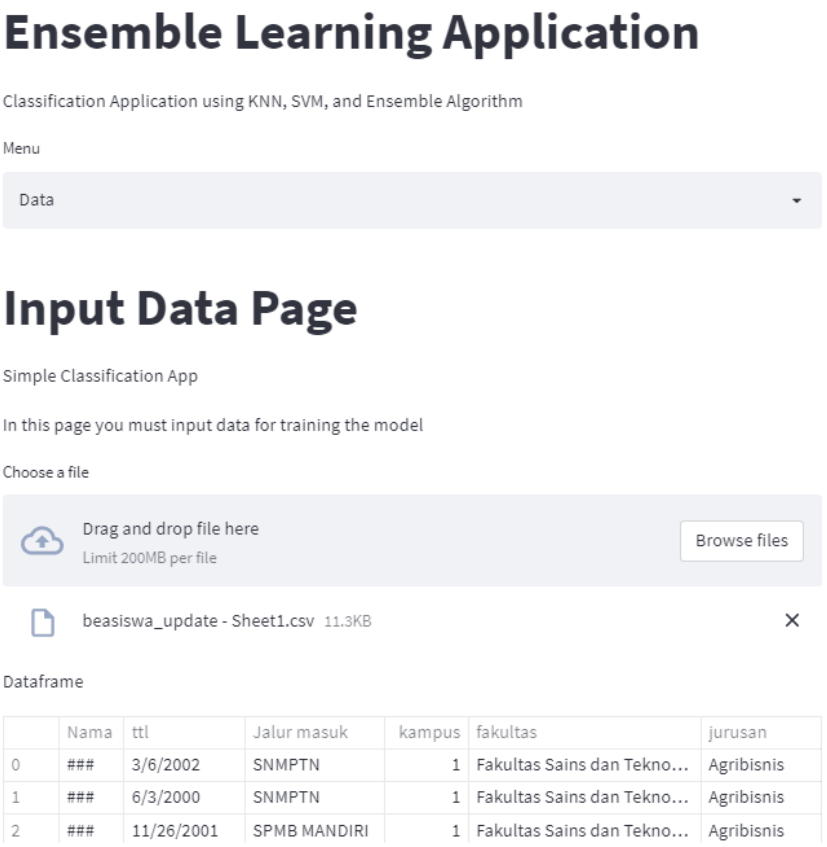
Table. 3. Ensemble Model Test Result

Number	**Ensemble Learning Model Performance**				
		Precision	**Recall**	**F1-Score**	**Support**
1	0	1.00	1.00	1.00	12
2	1	1.00	1.00	1.00	10
3	Accuracy			1.00	
4	Macro average	1.00	1.00	1.00	22
5	Weighted average	1.00	1.00	1.00	22

ISSN 2723-6471

From Table 3, the ensemble learning method itself, has a very high level of accuracy, which is 100% in this case. This indicates that the ensemble learning method is very good and can be used to predict students' scholarship acceptance decisions. After completing algorithm experiment, web-based application can be made which can accurately predict scholarship recipients using the "Streamlit" tool, and applied various machine learning models to obtain accurate results.

Application consists of 4 menus: Home, Data, Model and Report, each with its own unique function. The Home menu is the main menu that links to the main page of the application and is displayed via a drop-down of the full menu list.



	Nama	ttl	Jalur masuk	kampus	fakultas	jurusan
0	###	3/6/2002	SNMPTN	1	Fakultas Sains dan Tekno...	Agribisnis
1	###	6/3/2000	SNMPTN	1	Fakultas Sains dan Tekno...	Agribisnis
2	###	11/26/2001	SPMB MANDIRI	1	Fakultas Sains dan Tekno...	Agribisnis

Figure. 5. Application GUI - Input Process

The next menu is the data menu shown in Figure 5. Uploading desired dataset by dragging and dropping files from this menu. The file is in "csv" format and when data uploaded system showing the data frame.

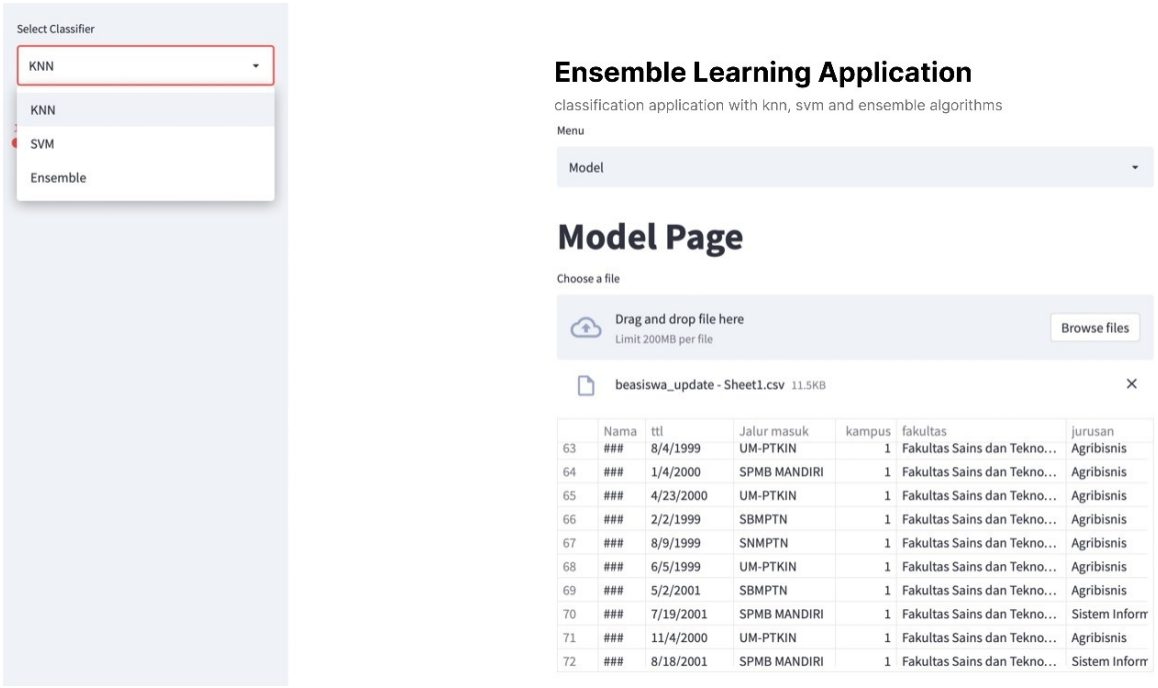


Figure. 6. Application GUI - Training Process

Model can be trained with an existing algorithm from the Models menu. There are three configurable algorithms. The first is the default model using ensemble, or choose the default algorithm, either KNN or SVM algorithm. To create the model, the user simply selects the desired algorithm and sets the parameters as shown in Figure 6.

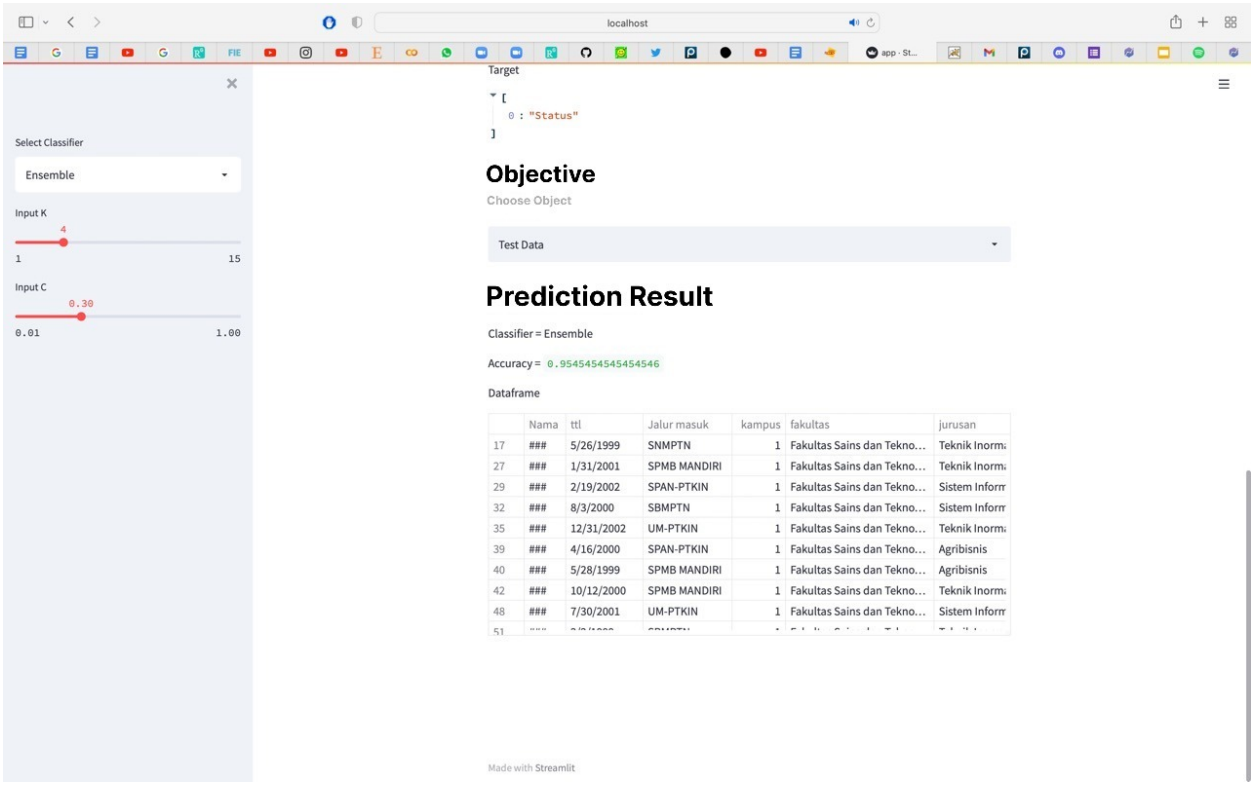


Figure. 7. Application GUI – Prediction Result

After training the model, system will allow to scroll down the page and finally see the scholarship prediction results in the "Prediction result" column. Diterima means accepted and ditolac means rejected. In Figure 7 showed whether the student has been accepted for the scholarship or not. In addition to the results, it also shows the accuracy comparison of the model, and the current ensemble learning model is good with 90% of accuracy.

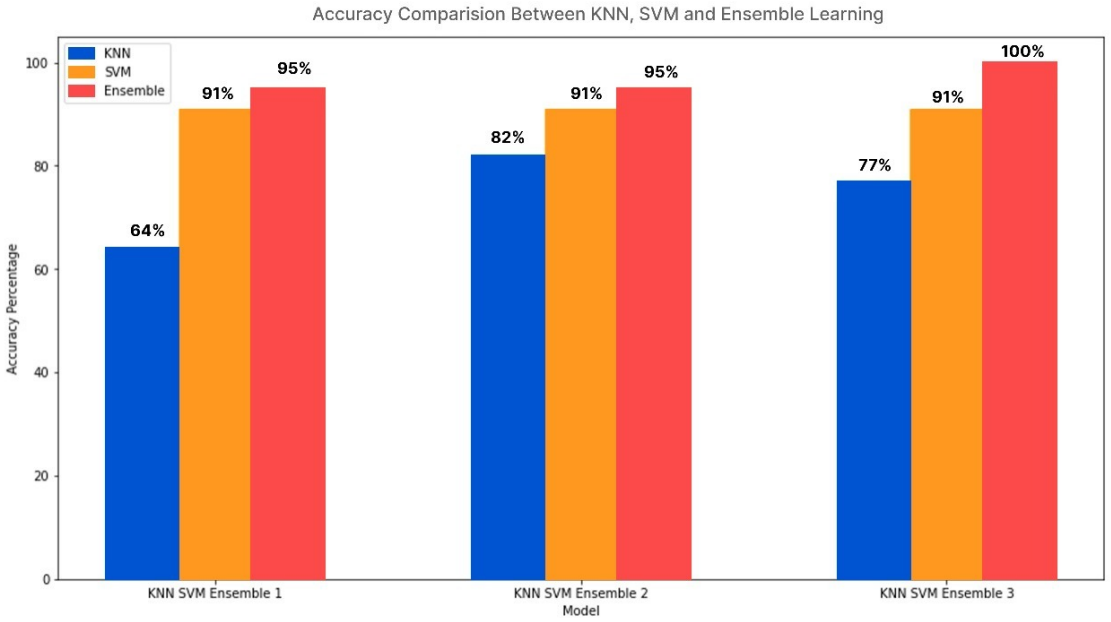


Figure. 8. Accuracy Comparison

After all process, model will be measured from accuracy of several predictive tests performed, as shown in Figure 8. As shown in this figure, Ensemble models always have the highest level of accuracy (over 90% to 100%) and Ensemble learning has the best accuracy compared to other basic algorithms.

6. Conclusion and Suggestion

The complexity of the scholarship selection process in an university has led into several problems: inefficient data, human-error, and consuming lot of time. This study proposed a new more-established kind of model to help predicting the selection process, whether the student will be accepted or not for the scholarship. Using Ensemble learning this study using two base learner which are K-Nearest Neighbor and Support Vector Machine to be combined into a Ensemble Learning model using voting classifier. Based from the result, ensemble learning can generate prediction result with 100% of accuracy, precision, recall, and F1-Score. According to the result, it can be said that the model can perfectly predict the selection process better than KNN and SVM Model. Furthermore, application also made to automatically predict the result based from the uploaded dataset using streamlit.

After completing the research, there are several things can be improved. By adding more data especially the training data into dataset, model performance can increased, both accuracy and precision and moveover the prediction result. Deployment will also work to reach broader audience, not only local access by utilizing cloud service in order to make real-time prediction application. Adding base learner will also considered and doing comparison to get the best algorithm combination.

In conclusion, the results of this study can be used as a DSS (Decision Support System) for scholarships at various universities. If the results of this study are utilized, application can be used not only in the prediction of scholarships but also in various similar fields in universities.

References

[1] N. Buslim, R. P. Iswara, and F. Agustian, "the Modeling of 'Mustahiq' Data Using K-Means Clustering Algorithm and Big Data Analysis (Case Study: Laz)," *J. Tek. Inform.*, vol. 13, no. 2, pp. 213–230, 2021, doi: 10.15408/jti.v13i2.19610.
[2] I. Ullah, K. Liu, T. Yamamoto, M. Zahid, and A. Jamal, "Electric vehicle energy consumption prediction using stacked generalization: an ensemble learning approach," *Int. J. Green Energy*, vol. 18, no. 9, pp. 896–909, 2021, doi: 10.1080/15435075.2021.1881902.
[3] M. Hosni, I. Abnane, A. Idri, J. M. Carrillo de Gea, and J. L. Fernández Alemán, "Reviewing ensemble classification methods in breast cancer," *Comput. Methods Programs Biomed.*, vol. 177, pp. 89–112, 2019, doi: 10.1016/j.cmpb.2019.05.019.
[4] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, no. 4, pp. 1–18, 2018, doi: 10.1002/widm.1249.
[5] I. K. Nti, A. F. Adekoya, and B. A. Weyori, "A comprehensive evaluation of ensemble learning for stock-market prediction," *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00299-5.
[6] P. Purwanto and D. Syarif Sihabudin Sahid, "Using KNN Algorithms for Determining the Recipient of Smart Indonesia Scholarship Program," *J. Komput. Terap.*, no. Vol. 7 No. 2 (2021), pp. 163–173, 2021, doi: 10.35143/jkt.v7i2.4962.
[7] W. D. Ahmad and A. Abu Bakar, "Classification Models for Higher Learning Scholarship Award Decisions," *Asia-Pacific J. Inf. Technol. Multimed.*, vol. 07, no. 02, pp. 131–145, 2018, doi: 10.17576/apjitm-2018-0702-10.
[8] U. Pasupulety, A. Abdullah Anees, S. Anmol, and B. R. Mohan, "Predicting stock prices using ensemble learning and sentiment analysis," *Proc. - IEEE 2nd Int. Conf. Artif. Intell. Knowl. Eng. AIKE 2019*, pp. 215–222, 2019, doi: 10.1109/AIKE.2019.00045.
[9] A. K. Patra, R. Ray, A. A. Abdullah, and S. R. Dash, "Prediction of Parkinson's disease using Ensemble Machine Learning classification from acoustic analysis," *J. Phys. Conf. Ser.*, vol. 1372, no. 1, 2019, doi: 10.1088/1742-6596/1372/1/012041.
[10] D. A. Pisner and D. M. Schnyer, *Support vector machine*. Elsevier Inc., 2019.
[11] P. Barberá, A. E. Boydstun, S. Linn, R. McMahon, and J. Nagler, "Automated Text Classification of News Articles: A Practical Guide," *Polit. Anal.*, vol. 29, no. 1, pp. 19–42, 2021, doi: 10.1017/pan.2020.8.
[12] S. Muthukrishnan, H. Krishnaswamy, S. Thanikodi, D. Sundaresan, and V. Venkatraman, "Support vector machine for modelling and simulation of heat exchangers," *Therm. Sci.*, vol. 24, no. 1PartB, pp. 499–503, 2020, doi: 10.2298/TSCI190419398M.
[13] S. Srisakaokul, Z. Wu, A. Astorga, O. Alebiosu, and T. Xie, "Multiple-Implementation Testing of Supervised Learning Software," *Work. Thirty-Second AAAI Conf. Artif. Intell.*, pp. 384–391, 2018, [Online]. Available: https://www.aaai.org/ocs/index.php/WS/AAAIW18/paper/viewPaper/17301.
[14] N. Buslim, L. K. Oh, M. H. Athallah Hardy, and Y. Wijaya, "Comparative Analysis of KNN, Naïve Bayes and SVM Algorithms for Movie Genres Classification Based on Synopsis.," *J. Tek. Inform.*, vol. 15, no. 2, pp. 169–177, 2022, doi: 10.15408/jti.v15i2.29302.
[15] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, 2020, doi: 10.1016/j.neucom.2019.10.118.
[16] E. K. Hashi, M. S. Uz Zaman, and M. R. Hasan, "An expert clinical decision support system to predict disease using classification techniques," *ECCE 2017 - Int. Conf. Electr. Comput. Commun. Eng.*, pp. 396–400, 2017, doi: 10.1109/ECACE.2017.7912937.
[17] H. A. Abu Alfeilat *et al.*, "Effects of Distance Measure Choice on K-Nearest Neighbor Classifier Performance: A Review," *Big Data*, vol. 7, no. 4, pp. 221–248, 2019, doi: 10.1089/big.2018.0175.
[18] S. Bagga, A. Goyal, N. Gupta, and A. Goyal, "Credit Card Fraud Detection using Pipeling and Ensemble Learning," *Procedia Comput. Sci.*, vol. 173, no. 2019, pp. 104–112, 2020, doi: 10.1016/j.procs.2020.06.014.
[19] R. Atallah and A. Al-Mousa, "Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method," *2019 2nd Int. Conf. New Trends Comput. Sci. ICTCS 2019 - Proc.*, pp. 1–6, 2019, doi: 10.1109/ICTCS.2019.8923053.
[20] M. Saqlain, B. Jargalsaikhan, and J. Y. Lee, "A voting ensemble classifier for wafer map defect patterns identification in semiconductor manufacturing," *IEEE Trans. Semicond. Manuf.*, vol. 32, no. 2, pp. 171–182, 2019, doi: 10.1109/TSM.2019.2904306.
[21] P. Singh, *Deploy Machine Learning Models to Production*. Bangalore, India: Apress, 2021.
[22] H. Koç, A. M. Erdoğan, Y. Barjakly, and S. Peker, "UML Diagrams in Software Engineering Research: A Systematic Literature Review," p. 13, 2021, doi: 10.3390/proceedings2021074013.
[23] Z. H. Muhamad, D. A. Abdulmonim, and B. Alathari, "An integration of uml use case diagram and activity diagram with Z language for formalization of library management system," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 4, pp. 3069–3076, 2019, doi: 10.11591/ijece.v9i4.pp3069-3076.
[24] J. Lang and D. Spišák, "Activity diagram as an orientation catalyst within source code," *Acta Polytech. Hungarica*, vol. 18, no. 3, pp. 127–146, 2021, doi: 10.12700/APH.18.3.2021.3.7.
[25] R. Fauzan, D. Siahaan, S. Rochimah, and E. Triandini, "Use case diagram similarity measurement: A new approach," *Proc.*

2019 Int. Conf. Inf. Commun. Technol. Syst. ICTS 2019, pp. 3–7, 2019, doi: 10.1109/ICTS.2019.8850978.

[26] M. Syarif and W. Nugraha, "UML Diagram Modeling of Cash Payment Systems in E-Commerce Transactions," *J. Tek. Inform. Kaputama*, vol. 4, no. 1, p. 70 halaman, 2020, [Online]. Available: http://jurnal.kaputama.ac.id/index.php/JTIK/article/view/240.

[27] S. Karlos, G. Kostopoulos, and S. Kotsiantis, "A Soft-Voting Ensemble Based Co-Training Scheme Using Static Selection for Binary Classification Problems," *Algorithms MDPI*, vol. 13, no. 26, 2020, doi: 10.3390/a13010026.