Prediction of Senior High School Students' Performance in a State University: An Educational Data Mining Approach

Jonardo R. Asor*, Gene Marck B. Catedrilla, Chester Alexis C. Buama, Monique E. Malabayabas, and Cristina E. Malabayabas

***Abstract*—The study was conducted to predict the performance of the senior high school students in one state university using machine learning algorithms. Data mining process was followed to develop a model for predicting the students' performance. The 4-year records of senior high school students of Laguna State Polytechnic University-Los Baños was gathered and used in the model development through well-known machine learning algorithms such as decision tree, naïve bayes, random forest, neural network and linear regression. Upon the development of the models, it is found in this study that naïve bayes performs well against the remaining algorithms and neural networks also shown a promising result in predicting student performance. The study also found that senior high school students have a high chance of not performing well upon entering the school based on the prediction of naïve bayes showing a high probability of satisfactory rating in Grade 11-1st semester applied subjects. Moreover, among the strands offered in Laguna State Polytechnic University-Los Baños, Accountancy, Business and Management students predicted to have the highest chance of having outstanding performance while Information and Communications Technology students predicted to have a high chance of satisfactory.**

***Index Terms*—Machine learning, naïve bayes, neural network, decision tree, random forest, linear regression**

I. Introduction

Education is considered to be the cornerstone for all future achievement in practically every country in the world. Even if every nation has its unique educational system, the majority of them share a similar idea or a similar level of knowledge. The Philippines' education system is distinctive in that it takes cues from its colonial past. In the Philippines, education is one of the most important social areas for addressing poverty and improving everyone's quality of life [1]. This is the rationale for the Philippine government's implementation of a Republic Act, which seeks to improve the nation's basic education system. Republic Act 10533 mandated to include an additional two years to the basic education system of the Philippines. The implementation aims to enhance the current education system of the country by strengthening the curriculum.

When the Department of Education (DepEd) issues DepEd Order No. 43 s. 2013—an order was issued in 2013 to put into effect the K–12 Curriculum Program for basic education, which aims to improve learners' fundamental skills, generate qualified graduates, and prepare them for the job market and the future. The Senior High School curriculum, which includes an additional two years and aims to help students learn and advance in a variety of subjects and fields of specialization, was expanded. K-12 Curriculum program includes four different tracks such as Academic, Sports, Technical-Vocational and Design which have different strands or specializations. Students are free to select and choose from among the various strands or specializations that make up each track based on their preferences or areas of strength [2].

The Laguna State Polytechnic University (LSPU) is a state university in the Philippines with four regular campuses and several auxiliary sites. It is a known university that offers academic programs in different colleges. In 2016, LSPU added the Senior High School under the College of Teacher Education (CTE) that offers different tracks namely Academic and Technical-Vocational-Livelihood. One of the missions of the said university was to deliver quality education through responsive instruction, distinctive research, and sustainable extension which leads the administration and faculty members of the university to leverage the academic performance of its students through different scientific-based strategy.

According to SAS Institute Inc. [3], the process of discovering, categorizing, predicting, recognizing patterns, inconsistencies, and correlations of data is known as data mining, and it is used to solve specific problems in various fields through modeling, analysis, and interpretation. This method uses a variety of algorithms to aid in analysis, modeling, and interpretation. Classification approach is the most famous data mining technique used in data analysis and pattern recognition. Additionally, it is used to categorize items into groups or categories and predict items based on observations made on a dataset. To evaluate and gauge the accuracy for the extraction of the best outcomes, a set of data was divided into training and testing data sets as stated in the research [4]. Moreover, data mining is considered to a powerful technology in discovering substantial information in a dataset which can be used in different problems like students' performance prediction and analysis [5]. Educational data mining (EDM) is a state-of-the-art method in discovering knowledge and/or patterns regarding students' performance which can help policy makers in creating new policies substantial to enhancing students' learning process [6]. While Phauk and Okazaki [7], found that educational data mining can be used to predict student performance using

Manuscript received August 28, 2022; revised September 21, 2022; accepted November 25, 2022.

J. R. Asor and G. M. B. Catedrilla are with the College of Computer Studies, Laguna State Polytechnic University, Los Baños Campus, Los Baños, Laguna, Philippines.

C. A. C. Buama is with College of Arts and Sciences, Laguna State Polytechnic University, Los Baños Campus, Los Baños, Laguna, Philippines.

M. E. Malabayabas and C. E. Malabayabas are with the College of Teacher Education, Laguna State Polytechnic University, Los Baños Campus, Los Baños, Laguna, Philippines.

*Correspondence: asor.jonardo@lspu.edu.ph (J.R.A.)

doi: 10.18178/ijiet.2023.13.6.1888

statistics, and machine learning. Further, extracting useful knowledge and information for the purpose of predicting and emerging the educational environment becomes an engaging strategy to higher education institutions [8]. On its peak as an emerging technology in solving complex problem, data mining, EDM in particular frequently showed promising results in discovering patterns in students' performance which can be used in enhancing or modifying the curriculum design [9].

This research paper aims to predict the performance of senior high school of Laguna State Polytechnic University–Los Baños (LSPU–LB) through datamining. Machine learning algorithms were implemented and evaluated in the development of model to ensure the reliability of prediction. Hence, LSPU–LB Senior High School records were used in the development as dataset.

II. Materials and Methods

In this paper, knowledge discovery in databases model was adopted to discovery new substantial information regarding senior high school students' performance. This well-known model for discovering knowledge is famous in data mining and in even artificial intelligent model development which is widely used around the globe with academicians and researchers. Hence, maximizing its process as main method in this study like this will be advantageous.

A. Data Acquisition and Preprocessing

Senior high school of LSPU-LB class records was collected. The class record is collected through the submitted reports of the teachers to the principal's office which contains all the rated performance of the senior high school students. The dataset contains the four batches of senior high school (junior and senior) students which also spans four years from 2018–2022. The students' performance record contains the following attributes: sex, age, address, core courses, specialization course, elective course, senior high school strands and general weighted average of each student (see Table I).

TABLE I: Attributes of Senior High School Performance Record

Attributes	Description
Sex	Sex of the student.
Age	Age of the student.
Address	Location where the student is living.
Core Courses	Grade in core courses taken by the student.
Specialization Courses	Grade in specialization courses taken by the student.
Elective Courses	Grade in elective courses taken by the student.
Descriptor	Students' performance rating.
Specialization	Strand pursued by the student.

TABLE II: Senior High School Students' Perfromance Ratings

Descriptor	Grading Scale	Remarks
Outstanding	90–100	Passed
Very Satisfactory	85–89	Passed
Satisfactory	80–84	Passed
Fairly Satisfactory	75–79	Passed
Did Not Meet Expectations	Below 75	Failed

A summary of learner progress is presented to parents or guardians quarterly during a parent-teacher conference, wherein the report card is discussed. Table II shows the grading scale and its corresponding descriptors. At the end of each grade level, remarks are given. For example, a learner received a transmuted grade of 90–100, it will be equivalent to Outstanding (O). If a learner received a transmuted grade of 85-89, which is equivalent to Very Satisfactory (VS), Satisfactory (S) for 80-84, and fairly satisfactory (FS) if the students received a grade of 75–79. Thus, if a learner received a grade of 74 below, it means that the learner Did Not Meet Expectations (DNME) in the specific grade level.

The senior high school specialization and elective courses varies depending on the track that the students were pursuing. Thus, the numbers of courses for specialization and elective were not the same in numbers for each track. Academic track like Accountancy, Business and Management (ABM) and Humanities and Social Sciences (HUMSS) contains two to three specialization courses while technical-vocation track like Information and Communications Technology (ICT) and Agri-Fishery Arts (AFA) has only one specialization course for each semester. This instance will greatly affect the performance of machine learning algorithms in pattern recognition and/or knowledge discovery which may give a biased result in analyzing students' performance. In addressing this problem, tracks that have more than one specialization track was converted into one through percentage method. The process will sustain and still show the performance of the students in their specialization courses since percentage method scopes all the rates of the students through average rate.

B. Model Development

Different well-known machine learning algorithms for students' performance prediction and analytics were implemented in this study to develop a model based from [10] methodology as applied in the same study. Hence, decision tree (DT) for classification algorithm was chosen for clear visualization of performances, naïve bayes (NB) for probabilistic prediction approach, random forest (RF) for ensemble algorithms, linear regression (LR) for regression problem and neural network (NN) for a more sophisticated prediction. Moreover, the models were developed using RapidMiner—a software known for data mining solutions.

1) Decision tree

DT is a tree like algorithm which consists of internal and external nodes. Internal nodes represent the condition of the tree while external nodes depict the outcome of the decision or simply the class. The branches which are represented by arrows show the conditions on how a class may be encountered or predicted [11].

2) Naïve bayes

NB algorithm is a probabilistic machine learning algorithm derive from Bayesian theorem which assumed independence between predictors. The algorithm is well-known to be a simple and easy to build model since it has no complex iterative parameter estimation which makes it efficient to use in large dataset. The simplicity and effectivity of NB made it famous and widely used in students' performance prediction [12–14]. In NB, it is assumed that the effect of the value of predictor (x) on a given class (c) is independent and does not affect other predictors. The assumption is called class conditional independence.

$$P(C=c_i|X=x)=\frac{P(X=x|C=c_i)P(C=c_i)}{P(X=x)} \propto P(X=x|C=c_i) \quad for\ i=1,2,...L \tag{1}$$

This implementation works depending on the dataset available which in this case is the senior high school rating. The dataset is considered for analyzing and predicting students' performance. NB implementation in this study may help on finding factors affecting the performance of the students.

3) Random forest

RF is an ensemble machine learning algorithm that is frequently used in students' performance and found to give a high accuracy rate in prediction [15]. This ensemble algorithm is composed of numerous structured-trees $\{h(x,\theta_k), k=1 \ldots\}$ where $\{\theta_k\}$ is the independent but identical vector distributed to cast vote in searching common class for each x. Furthermore, in the experiment conducted by [16] random forest shows a promising result in classification by regression problems which leads to be an interesting area to study for prediction.

4) Linear regression

LR is another simply yet effective machine learning algorithm in developing model for students' performance prediction. Thus, unlike NB, LR assumes that the dependent variables are related to the independent variables or the class in a linear manner. Stated in the study of Bum *et al*. [17] that LR is the quickest machine learning algorithm in developing model without compromising the accuracy of its prediction hence, it was proven on the experiment of research [18] after comparing the results of shortest time in building model and prediction accuracy to another two sophisticated machine learning algorithms (NN and support vector regressor). In this manner, LR can be used through this formula: $y=B0+B1*x1$ where y is the class and x are the independent variables.

5) Neural network

In this study, NN was developed containing an input layer with seven neurons, hidden layer with five neurons and output layer with four neurons. This common structure of NN found to give a promising result in predicting students' performance [19]. Hence, NN predictor model can be used to recognized pattern and trends on the students' performance which can be used by educators and school administrator for strategic planning and policy making [20].

6) Training and testing dataset

Division of dataset into two parts is one of the best practices in model development. The first part of the whole dataset is called training dataset which is commonly the 70% or 80% of the whole dataset while the second part which are the remaining 30% or 20% called testing dataset. In this study, 80% was used as training and the remaining 20% of the dataset used as testing. Training dataset is for training the five-model discussed in this paper while testing dataset is used to test and evaluate the performance of each model through cross validation. Table III shows the actual number of instances for training and testing dataset based on the class attribute which is the strand of the students.

TABLE III: NUMBER OF INSTANCES FOR TRAINING AND TESTING DATASET

Class	Training Dataset *Number of Instances*	Testing Dataset *Number of Instances*
ABM	38	10
HUMSS	45	11
ICT	46	11
AFA	14	4
Total	143	36

C. Model Evaluation

The training and testing dataset were divided into two sets for validation as stated in the model development phase. These results in the creation of the crucial validation measures required for a reliable validation of any machine learning algorithm's performance. Additionally, additional essential metrics were produced using both classification and regression approaches in addition to k-fold validation. The performance of the algorithms was measured for classification using metrics including accuracy, kappa, recall, precision, specificity, false positive rate (FPR), false negative rate (FNR), Matthew's coefficient correlation (MCC), and f-score. Furthermore, confusion matrix is used in the study for easier validation of prediction performance since it shows actual numbers of correct and incorrect prediction.

III. RESULTS AND DISCUSSION

As shown in Table IV, all the algorithms except LR works well in predicting the performance of senior high school of LSPU-LB. Hence, it shown in the table that NB and NN have an almost the same accuracy however, NB outperformed NN gaining an accuracy rate of 90.46% and 83.83% kappa. LR in this manner gained the lowest accuracy and kappa rate and even showed a non-learning rate after gaining a kappa rate lower than 50%.

TABLE IV: PERFORMANCE OF THE ALGORITHMS BASED ON ACCURACY AND KAPPA

Classification Algorithms	Accuracy	Kappa
DT	81.96	68.5
NB	90.46	83.3
RF	88.20	79.1
NN	89.87	82.1
LR	66.24	43.0

Confusion matrix showed true negative (TN), true positive (TP), false negative (FN), and false positive (FP) which can be used to calculate recall, precision, specificity, MCC, FPR, FNR, F-score, and accuracy among others. It is noticeable in Table V that NB and NN have the lowest prediction error. Through counts, it can determine that NB only has seventeen (17) prediction error while NN has a total of eighteen (18) prediction error. Thus, it is now harder to decide which among the two classification algorithms is better since there is only 0.56% of difference in terms of prediction error between them.

Since NB and NN were both giving promising result in accuracy, kappa and confusion matrix, a thorough evaluation was done using different substantial metrics for prediction (shown in Table VI). In most practices, precision, recall and

F1-score were always calculated in selected the most viable machine learning algorithms in terms of prediction. However, in some case like this study, other metrics must be included to avoid misjudgment between algorithms which gives an almost the same ratings.

Shown in Table VI are the metrics that can help to decide which has given a better prediction between NB and NN. Thus, RF has a higher recall than NB and NN which need to be address since recall shows how many classes are recalled correctly by the algorithm. Nonetheless, in precision and F1-score, NB and NN outperformed RF. Whilst, NB still outperforming NN in other metrics including FPR and FNR (error rates) where it has lower rates which show that there is a lower chance of NB to predict errors than NN. Therefore, NB is better than NN in predicting the students' performance in this study and it is necessary to use this algorithm in analyzing the dataset.

TABLE V: CONFUSION MATRIX OF THE CLASSIFICATION ALGORITHM MODELS

Classification Algorithms		actual O	actual VS	actual S
	pred. O	72	10	0
DT	**pred. VS**	13	63	4
	pred. S	0	5	11
	pred. O	80	7	0
NB	**pred. VS**	5	68	2
	pred. S	0	3	13
	pred. O	76	7	0
RF	**pred. VS**	9	70	4
	pred. S	0	1	11
	pred. O	78	5	0
NN	**pred. VS**	7	70	3
	pred. S	0	3	12
	pred. O	57	15	0
LR	**pred. VS**	26	52	6
	pred. S	2	11	9

TABLE VI: PERFORMANCE OF CLASSIFICATION ALGORITHM MODELS

Classification Algorithms	Recall	Precision	Specificity	FPR	FNR	MCC	F1-score
DT	78.07	79.96	89.41	10.59	21.93	68.06	78.98
NB	89.32	87.96	94.35	5.64	10.68	82.99	88.59
RF	90.27	83.08	92.13	7.87	9.72	79.42	86.31
NN	87.17	87.16	94.09	5.91	12.83	81.20	87.15
LR	60.66	64.58	78.98	21.01	35.42	41.10	61.82

In Fig. 1 shown the actual and predicted number of the students' performance using the descriptor prescribed by the Department of Education in the Philippines. It is shown that the predicted O and S performing students is higher than the actual number while predicted VS performance got a lower number. These may imply that some of the VS performing students may be an O or a S.

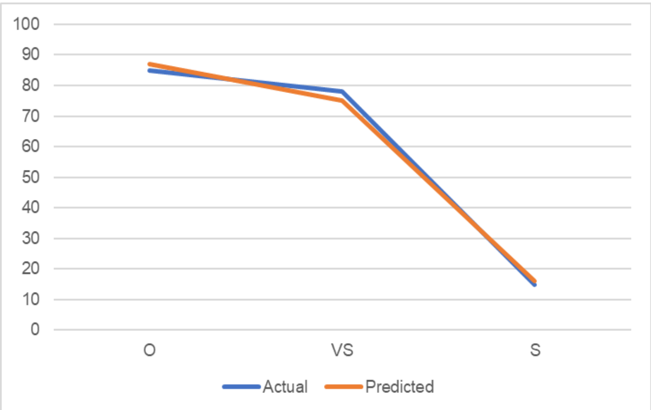

Fig. 1. The actual and predicted performance of the students using naïve bayes.

Fig. 2 shows that based on the prediction of NB, ABM and HUMSS have a higher chance of producing outstanding students. Hence, it shown in the prediction that majority of ABM were outstanding performing students and there is a lower chance of them will finished the strand with only satisfactory performance. HUMSS reflects to be an average performing students after gaining an almost the same rating in all the performance descriptor. Aquaculture (AC) students to be an average-to-outstanding students after not having any satisfactory performing students in the prediction. On the other hand, ICT students got the highest rates in acquiring S students while gaining the lowest O rate in NB prediction.

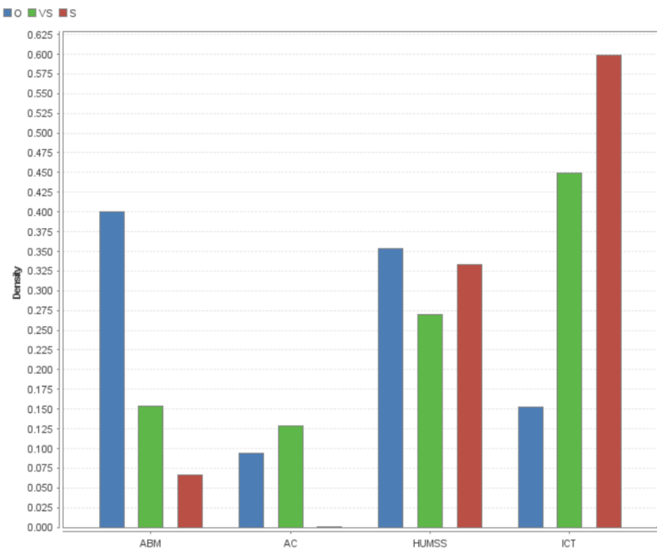

Fig. 2. Naïve bayes prediction on students' performance according to their strands.

It is noticeable in Fig. 3 that grade 11 during the 1st semester (AS-11-1) did not perform well after having a high S prediction, thus, O students were higher than the very satisfactory. In the 2nd semester (AS-11-2) of their grade 11, it still predicted that gaining a S rate is higher than O. However, unlike the 1st semester, VS is higher in 2nd semester than the O. Therefore, it can be assumed that students' performance on their applied subjects were not good at their first year in the school.

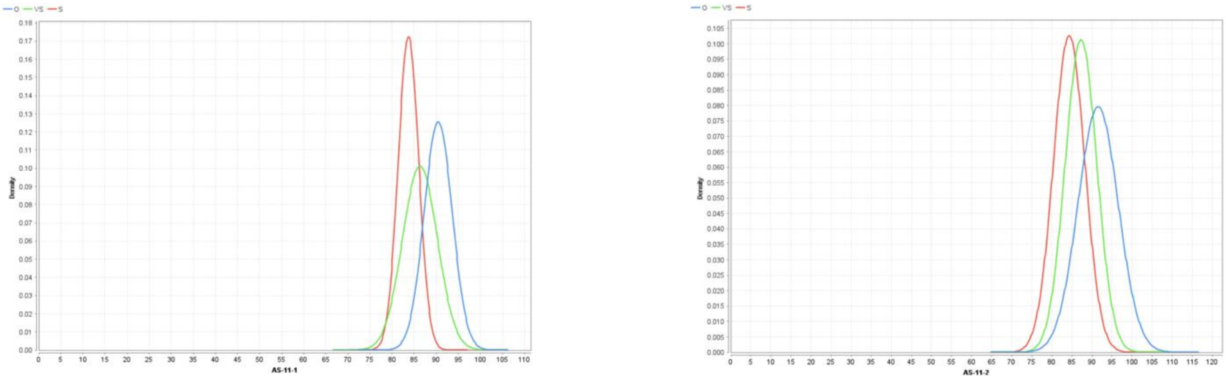

Fig. 3. Results of prediction of grade 11 students' performance in applied subjects for 1st and 2nd semesters.

Fig. 4 is the result of predicting the students' performance on their applied subjects for their 2nd year in the school. In the 1st semester (AS-12-1), it is perceptible that the students have a higher chance of getting O performance rate. However, the chance of getting a S rating was not far from the O. Nevertheless, for 2nd semester (AS-12-2) getting an O rate has higher chance since S prediction rating was very low compare to the latter. In addition, VS is higher than S performance which means that producing S students can be avoided during the 2nd semester of grade 12.

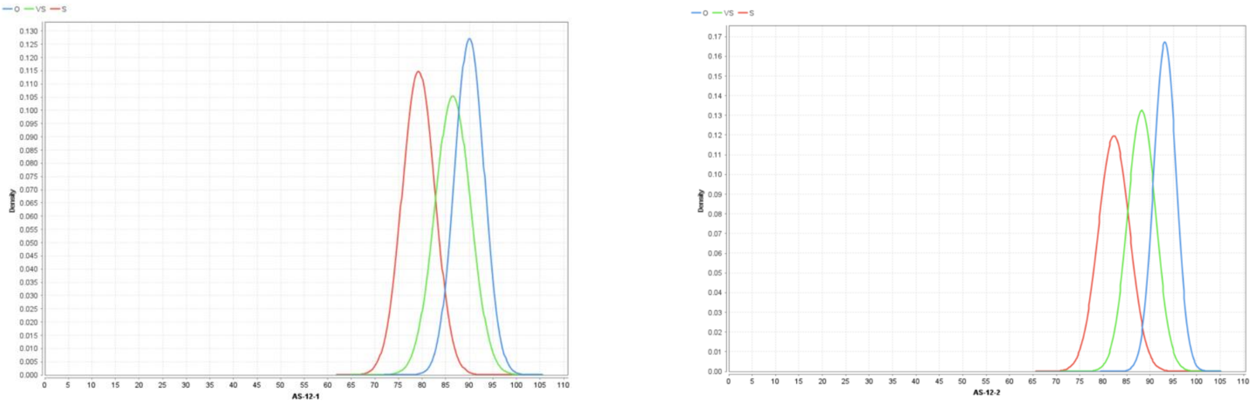

Fig. 4. Results of prediction of grade 12 students' performance in applied subjects for 1st and 2nd semesters.

In Fig. 5 shown the prediction of how grade 11 will perform on their specialized subjects. As reflected in the figure SB-11-1, grade 11 in their 1st semester has a high chance of getting an O performance. However, the chance of students getting S performance is almost the same as O. Furthermore, in the 2nd semester (SB-11-2) of grade 11's specialized subject performance, it is expected that the students will perform well after predicting a high rate of O performance than S. Hence, it is also noticeable that VS still have the same chance of occurrence in both 1st and 2nd semester.

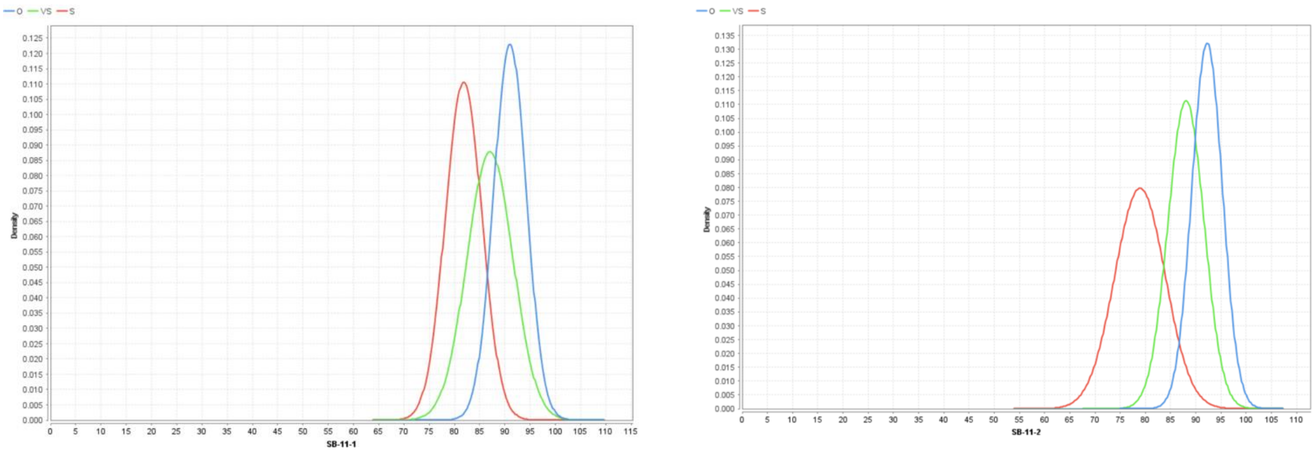

Fig. 5. Results of prediction of grade 11 students' performance in specialized subjects for 1st and 2nd semesters.

Fig. 6 shows the predicted performance of grade 12 students on their specialized subject. Here shown that in the 1st semester (SB-12-1) of grade 12 on their specialized subject, students have a high possibility to get a VS performance rating thus, getting a S rating has the lowest probability. For the 2nd semester (SB-12-2), O has the highest probability rate while VS and S rate have the same chance of occurrence.

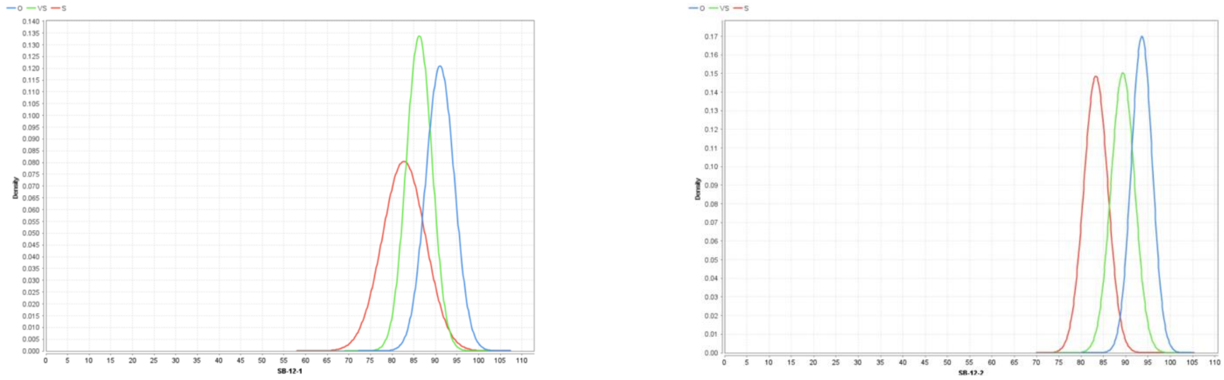

Fig. 6. Results of prediction of grade 12 students' performance in specialized subjects for 1st and 2nd semesters.

IV. Conclusion and Recommendations

In this paper model development using machine learning algorithms was conducted to predict the performance of the senior high school students of LSPU-LB. Among the five well-known machine learning algorithms in students' performance prediction, NB was found to be the most accurate algorithm to use after outperforming NN in different classification metrics. Thus, NN still shown a promising result in predicting students' performance after gaining 89.87% accuracy and kappa of 82.1% which was not far from the results of NB (accuracy=90.46, kappa=83.3). Moreover, NB is used to predict students' performance and found that senior high school students can be more outstanding on their subjects. In particular, ABM students have the high probability to get an outstanding performance against other tracks offered in LSPU-LB. Aquaculture track students on the other hand predicted to perform very well after predicting no satisfactory performance while the HUMSS students predicted to be average students after predicted to have an almost balanced performance of outstanding and satisfactory. Furthermore, ICT predicted to have a high probability of satisfactory students and low chances to have outstanding performance.

It is also found in this study that grade 11 students of 1st semester have high chances in not performing well in their applied and specialized subjects. While, there is still a high chance of getting a satisfactory performance of grade 11 for 2nd semester in applied subjects, it is found in this study that grade 11 for 2nd semester of specialized subject will perform very well, after gaining a high probability of outstanding performance through NB prediction. For grade 12 students, it is predicted in this study that despite of high chance of getting a satisfactory performance in the 1st semester, outstanding performance has much higher probability for applied subjects. Likewise, in the 2nd semester of grade 12, it is predicted that the students will have outstanding performance. Furthermore, outstanding performance has high chances for both grade 11 and 12 on their specialization subjects except grade 12 1st semester where very satisfactory has probability against outstanding and satisfactory. Therefore, it is concluded that new students were the ones who have a higher chance of getting a satisfactory performance, while senior high school students outstand more on their specialization subjects than applied. Based on the results and findings of this study, the following are recommended:

1) The usage of NN and NB in this kind of study.
2) Senior high school of LSPU–LB must device a procedure or technique to enhance the performance of ICT students.
3) It is also recommended to the school to have an effective strategy in learning adjustment of new students particularly in their applied subjects.

Conflict of Interest

This piece of work is funded by the Laguna State Polytechnic Universit, Los Baños and conducted in the funding agency therefore, the authors declare no conflict of interest.

Author Contributions

The data used in this research study was gathered by M.E. Malabayabas and Principal C.E. Malabayabas. Dr. C.A.C. Buama identify the requirements and conceptualized the study. G.M.B. Catedrilla preprocessed and prepare the data while J.R. Asor processed and developed the classification models. All the authors contribute in the content of the paper which includes editing and finalization.

Funding

This work is supported and funded by the Laguna State Polytechnic University.

Acknowledgment

The authors wholeheartedly acknowledged the Laguna State Polytechnic University for supporting this research works. Hence, the Senior High School of Laguna State Polytechnic Univeristy—Los Banos for trusting the authors in handling and processing their data. This will never be feasible without your support and trust.

References

[1] F. J. G. Castillo, D. F. Antiado and J. R. P. Rebalndo, "Philippine education system: Are we moving forward?" *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 12, pp. 844–851, 2019.

[2] G. M. Ducanes and D. J. Ocampo, "The Impact of the new senior high school program on the school participation of 16 and 17-year old learners in the Philippines," *The Philippine Statistician*, vol. 68, no. 2, pp. 111–130, 2019.

[3] SAS, "Data Mining: What it is and why is matters," *SAS Insights*. 2022..

[4] Geeks for Geeks, "Basic concept of classification (data mining)," 23 August 2022.

[5] S. Kumari and A. Bhatia, "Predicting students performance: An EDM approach," *International Journal of IT & Knowledge Management*, vol. 11, no. 1, pp. 1–8, 2017.

[6] A. F. Meghji, N. A. Mahoto, M. A. Unar, and M. A. Shaikh, "Analysis of student performance using EDM methods," presented at 5th International Multi-Topic ICT Conference (IMTIC), Pakistan, 2018.
[7] S. Phauk and T. Okazaki, "Integration of educational data mining models to a web-based support system for predicting high school student performance," *International Journal of Computer and Information Engineering*, vol. 15, no. 2, pp. 131–144, 2021.
[8] B. Albreiki, N. Zak, and H. Alashwal, "A systematic literature review of student' performance prediction using machine learning techniques," *Education Science*, vol. 11, no. 552, 2021.
[9] H. Luan and C.-C. Tsai, "A review of using machine learning approaches for precision education," *Educational Technology & Society*, vol. 24, no. 1, pp. 250–266, 2021.
[10] R. Alshabandar, A. Hussain, R. Keight, and W. Khan, "Students performance prediction in online courses using machine learning algorithms," presented at 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 2020.
[11] S. Zulfiker, N. Kabir, A. A. Biswas, P. Chakraborty, and R. Mahfujur, "Predicting students' performance of the private universities of bangladesh using machine learning approaches," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 3, pp. 672–679, 2020.
[12] J. L. Lerios and J. D. Niguidula, "Analysis of students' learning outcomes using educational data mining approach," *International Journal of Advanced Computational Engineering and Networking*, vol. 6, no. 2, pp. 41–50, 2018.
[13] V. Patil, S. Suryawanshi, M. Saner, V. Patil, and B. Sarode, "Student performance prediction using classification data mining techniques," *International Journal of Scientific Development and Research*, vol. 2, no. 6, pp. 163–167, 2017.
[14] A. Tripathi, S. Yadav, and R. Rajan, "Naïve Bayes Classification Model for the Student Performance Prediction," presented 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies, Kannur, India, 2020.
[15] S. Batool, J. Rashid, M. W. Nisar, J. Kim, T. Mahmood, and A. Hussain, "A random forest students' performance prediction (RFSPP) model based on students' demographic features," presented at 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), Pakistan, 2021.
[16] F. Balahadia, M. Villarica, G. M. Catedrilla, and J. Asor, "Building model for crime pattern analysis through machine learning using predictive analytics," *International Journal of Science, Technology, Engineering and Mathematics*, vol. 2, no. 1, pp. 61–73, 2022.
[17] S. Bum, I. B. Iorliam, E. O. Okube, and A. Iorliam, "Prediction of student's academic performance using linear regression," *Nigerian Annals of Pure and Applied Science*, vol. 2, pp. 259–264, 2019.
[18] E. Y. Obsie and S. A. Adem, "Prediction of student academic performance using neural network, linear regression and support vector regression: A case study," *International Journal of Computer Applications*, vol. 180, no. 40, pp. 39–47, 2018.
[19] M. A. Umar, "Student academic performance prediction using artificial neural networks: A case study," *International Journal of Computer Applications*, vol. 178, no. 48, pp. 24–29, 2019.
[20] P. M. Arsad, N. Buniyamin, and J.-L. Ab Manan, "A neural network students' performance prediction model (NNSPPM)," presented at 2013 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), Malaysia, 2013.

Copyright © 2023 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).