Predicting University Students' Academic Success Using Different Tree Classifiers And Ensemble Approaches To Suggest Suitable Program

Frederick F. Patacsil

Abstract: One of the most important issues in succeeding the academic life is to assign students to the right college program when they arrive to the end of the basic education stage. Data mining is one of the best alternative solution to find patterns and to provide a suitable program which enhance the student survival in their college life. This research experimented different tree classification algorithms and supplemented by ensemble approach to suggest the best classifier model based on predictive accuracy performance. The result reveals a very interesting prediction that the average grade in math is a very important predictor for programs like Mathematics and Engineering. For non-engineering program, the average grade in English is the dominant predicting variable. However, this study produce a conflating result which reveals that average grade in English is not a significant predictor on the program that offer English as a major subject (AB English Language) instead general grade average is the predicting variable. In the case of IT, the average grade in TLE and Science is an important predicting variable. The result of evaluation in the final stage of the model construction process reveals that bagging + j-48 tree classifier obtained highest accuracy as compared with other tree classifiers and forest tree classifier obtained the highest value in dropout precision and graduated recall. The experimental result also reveals that combining these tree learning algorithms using ensemble approach has minimal increase in classification performance.

Index Terms: tree classifier, decision tree, forest tree, j-48, ensemble approach, predicting, academic success

1. INTRODUCTION

The Universal Access to Quality Tertiary Education Act, which Philippine Congress ratified and transmitted to the Office of the President, was signed by President Rodrigo Duterte. The law covers a total of 112 state universities and colleges, and 78 local universities and colleges nationwide which gives full tuition subsidy for students. People tax is used to finance this Law and the government should ensure that student retention or persistence is attained throughout the duration of their stay. This is to assure that the money invested will not go to waste. With this scenario, the Philippine higher education generation for the next years will reverse the current situation from 80 percent of college students enrolled in private schools and 20 percent in state universities and colleges (SUCs) to 20 percent, private colleges, and 80 percent SUCs [1]. Free tuition also translates the increasing enrollment rates among students in the SUC and the backdrop, government, are spending large amounts of money per student enrolled. However, many of the students studying at the SUCs face several difficulties during the first year and thus, the performance of the first year has been identified as an important predictor of timely graduation rate. In terms of keeping the students in the university, the retention rate is a factor that has been studied extensively. An early identification of the students at high risk of failing will enable a timely intervention with the necessary measures by the educators that would increase the graduation rate. Preventing students' failure depends on the identification of the factors affecting success. However, very few studies are made to investigate the success of career path and the factors that affect the career choice of Filipino students. This provides us with limited information on how to help our students identify the proper career options and course choice they have to pursue in the future. Students are not properly oriented as to what program to choose out of their interests and skills, but because of the thought that these courses will provide jobs in the future. Many researches focus on the academic success, such as the success of the students to their courses or the success of students in their phases of studying, all in terms of current variables such as mental capability variables, psychomotor variables and different socio-demographic variables. Researchers rarely undertake scientific observation on student high school accomphishment. One reason for high drop out rates as reported in some researches, were poor career choices and lack of personal interest [2]. Yet, few studies investigated the success of career path used in the students, even the factors that affect the career choice of students. This provides us with limited information on how to help our students identify the proper career options and course choice they have to pursue in the future. Tree classification is the most familiar and most effective classifiying technique used to classify and predict values, hence, it was also used in this research paper to classify collected students' information and provide classifications based on the collected data. However, individual trees are prone to overfitting thus, too much reliance on the training data. In this data mining ira, researchers are continually advocating for the use of multiple classifiers to solve classification problems. The concept of combining classifiers is based on the assumption that the different classifiers, which use different data representations, different concepts and different modelling techniques are most likely to arrive at classification results with different patterns of generalization. Researchers have proved that integration of diverse classifiers reduces the classification errors[3]. Therefore, this study developed a proposed hybrid model that can be used to suggest student suitable program based on their enrollment dataset. Students success in their high school is in the suggested model in order to investigate its influence on the output variable (suggested course). This study has explored and analyzed the enrollment data that may have an impact on the study outcome of the students and experiment the performance accuracy and efficiency of single classifiers and add ensemble approach classifiers to propose a predictive model to help them choose a suitable program. A set of classification rules was extracted from the generated predictive model to predict suitable program for the students based on their enrollment data set. Furthermore, this study was proposed to investigate and compare the use of different tree classifiers and ensembles of classifiers technique to improve the results of different classifiers.

2. RELATED LITERATURE

2.1 Student High School Profiles

The determinants of graduation and academic performance of college students were studied taking into account the student's pre-collegiate endowment of knowledge and various factors associated with the high school profile.

2.2. Predicting Factors

High school GPA, admissions test scores, gender, and race/ethnicity are the factors in which the researcher has consistently found to be significant predictors of retention.

2.2.1 High School Grade Point Average

The GPA used in this analysis was the weighted grade-point average. Previous researches have demonstrated that GPA is a consistently better predictor of college performance than other variables [4,5,6,7,8,9, 10, 11, 12, 13, 14, 15, 16, 17, 18]. The main idea of previous researchers in using GPA is because it has a tangible value for future educational and career mobility. It can also be considered as an indication of realized academic potential [19]. Further, Grade Point Average is an indicator used to evaluate the success of the students during their high school years.

2.2.2 Grade in Math, Science, English and TLE

Some of the predictor variables considered in the study were high school grade-point average and grades in Science, Math, English and TLE. Other researcher utilized the high school final grades in English, Math and major subjects as predictor variables [20]. Beaulac, C., & Rosenthal, J. S. [21] observed that grades in Mathematics, Biology and Chemistry are consistently among the most important variables for predicting if a student will complete their program. Studies varied in identifying factors that affect student retention the most in their freshmen year. Zhang [22], Veenstra [23] claimed that high school GA and to grades in math, chemistry, and physics, are all strong predictors for engineering student retention [24],[25]. The study of Aulck, Velagapudi, Blumenstock, & West [26], Mesarić & Šebalj [8] revealed that GA in Math, English, and Chemistry were among the strongest individual predictors of student retention.

2.2.3 Gender

Previous studies explore the socio-demographic variables that may influence persistence or drop-out or pre-identifying successful and unsuccessful students [26],[27]. Cengiz, N., & Uka [24] found based on their research utilizing data mining techniques that the most important predictors for student success were the students' high school GPA and gender.

2.2.4 Machine Learning Building Predictive Model

There is a significant number of classification/prediction machine learning available with difference approach to determine student performance which was reported by many researchers. Decision tree algorithms were applied to generate a model that will predict the performance of students. Kabra & Bichkar [28], Kovacic[26], Quadril & Kalyankar[7] Al-Barrak & Al-Razgan [8], Mesarić & Šebalj [29] made use of decision trees in educational data mining and applied to predict the performance of students using their past performance data. The model enables to identify the students in advance who are likely to fail/drop-out/unsuccessful and allow the teacher to provide appropriate inputs. Other researchers utilized different learning machine to create their predictive model. Kabakchieva [29] applied different decision trees algorithms for predicting the academic success of students and found that decision is obtained by having high accuracy rate. Sembiring, Zarlis, Hartama, Ramliana, & Wani [30], Shovon, M. H. I., & Haque, M. [31] utilized Smooth Support Vector Machine (SSVM) classification and kernel k-means clustering techniques to determine/predict the of success students and develop a model of student performance predictors. Yadav and Pal [32] utilized decision tree algorithms to classify engineering students to predict their performance in the final exam. The dataset comprised 90 student records with 16 attributes. The C4.5 algorithm produced the best accuracy standing at 67.78%. Kovačić [33], Simeunović and Preradović [34], Shah's [35], Cheewaprakobkit [36], Nghe et al. [37] conducted a research in predicting student success using the different classifying learning machine. Decision trees were the most successful in terms of correct classifications among growing method of classifying learning machines. Decision tree was utilized in this study because it is so simple to understand the results and easy to make good interpretations. In addition, it is easier to be understood by a reader of this study. Most of researchers have used this technique because of its simplicity and comprehensibility to uncover small or large data structure and predict the value [22, 24, 38]. Further, based on the previous studies, decision tree is one of the most commonly used techniques in predicting student's performance [39]. However, the main disadvantage of decision trees is that they tend to overfit, but there are ensemble methods to counteract this.

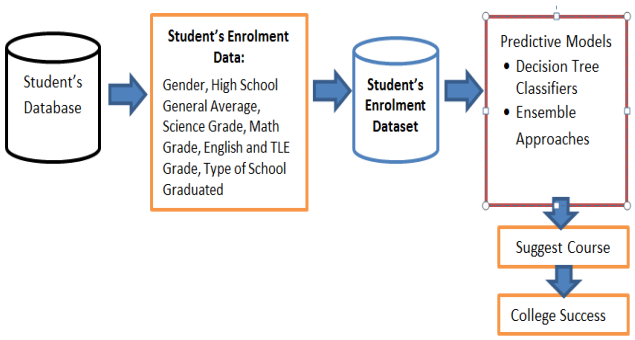
3. METHODS

The educational data mining (EDM) is defined as: "Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in. [40]"

3.1 Predictive Analytics

Predictive analytics is the act of predicting future events and behaviors present in previously unseen data, using a model built from similar past data [41], [38]. The method use previously collected data, a machine learning algorithm finds the relations between different properties of the data. The result is a model that will able to predict future data based on the properties of the collected data.

Conceptual Model



3.2 Data Selection and Preprocessing Data

The dataset was composed of 2401 admission records of 956 graduated and 1445 dropout students of Pangasinan State University - Urdaneta City Campus was collected who finish their schooling from the year 2012-16 as shown in table 1. The data were collected through the enrollment form filled by the student at the time of admission. The student enters their demographic data such as the gender, grade in English, math, science, TLE, general average and type of school graduated (public and private). From these, the attributes that possibly influence their result are selected as shown in Table 2. Most of the attributes reveal the past performance of the students. Reasons behind concentrating on the past performance data are 1. Data is easily available in the Registrars Office of the campus. 2. If a student has performed well in the high school, it is most likely that he will perform better during their college years as well or the other way around.

Figure 1 *Total Number of Graduated per program*

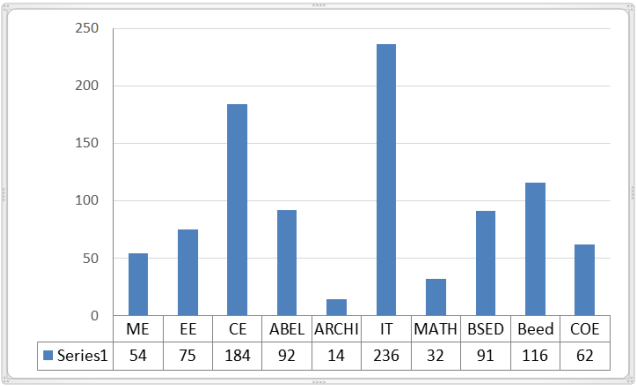


Figure 2 *Total Number of dropouts per program*

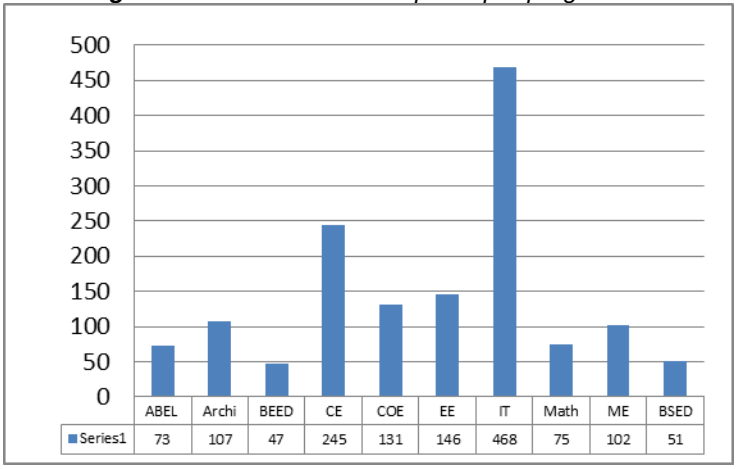


Table 2 *List Predicting Variables*

Variable	Description	Type	Values
Course	Program were the students enrolled	Character	ABEL - AB English Archi -BS Architecture BEED - BS Elementary Education CE - BS Civil Engineering COE - BS Computer Engineering EE - BS Electrical Engineering IT - BS Information Technology Math - BS Mathematics ME - BS Mechanical Engineering BSED - BS Secondary Education
Gender	Students Gender (Male	Nominal	1 - Male 2 - Female
Grade in Science	Average Grade in	Numeric	0 - 100
Grade in Math	Average Grade in Math	Numeric	0 - 100
Grade in English	Average Grade in	Numeric	0 - 100
Grade in TLE	Average Grade in TLE	Numeric	0 - 100
High School Grade	high-school grade point average	Numeric	0 -100
Type School	Type of School	Character	1 - Public 2 - Private

3.3 Predictor Variables

The main predictor variables considered in the study were high-school grade-point average(HSGPA) and grade average in the subjects, English, Math, Science and TLE. The HSGPA used in this analysis was weighted grade-point average, that is, a HSGPA at 85. This is the minimum GPA needed to be admitted in the majority of programs offered in the campus. Grades in Science, Math and TLE is also considered in the analysis because these are considered as prerequisites in the subjects in college. These subjects are considered preparatory subjects where students have to take general course in Mathematics, and English ommunication. In addition, this study considered the school were students graduated and their gender.

3.4 Model the Predictive Building

3.4.1 Classifying and Predictive tool

According to previous studies, the most widely used and supervised classification technique is Decision Tree which also includes learning and classifying. The steps involve were simple and fast and thus, very intuitive and easy to explain. Decision Tree can be applied to any domain [42]. The main goal of the decision tree is to create a model that will predict the value of a target variable by applying several inputs. In addition, it tree is a decision support tool that uses a tree-like graph or model of decisions and their possible outcomes. A decision tree is a classifier in the form of a tree structure where each node is either: Leaf node- Leaf node is an indicator of the value of target attribute(class) of examples, or a decision node- A decision node specfies all possible tests on a single attribute-value, with one branch and sub-tree for each possible outcome of the test [43]. Sample result is shown in figure 2.

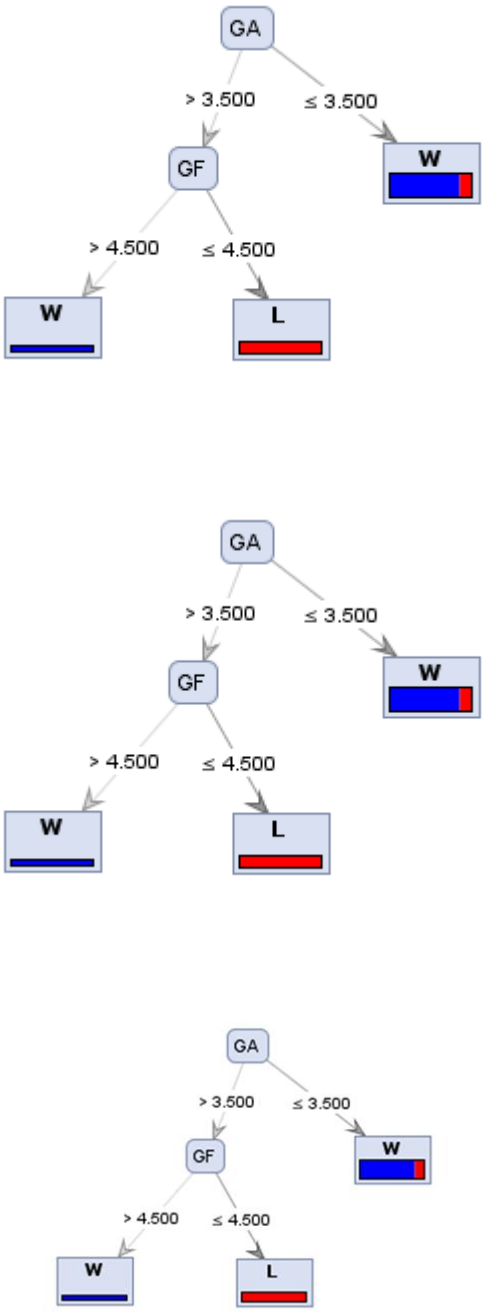


***Figure 2**. Sample tree structure result*

In this study, node in the tree is a predicting variables, and its branches are drawn on the basis of the suitable program for each students. Every node provides a decision predicting the success of students. However, to strengthen the predicting performance of the proposed tree model, ensemble methods was applied through combining several decision trees classifiers, boosting decision tree classifiers and bagging decision tree classifiers.

3.4.2 Building Model Process

Building the predictive/classification model is the next step. In this step, decision tree has been selepcted as a classifier under the cross validation method. The proposed model used 8 input variables, as shown in Table 2. Data was collected from enrollment information of all the students who graduated from school years 2013 - 2017. The attribute having maximum gain ratio value is the basis of splitting the nodes. This process continues till the complete tree is constructed. The RapidMiner tool kit was used to select the attributes and construct the decision tree. Moghimipour, I., & Ebrahimpour, M. [44] study reveals that Rapidminer obtained the highest accuracy as compared over three Data Mining Software. Fig 2 shows the decision tree construction. Each leaf node is represented by the rectangle and root node/splitting node is represented by an oval.

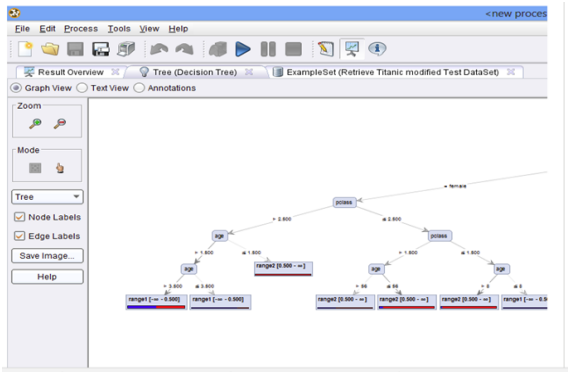


***Figure 3**: Decision Tree Construction*

This study used decision tree algorithms lto generate predictive model to suggest program based on the high school academic records of the students. According to Rokach and Maimon [45] Decision tree provides many advantages and some of which are the following:

• it is simple and can be clearly understood by the reader, end user and the analyst.
• it can accommodate different kinds of input data such as textual, nominal, and numeric.
• it can continue to process data that are erroneous or missing or uncompleted values.
• with a minimal amount of effort and time, it produces a high level of performance.
• it can work in data mining applications over a multi-variety of platforms .

However, there are drawbacks of decision trree classifiers such as the following:

- There is a high probability of overfitting in decision Tree.
- Generally, it gives low prediction accuracy for a dataset as compared to other machine learning algorithms.
- Information gain in a decision tree with categorical variables gives a biased response for attributes with greater no. of categories.
- Calculations can become complex when there are many class labels [46].

However, the study experimented ensemble models to counter the weaknesses of decision tree classifier to form a strong learner, thus, increasing the accuracy of the model.

3.4.2.1 Ensemble Model

Ensemble model is composed of meta-algorithms that creates n leaners from one algorithm sequentially or combine several machine learning techniques into one predictive model in order to decrease variance, bias , and improve predictions [47] . Furthermore, ensemble methods combine several decision trees classifiers to produce better predictive performance than a single decision tree classifier. The main principle behind the ensemble model is that a group of weak learners come together to form a strong learner, thus, increasing the accuracy of the model [48]. Previous research results clearly show that our ensemble model achieves better accuracies and better rules than other component models and other conventional forecasts combination schemes [48] [50]. According to Adejo, Connonlly, the results of their study show that heterogeneous ensemble techniques are very efficient and accurate in the prediction of student performance and help in proper identification of student at risk of attrition [51].

3.4.2.2 Vote (Stacking)

The Vote (stacking) operator is a nested operator that has a subprocess with at least two learners, called base learners. Furthermore, all the operators in the subprocess of the vote operator accept the given data set and generate a classification model. This study experimented ensemble (vote) to handle imbalance data (graduate / not graduated data) and to improve the performance of predicting suitable program for the student. This study run the three tree classification model and then combining the results into a sole score in order to improve the accuracy of predictive analytics and data mining applications.

3.4.2.3 Bagging

Bagging is an ensemble meta algorithm that creates n learners from one algorithm sequentially. The dataset is randomly sampled with replacement and created n datasets in a given ratio. There can be data points that are misclassified by a given learner [52]. The error of the previous classifier is considered and a new weight is given to the misclassified data element letting that data element to appear in new datasets more often. Therefore, bagging is used to help reduce the bias because it reduces variance and overfitting. According to [53], for unstable learning algorimths and unbalance data, bagging is an effective ensemble technique where there is a big changes in predictions result with small changes in the training data set. In addition, [54] research shows that ensemble decision tree classfiers using Adaboost and Bagging improves the performance of selected data sets.

3.4.2.4 AdaBoost

Boosting refers to a family of algorithms that are able to convert weak learners to strong learners. The main principle of boosting is to fit a sequence of weak learners models that are only slightly better than random guessing, such as small decision trees− to weighted versions of the data [55]. According to [56], Adaboost is found to be the best meta decision classifier for predicting the student's result based on the marks obtained in the semester. Furthermore, [57] found that adaboost improve the performances of tree algorithms.

3.5 Model Evaluation and Interpretation

To evaluate the predictive models 10 fold cross validation and percentage, qsplit methods have been used. 10 Fold Cross validation was utilized because the data is small and it is not feasible to split into two subsets. Thus, in order to minimize the bias, it makes full use of the dataset for training and for testing. Results were presented using confusion matrix that contains information about the actual and predicted classification done by the predictive models [47]. In terms of comparing their performance, the confusion matrix that contains the precision, recall, and accuracy [57] was employed.

Table 1 *Confusion Matrix for Two-Class Classifier*

Predicted Class	Graduated	Drop out	Class Precision
Graduated	tn	fp	Not Graduated
Not Graduated	fn	tp	Graduated
Class Recall	Drop out	Graduated	

Accuracy:

- The accuracy (AC) is the proportion of the total number of predictions that were correct. It is determined using equation (1) :

$$AC = \frac{tn + tp}{tn + fp + fn + tp} \quad (1)$$

- The recall (in the case of positive cases) is the proportion of positive cases that were correctly identified, as calculated using equation 2(2):

$$Recall = \frac{tp}{tp + fp} \quad (2)$$

- The precision (in the case of positive cases) is defined as the proportion of negative cases that were classified correctly, as calculated using equation (3):

$$Precision = \frac{tp}{tp + fn} \quad (3)$$

In addition, filters were implemented in the data and compared the result of other model without applying filters. This procedure provides a mechanism to gradually modify and correct noisy data may lead to better classification.

4. Results and Discussions

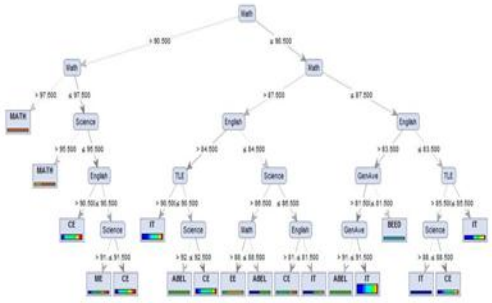


Figure 1*. Decision Tree Produced by Decision Tree Algorithm*

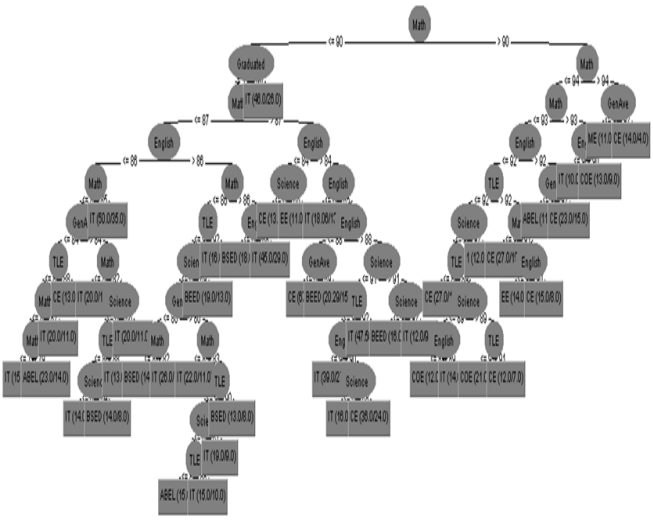


***Figure 2**. Decision Tree Produce by J-48 Algorithm*

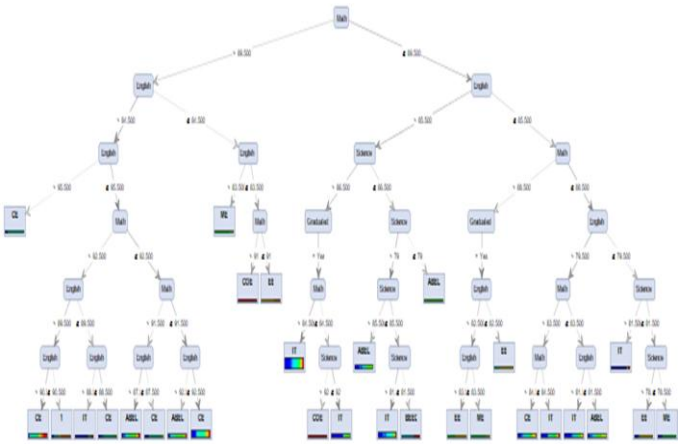


***Figure 3.** Figure 1. Decision Tree Produced by Forest Tree Algorithm*

The models in figure 1, 2, and 3 give an interesting information about student and provides guidance to parents to choose a suitable track for them to be successful. The result indicates that grade in Math is the best predictor of whether or not a student is going to graduate successfully or not. In addition, grades in Science and English become the next best predictor as shown in the figure above.

- For Math program. Students should have an average Math grade greater than 90.50 and a greater than 95 grades in science. The results indicate that to be able to graduate under Math program, students should have good grades in Math and Science.

- For Information Technology program. TLE is the main predictor for IT students and their average grade should be greater than 90, followed by Science and English. This result shows that IT students should need to have skills, and understanding on Science and Technology before enrolling IT program.

- For engineering programs. Students who intend to enroll engineering course should have an average Math grade of 88. The result also reveals that Science and English are also important predictors for engineering courses. This result confirmed the study of [21] [22] that the confidence in quantitative skills were significant predictors for engineering students' success.
- For ABEL and Education programs. High School General Average Grade Point and English is the main predictor for ABEL and BS Education courses. Students should have an High School General Average Grade Point of 91 and an average grade of 87 in English. The result obviously reveals that English majors and future teachers should have a high grade in English.

- For Architecture program. Science and TLE are the most important predictor for architecture students. Students should obtain a grade of more 87 in Science and more than 81 in TLE. The Science is the most significant predictors for architecture program followed by TLE. The result reveals that architecture students should need to have a skill (drawing which is included in the TLE course) and understanding of Science and Technology if they intend to enroll in the architecture program.

The result reveals a very interesting prediction that the average grade in Math is a very important predictor for courses like Mathematics and engineering programs. In addition, non engineering courses average grade in English is the dominant predicting variable. In the case of IT, the average grade in TLE and Science is important predicting variable. However, the result reveals that general high school point grade average is not a significant predictor in some program. Variables gender and type of school were graduated are not influential variables for predicting student success as shown in the result.

Rules Extraction A set of rules can be extracted from the decision tree. These rules are used to predict and classify the suitable program for each student. The class label in this tree acts as the suitable classified program after the end of a high school. The set of extracting rules from the decision tree is shown in Table 2.

***Table 2**. Extracted Rules from the result of decision tree generated rules*

Program	Rules
Mathematics	Math > 97.500 Science > 95.500: Mathematics
Civil Engineering	Math > 90.500 Science > 86.500 English > 90.500: Civil Engineering
Mechanical Engineering	Math > 90.500 English ≤ 90.500 Science > 91.500: Mechanical Engineering
Electrical Engineering	Math > 88.500 English > 81.500 Science > 86.500: Electrical Engineering
Information Technology	TLE > 90.500 Math > 87.500 Science > 88.50 English > 84.500: Information Technology
AB English Language	GenAve > 91.500 English > 83.500: AB English Language
Elementary Education	English > 90 Science > 80.500: Elementary Education
Secondary Education	English > 92.500 Science > 89.500 : Secondary Education
Computer Engineering	Math > 94.500: COE GenAve > 91.500 English > 92.500:Computer Engineering
Architecture	Science > 87.500 TLE > 81.500

The table 2 reveals that average grade in Math and Science is very important predictor if student want to enrol engineering and Mathematics programs. However, Mathematics program requires a very high grade in Math followed by computer engineering program and electrical engineering program has the lowest requirement for Math grade. In the case of education, English is the main predictor however, the table reveals an interesting result that grade in English is not a significant predictor on the program that offer English as major subject (ABEL) instead general grade average is its predicting variable.

4.1 Model Evaluation

***Table 3**. Comparison Accuracy Results of Tree Classifier and Applying Ensemble Approach*

Classifier	Precision		Recall		Accuracy
	Graduated	Dropout	Graduated	Dropout	
Decision Tree	60.98	74.23	61.05	74.2	68.97
Forest Tree	56.8	77.44	62.49	73.04	69.22
J-48	58.16	76.82	62.4	73.51	69.39
Ensemble Model (Vote)	62.32	73.41	57.95	76.82	69.31
Ensemble Model (Bagging + Decision Tree)	57.19	71.20	55.75	72.39	65.76
Ensemble Model (Bagging + Forest Tree)	61.29	72.64	55.65	76.75	68.35
Ensemble Model (Bagging+j-48)	64.31	73.93	58.05	78.69	70.47
AdaBoost+Decision Tree	53.65	72.06	62.24	64.43	63.56
AdaBoost+Forest Tree	60.54	71.62	54.08	76.68	67.68
AdaBoost+j-48	61.57	73.00	57.32	76.33	68.73

Table 3 shown the result of evaluation in the final stage in the model construction process. Table reveals that bagging + j-48 tree classifier obtained highest accuracy as compared with other tree classifiers. Furthermore, significant increase was attained in dropout recall and graduated precision however, in terms of precision (dropout) and recall (graduated) bagging + J-48 did not show any increase in its performance as compared with j48 classifier. Another interesting result is the classifying performance of forest tree which obtained the highest value of 76.62 in the case dropout precision and 62.49 was obtained for graduated recall. In the case of ensemble model, it can been seen in the table that it is a minimal increase of their performances in all measured areas. The result reveals that combining these machine learning algorimhms have minimal increase in classification performance which can be attributed to the average of performances of true models and bad models. According to [58] that building ensembles with the most accurate models may not result in better ensembles. Furthermore, experiments by [59] show that in the case of substantial classification noise, bagging is much better than boosting, and sometimes better than randomization.

5 CONCLUSION

This study has explored and analysed the enrollment data that may have impact on the study outcome of the students and proposes a tree classification algorithms to help them choose a suitable program when they enter PSU-Urdaneta City Campus. The experimentation of tree classification algorithms and applying ensemble approach reveals that bagging + j-48 tree classifier obtained the highest accuracy as compared with other classifiers. However, forest tree classifier obtained the highest value of 76.62 in the case dropout precision and 62.49 was obtained for graduated recall. Minimal increase in their performance were recorded when ensemble approach was applied in all measure areas. The result reveals a very interesting prediction that the average grade in Math is a very important predictor for courses like Mathematics and engineering courses. In addition, for non engineering courses, the average grade in English, is the dominant predicting variable. In the case of IT, the average grade in TLE and Science is an important predicting variable. However, the result reveals that general high school point grade average is not a significant predictor for engineering programs however, the result reveals an interesting outcome that grade in English is not a significant predictor on the program that offer English

as major subject (ABEL) instead general grade average is its predicting variable. In addition, variables gender and type of school were graduated are not also significant predictors. The study shows the potential of data mining in higher education, especially when used to improve students' performance and detect early predictor of their success.

REFERENCES

[1] Macha, W., Mackie, C. , and Magaziner, J., Education in the Philippines, https://wenr.wes.org/2018 /03/education-in-the-philippines.

[2] Kotsiantis, S., & Pintelas, P. (2005). Local voting of weak classifiers. International Journal of Knowledge-based and Intelligent Engineering Systems, 9(3), 239-248.

[3] Tumapon, T. , Creating a culture of student engagement, http://www.manilatimes.net/creating-culture-student-engagement/347968/, Sept. 2017

[4] Radunzel, J., & Noble, J. (2012). Predicting Long-Term College Success through Degree Completion Using ACT [R] Composite Score, ACT Benchmarks, and High School Grade Point Average. ACT Research Report Series, 2012 (5). ACT, Inc.

[5] Geiser, S., & Santelices, V. (2006). The role of advanced placement and honors courses in college admissions. Expanding opportunity in higher education: Leveraging promise, 75114.

[6] Estrera, P. J. M., Natan, P. E., Rivera, B. G. T., & Colarte, F. B. Student Performance Analysis for Academic Ranking Using Decision Tree Approach in University of Science and Technology of Southern Philippines Senior High School.

[7] M. N. Quadril and N. V. Kalyankar, “Drop Out Feature of Student Data for Academic Performance Using Decision Tree Techniques,” Global Journal of Computer Science and Technology, vol. 10, no. 2, pp. 2 - 5, April 2010.

[8] Al-Barrak, M. A., & Al-Razgan, M. (2016). Predicting students final GPA using decision trees: A case study. International Journal of Information and Education Technology, 6(7), 528.

[9] D. M. D. Angeline, Association rule generation for student performance analysis using apriori algorithm, The SIJ Transactions on Computer Science Engineering & its Applications (CSEA) 1 (1) (2013) p12–16.

[10] M. M. Quadri, N. Kalyankar, Drop out feature of student data for academic performance using decision tree techniques, Global Journal of Computer Science and Technology 10 (2).

[11] E. Osmanbegovic, M. Sulji ´ c, Data mining approach for predicting student performance, Economic Review 10 (1).

[12] W. Ham ¨ al ¨ ainen, ¨ M. Vinni, Comparison of machine learning methods for intelligent tutoring systems, in: Intelligent Tutoring Systems, Springer, 2006, pp. 525–534.

[13] M. M. A. Tair, A. M. El-Halees, Mining educational data to improve students performance: a case study, International Journal of Information 2 (2).

[14] M. Mayilvaganan, D. Kalpanadevi, Comparison of classification techniques for predicting the performance of students academic environment, in: Communication and Network Technologies (ICCNT), 2014 International Conference on, IEEE, 2014, pp. 113–118.

[15] S. Natek, M. Zwilling, Student data mining solution–knowledge management system related to higher education institutions, Expert systems with applications 41 (14) (2014) 6400–6407.

[16] T. M. Christian, M. Ayub, Exploration of classification using nbtree for predicting students’ performance, in: Data and Software Engineering (ICODSE), 2014 International Conference on, IEEE, 2014, pp. 1–6.

[17] K. F. Li, D. Rusk, F. Song, Predicting student academic performance, in: Complex, Intelligent, and Software Intensive Systems (CISIS), 2013 Seventh International Conference on, IEEE, 2013, pp. 27–33.

[18] S. T. Jishan, R. I. Rashu, N. Haque, R. M. Rahman, Improving accuracy of students final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique, Decision Analytics 2 (1) (2015) 1–25.

[19] U. bin Mat, N. Buniyamin, P. M. Arsad, R. Kassim, An overview of using academic analytics to predict and improve students’ achievement: A proposed proactive intelligent intervention, in: Engineering Education (ICEED), 2013 IEEE 5th Conference on, IEEE, 2013, pp. 126–130. Amirah Mohamed Shahiri et al. / Procedia Computer Science 72 (2015) 414 – 422 421

[20] Using Decision Trees to Predict Student Placement and Course Success CAIR Conference November 21, 2014.

[21] Beaulac, C., & Rosenthal, J. S. (2018). Predicting University Students' Academic Success and Choice of Major using Random Forests. arXiv preprint arXiv:1802.03418

[22] Zhang, G., Anderson, T. J., Ohland, M. W., & Thorndyke, B. R. (2004). Identifying factors influencing engineering student graduation: A longitudinal and cross-institutional study. Journal of Engineering education, 93(4), 313-320.

[23] Veenstra, C. P. (2008). Modeling Freshman Engineering Success.

[24] Cengiz, N., & Uka, A. (2014). Prediction of Student Success Using Enrolment Data. KOS, 14(17), 45-2.

[25] Al-Radaideh, Q. A., Al Ananbeh, A., & Al-Shawakfa, E. (2011). A classification model for predicting the suitable study track for school students. Int. J. Res. Rev. Appl. Sci, 8(2), 247-252.

[26] Kovacic, Z. (2010). Early prediction of student success: Mining students' enrolment data.

[27] Aulck, L., Velagapudi, N., Blumenstock, J., & West, J. (2016). Predicting student dropout in higher education. arXiv preprint arXiv:1606.06364.

[28] Kabra, R. R., & Bichkar, R. S. (2011). Performance prediction of engineering students using decision trees. International Journal of Computer Applications, 36(11), 8-12.

[29] Kabakchieva, D. (2012). Student performance prediction by using data mining classification algorithms. International Journal of Computer Science and Management Research, 1(4), 686-690.

[30] Sembiring, S., Zarlis, M., Hartama, D., Ramliana, S., & Wani, E. (2011, April). Prediction of student academic performance by an application of data mining techniques. In International Conference on Management and Artificial Intelligence IPEDR(Vol. 6, No. 1, pp. 110-114).

[31] Shovon, M. H. I., & Haque, M. (2012). Prediction of student academic performance by an application of k-means clustering algorithm. International Journal of

Advanced Research in Computer Science and Software Engineering, 2(7).
[32] Yadav, S. K. and Pal, S. (2012). Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification. World of Computer Science and Information Technology Journal (WCSIT), 2(2), 51-56.
[33] Kovačić, Z. (2012). Predicting student success by mining enrolment data. Research in Higher Education Journal, 15(1).
[34] Simeunović, V. and Preradović, Lj. (2014). Using data mining to predict success in studying. Croatian Journal of Education, 16(2), 491-523.
[35] Cheewaprakobkit, P. (2015). Predicting student academic achievement by using the decision tree and neural network techniques. Catalyst, 12(2), 34-43
[36] Shah, N. S. (2012). Predicting Factors That Affect Students' Academic Performance by Using Data Mining Techniques. Pakistan Business Review, 13(4), 631-638.
[37] Nghe, N. T., Janecek, P. and Haddawy, P. (2007). A comparative analysis of techniques for predicting academic performance. 37th ASEE/IEEE Frontiers in Education Conference, 10-13th October 2007, Milwaukee.
[38] Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. Mis Quarterly, 553-572.
[39] M. Pandey and V. K. Sharma, "A Decision Tree Algorithm Pertaining to the Student Performance Analysis and Prediction," Int.J. Comput. Appl., vol. 61, no. 13, pp. 2–6, 2013.
[40] http://educationaldatamining.org/
[41] Nyce, C., & Cpcu, A. (2007). Predictive analytics white paper. American Institute for CPCU. Insurance Institute of America, 9-10.
[42] Lakshmi, T. M., Martin, A., Begum, R. M., & Venkatesan, V. P. (2013). An analysis on performance of decision tree algorithms using student's qualitative data. International Journal of Modern Education and Computer Science, 5(5), 18.
[43] Hemlata Chahal, "ID3 Modification and Implementation in Data Mining", International Journal of Computer Applications (0975-8887), Volume 80-No7, October 2013.
[44] Moghimipour, I., & Ebrahimpour, M. (2014). Comparing decision tree method over three data mining software. International Journal of Statistics and Probability, 3(3), 147.
[45] L. Rokach and O. Maimon. "Data mining with decision trees: theory and applications." World scientific, 2014.
[46] Gupta, Bhumika, Aditya Rawat, Akshay Jain, Arpit Arora, and Naresh Dhami. "Analysis of various decision tree algorithms for classification in data mining." Int J Comput Appl 8 (2017): 15-9.
[47] Hamilton, H. (2009). Computer Science 831: Knowledge discovery in databases, Retrieved from http://www2.cs.uregina.ca /~dbd/cs831/notes/ confusion_matrix/confusion_matrix.html
[48] Garg, R. (2018), A Primer to Ensemble Learning – Bagging and Boosting , https://www.analyticsindiamag.com/primer-ensemble-learning- bagging-boosting/
[49] Adhikari, R., Verma, G., & Khandelwal, I. (2015). A model ranking based selective ensemble approach for time series forecasting. Procedia Computer Science, 48, 14-21.
[50] Satyanarayana, A., & Nuckowski, M. (2016). Data mining using ensemble classifiers for improved prediction of student academic performance.
[51] Adejo, O. W., & Connolly, T. (2018). Predicting student academic performance using multi-model heterogeneous ensemble approach. Journal of Applied Research in Higher Education, 10(1), 61-75.
[52] Sewwandi, U. , (2018)How to create ensemble models using rapid miner, https://towardsdatascience.com/how-to-create-ensemble-models-using-rapid-miner-72a12160fa51
[53] Breiman L.1996. Bagging predictors, Machine Learning, 24(2):123-140.
[54] Hasan, M. R., Siraj, F., & Sainin, M. S. (2015, December). Improving ensemble decision tree performance using Adaboost and Bagging. In AIP Conference Proceedings (Vol. 1691, No. 1, p. 030008). AIP Publishing.
[55] Patel, S. , (2017). Machine Learning 101, https://medium.com/machine-learning-101/https-medium-com-savanpatel-chapter-6-adaboost-classifier-b945f330af06
[56] Shanthini A., G. Vinodhini and R.M. Chandrasekaran (2018), Predicting Students' Academic Performance in the University Using Meta Decision Tree Classifiers, Journal of Computer Science
[57] Smolyakov, V.,(2017) "Ensemble Learning to Improve Machine Learning Results How ensemble methods work: bagging, boosting and stacking", https://blog.statsbot.co/ensemble-learning-d1dcd548e936.
[58] Wang, W. (2008, June). Some fundamental issues in ensemble methods. In 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence) (pp. 2243-2250). IEEE.
[59] Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. Machine learning, 40(2), 1