

Article

Using Decision Trees and Random Forest Algorithms to Predict and Determine Factors Contributing to First-Year University Students' Learning Performance

Thao-Trang Huynh-Cam ^{1,*}, Long-Sheng Chen ^{1,*}  and Huynh Le ²

¹ Department of Information Management, Chaoyang University of Technology, Taichung 413310, Taiwan; thaotrang74@gmail.com or s10914905@gm.cyut.edu.tw

² Faculty of Medicine, Vo Truong Toan University, Tan Phu Thanh 930000, Vietnam; lehuynh10h@gmail.com

* Correspondence: lschen@cyut.edu.tw

Abstract: First-year students' learning performance has received much attention in educational practice and theory. Previous works used some variables, which should be obtained during the course or in the progress of the semester through questionnaire surveys and interviews, to build prediction models. These models cannot provide enough timely support for the poor performance students, caused by economic factors. Therefore, other variables are needed that allow us to reach prediction results earlier. This study attempts to use family background variables that can be obtained prior to the start of the semester to build learning performance prediction models of freshmen using random forest (RF), C5.0, CART, and multilayer perceptron (MLP) algorithms. The real sample of 2407 freshmen who enrolled in 12 departments of a Taiwan vocational university will be employed. The experimental results showed that CART outperforms C5.0, RF, and MLP algorithms. The most important features were *mother's occupations, department, father's occupations, main source of living expenses, and admission status*. The extracted knowledge rules are expected to be indicators for students' early performance prediction so that strategic intervention can be planned before students begin the semester.

Keywords: students' learning performance; prediction model; random forest (RF); decision tree (DT); feature selection; technological and vocational education



Citation: Huynh-Cam, T.-T.; Chen, L.-S.; Le, H. Using Decision Trees and Random Forest Algorithms to Predict and Determine Factors Contributing to First-Year University Students' Learning Performance. *Algorithms* **2021**, *14*, 318. <https://doi.org/10.3390/a14110318>

Academic Editor: Frank Werner

Received: 20 September 2021

Accepted: 28 October 2021

Published: 30 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Institutional research (IR) comprises a set of activities that support institutional planning, policy development, and decision making within higher education institutions (HEIs) [1]. In recent years, the urge to achieve excellence in research has led HEIs to have greater awareness of their roles in the entire educational management process and to place more strategic emphasis on the development of assessment tools for monitoring and evaluating the research quality [2]. In the USA and Japan, IR has been widely and successfully applied to evaluation, strategic planning, budget analysis, enrollment management, and research studies. Their studies focus on income analysis, research activities, and some issues reflecting strategic targets of HEIs. These studied issues might have some diversities from technical and vocational universities and colleges in Taiwan [3]. Thus, Taiwanese technical and vocational universities need to discover their own IR issues for specific targets and constraints.

Students, the indispensable participants in universities, their learning performance, and their attitudes towards these campuses should be seriously evaluated since they not only impact students' motivation, but also affect teaching quality and shape the design and delivery of university courses [4]. Specially, students' early performance prediction is important to academic communities so that strategic intervention can be planned before students reach the final semester. If universities in general, and Taiwanese technical and

vocational universities can, in particular, analyze students' learning data to understand the important variables of learning effectiveness, they not only can predict the strength and weakness of students' learning conditions, but can also propose preventive measures at the early stages. For students who may have outstanding academic performance, educational teams can invest resources to encourage students to strengthen their language, employment and research skills, help them find better opportunities, and set an example in order to help universities recruit more outstanding students. For students whose learning effectiveness is lagging behind, universities can provide additional remedial teaching and provide other measures to enhance schoolwork, such as providing teaching assistants and strengthening basic subjects and skills. In addition, HEIs need to continually increase the quality of teaching and the academic performance of their students [5].

In practice, students of Taiwanese technical and vocational universities often suffer from relatively low academic performance and rather high drop-out rates due to their fairly poor financial situation. However, finance is not the only factor affecting students' learning performance. According to the statistics of the drop-out rate in the 2019 academic year of the Ministry of Education of Taiwan government, the rate is 6.3% for general universities and 8.2% for technical colleges. In addition, there are 186,446 people who leave school each year, accounting for 15.3% of all tertiary students. Among them, the majority drop-out students leave schools after the first year. The biggest factor for leaving tertiary education, aside from lack of interest, is poor academic performance. Therefore, to build a prediction model for learning to avoid dropping out is extremely important.

In recent years, machine learning algorithms, and artificial intelligence (AI) [6] have been widely applied to predict students' learning performance and to find the important features that have high impact on students' academic performance. Machine learning techniques were employed in [7] to examine the effect of co-curricular activities on a student's academic performance. Tree-based models and artificial neural networks (ANN) [8] were built in [5] to analyze students' academic performance in virtual learning. In the latest research, explainable artificial intelligence (AI) refers to methods, which can produce accurate and explainable models of AI algorithms [9]. Thus, AI solution results can be understood by humans. Following this trend, this study will use machine learning algorithms including decision trees (DT) and random forests (RF) algorithms, which can generate explainable results, to predict freshmen's academic performance. Except for DT and RF, multilayer perceptron (MLP) [10] will be performed as our comparison base.

The prediction of first-year student academic achievements has received substantial attention in educational practice and theory [11]. Previous works used some variables, such as resilience, engagement [12], scores of quizzes and assignments [13,14], students' academic self-concept [15], motivation, social relationships [16], and participation [11], to construct prediction models of first-year academic achievements. However, the information on these variables in the research can only be obtained during the course or in the progress of the semester. Some information also needs to be obtained through survey questionnaires and interviews. This is not enough to improve students' learning performance in time, which is especially true for those students who are performing poorly due to economic factors.

In Taiwanese vocational universities, the majority of students are economically disadvantaged. They often need to rely on government tuition, miscellaneous fee waivers, and student loans to register. In addition, they must work part-time every month to support themselves and their family's living expenses. In addition to lack of interest, the biggest reason for dropping out is due to poor learning results. Therefore, the models established in published works [11–16] and the prediction results are often less time-sensitive. A predictive model needs to be established before the semester begins to provide student counseling, financial assistance, and supplements. The annotation of teaching resources could be more accurate and more immediate. Therefore, this study attempts to use family background variables, including department, gender, address, admission status, Aboriginal status, child of new residents, family children ranking, on-campus accommodation,

main source of living expenses, student loan, tuition waiver, parents' average income, status, occupations, and education. These variables can be obtained before the start of the semester, in order to construct predictions before the freshmen students start to learn, and thus buy more time for student guidance or investing learning resources in technological and vocational education. In sum, this paper aims to build a prediction model that can be used to predict freshmen students' learning performance based on decision trees and random forest algorithms. The sample was 2407 freshmen who enrolled in 12 departments of a university in Taiwan. From this constructed model, we can determine which students will succeed and which students indicate to be poor; the university is then able to offer them necessary assistance before they start their sophomore year. Based on experimental results, we can highlight some factors, which highly affect the first-year undergraduates' learning performance.

2. Literature Review

2.1. *The Learning Performance of First-Year Students*

Students' learning performance plays a vital role in universities since it affects both individual and organizational performance [17,18]; therefore, studies on factors and variables affecting students' learning performance have been in existence for decades and have continuously attracted an increasing number of diverse researchers. In 1975, in [19,20], four factors were identified as causing poor students' academic performance: (1) society, (2) school, (3) family, and (4) student. In contrast, general factors affecting successful learning performance were highlighted in [21]. Particularly, authors in [22] reported that the factors, such as gender, students' ages, and students' high school scores in mathematics, English, and economics affected university students' scores and they also concluded that students with high scores in their high schools performed better in their university level. Additionally, authors in [23] studied the relationship between students' matriculation exam scores and their academic performance and found that a student's admission scores positively affected their undergraduates performance.

The idea of applying data mining in the educational system attracted authors in [24] in 2007 since data mining can show discovered knowledge to educators and academic teams, and show recommendations to students. Moreover, authors in [25] used ANN for university educational systems while the authors in [19] applied ANN in a narrower field of academic performance prediction in university. Particularly, Oladokun et al. [19] utilized an ANN model to predict students' academic performance based on factors, such as ordinary level subjects' scores and subjects' combination, matriculation exam scores, age on admission, parental background, types and location of secondary schools attended, and gender. Students' learning performance was predicted based on their average point scores (APS) of Grade 12 in [8], on high school scores in [17], and on cumulative grade point average (CGPA) in fundamental subjects [18].

The predictors of first-year student success has received much attention in educational practice and theory [11]. Consequently, many researchers have paid attention to this issue. For example, Ayala and Manzano [12] investigated whether or not a relationship between the dimensions of resilience and engagement, and the academic performance of first-year university students. Baneres et al. [13] (2019) aimed to identify at-risk students by building a predictive model using students' grades. Their model can predict at-risk students during the semester on a first-year undergraduate course in computer science. Neumann et al. [15] focused on first year international students in undergraduate business programs at an English-medium university in Canada. They found there to be a positive relationship between students' academic self-concept and subsequent academic achievement. In the work of Anderton [26], he indicated gender and the Australian Tertiary Admissions Rank as significant predictors of academic performance. After surveying 80 published articles, Zanden et al. [11] found that some predictors contributed to multiple domains of success, including students' previous academic performance, study skills, motivation, social relationships, and participation in first-year programs.

We can establish from these published works the variables used, such as resilience, engagement, scores of quizzes and assignments, students' academic self-concept, motivation, social relationships, and participation to build prediction models. However, the information on these variables used in the literature can only be obtained during the course or in the progress of the semester. As well, some information needs to be obtained through questionnaires and interviews. This shortens the time for universities to take remedial measures, especially for some students of poor learning performance, caused by economic factors. In practice, obtaining this information and then making predictions based is too slow to prevent students from dropping out due to poor academic performance. Therefore, this study attempts to use family background variables, including department, gender, address, admission status, Aboriginal status, child of new residents, family children ranking, on-campus accommodation, main source of living expenses, student loan, tuition waiver, parent's average income, status, occupations, and education. These variables can be obtained before the start of the semester, allowing to make predictions before the freshmen students start to learn, and providing more time for student guidance or investing in learning resources.

2.2. Decision Trees

Decision trees (DT) are widely applied for prediction and classification in domain of machine learning [27]. DT have the advantages of simple use, easy understanding, high accuracy, and high prediction ability [28–30]. In recent years, decision trees have been successfully applied in education areas [6,29–38]. For example, Wang et al. [33] proposed a higher educational scholarship evaluation model based on a C4.5 decision tree, while Hamoud et al. [34] used DT to predict and analyze student behaviors. Their results indicated that students' health, social activities, interpersonal relationships, and academic performance affected learning performance. Furthermore, authors in [27] used the DT method to conduct research on students' employment wisdom courses in order to provide solutions for training professionals and employment courses, and to solve the contradiction between training plans and enterprise needs. A semi-automated assessment model was built by using DT in [35].

There are a variety of DT algorithms, such as ID3, C4.5, C5.0 (a commercial version of C4.5), and CART (classification and regression tree). Among them, C4.5 and CART algorithms are the most popular and have many useful applications [33]. Compared with other classification methods, such as ANN and support vector machines, the decision tree can extract readable knowledge rules, which is helpful for university-side decision-making reference [34,35]. Therefore, this study will use decision trees algorithms, including C5.0 and CART, to build DT prediction models.

2.3. Random Forests

Random forests (RF) are regarded as an effective method in machine learning since RF can solve the problems of over-training [39,40], which decision trees may face. RF operates classification, regression, and other tasks by constructing multiple decision trees during training [41–43]. The calculation method is to evaluate multiple independent DT and determine the result through their voting results. When each node in DT is split using the best among the attributes, "each node in RF is split using the best among the subset of predictors randomly chosen at the node" [40]. RF has been widely applied to IR in universities. For example, in the work of [38], they used RF to predict if a student would obtain an undergraduate degree or not using the learning performance of the first two semesters of courses completed in Canada. Ghosh and Janan [16] utilized 24 variables, including creating good notes, group study, adaptation to university, and self-confidence, which were obtained from a questionnaire survey. RF was then employed to predict the first-year student performance of a university in Bangladesh. From the above literature, we can establish that RF has been successfully applied to predict students' learning performance. Therefore, this study also applied RF as one of the candidate algorithms to predict the

learning performance and to identify features, which importantly affect first-year students' learning performance.

2.4. Artificial Neural Networks

An artificial neural network (ANN) is a computational system which mimics the neural structures and the process of human brains, including biological structure, processing capacity, and learning ability. ANNs can receive input data, analyze, and process information, and provide output data/actions through a large number of interconnected "neurons" or nodes. It is the foundation of artificial intelligence (AI) and solves problems, which are difficult and/or impossible to be carried out by humans. However, ANNs must be trained with a large amount of data/information through mathematical models and/or equations because ANNs cannot understand, think, know, and process data like the human nervous system. There are two types of ANN: supervised learning and unsupervised learning. Supervised learning is a process of supervising or teaching a machine/computer by feeding it input data and correct output data, which is referred to as a "labelled dataset" so that the machine/computer can predict the outcome of sample data. Supervised learning is the machine learning task of learning that maps an input to an output based on sample input–output pairs. Unsupervised learning uses machine learning algorithms, which draw conclusions on an "unlabeled dataset". Data must then be determined based only on input data.

ANN has been applied in numerous applications with considerable attainment. ANN have been effectively and efficiently applied in the area of prediction [44,45] since ANN can be used to predict future events based on historical data. In addition, a deep learning algorithm and neural network [46–50] have been proposed for university student performance prediction. Dharmasaroja and Kingkaew [49] used ANN to predict learning performance in medical education. In their work, they used demographics, high-school backgrounds, first-year grade-point averages, and composite scores of examinations during the course to be input variables. Sivasakthi [50] utilized MLP, Naïve Bayes, and DT to predict introductory programming performance of first year bachelor students.

In the works of [20,39], MLP was applied to build a model for predicting student performance and had good results. Therefore, we use MLP to be our comparison base in this study.

3. Methodology

The experimental process of this study included five steps as shown in Figure 1.

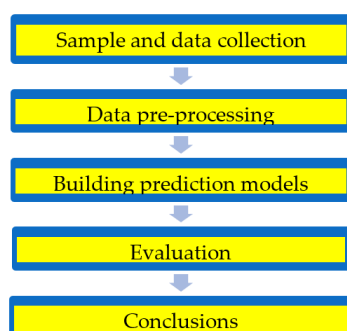


Figure 1. The experimental process.

3.1. Sample and Data Collection

This research was conducted at the end of the first semester of the academic year 2020–2021 at one technical and vocational university in Taiwan. The data for the experimental models were collected through the school register system and school grading system. When students first enroll in this university, they were required to fill in their personal

information in an electronic form through the school register system. Then, during the learning process, all subjects' grades and achievements of every student were recorded in the school grading system. Therefore, at the research time, each student's registered profile included 18 personal information variables and one variable of average scores of all the subjects' grades, which they learned in the first semester.

3.2. Data Pre-Processing

In the data pre-processing phase, we performed data clean and data normalization steps. In the data clean step, we dealt with missing value examples and processed category data, after determining the 18 input and output variables (learning performance). In this step, we removed all examples that contain missing values, and encoding category data.

In data normalization step, the data was normalized according to Equation (1).

$$X_{\text{mon}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad (1)$$

where X_{max} is the maximum value, X_{min} is the minimum value, and X_{mon} is the normalized value.

3.3. Building Prediction Models

The study employed the experiments on Windows Operating Systems with a 3.80 GHz Intel(R) Xeon(R) E-2174G CPU and 64 GB of RAM. Four supervised learning models based on MLP, random forest (RF) and decision tree (DT) algorithms were developed. C5.0 and CART algorithms were used to build DT prediction models while the python (version 3.7.1) programming language was used to build RF prediction models. The experiment was carried out five times on each model. The mean values and standard deviation of the classification performance in each model were then taken and used as the benchmark for measuring the DT and RF models. The aims of various experiments were to investigate and benchmark their performance in predicting freshmen's learning performance on the dataset and to select features which highly affect students' learning performance.

Furthermore, there are three cases of output data in this experimental study as follows:

- Case 1 is the origin case for the output: Excellent, Very Good, Good, Average, and Poor class to measure the four models' prediction performance originally and generally.
- Case 2 is to combine the majority output: Very Good, Good, Average classes into the Normal class to investigate whether the four models predict the minority or not.
- Case 3 is to focus only on the minority output: Excellent and Poor classes.

3.3.1. Decision Trees (DT)

The experimental process of C5.0 algorithm for all the three cases in this study included the following steps.

- (1) Create training and testing data
- (2) Set decision tree parameters
- (3) Create an initial rule tree
- (4) Prune this tree
- (5) Process the pruned tree to improve its understandability
- (6) Pick a tree whose performance is the best among all constructed trees
- (7) Repeat steps 1–6 for 10 experiments
- (8) Take the mean values and standard deviation of the classification performance in 10 experiments for benchmarking.

We used a 10-fold cross validation (CV) experiment and constructed a DT for each fold of the data set based on the C5.0 algorithm. The collected data sets were divided into 10 equal sized sets and each set was then in turn used as the test set. Beside the test set, we used 9 other sets as our training set to build DT. Therefore, we had 10 trees. The tree,

which had the best performance, was picked out and all attributes left in this tree were considered as important.

Apart from the C5.0 algorithm, after extracting the DT experimental results, this study utilized the CART algorithm by python as the other technique to test, compare and measure the prediction accuracy and feature importance selection between C5.0 and CART. The experimental process of CART algorithm for all the three cases was as follows:

- (1) Create training and testing data.
- (2) Set DT parameters.
- (3) Process the DT with training, testing, and cross validation for prediction accuracy.
- (4) Plot the Gini feature importance results.
- (5) Repeat steps 1–4 for 10 experiments.
- (6) Take the mean values and standard deviation of the classification performance in 10 experiments for benchmarking.

3.3.2. Random Forest (RF)

The RF experimental process in this study consists of the following steps.

- (1) Create training and testing data.
- (2) Set random forest parameters.
- (3) Process the RF with training, testing, and cross validation for prediction accuracy.
- (4) Plot the Gini feature importance results.
- (5) Repeat steps 1–4 for 10 experiments.
- (6) Take the mean values and standard deviation of the classification performance in 10 experiments for benchmarking.

3.3.3. Multilayer Perceptron (MLP)

MLP [39] is a multi-layer structure composed of an input layer, a hidden layer, and an output layer, the input layer receives data, the hidden layer processes the data, and the output layer is responsible for the final output of the model. The MLP experimental process in this study consists of the following steps.

- (1) Set the initial weight and deviation value
- (2) Input training data and target data
- (3) Calculate the error between the expected output and the target
- (4) Adjust the weight and update the network weight
- (5) Repeat step (3)~step (4) until the end of learning or convergence.

4. Experimental Results

After pre-processing, the dataset was imported to both See5 software to implement C5.0 algorithm and jupyter software to implement MLP, and both RF and CART algorithms, i.e., DT models were conducted in two different algorithms: C5.0 and CART. Every model was implemented 10 times in each software with 10 different training and testing dataset in which the students' learning performance variables were divided into three different cases (Table 1):

- Case 1: EX-VG-G-AVG-Poor (Excellent-Very Good-Good-Average-Poor) classification,
- Case 2: EX-Normal-Poor (Excellent-Normal-Poor) classification, and
- Case 3: Ex-Poor (Excellent-Poor) classification.

The experimental results of four models in each case will be presented in the following sections.

Regarding parameter settings, in RF, and the number of trees in the forest was set to 100. For the decision tree, in C5.0 and CART, pruning CF affects the way of estimating the error rate, thereby affecting the severity of pruning, in order to avoid overfitting of the model. In this study, pruning CF was set to 25%. In MLP, the learning rate was set to 0.3, and the training stop condition was set to the number of learning iterations to 1000. At

this time, the RMSE (root-mean-square error) has been flattened, representing the network has converged.

Table 1. The output data transformation.

Case No.	Transferred Number	Output Variable	Average Score	Distribution/Classification
Case 1: Origin	5	Excellent (EX)	90–100 points	Origin case: EX-VG-G-AVG-Poor classification
	4	Very Good (VG)	80–89 points	
	3	Good (G)	70–79 points	
	2	Average (AVG)	50–69 points	
	1	Poor	0–49 points	
Case 2: Combination of majority	3	Excellent	90–100 points	Combination of majority: EX-Normal-Poor classification
	2	Normal	50–89 points	
	1	Poor	0–49 points	
Case 3: Focus on minority	2	Excellent	90–100 points	Focus on minority: EX-Poor classification
	1	Poor	0–49 points	

4.1. Data Preprocessing

The learning performance prediction data set had 4375 first-year students enrolled in 12 departments of a Taiwanese university during the first semester of the academic year 2020–2021. These departments were selected randomly. However, after data cleaning, only 2407 usable numbers of students were selected for the experimental sample data since all variables were fulfilled in students' profiles, resulting in a return rate of 55%. The remaining 1968 students (45%) who had missing variables in their profiles, dropped out, and/or were suspended, were excluded in this study.

After relevant data sets were processed, a total number of 18 factors, which were predicted to influence the learning performance of freshmen students, were used as input (independent) variables for the prediction model (Table 2). These proposed factors included "Department", "Gender", "Address", "Admission status", "Aboriginal", "Child of new residents", "Family children ranking", "Parent average income per month", "On-campus accommodation", "Main source of living expenses", "Students' loan", "Tuition waiver", "Father live or not", "Father's occupations", "Father's education", "Mother live or not", "Mother's occupations", and "Mother's education". The factor "Average scores" of all the subjects' grades in the first semester recorded in the school grading system was used as output (dependent) variable for the model (Table 1).

Table 2. The input variables.

No.	Feature Name	Feature Description	No.	Feature Name	Feature Description
1	Department	Students' majored department	10	Main source of living expenses	Students' living expenses support
2	Gender	Students' sex	11	Student loan	Students borrow money in- and out-school or from friends/relatives
3	Address	Students' home address type	12	Tuition waiver	Free or reduce tuition fee
4	Admission status	School admission offered to student	13	Father live with or not	Students' father lives in the family or separates with their mother
5	Aboriginal	Students' origin	14	Father's occupation	Careers of students' father
6	Child of new residents	Immigration status of students' family	15	Father's education	Father's highest education status
7	Family children ranking	Student's born time in the family	16	Mother live with or not	Students' mother lives in the family or separates with their father
8	Parent average income per month	The average income of student's parents in each month	17	Mother's occupations	Careers of students' father
9	On-campus accommodation	Students live in- or out of the school.	18	Mother's education	Mother's highest education status

4.2. Definition of the Input and Output Variables

Table 2 reports 18 selected factors for input (independent) variables, including feature names and their description. Table 1 shows the output (dependent) variable, the classification of the chosen output variables, which follow the grading system, and how the output was distributed in this study. For the scope of this paper, the domain of the output variable represents the average score of all the subjects' grades in the first semester of the academic year 2020–2021 of the freshmen.

4.3. Experiment Results

4.3.1. Results of Case 1 and Case 2

Table 3 shows results of Case 1, which is our original data. For Case 1, the mean values (standard deviation) of overall accuracy are 51.20% (0.44%), 47.86% (0.68%), 52.61% (0.7%), and 41.67% (1.70%) for CART, C5.0, RF, and MLP, respectively. From Table 3, we can find all models built by these algorithms cannot achieve an acceptable performance. The reason may be that we divided too many class labels (EX, VG, G, AVG, Poor). Therefore, we combined the majority (VG, G, AVG) into a new class label (Normal) for Case 2 because we expected the models can predict the minority.

Table 3. Results of Case 1 and Case 2.

Experiment No.	Case 1: Origin EX-VG-G-AVG-Poor Classification				Case 2: Combination of Majority EX-Normal-Poor Classification			
	CART (%)	C5.0 (%)	RF (%)	MLP (%)	CART (%)	C5.0 (%)	RF (%)	MLP (%)
1	51.86	47.80	52.69	42.32	87.75	91.60	89.62	90.66
2	51.24	47.30	51.86	39.83	86.92	91.60	89.62	90.66
3	50.62	47.90	52.48	41.28	87.34	91.60	89.62	90.66
4	50.82	47.20	52.28	40.87	87.55	91.60	89.62	90.66
5	51.45	49.10	53.73	43.15	87.13	91.60	89.62	90.66
6	52.28	46.20	52.28	40.66	88.17	91.60	89.62	87.75
7	52.07	49.10	52.69	41.70	88.17	91.60	89.62	90.24
8	51.86	47.90	52.07	41.28	88.38	91.60	89.62	90.04
9	50.82	47.30	53.11	45.58	86.92	91.60	89.62	88.58
10	52.07	48.60	52.28	40.04	87.75	91.60	89.62	89.21
Mean	51.51	47.84	52.55	41.67	87.61	91.60	89.62	89.91
Standard Deviation	0.60	0.91	0.55	1.70	0.53	0.00	0.00	1.05

Table 3 also lists results of Case 2. For Case 2, the mean values (standard deviation) of overall accuracy are 87.50% (0.44%) for CART, 91.60% (0%) for C5.0, 89.62% (0%) for RF, and 89.91% (1.05%) for MLP. The prediction accuracies have been significantly improved. Among these four algorithms, C5.0 outperforms MLP, CART, and RF.

Table 4 reports the confusion matrix of C5.0 in Case 2. It is obvious that C5.0 algorithm cannot recognize the minority classes (EX and Poor). In other words, the constructed prediction models by C5.0 algorithm cannot identify excellent and poor students. Those minority are usually important for HEIs' management to invest teaching resources and offer special assistance.

Table 4. Confusion matrix of C5.0 in Case 2 (fold 1).

Actual \ Predicted	EX	Normal	Poor
EX	0	93	0
Normal	0	2206	0
Poor	0	108	0

For our research purposes, this prediction model can only find normal students. The students who need tutoring with poor learning effectiveness and the gifted students who need additional teaching resources to achieve higher achievements will not be identified. Therefore, we implemented another experiment similar to Case 3 in which we focused only the minority classes: Excellent and Poor.

4.3.2. Results of Case 3

In Case 3, we only used two class labelled samples to build prediction models. For Case 3, we focused on the Excellent and Poor classes. Table 5 lists results of Case 3. From this table, we can find that the mean values (standard deviation) of overall accuracy are 79.82% (0.91%) for CART, 74.52% (0.41%) for C5.0, 79.02% (4.43%) for RF, and 69.02% (7.28%) for MLP.

Table 5. Results of Case 3.

Experiment No.	Case 3: Focus on Minority			
	EX-Poor Classification			
	CART (%)	C5.0 (%)	RF (%)	MLP (%)
1	80.04	75.20	82.92	73.17
2	80.04	74.60	73.17	60.97
3	80.48	74.10	75.60	63.41
4	78.04	74.60	80.48	60.97
5	80.48	74.10	82.92	63.41
6	78.04	72.70	82.92	65.85
7	80.48	72.70	82.92	82.92
8	81.48	75.20	82.92	73.17
9	82.92	76.60	80.48	70.73
10	78.04	76.10	75.60	75.60
Mean	80.00	74.59	79.99	69.02
Standard Deviation	1.60	1.28	3.78	7.28

In order to validate the difference between CART, RF, C5.0, and MLP, we implemented one way ANOVA. Null hypothesis is “All means are equal” and alternative hypothesis is “At least one mean is different”. The significance level (α) is set as 0.05. From Table 6, we can reject null hypothesis due to the p -value (0.000) is less than 0.05.

Table 6. Analysis of Variance.

Source	DF	Adj SS	Adj MS	F-Value	p -Value
Factor	3	826.5	275.49	15.43	0.000
Error	36	642.6	17.85		
Total	39	1469.0			

To find the best prediction models, 6 statistical hypotheses under 95% confidence level have been carried out using two-sample t -test. Table 7 lists the results of statistical hypotheses tests. From the results of H1 and H2, we can find CART has no significant difference compared to RF. From H3 to H6, the p -values are all less than 0.05. Consequently, for these four hypotheses, we reject all null hypotheses. It means CART is better than C5.0 and MLP; RF is better than C5.0 and MLP. In sum, it can be concluded that CART is slightly better than RF since the difference is not significant. And both CART and RF are significantly superior to C5.0 and MLP. In this case, CART is superior to MLP, C5.0, and RF.

Table 7. Results of statistical hypotheses for comparison.

No.	Hypothesis	<i>p</i> -Value	Conclusion
H1	$H_0 : CART \leq RF$ $H_1 : CART > RF$	0.497	Accept H_0
H2	$H_0 : CART = RF$ $H_1 : CART \neq RF$	0.993	Accept H_0
H3	$H_0 : CART \leq C5.0$ $H_1 : CART > C5.0$	0.000	Reject H_0
H4	$H_0 : CART \leq MLP$ $H_1 : CART > MLP$	0.001	Reject H_0
H5	$H_0 : RF \leq C5.0$ $H_1 : RF > C5.0$	0.001	Reject H_0
H6	$H_0 : RF \leq MLP$ $H_1 : RF > MLP$	0.000	Reject H_0

To make a fair comparison, we also provided a confusion matrix of C5.0 in Table 8. From Tables 4 and 8, we can establish the ability of identifying excellent students is increased from 0% to 69.59%. The accuracy of predicting the learning poor has been improved from 0% to 83.33%. These four algorithms can precisely predict Excellent and Poor classes.

Table 8. Confusion matrix of C5.0 in Case 3 (fold 1).

Actual \ Predicted	EX	Poor
EX	61	32
Poor	18	90

4.3.3. Results of Importance Feature Selection

In DT algorithms, the nodes left in the constructed trees will be considered as important. Table 9 provides the extracted top five important features for three cases in the three models. However, in Case 1 and Case 2, the extracted features only can be used to identify the majority students. In Case 3, the discovered features could be used to predict excellent and poor students.

Table 9. Comparisons of DT and RF in the top 5 important features.

Algorithm	Case 1: EX-VG-G-AVG-Poor Classification	Case 2: EX-Normal-Poor Classification	Case 3: EX- Poor Classification
C5.0	Father's occupation Mother's occupation Department Admission status Main source of living expenses	X	Mother's occupations Main source of living expenses Admission status Department Family children ranking
CART	Father's occupation Mother's occupation Department Parent average income per month Fathers' education	Father's occupation Mother's occupation Department Parent average income per month Main source of living expenses	Mother's occupation Department Fathers' occupation Main source of living expenses Admission status
RF	Father's occupation Mother's occupation Department Parent average income per month Fathers' education	Mother's occupation Department Parent average income per month Main source of living expenses	Mother's occupation Department Main source of living expenses Fathers' occupation Admission status

In Case 3, CART algorithm had the best performance. Consequently, we used results of CART to select important features. Figure 2 shows the rank of Gini importance of CART for Case 3. From Table 9 and Figure 2, we can find the top five important features. They are “Mother’s occupations”, “Department”, “Fathers’ occupations”, “Main source of living expenses”, and “Admission status”.

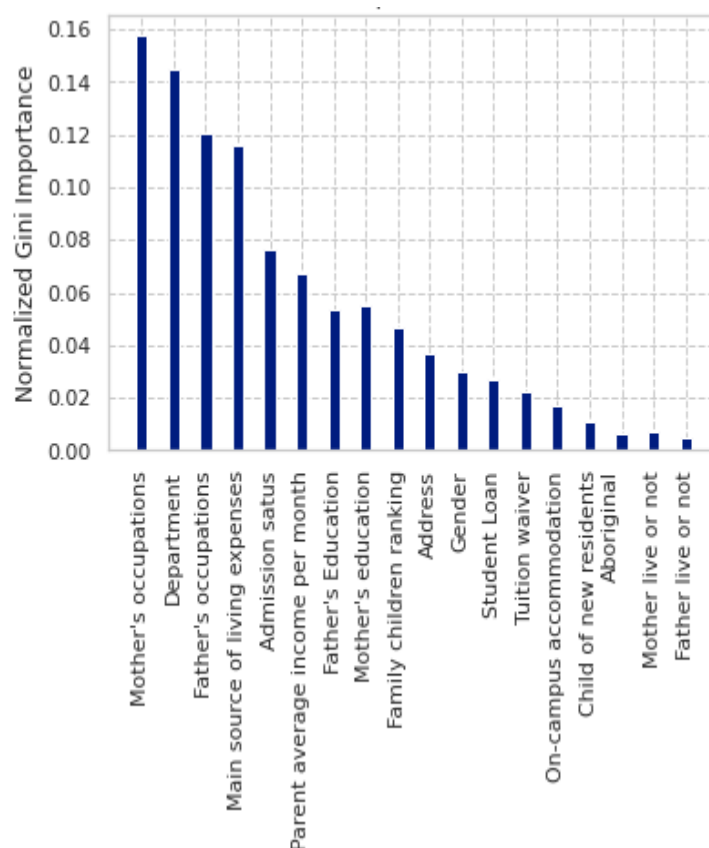


Figure 2. Gini importance from CART for Case 3: Ex- Poor classification.

4.4. Extracted Rules from Decision Trees

Table 10 summarizes all the knowledge rules extracted from decision trees. Rules 1 to 13 can be used to predict the freshman academic performance. These rules will be discussed in details in the following sections.

Rule 1 to Rule 9 are for predicting students with excellent academic performance.

- Rule 1 shows that the on-the-job students are hardworking and have excellent academic performance.
- Rule 2 reports that if the main source of living expenses comes from family support, and the mother is a housewife who does not need to earn money for living can pay full attention to her children’s education, it is not surprising that such students will perform well in their studies.
- Rule 3 displays that when students of TF2 department live in the dormitory on campus, their academic performance will be excellent because the on-campus dormitory is mainly provided for economically disadvantaged students. Therefore, living in the dormitory inside the school is less expensive. Moreover, there is an unnecessary daily commute, students can fully use the on-campus library and other learning resources, thus the learning performance is naturally excellent. In the future, the accommodation for the TF2 department students should be arranged for the on-campus dormitory.

- Rule 4 points out that if students' sources of living expenses come from their families, and the occupation of their mothers is as a government employee, they will have excellent academic performance.
- Rule 5 is also for specific departments. If students of TD5 department pay for student loans, their academic performance will be very good.
- Rule 6 points out that if the father's occupation is a government employee, the students' academic performance will be excellent.
- In Rule 7, if the source of living expense comes from scholarships and grants from inside and outside the school, students will perform very well.
- Regarding Rule 8, for female students, if the mother is a full-time housewife, they will perform well.
- Rule 9 also indicates that if the mother's occupation is an educator, the student's performance will also be very good.

Table 10. Extracted rules by decision trees.

No.	Rules
1	IF Admission status = <i>On-the-job student</i> THEN Learning performance = Excellent [0.909]
2	IF Main source of living expenses = <i>Family provided</i> AND Mother's occupations = <i>Housewife</i> THEN Learning performance = Excellent [0.900]
3	IF On-campus accommodation = <i>Yes</i> AND Department = <i>TF2</i> THEN Learning performance = Excellent [0.857]
4	IF Main source of living expenses = <i>Family provided</i> AND Mother's occupations = <i>Government employees</i> THEN Learning performance = Excellent [0.850]
5	IF Student Loan = <i>Yes</i> AND Department = <i>TD5</i> THEN Learning performance = Excellent [0.833]
6	IF Father's occupations = <i>Government employees</i> THEN Learning performance = Excellent [0.800]
7	IF Main source of living expenses = <i>Scholarships and grants inside and outside the school</i> THEN Learning performance = Excellent [0.800]
8	IF Gender = <i>Female</i> AND Mother's occupations = <i>Housewife</i> THEN Learning performance = Excellent [0.769]
9	IF Mother's occupations = <i>Education</i> THEN Learning performance = Excellent [0.750]
10	IF Gender = <i>Male</i> AND Mother's occupations = <i>Housewife</i> THEN Learning performance = Poor [0.889]
11	IF Student Loan = <i>No</i> AND Department = <i>TD5</i> THEN Learning performance = Poor [0.857]
12	IF Main source of living expenses = <i>Income from part-time job</i> THEN Learning performance = Poor [0.850]
13	IF Admission status = <i>Transfer student</i> THEN Learning performance = Poor [0.800]

From the above rules, we can see that the occupation of parents can determine the academic performance of freshmen students, especially government employees and educators who have a high education level. In addition, if the mother is a full-time housewife,

she can devote all her energy to student learning. It can also contribute to outstanding performance in learning. We can also see that if the financial resource is intact, whether it comes from family supply or scholarships inside and outside the school, it will also be quite helpful for students's learning.

Rule 10 to Rule 13 are for predicting students with extremely poor academic performance.

- Compared with Rule 2, Rule 10 has a clear contrast for male students, if the mother is a housewife, the academic performance will be poor. This results from the patriarchal tradition of Taiwanese society. Housewife mothers spoil their sons, which can cause this phenomenon. Therefore, it is necessary to carry out stricter learning supervision for the male students before the senior years.
- Rule 11 is for the TD5 department. If students in that department do not have student loans, i.e., they have better family background, their academic performance will be quite poor. This can be inferred that if the rich families do not have strict requirements for their children's education, their family member's academic performance will be poor. In this case, more than 50% of the students, who paid for student loans, received government financial subsidies, and tuition reductions or exemptions over the years are consistent among Taiwanese private vocational universities. The students enrolled in TD5 also have low admission scores. Therefore, the university can provide intensive study guidance and strict schoolwork supervision for those students who are not doing well financially, in the departments with low admission scores.
- Rule 12 reflects the general situation of students in private vocational universities in Taiwan. If the source of living expenses is mainly from part-time jobs, then those students' academic performance will also be poor. At this point, the government has launched a program of "purchasing working hours", which allows economically disadvantaged students to invest in studies by paying work-study fees. They can get financial support and promote social class mobility as with doing part-time jobs.
- Rule 13 states that if a freshman is a transfer student, academic performance will be quite poor. Therefore, for the transfer students who enter the school in the first year, the student guidance system will help them integrate into class and establish contacts. After solving the possible problems, the school's remedial teaching methods can be effective.

Since most students in Taiwanese private vocational universities are economically disadvantaged, these rules have a good reference value for Taiwanese private vocational universities.

5. Discussion and Conclusions

In practice, the prediction models built in Case 3 are more meaningful than models of Case 1 and Case 2. Therefore, we focus on results of Case 3. In this case, the experimental results showed that prediction accuracy mean rate of RF 10-fold experiments was nearly 79.99%, that of DT 10-fold experiments was 74.59% by C5.0 algorithm and 80.00% by CART algorithm, and that of MLP 10-fold experiment was 69.02%. CART outperforms C5.0, RF, and MLP algorithms.

For Case 3, the selected factors, which most influenced freshmen's learning performance were "Mother's occupations", "Department", "Father's occupations", "Main source of living expenses", and "Admission status". Importantly, the two factors: "Mother's occupations" and "Department" had the highest significant impact on first-year students' learning performance; whereas four factors: "Father live or not", "Mother live or not", "Child of new residents", and "Aboriginal" had the least effect on freshmen's learning performance. The analysis results are expected to be a roadmap for students' early performance prediction so that strategic intervention can be planned before students reach the final semester. The results of prediction model and those discovered to be important factors also can be used as leading indicators to prevent students from being dropped out due to poor learning performance.

From the extracted knowledge rules of decision trees, we have discovered some useful information. To predict excellent students, the occupation of parents can determine the

academic performance of freshmen students, especially when parents' occupations are government employees and teachers who have higher education backgrounds. Moreover, if the mother is a housewife, it can also contribute to outstanding academic performance. It also could be found that if the financial resource is intact, whether it comes from family supply or scholarships, it will also be quite helpful for students' learning.

To predict students of extremely poor academic performance, we also discovered some rules. The technological and vocational universities should focus on transfer students and those students whose living expenses is mainly from part-time jobs. Generally, their learning performance will be poor and they require additional guidance.

In this study, we used family background variables, which can be obtained in the beginning of freshmen semester to predict students' learning performance. We can use the established models to predict the academic performance of freshmen as soon as they enter the school. If a student is predicted with poor learning performance, educational teams can carry out early-warning counseling measures, such as reminding class tutors to pay more attention to them. In the case of negative influence of part-time jobs on the absences and poor learning situations, educational teams can offer early remedial teaching resources or teaching assistants for individual tutoring. These proposed measures can effectively prevent these poor students from falling behind in their learning process.

For students who are predicted for excellent academic performance, universities can focus on elite-style tutoring, such as special classes for professional and technical advancement, license examination training, entrepreneurial competitions and other employment skills enhancement. For undergraduates who are planning to enter higher education programs, universities can offer more support for foreign language skills development and entrance examinations.

In sum, this study successfully built prediction models for freshmen's academic performance using CART, C5.0, RF, and MLP algorithms in a Taiwanese vocational university. Five important features have been determined to take advanced actions for HEIs management. For potential direction of future works, other machine learning algorithms could be applied. In addition, more input variables could be included in the future. Regarding techniques of solving class imbalance problems, such as under-sampling, over-sampling (synthetic minority oversampling technique, SMOTE), and cost adjust methods, future works can introduce those techniques to deal with imbalanced data. Furthermore, this study used an off-line training mode, which means we can have time to build high accuracy prediction models and determine the important variables based on them. Therefore, we focus on prediction accuracy without considering computational time and complexity. In future works, computational complexity and time could be considered to evaluate the constructed models.

Author Contributions: Conceptualization, T.-T.H.-C. and L.-S.C.; methodology, T.-T.H.-C.; software, T.-T.H.-C.; validation, T.-T.H.-C. and H.L.; formal analysis, T.-T.H.-C.; writing—original draft preparation, T.-T.H.-C.; writing—review and editing, L.-S.C.; visualization, H.L.; supervision, L.-S.C.; project administration, L.-S.C.; funding acquisition, L.-S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by Ministry of Science and Technology, Taiwan (Grant No. MOST 110-2410-H-324-003) and Chaoyang University of Technology (Granted No. 110F0021109).

Institutional Review Board Statement: Not applicable.

Acknowledgments: Authors are grateful for the financial assistance provided by the Ministry of Science and Technology, Taiwan and Chaoyang University of Technology.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Haskell, C.D. Institutional research as a bridge. *High. Educ. Eval. Dev.* **2017**, *11*, 2–11. [\[CrossRef\]](#)
2. Iyer, N.C.; Gireesha, H.M.; Shet, R.M.; Nissimgoudar, P.; Mane, V. Autonomous Driving Platform: An Initiative under Institutional Research Project. *Procedia Comput. Sci.* **2020**, *172*, 875–880. [\[CrossRef\]](#)

3. Cheng, T.M.; Hou, H.Y.; Agrawal, D.C.; Chen, L.S.; Chi, C.J. Factors affecting starting wages of master's degree-graduates in Taiwan. *J. Inst. Res. South East Asia* **2020**, *18*, 136–154.
4. Bai, S.; Hew, K.F.; Sailer, M.; Jia, C. From top to bottom: How positions on different types of leaderboard may affect fully online student learning performance, intrinsic motivation, and course engagement. *Comput. Educ.* **2021**, *173*, 104297. [[CrossRef](#)]
5. Rivas, A.; González-Briones, A.; Hernández, G.; Prieto, J.; Pablo Chamoso, P. Artificial neural network analysis of the academic performance of students in virtual learning environments. *Neurocomputing* **2021**, *423*, 713–720. [[CrossRef](#)]
6. Tarik, A.; Aissa, H.; Yousef, F. Artificial Intelligence and Machine Learning to Predict Student Performance during the COVID-19. *Procedia Comput. Sci.* **2021**, *184*, 835–840. [[CrossRef](#)]
7. Rahman, S.R.; Islam, A.; Akash, P.P.; Parvin, M.; Moon, N.N.; Nur, F.N. Effects of co-curricular activities on student's academic performance by machine learning. *Curr. Res. Behav. Sci.* **2021**, *2*, 100057. [[CrossRef](#)]
8. Kanakana, G.; Olanrewaju, A. Predicting student performance in engineering education using an artificial neural network at Tshwane University of Technology. In Proceedings of the International Conference on Industrial Engineering, Systems Engineering and Engineering Management for Sustainable Global Development, Stellenbosch, South Africa, 21–23 September 2011; Volume 2123, p. 17.
9. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Benetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2019**, *58*, 82–115. [[CrossRef](#)]
10. Heidari, A.A.; Faris, H.; Mirjalili, S.; Aljarah, I.; Mafarja, M. Ant lion optimizer: Theory, literature review, and application in multi-layer perceptron neural networks. In *Nature-Inspired Optimizers*; Springer: Cham, Switzerland, 2020; pp. 23–46.
11. van der Zanden, P.J.; Denessen, E.; Cillessen, A.H.; Meijer, P.C. Domains and predictors of first-year student success: A systematic review. *Educ. Res. Rev.* **2018**, *23*, 57–77. [[CrossRef](#)]
12. Ayala, J.C.; Manzano, G. Academic performance of first-year university students: The influence of resilience and engagement. *High. Educ. Res. Dev.* **2018**, *37*, 1321–1335. [[CrossRef](#)]
13. Baneres, D.; Rodriguez-Gonzalez, M.E.; Serra, M. An Early Feedback Prediction System for Learners At-Risk Within a First-Year Higher Education Course. *IEEE Trans. Learn. Technol.* **2019**, *12*, 249–263. [[CrossRef](#)]
14. Beaulac, C.; Rosenthal, J.S. Predicting University Students' Academic Success and Major Using Random Forests. *Res. High. Educ.* **2019**, *60*, 1048–1064. [[CrossRef](#)]
15. Neumann, H.; Padden, N.; McDonough, K. Beyond English language proficiency scores: Understanding the academic performance of international undergraduate students during the first year of study. *High. Educ. Res. Dev.* **2018**, *38*, 324–338. [[CrossRef](#)]
16. Ghosh, S.K.; Janan, F. Prediction of Student's Performance Using Random Forest Classifier. In Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management, Singapore, 7–11 March 2021.
17. Abu Naser, S.; Zaqout, I.; Abu Ghosh, M.; Atallah, R.; Alajrami, E. Predicting Student Performance Using Artificial Neural Network: In the Faculty of Engineering and Information Technology. *Int. J. Hybrid Inf. Technol.* **2015**, *8*, 221–228. [[CrossRef](#)]
18. Arsad, P.M.; Buniyamin, N. A neural network students' performance prediction model (NNSPPM). In Proceedings of the 2013 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), Kuala Lumpur, Malaysia, 25–27 November 2013; pp. 1–5.
19. Bakare, C.G. Some psychological correlates of academic success and failure. *AJER* **1975**, *2*, 11–22.
20. Oladokun, V.O.; Adebajo, A.T.; Charles-Owaba, O.E. Predicting students' academic performance using artificial neural network: A case study of an engineering course. *Pac. J. Sci. Technol.* **2008**, *9*, 72–79.
21. Dynan, K.E.; Rouse, C.E. The underrepresentation of women in economics: A study of undergraduate economics students. *J. Ecol. Educ.* **1997**, *28*, 350–368. [[CrossRef](#)]
22. Anderson, G.; Benjamin, D.; Fuss, M.A. The determinants of success in university introductory economics courses. *J. Ecol. Educ.* **1994**, *25*, 99–119. [[CrossRef](#)]
23. Adedeji, O.B. *A Study of the Relationship between Students Ume Results and Their Undergraduate Performance*; Department Of Industrial and Production Engineering, University Of Ibadan: Ibadan, Nigeria, 2001; Unpublished work.
24. Romero, C.; Ventura, S.; Espejo, P.; Hervás, C. Data mining algorithms to classify students. In Proceedings of the Educational Data Mining, Montréal, QC, Canada, 20–21 June 2008; pp. 20–21.
25. Chhachhiya, D.; Sharma, A.; Gupta, M. Designing optimal architecture of neural network with particle swarm optimization techniques specifically for educational dataset. In Proceedings of the 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence, Noida, India, 12–13 January 2017; pp. 52–57. [[CrossRef](#)]
26. Anderton, R.S. Identifying factors that contribute to academic success in first year allied health and science degrees at an Australian University. *Aust. J. Educ.* **2017**, *61*, 184–199. [[CrossRef](#)]
27. Chen, M.-Y.; Chang, J.-R.; Chen, L.-S.; Shen, E.-L. The key successful factors of video and mobile game crowdfunding projects using a lexicon-based feature selection approach. *J. Ambient. Intell. Humaniz. Comput.* **2021**, 1–19. [[CrossRef](#)]
28. Kabakchieva, D. Student performance prediction by using data mining classification algorithms. *J. Comput. Sci. Manag. Res.* **2012**, *1*, 686–690.
29. Zhu, W.; Zeng, X. Decision Tree-Based Adaptive Reconfigurable Cache Scheme. *Algorithms* **2021**, *14*, 176. [[CrossRef](#)]

30. Wang, C.; Bi, J.; Sai, Q.; Yuan, Z. Analysis and Prediction of Carsharing Demand Based on Data Mining Methods. *Algorithms* **2021**, *14*, 179. [\[CrossRef\]](#)
31. Wijenayake, S.; Graham, T.; Christen, P. A Decision Tree Approach to Predicting Recidivism in Domestic Violence. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*; Springer: Cham, Switzerland, 2018; pp. 3–15. [\[CrossRef\]](#)
32. Roy, A.G.; Urolagin, S. Credit risk assessment using decision tree and support vector machine based data analytics. Creative Business and Social Innovations for a Sustainable Future. In Proceedings of the 1st American University in the Emirates International Research Conference, Dubai, United Arab Emirates, 15–16 November 2017; pp. 79–84.
33. Wang, X.; Zhou, C.; Xu, X. Application of C4.5 decision tree for scholarship evaluations. *Procedia Comput. Sci.* **2019**, *151*, 179–184. [\[CrossRef\]](#)
34. Hamoud, A.K.; Hashim, A.S.; Awadh, W.A. Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis. *Int. J. Interact. Multimedia Artif. Intell.* **2018**, *5*, 26. [\[CrossRef\]](#)
35. Al-Hoqani, W.M.A.; Regula, T. A semi automated assessment and marking approach of decision tree diagram. *Mater. Today Proc.* **2021**, in press. [\[CrossRef\]](#)
36. Villavicencio, C.; Macrohon, J.; Inbaraj, X.; Jeng, J.-H.; Hsieh, J.-G. COVID-19 Prediction Applying Supervised Machine Learning Algorithms with Comparative Analysis Using WEKA. *Algorithms* **2021**, *14*, 201. [\[CrossRef\]](#)
37. Chang, J.-R.; Chen, M.-Y.; Chen, L.-S.; Chien, W.-T. Recognizing important factors of influencing trust in O2O models: An example of OpenTable. *Soft Comput.* **2019**, *24*, 7907–7923. [\[CrossRef\]](#)
38. Duke, C.R. Learning Outcomes: Comparing Student Perceptions of Skill Level and Importance. *J. Mark. Educ.* **2002**, *24*, 203–217. [\[CrossRef\]](#)
39. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
40. Abubakar, Y.; Ahmad, N.B.H. Prediction of students' performance in e-learning environment using random forest. *IJIC* **2017**, *7*. [\[CrossRef\]](#)
41. Chen, S.-H.; Pai, F.-Y.; Yeh, T.-M. Using the Importance–Satisfaction Model and Service Quality Performance Matrix to Improve Long-Term Care Service Quality in Taiwan. *Appl. Sci.* **2019**, *10*, 85. [\[CrossRef\]](#)
42. Chen, W.-K.; Chen, L.-S.; Pan, Y.-T. A text mining-based framework to discover the important factors in text reviews for predicting the views of live streaming. *Appl. Soft Comput.* **2021**, *111*, 107704. [\[CrossRef\]](#)
43. Tsai, S.-C.; Chen, C.-H.; Shiao, Y.-T.; Ciou, J.-S.; Wu, T.-N. Precision education with statistical learning and deep learning: A case study in Taiwan. *Int. J. Educ. Technol. High. Educ.* **2020**, *17*, 12. [\[CrossRef\]](#)
44. Nasser, I.M.; Al-Shawwa, M.O.; Abu-Naser, S.S. Developing Artificial Neural Network for Predicting Mobile Phone Price Range. *Int. J. Acad. Inf. Syst. Res.* **2019**, *3*, 1–6.
45. Chang, J.-R.; Chen, M.-Y.; Chen, L.-S.; Tseng, S.-C. Why Customers Don't Revisit in Tourism and Hospitality Industry? *IEEE Access* **2019**, *7*, 146588–146606. [\[CrossRef\]](#)
46. Bosch, E.; Seifried, E.; Spinath, B. What successful students do: Evidence-based learning activities matter for students' performance in higher education beyond prior knowledge, motivation, and prior achievement. *Learn. Individ. Differ.* **2021**, *91*, 102056. [\[CrossRef\]](#)
47. Osmanbegovic, E.; Suljic, M. Data mining approach for predicting student performance. *Econo. Rev. J. Econo. Busin* **2012**, *10*, 3–12.
48. Kim, B.H.; Vizitei, E.; Ganapathi, V. GritNet: Student performance prediction with deep learning. *arXiv* **2018**, arXiv:1804.07405.
49. Dharmasaroja, P.; Kingkaew, N. Application of artificial neural networks for prediction of learning performances. In Proceedings of the 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Changsha, China, 13–15 August 2016; pp. 745–751. [\[CrossRef\]](#)
50. Sivasakthi, M. Classification and prediction based data mining algorithms to predict students' introductory programming performance. In Proceedings of the 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 23–24 November 2017; pp. 346–350. [\[CrossRef\]](#)