Predictive Modelling of Student Performance in MMU Based on Machine Learning Approach

Jun Yang Chan[1], Hu Ng[1](✉), Timothy Tzen Vun Yap[1], and Vik Tor Goh[2]

[1] Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia
nghu@mmu.edu.my
[2] Faculty of Engineering, Multimedia University, Cyberjaya, Malaysia

Abstract. This research work is to identify and forecast student performance in Multimedia University (MMU) based on machine learning approach. Surveys had been carried out to collect students' cumulative grade point average, background, history grades, opinion towards MMU environment, and lifestyle. In this study, two datasets, namely Dataset 1 (1288 records) and Dataset 2 (945 records) have been collected from two trimesters. Following that, data pre-processing and data transformation are carried out. Next, feature selection is performed to eliminate the irrelevant features. Finally, the classification models of Support Vector Machine (SVM), Naive Bayes (NB), Random Forest (RF), K-Nearest Neighbor (KNN), Multilayer Perceptron (MLP), and Convolutional Neural Network (CNN) origins are designed, and their performances evaluated. The classification was performed with 80–20, 70–30, 60–40 and 50–50 train-test splits and 10-fold cross validation on Merge Dataset. GridSearchCV was applied to perform hyperparameter tuning. The performance metrics are accuracy, precision, recall and F1-Score. RF obtains the best accuracy results among the six models in Merged dataset classification, RF scores 0.9093 for 5-Level classification with oversampling and 1.0 for binary classification with oversampling.

Keywords: Student performance · Boruta feature selection · Machine learning · Data mining

1 Introduction

The term 'student' refers to a person who gets educated in a school or any other learning. Every one of us has been a student at some point in our lives. Learning is our principle of life to grow into a well-educated person. Every student has the duty to focus to their studies and learning more fresh knowledge to attain their ultimate goals. Hence, students will not only gain self-esteem but will also be able to contribute to the betterment of our community.

Student performance is a key feature in education [1]. Student performance can be fulfilled with exam results, accomplishment in academic or co-curriculum and interaction between other students, examples like communication and activity participation in the classroom. Teachers attempt to implement strategies to improve student performance

S.-C. Haw and K. Sonai Muthu (Eds.): CITIC 2022, 10, pp. 258–278, 2022.
https://doi.org/10.2991/978-94-6463-094-7_21

who are struggling during class activity [2]. Hence, understanding the elements that influence student performance is an essential consideration.

Factors such as students background, history of grades, school or university environment, lifestyle are playing a significant role to improve student performance. This gives ways to machine learning (ML) to predict student performance by detecting which distractions may influence student performance.

The objectives of this paper are to find out factors that affect student performance in MMU by developing various predictive models to classify student performance.

2 Literature Review

2.1 Factors Affect Student Performance, EDA and Machine Learning Techniques

Atlantic International University [3] discovered children from poor socioeconomic backgrounds respond incomprehensibly to classroom instruction and poor student performance. Hartawan et al. [4] claimed student achievement is also one of the factors that influence student performance to work hard to achieves their goals. Oselumese et al. [5] found out that a good learning environment is able to help on improving student performance. Hartawan et al. [4] located that motivation of learning is the key to student performance success.

Based on Bhatia and Jaggi [6] Exploratory Data Analysis (EDA) is used to analyse the sample and the measures of the data. EDA help to define errors and maximise the understanding of dataset, identify critical variables, outliers and inconsistencies data and also identify the relationship between variables.

SVM is a classification method for both linear and nonlinear data. It can separate different classes of data by finding out their decision boundaries or hyperplanes using support vectors and margins [7].

NB is a classifier to calculate probabilities and conditional probabilities based on Bayes Theorem. It is a statistical classifier that predicts which class label belongs to, based on the currently existing objects [8].

RF is a machine learning method that employs many decision trees during the training stage to produce an average output [9].

KNN is an algorithm for categorising unknown or new data points to the nearest neighbor based on Euclidean space with positive integer dimension [10].

ANN is an adaptive system that modifies structure based on external or internal data that goes through the network of the learning process. Basically, ANN creates a set of outputs from a set of different inputs. ANN contains three layers; the first is the Input layer, Hidden layer and Output layer [11].

CNN is similar to ANN. CNN produces highly accurate recognition results, CNN is also able to be retrained to do additional recognition tasks and expand on pre-existing networks. CNN contains more hidden layers in neural system architecture to improve prediction results [12].

2.2 Related Work

Al-Shehri et al. [13] predicted student performance using 395 data with 33 attributes collected from the University of Minho. After 10-fold cross-validation, they obtained correlation coefficient score of 0.96 and 0.95t by applying SVM and KNN respectively.

Yaacob et al. [8] applied the Decision Tree (DT), NB, KNN and Logistic Regression (LR) to classify the student performance from 631 data of Universiti Teknologi MARA Kelantan and Universiti Teknologi MARA Negeri Sembilan. The NB model recorded 89.26% as the top accuracy score among all the ML models.

Amra and Maghari [10] forecasted student performance using the KNN and NB. 500 students' data with 8 attributes selected from the dataset collected by ministry of education in Gaza. NB received the highest accuracy between both classifier with 93.17% of accuracy score.

Lau et al. [14] applied ANN model with 11 input features to predict student performance of 1000 undergraduates' data that collected from a university in China. They obtained 84.8% of accuracy score.

Akour et al. [15] applied the CNN with 480 student data that collected from experience API (xAPI) to predict the student performance. After the 200 epochs, CNN achieved 100% accuracy score with two layers of model.

3 Methodology

In this research work, the approach consists of five stages, namely data acquisition, data pre-processing, data transformation, data oversampling, feature selection and classification.

3.1 Data Acquisition

Surveys were performed via google form to obtain attributes that influencing MMU students from 7 faculties, across Cyberjaya and Melaka campuses. Multiple options of questions have been set to gather information for their backgrounds, grades, opinions towards MMU environment, and lifestyle via Google Form. Questions are designed based on the works from Ng et al. [16] and Cortez and Silva [17]. Data collected in trimester 1, September 2021 with 1288 responses as Dataset 1 and trimester 2, December 2022 (945 data) as Dataset 2. Only 1260 data from Dataset 1 and 888 data from Dataset 2 with containing CGPA detail from participants were used in this research.

3.1.1 Exploratory Data Analysis Dataset 1

Figure 1 shows surveys were participated by 773 female students and 487 male students in Dataset 1. Figure 2 shows 628 of students who get high grad4e of CGPA didn't fail any subject.

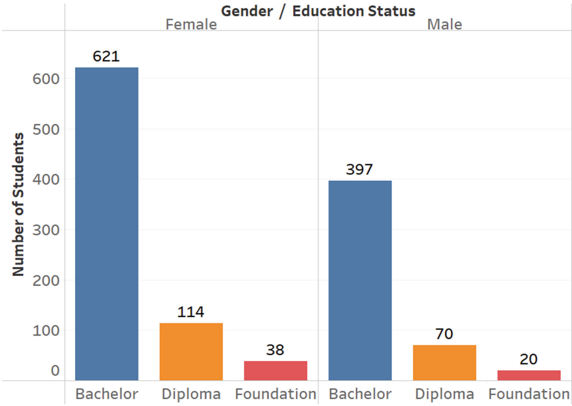

Fig. 1. Number of Student Gender and Education Status

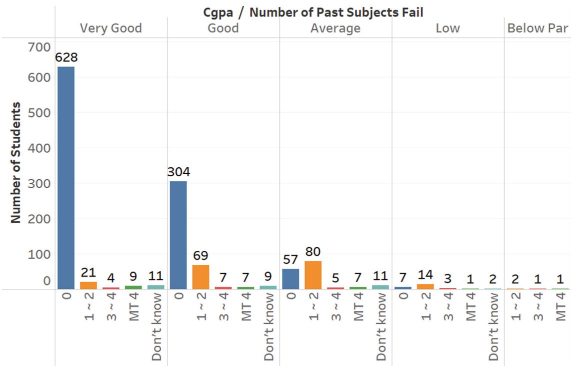

Fig. 2. Distribution of Number on Subjects with Passed or Failed

Figure 3 exhibits duration spending of student in studying and playing game. It can be observed that students spend more time in studying and lesser time in game played an important role in getting a good CGPA. The CGPA ranking are defined into 5 classes; below par (less than <2.0), poor ($\leqslant 2.5$ and >2.0), average ($\leqslant 3.0$ and >2.5), good ($\leqslant 3.5$ and >3.0) and very good ($\leqslant 4.0$ and >3.5).

3.1.2 Exploratory Data Analysis Dataset 2

Figure 4 shows distribution of reasons in enrolling into MMU and daily travel time spend. It can be observed that the university's course preferences and shorter travel time are the major attraction for enrolling to study in MMU. Also, it can be found a high distribution of zero travel time, this is mainly due to many students are staying at home for virtual learning during the pandemic Covid-19.

Cgpa	Game Time	Study Time < 2 hours	2 < 5 hours	5 < 8 hours	>= 8 hours
Below Par	< 1 hours				1
	1 < 3 hours	1			
	>= 5 hours	2			
Low	< 1 hours	2	7	1	1
	1 < 3 hours	7	2	2	
	3 < 5 hours	2			
	>= 5 hours	3			
Average	< 1 hours	19	39	6	1
	1 < 3 hours	21	22	6	1
	3 < 5 hours	5	10		
	>= 5 hours	24	4	2	
Good	< 1 hours	48	83	25	5
	1 < 3 hours	50	70	12	6
	3 < 5 hours	27	30	6	3
	>= 5 hours	15	9	6	1
Very Good	< 1 hours	76	121	59	13
	1 < 3 hours	88	105	51	7
	3 < 5 hours	45	47	10	1
	>= 5 hours	33	14	3	

Fig. 3. Distribution of Study and Game Time

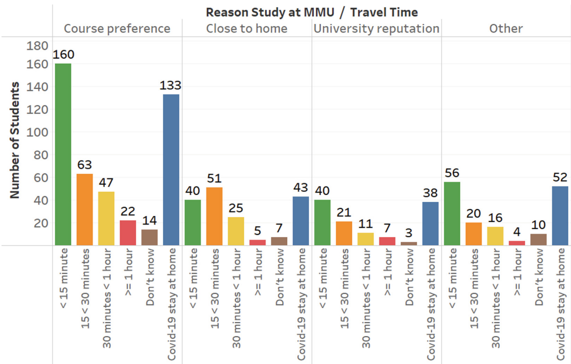

Fig. 4. Distribution of Reasons in Enrolling into MMU and Daily Travel Time Spend

3.1.3 Student Attributes

A total of 2148 responses with CGPA were used in classification tasks. There are 38 attributes for each responded data from two surveys. The attributes are split into five categories that comprise of student background (SB), student lifestyle (SL), student history of grades (SHG), student opinion towards MMU environment and all the 38 attributes (ALL). The Tables 1, 2, 3 and 4 provide all the attributes with the related definitions.

Table 1. Student Background (SB) Attributes

Attribute	Definition
Gender	Student's gender
Age	Student's age
Faculty	Student study in which faculty?
Living area	Student living area type
Family size	Number family members (including student)
Living status	Students living where before Covid-19?
Parent's education [Father]	Father education level
Parent's education [Mother]	Mother education level
Father's job	Father occupation
Mother's job	Mother occupation
Financial problem	Do student family have financial problem?
Reason study at MMU	Why do student study at MMU

Table 2. Student History Grades (SHG) Attributes

Attribute	Definition
Education status	Students study in which programme?
CGPA	Student current CGPA
SPM result (A)	How many A did student score in SPM
SPM result (B)	How many B did students score in SPM
SPM result (C)	How many C did students score in SPM

Table 3. Student Opinion Towards MMU Environment (SOTME) Attributes

Attribute	Definition
Size of class	Most suitable number of students in one class?
Facilities	How do students rate the facilities at MMU?
Class environment	How comfortable studying in MMU's classroom?
Overall environment	How the environment outside MMU's classroom?

3.2 Data Pre-processing

A value of zero is placed to those attributes with null content to avoid any error during the later classification process.

Table 4. Student Lifestyle (SL) Attributes

Attribute	Definition
Travel time	How long do students need to travel to MMU?
Study time	How long do students study daily?
Study type	Students like study alone or in group
Game time	How long do students play video games daily
Number of past subject fail	How many subjects did students fail before in MMU? (Students who join MMU during the pandemic just count subject failed during SPM)
Scholarship	A payment that supports a student's education
PTPTN	Loans for students
Family support	Student family members pay for student's expenses?
Extra paid classes	Extra paid classes within the course subject.
Extra curricular activities	Extra activities that are not related to the course subject. Example: club activity, sports training, etc.
Dating	Student in romantic relationship?
Part time job	Do student work part time?
Free time after class	How student rate their free time after per day class
Going out with friends	How often do students go out with friends in a trimester before the Covid-19 pandemic?
Social interaction	Students like making friends or staying alone
Alcohol consumption	How often do students drink alcohol
Skipping class	How often do students skip class

3.3 Data Transformation

Data transformation is carried out to ensure all the value in attributes are in numeric format. For example, for Gender attribute, the Male option will be set to value of 0 and the Female option will be value of 1.

3.4 Data Oversampling

Due to the imbalanced of the CGPA classes, for example: there is only 4 records for under par and poor, but 1260 for average, good and very good. RandomOverSampler is applied to over-sample the minority classes by duplicating original data. RandomOverSampler is chosen as it is proven effective in many researchers' findings [18, 19]. The experiments are carried out on dataset with oversampling (WO) process and without oversampling (WTO). Table 5 shows the capacity of records with and without oversampling process.

Table 5. Before and After Oversampling

Dataset	Before oversampling (row, column)	After oversampling (row, column)
Dataset 1 (D1)	(1260, 38)	(3365, 38)
Dataset 2 (D2)	(888, 38)	(2240, 38)
Merge (D1 + D2)	(2148, 38)	(4276, 38)

3.5 Feature Selection

Based on finding in Tang and Zhang [18], Boruta is a wrapper algorithm that is constructed around the ML model. Boruta distinguishes the essential attributes that positively contribute to accuracy rate during the classification process. All the 38 attributes have been identified important and being selected by the Boruta during the feature selection process.

3.6 Classification

The classification is performed by SVM, NB and MLP. Two datasets (Dataset 1 and Dataset 2) are involved in the classification process. Experiments are carried out on two model: Merged dataset (merging Dataset 1 and Dataset 2) and D1-D2, where Dataset 1 is used as training set and Dataset 2 is used the evaluation test set. Only the merged dataset (Merged) is evaluated under the train-test splits ration for 50–50, 60–40, 70–30, and 80–20.

GridSearchCV is applied to perform hyperparameter tuning. The performance metrics were accuracy, precision, recall and F1-Score. The classification is carried out by applying Scikit-learn package that written in Python. Due to the limitation of space, only accuracy scores are shown in the later 4.0 Result and Analysis section.

4 Result and Analysis

4.1 Classification Results

Various experiments have been carried out to identified attributes that influencing student performance, those attributes are split into five categories that comprise of student background (SB), student lifestyle (SL), student history of grades (SHG), student opinion towards MMU environment and all the 38 attributes (ALL).

SVM, NB, RF, KNN, ANN, and CNN models are applied to perform 5-Level classification (5LC) and binary classification (BC). BC is indicated fail (below par and poor) or pass status (average, good and very good) meanwhile for 5-level classification is for student scores in F (below par), D (poor), C (average), B (good) and A (very good).

4.1.1 Support Vector Machine

Table 6 shows the best parameters obtained for SVM. Tables 7 and 8 show the classification results. TT represents the train-test split ratio. While categories indicate the five categories of attributes grouping.

Table 6. Best Parameters of SVM

Merge dataset			
Classification	Parameter		
	C	gamma	kernel
5 level	1	1	rbf
Binary	0.1	1	rbf
Train-test split dataset			
5 level	1	1	rbf
Binary	0.1	1	rbf

Table 7. 5LC & BC using SVM with Merge Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
80–20	0.5248	0.5279	0.5357	**0.5876**	0.5651
70–30	0.5452	0.5333	0.5426	0.5871	0.5643
60–40	0.5407	0.5313	0.5291	0.5744	0.5640
50–50	0.5400	0.5155	0.5394	0.5813	0.5717
5LC train-test WO					
80–20	0.8958	0.6713	0.4344	0.9173	**0.9231**
70–30	0.8802	0.6615	0.4162	0.9020	0.9049
60–40	0.8698	0.6558	0.4117	0.8843	0.8837
50–50	0.8409	0.6468	0.4175	0.8848	0.8673
BC train-test WTO					
80–20	0.9953	0.9953	0.9953	0.9953	**1.0**
70–30	0.9948	0.9948	0.9948	0.9948	**1.0**
60–40	0.9957	0.9957	0.9957	0.9957	**1.0**
50–50	0.9957	0.9957	0.9957	0.9957	**1.0**
BC train-test WO					
80–20	0.9996	0.9824	0.8057	0.9173	0.9953
70–30	**0.9997**	0.9831	0.8028	0.9020	0.9953
60–40	0.9996	0.9836	0.7941	0.8843	0.9957
50–50	**0.9997**	0.9844	0.7965	0.8848	0.9956

Table 8. 5LC & BC using SVM with Train-Test Split Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
D1–D2	0.5496	0.5507	0.5496	**0.5912**	0.5664
5LC train-test WO					
D1–D2	0.2184	**0.2680**	0.1861	0.2071	0.2
BC train-test WTO					
D1–D2	**0.9932**	**0.9932**	**0.9932**	**0.9932**	**0.9932**
BC train-test WO					
D1–D2	**0.5**	0.4938	0.4501	**0.5**	**0.5**

4.1.2 Naïve Bayes

Naive Bayes classifier does not allow for hyperparameter tuning. GaussianNB has its own ML algorithm. Tables 9 and 10 shows the classification results. Due to no hyperparameter tuning, most of accuracy score are considered low.

4.1.3 Random Forest

Table 11 shows the best parameter based on each classification. All classification results are show in Tables 12 and 13.

4.1.4 K-Nearest Neighbor

Table 14 shows the best parameter based on each classification. All classification results are show in Tables 15 and 16.

4.1.5 Multilayer Perceptron

Table 17 shows the best parameter based on each classification. All classification results are show in Tables 18 and 19.

Table 9. 5LC & BC using NB with Merge Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
80–20	0.4078	0.5038	0.5209	0.2031	0.2833
70–30	0.3261	0.4883	**0.5307**	0.2879	0.3535
60–40	0.3225	0.4895	0.5190	0.2860	0.3426
50–50	0.3051	0.4984	0.5248	0.3612	0.4031
5LC train-test WO					
80–20	**0.3595**	0.3348	0.2759	0.2991	0.3348
70–30	0.3538	0.3366	0.2594	0.2886	0.3264
60–40	0.3517	0.3356	0.2677	0.2880	0.3203
50–50	0.3560	0.3454	0.2786	0.2834	0.3186
BC train-test WTO					
80–20	0.3186	0.9953	0.9953	0.3713	0.5000
70–30	0.6641	0.9948	0.9948	0.5111	0.6279
60–40	0.6733	**0.9957**	**0.9957**	0.5093	0.6260
50–50	0.6350	**0.9957**	0.9956	0.6291	0.7157
BC train-test WO					
80–20	0.7274	0.7060	0.6561	0.6090	0.6511
70–30	**0.7295**	0.7118	0.6471	0.6137	0.6537
60–40	0.7267	0.7093	0.6452	0.6123	0.6523
50–50	0.7266	0.7128	0.6448	0.6135	0.6561

Table 10. 5LC & BC using NB with Train-Test Split Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
D1–D2	0.1306	**0.5169**	0.5135	0.1644	0.2770
5LC train-test WO					
D1–D2	0.1541	**0.3275**	0.2266	0.1623	0.2209
BC train-test WTO					
D1–D2	0.3243	**0.9932**	**0.9932**	0.8401	0.8682
BC train-test WO					
D1–D2	0.4048	**0.7273**	0.4512	0.5074	0.5068

Table 11. Best Parameters of RF

Merge dataset			
Classification	Parameter		
	n_estimators	max_features	Criterion
5 level	400	auto, sqrt	gini
Binary	100	auto	gini
Train-test split dataset			
5 level	300	auto	entropy, gini
Binary	100	auto	gini

Table 12. 5LC & BC using RF with Merge Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
80–20	0.5016	0.5217	0.4969	0.5682	0.5930
70–30	0.5075	0.5090	0.5023	0.5649	**0.5943**
60–40	0.5047	0.4845	0.5081	0.5523	0.5713
50–50	0.5205	0.4929	0.5130	0.5698	0.5891
5LC train-test WO					
80–20	0.8820	0.6615	0.4200	0.8909	**0.9093**
70–30	0.8567	0.6621	0.4198	0.8819	0.8984
60–40	0.8415	0.6550	0.4208	0.8658	0.8780
50–50	0.8190	0.6519	0.4205	0.8491	0.8677
BC train-test WTO					
80–20	0.9953	0.9946	0.9953	0.9953	0.9953
70–30	0.9948	0.9928	0.9948	0.9948	0.9948
60–40	0.9953	0.9942	**0.9957**	**0.9957**	**0.9957**
50–50	0.9953	0.9941	**0.9957**	**0.9957**	**0.9957**
BC train-test WO					
80–20	0.9992	0.9809	0.8030	**1.0**	**1.0**
70–30	0.9992	0.9816	0.7976	**1.0**	**1.0**
60–40	0.9990	0.9821	0.7962	**1.0**	**1.0**
50–50	0.9991	0.9827	0.7951	**1.0**	**1.0**

Table 13. 5LC & BC using RF with Train-Test Split Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
D1–D2	0.5214	0.5364	0.4944	0.5683	**0.6051**
5LC train-test WO					
D1–D2	0.2477	0.2351	0.1769	0.2967	**0.3254**
BC train-test WTO					
D1–D2	**0.9932**	**0.9932**	**0.9932**	**0.9932**	**0.9932**
BC train-test WO					
D1–D2	**0.5**	0.4932	0.4501	**0.5**	**0.5**

Table 14. Best Parameters of KNN

Merge dataset			
Classification	Parameter		
	n_neighbors	Weights	Metric
5 level	10	Distance	Manhattan
Binary	5	Uniform	Euclidean
Train-test split dataset			
5 level	8, 10	Distance	Manhattan
Binary	5	Uniform	Euclidean

4.1.6 Convolutional Neural Network

All classification results are show in Tables 20 and 21.

4.2 Analysis

RF obtains the best accuracy results among the six models in merged dataset with oversampling classification, RF scores 0.9093 for 5-Levl classification and 1.0 for binary classification. The analysis on the selected features are presented from Figs. 5, 6 and 7.

Based on Fig. 5, students with Very Good CGPA are highly related with scoring nine or more A's in SPM in the past. This concludes that those students who obtained more A's in SPM have a better probability of scoring Very Good CGPA during the study in university.

Based on Fig. 6, majority of students with Very Good CGPA do not have a scholarship and financial problems. It is also can be observed that getting Good grades in CGPA does not related with financial status, but rather on student own hard work and effort. Where

Table 15. 5LC & BC using KNN with Merge Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
80–20	0.4985	0.4985	0.4798	0.5411	0.5442
70–30	0.4946	0.4873	0.4724	0.5323	**0.5638**
60–40	0.4934	0.4810	0.5147	0.5198	0.5519
50–50	0.5065	0.4894	0.4851	0.5214	0.5571
5LC train-test WO					
80–20	0.8252	0.6604	0.3856	0.8352	**0.8619**
70–30	0.8071	0.6554	0.4064	0.8232	0.8444
60–40	0.7940	0.6493	0.3929	0.8088	0.8338
50–50	0.7747	0.6397	0.3917	0.7953	0.8175
BC train-test WTO					
80–20	0.9953	0.9953	0.9953	0.9946	0.9953
70–30	0.9948	0.9948	0.9948	0.9948	0.9948
60–40	**0.9957**	**0.9957**	**0.9957**	0.9953	**0.9957**
50–50	**0.9957**	**0.9957**	**0.9957**	**0.9957**	**0.9957**
BC train-test WO					
80–20	0.9922	0.9770	0.7905	0.9926	**0.9961**
70–30	0.9925	0.9790	0.7662	0.9909	0.9948
60–40	0.9901	0.9772	0.7894	0.9905	0.9942
50–50	0.9905	0.9740	0.7710	0.9880	0.9924

there are found 86 students able to achieve high CGPAs despite financial difficulties and a lack of scholarship funding.

Figure 7 shows three types of habits that highly impact student performance, namely students' study time, how frequently students skip class, and students' game time. Based on the statistics in Fig. 7, the ideal habits for the students to obtained a get Very Good in CGPA are play gaming for less than one hour per day, never skip class, and study for $2 < 5$ h each day.

5 Conclusion and Future Work

This paper manages to find out the factors that affect MMU student performance and develop some predictive models to classify student performance. The study looks at data only from one university, Multimedia University, Malaysia and may not reflect a general view of the case. Future work will include more datasets from different countries. Also, other classifiers will be explored and investigated.

Table 16. 5LC & BC using KNN with Train-Test Split Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
D1–D2	0.4910	0.4944	0.4583	0.5383	**0.5574**
5LC train-test WO					
D1–D2	0.2828	0.2574	0.1873	0.2689	**0.3439**
BC train-test WTO					
D1–D2	**0.9932**	**0.9932**	**0.9932**	**0.9932**	**0.9932**
BC train-test WO					
D1–D2	0.4960	0.4972	0.5	**0.5845**	0.4994

Table 17. Best Parameters of MLP

Merge dataset					
Classification	Parameter				
	alpha	hidden_layer_size	Activation	Solver	learning_rate
5 level	1e-06	(16, 8, 16)	tanh	adam	constant
Binary	0.1	(16, 8, 16)	logistic	adam	constant
Train-test split dataset					
5 level	0.0001	(16, 8, 16)	tanh	adam	adaptive
Binary	0.1	(16, 8, 16)	logistic, relu	sgd, adam	constant

Authors' Contributions

Hu Ng, roles: Corresponding author, Conceptualization, Project Administration, Supervision, Validation, Visualization, Writing—Review & Editing.

Jun Yang Chan, roles: Data Curation, Formal Analysis, Investigation, Resources, Writing—Original Draft Preparation.

Timothy Tzen Vun Yap, roles: Conceptualization, Investigation, Project Administration, Supervision, Validation, Writing—Review & Editing.

Vik Tor Goh, roles: Conceptualization, Validation, Visualization, Writing—Review & Editing.

Table 18. 5LC & BC using MLP with Merge Dataset

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
80–20	0.5302	0.5357	0.5349	0.5643	0.5612
70–30	0.5442	0.5189	0.5416	**0.5773**	0.5674
60–40	0.5357	0.5178	0.5360	0.5667	0.5543
50–50	0.5279	0.5236	0.5397	0.5630	0.5587
5LC train-test WO					
80–20	0.6667	0.5823	0.3833	0.7855	**0.8277**
70–30	0.6841	0.5656	0.3781	0.7849	0.8138
60–40	0.6454	0.5474	0.3906	0.7802	0.8145
50–50	0.6677	0.5598	0.3865	0.7705	0.8068
BC train-test WTO					
80–20	0.9953	0.9953	0.9953	0.9953	0.9953
70–30	0.9948	0.9948	0.9948	0.9948	0.9948
60–40	**0.9957**	**0.9957**	**0.9957**	**0.9957**	0.9950
50–50	**0.9957**	**0.9957**	**0.9957**	**0.9957**	**0.9957**
BC train-test WO					
80–20	0.9973	0.9809	0.8026	**0.9992**	0.9957
70–30	0.9958	0.9792	0.7612	0.9971	0.9964
60–40	0.9967	0.9776	0.7855	0.9971	0.9955
50–50	0.9956	0.9758	0.7930	0.9961	0.9956

Table 19. 5LC & BC using MLP with Train-Test Split Dataset

5LC train-test WTO					
	Categories				
TT	SB	SHG	SOTME	SL	All
D1–D2	0.5480	0.5387	0.5496	**0.5755**	**0.5755**
5LC train-test WO					
D1–D2	0.2455	0.2398	0.1924	**0.2799**	0.2657
BC train-test WTO					
D1–D2	**0.9932**	**0.9932**	**0.9932**	**0.9932**	**0.9932**
BC train-test WO					
D1–D2	**0.5253**	0.4921	0.4501	0.4991	0.4987

Table 20. 5LC & BC using CNN with Merge Dataset.

5LC train-test WTO					
	Categories				
	SB	SHG	SOTME	SL	All
80–20	0.4543	0.5023	0.5194	0.5155	0.5178
70–30	0.4605	0.5147	0.5168	0.5106	0.5178
60–40	0.4690	0.5093	0.5217	0.4864	0.5151
50–50	0.4544	0.5075	**0.5282**	0.4916	0.5314
5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
80–20	0.7936	0.5759	0.4215	0.8556	**0.8840**
70–30	0.7813	0.5624	0.4089	0.8523	0.8724
60–40	0.7763	0.5667	0.4018	0.8349	0.8587
50–50	0.7539	0.5588	0.4045	0.8198	0.8451
BC train-test WTO					
80–20	0.9945	0.9953	0.9953	0.9891	0.9938
70–30	0.9912	0.9927	0.9948	0.9901	0.9928
60–40	0.9899	0.9946	**0.9957**	0.9903	0.9949
50–50	0.9910	0.9922	**0.9957**	0.9935	0.9938
BC train-test WO					
80–20	0.9911	0.9634	0.7843	0.9977	**0.9996**
70–30	0.9907	0.9540	0.7937	0.9974	0.9995
60–40	0.9879	0.9468	0.7801	0.9944	0.9984
50–50	0.9869	0.9487	0.7814	0.9961	0.9991

Table 21 5LC & BC using CNN with Train-Test Split Dataset.

5LC train-test WTO					
TT	Categories				
	SB	SHG	SOTME	SL	All
D1–D2	0.4591	0.5206	0.5330	0.5060	**0.5574**
5LC train-test WO					
D1–D2	0.2284	0.2398	0.1683	0.2712	**0.3014**
BC train-test WTO					
D1–D2	0.9921	0.9921	**0.9932**	0.9887	0.9910
BC train-test WO					
D1–D2	**0.5508**	0.4847	0.4478	0.4996	0.5

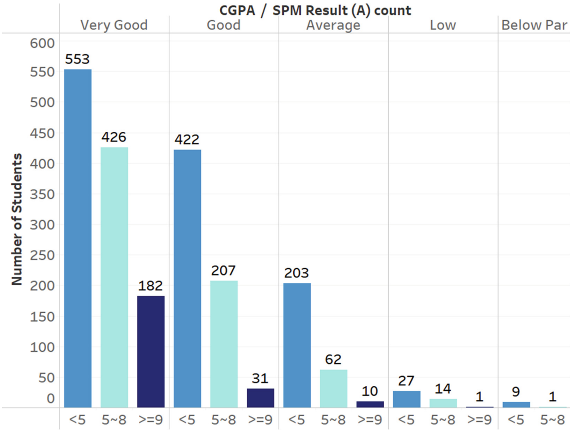

Fig. 5. Distribution of Student SPM Result (A) for CGPA

CGPA	Scholarship	Financial Problem		
		Don't know	No	Yes
Very Good	No	107	488	86
	Yes	75	312	93
Good	No	95	351	90
	Yes	16	88	20
Average	No	46	150	47
	Yes	10	16	6
Low	No	9	24	6
	Yes		1	2
Below Par	No	3	1	4
	Yes		1	1

Fig. 6. Distribution of Student Scholarship and Financial Problem for CGPA

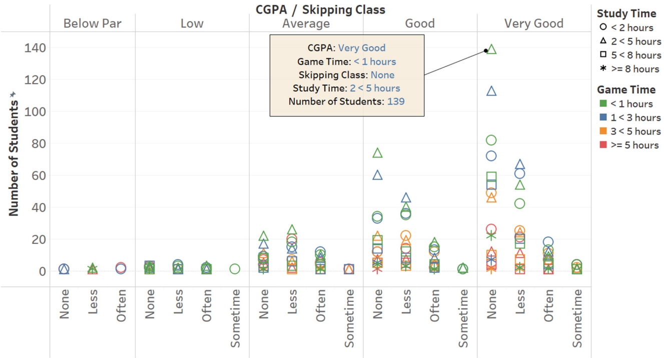

Fig. 7. Distribution of Student Study Time, Skipping Class, Game Time for CGPA

Acknowledgments. The authors would like to thank you the Management of Multimedia University (MMU) for granting permission to conduct survey on MMU's students. Heartiest appreciation to those participants in providing the precious data for this research work.

References

1. Abaidoo, A. (2018). Factors contributing to academic performance of students in a junior high school. GRIN. DOI: https://www.grin.com/document/450284
2. Thiede, K. W., Brendefur, J. L., Osguthorpe, R. D., Carney, M. B., Bremner, A., Strother, S., Oswalt. S., Snow, J.L., Sutton, J., & Jesse, D. (2015). Can teachers accurately predict student performance?. Teaching and Teacher Education, 49, 36–44. DOI: https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=1146&context=cifs_facpubs
3. Atlantic International University. (2017). GRIN - The Family Effect on Academic Performance in School. A Case Study of selected Schools in Kabale District. GRIN. DOI: https://www.grin.com/document/384998v
4. Hartawan, F. O., Lawrence, C., Chrisanto, A. Y., Cornelius, L., Kurniawan, D., Surya, W., ... & Soetanto, C. B. (2021). Relationship between learning motivation and student achievement of Bina Nusantara University. Journal of Economics and Business Letters, 1(1), 1–4. DOI: https://journal.privietlab.org/index.php/JEBL/article/download/10/1
5. Oselumese, I. B., Omoike, D., & Andrew, O. (2016). Environmental influence on students' academic performance in secondary school. International journal of fundamental Psychology and Social Sciences, 6(1), 10–14. DOI: https://web.archive.org/web/20170516184220id_/http://fundamentaljournals.org:80/ijfpss/downloads/58_IJFPSS_March_2016_10_14.pdf
6. Bhatia, V. K., & Jaggi, S. (2018). Exploratory Data Analysis. DOI: http://bioinformatics.iasri.res.in/ePublication/book/module1/VKBhatia-Exploratory_data_analysis.pdf
7. Javatpoint. (n.d.). Support Vector Machine (SVM) Algorithm - Javatpoint. www.javatpoint.com. Retrieved September 12, 2021, from https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm

8. Yaacob, W. F. W., Nasir, S. A. M., Yaacob, W. F. W., & Sobri, N. M. (2019). Supervised data mining approach for predicting student performance. Indonesian Journal of Electrical Engineering and Computer Science, 16(3), 1584–1592. DOI: https://pdfs.semanticscholar.org/aed4/a5c483168dc3e494e67f28ed1c176b4d78f9.pdf?_ga=2.255710216.1612971942.1635568391-397263090.1635568391
9. Edla, D. R., Mangalorekar, K., Dhavalikar, G., & Dodia, S. (2018). Classification of EEG data for human mental state analysis using Random Forest Classifier. Procedia computer science, 132, 1523–1532. DOI: https://sci-hub.hkvisa.net/, https://doi.org/10.1016/j.procs.2018.05.116
10. Amra, I. A. A., & Maghari, A. Y. (2017, May). Students performance prediction using KNN and Naïve Bayesian. In 2017 8th International Conference on Information Technology (ICIT) (pp. 909–913). IEEE. DOI: https://sci-hub.mksa.top/, https://doi.org/10.1109/icitech.2017.8079967
11. Abu-Naser, S. S., Zaqout, I. S., Abu Ghosh, M., Atallah, R. R., & Alajrami, E. (2015). Predicting student performance using artificial neural network: In the faculty of engineering and information technology. DOI: http://dstore.alazhar.edu.ps/xmlui/bitstream/handle/123456789/391/27-05-2019-12.pdf?sequence=1&isAllowed=y
12. MathWorks. (n.d.). What is a Convolutional Neural Network? MATLAB & Simulink. Retrieved September 12, 2021, from https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html
13. Al-Shehri, H., Al-Qarni, A., Al-Saati, L., Batoaq, A., Badukhen, H., Alrashed, S., ... & Olatunji, S. O. (2017, April). Student performance prediction using support vector machine and k-nearest neighbor. In 2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE) (pp. 1–4). IEEE. DOI: https://sci-hub.mksa.top/, https://doi.org/10.1109/CCECE.2017.7946847
14. Lau, E. T., Sun, L., & Yang, Q. (2019). Modelling, prediction and classification of student academic performance using artificial neural networks. SN Applied Sciences, 1(9), 1–10. DOI: https://link.springer.com/content/pdf/, https://doi.org/10.1007/s42452-019-0884-7.pdf
15. Akour, M., Alsghaier, H., & Al Qasem, O. (2020). The effectiveness of using deep learning algorithms in predicting students achievements. Indonesian Journal of Electrical Engineering and Computer Science, 19(1), 387–393. DOI: https://bit.ly/3uLrUo2
16. Ng, H., bin Mohd Azha, A. A., Yap, T. T. V., & Goh, V. T. (2022). A Machine Learning Approach to Predictive Modelling of Student Performance. F1000Research, 10, 1144.
17. Cortez P, Silva A: Using data mining to predict secondary school student performance. 15th European Concurrent Engineering Conference 2008, ECEC 2008–5th Future Business Technology Conference, FUBUTEC 2008. 2008;2003(2000):5–12
18. Tang, R., & Zhang, X. (2020, May). CART Decision Tree Combined with Boruta Feature Selection for Medical Data Classification. In 2020 5th IEEE International Conference on Big Data Analytics (ICBDA) (pp. 80–84). IEEE. DOI:https://sci-hub.mksa.top/, https://doi.org/10.1109/ICBDA49040.2020.9101199
19. Ghorbani, R., & Ghousi, R. (2020). Comparing different resampling methods in predicting students' performance using machine learning techniques. IEEE Access, 8, 67899–67911. DOI: https://sci-hub.hkvisa.net/, https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9062549