

Review on Predicting Students' Graduation Time Using Machine Learning Algorithms

Nurafifah Mohammad Suhaimi

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor
Email: afifahsuhaimi01@gmail.com

Shuzlina Abdul-Rahman, Sofianita Mutalib, Nurzeatul Hamimah Abdul Hamid, Ariff Md Ab Malik
Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor
Email: {shuzlina, sofi, nurzeatul, arif215}@tmsk.uitm.edu.my

Received: 19 March 2019; Accepted: 23 May 2019; Published: 08 July 2019

Abstract—Nowadays, the application of data mining is widely prevalent in the education system. The ability of data mining to obtain meaningful information from meaningless data makes it very useful to predict students' achievement, university's performance, and many more. According to the Department of Statistics Malaysia, the numbers of student who do not manage to graduate on time rise dramatically every year. This challenging scenario worries many parties, especially university management teams. They have to timely devise strategies in order to enhance the students' academic achievement and discover the main factors contributing to the timely graduation of undergraduate students. This paper discussed the factors utilized by other researchers from previous studies to predict students' graduation time and to study the impact of different types of factors with different prediction methods. Taken together, findings of this research confirmed the usefulness of Neural Network and Support Vector Machine as the most competitive classifiers compared with Naïve Bayes and Decision Tree. Furthermore, our findings also indicate that the academic assessment was a prominent factor when predicting students' graduation time.

Index Terms—Graduate on Time, Prediction, Data Mining, Higher Education

I. INTRODUCTION

The government of Malaysia is keen to transform Malaysia into a high-income and highly industrialized country. To reach these ambitious plans, Malaysia needs to produce more employable higher-skilled graduates who are able to carry out work efficiently. However, as much as government is eager to produce skilful graduates, it is also very concern with students who do not make it graduate on time. Most of the students are graduating on time, but a few students do not do so, and the number is increasing every year. This issue has been brought up by The Ministry of Education, Professors, Academician and many more as the number of students who do not complete their study within the stipulated time is

increasing dramatically every year. This serious issue impacts the universities as good graduation rate is at stake that also the main key when measuring the position of a university in the education industry. Another key point from [1], the writers emphasized that student graduation rates were often used as an objective metric to measure institutional effectiveness. This challenging scenario is eye-opening that worries many parties, especially the management team of UiTM. Research reveals that some of the factors that contribute to the failure of students graduating on time are gender issue [2], marital status and age issue [3], family issues [4], and previous academic record [5]. When students are not able to focus on study due to many issues, they tend to have a problem in giving 100% commitment to their study. Hence, they are not performing well in their final examination or to make it worse, some of them fail it. Consequently, they have to carry the failed paper in the next semester that drag times which is one of the reasons they are unable to graduate on time.

According to [6] one of the areas that have been put a lot of attention among researchers is education, which is the way or mechanism to predict students' graduation time [7]. Researchers have pointed out that students' performance was the most vital thing to be considered if the intention was to predict students' graduation time [8,9,10]. By analysing students' performance, students' graduation time can be forecasted; whether it was on time or not. Hence, academicians and university management team can put more attention on those who are unable to graduate on time. However, analysing the students' performance is very complicated as it involves a huge amount of data that is continuously increasing year by year. Consequently, this huge amount of data is complex that needs to be analysed by computers. Data mining is the process that is able to analyse this huge data, so the interesting patterns of the data can be extracted and utilized to improve our education system [11].

Overall, the main motivation of this research is to improve the university's performance by being able to detect students who are most likely to not be able to graduate on time early and ahead of time. Hence, sturdy

plans and approaches can be done to help these students to improve their study performance, and able to graduate on time.

The objectives of this literature review were to analyse the current research on the prominent factors affecting students' graduation time as well as to examine the impact of different types of data on different classifiers. This paper was arranged as follows: First section summarized the overview of this research. Section Two described the higher education system in Malaysia and the discussion about issues pertaining to "graduate on time" among university students. The third section discussed data mining as well as the application of it in the education system. It was followed by a brief discussion of the machine learning algorithm in Section Four. Related works are in the fifth section. For the next section, the detail results of the existing prediction methods in predicting students' graduation time were elaborated. Lastly, the discussion and conclusion of this systematic literature review were outlined in Section Seven and Section Eight, respectively.

II. MALAYSIAN HIGHER EDUCATION SYSTEM

The education system is one of the areas that has evolved rapidly in Malaysia, as education is very important to society. Malaysian Higher Education System has made notable achievement in escalating student enrolment, gaining worldwide recognition based on the publication of research papers, scientific research, as well as catching international students' attention to choose Malaysia as their study hub. As stated in Malaysia Education Blueprint 2015-2025, Malaysia is committed to gratifying its vision to increase the enrolment numbers of international students by 250,000 students in 2025. These achievements are salient for the country's Gross Domestic Product and export earnings as well as to stay abreast with global trends [14]. The Ministry of Higher Education is known as MOHE which is responsible for higher education in Malaysia. As shown in Fig. 1, higher education institutions can be categorized into two categories which are government-funded institutions and private higher educational institutions.

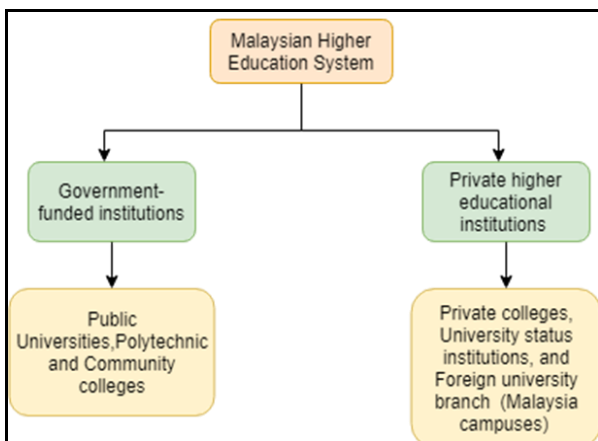


Fig.1. Malaysian Higher Education System

Public universities, polytechnic and community colleges are fall under the Government-funded institutions while private colleges, university status institutions, and foreign university branch campuses like Monash University Malaysia and The University of Nottingham Malaysia Campus fall under Private Higher Educational Institutions. Currently, there are 18 public universities and 47 private universities in Malaysia. The primary medium of instruction in most education institutions is English. The Ministry has classified 20 Malaysian public universities into 3 groups according to the university's focus area and strength. The 3 groups are Malaysian Research Universities (MRU), Malaysian Technical University Network (MTUN) and Malaysian Comprehensive Universities (MCU2). Fig. 2 explains in details about these groups together with the selected Universities of each group.




	Profile	Potential aspirations	Universities in grouping
 Malaysian Research Universities (MRU)	<ul style="list-style-type: none"> High population of post-graduate students High priority on R&D and engage in extensive research activities across multiple faculties and multi disciplinary areas 	<ul style="list-style-type: none"> Aspire for overall research excellence with research partnerships across broad set of leading universities globally 	UKM UM UPM UTM USM
 Malaysian Technical University Network (MTUN)	<ul style="list-style-type: none"> Specialisation in technical programmes; including technical vocational education and training (TVET) Strong focus on employer linkages in technical and professional fields 	<ul style="list-style-type: none"> Typically aspire for excellence in technical or professional fields, whether in teaching and instruction or in niche areas of technical research 	UTeM UTHM Unimap UMP
 Malaysian Comprehensive Universities (MCU)	<ul style="list-style-type: none"> Primary focus on undergraduate and postgraduate instruction across broad range of subjects High priority on high quality education and training, innovations in programme design 	<ul style="list-style-type: none"> Typically aspire for excellence in overall teaching and instruction in areas of specialisation 	UiAM UiTM UMK UMT UUM USIM UniSZA UPSI UMS UPNM UNIMAS

Fig.2. Public universities category (Enhancing Academic Productivity and Cost Efficiency, 2017)

As can be seen in Fig. 2, we can conclude that MRU category focuses more on postgraduate students as it involves research program while MTUN focuses on technical programs only. As expected, the number of postgraduate students who have graduated in Universiti Putra Malaysia (UPM) and Universiti Teknologi Malaysia (UTM) are 13214 and 10573 respectively, contrary to Universiti Teknikal Malaysia (UTeM) which only has 1358 number of postgraduate students, as UTeM's prime focus is not on research program, hence the little number of postgraduate students. The Ministry mentioned that by categorizing these universities into its focus area, the productivity of a university can be improved as each university will focus on one thing at once, so the work's efficiency will increase, hence, an excellent final outcome can be produced [65].

Technically, undergraduate and postgraduate students undergo different ways of learning and requirements for students to complete their respective degrees. Undergraduate students tend to participate more in out-of-classroom activities such as recreation and sports. Therefore, the university has to invest more in the necessary infrastructure to support those activities. In contrast to undergraduate students, postgraduate students usually spend their time more in the research lab, with a one-to-one meeting with the supervisor. Here, the university needs to invest more in providing research equipment, as well as providing research grants to

support projects carried out by postgraduate students which will result in the manuscript publication. Ministry aims to expand the enrolment in higher education to 53% by 2025, which will require an additional 1.1 million places.

On top of that, each student who pursues their study necessitates several phases that require involvements from many parties. As illustrated in Fig. 3, there are 5 main phases involved, starting from the application process to the students' graduation.

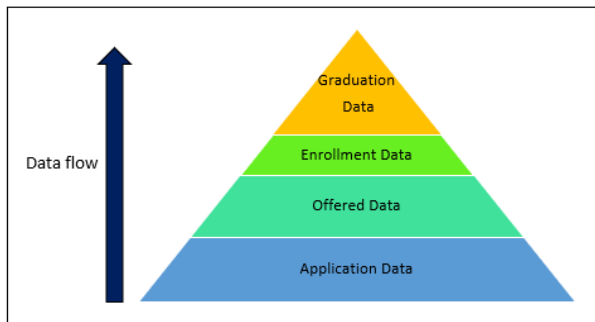


Fig.3. Overview of Data Flow from UPU to Graduation

As can be seen in the figure above, we illustrated the phases in a pyramid chart because it has the structure of a triangle with lines dividing it into sections. Each position represents the data type that illustrated different roles in MOHE. The different width of the section indicates the hierarchy level and the volume of the data. The widest section which is the application data from Unit Pengambilan Universiti (UPU) contains general data but the volume of the data is the largest among other sections. UPU is a centralized platform that manages the application process for Public Universities (UA), Polytechnics, Community Colleges and Institut Latihan Kemahiran Awam (ILKA) in Malaysia. Instead of sending multiple applications directly to various institutions, UPU allows the student to apply for several courses and institutions at the same time via its system. UPU accepts applications from Sijil Pelajaran Malaysia (SPM) leavers, as well as those who have completed Sijil Tinggi Pelajaran Malaysia (STPM), Matrikulasi, Asasi or Diplomas from public universities. Basically, UPU holds extensive data of all type of student across Malaysia who is eligible to apply university application. Moving up, in "offered data" the position becomes more specific as it only holds data regarding students that have successfully been selected to enroll in the university. The volume of data has shrunk to a smaller size compared to the previous section as a university will only capable to cater only a number of students, depends on the quota that has been set by the university. Offer letter then will be produced to the selected students. When the students accept the offer and register at the university, this data will be stored as "enrolment data", which now become more specific as not all of the students will accept university's offer to pursue their study in its institution. Once students have enrolled in the university, they usually will spend approximately 3 to 4 years to complete

their study, depends on what course they enroll. Finally, when students have successfully completed their study, students now will be able to graduate. The data of students who have graduated will be stored in the university's database. This is the narrowest section as it contains the most specific data because it is the last phase, and the volume of the data is the smallest as not all students in the university able to graduate. Specifically, graduates are classified into two categories, which are on time graduation and delay graduation. These categories will be discussed briefly in Section A.

A. An overview of "Graduate on Time" in the Higher Education System

The concept of graduate on time (GOT) can be best explained as a condition where students finish or complete their studies within the stipulated time. Another definition by [15], GOT is a state where students accomplished their studies in respective time that has been set by the university. Furthermore, according to Texas Tech University, timely graduation is defined as: "When the students graduated within the university expected timeline". As a matter of fact, a report from [65] the reporter claimed that for Bachelors program, all programs except architecture, medicine, dentistry and veterinary with Diploma Certificate takes 3 years to graduate, while with other qualifications, it takes 4 years to graduate. However, for architecture, medicine, dentistry and veterinary programs, it takes 4 years to graduate with a diploma certificate and five years without a diploma certificate. In sum, the estimated time for undergraduate students to complete their studies is around 3 to 4 years, depending on the programs that they take. Moreover, as mentioned by [16] most of the institutions in other countries in Europe have set the time for undergraduate students to graduate is four years, but regrettably, a lot of students delay in finishing their degree (40% completed within four years while 60% in six years). Nevertheless, almost all institutions nowadays are having a problem in dealing with students who fail to graduate on time. A report from "Statistik Pendidikan Tinggi 2017 - Bab 2: Universiti Awam", the number of graduates from public universities is decline drastically [66]. The increasing number of students who are unable to graduate on time will significantly affect the institution to produce the quality outputs each year and contribute the low score on the graduation rates. Consequently, it will give bad impact to the university's performance, hence affect the university's ranking.

According to Minister of Higher Education, students who take a longer time to graduate will disturb the university's budget as the university has to spend more money to those who extend their study since they need to hire more lecturers to solve this issue. As written in afterschool.my, there are about 25% of students in Malaysian universities overstay their course duration [67]. The author also suggested that the government should penalize students that fail to graduate on time so this action could be as a reminder to others to always take their study seriously. This challenging scenario worries

many parties, especially the University's management team. The team has to think outside of the box and come out with a sturdy plan on how to overcome such problem as well as improving the number of graduation rates. They need to handle this issue brilliantly and proactively as the university's achievement depends highly on the graduation rate. One of the solutions to handle this issue is by analyzing students' performance since it can be the indicator to predict the students' graduation time. However, analysing the performance of students is very complicated and tedious as it involves a lot of data that is continuously increasing year by year. Alternatively, data mining is applicable to perform analysis in solving this issue and is discussed in detail in the next section.

III. DATA MINING

Data mining is a process that can make a huge amount of data turns into meaningful information or knowledge. Kantardzic defined data mining as the "process of discovering various models, summaries, and derived values from a given collection of data" [18]. Data mining is also can be defined as a mining process to analyze and categorize the information as well as summarize the knowledge from different kind of data stored in database and data warehouse [17]. Data mining is used for acquiring patterns and summaries of new and important information from the present data which can be referred to as data description. Then, the described data can be used to make a prediction of a certain problem or area, which now is referred to as data prediction. [18] stated that in the university context, most of the present system only maintain the information of students by only enabling university management to store and retrieve that information. Technically, with the utilization of the data mining process, it makes the present system becomes smarter [19]. This statement seems to be realistic as [20] stated that Data mining is a perfect tool for gleaning and extracting useful information from the meaningless data. In short, the vast amount of university data needs to be explored, extracted and analyzed so top management is able to come out with a set of meaningful information that can be understood for further quality improvement. Thanks to technology, these large data can be mined using Data mining Techniques, hence, the quality of the educational processes can be ameliorated [4,20]. Technically, mining in the education area is called Educational Data Mining (EDM). Fig. 4 illustrated the iterative phases involved in the application of Educational Data Mining systems.

As illustrated in Fig. 4, lecturers are the ones who are responsible for designing, planning, building, and maintaining the educational systems. For example, lecturers here provide the learning materials to the students as well as supplying students assessment report such as their test marks to the educational system. Students, on the other hand, utilize all the facilities provided by the university, communicate with lecturers and other students as well as participating in the classroom, event or workshop. Students also provide their

personal information like their age, gender, marital status and many more to the educational system.

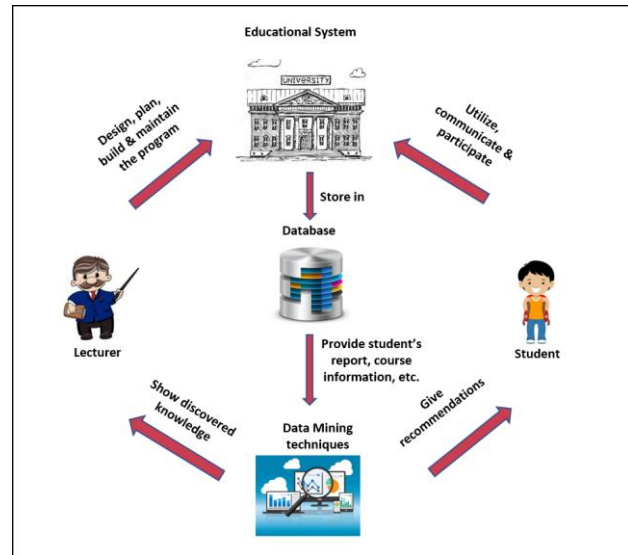


Fig.4. Educational Data mining data phases model

Next, all the collected data from the educational system will be stored in database and mined using data mining technique to present the discovered knowledge like students' performance to the lecturers. Besides, it will also provide useful recommendations to the student such as which subject that student should enroll, based on their previous academic achievement. Technically, this analysis provides new information to both parties that would be difficult to discern by simply looking at the raw data. Moreover, a study by [3] concluded that Data mining techniques are applicable in supplying students' performance to lecturers in order to assist them to identify students with low academic achievements at an early stage. Hence, initiatives can be made to help those students to improve their academic performance. To sum up, the application of data mining nowadays is widely prevalent in the education system as it is able to obtain meaningful information from meaningless data. It also is very useful to analyze students' performance, forecast students' graduation time and other related issues [23,24]. The accuracy of data mining can be improved as well as performing a deeper analysis of the mined data with the use of machine learning algorithms. Further explanation regarding machine learning algorithm will be discussed in the next section.

IV. MACHINE LEARNING ALGORITHM

In 1959, Arthur Samuel defines machine learning as "A field of study that gives computers the ability to learn without being explicitly program." Three decades later, Tom Mitchell outlines a new improvised definition which makes more sense in engineering field which is "A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E." Nowadays, machine learning algorithms

are the methods usually employed by researchers to discover patterns from data sets by letting them learn on their own [26]. In addition, with the advancement of extensive computing capabilities, learning through a tremendous amount of data is now seems possible. In other words, machine-learning is practically suitable for analysis as humans are likely to make errors. Machine

learning uses self-learning algorithms to improve its performance at a task with experience over time. It can be used to reveal insights and provide feedback in near real-time. Machine-learning can be classified into various types of learning such as supervised learning and unsupervised learning. Fig. 5 shows the taxonomy of machine learning algorithms.

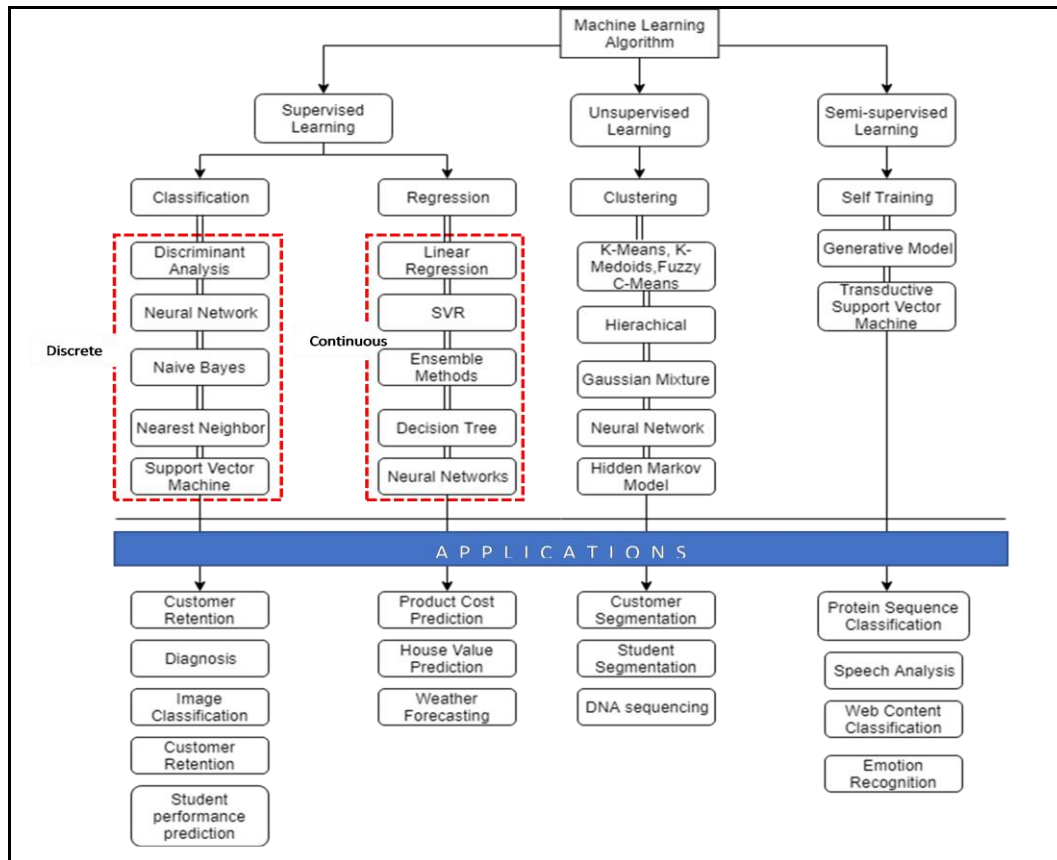


Fig.5 Taxonomy of Machine Learning Algorithm

Supervised machine learning aims to create a model that can predict the future, based on the evidence in the presence of uncertainty. A supervised machine learning algorithm will train a model to make new, yet reasonable predictions based on what it has learned before [27]. Classification and regression are famous techniques in supervised learning. In the context of classification, this technique can be applied to map the input to discrete output variables which are usually known as labels and categories. In brief, class or category for a given observation is predicted by the mapping function. For example, the performance of a group of students can be classified as “perform” or “not perform”. There are many techniques that fall under classification such as Support Vector Machine, Naïve Bayes, Discriminant Analysis, Nearest Neighbor, Neural Network, and Naïve Bayes. For regression, we apply this technique when we want to map input to continuous output variables. A continuous output variable is an integer real-value such as amounts, sizes and temperature. For example, a house may be predicted to sell for a specific Ringgit value, perhaps in the range of RM 400,000 to RM600,000. Regression, SVR, Ensemble

Methods, Decision Tree and Neural Network are the common techniques of regression.

The discovery of the hidden patterns from data sets is the main objective of unsupervised learning as well as producing an inference from it. Clustering is the most common unsupervised learning technique that works to explore data in the data analysis process to find hidden patterns, corresponding to the objective of unsupervised learning. As defined by [28], clustering is when the data is divided into groups of similar objects. In some applications, clustering is also known as data segmentation as it divides huge data set into many groups, where each group shares a similar characteristic [29,30]. On top of that, clustering nowadays is commonly deployed in education Data mining to group students based on their characteristics. [31] applied unsupervised clustering method in his research to predict the status of students' graduation within five major departments at California State University Northridge (CSUN). In the discussion part, he pointed out that identifying common traits among students within each cluster is very handy when cluster analysis is used. Another work by [32], the

writers are concerned with the huge size of sequenced data of DNA, hence, they come out with a machine learning approach to cater to the DNA copy number issue. In their project, they utilized with Gaussian Mixture (GM) together with Bayesian hidden Markov model (HMM) Clustering approach in order to replicate the change of DNA copy number across the genome. [31] also claimed that the achievement of GM and HMM have smashed the existing method like Binary Segmentation. On the other hand, another famous clustering algorithm is k-means. It is known as the simplest clustering algorithms in solving the clustering problem. This statement concurs well with [33] as they also outline that the k-means algorithm has been extensively used among researchers as it is a straightforward algorithm that requires minimum effort to implement. K-means is categorized in the non-hierarchical clustering method which works in finding to group the data into the form of one or more clusters [34]. The data that have the same criteria are categorized into one cluster so that the other data that have different criteria are categorized into another cluster. To sum up, in supervised learning, data is needed to be labeled to its expected output for it to learn, while in unsupervised learning, it can learn without any label on the data. As can be seen from Fig. 5, Neural Network learning can be either supervised or unsupervised. According to [35] with supervised training, neural network undergoes iterations, until the actual output of the neural network matches the expected output. Similarly, unsupervised training works like supervised training but there is no expected output given. Unsupervised training usually is used in Neural Network when we want to classify the inputs into various groups.

Semi-supervised learning is catching on quickly among developers and researchers, as the numbers of unlabelled data are bigger than labeled data in a dataset and this issue is growing speedily [36,37]. Basically, semi-supervised learning is trained on a combination of labeled and unlabeled data. As mentioned previously, supervised learning deals with labeled data but it spends longer time to complete its process, hence, it costs more money. Besides, when the model is labeled with too much labeling, it urges human biases on the model. To cater to such problems, the deployment of unlabelled data has improved the accuracy of the classifier as well as to reduce the cost and time taken to build it. There are several categories that fall under semi-supervised learning algorithms such as Self-Training, Generative Model and Transductive Support Vector Machine. Self-Training is a wrapper algorithm where the labeled data are trained by the classifier and later will classify unlabelled data [38]. The unlabelled data that is associated with the highest confidence score will be added to the training set. As stated in [38], self-Training is known as the simplest algorithm in semi-supervised learning but still gives a good solution to the classification problem. Contrary with Self-Training,

Generative Model has difficulties in providing good solutions to the classification problem, especially when the unlabelled data is more than labeled data, as it has a problem to balance the effect of unlabelled and labeled data. Generative Model technically works best with sentiment analysis and predicting handwritten digit number, as stated in [40] Transductive Support Vector Machine on the other hand is the extension of Support Vector Machine that works with unlabelled data [41]. The application of semi-supervised learning is protein sequence classification, speech analysis, web content classification and emotion recognition.

In this section, we have discussed briefly machine learning algorithms. To sum up, supervised learning works with labeled data, unsupervised learning works with unlabelled data, while semi-supervised learning works with both type of data. For the next section, we discussed the principal factors that have been utilized by researchers to predict students' graduation time.

V. RELATED WORKS

A total of 33 papers on student performance predictions were studied and analyzed to gather all the factors that influenced the performance of students. Our research bears a close resemblance to a study by [1], as they utilized student's performance to predict the students' graduation delay. They believed that the graduation time of a student depends highly on the performance of that student itself. There are several studies that discuss the factors affecting the student's performance. According to [42], the performance of students has become a concern for educators. In the paper, the importance of determining the factors was highlighted as the factors are significant to analyze the performance of students. He also mentioned in his study that the top factors that act as the main contribution in analyzing students' performance are a family expenditure, family income, personal information, and the family's assets. Based on his research, he concluded that family expenditure and personal information are showing a remarkable effect on the student's performance. From our reading, we found out that there are 26 attributes that have been frequently used by other researchers, for measuring student performance As illustrated in Fig. 6, we combine all the factors from 33 articles and tabulate the data into a bar graph, so we can see clearly which factor has been extensively used among researchers in order to analyze students' performance. The number in the figure which is above the bar represents the total number of papers that employed the factor as the predictor. As previously mentioned, analyzing student performance is one of the efficient ways to forecast students' graduation time, whether they will graduate on time or not.

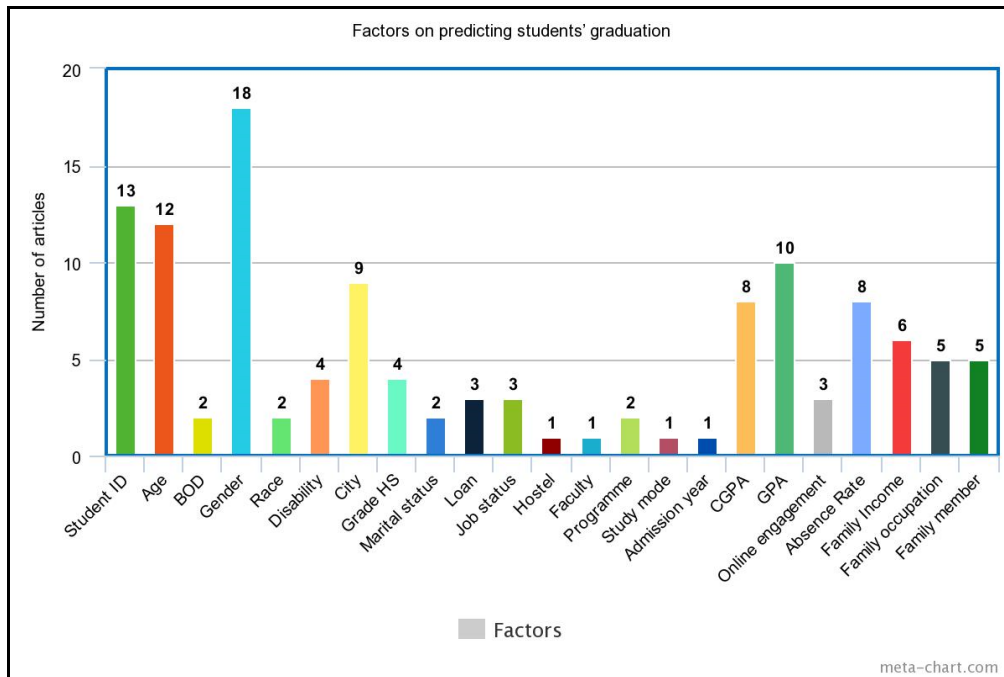


Fig.6. Factors on predicting students' graduation

From the graph above, we can see clearly that gender is the main factor that can be used to analyze students' performance. Discussions regarding the effect of gender on study performance have dominated research in recent years. [42] have identified that the study style of male and female is different, where female tends to study in the more systematic way while male, on the other hand, tends to study at the last minute of time. In fact, [43] mentioned that from their research, most of the female students have more positive behavior compared to male. A survey conducted by [44] she distributed 54.9% of workload to women and 45.1% to men, but both genders finish the tasks assigned to them at the same rate. From this survey, it shows clearly that woman able to complete their task efficiently within the stipulated time, compared to men. The main reason why female works better than male is because the female has better self-discipline and more focus in their study or work compared to men [45]. Above all, this statement can be strengthened by research done by [15]. They proposed a predictive model to predict the graduation status of PHD students. The result shows that there are a total of 79 students who are predicted to graduate on time. Based on their research, they outlined that female tend to graduate on time compared to male as the number of female students who are predicted to graduate on time is 56% higher compared to male. As male and female have different learning rate and behavior, gender is important as it gives an impact on students' performance. Apart from that, 12 out of 23 articles used age as the factor to analyze student performance. [46] explained that as we are aging, there are some physical changes in our brain that make us have more difficulty to remember or learn efficiently. Howbeit, the statement made by [46] raises doubts as several authors believe that age is not a significant predictor of

students' performance, as outlined by [3]. This finding is consistent with [47] who discovered that there is no significant relationship between academic performance and age. In a similar vein, Murray - Harvey (1994) emphasized that age does not give an impact on students' GPA, which means, age is not a significant predictor to predict students' graduation time.

Apart from gender, we can see that a lot of researchers used GPA to analyze students' performance as GPA has a high influence to predict students' graduation time. If the students are seen to have low GPA during the first and second year of their studies, they might probably not be able to complete their study within the study plan. The reason is that they might have to carry any failed paper to the next semester, which most probably, will extend their study time. However, even though just a few researchers opted marital status as the factor to analyze student performance, we believe that marital status also gives a huge impact on students' performance. [45] stated in his report that married students have limited time to study or to do revision as they have other responsibilities to carry out like cooking and cleaning the house. Plus, when they are not able to give 100% commitment to their study, they tend to neglect their responsibilities as a student, such as do the assignments given or fully prepare themselves for test or examination. With such behaviors, the probability for them to not excel in their study is high, which will result in a higher chance to not graduate on time. After we have done analysing the prominent factors that contribute to the students' graduation time, we will utilize data mining technique to make a prediction to predict students' graduation time. Next section discusses in brief the prediction methods to predict students' graduation based on their performance.

VI. PREDICTION METHODS

In order to predict students' graduation time, predictive modeling is employed. Predictive modeling utilizes data mining techniques to make a prediction on what it has learned during the learning process. There are four common prediction methods or more known as classifiers that have been used by researchers to build this predictive model which are Decision Tree, Naïve Bayes, Neural Network and Support Vector Machine classifier. The detailed mapping of the factors and accuracy of each classifier will be discussed in the next section.

1) Decision Tree

Decision Tree (DT) is one of the powerful classifiers that has mostly been employed among researchers in their study for prediction [46]. Specific entities are classified into classes and represent in a tree structure form [47]. This classifier basically accepts a set of inputs and able to produce output. [48] used data mining technique to evaluate students' data using the J48 decision tree in Weka tool. Based on their findings, they stated that DT is helpful in predicting students' academic result and easy to interpret. This fits well with research done by [49] as they also found out that by using DT, the result is easy to understand especially for teacher or lecturer who has no insight about data mining since the prediction result is explained in the form of IF-THEN rules. Another work by [50], they identified students who are likely to fail in their study based on factors like previous academic result and family issue, so efficient actions can be taken on those students. They employed several classifiers in their work which are Naïve Bayes (NB), and J48 decision tree. The result shows that DT had a higher accuracy which is 73.92% followed by NB which is 68.60 %.

2) Naïve Bayes

NB is a classifier that utilizes probability and Bayesian Theorem to make a prediction based on certain features that have been set in the program. [51] utilized NB algorithm to predict the academic performance and behavior of students at secondary school. The students are classified into two categories which are pass and fail. The factors used in this model to make classification is mostly related to academic performance and WEKA tool is used to calculate the accuracy of the classifier which is 87%. Other than that, [28] applies the NB model to predict students' performance and claimed that NB is a straightforward classifier and able to works well with huge data sets. Even though it is a straightforward classifier that works in a simple way, NB outperformed even highly sophisticated classification methods like SVM and DT in his research. Berkhin's claim is correlated favorably with an experiment done by [18], as they discovered that NB outperformed other classification methods such as Regression, DT and Neural networks (NN) in predicting students' performance. A further experiment carried out by [24] they mentioned that NB tends to converge faster, hence it takes a shorter time to

make a decision and perform well compared to another classifier.

3) Support Vector Machine

Support Vector Machine (SVM) is one of the famous prediction methods to predict student performance. SVM constructs hyperplanes in multidimensional space to perform its classification, which separates class levels into different cases. [22] concerns with the importance of enhancing student pass rates as it reflects the school teaching level. Therefore, they have analysed the prominent features of students and forecasted their pass rate using SVM and DT. Their finding pointed out that SVM achieved higher accuracy than DT because SVM is beneficial for binary classification problems, which in this context, pass or fail. Another key point from [68] he mentioned that SVM tends to not perform well with large data set as the time for training is longer. This statement lends support to [55] as based on their analysis, they concluded that SVM works amazingly with small data sets for prediction.

4) Neural Network

Neural network (NN) is growing at a fast pace and has been employed by researchers for classification in many applications area [52]. It is a system of interconnected neurons that has input layer, hidden layer and output layer [53]. The neuron is associated with other neurons for a network to develop. The input value from the input layer will be computed to produce the expected output [54]. All the potential interactions between predictors variable can be discovered by Neural Network [55]. A work is done by [56] they analyse the impact of different attributes of students' performance on classifier by using Neural Network, Naïve Bayes, Decision Tree, and Nearest Neighbour. Their finding shows that when NN is used to classify all attributes, the accuracy is 78.6, while without behavioral attributes, the result decreased to 52.3. However, when the student absent days attribute is eliminated, the accuracy went up again to 64.5. These findings show clearly that the behavioral factor affects the accuracy of the classifier. Moreover, [57] investigate the performance of NN to make a prediction of the students' performance in SPAECE. They considered several factors that are applicable to predict students' performance which are grade, length of study and previous academic score and used these variables as the inputs in their experiment. In the test data evaluation, the result shows that NN is able to correctly predict 95% of the students' performance correctly predict 95% of the students' performance.

VII. DISCUSSION

In this section, the relationships between the factors for predicting students' graduation and classifiers were discussed. Besides that, the impacts of different factors on different classifier were also studied and discussed. As illustrated in Fig. 7, the researchers have categorized all the factors into five groups.

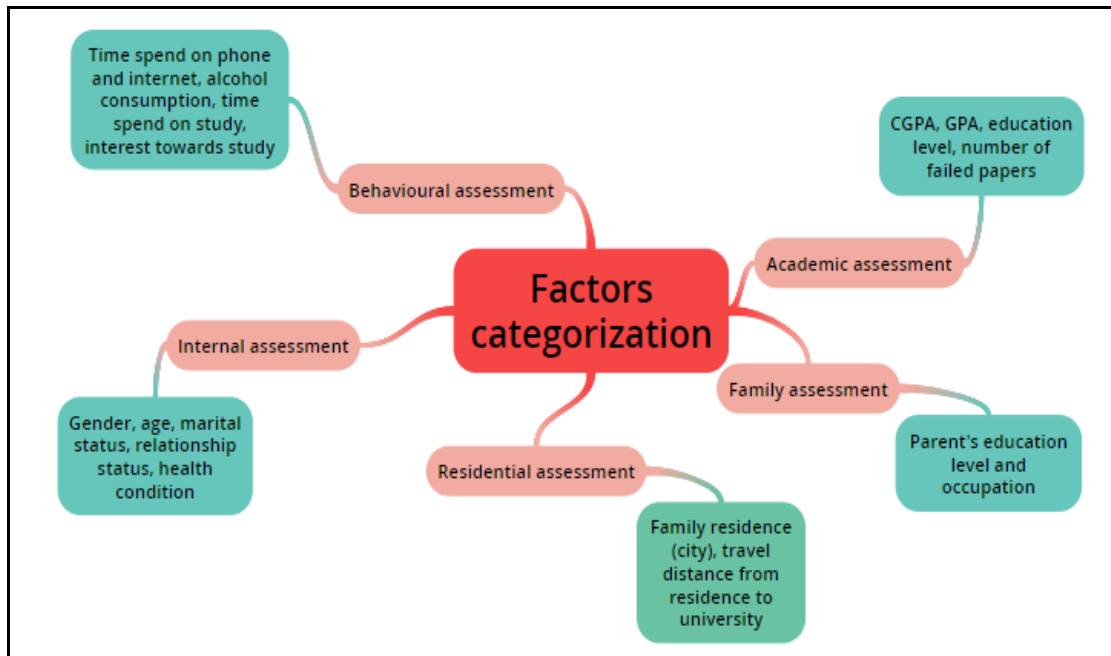


Fig.7. Mind Map of Factor Categorization

Based on our analysis of different research papers, we discovered many factors that have been used by other researchers in their work to predict students' graduation time. We have collected all the factors and grouped them based on their characteristics. We came out with five categories of factors which were internal assessment, academic assessment, behavioural assessment, family assessment and residential assessment. The last nodes of the mind map above showed the characteristics of their factor. For example, CGPA, GPA, education level and

number of failed papers were the characteristics of academic assessment. In brief, all factors regarding student's academic were classified as "academic assessment". Table 1 shows the way we classified those factors. The classification made the task of analysing the impact of different factors on different classifier became easier. From our evaluation of different articles from 2014 until 2018, we have tabulated our findings in Table I.

Table I Analysis of Students' Performance Factors and Classifier Accuracy

Classifier	Factors	Accuracy	Author/Title
Naïve Bayes	Internal assessment, academic assessment, family assessment, behavioural assessment	75.9%	[57]
	Internal assessment, residential assessment, academic assessment, family assessment	69.07%	[58]
	Academic assessment	83.65%	[59]
	Academic assessment, behavioural assessment	70%	[51]
Decision Tree	Behavioral assessment, internal assessment, residential assessment, family assessment, academic assessment	73.92%	[50]
	Academic assessment	79.5%	[60]
	Internal assessment, academic assessment, family assessment, residential assessment	72.5%	[61]
Neural Network	Academic assessment, behavioral assessment	78.6%	[56]
	Academic assessment	95%	[57]
	Internal assessment, academic assessment	84.6%	[35]
Support Vector Machine	Internal assessment, academic assessment, family assessment, behavioural assessment	93.9%	[62]
	Academic assessment	87.5%	[63]
	Internal assessment, academic assessment, behavioural assessment	80%	[64]

By looking at the table above, we can see that NN has the highest prediction accuracy which is 95%. Technically, the accuracy result on prediction depends highly on the features that were used during the prediction process. NN scored the highest prediction accuracy because of the influenced of a particular factor during the running process. In this context, the factor involved is an academic assessment. Based on our analysis, academic assessment is the main factor that has been utilized with NN. With the presence of academic assessment only, the prediction accuracy increased dramatically. However, when the behavioral assessment is combined with academic assessment as the inputs to make a prediction, the accuracy score decreased. Findings from [55] has further strengthened our confidence in our finding as they also highlighted that the behavioral assessment declined the accuracy level of NN. This happened because NN has difficulty to make a prediction on qualitative data, which is non-numerical in nature.

Next, SVM scored the second highest of the prediction accuracy which is 93.95%. The reason why SVM worked well is that it is able to solve high-dimensional data which is data that associates with many attributes and features [22]. SVM also can perform efficiently, with only small tuning. Findings showed that internal, academic, family and behavioural assessment are the relevant factors that have been utilized for the classifier to make a prediction. Contrary to NN, when the academic assessment is used alone in SVM, the accuracy level decreased. This indicates that academic assessment needs to be combined with other factors in order for SVM to achieve high accuracy.

On the other hand, DT and NB are seen to be able to work with all type of factors and produce an average accuracy level. Based on table 1, DT and NB can work with the categorical and numerical data as the accuracy score of this classifier did not give any huge differences with both types of data. However, when the academic assessment is the only factor that is used as input, the accuracy score increase. Thus, it shows that NB and DT worked efficiently with academic assessment factor, but still produce an average accuracy score when works with other factors. Overall, our finding validates the usefulness of NN and SVM as the most competitive classifiers compared to NB and DT, while academic assessment is the prominent factor to make a prediction on students' graduation time.

VIII. CONCLUSION

This paper reviewed and analysed the current research on the prominent factors affecting students' graduation time as well as analysed the data mining techniques utilized by previous researches to predict students' graduation time. In brief, the researchers analyzed the factors of students' performance, and appointed them as indicators for predicting students' graduation time. Technically, students with good academic performance most likely graduate on time, and vice versa. The accurate prediction will help university management to

manage students who are not likely graduate on time and come out with plans to assist them in improve their performance and eventually graduate on time. In fact, good graduation rates are often acting as an objective metric to measure the effectiveness of university, hence, good graduation rates will improve the university's rank. This paper has discussed the factors utilized by other previous researchers to predict students' graduation time with the prediction methods involved. Taken together, findings of this study indicated the possibilities and potentials of NN and SVM to predict students' graduation time in higher education institutions in Malaysia. These models are developed based on several inputs factors and they work significantly positive with academic assessment factors. Academic assessment factors related to student's past and current performance in their study were very relevant to make a prediction on their graduation time. In conclusion, the impact of this work laid in intention to help and assist other researchers in developing a real model that can predict students' graduation time easily and accurately. It will also help the educational system in predicting students' graduation time systematically and assist lecturers in identifying students who need more attention in their study. Further work needs to be done on a real model with the utilization of our findings in this research to confirm them.

REFERENCES

- [1] T. Ojha, G. L. Heileman, M. Martinez-Ramon, and A. Slim, "Prediction of graduation delay based on student performance," 2017.
- [2] M. Dayio, "Gender Differences in Academic Performance in a Large Public University in Gender Differences in Academic Performance in a Large Public University in Turkey Department of Economics Serap Tütüç-A ik," no. April, 2015.
- [3] B. G. Amuda, A. K. Bulus, and H. P. Joseph, "Marital Status and Age as Predictors of Academic Performance of Students of Colleges of Education in the North-Eastern Nigeria," *Am. J. Educ. Res.* Vol. 4, 2016, Pages 896-902, vol. 4, no. 12, pp. 896-902, 2016.
- [4] R. Asif, A. Merceron, S. A. Ali, and N. G. Haider, "Analyzing undergraduate students' performance using educational data mining," *Comput. Educ.*, vol. 113, pp. 177-194, 2017.
- [5] D. Herath, "Educational Data Mining to Investigate Learning Behaviors: A Literature Review," no. May, 2018.
- [6] R. S. Agrawal and M. H. Pandya, "Data Mining With Neural Networks to Predict Students Academic Achievements," vol. 8491, pp. 100-103, 2016.
- [7] O. D. Nurhayati, O. S. Bachri, A. Supriyanto, and M. Hasbullah, "Graduation prediction system using artificial neural network," *Int. J. Mech. Eng. Technol.*, vol. 9, no. 7, pp. 1051-1057, 2018.
- [8] A. Gopalakrishnan, R. Kased, H. Yang, M. B. Love, C. Graterol, and A. Shada, "A multifaceted data mining approach to understanding what factors lead college students to persist and graduate," *Proc. Comput. Conf.* 2017, vol. 2018-Janua, no. July, pp. 372-381, 2018.
- [9] R. Ahuja and Y. Kankane, "Predicting the probability of student's degree completion by using different data

- mining techniques,” 2017 4th Int. Conf. Image Inf. Process. ICHIP 2017, vol. 2018–Janua, pp. 474–477, 2018.
- [10] A. Slim, G. L. Heileman, M. Hickman, and C. T. Abdallah, “A geometric distributed probabilistic model to predict graduation rates,” 2017 IEEE SmartWorld Ubiquitous Intell. Comput. Adv. Trust. Comput. Scalable Comput. Commun. Cloud Big Data Comput. Internet People Smart City Innov. SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI 2017 - , pp. 1–8, 2018.
- [11] D. Kabakchieva, “Student Performance Prediction by Using Data Mining Classification Algorithms,” Int. J. Comput. Sci. Manag. Res., vol. 1, no. 4, pp. 686–690, 2012.
- [12] P. Kamal and S. Ahuja, Academic Performance Prediction Using Data Mining Techniques: Identification of Influential Factors Effecting the Academic Performance in Undergrad Professional Course, vol. 741. Springer Singapore, 2019.
- [13] F. Ahmad, N. H. Ismail, and A. A. Aziz, “The prediction of students' academic performance using classification data mining techniques,” Appl. Math. Sci., vol. 9, no. 129, pp. 6415–6426, 2015.
- [14] M. Hamiz, M. Bakri, N. Kamaruddin, and A. Mohamed, “Assessment analytic theoretical framework based on learners' continuous learning improvement,” Indones. J. Electr. Eng. Comput. Sci., vol. 11, no. 2, pp. 682–687, 2018.
- [15] S. S. R. Shariff, N. A. M. Rodzi, K. A. Rahman, S. M. Zahari, and S. M. Deni, “Predicting the ‘graduate on time (GOT)’ of PhD students using binary logistics regression model,” AIP Conference Proceedings, 2016. .
- [16] W. K. Mujani, A. Muttaqin, and K. A. Khalid, “Historical development of public institutions of higher learning in Malaysia,” Middle-East J. Sci. Res., vol. 20, no. 12, pp. 2154–2157, 2014.
- [17] L. Jing, “Data Mining and Its Applications in Higher Education,” New Dir. Institutional Res., vol. 2002, no. 113, p. 17, 2002.
- [18] T. Devasia, T. P. Vinushree, and V. Hegde, “Prediction of students performance using Educational Data Mining,” 2016 Int. Conf. Data Min. Adv. Comput., pp. 91–95, 2016.
- [19] S. Hussain, R. Atallah, and A. Kamsin, Classification, Clustering and Association Rule Mining in Educational Datasets Using Data Mining Tools: A Case Study, vol. 765. Springer International Publishing, 2019.
- [20] A. Dutt, M. A. Ismail, and T. Herawan, “A Systematic Review on Educational Data Mining,” IEEE Access, vol. 5, pp. 15991–16005, 2017.
- [21] A. Peña-ayala, “Educational data mining : A survey and a data mining-based analysis of recent works,” vol. 41, pp. 1432–1434, 2014.
- [22] X. Ma and Z. Zhou, “Student pass rates prediction using optimized support vector machine and decision tree,” 2018 IEEE 8th Annu. Comput. Commun. Work. Conf. CCWC 2018, vol. 2018–Janua, pp. 209–215, 2018.
- [23] S. S. Athani, S. A. Kodli, M. N. Banavasi, and P. G. S. Hiremath, “Student academic performance and social behavior predictor using data mining techniques,” Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA 2017, vol. 2017–Janua, pp. 170–174, 2017.
- [24] H. Al-Shehri et al., “Student performance prediction using Support Vector Machine and K-Nearest Neighbor,” Can. Conf. Electr. Comput. Eng., pp. 17–20, 2017.
- [25] P. Saraswathi and N. Nagadeepa, Predicting the Performance of Disability Students Using Assistive Tools with the Role of ICT in Mining Approach, vol. 104. Springer Singapore, 2019.
- [26] A. S. Asyraf, S. Abdul-Rahman, and S. Mutalib, “Mining textual terms for stock market prediction analysis using financial news,” Commun. Comput. Inf. Sci., vol. 788, pp. 293–305, 2017.
- [27] F. S. Ismail and N. A. Bakar, “Adaptive mechanism for GA-NN to enhance prediction model,” ACM IMCOM 2015 - Proc., no. October 2016, p. 101:1--101:5, 2015.
- [28] P. Berkhin, “A Survey of Clustering Data Mining,” Group. Multidimens. Data, no. c, pp. 25–71, 2006.
- [29] L. Cahaya, L. Hiryanto, and T. Handhayani, “Student Graduation Time Prediction Using Intelligent K-Medoids Algorithm,” pp. 263–266, 2017.
- [30] A. Ahmad, R. Yusoff, M. N. Ismail, and N. R. Rosli, “Clustering the imbalanced datasets using modified Kohonen self-organizing map (KSOM),” Proc. Comput. Conf. 2017, vol. 2018–Janua, no. July, pp. 751–755, 2018.
- [31] S. Sajjadi, B. Shapiro, C. Mckinlay, A. Sarkisyan, C. Shubin, and E. Osoba, “Finding bottlenecks: Predicting student attrition with unsupervised classifier,” 2017 Intell. Syst. Conf. IntelliSys 2017, vol. 2018–Janua, pp. 1166–1172, 2018.
- [32] G. Manogaran, V. Vijayakumar, R. Varatharajan, P. Malarvizhi Kumar, R. Sundarasekar, and C. H. Hsu, “Machine Learning Based Big Data Processing Framework for Cancer Diagnosis Using Hidden Markov Model and GM Clustering,” Wirel. Pers. Commun., vol. 102, no. 3, pp. 2099–2116, 2018.
- [33] N. Nidheesh, K. A. Abdul Nazeer, and P. M. Ameer, “An enhanced deterministic K-Means clustering algorithm for cancer subtype prediction from gene expression data,” Comput. Biol. Med., vol. 91, no. October, pp. 213–221, 2017.
- [34] M. Singh, H. Nagar, and A. Sant, “K-mean and EM Clustering algorithm using attendance performance improvement Primary school Student,” vol. 1, no. 1, pp. 131–133, 2016.
- [35] S. A. Naser, I. Zaqout, M. A. Ghosh, and R. Atallah, “Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology,” Int. J. Hybrid Inf. Technol., vol. Vol.8, No., no. February, 2015.
- [36] H. Pan and Z. Knag, “Robust Graph Learning for Semi-Supervised Classification,” Int. Conf. Intell. Human-Machine Syst. Cybern., 2018.
- [37] P. H. M. Braga and H. F. Bassani, “A Semi-Supervised Self-Organizing Map for Clustering and Classification,” 2018 Int. Jt. Conf. Neural Networks, pp. 1–8, 2018.
- [38] N. Fazakis, S. Karlos, S. Kotsiantis, and K. Sgarbas, “Self-Trained LMT for semisupervised learning,” Comput. Intell. Neurosci., vol. 2016, 2016.
- [39] I. E. Livieris, A. Kanavos, V. Tampakas, and P. Pintelas, “An auto-adjustable semi-supervised self-training algorithm,” Algorithms, vol. 11, no. 9, pp. 1–16, 2018.
- [40] D. D. F. Adiwardana, A. Matsukawa, and J. Whang, “Using Generative Models for Semi-Supervised Learning,” Med. Image Comput. Comput. Interv. – MICCAI 2016, pp. 106–114, 2016.
- [41] D. P. Kingma, D. J. Rezende, S. Mohamed, and M. Welling, “Semi-supervised Learning with Deep Generative Models,” pp. 1–9, 2009.
- [42] G. Kapil, A. Agrawal, and R. A. Khan, “A Study of Big Data Characteristics,” 2016 Int. Conf. Commun. Electron. Syst., 2014.

- [43] B. Cubic and H. Seibel, "Personality Differences in Incoming Male and Female Medical Students," vol. 9152, no. 304, pp. 1–6.
- [44] R. Murray-Harvey, "Learning styles and approaches to learning: distinguishing between concepts and instruments," 1994.
- [45] O. C. Potokri, "The Academic Performance Of Married Women Students In Nigerian Higher Education Doctor Of Philosophy (Phd) In Management And Policy Studies University Of Pretoria , South Africa Promoter : Prof . Venitha Pillay," 2011.
- [46] N. W. Zamani, S. Shaliza, and M. Khairi, "A Comparative Study on Data Mining Techniques for Rainfall Prediction in Subang," Proceeding Int. Conf. Math. Eng. Ind. Appl. 2018 (ICoMEIA 2018), vol. 020042, 2013.
- [47] M. K. Najafabadi, A. H. Mohamed, and M. N. Mahrin, "A survey on data mining techniques in recommender systems," *Soft Comput.*, pp. 1–28, 2017.
- [48] P. Guleria, N. Thakur, and M. Sood, "Predicting Student Performance Using Decision Tree Classifiers and Information Gain," pp. 126–129, 2014.
- [49] S. Paul and S. K.P, "Data Mining Techniques for Predicting Student Performance," *IEEE*, no. March, pp. 0–2, 2015.
- [50] S. Roy and A. Garg, "Predicting Academic Performance of Student Using Classification Techniques," *IEEE*, pp. 568–572, 2017.
- [51] U. Pujianto, E. N. Azizah, and A. S. Damayanti, "Naive Bayes Using to Predict Students ' Academic Performance at Faculty of Literature," pp. 163–169, 2017.
- [52] A. Halinka, P. Rzepka, and M. Szablicki, "Agent model of multi-agent system for area power system protection," *Proc. - Int. Conf. Mod. Electr. Power Syst. MEPS 2015*, pp. 1–4, 2015.
- [53] N. N. Hamadneh, W. S. Khan, and W. A. Khan, "Prediction of thermal conductivities of polyacrylonitrile electrospun nanocomposite fibers using artificial neural network and prey predator algorithm," *J. King Saud Univ. - Sci.*, 2018.
- [54] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. De Jesús, *Neural Network Design (2nd Edition)*. 2014.
- [55] A. M. Shahiri, W. Husain, and N. A. Rashid, "A Review on Predicting Student's Performance Using Data Mining Techniques," *Procedia Comput. Sci.*, vol. 72, pp. 414–422, 2015.
- [56] M. Hasibur Rahman and M. Rabiul Islam, "Predict Student's Academic Performance and Evaluate the Impact of Different Attributes on the Performance Using Data Mining Techniques," *2nd Int. Conf. Electr. Electron. Eng. ICEEE 2017*, no. December, pp. 1–4, 2018.
- [57] R. M. De Albuquerque, A. A. Bezerra, D. A. De Souza, L. B. P. Do Nascimento, J. J. De Mesquita S á and J. C. Do Nascimento, "Using neural networks to predict the future performance of students," *2015 Int. Symp. Comput. Educ. SIIIE 2015*, pp. 109–113, 2016.
- [58] A. Kesumawati and D. T. Utari, "Predicting patterns of student graduation rates using Na ïve bayes classifier and support vector machine," vol. 060005, p. 060005, 2018.
- [59] R. Asif, A. Merceron, and M. K. Pathan, "Predicting Student Academic Performance at Degree Level: A Case Study," *Int. J. Intell. Syst. Appl.*, vol. 7, no. 1, pp. 49–61, 2014.
- [60] A. O. Ogunde, "A Data Mining System for Predicting University Students ' Graduation Grades Using ID3 Decision Tree Algorithm A Data Mining System for Predicting University Students ' Graduation Grades," no. August, 2014.
- [61] C. Anuradha and T. Velmurugan, "A Comparative Analysis on the Evaluation of Classification Algorithms in the Prediction of Students Performance," *Indian J. Sci. Technol.*, vol. 8, no. July, pp. 1–12, 2017.
- [62] V. L. Miguás, A. Freitas, P. J. V Garcia, and A. Silva, "Early segmentation of students according to their academic performance: A predictive modelling approach," *Decis. Support Syst.*, vol. 115, no. September, pp. 36–51, 2018.
- [63] Pushpa, Manjunath, Mrunal, M. Singh, and C. Suhas, "Class Result Prediction using Machine Learning," vol. 6, pp. 1208–1212, 2017.
- [64] M. Mayilvaganan and D. Kalpanadevi, "Comparison of Classification Techniques for predicting the performance of Students Academic Environment," *Int. Conf. Commun. Netw. Technol.*, pp. 113–118, 2014.
- [65] *Enhancing Academic Productivity and Cost Efficiency*. (2017). Putrajaya: Ministry of Higher Education Malaysia.
- [66] *Statistik Pendidikan Tinggi 2017 : Kementerian Pendidikan Tinggi*. Putrajaya: Kementerian Pendidikan Tinggi.
- [67] *Graduating on time is Malaysia's target*. (2015, October 30).
- [68] Ray, S. (2017, September 13). *AnalyticsVidhya*. Retrieved from Understanding Support Vector Machine algorithm from examples.

Authors' Profiles



Nurafifah Binti Mohammad Suhaimi received her BSc. degree of Intelligent System in 2017 from University Teknologi Mara. Currently, she is a master's degree student of Information System (Intelligent System) at the Information Systems Department in the same university.



Shuzlina Abdul Rahman is an Associate Professor of Information System in Universiti Teknologi Mara. She received a master's degree in information technology in Universiti Utara Malaysia in 2000. In 2012, she received her PhD degree from the Universiti Kebangsaan Malaysia. She teaches courses related to the philosophy of artificial intelligence, data mining, advanced decision support systems, business intelligence and intelligent systems development. Her primary research interests involve the computational intelligence, data mining and optimization and intelligent data analytics.



Sofianita Mutalib is a senior lecturer of Information System in Universiti Teknologi Mara. She received a master's degree in Information Technology in Universiti Kebangsaan Malaysia in 1998. She teaches courses related to intelligent system development, decision support systems and

data mining. Her primary research interests involve the intelligent systems, data mining as well as machine learning



Nurzeatul Hamimah Abdul Hamid is a senior lecturer of Information System in Universiti Teknologi Mara. She received a master's degree in Intelligent Systems at the University of Sussex, UK in 2005. She teaches courses related to fundamentals of artificial intelligence, artificial intelligence programming paradigm and intelligent agent.

Her primary research interests involve the software agents, normative multi-agent systems, trust and reputation systems.



Ariff Md Ab Malik is a senior lecturer of Faculty of Business Management. He is currently acting as the Head of Strategic Planning (Information Management) of UiTM. He received his PHD from University Kebangsaan Malaysia in Computer Science. His primary research interest involves Information System Management, Heuristic, and Benchmarking.

How to cite this paper: Nurafifah Mohammad Suhaimi, Shuzlina Abdul-Rahman, Sofianita Mutalib, Nurzeatul Hamimah Abdul Hamid, Ariff Md Ab Malik, "Review on Predicting Students' Graduation Time Using Machine Learning Algorithms", International Journal of Modern Education and Computer Science(IJMECS), Vol.11, No.7, pp. 1-13, 2019.DOI: 10.5815/ijmeecs.2019.07.01