Data Mining for Student Advising

Hosam Alhakami[1], Tahani Alsubait[2], Abdullah Aljarallah[3]
College of Computer and Information Systems
Umm Al-Qura University, Makkah, Saudi Arabia

***Abstract*—This paper illustrates how to use data mining techniques to help in advising students and predicting their academic performance. Data mining is used to get previously unknown, hidden and perhaps vital knowledge from a large amount of data. It combines domain knowledge, advanced analytical skills, and a vast knowledge base to reveal hidden patterns and trends that are applicable in virtually any sector ranging from engineering to medicine, to business. However, it is possible for educational institutes to use data mining to find useful information from their databases. This is usually called Educational Data Mining (EDM). Advancing the field of EDM with new data analysis techniques and new machine learning algorithms is vital. Classification and clustering techniques will be used in this project to study and analyse student performance. The key importance of this project is that it discusses different data mining techniques in the literature review to study student behaviour depending upon their performance. We tried to identify the most suitable algorithms from the existing research methods to predict the success of students. Various data mining approaches were discussed and their results were evaluated. In this paper, the J48 algorithm was applied to the data set, gathered from Umm Al-Qura University in Makkah.**

Keywords*—Data mining; performance prediction; student analytics; academic advising; classification algorithms; decision tree; J48; neural network; Weka

I. Introduction

Educational institutions are very important for the socio-economic development of a country and students are the building blocks of any educational institute. They are very important for society because they are responsible for the future development. Depending upon the role of students in a country, their performance measure is an important aspect for any institution. The performance of a student is not only important to the institution, but it also affects the corporate sectors and job markets. High ranked universities produce great leaders in their particular domains because they put their resources to judge student's abilities and predict their field depending upon their performance. Several factors influence how students perform academically, and they include socio-economic factors as well as other environmental variables [1]. The impact of such factors can be better managed when people know about them and how they affect the performance of students. The tools, approaches, as well as the investigation that is designed to extract meaning from huge data sources of data that people learning activities generate automatically in an educational environment, is referred to as Educational Data Mining [2]. Investigations on educational mining have been focused on, to a large extent, in recent times. There is a special emphasis on the area of explaining and predicting academic performance. In fact, the big data contained in educational databases makes it more difficult to predict the performance of students. It is suggested that the model based on institutional internal databases and external open data sources performs better than the model based on only institutional internal databases [3]. Nevertheless, it is crucial for one to be able to predict the performance of students in an educational environment. Every academic institution has a long-term goal of ensuring that the success of students increases. The ability of an educational institution to predict the academic performance of students on time before their final examination makes it possible to put in additional efforts to make arrangements to help students that are not performing well and ensure they succeed. However, it is possible to aid improvement in courses by identifying the characteristics that influence the success rate of students in these courses. Investigators have the rare privilege of studying students' learning behaviours, as well as the methods that can help achieve success by using technology-based educational tools that are developed recently and by applying quality standards.

This paper aims to study the performance of students using different data mining techniques such as classification and clustering and provides a suitable technique that could be used by student advisors. This study will help the universities to improve the performance of the students. It introduces student marks prediction models using predictive model approaches based on student behaviour. We used data sources from Umm Al-Qura University that were used in practical settings to predict students' academic performance. Also, it focuses on identifying the variables that can be used to predict the future performance of students.

II. Background

The explosion of database management systems has led to a massive collection of every kind of information nowadays [4]. To date, the available information from military intelligence, text reports, satellite pictures, scientific data, and business transactions, is more than can be handled. It is no longer enough to retrieve information for decision-making [5]. There is currently the creation of new needs that will assist in making better managerial decisions. The needs include the discovery of patterns in raw data, extracting the significance of stored information and automatic summarization of data.

A. Data Mining Techniques

Data mining is also commonly referred to as Knowledge Discovery in Databases (KDD), referring to the knowledge discovery process, knowledge extraction, knowledge mining from data or data/pattern analysis [5]. It refers to the nontrivial extraction of implicit, previously unknown and possibly valuable information from data in databases [6]. Ramageri et

al. [7] explain that data mining can be described as a process through which useful information and patterns are extracted from big data.

According to Ramaraj et al. [8], classification refers to the task of generalising a known structure and applying it to a new one. It is possible for data mining classification techniques to process big data [9]. It helps to predict categorical class labels and categorises data by the training set and class labels, and it is also useful for categorising data that is available recently. Nikam [10] stated that the classification procedure is an established technique, which constantly makes those types of decisions in new situations. Decision Tree, Neural Networks, Naïve Bayesian Classification, Support Vector Machines, and K-Nearest Neighbour are common algorithms used to classify data.

1) A decision tree classifier: It is a classifier that uses the instance space's recursive partition. It is made of nodes that form a rooted tree, which means that it is a directed tree with a node referred to as roots, which have no incoming edges [8]. Intermediate nodes generate outgoing edges, and they test the nodes after performing gates. It consists of a decision tree that is generated according to instances. Nodes that do not consist of an outgoing branch are referred to as terminal or decision nodes [11]. Individual internal nodes split the instance space into two or more sub-spaces in a decision tree. The internal nodes, as well as the root, relate to features, while leaf nodes relate to classes [8]. All in all, there is an outgoing branch of individual non-leaf nodes, for every probable value of the characteristics that are related to the node. Sequential nodes are visited pending the time that a leaf node is reached, when using a decision tree to decide the class for a new instance, beginning with the root [8]. A test is applied at the root node and every internal node. The crisscrossed branch, as well as the next node visited, is determined by the result of the test.

2) Neural networks: It uses the gradient descent technique of the biological nervous system that has several interrelated processing elements. Such elements are referred to as neurons. The learned network's operability is improved using the rules that are extracted from the trained neural network [11]. Put differently; neural networks can be described as an emulation of the biological neural system [12]. It comprises an interconnected group of artificial neurons as well as processes information through a connectionist technique to computation. Generally, neural networks are adaptive systems in which internal or external information flowing through the network during the learning phase changes its structure [12].

3) Naïve bayesian classification: A Naïve Bayes classier refers to a simple probabilistic classier that functions with Bayes theorem (from Bayesian statistics) that has a strong (naïve) independence assumption.

Put differently; a naïve Bayes classifier assumes the presence or absence of a specific attribute of a class is not related to any other attribute's presence or absence [13]. It does not matter if these attributes rely on one another or the presence of other attributes. It is possible to train naïve Bayes classifiers efficiently in a supervised learning system, based on the probability model's exact nature. It is the technique of maximum probability that naïve Bayes models use an estimation parameter in several practical applications. This type of classifier has worked well in several real-world complex situations according to Irina et al. [14], even though it has naïve designs as well as over-simplified presumptions.

4) Support Vector Machines (SVM): The first person to introduce the modern method of classification called the support vector machine is Vapnik [15]. It is commonly used in bioinformatics because it is highly accurate, and it can handle high-dimensional data such as flexibility in modelling different data sources, as well as gene expression [16]. It belongs to the general group of kernel techniques [17]. The support vector machine (SVM) has been an effective technique for general pattern recognition, classification as well as regression. It can be said to be an excellent classifier as it has a high generalisation performance that does not require the addition of previous knowledge, regardless of whether it has an extremely high dimension of the input space. The SVM aims at looking for the best classification function that helps to differentiate between members of both classes in the training data [8]. It is possible to use geometrics to determine the best classification function. Regarding a dataset that can be separated linearly, a linear classification function matches a separating hyperplane f(x) which goes through the middle of both classes, separating them [11].

5) K-Nearest neighbour classifiers: This type of classifier is based on learning by analogy. It is the n-dimensional numeric characteristic that describes the training samples. Every sample signifies a point in n-dimensional space [8]. Along these lines, the entire training samples are stored in n-dimensional pattern space. The k-nearest neighbour classifier looks for the k training samples' training samples which are closest to the unknown sample when it is given an unknown sample. It is the Euclidean distance that determines the closeness. Nearest neighbour classifiers give every attribute equal weight.

Also, it is possible to use it for prediction; in other words, it can be used to return a real-valued prediction for a given sample that is unknown. The algorithm of k-nearest neighbours is among the easiest machine learning algorithms. It is when there is a majority vote of the neighbours of an object [8]. It is crucial to choose an appropriate I value when using a k-nearest neighbour algorithm [8].

B. Analytical Tools Weka

The formatting of datasets in Weka should be in the ARFF or CSV format. These would be used automatically in Weka Explorer when a particular file is not recognised. Data can be imported from a database using the facilities contained in the pre-process panel; this data utilises a filtering algorithm when it comes to pre-processing. It is possible to convert this data using the filters, which allows the deletion of cases and assigns based on specified principles.

WEKA tool's Classification algorithm is used to carry out experiments on the data set. The first step involves searching for an aggregate number of cases of the particular data using the j48 classification algorithm as well as Naïve Bayes. The following step requires the experiment to conduct the cost analysis as well as find out the correctness of the Classification.

III. Related Work

Johnson (2018) [18] stated that there is a high demand for enrolments in computer science, and the graduations of successful computer science undergraduates are, without a doubt, very significant. He, therefore, proposed building upon the current data mining and modelling, learning analytics, and machine learning applications for predicting the success of students that goes beyond retention. To Khare et al. (2018) [19], educational data mining (EDM) refers to an applied field of research, which combines data mining, statistics, and machine learning in the field of education, but not restricted to MOOCs, intelligent tutoring systems, universities and schools. They decided to explore the essence of data mining in the online education setting and discover its ways of improving the learning experience of the student. They intend to achieve their objective by reviewing some of the basic data mining algorithms used in education and the innovations that will come up in the future. The proposal presented by Brooks (2013) [20] states that EDM community research's distinction emanates from intelligent tutoring systems, which is the interaction between the student, domain material and the system, whereas, the focus of the learning analytics researchers is on enterprise learning system such as classroom management systems, which pile up data from all the courses.

However, according to Thomas and Gelan [21], the definition of learning analytics is that they collect, measure, evaluate and report data regarding the learners in their context to understand and optimise learning and learning environments. They believe that learning analytics presents some potential opportunities because, through it, teachers can obtain valuable data on learners that are succeeding and failing. On the other hand [22] stated that students that take transfer in community colleges face challenges in their pursuit of bachelor's degrees, and this usually leads to credit loss. To them, this may consequently reduce the students' chances of completing their credential, and increase the costs and time for the students, their families, and taxpayers.

According to Lacefield [20], accountability appears to be permanently rooted in the environment of K-12, like the expectation of delivering quality education to children of school and adolescents. However, this expectation has failed repeatedly and has drawn the attention of policymakers and the public to the drawbacks of major accountability systems. Therefore, they tried to show how to use the predictive analytics applied to school student system (SIS) records, to make advising of students and activities such as mentoring or coaching at-risk students easier.

The argument of Dash and Vaidhehi [23] stated that academic advising requires much time, responsibility and skills. They believe that it is necessary for the computerised advising system to be forthcoming so that human advisors can be assisted effectively. Additionally, Olaniyi [24] said that Educational Data Mining (EDM) focuses on developing and modelling techniques that discover knowledge from data emanating from educational environments. Moreover, based on the perception of Pal, and Chaurasia [25], the consumption of alcohol in higher education institutions is not new; India's legal drinking age is 18 years, but it is dangerous for underage students and those that are 18 years and above to drink heavily. Therefore, four popular data mining algorithms; REP Tree, Bagging, Sequential Minimal Optimisation (SMO) and Decision Table (DT), obtained from a rule-based classifier or a decision tree to improve academic performance's efficiency in the educational institutions for alcohol-consuming students were discussed.

El-Halees and Abu-Zaid [26] stated that presently, websites are perceived as a vital tool in several real-life applications including, entertainment, industry, education, and business, and this has brought many concerns regarding the quality of these websites. Nevertheless, Hussain [27] pointed out that Google has one million search queries every one minute on the internet; more than two million emails are sent, there are 100,000 tweets, thousands of photos are uploaded, and much more traffic. To them, terabytes of data, which has a grade value that can shape higher education institutions generate the future of nations. Moreover, Yadav et al. (2012) [27] believe that Knowledge Discovery and Data Mining (KDD) is a multifaceted discipline that focuses on the methods used to extract valuable knowledge from data. Education's quality can be increased using this knowledge.

According to Tair and El-Halees [28], educational mining focuses on developing methods of knowledge discovery using data collected from the field of education. To Baepler and Murdoch [29], the areas of academic analytics and educational data mining have experienced rapid development, and the outcome has resulted in new potentials for collecting, analysing and presenting student data.

Weakley et al. [30] said there are divergent views that the use of predictive measures violates a professional ethical principle to develop a comprehensive understanding of their advice, according to professional advisers at the Public Higher Education Institute, which may contribute to the enrichment of content in the exploration of Educational data mining from a social and ethical perspective.

IV. Methodology and Research Design

A. Procedure

This project involves some major steps, including:

- Step 1: Collect student-related information from the university for at least five years for the complete information regarding the dataset used.
- Step 2: Clean and verify the student information collected from Step 1. Microsoft Excel functionalities such as data filtering, data sorting, and so on, would be used to verify, validate and clean the data manually.
- Step 3: To allow us to study the student's performance, once the data is cleaned, verified and categorised. Step

1 to Step 3 will be repeated for the entire datasets received from the university.

- Step 4: To use the Weka data mining tool to perform the classifications, to analyse the appropriate algorithm.
- Step 5: Analyse the performance of the algorithms from the results received from step 4.
- Step 6: To evaluate the recommended classification method in this project as well as the prediction results with experienced lecturers. This method is key to ensure that the suggestions we provided are appropriate for real-life usage.

B. Classification Algorithms

The following classification algorithms were compared according to their performance on the dataset:

Decision tree (J48): It is a predictive machine-learning model whose function is to determine the new sample's target value using various attribute values of the available data. The decision tree's internal nodes represent the various attributes. The branches between the nodes indicate these attributes' possible values in the experimental samples, and the final value of the dependent variable is indicated by the terminal nodes [31]. J48 algorithm has been used to generate decision tree using the Weka tool.

Naive Bayes: This type of classifier is based on the rule of Bayes that expresses an event's possibility before there is a clear proof, as well as an event's possibility after the proof becomes apparent [32]. There are several reasons for using this classifier, and they include building the easiest way without requiring any complicated iterative parameter evaluation schemes.

C. Performance Measurement Factors

A brief discussion of the performance measurement factors is provided below:

TP Ratio: TP stands for True positive. The number of dataset rows that are predicted as positive that are truly positive is known as TP ratio.

FP Ratio: FP stands for False positive. The number of dataset rows that are predicted as positive that are truly negative is known as FP ratio.

Accuracy: This is measured based on the ratio of the correct observation over the particular dataset's overall observations.

Precession: This is measured based on the ratio of the positive observation that has been predicted accurately over the overall positive predicted observations.

Recall: It is measured based on the ratio of the positive observation that has been predicted over all the observations in actual yes class.

F-measure: This refers to an average between the recall and accuracy.

Classification Matrix: This refers to a table whose primary purpose is to represent the performance of a particular classification model that can be any algorithm.

D. Dataset

The dataset was gathered from the information of graduate students from Umm Al-Qura University in the last 5 years in Makkah. 26711 student records were used for training and 11960 students for testing. The total number of students is 38671. Some records with incomplete data were discarded, the total number of gathered records is 59699.

The data fields in the dataset are provided below:

- NATIONALITY: This is the field column in the dataset which provides the student nationality is where they are a legal citizen.
- SCHOOL GOVERNATE: This is the field column in the dataset which provides the student school governate is an administrative division of a country.
- GRADE: This is the field column in the dataset which provides the student Final student grade at the university (Excellent, Very Good, Good, Pass).
- TAHSILI MARK: This is the field column in the dataset which provides the student mark in Tahsili exam (i.e., standardised national exam for student's academic performance).
- QDRAT_MARK: This is the field column in the dataset which provides the student mark in Qdrat exam (i.e., standardised national exam for student's skills measurement).
- CAMPUS: This is the field column in the dataset which provides the grounds and buildings of a university, college.
- GENDER: This is the field column in the dataset which provides the student state of being male or female.
- SCHOOL AVERAGE: This is the field column in the dataset which provides the student school average.
- SCHOOL BRANCH: This is the field column in the dataset which provides the student internal specialization in school. Many of these specializations are special to Makkah district such as Dar Al Tawheed, Literary Section, Memorization of the Koran, scientific department, Dar Al Hadith, scientific department Courses, Commercial Secondary, Literary Section, Al - Haram Institute, Literary Section Courses, Noor Institutes, Holy Quran Institute for National Guard, Health Institute, Visual impairment, Teachers Training Institute, Secondary Teacher Institutes, Secondary professional, Secondary Alsolatyh, Scientific Institute.
- FACULTY: This is the field column in the dataset which provides the student colleges at universities.
- AGE: This is the field column in the dataset which provides the student age.

V. Results and Discussion

A. Weka Analysis

The Weka result for decision tree (J48) algorithm is shown in Fig. 1.

As shown in Fig. 1 for J48, the percentage of correctly classified instances is 84.38%, while the percentage of the incorrectly classified instances is 15.61%. On the other hand, as shown in Fig. 2 for Naive Bayes algorithm, the percentage of correctly and incorrectly classified instances are 46.68% and 53.31%, respectively. In this scientific paper, the logarithm of J48 is chosen compared to the logarithm of Naive Bayes, due to the fact that it is correctly classified as the highest in the algorithm of J48.

B. Analysis of the Factors that affect Students' Performance

The basic dataset analysis was performed using pivot table functionality and the basic statistical functions of the MS Excel.

Fig. 3 shows that the percentage of female is the highest in the excellent score (78%) and the score is very good for 65% females, while male students are 22% excellent and 35% very good.

Table I contains four elements (PASS, GOOD, VERY GOOD and EXCELLENT) to describe the student's performance in the final rate (GPA) based on gender.

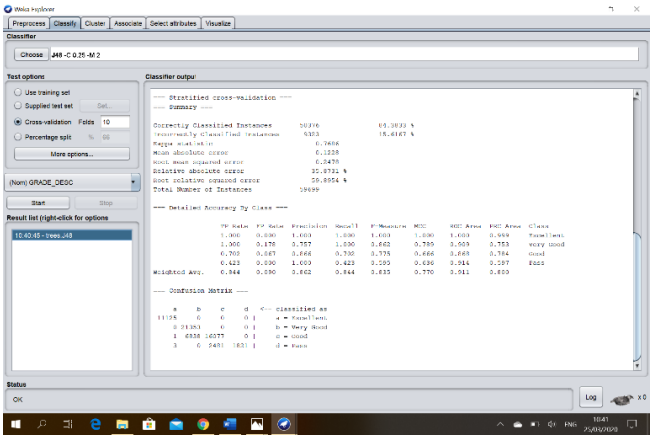

Fig. 1. Use Training Set to Learn the Algorithm J48.

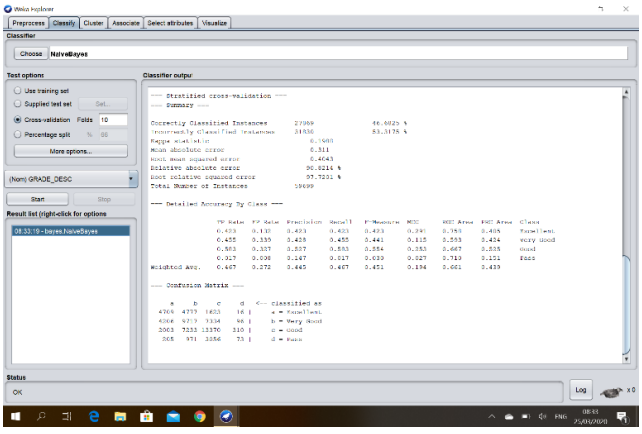

Fig. 2. Use Training Set to Learn the Algorithm Naïve Bayes.

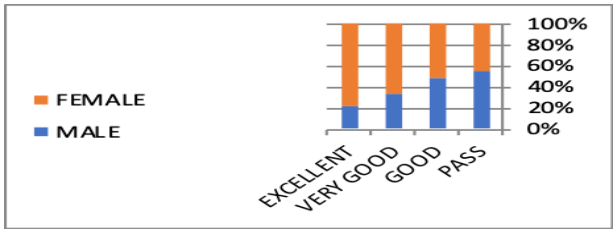

Fig. 3. The Grade for all Student based on Gender.

TABLE I. The Grade for All Student based on Gender

GENDER_DESC	*EXCELLENT*	*VERY GOOD*	*GOOD*	*PASS*
MALE	2458	7245	11085	2341
FEMALE	8667	14108	11831	1964

Depending on the percentage, we can focus on the development and training process on male students and motivate them through the academic supervisor.

Fig. 4 indicates that the percentage of Jamoum and Laith students is the highest in the good grade of (39%). The percentage of Business and IT is the highest in the very good grade of 40%. The percentage of Islamic is the highest in the good grade of 45%. The percentage of Education is the highest in the very good grade of 48.5%. The percentage in pass grade is 1%. The percentage of Medical is the highest in the very good grade of (49%)

Students studying in Laith and Jaumum colleges are the lowest in the percentage of the university average, since they have studied high school in Makkah, suggesting that we should focus on the development and training process of students and motivate them through academic supervision compared to other colleges.

Fig. 5 indicates that the percentage of Memorization of the Koran student is the highest in the excellent grade of 46%. The percentage of the Literary Section is the highest in the good grade of 44%. The percentage of the scientific department is the highest in the very good grade of 37%. The percentage of scientific department Courses is the highest in a very good grade of 42%. The percentage of Literary Section Courses is the highest in the very good grade of 41%. The percentage of Scientific Institute is the highest in the good grade of 42%. The percentage of other departments is the highest in the excellent grade of 32%.

The percentage shows that we can focus on the development and training process on the student who studied in high school in the literary section or scientific department and motivate them through the academic supervisor.

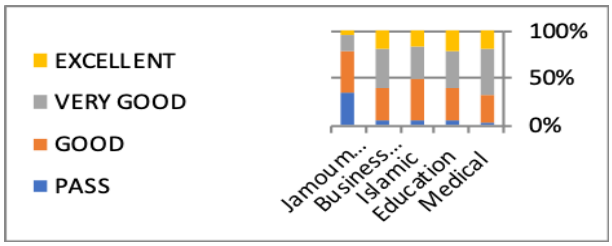

Fig. 4. The Grade based on College for Students who Study High School in Makkah.

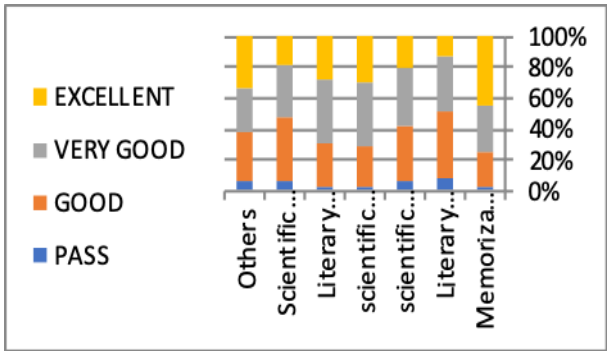

Fig. 5. The grade for student based on department in the high school who study high school in Makkah.

Fig. 6 indicates that the percentage of the Tahsili exam (100-95%) is the highest in the excellent grade of 88% and the grade of very good is 12%. The percentage of the Tahsili exam (94-90%) is the highest in the excellent grade of 60%. The percentage of the Tahsili exam (89-80%) is the highest in the excellent grade of 46% and the grade of very good is 34%. The percentage of the Tahsili exam (79-70%) is the highest in the very good grade of 42% and the grade of good is 28%.

Students who have a rate of (38-54%) and a rate of (69-55%) in the Tahsili exam are the lowest in the percentage in the university average, suggesting that we should focus on the process of developing and training students and motivate them through the academic supervisor, at the beginning of studying at the university.

Fig. 7 indicates that the percentage of the Qdrat exam (100-95%) is the highest in the excellent grade of 69%. The percentage of the Qdrat exam (94-90%) is the highest in the excellent grade of 55%. The percentage of the Qdrat exam (89-80%) is the highest in the excellent grade of 37%. The percentage of the Qdrat exam (79-70%) is the highest in the very good grade of 42%.

Understudies who have a percentage of (42-54%) and (69-55%) in the Qdrat test are the least in the rate in the GPA, we can help them during the time spent creating and preparing understudies and spur them through the academic supervisor, toward the start of learning at the college.

Fig. 8 indicates that the percentage of student age (16-22 y) is the highest in the excellent grade of 13% and the grade of very good is 11%. The percentage of student age (23-29 y) is the highest in the excellent grade of 78%. The percentage of student age (30-39 y) is the highest in the passing grade of 21%. The percentage of student age (40-54 y) is the highest in a very good grade of 3%.

The largest percentage of students is confined to the age group (23-29 y), and we can focus on students with the age group (30-39 y) as the lower percentage increases in it, and this helps the academic supervisor to guide students and increase training and development for them.

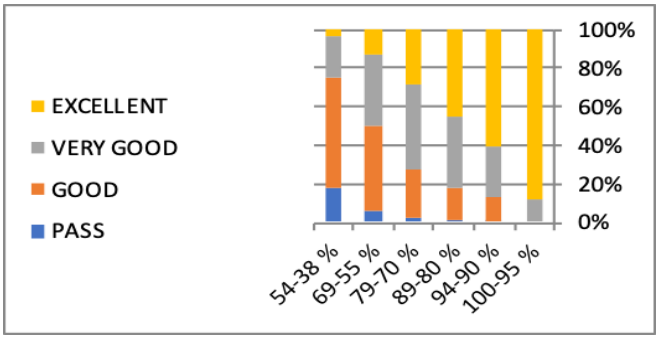

Fig. 6. The grade for all student who take the Tahsili exam.

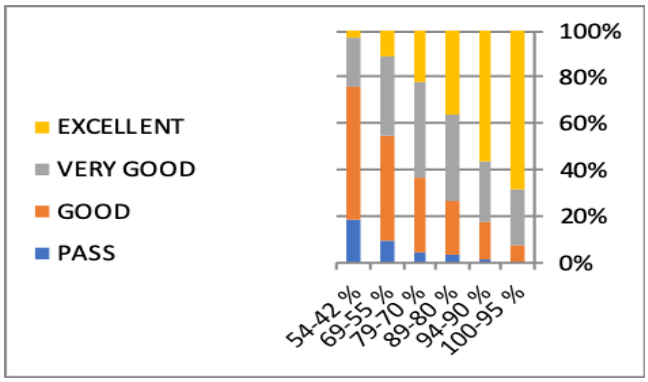

Fig. 7. The grade for all student who take the Qdrat exam.

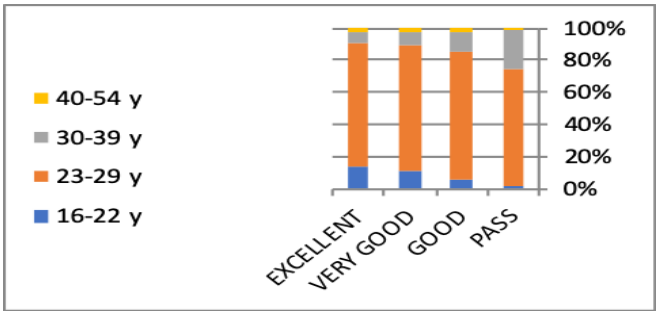

Fig. 8. The grade for all student based on age.

VI. Conclusion and Future Work

In conclusion, the decision tree (J48) classification algorithm was used to analyse student performance. Moreover, the project introduced a model that can predict the college and the GPA of a specific student by analysing the exams and other features. The accuracy of the prediction is high because it has identified that GPA and extra tests are not the only factors that affect the final results of the student. This project has identified additional factors that can be influencing student performance which are School, Sex, Age, Nationality, and City. Additional factors that this project has identified as factors influencing the student performance, as would be expected, are High School Ratio, Qdrat exam, Tahsili Exam, and department in high school.

REFERENCES

[1] Surjeet Kumar Yadav, Brijesh Bharadwaj, and Saurabh Pal. “Data mining applications: A comparative study for predicting student’s performance”. In: arXiv preprint arXiv:1202.4815 (2012).

[2] P Nithya, B Umamaheswari, and A Umadevi. “A survey on educational data mining in field of education”. In: International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) 5.1 (2016), pp. 69–78.

[3] Farhana Sarker, Thanassis Tiropanis, and Hugh C Davis. “Students’ performance prediction by using institutional internal and external open data sources”. Accessed from: http://eprints.soton.ac.uk/353532/1 /Students'%20mark%20prediction%20model.pdf (2013).

[4] Aliyar, F. Database Needs and Data Mining. International Research Journal of Management Science and Technology.2010.

[5] Atul Goyal, Dinesh Kumar, and Ms Sunita Baniwal. “Performance Analysis Of K-Mean Algorithm Using Markov Chain Lloyd”. Internatinal Journal of Research Review in Engineering Science & Technology, 4 (1) (2015), pp. 7–11.

[6] Osmar R Za¨ıane. “Principles of knowledge discovery in databases”. In: Department of Computing Science, University of Alberta 20 (1999).

[7] Ramageri, B. Data Mining Techniques And Applications . Indian Journal of Computer Science and Engineering, 1(4), pp.301-305.2010.

[8] Neelamegam and E Ramaraj.“Classification algorithm in data mining: An overview”. In: International Journal of P2P Network Trends and Technology (IJPTT) 4.8 (2013), pp. 369–374.

[9] Sahani, R., Rout, C., Badajena, J.C., Jena, A.K. and Das, H., Classification of Intrusion Detection Using Data Mining Techniques. In Progress in Computing, Analytics and Networking (pp. 753-764). Springer, Singapore.2018.

[10] Nikam, S.S., A comparative study of classification techniques in data mining algorithms. Oriental Journal of Computer Science and Technology, 8(1), pp.13-19.2015.

[11] Gorade, M., Deo, P. and Purohit, P. A Study of Some Data Mining Classification Techniques. International Research Journal of Engineering and Technology(IRJET), 4(4).2017.

[12] Zhang, G. Neural Networks For Data Mining. Data Mining and Knowledge Discovery Handbook, pp.419-444. 2009.

[13] Flach, P. and Lachiche, N. Naive Bayesian Classification of Structured Data. Machine Learning, 57(3), pp.233-269.2004.

[14] Irina, R. An empirical study of the naive Bayes classifier. IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence.. [online] Available at: http://www.research.ibm.com/people/r/rish [Accessed 2 Mar. 2019].(2001).

[15] Boser, B., Guyon, I. and Vapnik, V. A training algorithm for optimal margin classifiers. Pittsburgh, PA.: ACM Press, pp.144–152.1992.

[16] Sch olkopf, B., Tsuda, K. and Vert, J .Kernel Methods in Computational Biology. MIT press. 2004.

[17] Scholkopf, B. and Smola, A. 2002 Learning with Kernels. Cambridge, MA: MIT Press. 2002.

[18] Johnson, W.G., Data Mining and Machine Learning in Education with Focus in Undergraduate CS Student Success. In Proceedings of the 2018 ACM Conference on International Computing Education Research (pp. 270-271). ACM. 2018.

[19] Khare, K., Lam, H. and Khare, A., Educational Data Mining (EDM): Researching Impact on Online Business Education. In On the Line (pp. 37-53). Springer, Cham.2018.

[20] Warren E Lacefield and E Brooks Applegate.“Data Visualization in Public Education: Longitudinal Student-, Intervention-, School-, and District Level Performance Modeling.” In: Online Submission (2018).

[21] Michael Thomas and Anouk Gelan. Special edition on language learning and learning analytics. 2018.

[22] John Fink et al. “Using data mining to explore why community college transfer studentsearn bachelor’s degrees with excess credits” .In:(2018).

[23] Poulami Dash and V. Vaidhehi. “Enhanced Elective Subject Selection for ICSE School Students using Machine Learning Algorithms”. In: Indian Journal of Science and Technology 10.21 (2017). issn: 0974 - 5645. url: http://www.indjst.org/index.php/indjst/article/view/109551.

[24] Olaniyi, A.S., Kayode, S.Y., Abiola, H.M., Tosin, S.I.T. and Babatunde, A.N., Student’s Performance Analysis Using Decision Tree Algorithms. Annals. Computer Science Series, 15(1).2017.

[25] Saurabh Pal and Vikas Chaurasia. “Performance Analysis of Students Consuming Alcohol Using Data Mining Techniques”. In: International Journal of Advance Research in Science and Engineering 6.2 (2017), pp.238– 250.

[26] Alaa M El-Halees and Ibrahim M Abu-Zaid .“Automated Usability Evaluation on University Websites using Data Mining Methods”. In: AutomatedUsabilityEvaluationonUniversityWebsitesusingDataMiningMethods 6.11 (2017).

[27] Mohammed Hussain et al. “Mining educational data for academic accreditation: aligning assessment with outcomes”. In: Global Journal of Flexible Systems Management 18.1 (2017), pp. 51–60.

[28] Mohammed M Abu Tair and Alaa M El-Halees. “Mining educational data to improve students’ performance: a case study”. In: Mining educational data to improve students’ performance: a case study 2.2 (2012).

[29] Paul Baepler and Cynthia James Murdoch.“Academic analytics and data mining in higher education”. In: International Journal for the Scholarship of Teaching and Learning 4.2 (2010), p. 17.

[30] Weakley, Jonathon JS, et al. "Visual Feedback Attenuates Mean Concentric Barbell Velocity Loss and Improves Motivation, Competitiveness, and Perceived Workload in Male Adolescent Athletes." The Journal of Strength & Conditioning Research 33.9 (2019): 2420-2425.

[31] Sewaiwar, Purva and Kamal Kant Verma. “Comparative Study of Various Decision Tree Classification Algorithm Using WEKA.” (2015).

[32] Yoan Martínez López, Julio Madera and Ireimis Leguen de Varona. “Study Of The Performance Of The K* Algorithm In International Databases.” (2016).