Indonesian Journal of Electrical Engineering and Computer Science
Vol. 32, No. 1, October 2023, pp. 363~371
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v32.i1.pp363-371

The prediction of student's academic performance using RapidMiner

Muhammad Firdaus Mustapha, Alia Nur Izzah Zulkifli, Omar Kairan, Nur Nabila Sofea Mat Zizi, Nur Naim Yahya, Nur Maisarah Mohamad
College of Computing, Informatics and Media, Universiti Teknologi MARA, Machang, Malaysia

Article Info

Article history:

Received Sep 13, 2021
Revised Jun 4, 2023
Accepted Jun 17, 2023

Keywords:

CRISP-DM
Data analytics
Data visualization
Decision tree
RapidMiner
Tableau

ABSTRACT

Students' performance analysis basically consists of determining the factors influencing the performance and how it will give impact towards success. It will help us to understand students' behavior and how to improve their academic performance. The efficiency of this analysis depends on the information given by the user through learning management system (LMS). In order to improve the information, we have applied algorithms on the dataset and prepared a model by using Tableau and RapidMiner. Cross-validation with decision tree also has been applied on datasets. This can help in evaluating statistical computational results into a generalized data set. Based on the calculation of data mining, it can analyze that our model is quite stable since it has high accuracy with lower standard deviation. So, the processes like testing and validation, applying the model and decision tree on RapidMiner generates the output in a specific form. The result shows that the percentage of students who are absence is better than students who are absence more than 7 days. At last, a model is prepared, and it can help the schools, students, and the parents in adapting appropriate measures to ensure the success of students at school.

Corresponding Author:

Muhammad Firdaus Mustapha
College of Computing, Informatics and Media, Universiti Teknologi MARA
18500 Machang, Kelantan, Malaysia
Email: mdfirdaus@uitm.edu.my

1. INTRODUCTION

The amount of data on the internet is increasing, making it more challenging for users to get up-to-date and useful information. Finding data to meet the user's information needs is a key issue in the search for information. For this research, we have collected 480 different data of students' record. RapidMiner [1], [2] is used to extract data, which is then subject to additional analysis. Some of the advantages of RapidMiner are supports all computer environments and has a variety of data visualization output such as 3D graphs, maps, and scattered matrices. The data consists of 17 attributes and these attributes are then being analyzed and transformed into decision tree for further process of testing and validation.

Academic performance is important to produce more skilled student so that the job performance will be better [3], [4]. In order to understand student behaviour and capabilities, research in students' performance can produce a better prediction in the future. Accuracy gives ability to predict the class label correctly. It allows forecasters the ability to improve their predictions based on the data provided [5], [6]. The label for this research is students' absence days. Other than that, classification techniques like cross-validation [7], [8] with decision tree [9], [10] has been applied on datasets. It contains various types of information related to the students that might cause an impact towards students' performance. Incorporating cross-validation into a common data set can aid in the evaluation of the output of statistical calculations. This type of approach is used when the main objective of the model is prediction, and its accuracy needs to be determined.

Educational achievement is essential for an institution to develop high quality outcomes that will ultimately translate into future workplace success [11]. However, academic performance is not effected by age, gender, nationality, and birth of place but with their own studies [12]. For example, how many students raise their hand in the class, involves in discussion, visit resources, viewing announcement, and also number of students' absence from school. In fact, the efforts of the students themselves will affect their academic success [13]. Moreover, poor study habits will cause the delaying in study [14] and that will affect the cumulative grade points average (CGPA) of students. All the datasets can be used in RapidMiner [15] as a tool to test the accuracy of the datasets which is about factors that can influence the prediction of students' academic performance.

Thus, this research aims to determine the factors influencing students' performance so that the result obtained will be used to improve students' performance at school. By using RapidMiner, analysis report and a prediction model will be easily delivered. The decision tree in the prediction model will help in determining whether student success is influenced by their attendance at school or by their participation and effort in class. The analysis report will summarize the data and the prediction of factors influencing students' performance will be evaluated.

2. METHOD

Data mining processes based on cross-industry-standard process data mining (CRISP-DM) [16] is a method for data mining and analytics project. It provides a structured approach to guide the entire data mining process, from understanding the business problem to deploying the final solution. Some of the advantages of CRISP-DM are highly efficient [17] and make the result of data mining available more accurately and quickly [18]. It is often adapted to accommodate domain-specific requirements [19]. It has six phases which are business understanding, data understanding, data preparation, modelling, evaluation and deployment [17], [18], [20].

2.1. Business understanding

The first phase in the procedure is business understanding, where we determine the issue or query to be resolved. From this data, we want to identify the factors and the extent to which they affect a student's academic success. Imagine for example, the parents or the school really want to know what the appropriate measures are to ensure the success of students at school could be. In more accurate, we try to answer some of the business questions:

- Is the student's absence days affect his/her performance at the school?
- Is the student's class related to performance at the school?
- Can student's academic success be predicted based in its attribute with reasonable accuracy?

2.2. Data understanding

The second phase has to do with the information that will be used to help solve business challenges. The project that we are working on is using an educational data set which is collected from learning management system (LMS) that called as Kalboard 360 [21], [22]. Kalboard 360 is a multi-agent LMS created with the goal of facilitating learning through the use of advanced technologies. It is collected using the experience API (xAPI), a mechanism for tracking student activity. The training and learning architecture (TLA) enables learning progress and learning activities such as watching training videos or reading articles. The learner, activities and objects that define the learning experience can all be found from the results of the experience API. In this dataset, it involves of 480 of students records which is 305 of male students and 175 of female students. They have various backgrounds, with 179 students coming from Kuwait, 172 from Jordan, 28 from Palestine, 22 from Iraq, 17 from Lebanon, 12 from Tunis, 11 from Saudi Arabia, 9 from Egypt, 7 from Syria, 6 from the USA, Iran and Libya, 4 from Morocco, and only one from Venezuela.

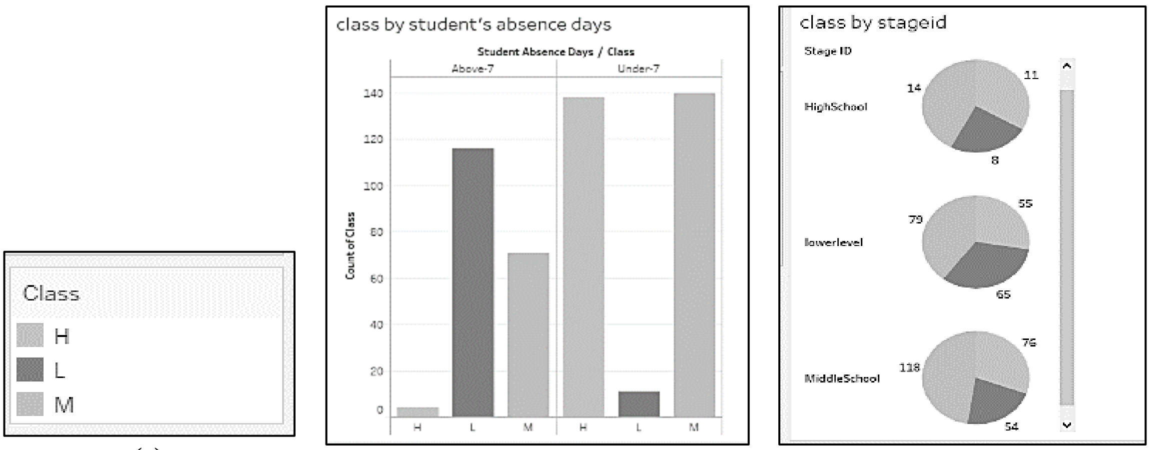
There are 17 aspects involved in this dataset as well as there are no missing values as shown in Table 1. Moreover, we used multivariate data characteristic, and all those attributes are either integer or categorical which is aligned with this project. The features are divided into three main groups. The first is demographic, gender or nationality characteristics. Next are the characteristics of the academic background, such as grade level, education level, and section. Students are divided into three numerical intervals based on their overall grade or mark which is low-level for students who score from 0 to 69, middle-level for students who score somewhere from 70 to 89 and high-level for most successful students who score more than 89. Lastly, the attribute based on behavioral characteristics including raising hands in class, visited resources, answering to parent surveys and parent school satisfaction. Additionally, the data set includes information about student attendance, with students divided into two groups based on their absence days: 191 students were absent for more than 7 days, while 289 students were absent for less than 7 days.

Table 1. Raw data

Attributes	Datatype	Description
Gender	string	Student's gender - ('Male' or 'Female')
Nationality	string	Student's nationality - ('Kuwait', 'Lebanon',' Egypt',' SaudiArabia',' USA',' Jordan', 'Venezuela',' Iran',' Tunis',' Morocco',' Syria',' Palestine',' Iraq',' Libya')
Place of birth	string	Student's place of birth - ('Kuwait', 'Lebanon',' Egypt',' SaudiArabia',' USA',' Jordan', 'Venezuela',' Iran',' Tunis',' Morocco',' Syria',' Palestine',' Iraq',' Libya')
Stage Id	string	Educational level that student belongs to - ('lowerlevel','MiddleSchool', 'HighSchool')
Grade Id	string	The grade that student is enrolled in - ('G-01', 'G-02', 'G-03','G-04','G-05','G-06','G-07','G-08','G-09','G-10','G-11','G-12 ')
Section Id	string	Classroom the student belongs to - ('A','B','C')
Topic	string	Course topic - ('English',' Spanish', 'French',' Arabic',' IT',' Math',' Chemistry', 'Biology', 'Science',' History',' Quran',' Geology')
Semester	string	School year semester - ('First',' Second')
Relation	string	Parent responsible for the student - ('mom', 'father')
Raised hands	integer	Number of times the student has raised hands in the classroom - (0-100)
Visited resources	integer	Number of times the student visited the course content - (0-100)
Viewing announcements	integer	Number of times the student checked the announcements - (0-100)
Discussion	integer	Number of times the student participated in group discussions - (0-100)
Parent answering survey	string	Did the parent answer the survey which are provided from school - ('Yes', 'No')
Parent school satisfaction	string	The degree of parent satisfaction from school - ('Good', 'Bad')
Student absence days	integer	Number of days a student was absent - ('above-7',' under-7')
Class	string	Indicator of the student's performance – ('L','M','H')

Next, we try to analyze the given information by using data visualization. The graphical presentation of information and data is referred to as data visualization. We chose Tableau [23] to interpret our data in visual elements such as bar graphs, staked bar graph and pie chart because Tableau has a very convenient feature when learning to create visualizations [24]. It can suggest a visualization based on the types of fields that we select. The main idea why we decide to create data visualization is to keep our eyes on the data or message. We quickly see trends and outliers based on a chart as depicted in Figure 1. In addition, from visualization it can tell us a story, removing the noise from data and the most important one is, highlighting the useful information.

From the Figure 1(a) orange color refer to low-level for students who score from 0 to 69, green color refers to middle-level for students who score somewhere from 70 to 89 and cream color refer to high-level for most successful students who score more than 89. Figure 1(b) shows students who are not absent from school under seven days will be more encouraged to be a successful student who scores more than 89 marks. From the pie chart in Figure 1(c), study grades do not tend to affect student's performance. In this dataset comprise of three grades which are high school, middle school and lower level. Middle-level performing students make up the majority at all three levels.



(a) (b) (c)

Figure 1. Data visualization for (a) performance-based student classes, (b) performance-based student attendance, and (c) breakdown of students according to study grade

Figure 2 shows the students involvement in class which is they raised hand to participate. According to the figures, more students from middle school grade are getting involved in class activities. Normally active students will perform better in academic results. At the same time, it proves that students who are active in class can also contribute to good performance [25] as shown by the high school grade group in Figure 1(c).

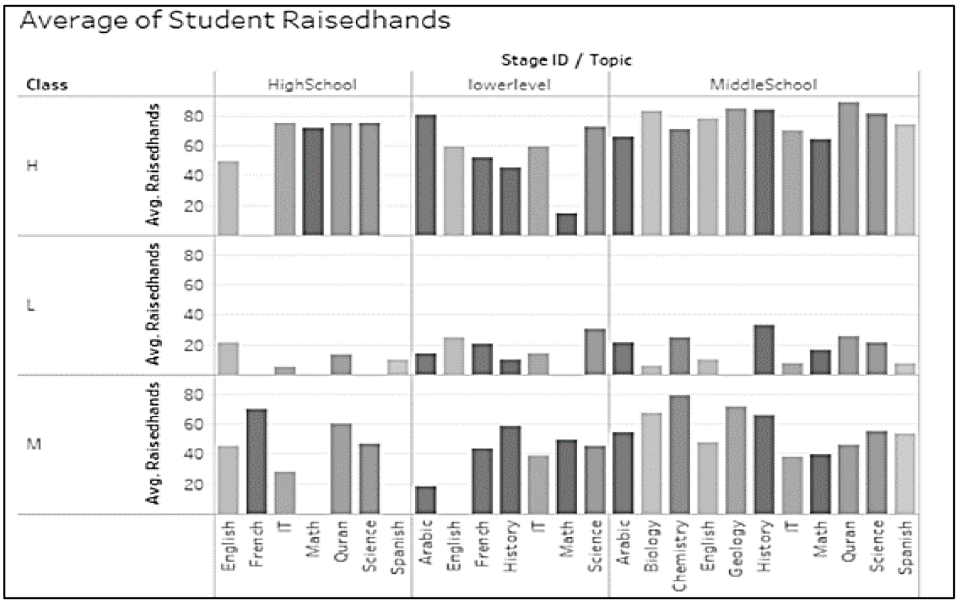


Figure 2. Students raised hands on every topic affecting student's class

2.3. Data preparation

In this step, the data will be prepared for the next step which is data modelling. Majority of the data is not in the desired format; therefore, conversion is required to achieve a format that can help in producing better results such as deletions, missing values, data cleaning and redundancy. But for our data, there are no missing values which means the dataset is clean. Hence, data cleaning is not necessary. We do a renovation for our data where we remove a few attributes which are not related to our model such as "Nationality" "PalaceOfBirth", "StageID", "GradeID" and more. We remove those unrelated attributes because they have no significant correlation to the model built. Plus, the deletion process for unnecessary attributes also can increase the accuracy of performance. Therefore, the attributes that will be built into the model consist of only eight attributes as shown in Figure 3.

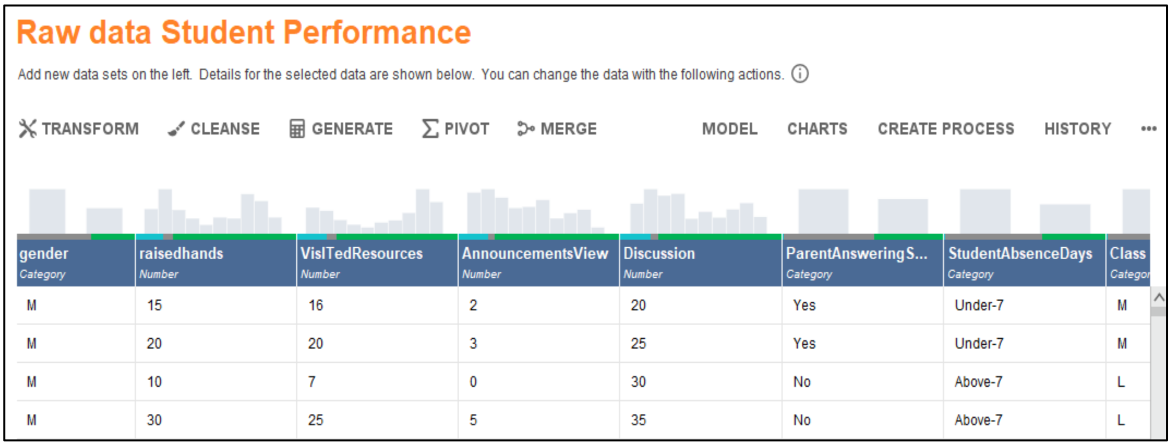


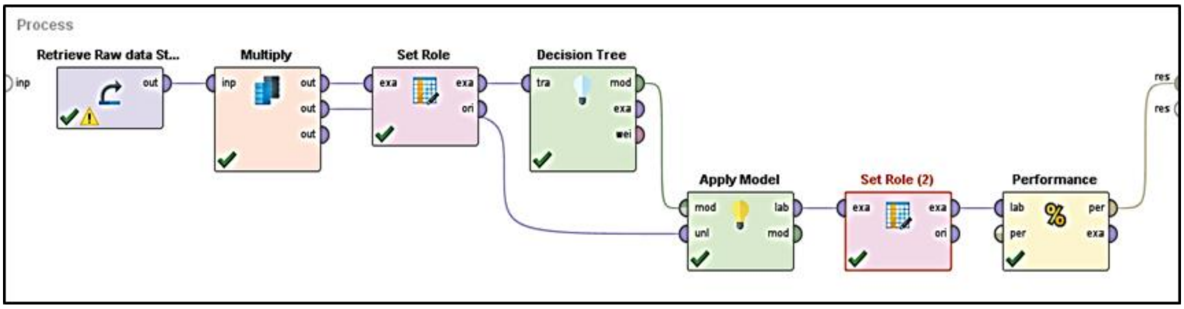
Figure 3. Eight attributes used for model

2.4. Modeling

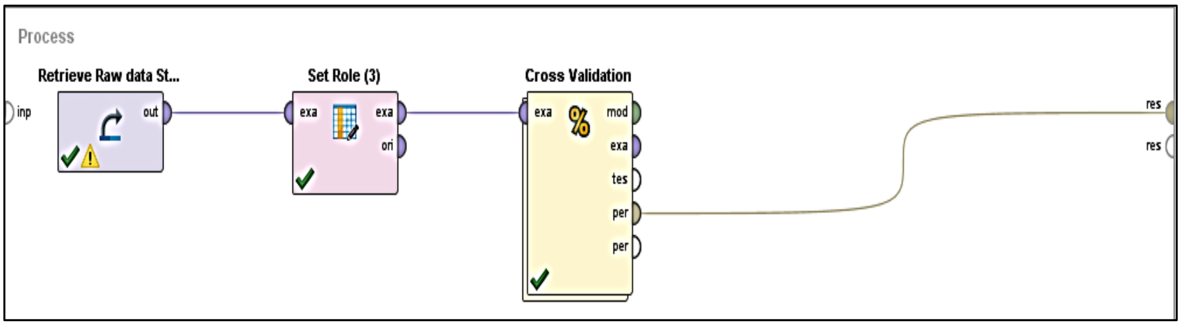
This is an important phase where we try to discover a model that accurately reflects the data. The main application of data mining techniques to data is modeling. To produce a model for predicting factors and the extent to which they affect a student's academic success, we use the algorithm of classification modeling technique to generate data design with the build model. Firstly, we create a decision tree. We then entered the operator set role after importing the updated student performance dataset, which was reviewed and corrected. We chose the attribute “StudentAbsencesDay” as our label. Lastly, we place an operator tree diagram to generate a tree diagram for our dataset. We do not have to use operator filter examples in our process to clean all the errors and the missing value since our dataset has no missing values. The process was done earlier to generate a tree like diagram.

Based on Figure 4, we use the operator multiply where it created the copy of data to generate the applying model. In this step, we generate the applying model to predict the accuracy of our dataset. The operator set role and decision tree also needed in this process. Then we insert an operator apply model to make it complete as shown in Figure 4(a). Lastly, we run the process and it provides us with the prediction by the system and the precision value.

After successfully using the model, the performance of the model should be evaluated. The Performance operator in RapidMiner helps us in calculating the performance of the selected model. Immediately after the process of testing the model, the results regarding the accuracy of the data set are displayed in a confusion matrix format. Validation is the final sub-process of the entire process, and the model will be validated in this process. As Figure 4(b) shows that operator set role and cross validation has been used to make sure this process is successful. In cross validation, there are two sub-process which are training and testing. The first step is to divide the data into 10 equal parts. The remaining 9 subsets are used as training data, and 1 subset is kept as test data. The cross-validation procedure was then run 10 times, using 1 sample of the test data from each of the 10 subgroups. Model accuracy was then calculated by averaging the accuracy results from the first 10 iterations.



(a)



(b)

Figure 4. Testing model process based on (a) applying decision tree model and (b) cross validation esting model process

Based on Figure 5, it shows that the prediction (StudentAbsenceDays) which is the green column, tells the prediction that has been done by the system. The two yellow columns are for the confidence column. The following column confidence (under-7) tells the precision for students who are absence less for 7 days while column confidence (above-7) for students who are absence more than 7 days.

Row No.	prediction(S...	confiden... ↓	confidence(...	gender	raisedhands	VisITedReso...	Announcem...	Discussion	ParentAnsw...	StudentAbs...	Class
11	Under-7	1	0	M	50	88	30	80	Yes	Under-7	H
20	Under-7	1	0	M	70	50	40	99	Yes	Under-7	H
48	Under-7	1	0	F	70	4	39	90	Yes	Under-7	H
49	Under-7	1	0	F	13	80	40	88	Yes	Under-7	H
54	Under-7	1	0	F	49	70	19	75	Yes	Under-7	H
63	Under-7	1	0	M	80	90	70	80	Yes	Under-7	H
68	Under-7	1	0	F	65	75	23	80	Yes	Under-7	H
69	Under-7	1	0	F	70	69	35	30	Yes	Under-7	H
80	Under-7	1	0	F	80	90	49	55	Yes	Under-7	H
87	Under-7	1	0	M	70	12	40	50	Yes	Under-7	H
92	Under-7	1	0	M	80	90	55	19	Yes	Under-7	H
93	Under-7	1	0	F	50	70	19	15	Yes	Under-7	H
94	Under-7	1	0	M	55	89	40	40	Yes	Under-7	H
96	Under-7	1	0	F	100	80	2	70	No	Under-7	H
97	Under-7	1	0	F	14	60	11	75	Yes	Under-7	H

ExampleSet (480 examples, 3 special attributes, 8 regular attributes)

Figure 5. Prediction analysis

2.5. Evaluation

The evaluation phase determines whether the business goals have been achieved or not. This evaluation process allows us to evaluate the performance of the model once it is built. In addition, we must build trust in the validity and reliability of the results because if we simply accept the findings of the data mining model without any review, it will be very dangerous and result in poor decision making. Here is the overall accuracy and precision of all predictions that were achieved from the process as shown in Figure 6. The prediction class achieved the most outstanding class precision (89.26%) and the highest-class recall (95.50%) in the real Under-7 category.

The accuracy of correct predictions is 80.00% has be done by the model while, the next one to accuracy, which is the '+/-4.63%' one is called as a standard deviation in Figure 6(a). We take ten accuracies of our individual model and determine the standard deviation. Figure 6(b) illustrates the precision of the correct prediction model, which is 89.47% with +/- 4.85% standard deviation. The most crucial item when evaluating a model is, the smaller standard deviation value that we receive, the more stable is our model. A stable model where it has high accuracy and precision is good as it will produce a smaller range of best and worst cases which we must consider for the quality of our predictions. Finally, we should decide on how to proceed with the obtained results.

accuracy: 80.00% +/- 4.63% (micro average: 80.00%)

	true Under-7	true Above-7	class precision
pred. Under-7	276	83	76.88%
pred. Above-7	13	108	89.26%
class recall	95.50%	56.54%	

(a)

precision: 89.47% +/- 4.85% (micro average: 89.26%) (positive class: Above-7)

	true Under-7	true Above-7	class precision
pred. Under-7	276	83	76.88%
pred. Above-7	13	108	89.26%
class recall	95.50%	56.54%	

(b)

Figure 6. Results for model performance based on (a) accuracy and (b) precision from cross validation

2.6. Deployment

A final report and presentation of results is produced in this step. Information or pattern recognition from the data mining process is fed into the final phase. According to the research carried out, new knowledge and information was created in the data mining process to determine the elements and the extent to which they affect the academic progress of students in compliance with the data mining goals. RapidMiner is a tool used in schools to assess the accuracy and precision of classifying student performance. Any areas for improvement will be found in the final evaluation of the project. The final stage is to build a model that we will implement as a finished product.

3. RESULTS

The above processes or steps such as applying the model, testing and validation in order to analyze the decision tree on RapidMiner, generates the output in the form shown in above sections. Thus, we can conclude that the model has an accuracy of 80% and ±5.6% for standard deviation for predicting the statements based on label, while the precision of model is 89.47% with +/- 4.85% of standard deviation. Based on the calculation of data mining, it can analyze that our model is quite stable since it has high accuracy and precision with lower standard deviation. The smaller standard deviation value that we receive, the more stable is our model. The bar chart in Figure 7 compares the percentage of student absence days. It shows that the percentage of students who are absence less for 7 days is better than students who are absence more than 7 days where the differences between those group is 56%.

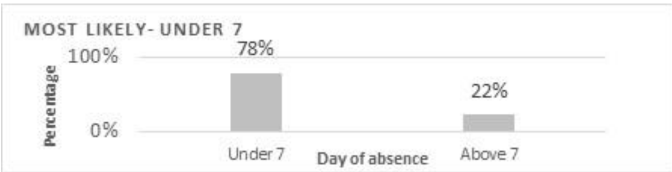


Figure 7. Percentage of student absence days

4. CONCLUSION

This study provides techniques for sentiment analysis and summarizing various documents using RapidMiner. Initially the dataset is collected from various school websites, blogs, and students' records. The collected data is processed, and their meaning must be understood. It is important to know the exact meaning of each attribute. This paper presents the analysis of students' performance using classification technique based on decision tree which is cross validation algorithm. The results showed the students' performance was affected by student absence days from school. Students who are not absent from school for under 7 days will be more encouraged to be a successful student where they can score more than 89 marks. However, students' performance also dependent on students' efforts factor where it was found that student who raised hands on every topic and getting involved in discussion are more prone towards success. When applying model students' performance and decision tree on RapidMiner tool, it gives out an accuracy of about 80.3% and ±5.6% for standard deviation. Other than that, the percentage of students who are absent for less than 7 days is better than students who are absent more than 7 days where the differences between those group is 56%. Based on the calculation of data mining, it can analyze that our model is quite stable since it has high accuracy with a lower standard deviation. In the near future, RapidMiner can be used as a tool to test the accuracy of the datasets and the analysis obtained will be able to help schools, students and also parents in adapting appropriate measures to improve the success of students at school.

ACKNOWLEDGEMENTS

This work was supported by Fundamental Research Grant Scheme (FRGS) under Ministry of Education (MOE) with grant number 600-IRMI/FRGS 5/3 (234/2019).

REFERENCES

[1] J. Santos-Pereira, L. Gruenwald, and J. Bernardino, "Top data mining tools for the healthcare industry," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 4968–4982, 2022, doi: 10.1016/j.jksuci.2021.06.002.

[2] J. Arunadevi, S. Ramya, and M. R. Raja, "A study of classification algorithms using RapidMiner," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 12, pp. 15977–15988, 2018.

[3] M. Noman, A. Kaur, and N. Nafees, “Children and youth services review COVID-19 fallout : Interplay between stressors and support on academic functioning of Malaysian University students,” *Children and Youth Services Review*, vol. 125, no. February, p. 106001, 2021, doi: 10.1016/j.childyouth.2021.106001.
[4] G. Czibula, A. Mihai, and L. M. Crivei, “S PRAR: a novel relational association rule mining classification model applied for academic performance prediction,” *Procedia Computer Science*, vol. 159, pp. 20–29, 2019, doi: 10.1016/j.procs.2019.09.156.
[5] J. Cotton, “Data: the importance of accuracy, integrity and real-time integration,” *New York: Information Builders*, vol. 20, 2017.
[6] R. Malhotra and K. Lata, “An empirical study to investigate the impact of data resampling techniques on the performance of class maintainability prediction models,” *Neurocomputing*, 2020, doi: 10.1016/j.neucom.2020.01.120.
[7] J. Wainer and G. Cawley, “Nested cross-validation when selecting classifiers is overzealous for most practical applications,” *Expert Systems With Applications*, vol. 182, no. December 2019, p. 115222, 2021, doi: 10.1016/j.eswa.2021.115222.
[8] T. T. Wong, “Parametric methods for comparing the performance of two classification algorithms evaluated by k-fold cross validation on multiple data sets,” *Pattern Recognition*, vol. 65, pp. 97–107, 2017, doi: 10.1016/j.patcog.2016.12.018.
[9] F. Mahan, M. Mohammadzad, and S. Meysam, “Chi-MFlexDT : Chi-square-based multi flexible fuzzy decision tree for data stream classification,” *Applied Soft Computing*, vol. 105, p. 107301, 2021, doi: 10.1016/j.asoc.2021.107301.
[10] H. L. Han, H. Y. Ma, and Y. Yang, “Study on the test data fault mining technology based on decision tree,” *Procedia Computer Science*, vol. 154, pp. 232–237, 2018, doi: 10.1016/j.procs.2019.06.035.
[11] N. R. Kuncel, M. Crede, and L. L. Thomas, “The validity of self-reported grade point averages, class ranks, and test scores: a meta-analysis and review of the literature,” *Review of Educational Research*, vol. 75, no. 1, pp. 63–82, 2005.
[12] C. F. Rodríguez-hern, M. Musso, E. Kyndt, and E. Cascallar, “Artificial neural networks in academic performance prediction: systematic implementation and predictor evaluation,” *Computers and Education: Artificial Intelligence*, vol. 2, no. December 2020, 2021, doi: 10.1016/j.caeai.2021.100018.
[13] K. R. Binning *et al.*, “Securing self-integrity over time : self-affirmation disrupts a negative cycle between psychological threat and academic performance,” *Journal of Social Issues*, no. March, pp. 1–23, 2021, doi: 10.1111/josi.12461.
[14] A. M. Ahmad, “Bad study habits of EFL learners as indicators of their poor performance : a case of the University of Bisha,” *International Journal of Applied Linguistics and English Literature*, vol. 7, no. 2, pp. 185-196, 2018, doi: 10.7575/aiac.ijalel.v.7n.2p.185.
[15] A. Naik and L. Samant, “Correlation review of classification algorithm using data mining tool : WEKA,” *Procedia - Procedia Computer Science*, vol. 85, no. Cms, pp. 662–668, 2016, doi: 10.1016/j.procs.2016.05.251.
[16] C. Schröer, F. Kruse, J. Marx, F. Kruse, and J. Marx, “A systematic literature review on applying CRISP-DM process model,” *Procedia Computer Science*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
[17] U. Gupta and R. Sharma, “Analysis of criminal spatial events in india using exploratory data analysis and regression,” *Computers and Electrical Engineering*, vol. 109, no. PA, p. 108761, 2023, doi: 10.1016/j.compeleceng.2023.108761.
[18] F. Morlock and M. Boßlau, “Concept for enabling customer-oriented data analytics via integration of production process improvement methods and data science methods,” *Procedia CIRP*, vol. 104, no. March, pp. 542–546, 2021, doi: 10.1016/j.procir.2021.11.091.
[19] V. Plotnikova, M. Dumas, and F. P. Milani, “Applying the CRISP-DM data mining process in the financial services industry: Elicitation of adaptation requirements,” *Data and Knowledge Engineering*, vol. 139, no. October 2021, pp. 1–17, 2022, doi: 10.1016/j.datak.2022.102013.
[20] U. Shafique and H. Qaiser, “A comparative study of data mining process models (KDD, CRISP-DM, and SEMMA),” *International Journal of Innovation and Scientific Research*, vol. 12, no. 1, pp. 217–222, 2014.
[21] A. Nouira, L. Cheniti-belcadhi, and R. Braham, “An enhanced enhanced xAPI xAPI data data model model supporting supporting assessment analytics,” *Procedia Computer Science*, vol. 126, pp. 566–575, 2018, doi: 10.1016/j.procs.2018.07.291.
[22] E. A. Amrieh, T. Hamtini, and I. Aljarah, “Mining educational data to predict student’s academic performance using ensemble methods,” *International Journal of Database Theory and Application*, vol. 9, no. 8, pp. 119–136, 2016, doi: 10.14257/ijdta.2016.9.8.13.
[23] S. Batt, T. Grealis, O. Harmon, and P. Tomolonis, “Learning Tableau: a data visualization tool,” *Journal of Economic Education*, vol. 51, no. 3–4, pp. 317–328, 2020, doi: 10.1080/00220485.2020.1804503.
[24] J. Hoelscher and A. Mortimer, “Using Tableau to visualize data and drive decision-making,” *Journal of Accounting Education*, vol. 44, no. May, pp. 49–59, 2018, doi: 10.1016/j.jaccedu.2018.05.002.
[25] N. Harb and A. El-Shaarawi, “Factors affecting students’ performance,” *Journal of Business Education*, vol. 82, no. 5, pp. 282–290, 2006.

BIOGRAPHIES OF AUTHORS

Muhammad Firdaus Mustapha is a senior lecturer at Universiti Teknologi MARA (UiTM) Kelantan Branch, Malaysia. He obtained his doctor of philosophy (Ph.D.) in computer science from UiTM Shah Alam, Malaysia in 2018. He received his bachelor of science (computer) and master of science (computer science) from Universiti Teknologi Malaysia (UTM), Malaysia in 2006 and 2009 respectively. His current research interests are artificial intelligence, computer vision, machine learning and deep learning. He can be contacted at email: mdfirdaus@uitm.edu.my.

Alia Nur Izzah Zulkifli acquired a bachelor of science (Hons.) mathemathics from Universiti Teknologi Mara (UiTM) Kelantan Branch, Malaysia in September 2022. Her final year project during her undergraduate degree was titled the effect of conversion rate on prey-predator model with disease in prey population. She is currently pursuing a master of science (mathematics) at the same university where she previously studied. Her master's thesis focuses on developing a new mathematical set for medical diagnosis. She can be reached via email: alianurizzah2704@gmail.com.

Omar Kairan received the master of applied science from Universiti Teknologi MARA (UiTM), Malaysia, in 2010 and completed his bachelor's in statistics in 2006. He is currently a senior lecturer at Universiti Teknologi MARA (UiTM) Kelantan Branch, Malaysia. His research interests are in statistical modelling and multivariate analysis. He can be contacted at email: omarkr@uitm.edu.my.

Nur Nabila Sofea Mat Zizi graduated with a bachelor of science (Hons) in mathematics in September 2022 from Universiti Teknologi MARA (UiTM) Kelantan Branch, Malaysia. She is now enrolled master of science (sports science and recreation) by research program at UiTM Shah Alam. She can be contacted at email: nurnabilasofea99@gmail.com.

Nur Naim Yahya obtained bachelor of science (Hons) mathematics in September 2022 from Universiti Teknologi Mara (UiTM) Kelantan Branch, Malaysia. She is currently a postgraduate student in diploma of teaching at Open University Malaysia (OUM) Kelantan Branch. She can be contacted at email: nurnaimm99@gmail.com.

Nur Maisarah Mohamad obtained bachelor of information technology (media informatics) in February 2020 from Universiti Sultan Zainal Abidin (UniSZA). Her final year project during her undergraduate degree was about educational courseware. She is currently a postgraduate student in master of computer science at Universiti Teknologi Mara (UiTM) Kelantan Branch, Malaysia. Her master thesis involves several methods in computer vision and deep learning includes convolutional neural network, transfer learning, and classification. She can be contacted at email: 2020273478@student.uitm.edu.my.