Mathematical Sciences and Informatics Journal
Vol. 3, No. 2, Nov. 2022, pp. 1-10
http://www.mijuitmjournal.com DOI: 10.24191/mij.v3i2.20117

Students' Attitude towards Video-based Learning: Machine Learning Analysis with Rapid Software

Abdullah Sani Abd Rahman
Faculty of Sciences and Information Technology, Universiti Teknologi Petronas, Perak, Malaysia
sani.arahman@utp.edu.my

Rita Meutia
Faculty of Economics and Business, Universitas Syiah Kuala, Banda Aceh, Indonesia
rita.mutia@unsyiah.ac.id

Yusnaliza Hamid
Faculty of Accounting, Universiti Teknologi MARA Perak Branch, Tapah Campus, Perak, Malaysia
yusna421@uitm.edu.my

Rahayu Abdul Rahman
Faculty of Accounting, Universiti Teknologi MARA Perak Branch, Tapah Campus, Perak, Malaysia
rahay916@uitm.edu.my

Article Info

Article history:

Received July 05, 2022
Revised Sept. 02, 2022
Accepted Oct 19, 2022

Keywords:

Machine learning
Rapid software
Video-based learning
Student attitude
Perceive useful
Perceive ease of use

ABSTRACT

The recent computer and Internet technologies have dramatically impacted many facets of education. There has been a rapid rise mainly since the COVID19 pandemic in the use of video-based learning implemented via online classroom setting. Regardless of its usefulness and practicality, the educational technologies adoption has its challenges faced by educators. The main challenge is to get intuition of the students' attitude that will influence the students' performances. In supporting the intervention approaches, machine learning techniques have been widely utilized. Another challenge is the difficulty to implement machine learning analysis by the educators. The emergence of rapid software platform can be useful for them, but the existence of these software is unrenowned. The goals of this paper are to: 1) provide fundamental experimental works of the machine learning implementation based on a new rapid software framework and 2) present the ability of machine learning in classifying students' attitude towards video-based learning. Data were collected from a university level accounting course (n=103), involving students who have different experienced or exposure on video-based online learning. Three machine learning algorithms (Support Vector Machine, Random Forest, and Decision Tree) have been tested on the dataset in a rapid software platform. The results show that all the three machine learning algorithms produced high accuracy (above 95%) prediction results based on the hold-out testing dataset. Furthermore, considering inputs of students perceive on the useful and ease of use of video-based learning as well as excluding the demography attributes in the machine learning models seems more useful in the tested case. This paper presents the fundamental design and easy implementation of machine learning for education domain useful for the inexpert data scientists in many fields.

Corresponding Author:
Rahayu Abdul Rahman
Faculty of Accounting, Universiti Teknologi MARA Perak Branch, Tapah Campus, Perak, Malaysia
email: rahay916@uitm.edu.my

PENERBIT PRESS
UNIVERSITI TEKNOLOGI MARA

1. Introduction

The world is on the rise of complexity and uncertainty mainly due to the COVID19 which have major impact on the education practices. In Malaysia, the declaration of COVID-19 lockdown or Movement Control Order (MCO) has led to the postponement of all face-to-face classroom activities. During the MCO periods, all universities implemented online learning that inspired most educators to utilize various innovative teaching due to the problems in attaining the students' attention and engagement. Video-based learning was one of the widely used approach during the MCO and is expected to be beneficial to be used continuously in the future of education setting either online or face-to-face[1].

Teaching with technology like video-based learning provides additional factors that need to be considered in terms of the teaching pedagogy and constructions of the learning contents[2]. It is highly important for the educators to get some insight on the students' behavior and perceptions towards the learning approaches and contents to ensure the most effective learning can be achieved at the end of every class. Additionally, to grasp the students' attitude at the early semester of teaching based on their demographic attributes and academic achievement as well as their background and exposure on video-based learning will help the teachers to plan the most suitable and appropriate teaching methods for their students[3],[4]. Nevertheless, to conduct the pre-class analysis can be a very difficult and hassle tasks for the educators. Recently, machine learning sound to be very useful in educational data mining but most of educators are inexpert data scientists that need to be supported with a very simple and easy tool.

The contribution of this paper is two-folds. First, it presents an easy framework for educators to implement machine learning analysis with rapid software tool. Second, the experimental works show performance comparison of three machine learning algorithms at three different settings. This paper can be as a useful guideline for the educators as well as other inexpert data scientists in conducting the basis tasks of machine learning analysis.

2. Literature Review

Machine learning is the one of the most promising application areas in a field of Artificial Intelligence (AI) where its application areas is unlimited[5]. This research focused on machine learning in education that has been expanding from the traditional classroom to the use of digital resources since the new era of Education 4.0. To date, there are plenty of education implementations with machine learning. One of the interesting applications is predicting student performances by learning about each students' historical data[6], [7]. Predicting students dropout with machine learning is another important application that have been conducted in [8], [9]. Researchers in [10] used multi-class machine learning to predict students; dropout, success or failure from the MOOCs learning contents. Academic procrastination is highly significant to the students' achievement where researchers in [11] used the students submission history data to predict the procrastination occurrence while in other interesting work was using the procrastination behavioral data to predict the students' achievement [12]. Acknowledging the importance of getting ideas on the students' learning styles, researchers in [13], [14] employed machine learning to achieve the objectives. All the mentioned works have characterized the important factors to be considered in the machine learning analysis including students' gender, academic results, and the students' historical attitudes on learning. This research focused on different aspects to contribute as a new machine learning predictor model by considering the students perceive useful and perceive ease of use to predict the video-based learning attitude. Perceive useful, perceive ease of use and attitude towards technology is the important attributes of Technology Acceptance Model (TAM)[15].

Although there are many software tools that can be used for implementing machine learning analysis, some of them remain difficult to be used by the inexpert users. Some of these tools is unknown and limited in the literature. The most popular tool is using R[16] or Python[17] programming but basic programming knowledge is required and complicated to the educators to conduct fast and frequent analysis. Python programming seem to be simpler than R but there involved variety of libraries[18] that need to be understood by the inexpert programmers. Using software tools with Graphical User Interfaces (GUIs) is better for them by some software like RapidMiner and Weka consists of huge components that need a comprehensive training. Thus, this research introduced a very simple rapid software tool with the minimal GUIs.

3. Methodology

3.1 Sample of data

This research gathered data from questionnaires survey for the machine learning analysis. The questionnaire consists of two sections: demographic and TAM constructs. In particular, the first section collected demographic characteristics of the respondents including gender, academic performance, residential area and monthly family income. This section also captured information on students' exposure of video-based learning prior covid-19 outbreak. The second section constructed based on TAM attributes namely perceived usefulness, perceived ease of use and attitude. The questionnaires were personally administered to undergraduate accounting students from the one public university in Malaysia during the second semester of 2021/2022 academic year. To ensure voluntary participation and honest responses from the students, the students were assured of confidentiality and that their responses were to be used solely for this research. Out of a total of 280 questionnaires administered, 103 valid responses were used for the machine learning analysis, representing a response rate of 36.78 percent. Table 1 lists the independent variables (IVs) in classifying the DV, which is the students' attitude toward video-based learning. Based on Pearson correlation test, the two TAM attributes present positive strong correlations (above 0.7 correlation coefficient) to the students' attitude. This is consistent with the finding revealed by researchers in [19]. Other demography and academic results present low correlations to the attitudes but remain included in the machine learning models. The results of demography inclusion and exclusion will be observed in the machine learning models.

Table 1. Pearson correlation of independent variables

Independent variables	Correlation coefficient
Perceived_ease_of_use	0.78
Perceived_usefulness_vbl	0.71
Cgpa	0.31
Gpa	0.23
Residential_area_urban	0.15
Prior_exposure_on_vbl	0.07
Monthly_household_income	0.06
Gender	0.02

3.2 The rapid software

To implement the machine learning, researchers used a rapid software introduced in [20]. The software allows researchers to implement three machine learning algorithms; Support Vector Machine (SVM), Random Forest and Decision Tree as presented in Figure 1. After clicking the selected algorithm, the next GUI is to set the testing ratio as depicted in Figure 2.

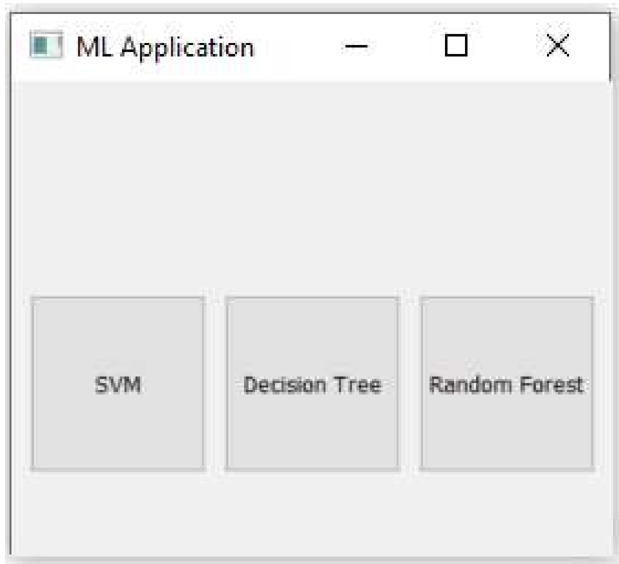

Figure 1. Main interface

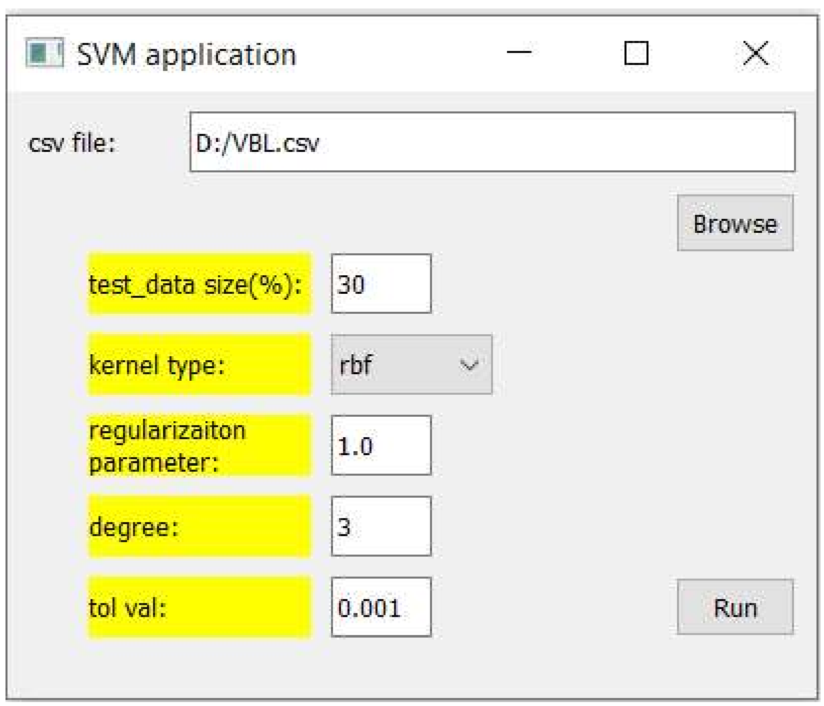

Figure 2. Configuration interface for SVM

Figure 2 shows the configuration interface when SVM is selected. There are three kernel types can be chosen *rbf, linear, poly, sigmoid.* Based on the data of this research, all the kernel types and others parameters did not influencing the SVM results. The best *degree* identified for the SVM is 3 and the best *test data size* is 30. The *test data size* used to set the machine learning training and testing ratio. For this research, 70 percent is used for training and 30 percent for testing. Thus, from the 103 dataset, 70 data used on the machine learning training while 31 for the machine learning testing for evaluating the performances. Different with Random Forest and Decision Tree, the configuration interface is depicted in Figure 3.

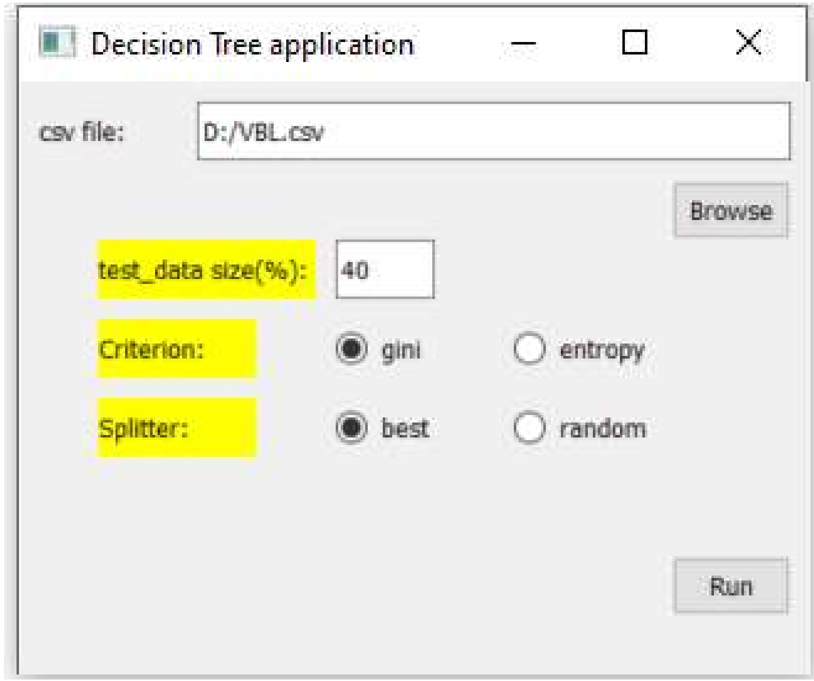

Figure 3. Configuration interface for Decision Tree and Random Forest

During the preliminary works, researchers have observed that the used of different *Criterion* and *Splitter* have not major effect the algorithms results. Like SVM, the best *test data size* was 30 for 70:30 ratio between training and testing datasets.

4. Results and Discussion

This section presents the results produced by the rapid software and the comparison analysis between machine learning models that used all IVs (included demography attributes) and only used two TAM attributes (perceive useful and perceive ease of use without demography).

4.1 Support Vector Machine (SVM)

Figure 4 presents the confusion matrix after selecting SVM and setting the mentioned configuration to be run. As *test data size* was set as 30, there are 31 data will be presented in the confusion matrix, where 30 of them has been correctly predicted as positive attitude (case 1) by the SVM. One real case with 0 value has been wrongly predicted as 1. Based on the confusion matrix, the metrics of accuracy, precision, recall and f1-score can be viewed as presented in Figure 5.

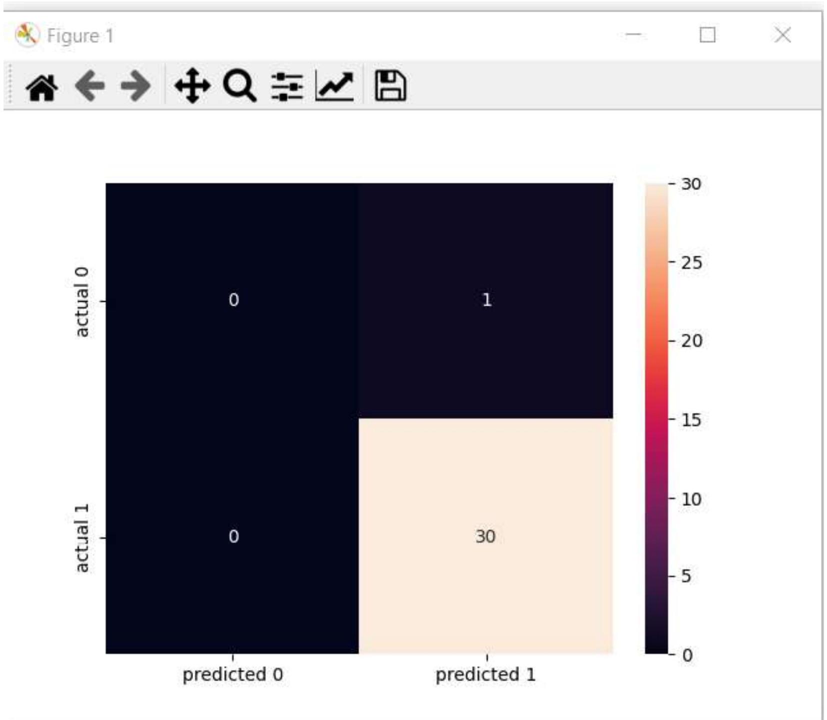

Figure 4. The Confusion Matrix of SVM

Results:

	precision	recall	f1-score	support
0	0.00	0.00	0.00	1
1	0.97	1.00	0.98	30
accuracy			0.97	31
macro avg	0.48	0.50	0.49	31
weighted avg	0.94	0.97	0.95	31

OK

Figure 5. The results of SVM performance

The accuracy is very high at 0.97 and the precision of detecting case 1 is 0.97. As the SVM can truly detected all the case 1, the recall ability reached to 1.0. Only one support data has been used for the case 0, and this is not correctly classified by SVM calculated as zero precision, recall and f1-score for the case. This issue will be a potential research to researchers by extending the variation of testing data with the 0 case in the future.

4.2 Decision Tree

Compared to SVM, Decision Tree produced slightly lower accuracy at 0.94 From the total 30 cases with positive attitude (case 1), there were 1 case has been predicted as negative attitude (case 0). The results of calculation for each performance metric are presented in Figure 7.

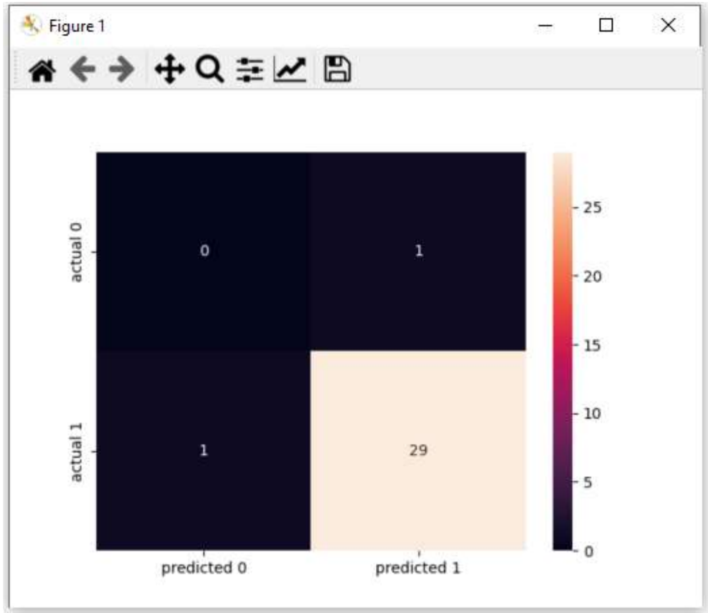

Figure 6. The Confusion Matrix of Decision Tree

Results:

	precision	recall	f1-score	support
0	0.00	0.00	0.00	1
1	0.97	0.97	0.97	30
accuracy			0.94	31
macro avg	0.48	0.48	0.48	31
weighted avg	0.94	0.94	0.94	31

OK

Figure 7. The results of Decision Tree performance

During the preliminary experiments, it has been observed that the results from Decision Tree is similar to SVM if the size of testing dataset was set to 20 at 95% accuracy, 0.98 recall and 0.95 f1-score for detecting case 1. Therefore, Decision Tree can reach better accuracy with lower size of testing set compared to SVM.

4.3 Random Forest

The output of results from Random Forest algorithm is not given in this paper as the results are exactly same to the SVM performances at both testing size 20 and 30. The following provides results analysis of the three machine learning algorithms in accordance with the attributes or IVs used in the models focused on the accuracy.

4.4 IVs analysis in the machine learning models

Table 2 lists the accuracy results from the three machine learning algorithms with two groups of IVs or attributes. The first group includes all the IVs from demography and TAM while the second group only used the two TAM attributes (perceive useful, perceive ease of use).

Table 2. Demography effect on the results

	All attributes including demography	**TAM attributes without demography**
SVM	0.97	0.86
Random Forest	0.97	0.86
Decision Tree	0.94	0.83

The results in Table 2 show that the six demography attributes have not contributing too much in the accuracy of the model as the correlation values from these attributes were very low (Refer Table 1). Therefore, by considering only on the two attributes from the TAM can be useful enough for the machine learning algorithms to classify the students' attitude.

5. Conclusion

This research presents different aspects of modelling for classifying students' attitude towards video-based learning from the context of attributes characterization and the rapid software platform. The fundamental steps used in the research can be easily replicated to be conducted for other kinds of machine learning classification problem from various application domain. Additionally, this research can be extended by considering all the three TAM attributes as the inputs model and predicting the pre-defined student achievement with the machine learning techniques. Other than demography attributes, other factors can be included such as the students' technology readiness.

Acknowledgements

The authors gratefully acknowledge the Universiti Teknologi Petronas, Universiti Teknologi MARA and Universitas Syiah Kuala for all the supports given to the research.

Conflict of Interest

The authors declare no conflict of interest in the subject matter or materials discussed in this manuscript.

References

[1] M. Sablić, A. Mirosavljević, and A. Škugor, "Video-based learning (VBL)—past, present and future: An overview of the research published from 2008 to 2019," *Technol. Knowl. Learn.*, vol. 26, no. 4, pp. 1061–1077, 2021.

[2] J. Gillett-Swan, "The challenges of online learning: Supporting and engaging the isolated learner," *J. Learn. Des.*, vol. 10, no. 1, pp. 20–30, 2017.

[3] L. Cheng, A. D. Ritzhaupt, and P. Antonenko, "Effects of the flipped classroom instructional strategy on students' learning outcomes: A meta-analysis," *Educ. Technol. Res. Dev.*, vol. 67, no. 4, pp. 793–824, 2019.

[4] E. P. Wagner, H. Sasser, and W. J. DiBiase, "Predicting students at risk in general chemistry using pre-semester assessments and demographic information," *J. Chem. Educ.*, vol. 79, no. 6, p. 749, 2002.

[5] D. Kučak, V. Juričić, and G. DJambić, "MACHINE LEARNING IN EDUCATION-A SURVEY OF CURRENT RESEARCH TRENDS.," *Ann. DAAAM Proc.*, vol. 29, 2018.

[6] J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido, and A. Durán-Dom\'\inguez, "Analyzing and predicting students' performance by means of machine learning: A review," *Appl. Sci.*, vol. 10, no. 3, p. 1042, 2020.

[7] R. Umer, T. Susnjak, A. Mathrani, and S. Suriadi, "On predicting academic performance with process mining in learning analytics," *J. Res. Innov. Teach. Learn.*, vol. 10, no. 2, pp. 160–176, 2017, doi: 10.1108/jrit-09-2017-0022.

[8] L. Kemper, G. Vorhoff, and B. U. Wigger, "Predicting student dropout: A machine learning approach," *Eur. J. High. Educ.*, vol. 10, no. 1, pp. 28–47, 2020.

[9] J. Y. Chung and S. Lee, "Dropout early warning systems for high school students using machine learning," *Child. Youth Serv. Rev.*, vol. 96, pp. 346–353, 2019.

[10] Y. Mourdi, M. Sadgal, H. El Kabtane, and W. B. Fathi, "A machine learning-based methodology to predict learners' dropout, success or failure in MOOCs," *Int. J. Web Inf. Syst.*, 2019.

[11] A. Akram *et al.*, "Predicting students' academic procrastination in blended learning course using homework submission data," *Ieee Access*, vol. 7, pp. 102487–102498, 2019.

[12] D. Hooshyar, M. Pedaste, and Y. Yang, "Mining educational data to predict students' performance through procrastination behavior," *Entropy*, vol. 22, no. 1, p. 12, 2019.

[13] K. Suresh, J. Meghana, and M. E. Pooja, "Predicting the E-Learners Learning Style by using Support Vector Regression Technique," in *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 2021, pp. 350–355.

[14] A. M. F. Yousef *et al.*, "Automatic Identification of Student's Cognitive Style from Online Laboratory Experimentation using Machine Learning Techniques," in *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 2021, pp. 143–149.

[15] R. Ibrahim, N. S. Leng, R. C. M. Yusoff, G. N. Samy, S. Masrom, and Z. I. Rizman, "E-learning acceptance based on technology acceptance model (TAM)," *J. Fundam. Appl. Sci.*, vol. 9, no. 4S, pp. 871–889, 2017.

[16] B. Bischl *et al.*, "mlr: Machine Learning in R," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 5938–5942, 2016.

[17] S. Raschka and V. Mirjalili, *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2*. Packt Publishing Ltd, 2019.

[18] J. Hao and T. K. Ho, "Machine learning made easy: a review of scikit-learn package in python programming language," *J. Educ. Behav. Stat.*, vol. 44, no. 3, pp. 348–361, 2019.

[19] D. Pal and S. Patra, "University students' perception of video-based learning in times of COVID-19: A TAM/TTF perspective," *Int. J. Human--Computer Interact.*, vol. 37, no. 10, pp. 903–921, 2021.

[20] A. S. Abd Rahman, S. Masrom, R. Abdul Rahman, and R. Ibrahim, "Rapid Software Framework for the Implementation of Machine Learning Classification Models," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 11, pp. 8–18, 2021, doi: 10.46338/ijetae0821_02.

Biography of all authors

Picture	**Biography**	**Authorship contribution**
	Ts. Abdullah Sani Abd Rahman obtained his first degree in Informatique majoring in Industrial Systems from the University of La Rochelle, France in 1995. He received a master's degree from Universiti Putra Malaysia in Computer Science, with specialization in Distributed Computing. Currently, he is a lecturer at the Universiti Teknologi PETRONAS, Malaysia and a member of the Institute of Autonomous System at the same university. His research interests are cybersecurity, data analytics and machine learning. He is also a registered Professional Technologist. He can be contacted at email: sani.arahman@utp.edu.my.	Drafting, machine learning and final checking the article

	Associate Professor Dr Rahayu Abdul Rahman is an Associate Professor at the Faculty of Accountancy, UiTM. She received her PhD in Accounting from Massey University, Auckland, New Zealand in 2012. Her research interest surrounds areas, like financial reporting quality such as earnings management and accounting conservatism as well as financial leakages including financial reporting frauds and tax aggressiveness. She has published many research papers on machine learning and its application to corporate tax avoidance. She is currently one of the research members of Machine Learning and Interactive Visualization Research Group at UiTM Perak Branch. She can be contacted through email at rahay916@uitm.edu.my.	Data collection and correlation analysis
	Madam Yusnaliza Binti Hamid is a Senior Accounting Lecturer at the Faculty of Accountancy, Universiti Teknologi MARA (UiTM), Tapah Campus, Perak. She earned her bachelor's degree and Master of Arts in Accounting and Financial Management at the University of Sheffield (1994) and University of Essex (1996), United Kingdom respectively. Her expertise lies in the field of Financial Accounting, Group Accounting and Financial Reporting involving the application of International Financial Reporting Standards (IFRS) and International Accounting Standards (IAS). Adapting with the development of electronic and online learning in the classroom, she has also created an educational YouTube channel. To date, the channel receives over 400,000 views from 22 countries globally, where it aims to provide accounting students' help online by way of comprehensive lectures, step-by-step answers to past exam questions, and detailed explanations of the selected IFRS principles and concepts. She can be contacted through email at yusna421@uitm.edu.my.	Abstract and introduction
	Associate Professor Dr. Rita Meutia, SE, MSi,Ak. is an Associate Professor at the Economic and Business Faculty, Universitas Syiah Kuala. She received her PhD in Accounting from Universitas Sumatera Utara, in 2019.She Started her career as a lecturer at USK after receiving her Bahelor's degree in Accounting from USK in 1995 , and then continued her Masters degree in Accounting from Universitas Padjadjaran, in 1999. She is an active researcher in Financial Accounting and Public Sector Accounting. She can be contacted through email at rita.mutia@unsyiah.ac.id	Literature review