Indonesian Journal of Electrical Engineering and Computer Science
Vol. 31, No. 2, August 2023, pp. 909~916
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v31.i2.pp909-916

Machine learning in predicting whistle-blowing intention of academic dishonesty with theory of planned behaviour

Suraya Masrom[1], Nor Hafiza Abdul Samad[2], Rahayu Abdul Rahman[3], Farah Husna Mohd Fatzel[3], Siti Marlia Shamsudin[3]
[1]College of Computing, Informatics and Media, Universiti Teknologi MARA, Perak Branch, Malaysia
[2]Faculty of Computing and Multimedia, Universiti Poly-Tech Malaysia, Kuala Lumpur, Malaysia
[3]Faculty of Accountancy, Universiti Teknologi MARA, Perak Branch, Malaysia

Article Info

Article history:

Received Jan 18, 2023
Revised Apr 10, 2023
Accepted Apr 16, 2023

Keywords:

Academic dishonesty
Area under curve
Machine learning prediction
Theory of planned behavior
Whistle-blowing

ABSTRACT

The COVID-19 pandemic and its aftermath have caused most higher educations to choose to implement remote learning as a new method of instruction and assessment. Nevertheless, remote learning has been criticized by having adverse impact on academic integrity. Whistle-blowing has been regarded as an effective mechanism in limiting such unethical behavior. Thus, the main objective of this study is to identify the influence attributes of whistle-blowing intention among university students. The effectiveness of the whistle-blowing attributes was observed in prediction models based on machine learning technique. This paper presents the fundamental knowledge on evaluations of tree-based machine learning algorithms namely decision tree, random forest, to be compared with logistics regression and gradient linear model. A rigorous evaluation reports are provided that includes the area under curve (AUC) as a supplementary metric to measure the model accuracy. Additionally, to provide a clearer insight on the whistle-blowing prediction models, the pattern of influences from the whistle-blowing attributes based on the adoption of theory of planned behavior (TPB) and demography are presented. The findings revealed that both TPB and demography attributes contain some degree of impressive knowledge for the machine learning to generate a good prediction result.

Corresponding Author:

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

1. INTRODUCTION

Remote learning has been implemented by higher education institutions globally in response to COVID-19 pandemic and its social confinement enforcement [1], [2]. Although remote learning provides some beneficial impact to the learning [3]-[5], there are some drawbacks that educators face [3], [6]. Prior studies in [7], [8] stressed that although remote learning is regarded as an effective strategy especially during COVID-19 pandemic to mitigate health risks for both educators and students, it has adverse impact on academic integrity. Using electronic examination as a student's assessment tool gives more opportunities for students to engage in academic dishonesty [9] and good for fostering their self-regulated learning [5]. Achmada *et al.* [10] define academic misconduct or dishonesty as an intentional act of fraud, in which a student seeks to claim credit for the work or efforts of another without authorization, or uses unauthorized materials or fabricated information in any academic exercise. Academic dishonesty includes forgery of academic documents, intentionally impeding or damaging the academic work of others, or assisting other students in acts of dishonesty.

In response, various strategies have been introduced by higher education institutions. One of widely used mechanisms to mitigate academic dishonesty among universities' students is whistle-blowing [11]-[14]. Whistle-blowing is defined as disclosure by organization members of illegal, immoral or illegitimate practices to persons or organizations that may be able to effect action [15]. Whistle-blowing plays an important role in uncovering frauds and organizational wrongdoing [16]. For example, in a corporate setting, by reporting dishonesty in place, whistle-blowing can help organizations to avoid financial losses due to employee embezzlement, lawsuits filed resulting from employee discrimination or moral assault cases, and reputation damages [17]. Whistle-blowing, however, is a risky moral duty. Most whistle-blowers face some form of retaliation from colleagues or supervisors after disclosing dishonesty [18], [19]. For instance, in a corporate environment, they suffer from termination, demotion, unfavorable job performance evaluation, involuntary transfer, assignment of unmanageable tasks, professional blacklisting and social ostracism. Meanwhile in academic settings, whistle-blowers face social ostracism, name-calling and other forms of social sanctions from their academic peers [20]. Due to various personal risks, many individuals choose to remain silent.

Given such a dilemma and social environment, it is important to predict whistle-blowing intentions and investigate factors that influence individuals to blow the whistle in an academic setting. Thus, this study aims to expand prior works by examining student's intentions to report wrongdoing in academic settings. Unlike prior studies [7], [14], [19], [21], [22] that employed traditional statistical methods, this study attempts to construct students' whistle-blowing intention model on academic dishonesty using computational intelligence approach or specifically with machine learning prediction technique. Further, despite widely use of machine learning in various domain of research including in education [23], business [24], fraud detection [25], energy management [26] and medical [27] that highlight the effectiveness of such methods to that of traditional statistical methods [28], [29], yet study on machine learning prediction and classification on whistle-blowing academic fraud is limited.

This study has three main contributions. First, it attempts to construct whistle-blowing academic dishonesty prediction model with machine learning algorithms. Second, in order to deepen current understanding on the acceptance of whistle-blowing as one of the universities mechanisms in mitigating academic dishonesty, this paper presents the inclusion of theory of planned behavior (TPB) [30] in the machine learning prediction models based on three constructs of TPB namely attitudes, subjective norm towards the behavior, and perception of behavior control. Third, this paper delivers a rigorous evaluation result of the machine learning models from the aspects of performance metrics and the attributes of whistle-blowing intentions.

The following section provides a brief description on the data set of the concerned problem and machine learning implementation methodology. Section 3 describes and discusses the experimental results for the representative compared algorithms. Finally, section 4 presents the conclusions and future research directions.

2. METHOD

2.1. Sample of data

This study gathered data from questionnaires survey, which consists of two sections; demographic and theory of planned behavior constructs. In particular, the first section collected demography including gender, age, type of university either public (IPTA) or private (IPTS), course and academic performance. The cumulative grade point average (CGPA) is the attribute to measure academic performance. This section also captured information on students' perception towards their university integrity culture and fear of retaliation perception. The second section aims to measure respondents' intention to report academic dishonesty. Based on the tenets of the TPB [13], [31] three constructs have been employed to measure student whistle-blowing intention that consists of attitudes, subjective norms towards the behavior and perception of behavior control.

Attitudes refers to the degree to which a person has a favorable or unfavorable measure of the whistle-blowing interest either through affective or instrumental attitude. Affective attitude emphasizes more the emotional aspects of behavior that reflects the enjoyment and negative feelings. On the other hand, an instrumental factor in attitude is a behavior that perceives to make desirable or undesirable outcomes. Instrumental attitude accentuates more the cognitive aspects of behavior.

Subjective norms towards the behavior refers to the belief on approving or disapproving the whistle-blowing behavior. It relates to a person's principles about whether their peers should engage in the behavior or not. The first type of subjective norms is descriptive norms, which are the perception towards other students that most commonly perform the whistle-blowing behavior. The second type is injunctive norms or social norms that refers to the social pressures in a group of peoples to perform the behavior.

Behavioral control or intention reflects the motivational factors that control the behavior such that the stronger the intention to perform the whistle-blowing behavior, the more likely the behavior will be performed.

Self-efficacy and perceived control-ability are the two attributes of behavior control. Self-efficacy is defined as the student's confidence to carry out the whistle-blowing behavior while to be able to control the whistle-blowing behavior is defined as perceived control-ability.

The specific indicators used to measure each of the three constructs were adapted from the works of [17]. The questionnaires were personally administered to undergraduate students from the three universities in Malaysia during the first semester of 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 300 questionnaires administered, 163 valid responses were used for the analysis, representing a response rate of 54.33%. Figure 1 presents the weights of correlation coefficients of each attribute used in the whistle-blowing classification model.

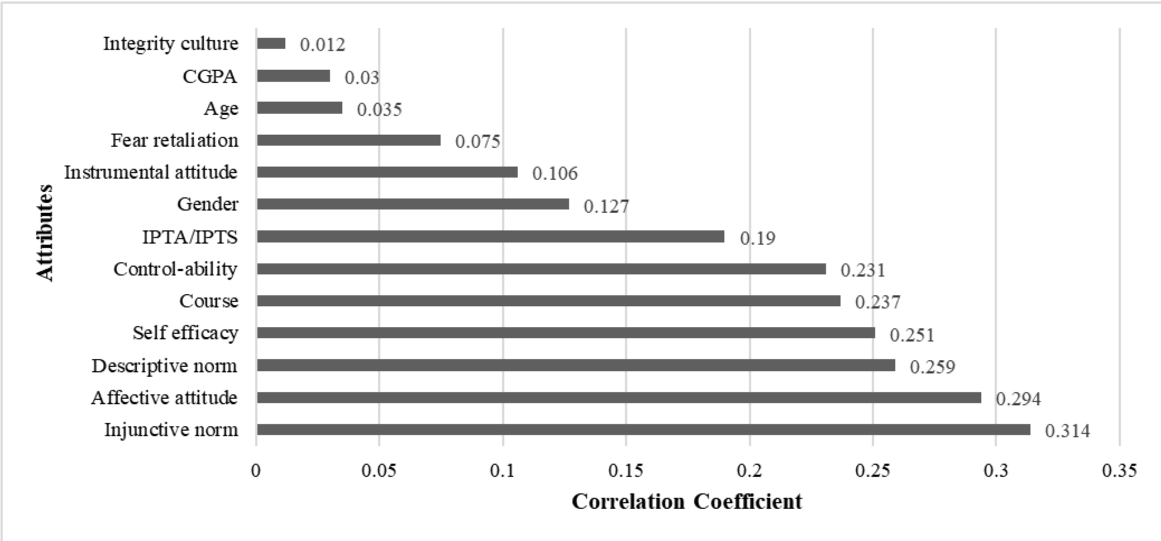

Figure 1. Weights by correlation outside machine learning algorithm

As seen in Figure 1, the main problem of the collected data-set is very weak association between each attribute to the dependent/target variable, which is the whistle-blowing intention. Thus, including all the attributes as features selection for the machine learning models is expected to be beneficial to increase the accuracy. Each of the attributes will contribute some degree of knowledge to the prediction models but it is important to understand how different their contribution is between the different machine learning algorithms.

2.2. The machine learning algorithms

This research used two types of the tree-based machine learning algorithms namely decision tree and random forest to be compared with other non-family tree-based algorithms (logistic regression, generalized linear model). Unlike logistic regression and generalized linear model, hyper-parameters preliminary analysis is essential for tree-based machine learning. As a tree-based algorithm, the common hyper-parameter is maximal depth and number of trees is an additional hyper-parameter for random forest. Table 1 lists the optimal setting for the hyper-parameters. The following Figure 2 and Figure 3 visualized the different error rates of decision tree and random forest based on the different hyper-parameters values.

Table 1. The optimal hyper-parameters for decision tree and random forest

Machine learning algorithm	Hyper-parameters	Error rate %
Decision tree	Maximal depth=4	44.9
Random forest	Number of trees=60 Maximal depth=4	36.7

As depicted in Figure 2, the highest error rate reached at 55.1% when the maximal depth was 2 and the lowest can be achieved when maximal depth was 4 as given in Table 1. In Figure 3, the Y-axis is to plot the maximal depth and the X-axis is number of trees for presenting the variation of error rates in random forest.

The size of the circle and colors representing the size of the error in such that the bigger the circle, the more error was generated by the random forest.

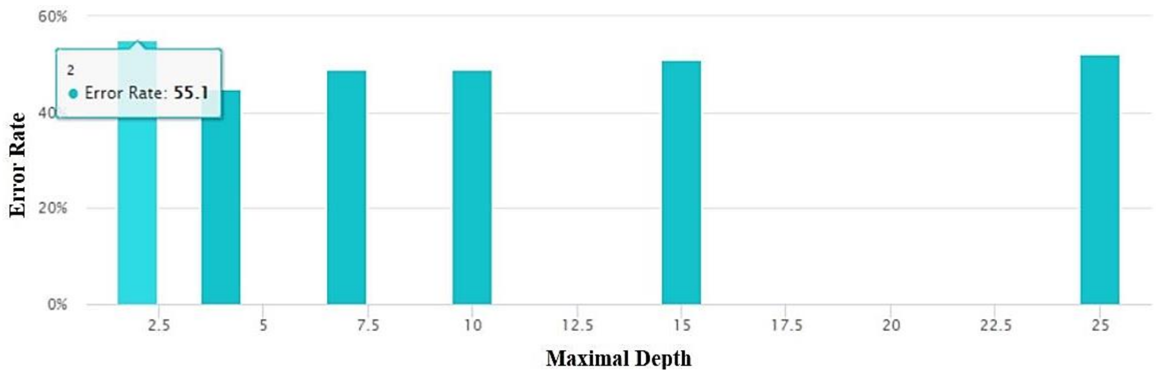

Figure 2. Error rates of decision tree at different maximal depth

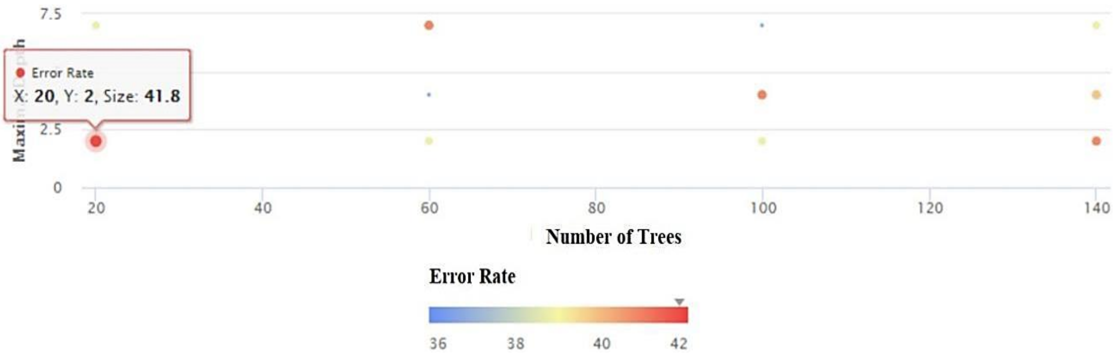

Figure 3. Error rates of random forest at different maximal depth and number of trees

The least optimal hyper-parameters values for random forest can be achieved when the maximal depth was set to 2 and the number of trees was 20. This setting generated the worst error rate at 41.8%. The lowest error rate (35.7%) is denoted with the smallest blue circle with 60 number of trees and 4 maximal depth.

2.3. Training approach and evaluation techniques

The research employed a 60:40 split training approach to separate the training and testing datasets, with 98 out of 163 data used for machine learning training and 65 for testing. Commonly used metrics for evaluating machine learning algorithms, such as accuracy and error, assess the model's overall prediction performance without specifying the class group. Additionally in this research, the area under curve (AUC) is a powerful metric used to evaluate the performance of the machine learning algorithms, which is more suitable for the whistle-blowing binary classification problem. Unlike accuracy, which only considers the overall number of correctly classified instances, AUC provides a comprehensive measure of the model's performance at all classification thresholds, taking into account the trade-off between sensitivity and specificity. This means that AUC is a more robust evaluation metric for classification models that can handle imbalanced datasets and account for the varying costs of false positives and false negatives.

3. RESULTS AND DISCUSSION

Table 2 lists the results of AUC, accuracy, classification error and total completing time (TCT) for each machine learning algorithm. A perfect classifier would have an AUC of 1, while a weak classifier that usually do the prediction only by chance without learning the data relationships and pattern would have an AUC of 0.5. Therefore, a higher AUC value indicates that the model has better predictive performance and is more capable of making correct predictions.

Table 2. The performance results

Algorithm	AUC	Accuracy	Classification error	TCT (ms)
Random forest	0.370	45.8%	54.2%	4000
Decision tree	0.485	50%	50.0%	798
Logistic regression	0.736	68.4%	31.6%	686
Generalized linear model	0.734	64%	36.0%	685

Results in the Table 2 show that the AUC values obtained by random forest, decision tree, and gradient boosted Trees are considerably lower than those of logistic regression and generalized linear model. The consistent performance of the AUC values with the accuracy and classification error results indicates the reliability of the machine learning algorithms. The total time taken by all the algorithms to complete the training and testing tasks is impressively short. The training and testing tasks for decision tree, logistic regression, and generalized linear model were completed in less than 1 second. Although random forest took significantly longer to complete, approximately 4 seconds, this can still be considered a good scoring time. Although random forest used a significantly longer time to complete in 4 seconds, it is considerable as a good scoring time.

In addition, gaining an understanding of how whistle-blowing intention attributes impact machine learning prediction is important. The purpose of comparing the correlations weight of each machine learning algorithm is to identify which attributes have the greatest impact on the prediction of whistle-blowing intention. The attributes have been grouped based on TPB and demography, and their respective weight contributions are presented in Table 3.

Table 3. The weights of correlations of each whistle-blowing intention attributes

Attributes	Random forest	Decision tree	Logistic regression	Generalized linear model
TPB				
Injunctive norm	0.168	**0.196**	**0.244**	**0.241**
Descriptive norm	0.080	0.029	0.073	0.005
Affective attitude	**0.225**	0.037	0.179	0.234
Integrity culture	0.061	0.023	0.146	0.148
Self-efficacy	0.072	0.037	0.114	0.162
Fear of retaliation	0.058	0.057	0.064	0.046
Instrumental attitude	0.068	0.018	0.087	0.058
Perceived control ability	0.050	0.044	0.105	0.043
Demography				
Course	0.022	0.032	**0.185**	**0.1267**
CGPA	0.032	0.052	0.042	0.085
Gender	0.027	0.045	0.039	0.072
IPTA/IPTS	0.057	0.036	**0.159**	0.068
Age	0.039	0.007	0.034	0.039

All machine learning models utilized all the selected attributes but they received very small weights of correlation from each of the attributes from the both groups (TPB and demography). Injunctive norm is the best used in most of the models (logistic regression, generalized linear model, decision tree). Affective attitude is the highest in random forest followed by an injunctive norm. Injunctive norm and affective attitude also have the biggest weights of correlation outside the machine learning models (Refer Figure 1) and remain its importance in the machine learning models. The course and gender present much slightly higher weights of contribution in logistics regression but in general, all the demography attributes worked with very low correlations in all the machine learning models.

As most of the attributes from TPB and demography presents very small weights of contributions, it will be useful to get more insight on how each group of attributes worked in the models. Table 4 lists the AUC of each machine learning algorithm that uses a different group of attributes from TPB and demography. The first group used all attributes while the only specific group of TPB and demography were used in the second and third group.

Table 4. The AUC of different group in whistle-blowing intention attributes

Algorithm	All attributes	TPB	Demography
Random forest	0.370	0.334	0.321
Decision tree	0.485	0.411	0.4500
Logistic regression	0.736	0.712	0.536
Generalized linear model	0.734	0.730	0.478

It can be observed from Table 4 that the inclusion of different groups of attributes does not present much impact on the tree-based machine learning models. However, ignoring the TPB attributes have affected the AUC of logistics regression and generalized linear models. In general, all machine learning models can provide better performance with a combination of all attributes.

4. CONCLUSION

This paper presents significant findings of research that concerned academic dishonesty that became more crucial due to the online learning implementation from COVID-19 pandemic. Whistle-blowing intention among students can be useful to educators but how the students can perceive this attitude as important to help their peer learning groups need to be apprehended. Acknowledging that machine learning techniques can be used to support fast and reliable prediction tasks, to identify which algorithms are suitable and which attributes are important in the algorithms is a valuable research initiative for further in-depth analysis. This research will be of great interest to researchers in education technology to expand the findings with different approaches of machine learning and various whistle-blowing perspectives of academic dishonesty.

ACKNOWLEDGEMENTS

We acknowledge the Accounting Research Institute, Universiti Teknologi MARA and Universiti Poly-Tech Malaysia for the full support of this research.

REFERENCES

[1] I. D. A. M. Budhyani, M. Candiasa, M. Sutajaya, and P. K. Nitiasih, "The effectiveness of blended learning with combined synchronized and unsynchronized settings on self-efficacy and learning achievement," *International Journal of Evaluation and Research in Education*, vol. 11, no. 1, pp. 321–332, Mar. 2022, doi: 10.11591/ijere.v11i1.22178.

[2] S. Sukirman, Y. Masduki, S. Suyono, D. Hidayati, H. C. A. Kistoro, and S. Ru'iya, "Effectiveness of blended learning in the new normal era," *International Journal of Evaluation and Research in Education*, vol. 11, no. 2, pp. 628–638, Jun. 2022, doi: 10.11591/ijere.v11i2.22017.

[3] A. M. Maatuk, E. K. Elberkawi, S. Aljawarneh, H. Rashaideh, and H. Alharbi, "The COVID-19 pandemic and E-learning: challenges and opportunities from the perspective of students and instructors," *Journal of Computing in Higher Education*, vol. 34, no. 1, pp. 21–38, Apr. 2022, doi: 10.1007/s12528-021-09274-2.

[4] M. Elhossiny, R. Eladly, and A. Saber, "The integration of psychology and artificial intelligence in e-learning systems to guide the learning path according to the learner's style and thinking," *International Journal of Advanced and Applied Sciences*, vol. 9, no. 12, pp. 162–169, Dec. 2022, doi: 10.21833/ijaas.2022.12.020.

[5] A. D. Alharthi and W. T. Elsigini, "Online learning vs blended learning in developing students' self-regulation at Umm Al-Qura University," *International Journal of Advanced and Applied Sciences*, vol. 9, no. 8, pp. 9–20, Aug. 2022, doi: 10.21833/ijaas.2022.08.002.

[6] H. Albaqawi, "Learning self-efficacy and barriers among students to online learning during the COVID-19 pandemic," *International Journal of Advanced and Applied Sciences*, vol. 9, no. 8, pp. 158–163, Aug. 2022, doi: 10.21833/ijaas.2022.08.020.

[7] S. Jatmika, J. Suwandi, J. T. B. Santoso, F. L. Oktaviana, and M. Karima, "Academic dishonesty on online learning among vocational high school students," *International Journal of Evaluation and Research in Education*, vol. 11, no. 4, pp. 1853–1860, Dec. 2022, doi: 10.11591/ijere.v11i4.22507.

[8] K. Srirejeki, A. Faturokhman, A. Praptapa, and B. S. Irianto, "Academic misconduct: Evidence from online class," *International Journal of Evaluation and Research in Education*, vol. 11, no. 4, pp. 1893–1902, Dec. 2022, doi: 10.11591/ijere.v11i4.23556.

[9] M. Amzalag, N. Shapira, and N. Dolev, "Two sides of the coin: lack of academic integrity in exams during the corona pandemic, Students' and Lecturers' Perceptions," *Journal of Academic Ethics*, vol. 20, no. 2, pp. 243–263, Jun. 2022, doi: 10.1007/s10805-021-09413-5.

[10] T. Achmada, I. Ghozali, and D. Pamungkas, "Detection of academic dishonesty: a perspective of the fraud pentagon model," *International Journal of Innovation, Creativity and Change. www.ijicc.net*, vol. 13, no. 12, pp. 266–282, 2020, [Online]. Available: www.ijicc.net.

[11] R. A. Bernardi, A. C. Landry, E. E. Landry, M. R. Buonafede, and M. E. Berardi, "What actions can be taken to increase whistle-blowing in the classroom?," *Accounting Education*, vol. 25, no. 1, pp. 88–106, Jan. 2016, doi: 10.1080/09639284.2015.1107496.

[12] R. A. Bernardi, M. B. Larkin, L. A. LaBontee, R. A. Lapierre, and N. C. Morse, "Classroom cheating: Reasons not to whistle-blow and the probability of whistle-blowing," in *Research on Professional Responsibility and Ethics in Accounting*, vol. 16, 2012, pp. 201–231.

[13] R. A. Bernardi, E. S. Goetjen, and J. M. Brax, "Whistle-blowing in the classroom: The influence of students' perceptions of whistleblowers," in *Accounting for the Public Interest: Perspectives on Accountability, Professionalism and Role in Society*, Dordrecht: Springer Netherlands, 2014, pp. 247–271.

[14] U. Radulovic and T. Uys, "Academic dishonesty and whistleblowing in a higher education institution: A sociological analysis," *African Journal of Business Ethics*, vol. 13, no. 2, Dec. 2019, doi: 10.15249/13-2-218.

[15] D. Z. Nayır, M. T. Rehg, and Y. Asa, "Influence of ethical position on whistleblowing behaviour: do preferred channels in private and public sectors differ?," *Journal of Business Ethics*, vol. 149, no. 1, pp. 147–167, Apr. 2018, doi: 10.1007/s10551-016-3035-8.

[16] T. D. Miethe and J. Rothschild, "Whistleblowing and the control of organizational misconduct," *Sociological Inquiry*, vol. 64, no. 3, pp. 322–347, Jul. 1994, doi: 10.1111/j.1475-682X.1994.tb00395.x.

[17] G. Lee and X. Xiao, "Whistleblowing on accounting-related misconduct: A synthesis of the literature," *Journal of Accounting Literature*, vol. 41, no. 1, pp. 22–46, Dec. 2018, doi: 10.1016/j.acclit.2018.03.003.

[18] F. Xiao and B. Wong-On-Wing, "Employee sensitivity to the risk of whistleblowing via social media: the role of social media strategy and policy," *Journal of Business Ethics*, vol. 181, no. 2, pp. 519–542, Nov. 2022, doi: 10.1007/s10551-021-04914-0.

[19] T. Iwai, L. Yeung, and R. Artes, "Voice or silence: antecedents of whistleblowing intentions," *RAUSP Management Journal*, vol. 56, no. 2, pp. 186–201, Jul. 2021, doi: 10.1108/RAUSP-06-2020-0126.

[20] K. A. Manaf, "Analysis on the Factors That influence students' perception towards whistleblowing intention," *Journal of Contemporary Social Science Research*, vol. 7, no. 1, pp. 20–26, 2022, [Online]. Available: https://ir.uitm.edu.my/id/eprint/72594/1/72594.pdf.

[21] M. G. Pacilli *et al.*, "Heroes or traitors? Perception of whistleblowers depends on the self-relevance of the group being reported," *Group Processes & Intergroup Relations*, p. 136843022211239, Oct. 2022, doi: 10.1177/13684302221123923.

[22] R. A. Bernardi, S. A. Bilinsky, C. H. Chase, L. D. Giannini, and S. A. MacWhinnie, "Decreasing cheating and increasing whistle-blowing in the classroom: A replication study," *Accounting Ethics Education: Making Ethics Real*, pp. 3–24, 2021.

[23] I. Tammouch, A. Elouafi, S. Eddarouich, and R. Touahni, "Centroid competitive learning approach for clustering and mapping the social vulnerability in Morocco," *International Journal of Advanced and Applied Sciences*, vol. 9, no. 9, pp. 70–77, 2022, doi: 10.21833/ijaas.2022.09.009.

[24] M. Makruf , A. Bramantoro, H. J. Alyamani, S. Alesawi, and R. Alturki., "Classification methods comparison for customer churn prediction in the telecommunication industry," *International Journal of Advanced and Applied Sciences*, vol. 8, no. 12, pp. 1–8, Dec. 2021, doi: 10.21833/ijaas.2021.12.001.

[25] L. Barik, "Data mining approach for digital forensics task with deep learning techniques," *International Journal of Advanced and Applied Sciences*, vol. 7, no. 5, pp. 56–65, 2020, doi: 10.21833/ijaas.2020.05.008.

[26] M. B. Slimene and A. Aljaloud, "Neural network estimation of a photovoltaic system based on the MPPT controller," *International Journal of Advanced and Applied Sciences*, vol. 7, no. 2, pp. 85–90, Feb. 2020, doi: 10.21833/ijaas.2020.02.012.

[27] N. K. W. Chan, A. S. H. Lee, and Z. Zainol, "A framework for predicting employee health risks using ensemble model," *International Journal of Advanced and Applied Sciences*, vol. 8, no. 9, pp. 29–38, 2021, doi: 10.21833/ijaas.2021.09.004.

[28] M. Ciolacu, A. F. Tehrani, R. Beer, and H. Popp, "Education 4.0-Fostering student's performance with machine learning methods," *2017 IEEE 23rd International Symposium for Design and Technology in Electronic Packaging, SIITME 2017 - Proceedings*, vol. 2018-January, pp. 438–443, 2017, doi: 10.1109/SIITME.2017.8259941.

[29] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/science.aaa8415.

[30] R. Dunn, J. Hattie, and T. Bowles, "Using the theory of planned Behavior to explore teachers' intentions to engage in ongoing teacher professional learning," *Studies in Educational Evaluation*, vol. 59, pp. 288–294, Dec. 2018, doi: 10.1016/j.stueduc.2018.10.001.

[31] I. Ajzen, "The theory of planned behaviour: Reactions and reflections," *Psychology & Health*, vol. 26, no. 9, pp. 1113–1127, Sep. 2011, doi: 10.1080/08870446.2011.613995.

BIOGRAPHIES OF AUTHORS

Associate Professor Ts. Dr Suraya Masrom is the head of Machine Learning and Interactive Visualization (MaLIV) Research Group at Universiti Teknologi MARA (UiTM) Perak Branch (https://machinelearningiv.wixsite.com/maliv). She received her Ph.D. in Information Technology and Quantitative Science from UiTM in 2015. She started her career as a lecturer at UTM after receiving her master's degree in computer science from Universiti Putra Malaysia in 2001. She transferred to the Universiti Teknologi MARA (UiTM), Seri Iskandar, Perak, Malaysia, in 2004. She is an active researcher in the meta-heuristics search approach, machine learning, and educational technology. She can be contacted at email: suray078@uitm.edu.my.

Nor Hafiza Abd Samad is Senior Lecturer at the Faculty of Computing and Multimedia, Kolej Universiti Poly-Tech MARA. She received her master's degree in Computer Networking from UiTM in 2014. Her research interest surrounds areas, like computer security, computer architecture and data communication and network. She also actively involves in publications, research and innovation projects activities in other related computer science and information technology. She can be contacted at email: hafiza@kuptm.edu.my.

Dr. Rahayu Abdul Rahman is an Associate Professor at the Faculty of Accountancy, UiTM. She received her Ph.D. 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 at email: rahay916@uitm.edu.my.

Farah Husna Mohd Fatzel is a lecturer at the Faculty of Accountancy, UiTM. She graduated from UiTM with a Master of Accountancy. She has more than 5 years of teaching experience in financial accounting and reporting, management accounting, as well as accounting information system. She is also active in research and publication and her areas of interest include accounting education, financial accounting and reporting, corporate governance and taxation. She can be contacted at email: farahhusna@uitm.edu.my.

Siti Marlia Shamsudin is a senior lecturer at the Faculty of Accountancy, UiTM. She holds a Bachelor’s in Accountancy and Master in Accountancy from UiTM. Her research interest is like taxation, green banking, corporate governance and financial reporting. She can be contacted at email: sitim008@uitm.edu.my.