SEMARAK ILMU PUBLISHING
202103268166(003316878-P)

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Tree-Based Pipeline Optimization Machine Learning in Classifying Whistleblowing of Academic Misconduct

Rahayu Abdul Rahman[1], Suraya Masrom[2,*], Jihadah Ahmad[3], Hafidzah Hashim[1], Evi Mutia[4]

1 Faculty of Accountancy, Universiti Teknologi MARA, Perak Branch, Tapah, Malaysia
2 College of Computing and Informatics, Universiti Teknologi MARA, Perak Branch, Tapah, Malaysia
3 Faculty of Computing and Multimedia, Universiti Poly-Tech Malaysia, Kuala Lumpur, Malaysia
4 Department of Accounting, Faculty of Economic and Business, Syiah Kuala University, Banda Aceh, Indonesia

ARTICLE INFO

Article history:
Received 29 June 2023
Received in revised form 14 December 2023
Accepted 29 December 2023
Available online 30 January 2024

Keywords:
Machine Learning; Tree-Based Pipeline Optimization; Genetic Programming; Whistleblowing; Academic Misconduct

ABSTRACT

The critical issue of academic misconduct is of utmost importance in the field of education and understanding whistleblowing behaviour can be a potential measure to effectively address this issue. This paper highlights the benefits of using the Tree-based Pipeline Optimization (TPOT) framework as a user-friendly tool for implementing machine learning techniques in studying whistleblowing behaviour among students in universities in Indonesia and Malaysia. The paper demonstrates the ease of implementing TPOT, making it accessible to inexpert computing scientists, and showcases highly promising results from the whistleblowing classification models trained with TPOT. Performance metrics such as Area Under Curve (AUC) are used to measure the reliability of the TPOT framework, with some models achieving AUC values above 90%, and the best AUC was 99% by TPOT with a Genetic Programming population size of 40. The paper's main contribution lies in the empirical demonstration and findings that resulted in achieving the optimal outcomes from the whistleblowing case study. This paper sheds light on the potential of TPOT as an easy and rapid implementation tool for AI in the field of education, addressing the challenges of academic misconduct and showcasing promising results in the context of whistleblowing classification.

1. Introduction

With the emergence of Education 5.0 in university-level education, educators are facing greater challenges compared to the era before the revolution of Industrial 4.0 [1-4]. To make the education ecosystem future-ready, educational technology with artificial intelligence (AI) is becoming prevalent in teaching and learning processes [5]. However, implementing AI technology is not just limited to expert computer scientists, but also needs to be accomplished by educators as novice end users. Unfortunately, educators often face difficulties in adapting AI as part of their usual practices [6]. Gaining adequate skills instantly is often impossible due to the varying levels of knowledge in

* *Corresponding author.*
E-mail address: suray078@uitm.edu.my

https://doi.org/10.37934/araset.38.2.165175

computer and programming skills required for implementation. Hence, this paper introduces a user-friendly and efficient implementation of AI, specifically machine learning, to address the significant issue of whistleblowing in the context of academic misconduct.

Academic misconduct, including fraud, forgery of documents, unauthorized use of materials, and plagiarism, has been a critical concern in the education sector for over a decade [7]. As described by researchers in [8], academic dishonesty refers to intentional acts of fraud that undermine the integrity of the educational process. Whistleblowing has proven to be instrumental in uncovering fraud in corporate settings [9,10]. However, its application in the field of education is still lacking, necessitating further research [11]. Understanding students' attitudes and perceptions towards whistleblowing can be valuable for university policymakers in promoting whistleblowing activities. Hence, conducting further research, particularly focusing on advanced data science techniques, is crucial to promote whistleblowing in cases of academic misconduct.

Many educators and researchers in the field of education may lack extensive experience and expertise in data science and machine learning, which can pose challenges when conducting research on topics such as students' attitudes towards whistleblowing in cases of academic misconduct. The complexity of selecting appropriate machine learning algorithms, hyperparameter tuning, and coding variations can be overwhelming for inexpert data scientists who are not well-versed in these areas. As a result, there is a need for easy frameworks that can automate the process of model selection, making machine learning more accessible and feasible for novice users in the education field.

In response to this need, researchers in the field of computing have developed software frameworks that are designed to be easy to use and efficient. For example, studies such as [13-15] have introduced rapid machine learning frameworks with Graphical User Interfaces (GUI) to enhance usability. These frameworks provide a range of machine learning algorithms, including Random Forest, Decision Tree, Support Vector Machine, and Logistic Regression [11], as well as deep learning algorithms for medical applications such as heart stroke stage detection and brain tumour detection [12,13]. However, GUI-based software tools often require significant computing resources, including High-Performance Computing (HPC), which may not be suitable for educators with limited resources. Cloud platforms like Microsoft Azure Machine Learning offer an alternative, but they can be expensive to subscribe. A more cost-effective option is Google Colab, which supports Jupyter Notebook for developing Python codes. However, given the numerous variations of Python codes available, selecting the appropriate machine learning algorithms and hyperparameter tuning can still be a time-consuming and experimental task.

To address this challenge, this paper introduces the Tree-based Pipeline Optimization (TPOT) framework [16], which is not yet widely adopted by non-expert data scientists, including educators. TPOT has shown promising results when evaluated on medical data [17], and it offers an efficient and cost-effective solution for automating the selection of suitable machine learning algorithms and hyperparameter tuning, making it accessible to educators and other non-expert users. This framework can bridge the gap between the potential of machine learning in solving real-world problems and the expertise of educators and researchers in the field of education, enabling them to leverage machine learning for valuable insights into issues such as whistleblowing and academic misconduct.

This paper presents the research framework that initially aims to demonstrate the promising results achieved by the whistleblowing classification models trained with TPOT. The objective of the research is to determine the optimal experimental settings and TPOT hyperparameters for the academic misconduct whistleblowing dataset. The main contribution of this paper is two-folds. Firstly, it introduces an easy and rapid framework for applying machine learning techniques to cases of academic misconduct and whistleblowing, by utilizing the Tree-based Pipeline Optimization (TPOT)

[12]. TPOT automates the laborious steps involved in machine learning while maintaining its effectiveness in detecting relevant cases. It employs Genetic Programming (GP) to search for the best pipeline, streamlining the process.

Secondly, this paper provides an extensive empirical evaluation of machine learning using validation and testing accuracy, as well as precision-recall scores, which are particularly suitable for imbalanced datasets. The dataset used in this study involved real data from students in four universities in Indonesia and Malaysia, resulting in an imbalanced dataset with 61% of records indicating whistleblowing intention and only 39% from the non-whistleblowing class.

The findings presented in this paper have implications for scholars across various domains, particularly in the fields of education and AI. The insights gained from this research can inform policymakers, educators, and researchers in developing strategies to promote whistleblowing activities and address academic misconduct effectively. Furthermore, the TPOT framework can serve as a valuable tool for researchers and practitioners in implementing machine learning techniques for similar cases in other contexts.

2. Methodology

2.1 The Dataset

The dataset comprises of 329 records obtained from the undergraduate students in two universities of Indonesia and two universities of Malaysia, capturing data related to academic misconduct and whistleblowing intention. The machine learning algorithms applied prediction probabilities to classify the class labels of whistleblowing intention as provided in Table 1.

Table 1
Dependent variable

DV	Class label
Whistleblowing intention	1- Whistleblowing 0- Not whistleblowing

The dependent variable or IV is a binary class where the value of class 1 representing whistleblowing intention while 0 for not whistleblowing. The independent variables (IVs) used to construct the whistleblowing classification model, as outlined in Table 2, were adapted from the Theory of Planned Behaviour [18,19].

Table 2
IVs of the whistleblowing classification

IV	Description
Attitude	The student instrumental and affective attitudes towards whistleblowing of academic misconduct.
Social norm	The degree of social norm, injunctive and descriptive towards whistleblowing of academic misconduct.
Behavioural control	Is a construct in psychology that refers to an individual's perception of their ability to control or perform a specific behaviour.
Demography	Gender, age, academic grading (recorded in different columns)

The Theory of Planned Behavior (TPB) is a well-known psychological theory that aims to explain human behaviour, particularly in the context of decision-making and behavioural intentions [20-22]. Instrumental attitudes refer to the cognitive evaluation or beliefs an individual holds about the practical or functional outcomes of a behaviour. It involves the assessment of the advantages or disadvantages, benefits or costs, and the expected outcomes of a particular behaviour. For example, if an individual believes that blowing the whistle on academic misconduct will result in positive outcomes such as protecting academic integrity or preventing further harm, they may have a positive instrumental attitude towards whistleblowing, and this may increase their likelihood of engaging in whistleblowing behaviour.

On the other hand, affective attitudes refer to the emotional or affective responses an individual has towards a behaviour. Affective attitudes are often based on emotions, values, or personal preferences, and they can influence an individual's intention and behaviour through emotional responses. For example, if an individual has a strong emotional aversion towards the act of whistleblowing due to fear of reprisal or negative consequences, they may have a negative affective attitude towards whistleblowing, and this may decrease their likelihood of engaging in whistleblowing behaviour.

Social norm refers to the unwritten rules, expectations, or standards that govern the behaviour of individuals within a particular social group or culture. In the context of whistleblowing of academic misconduct, social norm can influence an individual's attitude towards whistleblowing by shaping their perception of whether whistleblowing is socially accepted or frowned upon. For example, if a social norm within a particular academic community supports and encourages whistleblowing as a responsible and ethical behaviour to maintain academic integrity, it may positively influence an individual's attitude towards whistleblowing, leading to a more favourable attitude.

In the context of whistleblowing of academic misconduct, injunctive norm can influence an individual's attitude towards whistleblowing by shaping their perception of whether whistleblowing is socially approved or disapproved. For example, if an individual believes that their peers, faculty, or academic institution generally approve of whistleblowing as a responsible and necessary action to address academic misconduct, it may positively influence their attitude towards whistleblowing.

Descriptive involves the perception of what others are actually doing or have done in a given situation. In the context of whistleblowing of academic misconduct, descriptive norm can influence an individual's attitude towards whistleblowing by shaping their perception of how common or uncommon whistleblowing is among their peers or within their academic community. For example, if an individual perceives that whistleblowing is a common and accepted practice among their peers or academic community, it may positively influence their attitude towards whistleblowing.

Perceived behavioural control or self-efficacy, is an important factor in the TPB as it can significantly influence an individual's intention and actual behaviour. In the context of whistleblowing of academic misconduct, behavioural control can refer to an individual's perception of their ability to engage in whistleblowing behaviour when faced with a situation involving academic misconduct.

2.2 Experimental Method

Figure 1 illustrates the experimental methodology, which was implemented using Python programming in the Google Colab cloud platform. Each experiment method depicted in Figure 1 was repeated three times, with different split ratios (test-size of 0.2, 0.3, and 0.4) for each execution. Additionally, each execution involved four different population sizes (10, 20, 30, and 40) and three validations from three generations of Genetic Programming (GP). Consequently, the total number of runs amounted to 36, with a total of 108 validations from GP. The average accuracy of validations and testing accuracy from the three split ratio executions were recorded for further analysis using the Kruskal-Wallis statistical significance test [23,24].

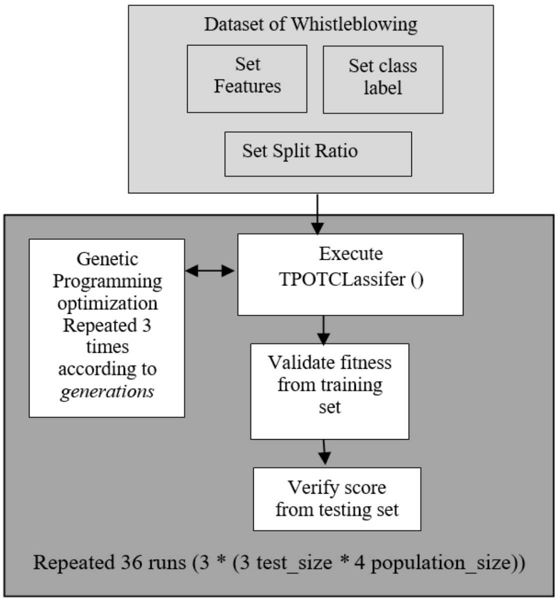

Fig. 1. The experimental method

Figure 2 depicts the Python code for implementing TPOT, a simple and efficient GP automated machine learning tool. The first line is to set the features while the second line is to set the class label as the dependent variable (DV). The "drop" keyword in the first line of code is used to exclude any features, including the DV, that are not required. In this case, all IVs in the dataset were utilized as features for the whistleblowing classification task. The "test size" keyword in the fifth line of code to determine the split ratio between the training and testing datasets from the original dataset. In this research, three different split ratios were utilized: 60:40, 70:30, and 80:20. Once the features are set using "X_train" and the DV is defined using a variable "Y_train", the TPOT classifier can be invoked with appropriate parameters, and the fitness is validated as well as the testing score is printed.

```
X_train = train_df.drop(["DV1WISTLEBLOWING"], axis=1).values
Y_train = train_df["DV1WISTLEBLOWING"]
Features = X_train
Class = Y_train
Feature_Train, Feature_Test, Class_Train, Class_Test  = train_test_split(X_train, Y_train, test_size=0.3)
tpot = TPOTClassifier(generations=3, population_size=40, mutation_rate=0.9,crossover_rate=0.1,verbosity=2, cv=5)
tpot.fit(Feature_Train, Class_Train)
print(tpot.score(Feature_Test, Class_Test))
```

```
Generation 1 - Current best internal CV score: 0.7

Generation 2 - Current best internal CV score: 0.7

Generation 3 - Current best internal CV score: 0.7

Best pipeline: DecisionTreeClassifier(PCA(input_matrix, iterated_power=5, svd_solver=randomized), criterion=entropy, max_depth=3, min_samples_leaf=2,
0.5959595959595959
```

Fig. 2. The Phyton codes to upload data

As shown in Figure 2, when the "test size" is set to 0.3, 230 out of 329 records were utilized as the training set, while the remaining 99 rows of data were used for testing, as illustrated in Figure 3. As shown in Figure 2, when the "test size" is set to 0.3, 230 out of 329 records were utilized as the training set, while the remaining 99 rows of data were used for testing.

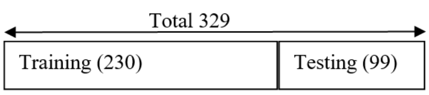

Fig. 3. Training and testing datasets

The TPOT classifier in Figure 2 explores the training dataset starting from the execution code at line 6. If not explicitly provided in the method call statement, TPOT utilizes default Genetic Programming parameters, such as generation, population size, mutation rate, crossover rate, and cross-validation (cv). In this research, the influence of population size on results was observed, with four different population sizes (10, 20, 30, 40) being used and compared. Cross-validation, determined by the "cv" parameter in TPOT, is used to separate the training dataset into training and validation sets at different resampling folds. Figure 4 displays the training and validation sets with a cv value of 5, as set in the codes in Figure 2.

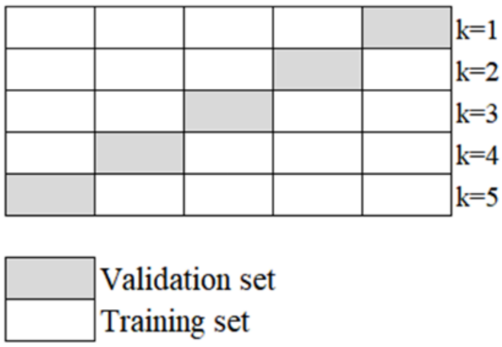

Fig. 4. Training and testing datasets

2.3 Performances Metrics

This research used validation and testing accuracy [25-27] as well as Area Under Curve (AUC) [28] to measure the performances of the TPOT. The accuracy of TPOT can be measured by counting the total correct prediction from the total dataset of validation and testing. As depicted in Figure 1, validation is to denote the fitness of the machine learning in the whistleblowing classification models while score verification is to measure the ability of the machine learning when tested on hold-out data sampling. Hold-out data sampling is the dataset that never been exposed to the machine learning.

Furthermore, AUC is more preferred over accuracy in the assessment of classification model such as for whistleblowing cases. This is because, AUC can present the trade-offs between sensitivity and specificity of the models. Sensitivity or True Positive Rate (TPR) is the number of correct whistleblowing classification from the total of real whistleblowing cases. measures the proportion of actual positive cases that are correctly identified as positive by the model. On the other hand, specificity or False Positive Rate (FPR) measures the proportion of not whistleblowing cases that are incorrectly classified as whistleblowing the model. Figure 5, presents the confusion matrix that can be used to calculate the FPR and TPR for the whistleblowing classification.

	Real class 1 (Whistleblowing)	Real class 0 (Not whistleblowing)
Predicted as 1 (Whistleblowing)	True Positive (TP)	False Positive (FP)
Predicted as 0 (Not whistleblowing)	False Negative (FN)	True Negative (TN)

Fig. 5. The confusion matrix of the classification of whistleblowing intention

TP refers to the cases where the actual class is positive (whistleblowing) and the predicted class is also positive (correctly identified as whistleblowing). FP is the cases where the actual class is negative (not whistleblowing) but the predicted class is positive (incorrectly identified as whistleblowing). TN denotes to the cases where the actual class is negative (not whistleblowing) and the predicted class is also negative (correctly identified as not whistleblowing) while FN refers to the cases where the actual class is positive (whistleblowing) but the predicted class is negative (incorrectly identified as not whistleblowing). Thus, TPR is defined as the ratio of TP to the sum of TP and FN while FRP is defined as the ratio of FP to the sum of FP and TN.

AUC can be observed from Receiver Operating Characteristic (ROC) plot that can depict the curve to map the relationship between the TPR and FPR.

Figure 5 presents the Python codes to plot the ROC and AUC of TPOT. For each "test size", the codes in Figure 2 and Figure 6 were executed four times, with varying settings of the GP "population size" parameter, using values of 10, 20, 30, and 40.

```
from sklearn.metrics import roc_curve
from sklearn import metrics
import numpy as np
import matplotlib.pyplot as plt
plt.figure(0).clf()

pred = tpot.predict_proba(Feature_Test)[:,1]
fpr, tpr, thresh = roc_curve(Class_Test, pred)
fpr, tpr, thresh = metrics.roc_curve(Class_Test, pred)
auc = metrics.roc_auc_score(Class_Test, pred)
plt.plot(fpr,tpr,label="Size=10, auc="+str(auc))
```

Fig. 6. Python codes to plot ROC and print the AUC

3. Results and Discussion

Table 3 lists the 12 set of results from the 4 settings of "population size". Each result was recorded based on the average of the 3 runs on each 3 "test size" and 4 "population size". GP used 3 generations based on the setting in Figure 2, hence the validation result for each "population size" is calculated based on the average from the 3 generations and 3 runs.

Table 3
Average of accuracy

Population size	Average of 3 validations from training set and from 3 runs	Average of testing set from 3 runs
Split Ratio (80:20) / "test size =0.2"		
10	0.67	0.67
20	0.69	0.63
30	0.73	0.62
40	0.68	0.70
Split Ratio (70:30) / "test size =0.3"		
10	0.69	0.67
20	0.67	0.66
30	0.67	0.73
40	0.70	0.70
Split Ratio (60:40) / "test size =0.4"		
10	0.66	0.67
20	0.69	0.64
30	0.67	0.61
40	0.68	0.61

The Kruskal-Walls test indicates that the p-values of all accuracy results from the 36 runs were below 0.05, leading to the rejection of the null hypothesis that the results do not have significant differences and are unreliable. Furthermore, Figure 7, Figure 8 and Figure 9 present the ROC Curve Plot and the AUC from the different size of testing dataset in relation to the "test size" parameter.

The findings demonstrate the effectiveness of TPOT machine learning, with accuracy levels exceeding 60%. Some models achieved accuracy of 70% or higher on the hold-out samples used for testing (results at row 4, 7 and 8). There is no noticeable pattern in the improvement of results based on the machine learning outcomes in relation to the split ratio and GP "population size".

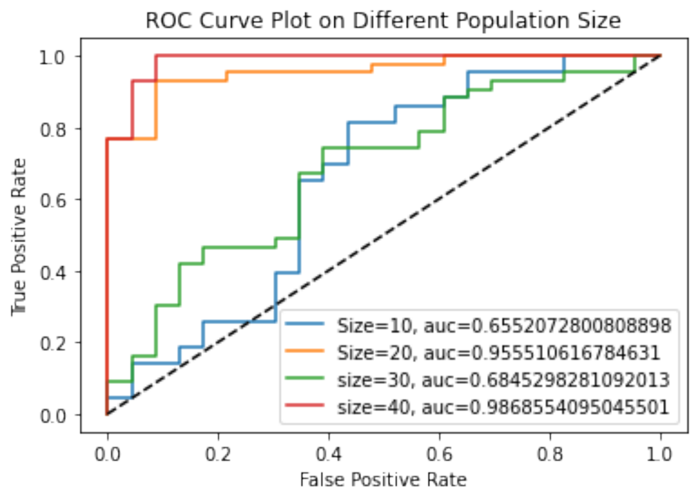

Fig. 7. The ROC-AUC of TPOT with 0.2 testing dataset

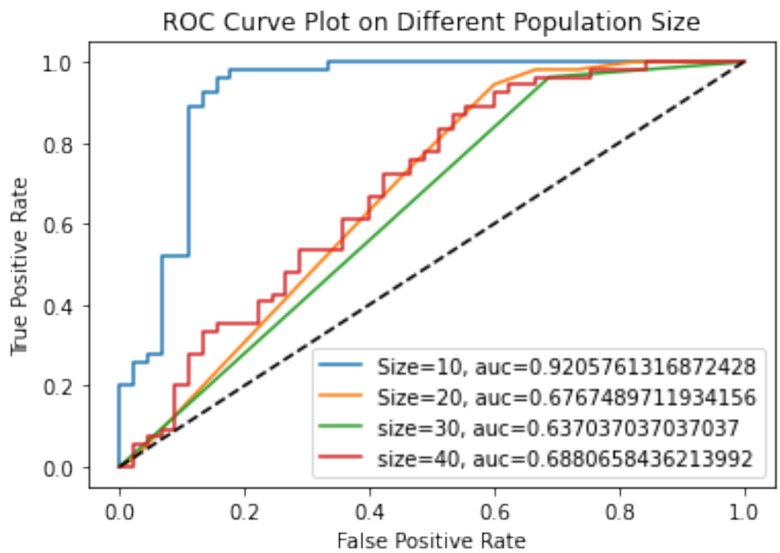

Fig. 8. The ROC-AUC of TPOT with 0.3 testing dataset

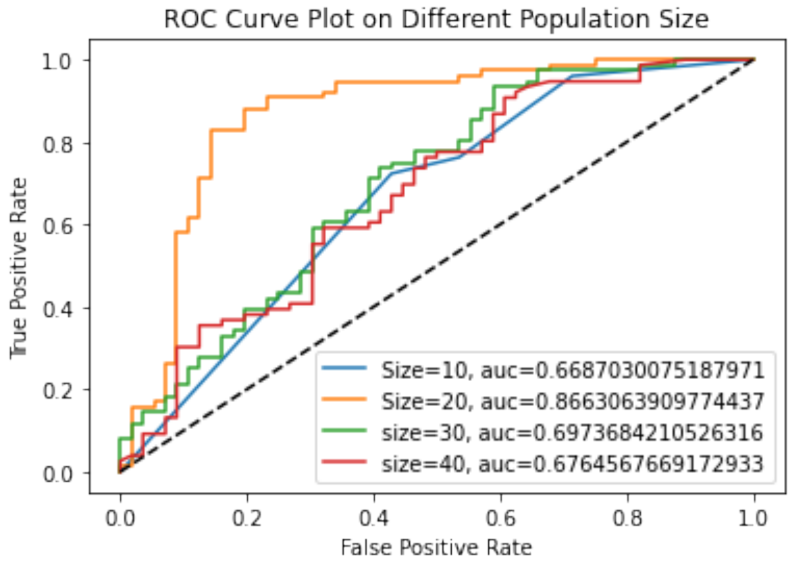

Fig. 9. The ROC-AUC of TPOT with 0.4 testing dataset

Among the 66 records in the 0.2 testing dataset, the best AUC result of 0.99 was achieved by TPOT with a GP population size of 40, slightly lower at 0.96 with a population size of 20. Moderate AUC values ranging from 0.6 to 0.7 were observed with population sizes of 10 and 30. In contrast, for the 0.3 testing dataset (99 records), the highest AUC of 0.92 was obtained with a population size of 10, while other settings yielded reasonable results with AUCs around 0.6. Lastly, with a 0.4 testing dataset (132 records), no models achieved an AUC above 0.9, and the highest AUC of 0.87 was observed with a population size of 20, while other settings yielded similar results around 0.6. Overall, these AUC results suggest that the split ratio with "test size" parameter and "population size" do not exert extreme influence on TPOT machine learning outcomes. This suggests that researchers can generally rely on the default settings when using TPOT for their machine learning tasks, as variations in these parameters do not lead to extreme changes in the performance of the model. This may simplify the implementation process and provide a level of stability in utilizing TPOT for automated machine learning.

4. Conclusions

The main contribution of this research is the novel application of machine learning, particularly TPOT, in predicting the whistleblowing intentions of undergraduate students towards academic misconduct. It introduces the untapped potential of AI technology, specifically machine learning, in transforming the higher education system by addressing the issue of academic misconduct through whistleblowing. By leveraging machine learning, AI-powered chatbots or virtual assistants can be equipped to create anonymous reporting systems, allowing students to report misconduct without fear of retaliation. The proposed TPOT machine learning approach shows promise in identifying key student features that are indicative of whistleblowing intention in combating academic misconduct among peers. It is important to note that the findings of this research are limited to the tested dataset and further exploration of whistleblowing attributes and TPOT machine learning techniques is warranted. While the focus of this research is on the educational domain, the methodology presented in this paper can be replicated by researchers in various domains, making it applicable to a wider range of contexts. One way to improve the accuracy of the model in future work is to collect a larger and more diverse dataset, which could provide the model with a wider range of examples and help it better identify patterns in the data. Additionally, it could be beneficial to include additional features or variables in the model, such as socio-demographic factors or academic performance, which could improve the predictive power of the model.

Acknowledgement
We acknowledge the financial support granted by the Universiti Poly-Tech Malaysia and Accounting Research Institute, Universiti Teknologi MARA for this research.

References

[1] Maatuk, Abdelsalam M., Ebitisam K. Elberkawi, Shadi Aljawarneh, Hasan Rashaideh, and Hadeel Alharbi. "The COVID-19 pandemic and E-learning: challenges and opportunities from the perspective of students and instructors." *Journal of computing in higher education* 34, no. 1 (2022): 21-38. https://doi.org/10.1007/s12528-021-09274-2
[2] Elhossiny, Mohammed, Eladly, Rania and Saber, Abdelnasser. “The integration of psychology and artificial intelligence in e-learning systems to guide the learning path according to the learner’s style and thinking Mohammed,” Int. J. Adv. Appl. Sci., 9(12) (2022): 162–169 https://doi.org/10.21833/ijaas.2022.12.020
[3] Alharthi, Ahmed D., and Waleed T. Elsigini. "Online learning vs blended learning in developing students' self-regulation at Umm Al-Qura University." (2022). https://doi.org/10.21833/ijaas.2022.08.002
[4] Albaqawi, H. "Learning self-efficacy and barriers among students to online learning during the COVID-19 pandemic." *International Journal of Advanced and Applied Sciences* (2022): 158-163. https://doi.org/10.21833/ijaas.2022.08.020
[5] Barik, Lalbihari, Omar Barukab, and A. Ahmed. "Employing artificial intelligence techniques for student performance evaluation and teaching strategy enrichment: An innovative approach." *International Journal of Advanced and Applied Sciences* 7, no. 11 (2020): 10-24. https://doi.org/10.21833/ijaas.2020.11.002
[6] Kaur, Harmanpreet, Harsha Nori, Samuel Jenkins, Rich Caruana, Hanna Wallach, and Jennifer Wortman Vaughan. "Interpreting interpretability: understanding data scientists' use of interpretability tools for machine learning." In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pp. 1-14. 2020. https://doi.org/10.1145/3313831.3376219
[7] McCabe, Donald L., and Linda Klebe Trevino. "Individual and contextual influences on academic dishonesty: A multicampus investigation." *Research in higher education* 38 (1997): 379-396. https://doi.org/10.1023/A:1024954224675
[8] Achmada, Tarmizi, I. Ghozali, I. D. Pamungkas, and U. D. Nuswantoro. "Detection of Academic Dishonesty: A Perspective of the Fraud Pentagon Model." *International Journal of Innovation, Creativity and Change* 13, no. 12 (2020): 266-282.
[9] Vinten, Gerald. "Whistleblowing towards disaster prevention and management." *Disaster Prevention and Management: An International Journal* 9, no. 1 (2000): 18-28. https://doi.org/10.1108/09653560010316032
[10] Brink, Alisa G., D. Jordan Lowe, and Lisa M. Victoravich. "The public company whistleblowing environment: Perceptions of a wrongful act and monetary attitude." *Accounting and the Public Interest* 17, no. 1 (2017): 1-30. https://doi.org/10.2308/apin-51681
[11] Abd Manaf, Khamisah. "Analysis on the factors that influence students’ perception towards whistleblowing intention." *Journal of Contemporary Social Science Research* 7, no. 1 (2022): 45-52.
[12] Olson, Randal S., and Jason H. Moore. "TPOT: A tree-based pipeline optimization tool for automating machine learning." In *Workshop on automatic machine learning*, pp. 66-74. PMLR, 2016.
[13] Abd Rahman, Abdullah Sani, Suraya Masrom, R. Abdul Rahman, and Roslina Ibrahim. "Rapid software framework for the implementation of machine learning classification models." *Int. J. Emerg. Technol. Adv. Eng* 11 (2021): 8-18. https://doi.org/10.46338/ijetae0821_02
[14] Kadtan, Yash Prakash, Aditya Pratap Singh Chauhan, and R. Brindha. "GUI based Prediction of Heart Stroke Stages by finding the accuracy using Machine Learning algorithm." *Annals of the Romanian Society for Cell Biology* (2021): 4571-4577.
[15] Sherawat, Deepika, Sonia, and Akshita Rawat. "Brain tumor detection using machine learning in GUI." In *Proceedings of Integrated Intelligence Enable Networks and Computing: IIENC 2020*, pp. 9-17. Springer Singapore, 2021. https://doi.org/10.1007/978-981-33-6307-6_2
[16] Gijsbers, Pieter, Joaquin Vanschoren, and Randal S. Olson. "Layered TPOT: Speeding up tree-based pipeline optimization." *arXiv preprint arXiv:1801.06007* (2018).

[17] Waring, Jonathan, Charlotta Lindvall, and Renato Umeton. "Automated machine learning: Review of the state-of-the-art and opportunities for healthcare." *Artificial intelligence in medicine* 104 (2020): 101822. https://doi.org/10.1016/j.artmed.2020.101822
[18] Balder, Nazdar, Harcharanjit Singh, Nur Naha Abu Mansor, Ayu Andirah Sakka, Morro Krubally, and Haseebullahet. "The impact of entrepreneurship education on attitude, subjective norm and self-efficacy among undergraduates." *International Journal of Advanced and Applied Sciences* 7, no. 10 (2020): 108-116. https://doi.org/10.21833/ijaas.2020.10.012
[19] Jusoh, N. M. "Empirical study of digital games acceptance in Malaysia: an extended theory planned behavior." *International Journal of Advanced and Applied Sciences* 3, no. 5 (2016): 59-64.
[20] Ru, Xingjun, Haibo Qin, and Shanyong Wang. "Young people's behaviour intentions towards reducing PM2. 5 in China: Extending the theory of planned behaviour." *Resources, Conservation and Recycling* 141 (2019): 99-108. https://doi.org/10.1016/j.resconrec.2018.10.019
[21] Ahmad, Wasim, Woo Gon Kim, Zaheer Anwer, and Weiqing Zhuang. "Schwartz personal values, theory of planned behavior and environmental consciousness: How tourists' visiting intentions towards eco-friendly destinations are shaped?." *Journal of Business Research* 110 (2020): 228-236. https://doi.org/10.1016/j.jbusres.2020.01.040
[22] Razali, Fitriyah, Dzurllkanian Daud, Choong Weng-Wai, and Wilson Rangga Anthony Jiram. "Waste separation at source behaviour among Malaysian households: The Theory of Planned Behaviour with moral norm." *Journal of Cleaner Production* 271 (2020): 122025. https://doi.org/10.1016/j.jclepro.2020.122025
[23] Johnson, Roger W. "Alternate Forms of the One-Way ANOVA F and Kruskal–Wallis Test Statistics." *Journal of Statistics and Data Science Education* 30, no. 1 (2022): 82-85. https://doi.org/10.1080/26939169.2021.2025177
[24] Clark, Jeremy SC, Piotr Kulig, Konrad Podsiadło, Kamila Rydzewska, Krzysztof Arabski, Monika Białecka, Krzysztof Safranow, and Andrzej Ciechanowicz. "Empirical investigations into Kruskal-Wallis power studies utilizing Bernstein fits, simulations and medical study datasets." *Scientific Reports* 13, no. 1 (2023): 2352. https://doi.org/10.1038/s41598-023-29308-2
[25] Adnan, Ahmed Yaseen, Mustaffa Kamal Iwan, and Hannan Mohammed Abdul. "Intelligent control for ship manoeuvering." *Journal of Advanced Research in Applied Mechanics* 67, no. 1 (2020): 1-9. https://doi.org/10.37934/aram.67.1.19
[26] Roslan, Nur Widad, Normaliza Abd Rahim, Nur Maisarah Roslan, and Siti Nur Aliaa Roslan. "Students' presupposition towards incooperating AI (Artifical Intelligence) technology in virtual and face-to-face classes." *International Journal of Advanced Research in Future Ready Learning and Education* 27, no. 1 (2022): 16-19.
[27] Yagoub, Sami Abdelrahman Musa, Gregorius Eldwin Pradipta, and Ebrahim Mohammed Yahya. "Prediction of bubble point pressure for Sudan crude oil using Artificial Neural Network (ANN) technique." *Progress in Energy and Environment* (2021): 31-39.
[28] Halimu, Chongomweru, Asem Kasem, and SH Shah Newaz. "Empirical comparison of area under ROC curve (AUC) and Mathew correlation coefficient (MCC) for evaluating machine learning algorithms on imbalanced datasets for binary classification." In *Proceedings of the 3rd international conference on machine learning and soft computing*, pp. 1-6. 2019. https://doi.org/10.1145/3310986.3311023