**International Journal of Advanced and Applied Sciences**, 10(1) 2023, Pages: 48-54

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

# Machine learning prediction of law enforcement officers' misconduct with general strain theory

Rahayu Abdul Rahman [1], Suraya Masrom [2,*], Jihadah Ahmad [3], Lilis Maryasih [4], Nor Balkish Zakaria [5], Mohd Auzan Md Nor [6]

[1]*Faculty of Accountancy, Universiti Teknologi MARA, Perak Branch, Tapah Campus, Shah Alam, Malaysia*
[2]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Perak Branch, Tapah Campus, Shah Alam, Malaysia*
[3]*Faculty of Computing and Multimedia, Kolej Universiti Poly-Tech MARA, Kuala Lumpur, Malaysia*
[4]*Faculty of Economics and Business, Universitas Syiah Kuala, Acheh, Indonesia*
[5]*Accounting Research Institute, Universiti Teknologi MARA, Selangor, Shah Alam, Malaysia*
[6]*Commercial Crime Investigation Division, Royal Malaysia Police, Perak, Malaysia*

A R T I C L E I N F O

*Article history:*
Received 16 June 2022
Received in revised form
21 September 2022
Accepted 22 September 2022

*Keywords:*
Employee misconduct
Machine learning
Prediction
Police forces

A B S T R A C T

The main objective of this study is to develop a machine learning prediction model on employee misconduct that signals the failure of the integrity of law enforcement officers in performing their duties and responsibilities. Using a questionnaire survey of two hundred eighty-six participants, from senior officers to rank and file police officers, this study presents the fundamental knowledge on the design and implementation of a machine learning model based on four selected algorithms; generalized linear model, random forest, decision tree and support vector machine. In addition to demographic attributes, the performance of each machine learning algorithm on the employee's misconduct has been observed based on the attributes of general strain theory namely financial stress, work stress, leadership exposure, and peer pressure. The findings indicated that peer pressure was the most influencer in the prediction models of all machine learning algorithms. However, random forest is the most outperformed algorithm in terms of prediction accuracy.



## 1. Introduction

Law enforcement bodies including the police institution, play an important role in maintaining and managing public safety. As the representative of the government in offering services to citizens, law enforcement officers must safeguard the public's trust by performing their jobs effectively and in an ethical manner. However, an increasing number of misconduct incidents among law enforcement officers exposed by the media highlight the abuse of their entrusted power (Ferdik et al., 2013), which in turn tarnish their reputations. Indeed, the wrongdoing cases of law enforcement officers signal the failure of integrity in performing their duties and responsibilities to the public.

Police misconduct is an example of the integrity failure of law enforcement agencies and has become a substantial issue across the globe (Ferdik et al., 2013; Weitzer and Tuch, 2004; Wu and Makin, 2021; Ouellet et al., 2019). Lofca (2002) defined police misconduct as inappropriate conduct and illegal actions taken by police officers that violate one or many criminal laws, departmental rules and regulations, and police ethical standards. Examples of police misconduct include false or misleading promises and information, violation of workplace rules, employment discrimination, sexual harassment, corruption, and offering or paying bribes.

Similar to other countries, police misconduct is also one of the main issues faced by the Royal Police of Malaysia (Sia Abdullah and Zamli, 2014), which exposed the institution to public confidence risk. The Royal Malaysia Police (RMP) receives the highest complaints of misconduct from the public among the law enforcement agencies in Malaysia. More than 75 percent of the institution's total misconduct complaints every year involved RMP officers. In addition, the 2021 Corruption Perception Index (TI,

* Corresponding Author.
Email Address: suray078@uitm.edu.my (S. Masrom)
https://doi.org/10.21833/ijaas.2023.01.007
Corresponding author's ORCID profile:
https://orcid.org/0000-0001-7957-8697

2021) measured the public perception of the corruption incidents perpetrated by the officials in a particular nation reports that Malaysia was in the 61st position out of 180 countries and obtained only 52 scores out of 100 (TI, 2021). The lower the CPI scores reflect the higher the public perception of the likelihood of public officials in Malaysia being entangled in corruption.

Ferdik et al. (2013) stressed that police misconduct erodes public trust and confidence in police officers and seriously tarnishes the image of government enforcement agencies. In addition, Hope (2015) argued that police misconduct that compromises the institutional integrity of a policing system and undermines its legitimacy could subsequently damage the country's reputation and weaken the ethical standards of the whole society (DCAF, 2012). According to DCAF (2012), the moral standard of society will be reduced as the public perceives the police as benefiting from unethical misconduct, and corruption, which in turn enhances the public's willingness to engage in crime.

Given the substantial adverse impact of police misconduct, it is important to predict such unethical behavior among the Royal Malaysia Police officers. To date, there are very limited studies on police misconduct in Malaysia except for asset misappropriation (Said et al., 2018) and corruption (Duasa, 2008). Thus, this study aims to expand on the existing body of knowledge by exploring the utilization of a machine learning classification approach for detecting police misconduct of RMP using four attributes of the general strain theory (Agnew, 1992); financial stress, work stress, as well as leadership and peers' exposure.

This study has two main contributions. First, it attempts to extend prior works (Burke, 1995; Cubitt et al., 2020a; Cubitt and Birch, 2021; Cubitt et al., 2020b) by providing evidence on police misconduct prediction models using machine learning in a developing country research setting; Malaysia. Second, it provides another design and implementation of machine learning prediction on police misconduct based on the general strain theory. The remainder of the paper is organized as follows. Section two provides a discussion of prior studies on police misconduct using machine learning algorithms. Section 3 elaborates on the data set of this study and the feature selection process. Section 4 presents and discusses the experimental results for each algorithm. The final section provides the summary and conclusions.

## 2. Literature review

A number of studies (Burke, 1995; Cubitt et al., 2020a; Cubitt and Birch, 2021; Cubitt et al., 2020b) have demonstrated that data mining, as well as artificial intelligence, approaches as alternative methodologies for classification and prediction problems of police misconduct. Burke (1995) was the first study that uses data mining classification techniques to predict police misconduct. The study constructs a police misconduct prediction model using a neural network called Brainmaker for the Chicago Police Department. The prediction model has been designed as an automated early warning system to identify and profile potential police misconduct. Further, the study by Cubitt et al. (2020a) aimed to predict misconduct among Australian police officers. Using 1200 respondents; 600 police officers who have committed instances of serious misconduct, and a matched sample of 600 comparison police officers across a 13-year period, the study constructs predictive models using a machine learning analysis, random forest. The findings show that secondary employment, performance issues as well as demographic variables were important predictors of police misconduct in the Australian setting.

Meanwhile, Cubitt and Birch (2021) used several input factors including officer age, officer gender, detective rank, complainant age, and prior management action to develop a police misconduct prediction model based on a machine learning algorithm, random forest. The sample consists of 3,830 officers who commit serious misconduct in the New York Police Department, between 2000 and 2019. The results show that inexperienced officer attributes; rank and age are significantly affecting the prediction of misconduct among New York police officials. In line with prior studies (Burke 1995; Cubitt et al., 2020a; Cubitt and Birch, 2021), Cubitt et al. (2020b) aimed to develop a robust predictive machine learning model on police misconduct. This study uses a theory of planned behavior's input factors and demographic attributes as predictors of serious police misconduct. The study utilized matched sample data; 600 police officers who have committed serious misconduct from 2005 to 2018, and a control group of 600 Australian police officers. Using a random forest algorithm, the results show that perceived behavioral control is a significant and good predictor for police serious misconduct.

Following previous research (Bishopp et al., 2020; Bishopp et al., 2016), this study utilized the general strain theory's input factors to construct a machine learning prediction model on police misconduct in Malaysia. The GST has been widely used to understand the phenomenon of white-collar crime including employee misconduct. According to the theory, stress/pressure is an important predictor of employee misconduct. As an element of stress/pressure is seen as a significant component of police work, thus GST might provide robust explanations of police behavior and intention towards misconduct. This study uses four stress/pressure indicators namely financial stress, work stress, peer pressure, and leadership pressure.

## 3. Research method

### 3.1. Sample of data

This study utilized a questionnaire survey, where the data was collected from 286 RMP officials. The questionnaire consisted of two sections. Section 1 consists of demographic information including gender, age, race, marital status, education level, work experience, department, and position level. Meanwhile, Section 2 contained four attributes of the general strain theory contributing to employee misconduct. Following prior studies (Reingold 2015; Parrouty, 2014; Arifin and Ahmad, 2017; Men, 2015; Kalshoven et al., 2011; Hart et al., 1994), several indicators have been used to measure each attribute. Each of these attributes was measured on a five-point Likert scale allowing respondents to decide whether they strongly disagree, disagree, neutral, agree, or strongly agree. Estimates for each construct were obtained using the average values of its indicators. The number of items/indicators in Sections 1 and 2 and the source of measurements of the questionnaire are shown in Table 1.

**Table 1:** General strain theory's constructs

| Features/constructs | Indicators | References |
|---|---|---|
| Work stress | 10 items | Reingold (2015) and Parrouty (2014) |
| Financial stress | 8 items | Arifin and Ahmad (2017) |
| Leadership pressure | 10 items | Men (2015) and Kalshoven et al. (2011) |
| Peer pressure | 8 items | Hart et al. (1994) |

### 3.2. Correlations between the independent variable to the dependent variable

Fig. 1 presents the independent variables (IVs) in predicting employee misconduct (dependent variable). Based on the Pearson correlation test, the attributes of general strain theory present a moderately strong correlation from the peer pressure attribute (above 0.5). Other IVs were having very low correlation mainly from the demographic attributes.

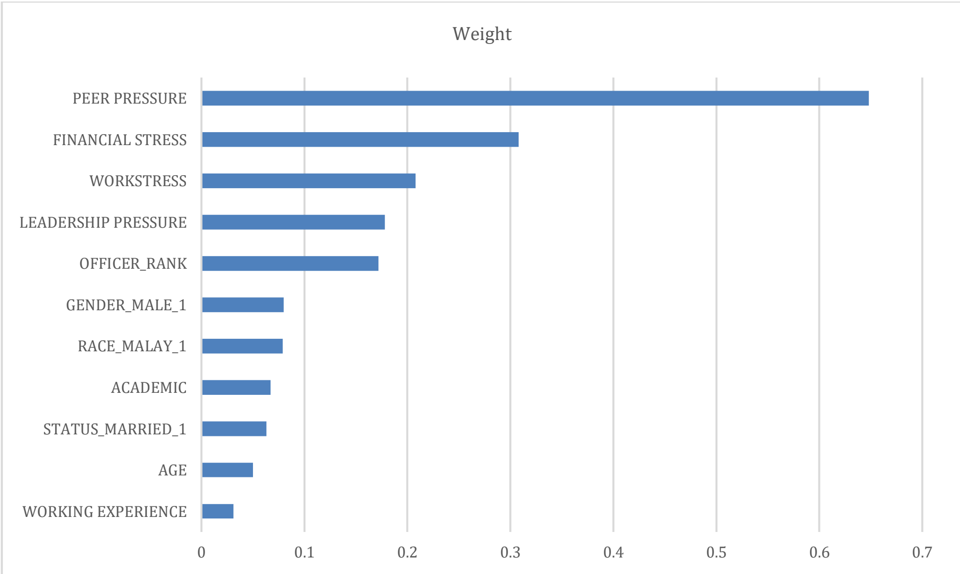


**Fig. 1:** Correlations of each independent variable to the dependent variable

### 3.3. The machine learning algorithms

Four machine learning algorithms namely generalized linear model, random forest, decision tree, and support vector machine (SVM) have been executed in a 16GB computer RAM. Table 2 and Table 3 show the different error rates generated by the decision tree and random forest respectively during the preliminary machine learning hyper-parameters tuning.

The most optimal maximal depth for the decision tree is when the value was set to 15 at the lowest error rate of 17.8%. While in a random forest, the algorithm worked at the most optimum when the maximal depth was 7 with 60 numbers of trees. The setting allowed the algorithm to achieve an 18.4% of error rate. Different from SVM, it used Gamma and C as two important parameters (Table 4). The most optimal setting was 0.5 for Gamma and 100 for C to reach the lowest error rate at 18.1%.

**Table 2:** Decision tree optimal setting

| Maximal depth | Error rates (%) |
|---|---|
| 10 | 21.6 |
| 4 | 19.1 |
| 7 | 18.9 |
| 10 | 18.9 |
| 15 | 17.8 |
| 25 | 18.8 |

**Table 3:** Random forest optimal setting

| Number of trees | Maximal depth | Error rates (%) |
|---|---|---|
| 20 | 2 | 21.2 |
| 60 | 2 | 21.3 |
| 100 | 2 | 21.2 |
| 20 | 4 | 20.0 |
| 60 | 4 | 19.3 |
| 100 | 4 | 19.3 |
| 20 | 7 | 18.6 |
| 60 | 7 | 18.4 |
| 100 | 7 | 18.5 |

For separating the training and testing datasets, the research split the training approach with a ratio of 60:40 percentages. Therefore, from the 100 data, 62 were used for the machine learning training, and 40 were used for the machine learning testing.

## 4. Results and discussion

The results were divided into two. Firstly, the results of performances for the machine learning prediction models of the employee misconduct are given in the following Fig. 2. The performances are R squared depicted in Fig. 2, Root Mean Square Error (RMSE) in Fig. 3, and relative error in Fig. 4.

**Table 4:** SVM optimal setting

| Gamma (RBF) | C | Error rates (%) |
|---|---|---|
| 0.005 | 10 | 19.5 |
| 0.050 | 10 | 20.3 |
| 0.500 | 10 | 18.3 |
| 5 | 10 | 19.2 |
| 0.005 | 100 | 22.7 |
| 0.050 | 100 | 74.5 |
| 0.500 | 100 | 18.1 |
| 5 | 100 | 19.2 |
| 0.005 | 1000 | 71.9 |
| 0.050 | 1000 | 47.8 |
| 0.500 | 1000 | 18.5 |
| 5 | 1000 | 19.2 |

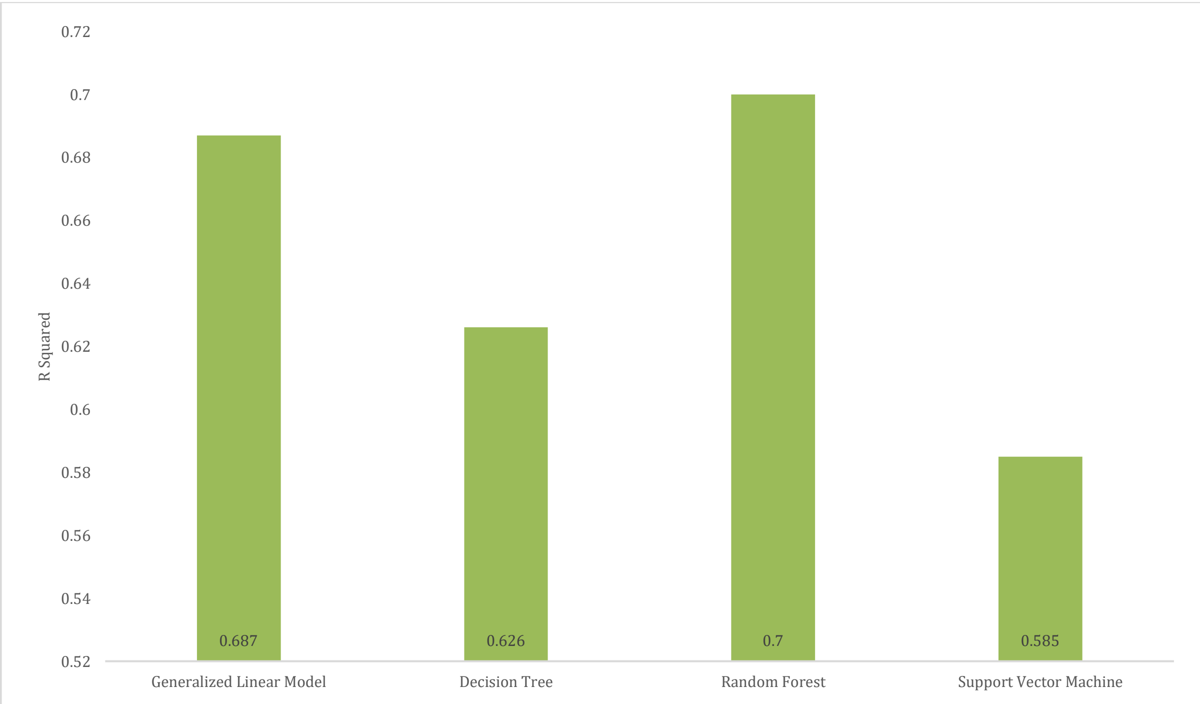


**Fig. 2:** Results of R-squared

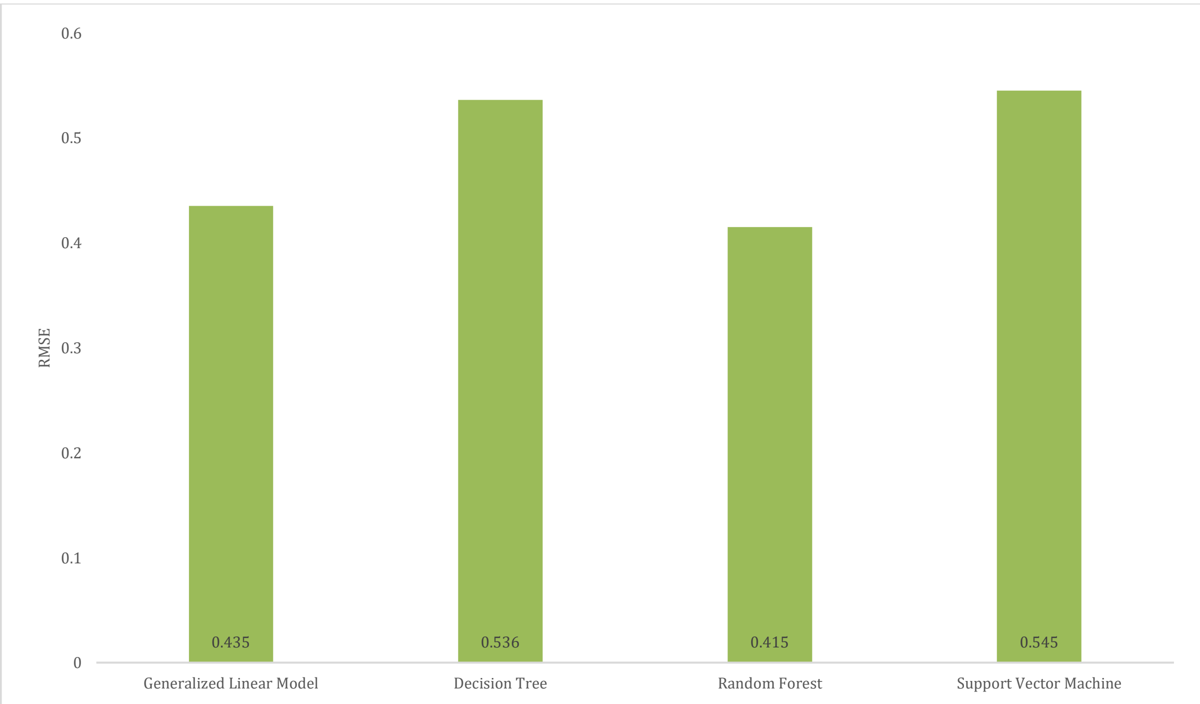


**Fig. 3:** Results of RMSE

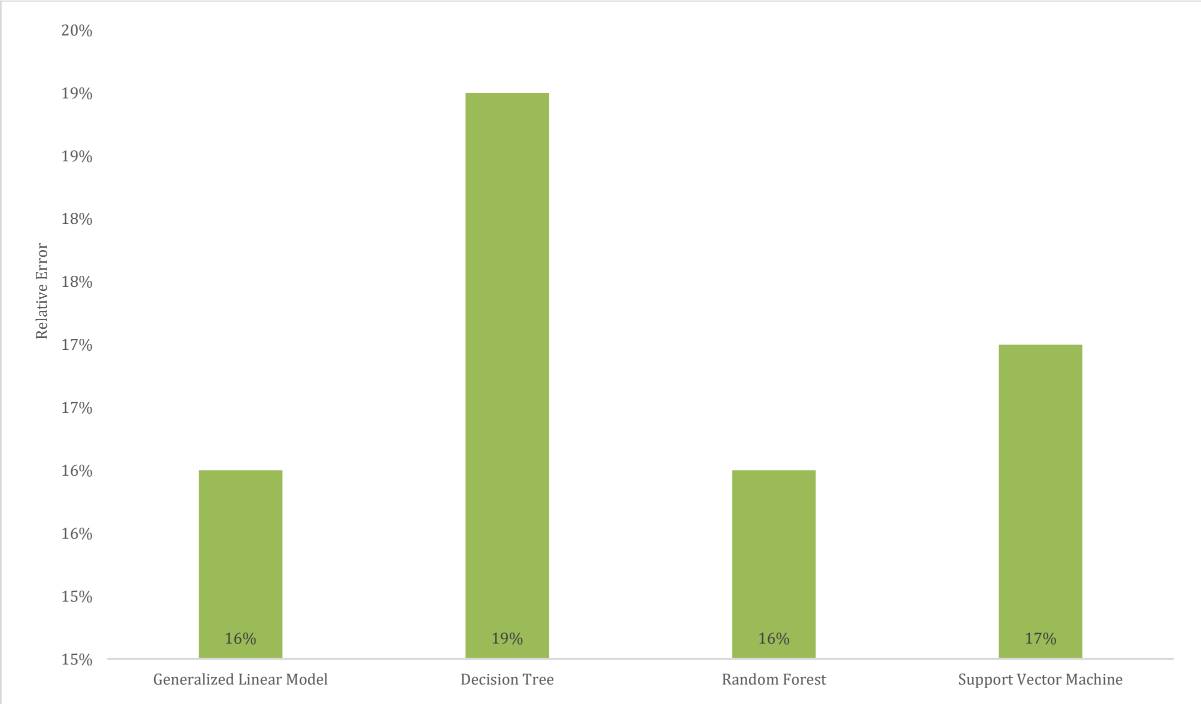


**Fig. 4:** Results of relative error

R squared (R2) presents the proportion of the variance in the prediction model that is explained by the IVs. The highest R squared was generated in a random forest. Besides, the lowest error of root mean square error (RMSE) of 0.415 was also generated by random forest. The relative error, which is defined as the ratio of the absolute error to the actual measurement, has been presented as lower (less than 20%) by all the algorithms. Furthermore, it is interesting to observe how the IVs influenced the prediction model. Fig. 5 presents the comparisons of the IVs' weight in each machine learning prediction model of employee misconduct.

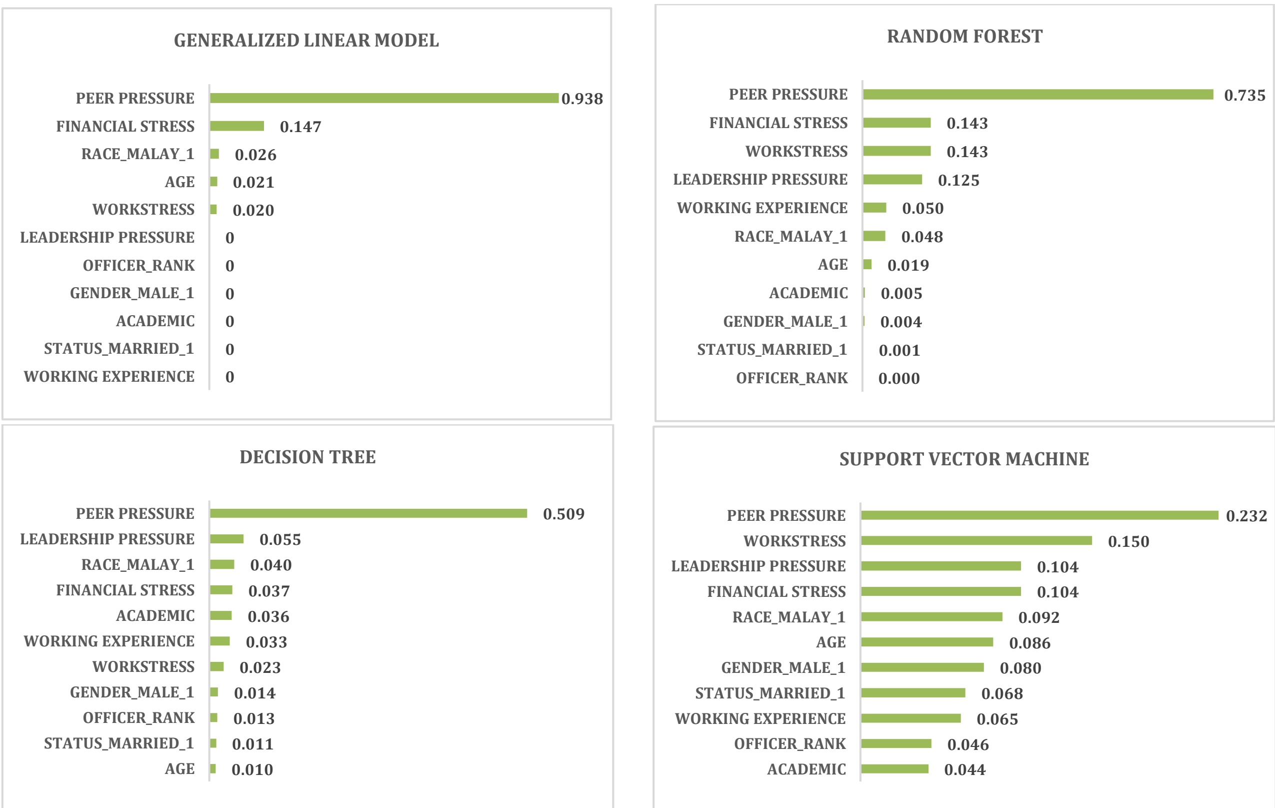


**Fig. 5:** Weights of IVs in the different machine learning algorithms

Peer pressure was the top influencer in the prediction of the employee misconduct model mainly in random forest and generalized linear model. This is consistent with prior research, which found a significant relationship between peer exposure and employees misconducts (Quispe-

Torreblanca and Stewart, 2019). Moreover, peers or a group of people with similar interests, backgrounds, social status, and ethicality provide the individual in the team with valuable information, guidance, and social support. Thus, peers can be served as a model that influences the behaviors and attitudes of others in the group (Palmonari et al., 1991).

All attributes from the general strain theory seem to be beneficial in all algorithms except in the generalized linear model that has zero effect from the Leadership Pressure. Demography attributes have a very low effect in all machine learning models, and extremely have no contribution in the generalized linear model. Fig. 6 and Fig. 7 depict the tree model generated by the decision tree and random forest (a number of trees equal to 60) respectively.

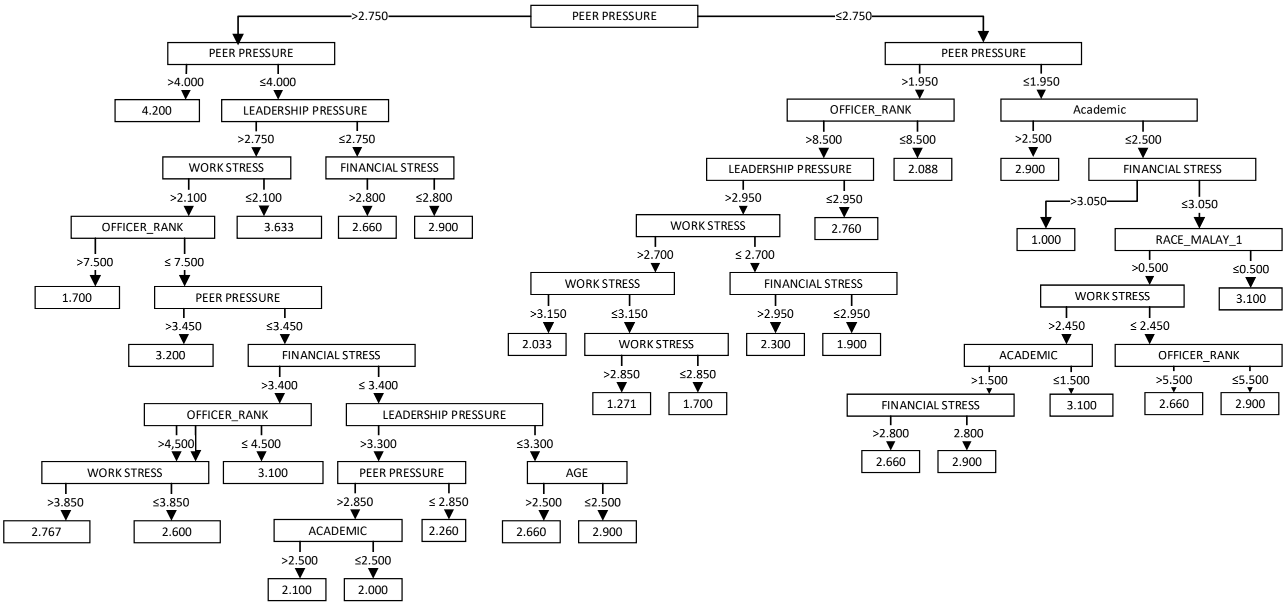


**Fig. 6:** The tree model from the decision tree

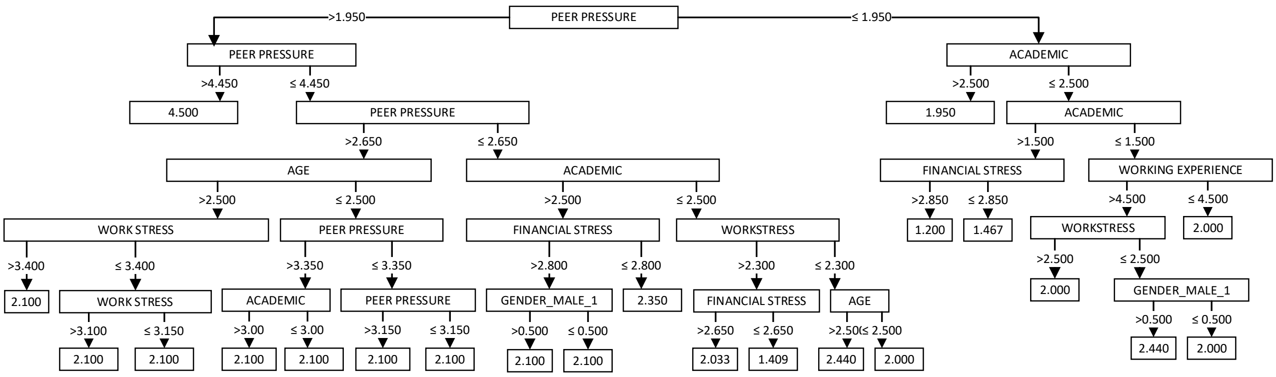


**Fig. 7:** The tree model from random forest

The decision tree and random forest showed that peer pressure has been considered as the important feature in the two models to be the root of the features.

## 5. Conclusion

This study presents the fundamental design and implementation for machine learning prediction of employee misconduct among police officers in Malaysia. The results of machine learning performances and the effect of different attributes from the general strain theory and demographic have been presented in this study. Within the scope of the tested dataset, the attributes of general strain theory were found to give more impact on police misconduct compared to demographic input factors. This study has revealed an interesting finding to be extended in future research works.

## Acknowledgment

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

## Compliance with ethical standards

### Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## References


Agnew R (1992). Foundation for a general strain theory of crime and delinquency. Criminology, 30(1): 47-88. https://doi.org/10.1111/j.1745-9125.1992.tb01093.x

Arifin MAM and Ahmad AH (2017). Peranan whistleblowing dalam meningkatkan integriti anggota polis Diraja Malaysia (Pdrm): Kajian Ke Atas Kontinjen Perak. Sains Insani, 2(1): 17-27. https://doi.org/10.33102/sainsinsani.vol2no1.15

Bishopp SA, Piquero NL, Piquero AR, Worrall JL, and Rosenthal J (2020). Police stress and race: Using general strain theory to examine racial differences in police misconduct. Crime and Delinquency, 66(13-14): 1811-1838. https://doi.org/10.1177/0011128720937641

Bishopp SA, Worrall J, and Piquero NL (2016). General strain and police misconduct: The role of organizational influence. Policing: An International Journal of Police Strategies and Management, 39(4): 635-651. https://doi.org/10.1108/PIJPSM-10-2015-0122

Burke TW (1995). Predicting police misconduct with a neural network program. Law Enforcement Technology, 22(6): 56-58.

Cubitt T, Wooden K, Kruger E, and Kennedy M (2020a). A predictive model for serious police misconduct by variation of the theory of planned behavior. The Journal of Forensic Practice, 22(4): 251-263. https://doi.org/10.1108/JFP-08-2020-0033

Cubitt TI and Birch P (2021). A machine learning analysis of misconduct in the New York Police Department. Policing: An International Journal, 44(5): 800-817. https://doi.org/10.1108/PIJPSM-11-2020-0178

Cubitt TI, Wooden KR, and Roberts KA (2020b). A machine learning analysis of serious misconduct among Australian police. Crime Science, 9(1): 1-13. https://doi.org/10.1186/s40163-020-00133-6

DCAF (2012). The armed forces: Roles and responsibilities in good security sector governance. Democratic Control of Armed Forces, Geneva, Switzerland.

Duasa J (2008). Tendency of corruption and its determinants among public servants: A case study on Malaysia. Available online at: https://mpra.ub.uni-muenchen.de/id/eprint/11562

Ferdik FV, Rojek J, and Alpert GP (2013). Citizen oversight in the United States and Canada: An overview. Police Practice and Research, 14(2): 104-116. https://doi.org/10.1080/15614263.2013.767089

Hart PM, Wearing AJ, and Headey B (1994). Perceived quality of life, personality, and work experiences: Construct validation of the police daily hassles and uplifts scales. Criminal Justice and Behavior, 21(3): 283-311. https://doi.org/10.1177/0093854894021003001

Hope Sr KR (2015). In pursuit of democratic policing: An analytical review and assessment of police reforms in Kenya. International Journal of Police Science and Management, 17(2): 91-97. https://doi.org/10.1177/1461355715580915

Kalshoven K, Den Hartog DN, and De Hoogh AH (2011). Ethical leadership at work questionnaire (ELW): Development and validation of a multidimensional measure. The Leadership Quarterly, 22(1): 51-69. https://doi.org/10.1016/j.leaqua.2010.12.007

Lofca I (2002). A case study on police misconduct in the United States of America and an applicable model for the Turkish National Police. University of North Texas, Ann Arbor, USA.

Men LR (2015). The role of ethical leadership in internal communication: Influences on communication symmetry, leader credibility, and employee engagement. Public Relations Journal, 9(1): 1-22.

Ouellet M, Hashimi S, Gravel J, and Papachristos AV (2019). Network exposure and excessive use of force: Investigating the social transmission of police misconduct. Criminology and Public Policy, 18(3): 675-704. https://doi.org/10.1111/1745-9133.12459

Palmonari A, Kirchler E, and Pombeni ML (1991). Differential effects of identification with family and peers on coping with developmental tasks in adolescence. European Journal of Social Psychology, 21(5): 381-402. https://doi.org/10.1002/ejsp.2420210503

Parrouty J (2014). Stress and burnout. Lulu, Paris, France.

Quispe-Torreblanca EG and Stewart N (2019). Causal peer effects in police misconduct. Nature Human Behaviour, 3(8): 797-807. https://doi.org/10.1038/s41562-019-0612-8 **PMid:31133678**

Reingold L (2015). Evaluation of stress and a stress-reduction program among radiologic technologists. Radiologic Technology, 87(2): 150-162.

Said J, Alam MM, Karim ZA, and Johari RJ (2018). Integrating religiosity into fraud triangle theory: Findings on Malaysian police officers. Journal of Criminological Research, Policy and Practice, 4(2): 111-123. https://doi.org/10.1108/JCRPP-09-2017-0027

Sia Abdullah NA and Zamli S (2014). An analysis of public trust and confidence shown towards the Royal Malaysian Police in the social media. Journal of Media and Information Warfare (JMIW), 6: 11-38.

TI (2021). Corruption perceptions index (CPI). Transparency International, Berlin, Germany.

Weitzer R and Tuch SA (2004). Race and perceptions of police misconduct. Social Problems, 51(3): 305-325. https://doi.org/10.1525/sp.2004.51.3.305

Wu G and Makin DA (2021). The differential role of stress on police officers' perceptions of misconduct. Asian Journal of Criminology, 16(3): 213-233. https://doi.org/10.1007/s11417-020-09324-1