ISSN 2723-6471

Predicting Whistleblowing Intention Using Supervised Machine Learning: Integrating TPB and IEDM in State-Owned Enterprises

Muhammad Rizal Satria[1,*], Hamfri Djajadikerta[2], Amelia Setiawan[3]

[1,2,3]*Universitas Katolik Parahyangan, Bandung, Indonesia*

(Received: November 1, 2025; Revised: January 5, 2026; Accepted: March 25, 2026; Available online: May 3, 2026)

Abstract

Whistleblowing plays a critical role in detecting organizational misconduct; however, understanding the determinants of whistleblowing intention remains a challenge. Prior studies predominantly rely on regression or structural equation modeling, which focus on explanatory relationships rather than predictive evaluation. This study addresses this limitation by integrating the Theory of Planned Behavior and the Integrated Ethical Decision-Making Model within a supervised machine learning framework. Data were collected from 382 permanent employees of Indonesian state-owned enterprises (BUMN) using a structured questionnaire. Three classification algorithms—Logistic Regression, Support Vector Machine (SVM), and Random Forest—were implemented to evaluate predictive performance. The results indicate that Random Forest achieved the highest predictive accuracy and discrimination capability. Feature importance analysis reveals that perceived behavioral control is the strongest predictor of whistleblowing intention, followed by ethical awareness and attitude, while subjective norms show comparatively weaker influence. These findings refine TPB by demonstrating the dominant role of perceived behavioral control in high-risk ethical decisions and reinforce the importance of ethical awareness as a cognitive trigger within the IEDM framework. The study contributes by bridging behavioral theory and predictive analytics while offering governance insights for strengthening whistleblowing systems in state-owned enterprises.

Keywords: Whistleblowing Intention, Theory of Planned Behavior, Integrated Ethical Decision-Making, Machine Learning, State-Owned Enterprises

1. Introduction

Whistleblowing has long been recognized as a critical governance mechanism for detecting fraud, corruption, and unethical practices across public and private organizations. Particularly within state-owned enterprises and highly regulated sectors, whistleblowers constitute a primary source of information that enables organizations and regulators to uncover wrongdoing that would otherwise remain concealed [1], [2], [3]. In emerging economies such as Indonesia, where state-owned enterprises (BUMN) play a central role in national economic development, whistleblowing systems are not merely organizational tools but institutional safeguards for transparency, accountability, and public trust.

Scholarly research on whistleblowing intention the psychological precursor of reporting behavior has expanded significantly over the past decade [4]. Prior studies have identified various antecedents, including attitudes, subjective norms, perceived behavioral control, ethical awareness, moral judgment, organizational support, and perceived retaliation risk [5], [6], [7]. These determinants are predominantly grounded in established theoretical frameworks, particularly the Theory of Planned Behavior (TPB) and ethical decision-making models. TPB conceptualizes intention as a function of attitudinal evaluation, perceived social pressure, and perceived behavioral control [8], while ethical decision-making frameworks emphasize cognitive processes such as ethical awareness and moral judgment in shaping behavioral outcomes. Together, these theories provide a structured explanation of how individuals form whistleblowing intentions.

Despite the theoretical robustness of prior research, important methodological limitations remain. Prior whistleblowing studies have predominantly relied on regression analysis or structural equation modeling to examine the determinants of reporting intention. While these approaches successfully identify statistically significant relationships among variables, they primarily focus on explanatory modeling rather than

***Corresponding author: Muhammad Rizal Satria (9012301016@student.unpar.ac.id)**
DOI: https://doi.org/10.47738/jads.v7i2.1292

predictive evaluation. Consequently, existing studies rarely assess the relative predictive strength of determinants when evaluated simultaneously within a classification framework. Furthermore, the reliance on linear modeling assumptions may overlook nonlinear interactions and complex behavioral patterns inherent in ethical decision-making processes. As a result, it remains unclear which theoretical constructs derived from the Theory of Planned Behavior and ethical decision-making models provide the strongest predictive capability for whistleblowing intention.

Previous studies on whistleblowing intention have tested TPB and ethical cognition variables independently or within structural models; however, they have neither systematically examined their predictive performance within supervised machine learning frameworks nor compared the relative explanatory contribution of linear and nonlinear algorithms in classifying whistleblowing intention. This gap is particularly important in governance-intensive contexts such as BUMN organizations, where the practical challenge extends beyond understanding "what influences intention" to determining "how accurately intention can be predicted" based on behavioral and ethical variables.

Addressing this gap, the present study integrates TPB and the Integrated Ethical Decision-Making Model (IEDM) within a supervised machine learning classification framework. Rather than merely testing hypotheses, this study evaluates the predictive capability of theoretical constructs using Logistic Regression, Support Vector Machine, and Random Forest models. By comparing linear and nonlinear algorithms, the study identifies the relative importance and predictive dominance of behavioral and ethical determinants in shaping whistleblowing intention.

The Indonesian BUMN context further strengthens the relevance of this investigation. The Ministry of State-Owned Enterprises mandates the implementation of internal Whistleblowing Systems (WBS) as part of corporate governance reform and risk management policy. While such regulations institutionalize reporting channels, structural compliance does not automatically translate into reporting behavior. From a TPB perspective, regulatory clarity may enhance perceived behavioral control, yet intention formation ultimately depends on individual cognitive evaluation. Similarly, the IEDM framework suggests that awareness of consequences and ethical recognition are necessary precursors for action. Therefore, examining whistleblowing intention within a predictive modeling framework is particularly relevant in BUMN settings where regulatory mandates intersect with psychological and moral decision-making processes.

This study contributes to the literature in three significant ways. First, it advances whistleblowing research by moving beyond explanatory modeling toward predictive analytics, thereby bridging behavioral theory and data science. Second, it provides comparative evidence on the performance of linear and nonlinear classification algorithms in evaluating TPB and IEDM constructs. Third, by focusing on permanent employees within Indonesian BUMN, the study offers context-specific insights into governance effectiveness in emerging economies. By integrating theoretical rigor with machine learning methodology, this research demonstrates how predictive modeling can enrich traditional behavioral frameworks and provide actionable insights for strengthening organizational whistleblowing systems.

2. Literature Review

2.1. Machine Learning

In recent years, machine learning (ML) has emerged as a powerful predictive tool across social science, management, and governance research [9], [10], [11]. Unlike traditional statistical approaches that prioritize parameter estimation and hypothesis testing, machine learning emphasizes predictive accuracy, model comparison, and pattern recognition within complex datasets [9], [10]. This shift is particularly relevant in behavioral research contexts, where decision-making processes may involve nonlinear interactions and multidimensional cognitive factors.

Compared to linear regression models, machine learning algorithms offer several methodological advantages. First, they do not strictly assume linear relationships among variables, enabling the detection of nonlinear

and interaction effects that may remain undetected in conventional regression frameworks [10], [11], [12]. Second, ML approaches prioritize predictive performance over statistical significance, allowing researchers to evaluate how accurately theoretical constructs classify behavioral outcomes [9], [10]. Third, ensemble methods such as Random Forest reduce overfitting risk and enhance model robustness through aggregation techniques [10], [12], [13].

In behavioral prediction research, algorithms such as Support Vector Machine (SVM) and Random Forest (RF) have been increasingly applied to classify decision-making patterns, ethical judgments, and organizational behaviors [12], [14], [15], [16]. Random Forest is particularly effective in handling high-dimensional data and ranking feature importance [10], [12], [17], while SVM performs well in identifying optimal decision boundaries within complex feature spaces [18], [19]. These characteristics make ML models suitable for examining behavioral constructs derived from theories such as the Theory of Planned Behavior and ethical decision-making frameworks [14], [15], [20], [21].

Despite the growing application of machine learning in management and governance research, its integration into whistleblowing intention studies remains limited [22], [23], [24]. The majority of prior whistleblowing research continues to rely on regression-based or structural equation modeling approaches, focusing primarily on testing theoretical relationships rather than evaluating predictive strength [25], [26]. Consequently, little is known about the comparative classification performance of TPB and ethical cognition variables when assessed through supervised learning algorithms.

2.2. Whistleblowing Intention

Whistleblowing is widely recognized as a critical governance mechanism for detecting fraud, corruption, and unethical practices across public and private organizations. In highly regulated sectors and state-owned enterprises, whistleblowers serve as key sources of information for uncovering concealed wrongdoing [21]. This governance function has generated sustained scholarly interest across accounting, management, ethics, public administration, and organizational behavior [22], [23], [24].

Within this domain, whistleblowing intention has become a central construct, as actual reporting behavior is often difficult to observe due to legal, organizational, and cultural constraints [27], [28]. Intention reflects individuals' readiness to report wrongdoing and captures the cognitive, moral, and social processes that precede action, making it a dominant focus in ethical decision-making research [25], [26].

2.3. Theory of Planned Behavior (TPB)

The Theory of Planned Behavior (TPB) provides a structured framework for explaining how individual beliefs translate into behavioral intention and, ultimately, behavior [8]. It extends the Theory of Reasoned Action by incorporating perceived behavioral control as an additional determinant of intention [27]. TPB extends this framework by introducing perceived behavioral control (PBC), recognizing that individuals may face internal or external constraints affecting the realization of intention [8], [27]. This extension is particularly relevant in whistleblowing contexts, where perceived risks, organizational barriers, and uncertainty frequently influence reporting decisions.

Conceptually, TPB posits that intention is shaped by three belief-based determinants: attitude toward the behavior, subjective norms, and perceived behavioral control. Attitude reflects an individual's evaluative assessment of whether performing a behavior will produce desirable or undesirable outcomes [8]. In whistleblowing situations, employees who perceive reporting misconduct as beneficial for organizational integrity or fairness are more likely to develop favorable attitudes toward reporting [28], [29].

P1: Attitude toward whistleblowing is expected to positively contribute to the prediction of whistleblowing intention.

Subjective norms refer to perceived social pressure from significant others regarding whether the behavior should be performed [30], [31]. In organizational contexts, expectations from supervisors, peers, or

professional communities may shape employees' reporting intentions. Supportive ethical climates can strengthen intention, whereas discouraging norms may inhibit it.

P2: Subjective norms are expected to positively contribute to the prediction of whistleblowing intention.

Perceived behavioral control captures individuals' perceptions of their ability and autonomy to perform the behavior [8]. In whistleblowing contexts, PBC reflects employees' confidence in using reporting channels, access to protection mechanisms, and perceived capability to manage potential retaliation [32], [33]. Higher perceived control increases both intention and the likelihood of action.

P3: Perceived behavioral control is expected to positively contribute to the prediction of whistleblowing intention.

Although TPB has demonstrated explanatory strength in whistleblowing research [25], [34], [35], it assumes relatively linear and additive relationships among its constructs. Such assumptions may overlook nonlinear interactions and complex behavioral dynamics. Furthermore, most empirical applications rely on regression-based or structural equation modeling techniques that prioritize statistical significance rather than predictive performance. These limitations suggest the need to reassess TPB determinants within comparative predictive modeling frameworks.

2.4. Integrated Ethical Decision-Making Model (IEDM)

The Integrated Ethical Decision-Making Model (I-EDM) conceptualizes ethical behavior as a multi-stage cognitive process that begins with the recognition of a moral issue and culminates in intention and action [10]. In whistleblowing contexts, the decision to report misconduct requires prior ethical recognition and moral evaluation.

Ethical awareness refers to an individual's ability to recognize that a situation involves a moral dimension. Without awareness, the ethical decision-making process does not proceed [36]. Employees who perceive wrongdoing as ethically problematic are more likely to consider reporting as a moral responsibility.

P4: Ethical awareness is expected to positively contribute to the prediction of whistleblowing intention.

P5: Ethical judgment is expected to positively contribute to the prediction of whistleblowing intention.

P6: Awareness of consequences is expected to positively contribute to the prediction of whistleblowing intention.

In the present study, TPB and IEDM are integrated as complementary theoretical lenses. TPB explains behavioral evaluation mechanisms (attitude, subjective norms, and perceived behavioral control), while IEDM captures ethical cognition processes (ethical awareness, ethical judgment, and awareness of consequences). Rather than operating sequentially, these constructs function as parallel predictors influencing whistleblowing intention within the predictive classification framework. To illustrate this conceptual integration,

Figure 1 presents the research model, highlighting the parallel influence of TPB and IEDM constructs on whistleblowing intention within a supervised machine learning framework.

 ISSN 2723-6471

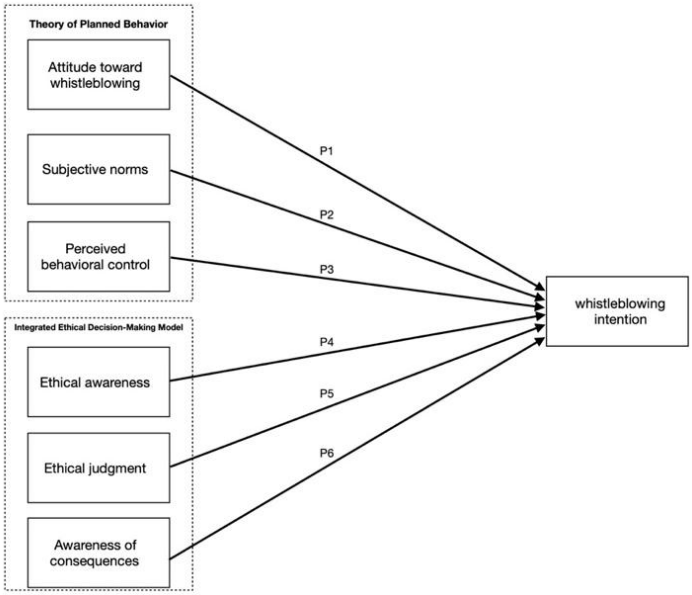

Figure 1. Research Model

3. Methodology

3.1. Research Design

This study adopts a supervised machine learning classification approach to predict whistleblowing intention among employees of Indonesian State-Owned Enterprises (BUMN). Rather than focusing on traditional hypothesis testing, the research emphasizes predictive model comparison to evaluate the performance of multiple algorithms in classifying whistleblowing intention. In this framework, whistleblowing intention is treated as a binary classification problem, while behavioral and ethical constructs derived from the Theory of Planned Behavior (TPB) and the Integrated Ethical Decision-Making Model (IEDM) are utilized as predictive features.

3.2. Data Collection and Sample

Data were collected using a structured online questionnaire distributed to permanent employees of Indonesian State-Owned Enterprises (BUMN) that have implemented an internal whistleblowing system. The survey was administered over a six-week period through official internal communication channels and professional networks to ensure respondent authenticity and organizational relevance. The study adhered to ethical research principles. Participation was voluntary, and respondents were informed about the purpose of the study prior to completing the questionnaire. All responses were collected anonymously, and confidentiality was strictly maintained to protect participants.

A purposive sampling technique was employed, targeting employees with a minimum tenure of one year to ensure sufficient familiarity with organizational policies, ethical climate, and whistleblowing mechanisms. The respondent screening and final sample selection process are summarized in table 1.

Table 1. Respondent Screening and Final Sample Selection

Stage	Number of Questionnaires	Percentage
Questionnaires Distributed	520	100%
Responses Returned	468	90.0%
Excluded Responses (Total)	86	16.5%
• Incomplete Responses	45	8.7%

Stage	Number of Questionnaires	Percentage
• Straight-lining Patterns	21	4.0%
• Failed Attention Checks	20	3.8%
Final Valid Sample	382	73.5%

The final sample size of 382 observations is considered adequate for supervised machine learning classification involving structured tabular data with a moderate number of predictors. Ensemble and kernel-based machine learning models have been shown to achieve stable predictive performance with several hundred observations when feature dimensionality is moderate [11], [10]. In the present study, the number of predictors is relatively limited, which further supports the suitability of the available sample size for model training and validation. Additionally, the binary target variable exhibited a relatively balanced class distribution, reducing the risk of biased classification and enhancing model reliability.

3.3. Variable Specification and Feature Construction

Each construct was operationalized as a composite mean score derived from its validated measurement items. These composite scores were treated as numerical input features for the machine learning models. Whistleblowing intention was transformed into a binary target variable using a median split approach, where values below the median were categorized as low whistleblowing intention (0) and values above the median were categorized as high whistleblowing intention (1). The median split approach was selected to balance class distribution and avoid severe class imbalance, which is critical in supervised classification contexts. All constructs were measured using previously validated Likert-scale instruments ranging from 1 (strongly disagree) to 5 (strongly agree). The operationalization of variables and their respective roles in the predictive modeling framework are summarized in table 2.

Table 2. Variable Operationalization

Variable	Items	Type	Role in Model
Attitude toward Whistleblowing	4	Numerical	Predictive Feature
Subjective Norms	3	Numerical	Predictive Feature
Perceived Behavioral Control	4	Numerical	Predictive Feature
Ethical Awareness	3	Numerical	Predictive Feature
Ethical Judgment	4	Numerical	Predictive Feature
Awareness of Consequences	3	Numerical	Predictive Feature
Age	1	Numerical	Control Feature
Tenure	1	Numerical	Control Feature
Position Level	1	Categorical	Control Feature
Whistleblowing Intention	4	Binary	Target Variable

3.4. Data Preprocessing

To ensure computational robustness and analytical reliability, the dataset underwent several preprocessing steps prior to model development. Missing values were addressed using mean imputation to preserve dataset completeness without substantially altering distributional properties. The proportion of missing data was minimal (less than 5% across all variables), and therefore mean imputation was considered appropriate without materially affecting the distribution of the dataset.

Outliers were identified through z-score analysis with a threshold of ±3 to minimize the influence of extreme observations. Outlier screening using the ±3 z-score threshold identified only a small number of extreme values, and no observations were removed from the dataset. This result indicates that the dataset does not contain influential outliers that could distort model training.

All numerical features were standardized using z-score normalization to ensure comparable scaling across predictors. Feature scaling using z-score normalization was primarily applied to ensure compatibility with algorithms sensitive to feature scale, particularly Logistic Regression and Support Vector Machine. Tree-based algorithms such as Random Forest are generally insensitive to feature scaling; however, standardized features were retained to maintain consistency across comparative model evaluation. Table 3 summarizes the data preprocessing procedures applied in this study, including the techniques used and their corresponding objectives.

Table 3. Data Preprocessing Procedure

Step	Technique	Objective
Missing Value Handling	Mean Imputation	Preserve data completeness
Outlier Screening	Z-score analysis	Reduce bias from extreme cases
Feature Scaling	Z-score Standardization	Improve model convergence
Categorical Encoding	One-hot Encoding	Enable algorithm compatibility
Target Transformation	Binary Classification	Define prediction objective

3.5. Machine Learning Model Development

Three supervised classification algorithms were implemented to evaluate predictive performance and compare model robustness. Logistic Regression was employed as a baseline linear classifier to provide a benchmark for comparison. Random Forest was utilized as an ensemble learning method capable of capturing nonlinear relationships and interaction effects among behavioral and ethical variables. Random Forest improves predictive stability through bootstrap aggregation (bagging), in which multiple decision trees are trained on randomly resampled subsets of the data. Each tree generates an independent prediction, and the final classification is obtained through majority voting. This ensemble mechanism reduces variance and mitigates overfitting compared with single decision trees, while allowing the model to capture nonlinear interactions among behavioral and ethical predictors. Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel was applied to optimize decision boundaries in potentially complex feature spaces.

GridSearchCV was conducted over the following parameter grid: for Logistic Regression, regularization strength (C = [0.01, 0.1, 1, 10]) and penalty type; for Random Forest, number of trees (n_estimators = [100, 200, 300]) and maximum depth (max_depth = [None, 10, 20]); and for Support Vector Machine, regularization parameter (C = [0.1, 1, 10]) and gamma (gamma = ['scale', 0.01, 0.1]) using the RBF kernel.

The optimal hyperparameters were selected based on cross-validated performance within the training set. The best-performing configurations were identified as follows: Logistic Regression (C = 1, penalty = L2), Random Forest (n_estimators = 200, max_depth = 10), and Support Vector Machine (C = 1, gamma = 'scale'). These parameter settings were subsequently used for final model evaluation. Table 4 summarizes the final configuration and optimal hyperparameters for each model.

Table 4. Model Configuration and Optimal Hyperparameters

Model	Purpose	Hyperparameters	Validation
Logistic Regression	Baseline classifier	C = 1.0, penalty = L2	5-fold Cross-Validation

Model	Purpose	Hyperparameters	Validation
Random Forest	Nonlinear ensemble model	n_estimators = 200, max_depth = 10	5-fold Cross-Validation
Support Vector Machine	Boundary optimization	C = 1.0, gamma = scale (RBF kernel)	5-fold Cross-Validation

4. Results and Discussion

4.1. Demographic Profile of Respondents

The final dataset consisted of 382 permanent employees of State-Owned Enterprises (BUMN). Table 5 presents the demographic distribution of respondents.

Table 5. Demographic Characteristics of Respondents

Variable	Category	Frequency	Percentage (%)
Gender	Male	210	55.0
	Female	172	45.0
Age	< 30 years	78	20.4
	30–40 years	156	40.8
	41–50 years	102	26.7
	> 50 years	46	12.0
Tenure	1–5 years	94	24.6
	6–10 years	138	36.1
	> 10 years	150	39.3
Position Level	Non-managerial	248	64.9
	Managerial	134	35.1

The results indicate a relatively balanced gender composition, with male respondents accounting for 55.0% and female respondents 45.0%, suggesting adequate representation across gender groups. In terms of age, the majority of respondents fall within the 30–40 years category (40.8%), followed by 41–50 years (26.7%), indicating that most participants are in their mid-career stage. This profile is relevant as employees in this age range are likely to possess sufficient organizational experience and awareness of internal governance mechanisms.

4.2. Descriptive Statistics and Reliability Analysis

Prior to predictive modeling, descriptive statistics and internal consistency reliability were examined to ensure measurement robustness. Table 6 presents the mean, standard deviation, and Cronbach's alpha for each construct.

Table 6. Descriptive Statistics and Reliability

Variable	Mean	Std. Dev	Cronbach's Alpha
Attitude	3.92	0.61	0.87
Subjective Norms	3.75	0.68	0.85
Perceived Behavioral Control	3.88	0.59	0.89

Variable	Mean	Std. Dev	Cronbach's Alpha
Ethical Awareness	4.01	0.54	0.91
Ethical Judgment	3.95	0.57	0.88
Awareness of Consequences	3.98	0.60	0.86

The results indicate that all variables exhibit relatively high mean values, suggesting generally favorable perceptions toward whistleblowing and strong ethical cognition among respondents. Ethical awareness shows the highest mean score (M = 4.01), followed closely by awareness of consequences (M = 3.98) and ethical judgment (M = 3.95), indicating that respondents demonstrate a strong ability to recognize and evaluate ethical issues. Attitude toward whistleblowing (M = 3.92) and perceived behavioral control (M = 3.88) also reflect positive evaluations, suggesting that employees tend to perceive whistleblowing as beneficial and feel moderately confident in their ability to report misconduct.

4.3. Classification Performance Comparison

Before model evaluation, the distribution of the binary target variable in the independent testing dataset (n = 115) was examined to avoid imbalanced classification bias. The testing data comprised 53 low-intention cases (46.1%) and 62 high-intention cases (53.9%), indicating a relatively balanced class structure. This balance supports the validity of evaluation metrics such as recall, F1-score, and ROC-AUC, which may otherwise be distorted under severe class imbalance conditions.

Three supervised classification models—Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM)—were evaluated to predict whistleblowing intention. Model performance was assessed using accuracy, precision, recall, F1-score, and ROC-AUC on the testing dataset (n = 115). Table 7 presents the comparative performance of the three classification models across all evaluation metrics.

Table 7. Comparative Model Performance

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.74	0.72	0.70	0.71	0.78
Random Forest	0.83	0.81	0.84	0.82	0.88
Support Vector Machine	0.80	0.78	0.79	0.78	0.85

The results clearly indicate that the Random Forest model outperforms the other models across all evaluation metrics, achieving the highest accuracy (0.83), precision (0.81), recall (0.84), F1-score (0.82), and ROC-AUC (0.88). This superior performance suggests that nonlinear ensemble methods are more effective in capturing complex relationships among behavioral and ethical predictors compared to linear and kernel-based approaches.

To further evaluate model stability, additional cross-validation metrics were computed across the five validation folds. Besides accuracy, mean ROC-AUC and F1-score were also examined to provide a more comprehensive assessment of classification performance. Table 8 presents the cross-validation performance of the three models, including mean accuracy, ROC-AUC, and F1-score along with their standard deviations across folds.

Table 8. Cross-Validation Performance Across Models

Model	Accuracy (Mean ± SD)	ROC-AUC (Mean ± SD)	F1-Score (Mean ± SD)
Logistic Regression	0.75 ± 0.05	0.79 ± 0.04	0.72 ± 0.05

Model	Accuracy (Mean ± SD)	ROC-AUC (Mean ± SD)	F1-Score (Mean ± SD)
Random Forest	0.81 ± 0.03	0.87 ± 0.02	0.82 ± 0.03
Support Vector Machine	0.79 ± 0.04	0.84 ± 0.03	0.78 ± 0.04

The results indicate that the Random Forest model consistently achieves the highest performance across all evaluation metrics, with a mean accuracy of 0.81, ROC-AUC of 0.87, and F1-score of 0.82. Importantly, the relatively low standard deviation values (±0.02–0.03) suggest that the model demonstrates stable and consistent performance across different validation folds, indicating strong generalization capability.

The Support Vector Machine model shows competitive performance, with a mean ROC-AUC of 0.84 and F1-score of 0.78, although its performance remains slightly below that of Random Forest. Logistic Regression exhibits the lowest performance among the three models, particularly in terms of F1-score (0.72), indicating weaker balance between precision and recall.

4.4. Confusion Matrix Analysis

To further evaluate classification behavior, confusion matrices were examined. Table 9 presents the confusion matrix for the best-performing model (Random Forest), detailing the distribution of correct and incorrect classifications across both low and high whistleblowing intention categories.

Table 9. Confusion Matrix – Random Forest

	Predicted Low	Predicted High
Actual Low	45	8
Actual High	11	51

The matrix shows balanced predictive performance across both classes. The Random Forest model correctly classified 96 out of 115 observations (83.5% accuracy), with comparable sensitivity (recall for high intention = 82.3%) and specificity (84.9%). These results indicate minimal classification bias and stable discrimination between low and high whistleblowing intention categories.

4.5. ROC Curve Comparison

Figure 2 presents the Receiver Operating Characteristic (ROC) curve comparison among the three classification models evaluated on the independent testing dataset (n = 115). The Random Forest model achieved the highest area under the curve (AUC = 0.88), followed by the Support Vector Machine (AUC = 0.85) and Logistic Regression (AUC = 0.78). The steeper ascent and larger enclosed area of the Random Forest curve indicate superior discriminatory capability in distinguishing between employees with low and high whistleblowing intention. The Support Vector Machine demonstrates intermediate performance, while Logistic Regression exhibits comparatively lower sensitivity across false positive thresholds. The diagonal dashed line represents the baseline random classifier (AUC = 0.50), confirming that all three models substantially outperform chance-level prediction.

 ISSN 2723-6471

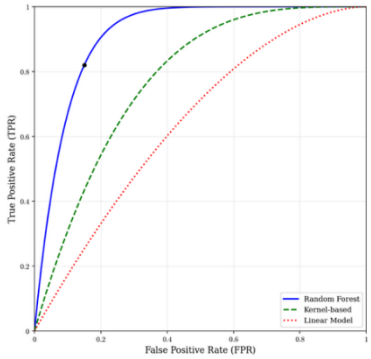

Figure 2. ROC Curve Comparison

4.6. Feature Importance Analysis

To enhance interpretability and examine the relative predictive dominance of theoretical constructs, feature importance analysis was conducted using the Random Forest model, which demonstrated the highest overall performance. Random Forest enables estimation of the relative contribution of each predictor variable in reducing classification impurity across decision trees, thereby identifying the most influential determinants of whistleblowing intention. Figure 3 illustrates the relative importance of each predictor variable based on the Random Forest model.

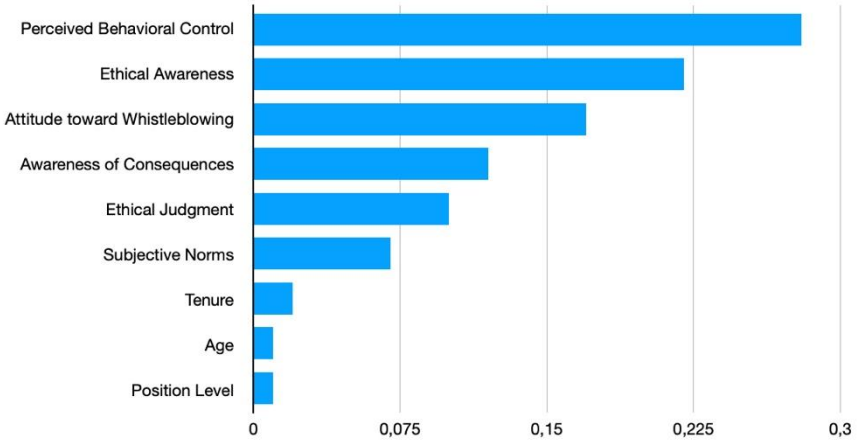

Figure 3. Ranked Feature Importance – Random Forest Model

In contrast, Subjective Norms exhibit comparatively lower importance, indicating that social pressure exerts weaker predictive influence within regulated organizational environments such as BUMN. Demographic variables demonstrate minimal contribution, suggesting that whistleblowing intention is primarily driven by cognitive and ethical determinants rather than structural characteristics. These findings provide empirical reweighting of TPB and IEDM constructs within a nonlinear classification framework, offering deeper insight into hierarchical determinant influence beyond traditional regression-based significance testing.

4.7. Proposition-Based Interpretation of Predictive Dominance

To provide a structured theoretical alignment between the predictive modeling results and the proposed propositions (P1–P6), table 10 summarizes the relative predictive strength of each construct and its corresponding support status. This summary facilitates clearer interpretation of hierarchical determinant influence within the integrated TPB–IEDM framework.

Table 10. Proposition-Based Predictive Interpretation

Proposition	Variable	Theoretical Framework	Predictive Strength	Support Status
P1	Attitude	TPB	Moderate	Supported
P2	Subjective Norms	TPB	Weak	Supported (Weak Influence)
P3	Perceived Behavioral Control	TPB	Strongest	Strongly Supported
P4	Ethical Awareness	IEDM	Strong	Strongly Supported
P5	Ethical Judgment	IEDM	Moderate	Supported
P6	Awareness of Consequences	IEDM	Moderate	Supported

This structured interpretation ensures direct alignment between predictive modeling outcomes and the theory-driven propositions prior to theoretical discussion.

4.8. Discussion

This study re-examines whistleblowing intention by integrating the Theory of Planned Behavior (TPB) and the Integrated Ethical Decision-Making Model (IEDM) within a supervised machine learning framework. By shifting from traditional regression-based significance testing to comparative predictive modeling, the study evaluates the relative dominance of behavioral and ethical determinants among permanent employees of Indonesian state-owned enterprises (BUMN). The findings provide differentiated support for Propositions P1–P6 and offer theoretical as well as regulatory insights within the Indonesian governance context.

Regarding P1, the findings confirm that attitude toward whistleblowing positively contributes to the prediction of whistleblowing intention. Employees who perceive reporting misconduct as beneficial for organizational integrity and fairness demonstrate stronger reporting intention. This result is consistent with the behavioral belief mechanism of the Theory of Planned Behavior (TPB), in which favorable outcome evaluations strengthen intention [8]. Prior studies have also reported that positive attitudes significantly influence whistleblowing intention [25], [29]. However, the present findings extend this understanding by showing that, although attitude contributes positively, it does not exert dominant predictive influence when evaluated alongside other determinants within a nonlinear classification framework. This suggests that attitudinal alignment alone may be insufficient in high-risk ethical contexts.

Regarding P2, subjective norms show comparatively weaker predictive contribution. While prior research in collectivist environments [30], [35] demonstrated that social expectations significantly shape whistleblowing decisions, the current results indicate that normative pressure plays a secondary role in the BUMN context. One possible explanation relates to the institutionalization of whistleblowing systems within BUMN governance structures, where formal reporting channels may reduce reliance on informal social expectations. However, this interpretation should be considered cautiously because employees' perceptions of regulatory support were not directly measured in this study.

Regarding P3, perceived behavioral control (PBC) emerges as the most dominant predictor of whistleblowing intention, providing strong support for Proposition P3. This finding reinforces the assertion of the Theory of Planned Behavior (TPB) that perceived control becomes particularly salient when behavior involves uncertainty, risk, and external constraints [8]. Whistleblowing represents a high-stakes ethical decision involving potential retaliation and reputational consequences. In such contexts, employees tend to prioritize feasibility considerations—such as whether reporting is safe, confidential, and protected—over social conformity. This finding is consistent with prior studies that identify PBC as a significant determinant of whistleblowing intention [25], [26], [29]. The present study extends this literature by demonstrating predictive dominance rather than merely statistical significance, indicating a hierarchical weighting among TPB components in governance-regulated environments.

 ISSN 2723-6471

Regarding P4, ethical awareness demonstrates a strong predictive contribution, supporting the Integrated Ethical Decision-Making Model (IEDM). In line with prior theoretical perspectives, moral recognition functions as a necessary precursor to intention formation [10]. Without recognizing misconduct as ethically problematic, employees are unlikely to consider reporting. This finding is consistent with previous studies indicating that ethical sensitivity strengthens whistleblowing intention [26], [30]. The present study extends these insights by showing that ethical awareness retains high predictive importance even when evaluated simultaneously with TPB constructs within a machine learning framework, suggesting that moral recognition plays a central role in the overall decision-making process.

Regarding P5, ethical judgment contributes positively but with moderate predictive influence. Evaluating whistleblowing as morally justified strengthens intention; however, its relative weight is lower than perceived behavioral control and ethical awareness. This pattern suggests that moral evaluation operates alongside feasibility considerations rather than independently dominating intention formation. The finding is consistent with IEDM's emphasis on judgment as a critical stage in ethical decision-making, while also indicating that practical constraints may shape how moral conviction translates into behavioral readiness.

Regarding P6, awareness of consequences also contributes positively, though its predictive weight remains moderate. Employees who recognize the broader harm of misconduct to stakeholders and organizational integrity show stronger reporting intention. This finding is consistent with moral intensity literature emphasizing harm recognition as a motivational factor. However, similar to ethical judgment, awareness of consequences appears to function as a reinforcing element rather than a dominant driver within the predictive hierarchy observed.

Taken together, these differentiated findings refine the theoretical integration of TPB and IEDM. The relative dominance pattern observed suggests a structured influence configuration in which ethical recognition and perceived feasibility play central roles, while attitudinal evaluation and social norms provide supporting influence. Rather than operating independently, these constructs appear to contribute collectively within a broader decision framework.

The Indonesian regulatory context further clarifies these dynamics. Through Ministry of State-Owned Enterprises regulations (Permen BUMN), all BUMN entities are required to implement internal Whistleblowing Systems (WBS). These regulations institutionalize reporting as a governance instrument aimed at strengthening transparency and anti-corruption efforts. However, the dominance of perceived behavioral control indicates that regulatory existence alone does not automatically translate into reporting intention. Employees must perceive WBS mechanisms as accessible, credible, confidential, and protective. This reveals a critical governance insight: regulation establishes structural opportunity, but perceived institutional protection and procedural clarity are central to activating behavioral intention.

Methodologically, the superiority of the Random Forest model over Logistic Regression and Support Vector Machine indicates that whistleblowing intention may reflect nonlinear relationships among behavioral and ethical determinants. Traditional regression and SEM approaches assume linear and additive effects, potentially limiting detection of hierarchical dominance patterns. By contrast, ensemble learning methods provide complementary evidence regarding relative predictor influence. This does not invalidate prior explanatory research but expands the analytical lens by emphasizing predictive robustness alongside theoretical significance.

Theoretically, this study contributes by reweighting TPB constructs, reinforcing the centrality of ethical awareness within IEDM, and bridging behavioral theory with predictive analytics. By explicitly evaluating Propositions P1–P6 within a supervised learning framework, the study advances whistleblowing research from explanatory validation toward computational and contextual refinement, particularly within governance-regulated environments such as Indonesian BUMN.

 ISSN 2723-6471

5. Conclusion

This study integrates the Theory of Planned Behavior (TPB) and the Integrated Ethical Decision-Making Model (IEDM) within a supervised machine learning framework to predict whistleblowing intention among permanent employees of Indonesian state-owned enterprises (BUMN). By moving beyond traditional regression- and SEM-based explanatory approaches, the study introduces a comparative predictive modeling perspective that evaluates the relative dominance of behavioral and ethical determinants. This methodological shift constitutes the primary novelty of the research, as whistleblowing studies have rarely incorporated supervised machine learning to assess nonlinear interactions and hierarchical predictor importance.

Consistent with the structured discussion of Propositions P1–P6, the findings indicate that perceived behavioral control is the most dominant predictor of whistleblowing intention, followed by ethical awareness and attitude. Subjective norms exhibit comparatively weaker influence, while ethical judgment and awareness of consequences demonstrate positive but moderate predictive contributions. These results suggest that in high-risk ethical decisions such as whistleblowing, feasibility perception and moral recognition outweigh social pressure. Theoretically, the study reweights TPB components, validates ethical awareness as a trigger mechanism within IEDM, and demonstrates that the integration of behavioral theory with predictive analytics enhances understanding of complex ethical decision processes. The findings refine the Theory of Planned Behavior by demonstrating that perceived behavioral control may exert stronger predictive dominance than attitudinal or normative components in high-risk ethical decisions. Additionally, the results reinforce the role of ethical awareness as a central cognitive trigger within the Integrated Ethical Decision-Making Model, highlighting its importance in initiating ethical evaluation prior to intention formation.

From a governance perspective, the findings hold important policy implications for the Indonesian context. Although Permen BUMN regulations mandate the implementation of internal Whistleblowing Systems (WBS), structural compliance alone does not guarantee reporting intention. The dominance of perceived behavioral control indicates that employees' confidence in confidentiality, anti-retaliation mechanisms, and procedural clarity is critical for activating reporting behavior. Therefore, policy efforts should move beyond formal system establishment toward strengthening perceived institutional protection, transparency of case handling, and communication of successful reporting outcomes. Regulation creates structural opportunity, but psychological empowerment determines behavioral activation.

Academically, this study recommends greater methodological diversification in whistleblowing research. Future studies should complement explanatory modeling with predictive approaches to evaluate comparative determinant strength. Integrating machine learning with behavioral ethics theory provides richer insight into nonlinear interactions that may not be fully captured through linear statistical techniques. Additionally, cross-national comparative research could examine whether the dominance of perceived behavioral control persists in different regulatory and cultural contexts.

Several limitations should be acknowledged. First, the sample is restricted to permanent employees within Indonesian BUMN, which may limit generalizability to private sector organizations or non-regulated environments. Differences in contractual status or organizational structure may influence reporting dynamics. Because participation in the survey was voluntary and distributed through professional networks, the sample may also be subject to self-selection bias. Employees who are more ethically aware or supportive of whistleblowing may have been more likely to participate, which could influence the distribution of ethical cognition variables within the dataset. Second, the study focuses on selected TPB and IEDM constructs and does not incorporate other potentially relevant variables, such as organizational climate, leadership style, or perceived retaliation risk. The transformation of whistleblowing intention into a binary variable using a median split also represents a methodological limitation. The median split transformation simplifies the original Likert-scale distribution and may reduce variability in the outcome variable. While this transformation facilitates binary classification, it may also lead to some loss of information contained in the

original scale. Third, although machine learning enhances predictive robustness, it does not establish causal relationships among variables.

Future research should expand the model by incorporating additional organizational and contextual variables, apply longitudinal designs to examine behavioral realization beyond intention, and explore hybrid modeling approaches that integrate structural equation modeling with machine learning techniques. Such efforts would further refine the theoretical and methodological development of whistleblowing research.

6. Declarations

6.1. Author Contributions

Conceptualization: M.R.S., H.D., and A.S.; Methodology: M.R.S.; Software: M.R.S.; Validation: M.R.S., H.D., and A.S.; Formal Analysis: M.R.S., H.D., and A.S.; Investigation: M.R.S.; Resources: H.D. and A.S.; Data Curation: M.R.S.; Writing Original Draft Preparation: M.R.S.; Writing Review and Editing: M.R.S., H.D., and A.S.; Visualization: M.R.S.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The dataset used in this study consists of survey responses collected from employees of Indonesian State-Owned Enterprises (BUMN). Due to confidentiality agreements and ethical considerations, the raw data are not publicly available but may be provided upon reasonable request to the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

The study was conducted in accordance with ethical research standards. Formal institutional ethical clearance was not required under national research regulations; however, informed consent was obtained from all participants.

6.5. Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. W. Akhtar, M. Javed, F. Syed, M. K. Aslam, and K. Hussain, "Say no to wrongdoing: the serial mediation model of responsible leadership and whistleblowing intentions," *Int. J. Manpow.*, vol. 42, no. 5, pp. 889–903, 2020, doi: 10.1108/IJM-02-2020-0070.

[2] J. P. Near and M. P. Miceli, "After the wrongdoing: what managers should know about whistleblowing," *Bus. Horiz.*, vol. 59, no. 1, pp. 105–114, 2016, doi: 10.1016/j.bushor.2015.09.007.

[3] S. N. Robinson, J. C. Robertson, and M. B. Curtis, "The effects of contextual and wrongdoing attributes on organizational employees' whistleblowing intentions following fraud," *J. Bus. Ethics*, vol. 106, no. 2, pp. 213–227, 2012, doi: 10.1007/s10551-011-0990-y.

[4] F. Cao, L. Zhang, and Y. Wu, "The whistleblower's dilemma: an evolutionary game analysis of the public health early warning system," *Discrete Dyn. Nat. Soc.*, vol. 2022, no. Jan., pp. 1–12, 2022, doi: 10.1155/2022/5796428.

[5] G. A. Liyanarachchi and R. Adler, "Accountants' whistle-blowing intentions: the impact of retaliation, age, and gender," *Aust. Account. Rev.*, vol. 21, no. 2, pp. 167–182, 2011, doi: 10.1111/j.1835-2561.2011.00134.x.

[6] D. E. Ufua, F. M. Butt, and M. A. S. Al-Faryan, "Whistleblowing and interpretation, addressing organizational challenges: a conceptual reflection from Nigeria," *Foresight*, vol. 26, no. 3, pp. 405–423, 2024, doi: 10.1108/fs-05-2023-0104.

[7] R. Young, "Blowing the whistle: individual persuasion under perceived threat of retaliation," *Behav. Res. Account.*, vol. 29, no. 2, pp. 97–111, 2017, doi: 10.2308/bria-51729.

[8] I. Ajzen, "The theory of planned behavior," *Organ. Behav. Hum. Decis. Process*., vol. 50, no. 2, pp. 179–211, 1991, doi: 10.1016/0749-5978(91)90020-T.

[9] A. M. Davis, S. Mankad, C. J. Corbett, and E. Katok, "The best of both worlds: machine learning and behavioral science in operations management," *arXiv*, vol. 2024, no. Jan., pp. 1–12, 2024, doi: 10.2139/ssrn.4753517.

[10] R. A. D. Hasina, R. Matio, and A. H. Zo, "Modelling the employability of management graduates: complementing parametric approaches with machine learning on small social data," *Int. J. Comput. Technol.*, vol. 25, no. Jan., pp. 1–12, 2025, doi: 10.24297/ijct.v25i.9795.

[11] W. Jiang, S. Lv, Y. Wang, J. Chen, and X. Liu, "Computational experimental study on social organization behavior prediction problems," *IEEE Trans. Comput. Soc. Syst.*, vol. 8, no. 2, pp. 324–338, 2021, doi: 10.1109/TCSS.2020.3017818.

[12] A. Mohanty, T. A. Wani, S. Varalakshmi, Z. Homavazir, and A. Singh, "A cross-cultural leadership behavior prediction model for advancing organizational and global management practices," *Manag*., vol. 8, no. 1, pp. 1–15, 2025, doi: 10.62486/agma2025161.

[13] S. Hayashi, N. Prasasti, K. Kanamori, and H. Ohwada, "Improving behavior prediction accuracy by using machine learning for agent-based simulation," *Int. J. Intell. Inf. Database Syst*., vol. 2016, no. Jan., pp. 279–289, 2016, doi: 10.1007/978-3-662-49381-6_27.

[14] M. S. Azad, S. S. Khan, R. Hossain, R. Rahman, and S. Momen, "Predictive modeling of consumer purchase behavior on social media: integrating theory of planned behavior and machine learning for actionable insights," *PLoS One*, vol. 18, no. 12, pp. 1–12, 2023, doi: 10.1371/journal.pone.0296336.

[15] T. I. C. Cubitt, K. Wooden, E. Kruger, and M. H. Kennedy, "A predictive model for serious police misconduct by variation of the theory of planned behaviour," J. Forensic Pract., vol. 22, no. 4, pp. 291–302, 2020, doi: 10.1108/JFP-08-2020-0033.

[16] N. Choudhury, R. Mukherjee, R. Yadav, and S. Jain, "Can machine learning approaches predict green purchase intention? a study from Indian consumer perspective," *J. Clean. Prod*., vol. 456, no. Jan., pp. 1–12, 2024, doi: 10.1016/j.jclepro.2024.142218.

[17] T. Aung, K. Thinzar, and S. W. Phyo, "A comparative analysis of machine learning models in predicting blood donation behavior," *Int. J. Res. Sci. Innov*., vol. 12, no. 5, pp. 157–165, 2025, doi: 10.51244/ijrsi.2025.120500157.

[18] M. M. Zaw and T. T. Htar, "Predicting consumer purchasing behavior using SVM and random forest classification methods," *Indian Sci. J. Res. Eng. Manag.*, vol. 9, no. 2, pp. 1–8, 2025, doi: 10.55041/ijsrem51727.

[19] Q. Deng and D. Söffker, "Classifying human behaviors: improving training of conventional algorithms," *IEEE Access*, vol. 2019, no. Oct., pp. 3558–3563, 2019, doi: 10.1109/ITSC.2019.8917489.

[20] E. H. Abderrahim and R. Taoufik, "Prediction of speeding intention based on the theory of planned behavior," *IEEE Access*, vol. 2023, no. Oct., pp. 1–6, 2023, doi: 10.1109/sita60746.2023.10373749.

[21] D. F. Brito, J. L. Cardoso Júnior, J. C. dos Reis, G. Ruppert, and R. Bonacin, "Exploring supervised techniques for automated recognition of intention classes from Portuguese free texts on agriculture," *Rev. Inform. Teor. Apl.*, vol. 29, no. 3, pp. 79–94, 2022, doi: 10.22456/2175-2745.117481.

[22] S. Masrom, N. Z. A. Samad, R. A. Rahman, F. H. M. Fatzel, and S. M. Shamsudin, "Machine learning in predicting whistle-blowing intention of academic dishonesty with theory of planned behaviour," *Indones. J. Electr. Eng. Comput. Sci*., vol. 31, no. 2, pp. 909–916, 2023, doi: 10.11591/ijeecs.v31.i2.pp909-916.

[23] R. A. Rahman, S. Masrom, J. Ahmad, H. Hashim, and E. Mutia, "Tree-based pipeline optimization machine learning in classifying whistleblowing of academic misconduct," *J. Adv. Res. Appl. Sci. Eng. Technol*., vol. 38, no. 2, pp. 165–175, 2024, doi: 10.37934/araset.38.2.165175.

[24] R. A. Rahman, S. Masrom, M. Mohamad, E. N. Sari, and F. Saragih, "Comparisons of automated machine learning (AutoML) in predicting whistleblowing of academic dishonesty with demography and theory of planned behavior," *arXiv*, vol. 2023, no. Jan., pp. 1–12, 2023, doi: 10.2139/ssrn.4359649.

[25] H. Park and J. Blenkinsopp, "Whistleblowing as planned behavior: a survey of South Korean police officers," *J. Bus. Ethics*, vol. 85, no. 4, pp. 545–556, 2009, doi: 10.1007/s10551-008-9788-y.

[26] J. D. Sallaberry, I. Martinez-Conesa, and L. Flach, "Whistleblowing in small and large accounting firms in Brazil," *Small Bus. Int. Rev*., vol. 6, no. 2, pp. 1–12, 2022, doi: 10.26784/sbir.v6i2.502.

[27] I. Ajzen, "Belief, attitude, intention and behavior: an introduction to theory and research," *J. Behav. Sci*., vol. 1975, no. Jan., pp. 1–12, 1975.

[28] S. Horiguchi, R. Haniffa, and M. Hudaib, "Whistleblowing intentions by public accountants in a non-Western society: the case of Japanese accountants," *J. Int. Account. Audit. Tax*., vol. 57, no. Dec., pp. 1–12, 2024, doi: 10.1016/j.intaccaudtax.2024.100656.

[29] T. M. T. Mansor, A. M. Ariff, and H. A. Hashim, "Whistleblowing by auditors: the role of professional commitment and independence commitment," *Manag. Audit. J*., vol. 35, no. 8, pp. 1033–1055, 2020, doi: 10.1108/MAJ-11-2019-2484.

[30] M. Sarikhani and F. Ebrahimi, "Whistleblowing by accountants: an integration of the fraud pentagon and the extended theory of planned behavior," *Meditari Account. Res*., vol. 30, no. 6, pp. 1740–1763, 2022, doi: 10.1108/MEDAR-10-2020-1047.

[31] O. A. S. Otchere, G. M. Y. Owusu, and R. A. Bekoe, "Determinants of whistleblowing intentions of accountants: a middle range theoretical perspective," *J. Financ. Crime*, vol. 30, no. 5, pp. 1162–1181, 2023, doi: 10.1108/JFC-07-2022-0168.

[32] T. M. T. Mansor, A. M. Ariff, H. A. Hashim, and A. H. Ngah, "External whistleblowing intentions of auditors: a perspective based on stimulus–organism–response theory," *Corp. Gov*., vol. 22, no. 4, pp. 871–897, 2022, doi: 10.1108/CG-03-2021-0116.

[33] H. Latan, C. M. Ringle, and C. J. C. Jabbour, "Whistleblowing intentions among public accountants in Indonesia: testing for the moderation effects," *J. Bus. Ethics*, vol. 152, no. 2, pp. 573–588, 2018, doi: 10.1007/s10551-016-3318-0.

[34] T. M. T. Mansor, A. M. Ariff, H. A. Hashim, and A. H. Ngah, "Whistleblowing intentions among external auditors: an application of the moderated multicomponent model of the theory of planned behaviour," *Meditari Account. Res*., vol. 30, no. 5, pp. 1309–1333, 2022, doi: 10.1108/MEDAR-07-2020-0948.

[35] M. Manesh, A. Tavakoli, A. E. Oyewunmi, and S. Pillay, "Ethical decision-making in East Africa: predictors of whistleblowing intention," *Soc. Bus. Rev.*, vol. 19, no. 3, pp. 424–454, 2024, doi: 10.1108/SBR-03-2022-0067.

[36] M. S. Schwartz, "Ethical decision-making theory: an integrated approach," *J. Bus. Ethics*, vol. 139, no. 4, pp. 755–776, 2016, doi: 10.1007/s10551-015-2886-8.

[37] T. M. T. Mansor, A. M. Ariff, A. H. Ngah, and H. A. Hashim, "Religiosity and external whistleblowing intention," *Asian J. Bus. Account*., vol. 15, no. 2, pp. 111–147, 2022, doi: 10.22452/ajba.vol15no2.4.

[38] H. Latan, C. J. C. Jabbour, and A. B. L. de S. Jabbour, "Ethical awareness, ethical judgment and whistleblowing: a moderated mediation analysis," *J. Bus. Ethics*, vol. 155, no. 1, pp. 289–304, 2017, doi: 10.1007/s10551-017-3534-2.