Prediction of Employee Loyalty Classification for Maintaining Company Sustainability using Combination of SEM-PLS and Machine Learning

Sudarmo Sudarmo [1*], Rachmie Sari Baso [2] and Muhammad Adenuddin Alwy [2]

[1]Department of Economics, Sekolah Tinggi Ilmu Ekonomi Balikpapan, Balikpapan, East Kalimantan, Indonesia
[2]Department of English Language Education, Foreign Language Academy of Balikpapan, Balikpapan, East Kalimantan, Indonesia
*Correspondence: sudarmo@stiebalikpapan.ac.id

Abstract: The company's sustainability is influenced by several internal and external factors, one of them is the turnover intention which has an impact on the weak loyalty of human resources. Based on the results of a survey, it shows the level of turnover intention is influenced by many factors. A high level of turnover intention will threaten the sustainability of the company. This should be of particular concern to human resource managers. Therefore, this study aims to analyze the prediction of turnover intention based on the factors affecting employees. The method used is a combination of SEM-PLS and Machine Learning based on the Random Forest algorithm. The empirical model developed in this study was based on four research variables. The variables used are Job Satisfaction (JS), Organizational Commitment (OC), Employee Performance (EP), Turnover Intention (TI). Then, analysis of the variable relationship using SEM-PLS resulted in JS and OC having a significant effect on the level of turnover intention. Thus, each indicator from JS and OC can be used to create a predictive model of turnover intention classification based on the Random Forest algorithm. The results of the Random Forest model obtained have a very good prediction accuracy of 98%. Based on the analysis of the important variables, it was found that OC01 was the most influential variable on the level of turnover intention, with a mean decrease accuracy of 39.51. The predictor variable OC01 means that employees have a good level of "strong belief in a career in the company". Therefore, human resource managers must pay better attention to career projections and the suitability of employee positions. So, the employees will give loyalty to the company, then it can better maintain the company's sustainability.

Keywords: Loyalty; Turnover Intention; Job Satisfaction; Organizational Commitment; SEM-PLS; Machine Learning; Random Forest.

---

1. Introduction

Management of Human Resources (HR) in a company can determine the sustainability of the company. The effectiveness and efficiency of HR management planning can be shown from the results of the employee recruitment process following the company's expectations (Cicek, 2011). Employees who have been recruited should be able to provide good performance in their work. In addition, good human resources are employees who are loyal to the company. If the loyalty of a worker is not good, it will have an impact on changes in the dynamic staffing structure. Too many employee changes shortly can affect the company's performance. Newly joined employees need a process and time to adapt so they can provide maximum performance. The process of changing too many employees needs to be avoided by the company's HR managers, as a step to streamline the time and cost of the recruitment process (Kremic et al., 2006).

The desire to switch jobs is an early indicator of employee turnover in a corporation. The degree of attitude propensity possessed by employees to hunt for new jobs elsewhere or their plans to resign from the company is referred to as Turnover Intention (TI). Another attitude that emerges in individuals when turnover intention arises is a desire to explore for alternative job openings, weighing the likelihood of obtaining a better job elsewhere (Carmeli & Weisberg, 2006). However, if the possibility to change jobs is not available, or if what is offered is not more appealing than what you presently have, employees who wish to quit the company are frequently late, frequently miss work, are less enthusiastic, or lack the will to strive hard. (Ertas, 2015) As a result, it is apparent that turnover intention will have a negative influence on the firm since it produces instability and uncertainty in labor circumstances, lowers employee productivity, and creates an unfavorable work environment. High employee turnover also reduces the effectiveness of the company's organization. Losing experienced employees and having to retrain new employees will add time and money to the process.

High employee turnover grabs the attention of the company because it disrupts operations, creates moral problems for employees who stay, and also inflates costs in recruitment, as well as administrative costs for processing new employees (Cho & Lewis, 2012). The turnover intention has an impact on the weak performance of employees. The way to anticipate is to know the factors that influence turnover intention. Then these factors can be used to reduce the level of turnover intention. Existing studies and literature show that a person's turnover intention is closely related to job satisfaction and organizational commitment where these three things will affect employee performance later (Nazir et al., 2016). Employee performance can reflect the sustainability of the company. In other words, good employee performance can provide good sustainability for the company (Dessler, 2013). Several approaches can be taken to achieve effective and efficient employee performance. One of the approaches is management needs to get commitment from employees to the company. Employee commitment is shown by accepting the company's goals, company interests, and loyalty to the company (Colquitt et al., 2014).

Based on organizational commitment can provide a strong prediction of the desire for voluntary turnover. There is a tendency for commitment before entering the company to be positively related to initial commitment, while commitment after entering the company will be negatively related to voluntary turnover. (Cohen, 2017). Then, commitment in the early stages of joining the company will influence employee job satisfaction. Companies that can make employees believe in the company vision will get employee commitment. So, employees will be loyal to the company and work well for the benefit of the company. This condition is ideal for achieving company goals because the company has complete support from its members, allowing them to focus solely on priority goals. Commitment is a key aspect for company businesses because of its impact on performance, which assumes that loyal employees put in more effort at work.

Strong commitment goes hand in hand with strong productivity. Commitment also gives an indication of lower turnover intention. Organizational commitment is useful for predicting the sustainability of the company related to employee performance (Meyer & Parfyonova, 2010). Analysis of the outcome variable in this study is related to the individual's turnover intention. Organizational commitment is strongly related to an individual's desire to leave the company's operations. Job satisfaction might also impact a person's decision to leave. Individuals who choose to leave the organization company may expect better satisfactory results elsewhere. The turnover decision is a result of their evaluation of numerous alternative positions. One of the reasons for looking for alternative work is satisfaction with the salary received. Employees feel a sense of justice (equity) to the salary received in accordance with the work done. Salary satisfaction can be interpreted that employees get a salary as expected. The level of job satisfaction of these employees is a factor that affects employee turnover. If the employee turnover rate is high, it indicates a fundamental problem in a company. In addition, the employee turnover process also takes a lot of time and costs. Therefore, companies need to reduce employee turnover rates according to company needs. So, it is necessary to know more deeply what affects the Turnover Intention of the employees. If the employee turnover rate can be controlled, then HR management will be better to increase company productivity.

The method approach that can be used to analyze the relationship between variables that affect the turnover intention of employees is SEM-PLS. SEM-PLS tests the validity and reliability of the indicators that make up the construct of the model. The SEM-PLS model can produce an analysis of the relationship between variables that significantly affect other variables. Furthermore, the variable which is a significant factor is used to determine the prediction of the classification of the level of the turnover intention of employees. The development method used is a machine learning based on the Random Forest model. Machine learning will practice understanding the classification of factors from the independent variable data on the dependent variable (Turnover Intention). The machine learning process will produce a Random Forest model which is used to classify the dependent variable based on the independent variable. So that can be predicted the level of turnover intention based on the influence of the independent variable.

Based on this description, the motivation and purpose of this study are to analyze the prediction of turnover intention classification using the SEM-PLS and machine learning methods. The development of this research topic is compared with differences from other studies that have been carried out, namely the combination of SEM-PLS and machine learning based on the Random Forest model. Therefore, the novelty of this study is the development of a new method to obtain a predictive analysis model for turnover intention classification. This is the advantage of this research because the Random Forest model uses independent variables whose significance has been analyzed first based on SEM-PLS. In addition, the results of machine learning with the Random Forest model can provide an analysis of the important variable to determine the best predictor. Thus, these results can be used as a reference for the focus of human resource management, so that companies can control the level of employee turnover intention. Thus, employees will give loyalty to the company. As a result, it can better maintain the company's sustainability.

2. Literature Overview

2.1. Basic Concepts

2.1.1. Turnover Intention

An employee can be said to have a turnover intention, if he wants to move or leave the company but has not yet been able to make it happen (Son, 2012). The turnover intention according to several definitions can be concluded as a person's desire to leave the company where they work to find another job that is better than before (Podsakoff et al., 2007). Desire to leave can be seen from the indicators that are addressed through behaviors such as thinking out of work, actively looking for another job, leaving work in the near future, always looking for opportunities to leave the company, lack of enthusiasm for working in the current job and leaving the company when getting a salary larger ones. There are several factors that cause employees to have a desire to leave, namely:

1. Structural Factors, namely factors related to work and organization, such as support from colleagues, support from superiors, work routines, equitable distribution of justice, role ambiguity, workload, employee skills, rewards, job security, and career development.
2. Pre-entry factors, namely factors that include positive personalities such as a tendency to be happy, and negative personalities such as a tendency to experience discomfort, and so on.
3. Environmental factors, namely factors related to things outside the work and organization. Environmental factors include job opportunities available outside the company, habits of moving from people around employees, and the number of family members covered.
4. Trade Union Factors, namely factors related to an employee's membership of a trade union that can influence an employee's decision to keep a job or decide to move.
5. Job Orientation, namely factors that include job satisfaction, organizational commitment, and employee activities or efforts to find alternative jobs outside the organization where they work today.

In addition to the things above, there are many causes of turnover intention, including work stress, work environment, job satisfaction, organizational commitment, and others experienced by employees (Qureshi et al., 2013).

2.1.2. Organizational commitment

Organizational or company management is largely determined by HR management. Based on studies related to HR management, organizational commitment is an important issue because it is an aspect that affects human behavior in organizations. The main reason is quite simple, because the achievement of the organization's vision, mission, and goals is carried out by members of the organization. So that if there is no commitment from members of the organization, it will not achieve organizational goals (Richard & Johnson, 2001). Employee commitment to the organization or company will affect the organization's capacity to achieve its goals and vision. At this time several companies dare to include an element of commitment as one of the requirements for job vacancies being offered. But not infrequently, there are still business actors who still do not understand the importance of this commitment properly. Several studies have provided views to understand employee commitment to the organization. Organizational commitment is identified by the match between employee goals and organizational goals. In addition, employees need to be willing to give all efforts for the sustainability of the organization. The level of employee loyalty can be seen from the interest in being able to continue to be part of the organization or company (Cohen, 2007). Another definition of organizational commitment is employee loyalty which is shown by their concern for the success of the organization on sustainable progress. (Yahaya & Ebrahim, 2016).

2.1.3. Job satisfaction

Job satisfaction has a negative impact on the level of turnover intention. In other words, if employees have a good level of job satisfaction, they will not want to change jobs. However, on the other hand, there are external factors that can influence such as labor market conditions and alternative job opportunities. This can make employees want to move because they are tempted by other offers even though their current job is fine (Aziri, 2011). The factors that encourage job satisfaction are:

1. Mentally challenging work.
2. Supportive working conditions.
3. Supportive co-workers.
4. Personality compatibility with work.

Employees who are satisfied with their work can be predicted to last long in the organization or company. Meanwhile, employees who are dissatisfied with their work will choose to move or leave the company. Perceived job satisfaction can be a factor for an employee to leave. In addition, some consideration of alternative jobs can also result in turnover intention. This is because employees expect more satisfaction jobs elsewhere. The researcher determines the indicators of each job satisfaction variable based on the opinions of experts regarding each of these variables which are then translated into questions in a questionnaire to the research sample.

1. Work Morale

Work morale includes the following three elements (Islam et al., 2011): (a) Loyalty and pride in the company. (b) Attitude towards superiors (c) Teamwork

2. Discipline

The factors that affect work discipline (Mangkunegara & Octorend, 2015) are as follows: (a) Punctuality at work. (b) Use of office facilities. (c) Responsibility for the completion of tasks. (d) Compliance with company regulations.

3. Work Performance

The indicators in work performance (Pratama, 2019) are (a) Work accuracy. (b) Work Tidiness. (c) Speed of completing tasks. (d) Initiative at work.

2.1.4. Employee Performance

Performance can be said to be a result of motivation, ability, and action. An employee must have a certain degree of motivation and level of ability to complete a task or job. In addition, when the level of motivation and ability is not balanced, performance will also be disrupted. For example, when an employee has good abilities but is not accompanied by goodwill or motivation, then his performance will be less than it should be. Some indicators of employee performance can involve factors of quality and quantity of output. Examples of employee performance indicators are attendance at work, helpfulness, cooperation, punctuality, and others (Shahzadi et al., 2014).

The definition of performance indicators can be said as aspects used to measure the performance of an employee. The form of indicators that measure employee performance (Mathis et al., 2015) are as follows:

a. Quantity is a measure of the number of products produced. This quantity can be expressed in the number of production units, the number of activity cycles completed by employees, and the number of activities produced to support the company.
b. The quality of work is expressed as the employee's perception of the quality of the work produced. This measure of work quality is related to the perfection of the work done. The factors that determine the quality of work are the skills and abilities of employees.
c. Timeliness is the employee's perception of an activity carried out within the specified period. For example, a job can be completed at the right time until it becomes the expected work result.
d. Attendance is also a performance indicator. This attendance consists of coming to work, leaving work, permission, and other information related to the employee's presence.
e. The ability to cooperate is the ability of employees to work together with a team or other people in their work. Cooperation skills will support the completion of tasks and jobs to achieve maximum results.

2.2. Relationship Between Variables

2.2.1. Effect of Job Satisfaction and Turnover Intention

Employee turnover can arise because of alternative jobs and or lack of job satisfaction. Based on the size of job dissatisfaction can be identified, how a worker has a desire to leave his job. In addition, job dissatisfaction indicators have a direct influence on the formation of a desire to move to other alternative jobs (Alsaraireh et al., 2014). In other words, employee turnover is also negatively related to job satisfaction. In addition, some of the important barriers in the decision to change a job are labor market conditions, spending on alternative employment opportunities, and the length of the work contract offered.

Identification of employee attitudes related to job satisfaction is often studied by identifying the relationship between satisfaction and turnover based on psychological variables (Liu et al., 2010). Employees who are dissatisfied with their jobs tend to do things that can interfere with the company's performance. Some of them are high turnover, slackness at work, high absenteeism, job complaints, and not working properly. It can be said, the higher the level of job satisfaction, the lower the desire to change jobs (Lai & Chen, 2012). Those who are dissatisfied with their jobs are more likely to depart and seek employment elsewhere. Several studies have indicated that work satisfaction is a potential predictor of intention to leave (Bonenberger et al., 2014; Kuo et al., 2014; Valentine et al., 2011).

2.2.2. Effect of Organizational Commitment and Turnover Intention

Organizational commitment is an attitude that demonstrates employee loyalty and is a continuing process by which an organization member communicates their interest in the organization's success and goodness. (Yahaya & Ebrahim, 2016). Individuals who have low organizational commitment tend to look for better work opportunities and leave their jobs because they have an embedded desire to leave the organization. Furthermore, increased job satisfaction and organizational commitment are likely to diminish employees' intent or purpose to leave the organization or company. Employees who are unsatisfied with aspects of their work and are uncommitted to their organization are more likely to seek employment elsewhere. Several empirical research has demonstrated that individuals who satisfy organizational commitment have a high level of job satisfaction and a lower propensity to leave (Aydogdu & Asikgil, 2011; Tarigan & Ariani, 2015; Yücel, 2012).

2.2.3. Effect of Turnover Intention and Employee Performance

Turnover intention is a precursor to staff turnover in a company. The attitude of propensity possessed by employees to hunt for new jobs elsewhere is referred to as turnover intention (Bothma & Roodt, 2013). Another attitude that emerges in employees when turnover intention arises is the desire to find alternative work vacancies, as well as evaluating the probability of getting a better job elsewhere. However, if the opportunity to change jobs is not available, or if what is available is not more appealing than what they currently have. Employees will leave the company emotionally and mentally, namely by frequently arriving late, skipping work, being less enthusiastic, or lacking the desire to try hard (Yi et al., 2011). As a result, turnover will have a negative influence on the organization since it generates insecurity in labor conditions. If the employees feel an unsuitable working environment, then has an impact on decrease human resource productivity (Balouch & Hassan, 2014).

2.2.4. Effect of Job Satisfaction and Employee Performance

Siengthai & Pila-Ngarm (2016) stated about the relationship between job satisfaction and employee performance by the state of the company. The employees who are more satisfied tend to be more effective than companies with less satisfied employees. Job satisfaction has a role in achieving better productivity and quality standards, avoiding the possibility of building a more stable workforce, and using human resources more efficiently (Mangkunegara & Octorend, 2015). Other research shows that job satisfaction is an independent variable that has a positive effect on management attitudes towards company strategy which is reflected through employee performance (Badrianto & Ekhsan, 2020; Inuwa, 2016).

2.2.5. Effect of Organizational Commitment and Employee Performance

Commitment is an employee's attitude with the desire to stay a member to attain the goal's company. Employees will be more motivated to attain organizational goals if they believe their attitudes and values are consistent with those of the organization. Several studies have found that corporate commitment has a major impact on employee performance (Ahmad et al., 2014; Fu & Deshpande, 2014).

3. Materials and Methods

This study was carried out to test the intended hypothesis by using research methods that have been designed according to the variables to be studied to obtain accurate research results.

3.1. Population, Sample, and Ethics Committee of Respondents

The scope of this research is employees of Tani Sejahtera Mandiri (TSM) Kalimantan Ltd., with a total of respondents is 500. Organizational commitment, job satisfaction, turnover intention, and employee performance are among the characteristics investigated. The sampling technique employed is stratified random sampling using a proportionate sampling approach, with the goal of representing employees in each portfolio. All subjects gave their informed consent for inclusion before they participated in the study.

The participant consent document was written in the cover letter in the survey. The form includes all elements of a regularly signed consent, including the confidentiality statement given below. The consent line says, "By completing the survey you are agreeing to participate in the research". The survey includes an "I agree" or "I disagree" option for participants to make their choice whether they agree to participate or not. Then, this study approved by the Ethics Committee of Tani Sejahtera Mandiri Ltd. (Date of approval: January 10th, 2021).

3.2. Research Variables and Indicators

In this study, an analysis of the relationship between the influence of the independent variable and the intervening variable was carried out on the dependent variable. In addition, this study also examines the indicators that influence these variables. The variables studied include organizational commitment, work satisfaction, turnover intention, and employee performance. Furthermore, the characteristics are simplified into research indicators, as shown in Table 1.

Table 1. Research variables and indicators

Latent Variables	Indicator
Organizational Commitment (OC)	OC1 Strong belief in a career in the company
	OC2 Level of involvement on company issues
	OC3 Level of interest in the company
	OC4 The feeling of being part of the company
	loyalty to the organization
Job Satisfaction (JS)	JS1 Satisfaction with salary
	Satisfaction with promotions and opportunities for advancement
	Satisfaction with colleagues and superiors in the company
	JS4 Satisfaction with working conditions/environment
Turnover Intention (TI)	The tendency of individuals to think about leaving the organization where they work now
	The probability that the individual will look for work in another organization
Employee Performance (EP)	EP1 Quality and quantity of employee work
	EP2 Attitude
	EP3 Cooperation and Communication between employees
	EP4 Creativity

3.3. Research Models and Hypotheses

In this section, the researcher proposes a model that is taken based on the results of a literature review and previous research. The proposed model uses organizational commitment, work satisfaction, turnover intention, and employee performance variables. The model studied is a model that was compiled based on a review of the literature and previous research, Figure 1 is a path diagram of the model studied in this study.

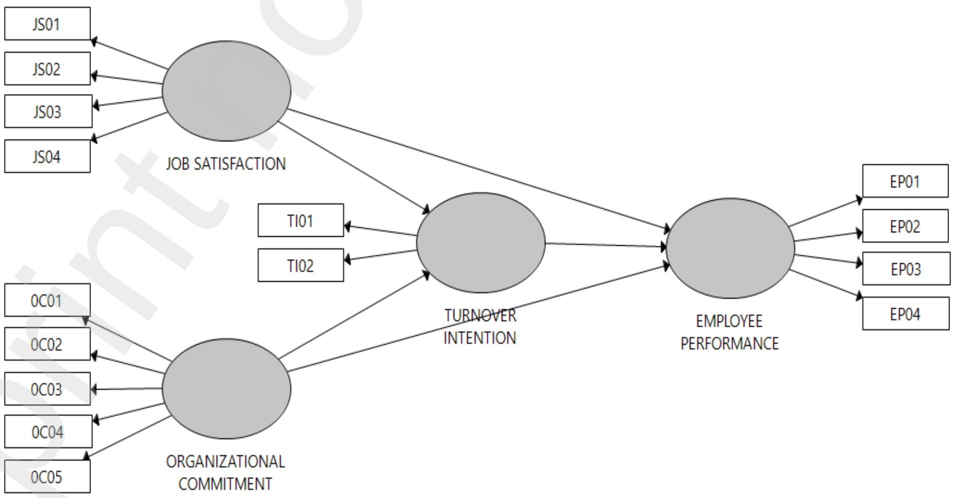

Figure 1. Path Diagram of the Research Model.

Based on Figure 1, along with the description in the literature review section, it can be seen that the hypotheses proposed in this study are:

H_1: JS has a positive effect on EP
H_2: JS has a negative effect on TI
H_3: OC has a positive effect on EP
H_4: OC has a negative effect on TI
H_5: TI has a negative effect on EP
H_6: OC affects EP through TI
H_7: JS affects EP through TI.

3.3. Structural Equation Modelling Partial Least Squares (SEM-PLS)

The analysis of this research uses the method of Structural Equation Modeling (SEM). SEM model processing is done using Smart PLS 3.0 software. The classification of SEM models is divided into two types; (1) Covariance-Based SEM (CB-SEM) and (2) Variance-Based SEM, often known as Partial Least Squares (SEM-PLS). There are various factors to consider while deciding which SEM approach to use. In terms of objectives, CB-SEM is employed if the researcher wants to test the theory, confirm the theory, or compare numerous alternative hypotheses. In contrast to CB-SEM, SEM-PLS is exploratory or an expansion of an existing theory. The SEM-PLS method was employed in this study (Hair Jr et al., 2017).

PLS is a powerful analytical method because it is not predicated on numerous assumptions. PLS can be used to confirm theory as well as to explain the presence or absence of a link between latent variables (Hair et al., 2019). Because it focuses more on data and with limited estimation procedures, the model specifications have little effect on parameter estimation.

In general, there are two main processes that must be completed when assessing the SEM method: (1) the evaluation process for the measurement model and (2) the evaluation process for the structural model (Ahmad et al., 2016). The two processes are carried out sequentially. Before conducting a structural model analysis, it is necessary to create a measurement model first. The measurement model was tested for the validity and reliability of latent variable indicators (Khatimah et al., 2019). The summary of the evaluation criteria for the measurement model is shown in Table 2.

Table 2. Summary of Measurement Model Evaluation Criteria

Validity and Reliability	Parameter	Criteria
Convergent Validity	Loading Factor	a. Greater than 0.70 for Confirmatory Research b. Greater than 0.60 for Exploratory Research
	Average Variance Extracted (AVE)	Greater than 0.50 for confirmatory and exploratory research
Discriminant Validity	Cross Loading	Loading to another construct is lower than its loading value on the construct
Reliability	Cronbach's Alpha	a. Greater than 0.70 for Confirmatory Research b. Greater than 0.60 still acceptable for exploratory research
	Composite Reliability	a. Greater than 0.70 for Confirmatory Research b. Greater than 0.60 still acceptable for Exploratory Research

In addition, the development of a structural model was used to investigate the relationship between latent variables in the measurement model. The preparation of these latent variables can be influenced by other construction factors. The form of the parameter indicating the regression of the exogenous latent variable is labeled γ ("gamma"), while those indicating the regression of endogenous latent variables are labeled β ("beta"). Assume that the random vectors $\boldsymbol{\eta}^T = (\eta_1, \eta_2, \dots, \eta_m)$ and $\boldsymbol{\xi}^{\mathrm{T}} = (\xi_1, \xi_2, \dots, \xi_n)$ are endogenous and exogenous variables, respectively, and that they form a simultaneous equation with a system of linear equations. As a result, the generic version of the structural equation model is defined as follows:

$$\eta = \mathrm{B}\eta + \Gamma\xi + \zeta \tag{1}$$

In structural equations, **B** and **Γ** are coefficient matrices, and $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \dots, \zeta_m)$ are error vectors. Element **B** represents the variable's effect on the other variables. Furthermore, the element has a direct influence on the variable **Γ** in the variable $\boldsymbol{\eta}$. It is believed that is Γ unrelated to $\boldsymbol{\zeta}$ and that **1-B** is nonsingular (Mondego & Gide, 2020).

The structural model's form is obtained in the following order:

$$\begin{aligned} \eta &= B\eta + \Gamma\xi + \zeta, \\ \eta - B\eta &= \Gamma\xi + \zeta, \\ (1 - B)\eta &= \Gamma\xi + \zeta, \\ \eta &= ((1 - B)^{-1}(\Gamma\xi + \zeta), \end{aligned} \tag{2}$$

with:

$\boldsymbol{B}$: endogenous latent variable coefficient matrix of size $m \times n$

Γ : coefficient matrix of exogenous latent variable with size $m \times n$

$\boldsymbol{\eta}$: endogenous latent variable vector of size $m \times 1$

$\boldsymbol{\xi}$: exogenous latent variable vector of size $n \times 1$

$\boldsymbol{\zeta}$: vector random of residual relationship between $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ of size $m \times 1$

PLS was used to evaluate structural models, and the R-Square value for each dependent variable was used to determine the structural model's predictive potential. Determination of latent variables that have a significant effect on the dependent variable is known based on the R-Square value. The criteria for the R-Square value used for the strong, medium, and weak models are 0.67, 0.33, and 0.19, respectively (Yang et al., 2021). An overview of the evaluation criteria for the structural model is shown in Table 3.

Table 3. Summary of the Structural Model Evaluation Criteria

Parameter	Criteria
R-square	0.67 (Strong); 0.33 (Moderate); and 0.19 (Weak)
Significance Level	5% (0.05)

After evaluating the model, it is necessary to analyze the connecting variables to determine the effect between the independent variable and the dependent variable. Analysis of the relationship was carried out by examining the effects of mediation. The influential relationship between the independent variables on the dependent variable can be direct or indirect. The nature of the relationship, directly or indirectly, is determined by the mediating effect of the connecting variable (Preacher et al., 2011). The SmartPLS 3.0 software uses a procedure devised by Baron and Kenny in 1986 to test the mediating effect. According to Baron and Kenny, testing the mediating effect is carried out based on the following three stages:

a. The first model is used to investigate the effect of the independent variable (X) on the dependent variable (Y). The relationship criteria must be significant 5% (0.05)

b. The second model investigates the effect of the independent variable (X) on mediation (M) and must be significant at 5% (0.05).

c. The third model investigates the effect of the independent variable (X) on mediation (M) on the dependent variable at the same time (Y).

The relationship between the independent variable (X) and the dependent variable (Y) is estimated to be insignificant at the final stage of testing. However, the effect of the mediating variable (M) on the dependent variable (Y) is expected to be significant at 5%. Significance analysis of the influence between variables using the mediating variable (M) was carried out in the final test. The significance test was carried out with the help of the SmartPLS 3.0 program. Based on the analysis of the SmartPLS 3.0 program, the indirect effect coefficient and the significance value of the indirect effect will be obtained (with a significant level of 5%).

3.3. Machine Learning

Machine learning is a method to make machines (computers) have the ability to learn and do a job automatically. The machine learning process is carried out through certain algorithms so the work order to the computer can be done automatically.

Machine learning is carried out through 2 phases, namely the training phase and the application phase. The training phase is the modeling process of the algorithm that will be learned by the system through training data. Meanwhile, the application phase is the process of using the output model of the training phase, to produce a certain decision using testing data.

Machine learning can be done in two ways, namely supervised learning and unsupervised learning. Unsupervised learning is the processing of sample data without requiring the final result to conform with a certain form, by using several data samples at once. The application of unsupervised learning can be found in the visualization process, or data exploration. Supervised learning is processing sample data x will be processed in such a way, to produce output that matches the final result y. Supervised learning can be applied to the classification process. One of the supervised processes is Random Forest.

3.3.1. Random Forest

Random Forest is a bagging method, which is a method that generates a number of trees from sample data where the creation of one tree during training does not depend on the previous tree, then decisions are taken based on the most votes. The Random Forest algorithm was first introduced by Breiman (2001). Random Forest has two problem-solving functions, namely classification, and regression (Svetnik et al., 2003). Random Forest can be used on several types of data such as discrete, continuous, multivariate combinations, and survival data (Wongvibulsin et al., 2020). Random Forest can detect interactions between dependent and independent variables, and is able to explore data with good flexibility (Auret and Aldrich, 2012). Analysis using Random Forest does not use certain assumptions that must be met.

The Random Forest method is used to build a decision tree consisting of root nodes, internal nodes, and leaf nodes by taking attributes and data randomly according to the applicable provisions (Prajwala, 2015). The root node is used to collect data, an inner node located at the root node contains questions about the data, and a leaf node is used to solve problems and make decisions. The decision tree begins by calculating the entropy value as a determinant of the level of attribute impurity and the value of information gain as in equations (3) and (4) (Reinders et al., 2019).

$$Entropy\ (Y) = -\sum_i p(c \mid Y) log_2\ p(c \mid Y) \tag{3}$$

where Y is set of cases and $p(c \mid Y)$ is value proportion of Y to class of c.

$$Information\ Gain\ (Y, a) = Entropy\ (Y) - \sum_{v \in\ Values\ (a)} \frac{|Y_v|}{|Y_a|}\ Entropy\ (Y_v) \tag{4}$$

with $Values(a)$ is all possible values in the case set a, Y_v is sub-class of Y with class of v have relationship with a, and Y_a is all values corresponding to a.

The random forest has two main parameters, namely: m is the number of trees to be used and k is the maximum number of features that are considered when branching (Kulkarni and Sinha). The more m values, then the better of classification results, while the recommended k value is the square root or logarithm of the total number of features. The following is an example of a plot tree from Random Forest in Figure 2.

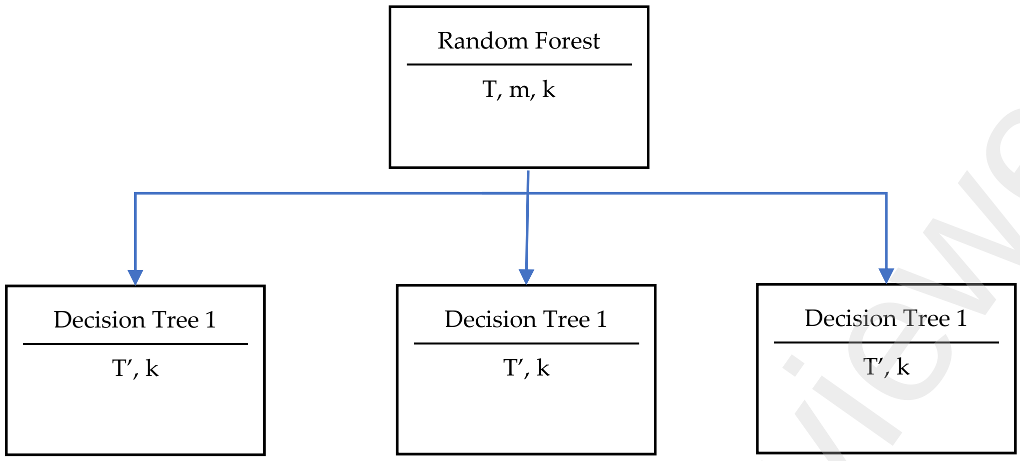

Figure 2. Illustration of Random Forest Plot Tree

Figure 2 shows the training process for a random forest using a dataset T with a number of m trees as the basic learner and k features chosen at random from the total features for branching in each tree. The training process on each tree uses the T' dataset which is the result of bootstrapping the dataset which is used as a parameter for a random forest.

3.3.2. Performance Evaluation of Random Forest Model

The evaluation was carried out for the selection of the best dataset distribution method and classification method seen through classification performance measures. The classification performance measure used in this study takes into account the confusion matrix. The confusion matrix is a useful tool to analyze how well or how accurately the classification method can recognize objects of observation from different classes. The following is an example of a confusion matrix for binary decision variables, which can be seen in Table 4 (Singh et al., 2021).

Table 4. Confusion Matrix

Confusion Matrix		Actual Class: True	Actual Class: False
Prediction Class	True	True Positive (TP) Correct Result	False Positive (FP) Unexpected Result
	False	False Negative (FN) Missing Result	True Negative (TN) Correct Result of "False"

Based on the prediction results from Table 4, then an evaluation of the performance of the classification model is carried out using the levels of accuracy, sensitivity, and specificity. The formula for measuring the performance evaluation is written as follows (Singh et al., 2021).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

Accuracy is the most commonly used method to assess classification performance. However, for the case of imbalanced data, accuracy places more weight on the majority class, so it is necessary to also pay attention to sensitivity and specificity. The analysis of the accuracy results can also be formed in the performance classification table as follows:

Table 5. Classification of Accuracy Level

Accuracy (%)	Classification
90 – 100	Very Good
80 – 90	Good
70 – 80	Enough
60 – 70	Bad
0 – 50	Very Bad

Based on Table 5, it is found that to achieve the model that is needed, the accuracy rate is more than 90%. The calculation of the accuracy, as well as the Random Forest algorithm process in this study, uses R software. The R packages used are *caret*, *e1071*, *randomForest*, *readxl*, *reprtree*, dan *ROSE*.

4. Results

This section describes the results of data processing and the interpretation of the results of data processing.

4.1. SEM-PLS Model Measurement

The model is measured by examining the validity and reliability of the indicators that compose the model's construct. Convergent validity with loading factor and AVE parameters is followed by discriminant validity with cross-loading parameters. In addition to Cronbach alpha and composite reliability metrics, reliability testing is performed too.

4.1.1. Convergent Validity

The process of evaluating the relationship between indicators used in a construct is called a convergent validity test. The requirements for convergent validity are met when the indicators used have a correlation in the same construct. The data used in this study came from a questionnaire filled out by 500 people. The data is used as input for the indicator value in testing the convergent validity. The convergent validity test was conducted based on the SEM model using SmartPLS 3.0.

The outer model is evaluated using three measurement criteria: convergent validity, discriminant validity, and composite validity. Convergent validity can be assessed using reflective indicators, as evidenced by the correlation between the indicators and their concept values. A loading factor value is stated to be valid/reliable if it has a correlation value more than 0.7; however, for research in the early phases of building a measurement scale, a loading value of 0.5 to 0.6 is regarded enough. However, if the final value is less than 0.5, the indicator is ruled incorrect and must be deleted from the model, requiring that data processing (running data) be repeated. Table 4 displays the results of assessing the convergent validity of this study's SEM model.

Table 4. Parameter Value of the First Iteration PLS SEM Loading Factor

	Organizational Commitment (OC)	Employee Performance (EP)	Job Satisfaction (JS)	Turnover Intention (TI)
OC1	0.800	-	-	-
OC2	0.784	-	-	-
OC3	0.880	-	-	-
OC4	0.268	-	-	-
OC5	0.757	-	-	-
EP1	-	0.868	-	-
EP2	-	0.788	-	-
EP3	-	0.868	-	-
EP4	-	0.937	-	-
JS1	-	-	0.946	-
JS2	-	-	0.890	-
JS3	-	-	0.920	-
JS4	-	-	0.767	-
TI1	-	-	-	0.948
TI2	-	-	-	0.932

Based on the results of the SEM-PLS Phase 1 study in Table 4, it is clear that several indications, such as OC04 with a loading factor value of 0.268, are not yet valid. An indication with a low loading factor value, such as OC04, indicates that the indicator's contribution to the construct is minor, and hence the indicator should be eliminated and re-analyzed. After removing OC04 from the model, the SEM-PLS analysis stage 2 was performed, based on the results of the stage 2 analysis, the loading factor value was generated as shown in Table 5.

According to the results of stage 2 SEM-PLS data processing in Table 5, all indicators were valid/had a loading factor value greater than 0.6, implying that all indicators met one of the criteria of convergent validity. Table 5 lists the prerequisites for loading factors greater than 0.6 that match the criteria for exploratory research.

Table 5. Parameter Value of Second Iteration PLS SEM Loading Factor

	Organizational Commitment (OC)	Employee Performance (EP)	Job Satisfaction (JS)	Turnover Intention (TI)
OC1	0.814	-	-	-
OC2	0.791	-	-	-
OC3	0.883	-	-	-
OC5	0.750	-	-	-
EP1	-	0.863	-	-
EP2	-	0.794	-	-
EP3	-	0.870	-	-

EP4	-	0.935	-	-
JS1	-	-	0.946	-
JS2	-	-	0.890	-
JS3	-	-	0.920	-
JS4	-	-	0.766	-
TI1	-	-	-	0.948
TI2	-	-	-	0.932

In addition to analysis the loading factor value, construct validity can be evaluated by testing the Average Variance Extracted (AVE) value. The loading factor value represents the original data score and shows the capacity of the latent variable value. The bigger the AVE value, the greater the ability to explain the value of the hidden variable's indicators. The AVE cut-off value used is 0.50, where an AVE value of at least 0.50 indicates a good measure of convergent validity, which means that the probability of an indicator in a construct entering another variable is lower (<0.50). So that the probability of the indicator converges and enters in constructs whose value in the block is greater than 50% of the value of convergent validity is greater than 50%.

Table 6. Parameter Values of AVE

	Average Variance Extracted (AVE)
Employee Performance (EP)	0.752
Job Satisfaction (JS)	0.780
Organizational Commitment (OC)	0.658
Turnover Intention (TI)	0.884

According to table 6, SEM-PLS data processing in the second stage of the test resulted in the AVE value of each variable being declared good because it exceeded the standards with a value greater than 0.5. This demonstrates that the latent variable can explain more than half of the variance in the indicators. As a result of Tables 5 and 6, it is possible to conclude that all indicators and constructs in the model have passed the convergent validity test. Following that, a discriminant validity test is performed to see whether the indicators of one concept are not highly associated with indications of other constructs.

4.1.2. Validity of Discriminant

The validity of discriminant test examines the relationship between indicators in one construct and indicators in other constructs. If the indications from distinct constructs are not associated, the discriminant validity conditions are met. This is tested using the criterion of the measured construct indicator loading being larger than or equal to the load-ing to other constructs or having a low cross-loading value. Table 7 displays the cross-loading parameter values of the indicators utilized in this investigation using the SmartPLS 3.0 software.

Table 7. Values of Cross Loading Parameter

	Organizational Commitment (OC)	Employee Performance (EP)	Job Satisfaction (JS)	Turnover Intention (TI)
OC1	0.814	0.142	0.336	-0.635
OC2	0.791	0.141	0.250	-0.521
OC3	0.883	0.360	0.522	-0.701
OC5	0.750	0.483	0.278	-0.546
EP1	0.345	0.863	0.321	-0.328
EP2	0.290	0.794	0.010	-0.333
EP3	0.268	0.870	0.140	-0.287
EP4	0.352	0.935	0.212	-0.354
JS1	0.450	0.157	0.946	-0.466
JS2	0.429	0.190	0.890	-0.509
JS3	0.382	0.181	0.920	-0.424
JS4	0.262	0.201	0.766	-0.271
TI1	-0.727	-0.419	-0.466	0.948
TI2	-0.676	-0.282	-0.448	0.932

An indication is also considered valid if its loading factor is greater than its cross-loading value. Table 7 displays the construct correlation of all loading values greater than cross-loading; in this example, the loading value is labeled yellow, indicating test passed with discriminant validity standards.

Another way to assess discriminant validity is to compare the square root value of each construct's AVE with the correlation value between the constructs and other constructs (latent variable correlation). If the AVE root for each construct is bigger than the correlation between the constructs and other constructs, as shown in Table 8, the model has acceptable for discriminant validity.

Table 8. Square Root of AVE

	Organizational Commitment (OC)	Employee Performance (EP)	Job Satisfaction (JS)	Turnover Intention (TI)
Employee Performance (EP)	0.867	-	-	-
Job Satisfaction (JS)	0.203	0.883	-	-
Organizational Commitment (OC)	0.365	0.441	0.811	-
Turnover Intention (TI)	-0.378	-0.487	-0.748	0.940

Table 8 demonstrates that the AVE root values of each construct are all bigger than the correlation between constructs. As a result of Tables 7 and 8, it is possible to conclude that all constructs in the estimated model and met the criteria for the discriminant validity test.

4.1.3. Reliability

A reliability test is performed to determine the consistency and reliability of product indicators when measuring a construct. Cronbach's alpha and composite reliability techniques were used in this work to examine the reliability, with Cronbach's alpha and composite reliability criteria > 0.6 for exploratory research. The SmartPLS 3.0 software is used to test the dependability of the SEM model under consideration; Table 9 displays the model's composite reliability parameter value.

Table 9. Composite Reliability Parameter Values

	Composite Reliability
Employee Performance (EP)	0.923
Job Satisfaction (JS)	0.934
Organizational Commitment (OC)	0.885
Turnover Intention (TI)	0.938

According to Table 9, all construct variables in this investigation showed a composite reliability value greater than the needed value minimum (0.6). Then, this demonstrates the Cronbach Alpha parameter value's reliability requirements in Table 10 were met by the variables in this study.

Table 10. Parameter Value of Cronbach's Alpha

	Cronbach's Alpha
Employee Performance (EP)	0.889
Job Satisfaction (JS)	0.906
Organizational Commitment (OC)	0.826
Turnover Intention (TI)	0.869

Based on the data in table 10, it is found that all construct variables have a Cronbach alpha value greater than 0.7. The minimum requirement of the Cronbach alpha parameter is 0.6 so that the results in Table 10 already meet the criteria of Cronbach alpha. In addition, this shows that the variables in this study meet both reliability standards. Then the measurement results of the three test parameters have also met the criteria. The fulfillment of the model's measurement criteria is based on the characteristics of validity and reliability. The validity test of this study resulted in a good level of accuracy and consistency for each indicator variable. In addition, the indicator variables also have good accuracy in assessing each construct.

4.2. Structural Model Evaluation

The structural model is evaluated by looking at the R-Square value for each dependent variable as the structural model's predictive power. Changes in the value of R-Square, in this case, can suggest some latent variables that have a significant effect on the dependent variable. Then to determine the level of significance use the bootstrapping process.

4.2.1. R-Square

R-square of the latent variable is one of the measurements used to assess the structural model's predictive potential. The R-square value states the percentage of variance of a variable that can be explained by other variables in the same model. R-square values of 0.67, 0.33, and 0.19 can be used to assess the model is strong, moderate, or weak. The R-Square value of the latent variable in the SEM model is shown in Table 11.

Table 11. R-Square Value

	R Square	Details
Employee Performance (EP)	0.158	Weak
Turnover Intention (TI)	0.590	Strong

According to the results of Table 11, there are two R-Square values in this study. Each indicator variable can be represented by an R-Square value. In addition, each indicator variable has unique characteristics. The characteristics of the variables in this study can be described as follows:

1) Employee Performance (EP) has an R-square value of 0.158, indicating the variable is weak. This demonstrates the factors in the model explained 15.8% of the variance in the EP variable, while the remainder is explained by variables outside the model.

2) The Turnover Intentions (TI) variable has an R-square value of 0.59, which is strong. This demonstrates that 59% of the variance in the TI variable can be explained by variables in the model, with the remainder explained by variables outside the model.

4.2.2. Hypothesis Testing

The final test of the structural model evaluation is to look at the level of significance using the bootstrapping process. The significance value of 5% was utilized as the requirement. Before understanding the significance test findings, the output of the PLS method and bootstrapping should be shown. The magnitude of the route coefficient values between latent variables is determined by the PLS algorithm output, but the size of the t-statistical significance value is determined by the bootstrapping output. The outcome of the PLS method and SEM bootstrapping using SmartPLS 3.0 is shown in Figure 2.

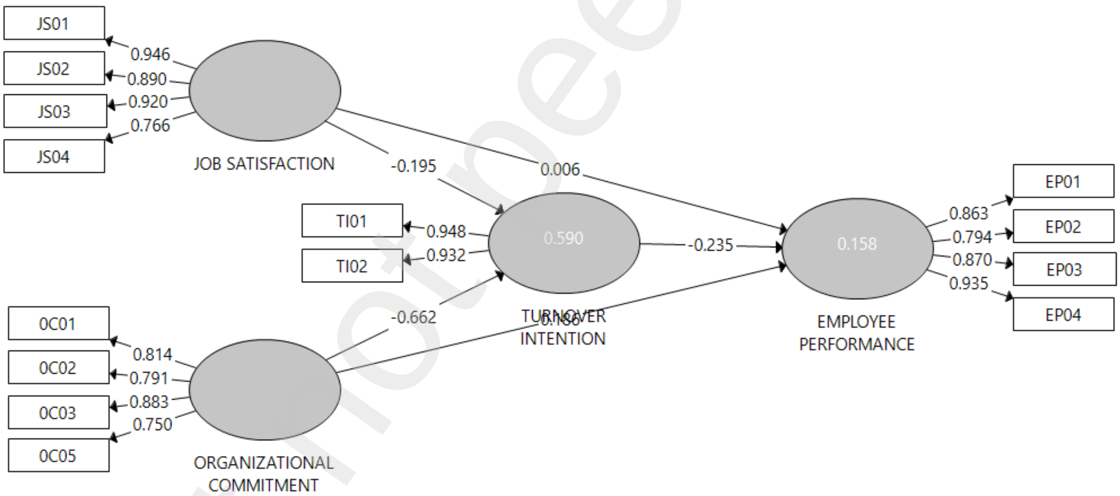

Figure 2. PLS Output

The magnitude of the route coefficient for each latent variable is observed in the results of Figure 2. Furthermore, the bootstrap results are shown in Figure 3 based on the significance value.

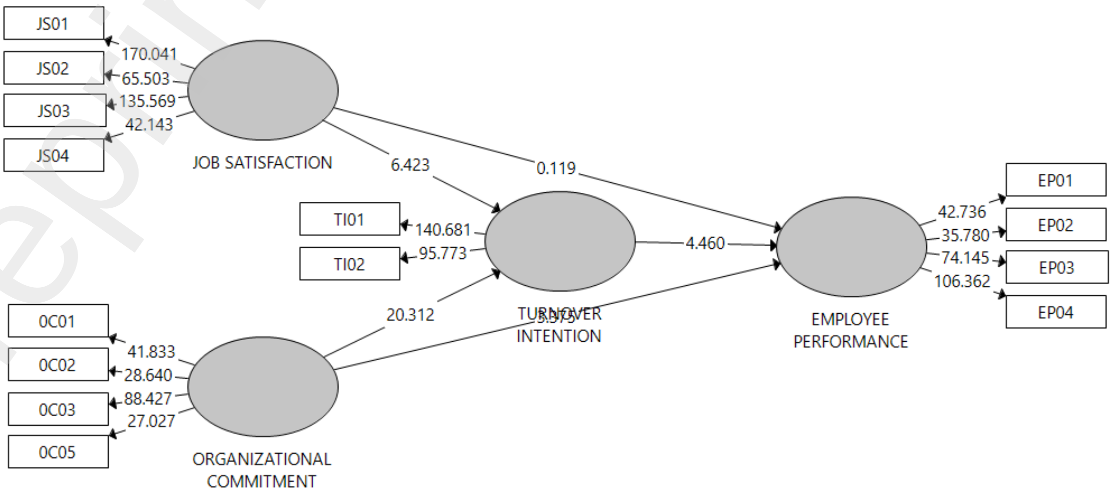

Figure 3. Bootstrap Output

The magnitude of the t-statistic value of the relationship of each latent variable is observed in Figure 3. The magnitudes of the path coefficient values and the t-statistical values can be seen in Figures 2 and 3. Then for a clearer analysis, it is shown in Table 12.

Table 12. Path Coefficient and P-Value

	Original Sample (O)	T Statistics (|O/STDEV|)	P Values
Job Satisfaction -> Employee Performance	0.006	0.117	0.907
Job Satisfaction -> Turnover Intention	-0.195	6.934	0.000
Organizational Commitment -> Employee Performance	0.186	3.375	0.001
Organizational Commitment -> Turnover Intention	-0.662	22.090	0.000
Turnover Intention ->Employee Performance	-0.235	4.544	0.000

Based on Table 12, it can be explained:

H_1: Job Satisfaction (JS) has a positive effect on Employee Performance (EP).

The results of the analysis in Table 12 shows the parameter coefficient of the influence of job satisfaction on employee performance is 0.006. The coefficient indicates a positive influence between the two variables. While the t-statistic value generated is 0.117. The t-statistic value is smaller than the t-table (0.117 < 1.96). Based on the t-statistic which is smaller than the t-table, it can be concluded that the hypothesis is rejected. In other words, job satisfaction has no significant effect on employee performance.

This can be interpreted as the greater level of job satisfaction can trigger an increase in employee performance. So, if an employee is satisfied with the job, then the results of the employee's performance will be good. But on the other hand, when the employee has a good level of job satisfaction, it can make employees feel that they can do work according to just their convenience. This can lead to stagnant employee performance because there is no demand for greater performance. Therefore, the effect of job satisfaction is not significant because other factors affect employee performance.

H_2: Job Satisfaction (JS) has a negative effect on Turnover Intention (TI)

The results of the analysis in Table 12 shows the parameter coefficient of the influence of job satisfaction on turnover intention is -0.195. The coefficient indicates a negative influence between the two variables. While the t-statistic value generated is 6.934. The t-statistic value is greater than the t-table (6.934 > 1.96). Based on the t-statistic which is greater than the t-table, it can be concluded that the hypothesis is accepted. In other words, job satisfaction has a significant effect on turnover intention.

This can be interpreted as the greater level of job satisfaction can trigger a decrease in turnover intention. So, if an employee is satisfied with his job, then the employee does not want to move or leave the company. Thus, the effect of job satisfaction is significant to suppress or reduce the level of the turnover intention of employees.

H_3: Organizational Commitment (OC) has a positive effect on Employee Performance (EP)

The results of the analysis in Table 12 shows the parameter coefficient of the influence of organizational commitment variable on employee performance is 0.186. The coefficient indicates a positive influence between the two variables. While the t-statistic value generated is 3.375. The t-statistic value is greater than the t-table (3.375 > 1.96). Based on the t-statistic which is greater than the t-table, it can be concluded that the hypothesis is accepted. In other words, organizational commitment has a significant influence on employee performance.

This can be interpreted as the greater level of organizational commitment of an employee can trigger an increase in employee performance. So, if an employee is fully committed to the work and organization (company), then the results of the employee's performance will be great. Thus, the effect of organizational commitment is significant to increase employee performance.

H_4: Organizational Commitment (OC) has a negative effect on Turnover Intention (TI)

The results of the analysis in Table 12 shows the parameter coefficient of the influence of organizational commitment variable on turnover intention is -0.662. The coefficient indicates a negative influence between the two variables. While the t-statistic value generated is 22.090. The t-statistic value is greater than the t-table (22.090 > 1.96). Based on the t-statistic which is greater than the t-table, it can be concluded that the hypothesis is accepted. In other words, organizational commitment has a significant effect on turnover intention.

This can be interpreted as the greater level of organizational commitment of an employee can trigger a decrease in turnover intention. So, if an employee is fully committed to the work and organization (company), then the employee does not want to move or leave the company. Thus, the effect of organizational commitment is significant to suppress or reduce the level of the turnover intention of employees.

H_5: Turnover Intention (TI) has a negative effect on Employee Performance (EP)

The results of the analysis in Table 12 shows the parameter coefficient of the influence of the turnover intention variable on employee performance (original sample) is -0.235, which means that there is a negative influence between the two variables. The t-statistic value generated is 4.544, which means that the result is greater than the t-table (4.544 > 1.96) so it can be concluded that the hypothesis is accepted, or Turnover Intention has a significant effect on Employee Performance.

This can be interpreted as the greater level of the turnover intention of an employee can trigger a decrease in employee performance. So, if an employee wants to move or leave the company, then the employee performance will decrease or stagnate because there is no passion for their work. Thus, the effect of turnover intention is significant to reduce the level of employee performance.

4.3. Testing of Mediation Effect

Testing of the mediation effect is a process of analyzing the connecting variables. Based on the connecting variable, it can be determined the relationship between the independent variable and the dependent variable. This indicates that the independent variable's influence on the dependent variable can occur directly or indirectly through a connecting variable or the media. Table 13 shows the results of the mediation effect test, including the size of the p-value, t-statistics, and the indirect effect value.

Table 13. Path Coefficient and P-Value

	Original Sample (O)	**T Statistics (|O/STDEV|)**	**P Values**
Organizational Commitment -> Turnover Intention -> Employee Performance	0.156	4.080	0.000
Job Satisfaction -> Turnover Intention -> Employee Performance	0.046	3.701	0.000

Based on Table 13, it can be explained:

H_6: Organizational Commitment (OC) affects Employee Performance (EP) through Turnover Intention (TI).

The results of the analysis in Table 13 shows the t-statistic value is 4.080, which means that the result is greater than the t-table (4.080 > 1.96) and the indirect effect is 0.156. This means that H_6 is accepted or the mediating model of the influence of organizational commitment on employee performance through turnover intention is accepted.

H_7: Job Satisfaction (JS) affects Employee Performance (EP) through Turnover Intention (TI).

The results of the analysis in Table 13 shows the t-statistic value is 3.701 which means that the result is greater than the t-table (3.701 > 1.96) and the indirect effect is 0.046. This means that H_7 is accepted or the mediation model of job satisfaction's influence on employee performance through turnover intention is accepted.

4.4. Analysis Factor using Machine Learning

The outputs of subsections 4.2 and 4.3 are the results of SEM analysis related to the relationship of the Turnover Intention variable with other independent variables. It was found that Turnover Intention was significantly influenced by Job Satisfaction and Organizational Commitment factors. Furthermore, the classification analysis of the Turnover Intention variable was carried out using the Random Forest algorithm. Turnover Intention classification analysis is based on Job Satisfaction and Organizational Commitment factors. Meanwhile, the classification of Turnover Intention is divided into 2 classes, namely: (1) "Stay"; employees who choose to stay working at the company and (2) "Leave"; employees who choose to leave/move from their current company.

The predictor variables used are indicators of Organizational Commitment and Job Satisfaction factors, namely OC01, OC02, OC03, OC05, JS01, JS02, JS03, and JS04. Based on these predictor variables, a Random Forest model will be obtained which can predict the classification of an employee's Turnover Intention. Then each predictor variable will be sorted based on the most important variable in predicting Turnover Intention.

4.4.1. Preparation of Training Data and Testing data

This study uses the percentage composition to divide between training data and testing data. The composition used is 80% for training data and 20% for data testing. The original data of this study was 500, so a lot of training data and data testing can be seen in Table 14 below.

Table 14. Summary of Training Data and Testing data

	Training	**Test**	**Total**
Data	400	100	500
Percentage	80%	20%	100%

Based on Table 14, it is found that the number of training data is 400. While the testing data is 100. Furthermore, the distribution of the training and testing data is carried out randomly using the RStudio software with the function of *sample(2,nrow(data),replace=TRUE,prob=c(0.8,0.2))*. The results of the random division produce training data and data testing containing each class "Stay" and "Leave" which can be seen in Table 15.

Table 15. Summary of Training Data and Testing data

	Classification of Turnover Intention		**Total**
	Stay	**Leave**	
Training Data	180	220	400
Testing data	47	53	100

4.4.2. Result of Random Forest Algorithm

Training data were obtained based on Table 15, then entered for the Random Forest algorithm process. The learning process carried out by machine learning based on the Random Forest algorithm will produce a classification model. This Random Forest classification model can be used as a predictor of whether an employee's Turnover Intention level belongs to the "Stay" or "Leave" class. The Random Forest Algorithm was performed with the help of R software using the function of *randomForest(Turnover~.,data=train.datarf)*. The results of the Random Forest algorithm using the R software can be seen in Figure 4.

```
Call:
 randomForest(formula = Turnover ~ ., data = train.datarf)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 3

        OOB estimate of  error rate: 5.74%
Confusion matrix:
    1   2 class.error
1 170  11  0.06077348
2  12 208  0.05454545
```

Figure 4. Random Forest Output

Based on Figure 4, it is found that the output of the Random Forest model type is classification. The number of trees grown is 500. Then the number of variables that are tested for each branch is 3. In addition, the OOB (Out Of Box) error rate is the estimated error that can be obtained if the data used is outside of the test data. The OOB obtained is 5.74%, this value is considered very good because the estimated error is less than 10%.

4.4.3. Plot of Random Forest Model

The Random Forest model obtained can be projected into a graph plot of each branching variable. The function used in R Software is *plot()*. The results of the Random Forest model graph projection can be seen in Figure 5.

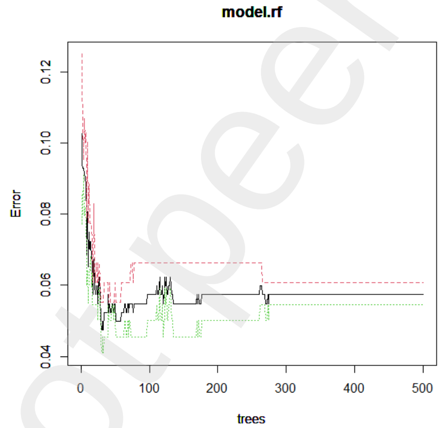

Figure 5. Plot of Random Forest Model

Based on Figure 5, it is found that the form of the Random Forest model has started to stabilize at the 300th tree. This indicates that the machine learning process for studying classification data training using the Random Forest model is stable or good. So that the Random Forest model can be used for general classification of other data.

4.4.4. Tree Plot of Random Forest Model

In addition to plotting the model graph against error, Random Forest can also be analyzed through tree plots. The function used in R Software is reprtree.:::plot.getTree(). The tree plot projection results from the Random Forest model can be seen in Figure 6.

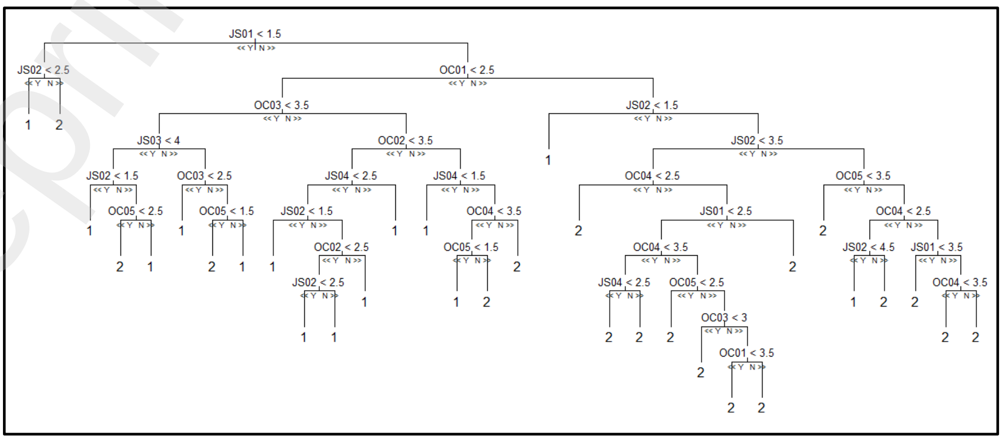

Figure 6. Tree Plot of Random Forest Model

Based on Figure 6, the tree plot form of the Random Forest model is stable. This tree plot shows the decision tree for each branching variable. The tree plot shows classifications 1 and 2, each representing the level of Turnover Intention for “Stay” and “Leave”. This indicates that the machine learning process for studying classification data training using the Random Forest model is stable or good. So that the Random Forest model can be used for the general classification of other data.

4.4.5. Result Evaluation of Random Forest Model

The result of applying the Random Forest model is to predict the classification based on the predictor variables. Prediction of the classification of the Random Forest model is carried out concerning the testing data. The function used in R software is *predict(model, test.data)*. The classification prediction results from the Random Forest model are presented in the form of a confusion matrix which can be seen in Table 16.

Table 16. Confusion Matrix Output

Confusion Matrix		Actual Class	
		Stay	Leave
Prediction Class	Stay	46	1
	Leave	1	52

Based on Table 16, it was found that the number of employees who belonged to the "Stay" class from the testing data was 47. The correct prediction results for the "Stay" class were 46 and the wrong predictions were 1. As for the employees belonging to the "Leave" class from the testing data is 53. The correct prediction results for the “Leave” class are 53 and the wrong prediction is 1. Furthermore, the measurement of the performance evaluation of the Random Forest model can be determined based on the level of accuracy, sensitivity, and specificity as follows.

$$\text{Accuracy} = \frac{46+52}{46+1+1+52} = \frac{98}{100} = 98\%$$

$$\text{Sensitivity} = \frac{46}{46+1} = \frac{47}{46} = 97.87\%$$

$$\text{Specificity} = \frac{52}{52+1} = \frac{52}{53} = 98.11\%$$

The results of the evaluation of the performance of the Random Forest model, each measure has exceeded 90%. The accuracy level is 98%, the sensitivity level is 97.87%, and the specificity level is 98.11%. Based on the classification of the accuracy level in Table 5, this Random Forest model is included in a very good value, which is more than 90%. Therefore, this model can already be used to determine a person's level of Turnover Intention in general on other real field data..

4.4.6. Importance Variabel Output

The Random Forest model produces an analysis of the importance of variables as the best predictor results that support the decision variables. The function used in R software is *View(importance(model))*. The results of the importance variable analysis to rank the best predictors for determining the level of Turnover Intention can be seen in Table 17.

Table 17. Importance Variable Output

Ranking	Predictor Variable	Level of Importance Variable (Mean Decrease Accuracy)
1	OC01	39.518173
2	JS02	33.334789
3	JS04	25.166420
4	OC02	23.092035
5	OC05	18.153701
6	JS01	15.089174
7	JS03	10.595010
8	OC03	6.644303

Based on Table 17, it is found that the best predictor to determine the level of Turnover Intention of an employee is OC01. OC01 has an importance variable level of 39.518173. The predictor variable OC01 means that employees have a high level of "strong belief in a career in the company" or believe that future career projections at the company will be better.

This can be said as a cause and effect, namely if the employee has confidence in the better career projections, then the employee will not have the desire to move or resign from the company.

5. Discussion

The analysis results can be used as a basis for discussion as follows:

Based on the results of the Random Forest model in Figure 6, it is found that the plot tree can be used as a way of classifying the level of Turnover Intention of an employee. The workings of the decision selection based on the plot tree is, following each branch of the variables that represent the decision tree. Each branch of the decision tree has a certain criterion value so the independent variable value affects the decision branch. The plot tree algorithm process is simplified with the help of R software using the *predict()* package. The *predict()* package can provide the final decision on the classification of the dependent variable, in this case, Turnover Intention. The classification of decisions taken is “Stay” or “Leave”. The two classifications describe the condition of an employee's expectations regarding his future. "Stay" has an interpretation that employees will stay at the company while showing loyalty, on the contrary, "Leave" means that employees are not comfortable with their current job.

The process of predicting the Turnover Intention level will be very useful for human resource management in a company. HR management can get an overview of the loyalty of its employees, thus helping to evaluate employee performance. In addition, HR management has a reference for making company policies that are following conditions that affect employee turnover decisions. Based on practical theory in HR management, a high level of Turnover Intention will disrupt the stability of the company's performance. This is due to the changing human resources and the teamwork process is disrupted. Therefore, HR management needs to reduce the level of Turnover Intention to maintain company sustainability.

In addition, the results of this study can also be used as a consideration for the process of recruiting new employees. Based on Table 17, the best predictor variable was obtained to determine the level of Turnover Intention. Therefore, the employee recruitment team can have a reference regarding the measurement of the loyalty of prospective employees. Implementation in the recruitment process can be done by giving more points to the OC01 indicator because it is the best predictor. Thus, if the recruiting team knows this measure, it is expected that the employees who are accepted have a good level of loyalty to the company. As a result, the human resources working in the company will have good performance, while at the same time maintaining the company's sustainability.

6. Conclusions

In this study, research was done to analyze the prediction model of employee turnover intention based on the factors that influence it, using a combination of SEM-PLS and machine learning methods. Based on SEM-PLS results were found that Turnover Intention was significantly influenced by Job Satisfaction (JS) and Organizational Commitment (OC) factors. Thus, each indicator from JS and OC is used to make a predictive model for the classification of turnover intention based on the Random Forest algorithm. The results of the Random Forest model have an accuracy of 98%, so it is classified as very good to be applied to other general data. Based on the analysis of the importance of variables, it was found that OC01 was the most influential variable on the level of turnover intention, with a mean decrease accuracy of 39.51. The predictor variable OC01 means that employees have a high level of "strong belief in a career in the company" or believe that future career projections at the company will be better. Therefore, human resource managers must pay attention to career projections and the suitability of employees' positions better, so that employees will give loyalty to the company. Then, it can better maintain the company's sustainability.

Acknowledgments

We would like to thank to PT. TSM Kalimantan and Sekolah Tinggi Ilmu Ekonomi Balikpapan, Indonesia, for supporting this research collaboration.

Informed Consent Statement

All subjects gave their informed consent for inclusion before they participated in the study.

Institutional Review Board Statement

References

Ahmad, N., Iqbal, N., Javed, K., & Hamad, N. (2014). Impact of organizational commitment and employee performance on the employee satisfaction. *International Journal of Learning, Teaching and Educational Research, 1*(1), 84-92.

Ahmad, S., Zulkurnain, N. N. A., & Khairushalimi, F. I. (2016). Assessing the validity and reliability of a measurement model in Structural Equation Modeling (SEM). *Journal of Advances in Mathematics and Computer Science*, 1-8.

Alsaraireh, F., Quinn Griffin, M. T., Ziehm, S. R., & Fitzpatrick, J. J. (2014). Job satisfaction and turnover intention among J ordanian nurses in psychiatric units. *International journal of mental health nursing*, *23*(5), 460-467.

Auret, L. and Aldrich, C. (2012). Interpretation of nonlinear relationships between process variables by use of random forests. *Minerals Engineering, 35,* 27-42.

Aydogdu, S., & Asikgil, B. (2011). An empirical study of the relationship among job satisfaction, organizational commitment and turnover intention. *International review of management and marketing*, *1*(3), 43-53.

Aziri, B. (2011). Job satisfaction: a literature review. *Management Research & Practice*, *3*(4).

Badrianto, Y., & Ekhsan, M. (2020). Effect of work environment and job satisfaction on employee performance in PT. Nesinak industries. *Journal of Business, Management, & Accounting*, *2*(1).

Balouch, R., & Hassan, F. (2014). Determinants of job satisfaction and its impact on employee performance and turnover intentions. *International Journal of Learning & Development*, *4*(2), 120-140.

Bonenberger, M., Aikins, M., Akweongo, P., & Wyss, K. (2014). The effects of health worker motivation and job satisfaction on turnover intention in Ghana: a cross-sectional study. *Human resources for health*, *12*(1), 1-12.

Bothma, C. F., & Roodt, G. (2013). The validation of the turnover intention scale. *SA Journal of Human Resource Management*, *11*(1), 1-12.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5–32.

Carmeli, A., & Weisberg, J. (2006). Exploring turnover intentions among three professional groups of employees. *Human Resource Development International*, *9*(2), 191-206.

Cho, Y. J., & Lewis, G. B. (2012). Turnover intention and turnover behavior: Implications for retaining federal employees. *Review of public personnel administration*, *32*(1), 4-23.

Cicek, I. (2011). The Effect of Outsourcing Human Resource on Organizational Performance: The Role of Organizational Culture. *International Journal of Business and Management Studies*, *3*(2), 131-143.

Cohen, A. (2007). Commitment before and after: An evaluation and reconceptualization of organizational commitment. *Human resource management review*, *17*(3), 336-354.

Cohen, A. (2017). Organizational Commitment and Turnover: A Met A-Analysis. *Academy of management journal, 36,* 1140–1157.

Colquitt, J., Lepine, J., & Wesson, M. (2014). *Organizational Behavior: Improving Performance and Commitment in the Workplace (4e)*. New York, NY, USA: McGraw-Hill.

Dessler, G. (2013). *Fundamentals of human resource management*. Pearson.

Ertas, N. (2015). Turnover intentions and work motivations of millennial employees in federal service. *Public personnel management*, *44*(3), 401-423.

Fu, W., & Deshpande, S. P. (2014). The impact of caring climate, job satisfaction, and organizational commitment on job performance of employees in a China's insurance company. *Journal of business ethics*, *124*(2), 339-349.

Gocheva-Ilieva, S., Ivanov, A., and Stoimenova-Minova, M. (2022). Prediction of Daily Mean PM10 Concentrations Using Random Forest, CART Ensemble and Bagging Stacked by MARS. *Sustainability*, *14*(2), 798.

Hair Jr, J. F., Matthews, L. M., Matthews, R. L., & Sarstedt, M. (2017). PLS-SEM or CB-SEM: updated guidelines on which method to use. *International Journal of Multivariate Data Analysis*, *1*(2), 107-123.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European business review*, *31*(1), 2-24.

Inuwa, M. (2016). Job satisfaction and employee performance: An empirical approach. *The Millennium University Journal*, *1*(1), 90-103.

Islam, J., Mohajan, H., & Datta, R. (2011). A study on job satisfaction and morale of commercial banks in Bangladesh.

Khatimah, H., Susanto, P., & Abdullah, N. L. (2019). Hedonic motivation and social influence on behavioral intention of e-money: The role of payment habit as a mediator. *International Journal of Entrepreneurship*, *23*(1), 1-9.

Kremic, T., Tukel, O. I., & Rom, W. O. (2006). Outsourcing decision support: a survey of benefits, risks, and decision factors. *Supply Chain Management, 11*, 467-482.

Kulkarni V. Y. and Sinha, P. K. (2013). Random Forest Classifiers: A Survey and Future Research Directions. *International Journal of Advanced Computing, 36*(1), 144-153.

Kuo, H. T., Lin, K. C., & Li, I. C. (2014). The mediating effects of job satisfaction on turnover intention for long-term care nurses in T aiwan. *Journal of nursing management*, *22*(2), 225-233.

Lai, M.-C., & Chen, Y.-C. (2012). Self-efficacy, effort, job performance, job satisfaction, and turnover intention: The effect of personalcharacteristics on organization performance. *International Journal of Innovation, Management and Technology*, *3*(4), 387.

Liu, B., Liu, J., & Hu, J. (2010). Person-organization fit, job satisfaction, and turnover intention: An empirical study in the Chinese public sector. *Social Behavior and Personality: an international journal*, *38*(5), 615-625.

Mangkunegara, A. P., & Octorend, T. R. (2015). Effect of work discipline, work motivation and job satisfaction on employee organizational commitment in the company (Case study in PT. Dada Indonesia). *Marketing*, *293*, 31-36.

Mathis, R. L., Jackson, J. H., & Valentine, S. R. (2015). *Human resource management: Essential perspectives*. Cengage Learning.

Meyer, J. P., & Parfyonova, N. M. (2010). Normative commitment in the workplace: A theoretical analysis and re-conceptualization. *Human resource management review*, *20*(4), 283-294.

Mondego, D. Y., & Gide, E. (2020). Exploring The Factors That Have Impact On Consumers'trust In Mobile Payment Systems In Australia. *JISTEM-Journal of Information Systems and Technology Management*, *17,* 1-19.

Nazir, S., Shafi, A.L., Qun, W., Nazir, N., & Tran, Q. (2016). Influence of organizational rewards on organizational commitment and turnover intentions. *Employee Relations, 38,* 596-619.

Podsakoff, N. P., LePine, J. A., & LePine, M. A. (2007). Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: a meta-analysis. *Journal of applied psychology*, *92*(2), 438.

Prajwala, T. R. (2015). A Comparative Study on Decision Tree and Random Forest Using R Tool. *International Journal of Advanced Research in Computer and Communication Engineering, 4*(1), 196-199.

Pratama, S. (2019). Effect of Organizational Communication and Job Satisfaction on Employee Achievement at Central Bureau of Statistics (BPS) Binjai City. *International Journal of Research and Review, 7*(11), 547-550.

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011). Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling, 18*(2), 161-182.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological methods*, *15*(3), 209.

Qureshi, M. I., Iftikhar, M., Abbas, S. G., Hassan, U., Khan, K., & Zaman, K. (2013). Relationship between Job Stress, Workload, Environment and Employees Turnover Intentions: What We Know, What Should We Know. *World Applied Sciences Journal*, *23*(6), 764-770.

Reinders, C., Ackermann, H., Yang, M. Y., Rosenhahn, B. (2019). *Chapter 4 - Learning Convolutional Neural Networks for Object Detection with Very Little Training Data*. Multimodal Scene Understanding: Academic Press. ISBN 9780128173589, 65-100,

Richard, O. C., & Johnson, N. B. (2001). Strategic human resource management effectiveness and firm performance. *International Journal of Human Resource Management*, *12*(2), 299-310.

Shahzadi, I., Javed, A., Pirzada, S. S., Nasreen, S., & Khanam, F. (2014). Impact of employee motivation on employee performance. *European Journal of Business and Management*, *6*(23), 159-166.

Siengthai, S., & Pila-Ngarm, P. (2016). *The interaction effect of job redesign and job satisfaction on employee performance*. Evidence-based HRM: a Global Forum for Empirical Scholarship.

Singh, P., Singh, N., Singh, K. K., & Singh, A. (2021). *Chapter 5 - Diagnosing of disease using machine learning*. Machine Learning and the Internet of Medical Things in Healthcare: Academic Press. ISBN 9780128212295, 89-111.

Son, J.-M. (2012). Job embeddedness and turnover intentions: an empirical investigation of construction IT industries. *International Journal of Advanced Science and Technology*, *40*, 101-110.

Svetnik, V., Liaw, A., Tong, C., Culberson, J., Sheridan, R., & Feuston, B. (2003). Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *Journal of chemical information and computer sciences, 43*, 47-58.

Tarigan, V., & Ariani, D. W. (2015). Empirical study relations job satisfaction, organizational commitment, and turnover intention. *Advances in Management and Applied Economics*, *5*(2), 21.

Valentine, S., Godkin, L., Fleischman, G. M., & Kidwell, R. (2011). Corporate ethical values, group creativity, job satisfaction and turnover intention: The impact of work context on work response. *Journal of business ethics*, *98*(3), 353-372.

Wongvibulsin, S., Wu, K.C. & Zeger, S.L. (2020). Clinical risk prediction with random forests for survival, longitudinal, and multivariate (RF-SLAM) data analysis. *BMC Med Res Methodol, 20*(1), 1-14.

Yahaya, R., & Ebrahim, F. (2016). Leadership styles and organizational commitment: literature review. *Journal of Management Development*, *35*(2), 190-216.

Yang, M., Mamun, A. A., Mohiuddin, M., Nawi, N. C., & Zainol, N. R. (2021). Cashless transactions: A study on intention and adoption of e-wallets. *Sustainability*, *13*(2), 831.

Yi, Y., Nataraajan, R., & Gong, T. (2011). Customer participation and citizenship behavioral influences on employee performance, satisfaction, commitment, and turnover intention. *Journal of Business Research*, *64*(1), 87-95.

Yücel, İ. (2012). Examining the relationships among job satisfaction, organizational commitment, and turnover intention: An empirical study. *International Journal of Business and Management, 7*(20), 45-58.

Prediction of Employee Loyalty Classification for Maintaining Company Sustainability using Combination of SEM-PLS and Machine Learning

Sudarmo Sudarmo [1*], Rachmie Sari Baso [2] and Muhammad Adenuddin Alwy [2]

[1]Department of Economics, Sekolah Tinggi Ilmu Ekonomi Balikpapan, Balikpapan, East Kalimantan, Indonesia
[2]Department of English Language Education, Foreign Language Academy of Balikpapan, Balikpapan, East Kalimantan, Indonesia
*Correspondence: sudarmo@stiebalikpapan.ac.id

Abstract: The company's sustainability is influenced by several internal and external factors, one of them is the turnover intention which has an impact on the weak loyalty of human resources. Based on the results of a survey, it shows the level of turnover intention is influenced by many factors. A high level of turnover intention will threaten the sustainability of the company. This should be of particular concern to human resource managers. Therefore, this study aims to analyze the prediction of turnover intention based on the factors affecting employees. The method used is a combination of SEM-PLS and Machine Learning based on the Random Forest algorithm. The empirical model developed in this study was based on four research variables. The variables used are Job Satisfaction (JS), Organizational Commitment (OC), Employee Performance (EP), Turnover Intention (TI). Then, analysis of the variable relationship using SEM-PLS resulted in JS and OC having a significant effect on the level of turnover intention. Thus, each indicator from JS and OC can be used to create a predictive model of turnover intention classification based on the Random Forest algorithm. The results of the Random Forest model obtained have a very good prediction accuracy of 98%. Based on the analysis of the important variables, it was found that OC01 was the most influential variable on the level of turnover intention, with a mean decrease accuracy of 39.51. The predictor variable OC01 means that employees have a good level of "strong belief in a career in the company". Therefore, human resource managers must pay better attention to career projections and the suitability of employee positions. So, the employees will give loyalty to the company, then it can better maintain the company's sustainability.

Keywords: Loyalty; Turnover Intention; Job Satisfaction; Organizational Commitment; SEM-PLS; Machine Learning; Random Forest.

---

1. Introduction

Management of Human Resources (HR) in a company can determine the sustainability of the company. The effectiveness and efficiency of HR management planning can be shown from the results of the employee recruitment process following the company's expectations (Cicek, 2011). Employees who have been recruited should be able to provide good performance in their work. In addition, good human resources are employees who are loyal to the company. If the loyalty of a worker is not good, it will have an impact on changes in the dynamic staffing structure. Too many employee changes shortly can affect the company's performance. Newly joined employees need a process and time to adapt so they can provide maximum performance. The process of changing too many employees needs to be avoided by the company's HR managers, as a step to streamline the time and cost of the recruitment process (Kremic et al., 2006).

The desire to switch jobs is an early indicator of employee turnover in a corporation. The degree of attitude propensity possessed by employees to hunt for new jobs elsewhere or their plans to resign from the company is referred to as Turnover Intention (TI). Another attitude that emerges in individuals when turnover intention arises is a desire to explore for alternative job openings, weighing the likelihood of obtaining a better job elsewhere (Carmeli & Weisberg, 2006). However, if the possibility to change jobs is not available, or if what is offered is not more appealing than what you presently have, employees who wish to quit the company are frequently late, frequently miss work, are less enthusiastic, or lack the will to strive hard. (Ertas, 2015) As a result, it is apparent that turnover intention will have a negative influence on the firm since it produces instability and uncertainty in labor circumstances, lowers employee productivity, and creates an unfavorable work environment. High employee turnover also reduces the effectiveness of the company's organization. Losing experienced employees and having to retrain new employees will add time and money to the process.

High employee turnover grabs the attention of the company because it disrupts operations, creates moral problems for employees who stay, and also inflates costs in recruitment, as well as administrative costs for processing new employees (Cho & Lewis, 2012). The turnover intention has an impact on the weak performance of employees. The way to anticipate is to know the factors that influence turnover intention. Then these factors can be used to reduce the level of turnover intention. Existing studies and literature show that a person's turnover intention is closely related to job satisfaction and organizational commitment where these three things will affect employee performance later (Nazir et al., 2016). Employee performance can reflect the sustainability of the company. In other words, good employee performance can provide good sustainability for the company (Dessler, 2013). Several approaches can be taken to achieve effective and efficient employee performance. One of the approaches is management needs to get commitment from employees to the company. Employee commitment is shown by accepting the company's goals, company interests, and loyalty to the company (Colquitt et al., 2014).

Based on organizational commitment can provide a strong prediction of the desire for voluntary turnover. There is a tendency for commitment before entering the company to be positively related to initial commitment, while commitment after entering the company will be negatively related to voluntary turnover. (Cohen, 2017). Then, commitment in the early stages of joining the company will influence employee job satisfaction. Companies that can make employees believe in the company vision will get employee commitment. So, employees will be loyal to the company and work well for the benefit of the company. This condition is ideal for achieving company goals because the company has complete support from its members, allowing them to focus solely on priority goals. Commitment is a key aspect for company businesses because of its impact on performance, which assumes that loyal employees put in more effort at work.

Strong commitment goes hand in hand with strong productivity. Commitment also gives an indication of lower turnover intention. Organizational commitment is useful for predicting the sustainability of the company related to employee performance (Meyer & Parfyonova, 2010). Analysis of the outcome variable in this study is related to the individual's turnover intention. Organizational commitment is strongly related to an individual's desire to leave the company's operations. Job satisfaction might also impact a person's decision to leave. Individuals who choose to leave the organization company may expect better satisfactory results elsewhere. The turnover decision is a result of their evaluation of numerous alternative positions. One of the reasons for looking for alternative work is satisfaction with the salary received. Employees feel a sense of justice (equity) to the salary received in accordance with the work done. Salary satisfaction can be interpreted that employees get a salary as expected. The level of job satisfaction of these employees is a factor that affects employee turnover. If the employee turnover rate is high, it indicates a fundamental problem in a company. In addition, the employee turnover process also takes a lot of time and costs. Therefore, companies need to reduce employee turnover rates according to company needs. So, it is necessary to know more deeply what affects the Turnover Intention of the employees. If the employee turnover rate can be controlled, then HR management will be better to increase company productivity.

The method approach that can be used to analyze the relationship between variables that affect the turnover intention of employees is SEM-PLS. SEM-PLS tests the validity and reliability of the indicators that make up the construct of the model. The SEM-PLS model can produce an analysis of the relationship between variables that significantly affect other variables. Furthermore, the variable which is a significant factor is used to determine the prediction of the classification of the level of the turnover intention of employees. The development method used is a machine learning based on the Random Forest model. Machine learning will practice understanding the classification of factors from the independent variable data on the dependent variable (Turnover Intention). The machine learning process will produce a Random Forest model which is used to classify the dependent variable based on the independent variable. So that can be predicted the level of turnover intention based on the influence of the independent variable.

Based on this description, the motivation and purpose of this study are to analyze the prediction of turnover intention classification using the SEM-PLS and machine learning methods. The development of this research topic is compared with differences from other studies that have been carried out, namely the combination of SEM-PLS and machine learning based on the Random Forest model. Therefore, the novelty of this study is the development of a new method to obtain a predictive analysis model for turnover intention classification. This is the advantage of this research because the Random Forest model uses independent variables whose significance has been analyzed first based on SEM-PLS. In addition, the results of machine learning with the Random Forest model can provide an analysis of the important variable to determine the best predictor. Thus, these results can be used as a reference for the focus of human resource management, so that companies can control the level of employee turnover intention. Thus, employees will give loyalty to the company. As a result, it can better maintain the company's sustainability.

2. Literature Overview

2.1. Basic Concepts

2.1.1. Turnover Intention

An employee can be said to have a turnover intention, if he wants to move or leave the company but has not yet been able to make it happen (Son, 2012). The turnover intention according to several definitions can be concluded as a person's desire to leave the company where they work to find another job that is better than before (Podsakoff et al., 2007). Desire to leave can be seen from the indicators that are addressed through behaviors such as thinking out of work, actively looking for another job, leaving work in the near future, always looking for opportunities to leave the company, lack of enthusiasm for working in the current job and leaving the company when getting a salary larger ones. There are several factors that cause employees to have a desire to leave, namely:

1. Structural Factors, namely factors related to work and organization, such as support from colleagues, support from superiors, work routines, equitable distribution of justice, role ambiguity, workload, employee skills, rewards, job security, and career development.
2. Pre-entry factors, namely factors that include positive personalities such as a tendency to be happy, and negative personalities such as a tendency to experience discomfort, and so on.
3. Environmental factors, namely factors related to things outside the work and organization. Environmental factors include job opportunities available outside the company, habits of moving from people around employees, and the number of family members covered.
4. Trade Union Factors, namely factors related to an employee's membership of a trade union that can influence an employee's decision to keep a job or decide to move.
5. Job Orientation, namely factors that include job satisfaction, organizational commitment, and employee activities or efforts to find alternative jobs outside the organization where they work today.

In addition to the things above, there are many causes of turnover intention, including work stress, work environment, job satisfaction, organizational commitment, and others experienced by employees (Qureshi et al., 2013).

2.1.2. Organizational commitment

Organizational or company management is largely determined by HR management. Based on studies related to HR management, organizational commitment is an important issue because it is an aspect that affects human behavior in organizations. The main reason is quite simple, because the achievement of the organization's vision, mission, and goals is carried out by members of the organization. So that if there is no commitment from members of the organization, it will not achieve organizational goals (Richard & Johnson, 2001). Employee commitment to the organization or company will affect the organization's capacity to achieve its goals and vision. At this time several companies dare to include an element of commitment as one of the requirements for job vacancies being offered. But not infrequently, there are still business actors who still do not understand the importance of this commitment properly. Several studies have provided views to understand employee commitment to the organization. Organizational commitment is identified by the match between employee goals and organizational goals. In addition, employees need to be willing to give all efforts for the sustainability of the organization. The level of employee loyalty can be seen from the interest in being able to continue to be part of the organization or company (Cohen, 2007). Another definition of organizational commitment is employee loyalty which is shown by their concern for the success of the organization on sustainable progress. (Yahaya & Ebrahim, 2016).

2.1.3. Job satisfaction

Job satisfaction has a negative impact on the level of turnover intention. In other words, if employees have a good level of job satisfaction, they will not want to change jobs. However, on the other hand, there are external factors that can influence such as labor market conditions and alternative job opportunities. This can make employees want to move because they are tempted by other offers even though their current job is fine (Aziri, 2011). The factors that encourage job satisfaction are:

1. Mentally challenging work.
2. Supportive working conditions.
3. Supportive co-workers.
4. Personality compatibility with work.

Employees who are satisfied with their work can be predicted to last long in the organization or company. Meanwhile, employees who are dissatisfied with their work will choose to move or leave the company. Perceived job satisfaction can be a factor for an employee to leave. In addition, some consideration of alternative jobs can also result in turnover intention. This is because employees expect more satisfaction jobs elsewhere. The researcher determines the indicators of each job satisfaction variable based on the opinions of experts regarding each of these variables which are then translated into questions in a questionnaire to the research sample.

1. Work Morale

Work morale includes the following three elements (Islam et al., 2011): (a) Loyalty and pride in the company. (b) Attitude towards superiors (c) Teamwork

2. Discipline

The factors that affect work discipline (Mangkunegara & Octorend, 2015) are as follows: (a) Punctuality at work. (b) Use of office facilities. (c) Responsibility for the completion of tasks. (d) Compliance with company regulations.

3. Work Performance

The indicators in work performance (Pratama, 2019) are (a) Work accuracy. (b) Work Tidiness. (c) Speed of completing tasks. (d) Initiative at work.

2.1.4. Employee Performance

Performance can be said to be a result of motivation, ability, and action. An employee must have a certain degree of motivation and level of ability to complete a task or job. In addition, when the level of motivation and ability is not balanced, performance will also be disrupted. For example, when an employee has good abilities but is not accompanied by goodwill or motivation, then his performance will be less than it should be. Some indicators of employee performance can involve factors of quality and quantity of output. Examples of employee performance indicators are attendance at work, helpfulness, cooperation, punctuality, and others (Shahzadi et al., 2014).

The definition of performance indicators can be said as aspects used to measure the performance of an employee. The form of indicators that measure employee performance (Mathis et al., 2015) are as follows:

a. Quantity is a measure of the number of products produced. This quantity can be expressed in the number of production units, the number of activity cycles completed by employees, and the number of activities produced to support the company.
b. The quality of work is expressed as the employee's perception of the quality of the work produced. This measure of work quality is related to the perfection of the work done. The factors that determine the quality of work are the skills and abilities of employees.
c. Timeliness is the employee's perception of an activity carried out within the specified period. For example, a job can be completed at the right time until it becomes the expected work result.
d. Attendance is also a performance indicator. This attendance consists of coming to work, leaving work, permission, and other information related to the employee's presence.
e. The ability to cooperate is the ability of employees to work together with a team or other people in their work. Cooperation skills will support the completion of tasks and jobs to achieve maximum results.

2.2. Relationship Between Variables

2.2.1. Effect of Job Satisfaction and Turnover Intention

Employee turnover can arise because of alternative jobs and or lack of job satisfaction. Based on the size of job dissatisfaction can be identified, how a worker has a desire to leave his job. In addition, job dissatisfaction indicators have a direct influence on the formation of a desire to move to other alternative jobs (Alsaraireh et al., 2014). In other words, employee turnover is also negatively related to job satisfaction. In addition, some of the important barriers in the decision to change a job are labor market conditions, spending on alternative employment opportunities, and the length of the work contract offered.

Identification of employee attitudes related to job satisfaction is often studied by identifying the relationship between satisfaction and turnover based on psychological variables (Liu et al., 2010). Employees who are dissatisfied with their jobs tend to do things that can interfere with the company's performance. Some of them are high turnover, slackness at work, high absenteeism, job complaints, and not working properly. It can be said, the higher the level of job satisfaction, the lower the desire to change jobs (Lai & Chen, 2012). Those who are dissatisfied with their jobs are more likely to depart and seek employment elsewhere. Several studies have indicated that work satisfaction is a potential predictor of intention to leave (Bonenberger et al., 2014; Kuo et al., 2014; Valentine et al., 2011).

2.2.2. Effect of Organizational Commitment and Turnover Intention

Organizational commitment is an attitude that demonstrates employee loyalty and is a continuing process by which an organization member communicates their interest in the organization's success and goodness. (Yahaya & Ebrahim, 2016). Individuals who have low organizational commitment tend to look for better work opportunities and leave their jobs because they have an embedded desire to leave the organization. Furthermore, increased job satisfaction and organizational commitment are likely to diminish employees' intent or purpose to leave the organization or company. Employees who are unsatisfied with aspects of their work and are uncommitted to their organization are more likely to seek employment elsewhere. Several empirical research has demonstrated that individuals who satisfy organizational commitment have a high level of job satisfaction and a lower propensity to leave (Aydogdu & Asikgil, 2011; Tarigan & Ariani, 2015; Yücel, 2012).

2.2.3. Effect of Turnover Intention and Employee Performance

Turnover intention is a precursor to staff turnover in a company. The attitude of propensity possessed by employees to hunt for new jobs elsewhere is referred to as turnover intention (Bothma & Roodt, 2013). Another attitude that emerges in employees when turnover intention arises is the desire to find alternative work vacancies, as well as evaluating the probability of getting a better job elsewhere. However, if the opportunity to change jobs is not available, or if what is available is not more appealing than what they currently have. Employees will leave the company emotionally and mentally, namely by frequently arriving late, skipping work, being less enthusiastic, or lacking the desire to try hard (Yi et al., 2011). As a result, turnover will have a negative influence on the organization since it generates insecurity in labor conditions. If the employees feel an unsuitable working environment, then has an impact on decrease human resource productivity (Balouch & Hassan, 2014).

2.2.4. Effect of Job Satisfaction and Employee Performance

Siengthai & Pila-Ngarm (2016) stated about the relationship between job satisfaction and employee performance by the state of the company. The employees who are more satisfied tend to be more effective than companies with less satisfied employees. Job satisfaction has a role in achieving better productivity and quality standards, avoiding the possibility of building a more stable workforce, and using human resources more efficiently (Mangkunegara & Octorend, 2015). Other research shows that job satisfaction is an independent variable that has a positive effect on management attitudes towards company strategy which is reflected through employee performance (Badrianto & Ekhsan, 2020; Inuwa, 2016).

2.2.5. Effect of Organizational Commitment and Employee Performance

Commitment is an employee's attitude with the desire to stay a member to attain the goal's company. Employees will be more motivated to attain organizational goals if they believe their attitudes and values are consistent with those of the organization. Several studies have found that corporate commitment has a major impact on employee performance (Ahmad et al., 2014; Fu & Deshpande, 2014).

3. Materials and Methods

This study was carried out to test the intended hypothesis by using research methods that have been designed according to the variables to be studied to obtain accurate research results.

3.1. Population, Sample, and Ethics Committee of Respondents

The scope of this research is employees of Tani Sejahtera Mandiri (TSM) Kalimantan Ltd., with a total of respondents is 500. Organizational commitment, job satisfaction, turnover intention, and employee performance are among the characteristics investigated. The sampling technique employed is stratified random sampling using a proportionate sampling approach, with the goal of representing employees in each portfolio. All subjects gave their informed consent for inclusion before they participated in the study.

The participant consent document was written in the cover letter in the survey. The form includes all elements of a regularly signed consent, including the confidentiality statement given below. The consent line says, "By completing the survey you are agreeing to participate in the research". The survey includes an "I agree" or "I disagree" option for participants to make their choice whether they agree to participate or not. Then, this study approved by the Ethics Committee of Tani Sejahtera Mandiri Ltd. (Date of approval: January 10th, 2021).

3.2. Research Variables and Indicators

In this study, an analysis of the relationship between the influence of the independent variable and the intervening variable was carried out on the dependent variable. In addition, this study also examines the indicators that influence these variables. The variables studied include organizational commitment, work satisfaction, turnover intention, and employee performance. Furthermore, the characteristics are simplified into research indicators, as shown in Table 1.

Table 1. Research variables and indicators

Latent Variables	Indicator
Organizational Commitment (OC)	OC1 Strong belief in a career in the company
	OC2 Level of involvement on company issues
	OC3 Level of interest in the company
	OC4 The feeling of being part of the company
	loyalty to the organization
Job Satisfaction (JS)	JS1 Satisfaction with salary
	Satisfaction with promotions and opportunities for advancement
	Satisfaction with colleagues and superiors in the company
	JS4 Satisfaction with working conditions/environment
Turnover Intention (TI)	The tendency of individuals to think about leaving the organization where they work now
	The probability that the individual will look for work in another organization
Employee Performance (EP)	EP1 Quality and quantity of employee work
	EP2 Attitude
	EP3 Cooperation and Communication between employees
	EP4 Creativity

3.3. Research Models and Hypotheses

In this section, the researcher proposes a model that is taken based on the results of a literature review and previous research. The proposed model uses organizational commitment, work satisfaction, turnover intention, and employee performance variables. The model studied is a model that was compiled based on a review of the literature and previous research, Figure 1 is a path diagram of the model studied in this study.

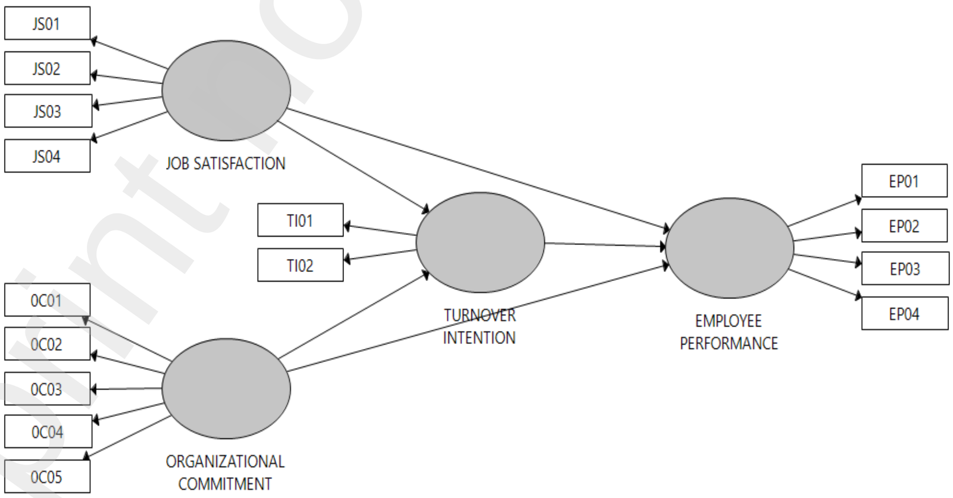

Figure 1. Path Diagram of the Research Model.

Based on Figure 1, along with the description in the literature review section, it can be seen that the hypotheses proposed in this study are:

H_1: JS has a positive effect on EP
H_2: JS has a negative effect on TI
H_3: OC has a positive effect on EP
H_4: OC has a negative effect on TI
H_5: TI has a negative effect on EP
H_6: OC affects EP through TI
H_7: JS affects EP through TI.

3.3. Structural Equation Modelling Partial Least Squares (SEM-PLS)

The analysis of this research uses the method of Structural Equation Modeling (SEM). SEM model processing is done using Smart PLS 3.0 software. The classification of SEM models is divided into two types; (1) Covariance-Based SEM (CB-SEM) and (2) Variance-Based SEM, often known as Partial Least Squares (SEM-PLS). There are various factors to consider while deciding which SEM approach to use. In terms of objectives, CB-SEM is employed if the researcher wants to test the theory, confirm the theory, or compare numerous alternative hypotheses. In contrast to CB-SEM, SEM-PLS is exploratory or an expansion of an existing theory. The SEM-PLS method was employed in this study (Hair Jr et al., 2017).

PLS is a powerful analytical method because it is not predicated on numerous assumptions. PLS can be used to confirm theory as well as to explain the presence or absence of a link between latent variables (Hair et al., 2019). Because it focuses more on data and with limited estimation procedures, the model specifications have little effect on parameter estimation.

In general, there are two main processes that must be completed when assessing the SEM method: (1) the evaluation process for the measurement model and (2) the evaluation process for the structural model (Ahmad et al., 2016). The two processes are carried out sequentially. Before conducting a structural model analysis, it is necessary to create a measurement model first. The measurement model was tested for the validity and reliability of latent variable indicators (Khatimah et al., 2019). The summary of the evaluation criteria for the measurement model is shown in Table 2.

Table 2. Summary of Measurement Model Evaluation Criteria

Validity and Reliability	Parameter	Criteria
Convergent Validity	Loading Factor	a. Greater than 0.70 for Confirmatory Research b. Greater than 0.60 for Exploratory Research
	Average Variance Extracted (AVE)	Greater than 0.50 for confirmatory and exploratory research
Discriminant Validity	Cross Loading	Loading to another construct is lower than its loading value on the construct
Reliability	Cronbach's Alpha	a. Greater than 0.70 for Confirmatory Research b. Greater than 0.60 still acceptable for exploratory research
	Composite Reliability	a. Greater than 0.70 for Confirmatory Research b. Greater than 0.60 still acceptable for Exploratory Research

In addition, the development of a structural model was used to investigate the relationship between latent variables in the measurement model. The preparation of these latent variables can be influenced by other construction factors. The form of the parameter indicating the regression of the exogenous latent variable is labeled γ ("gamma"), while those indicating the regression of endogenous latent variables are labeled β ("beta"). Assume that the random vectors $\boldsymbol{\eta}^T = (\eta_1, \eta_2, \ldots, \eta_m)$ and $\boldsymbol{\xi}^T = (\xi_1, \xi_2, \ldots, \xi_n)$ are endogenous and exogenous variables, respectively, and that they form a simultaneous equation with a system of linear equations. As a result, the generic version of the structural equation model is defined as follows:

$$\eta = \mathrm{B}\eta + \Gamma\xi + \zeta \tag{1}$$

In structural equations, $\mathbf{B}$ and $\boldsymbol{\Gamma}$ are coefficient matrices, and $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \ldots, \zeta_m)$ are error vectors. Element $\mathbf{B}$ represents the variable's effect on the other variables. Furthermore, the element has a direct influence on the variable $\boldsymbol{\Gamma}$ in the variable $\boldsymbol{\eta}$. It is believed that is Γ unrelated to $\boldsymbol{\zeta}$ and that **1-B** is nonsingular (Mondego & Gide, 2020).

The structural model's form is obtained in the following order:

$$\begin{aligned} \eta &= B\eta + \Gamma\xi + \zeta, \\ \eta - B\eta &= \Gamma\xi + \zeta, \\ (1 - B)\eta &= \Gamma\xi + \zeta, \\ \eta &= ((1 - B)^{-1}(\Gamma\xi + \zeta), \end{aligned} \tag{2}$$

with:

$\boldsymbol{B}$: endogenous latent variable coefficient matrix of size $m \times n$

$\boldsymbol{\Gamma}$: coefficient matrix of exogenous latent variable with size $m \times n$

$\boldsymbol{\eta}$: endogenous latent variable vector of size $m \times 1$

$\boldsymbol{\xi}$: exogenous latent variable vector of size $n \times 1$

$\boldsymbol{\zeta}$: vector random of residual relationship between $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ of size $m \times 1$

PLS was used to evaluate structural models, and the R-Square value for each dependent variable was used to determine the structural model's predictive potential. Determination of latent variables that have a significant effect on the dependent variable is known based on the R-Square value. The criteria for the R-Square value used for the strong, medium, and weak models are 0.67, 0.33, and 0.19, respectively (Yang et al., 2021). An overview of the evaluation criteria for the structural model is shown in Table 3.

Table 3. Summary of the Structural Model Evaluation Criteria

Parameter	Criteria
R-square	0.67 (Strong); 0.33 (Moderate); and 0.19 (Weak)
Significance Level	5% (0.05)

After evaluating the model, it is necessary to analyze the connecting variables to determine the effect between the independent variable and the dependent variable. Analysis of the relationship was carried out by examining the effects of mediation. The influential relationship between the independent variables on the dependent variable can be direct or indirect. The nature of the relationship, directly or indirectly, is determined by the mediating effect of the connecting variable (Preacher et al., 2011). The SmartPLS 3.0 software uses a procedure devised by Baron and Kenny in 1986 to test the mediating effect. According to Baron and Kenny, testing the mediating effect is carried out based on the following three stages:

a. The first model is used to investigate the effect of the independent variable (X) on the dependent variable (Y). The relationship criteria must be significant 5% (0.05)

b. The second model investigates the effect of the independent variable (X) on mediation (M) and must be significant at 5% (0.05).

c. The third model investigates the effect of the independent variable (X) on mediation (M) on the dependent variable at the same time (Y).

The relationship between the independent variable (X) and the dependent variable (Y) is estimated to be insignificant at the final stage of testing. However, the effect of the mediating variable (M) on the dependent variable (Y) is expected to be significant at 5%. Significance analysis of the influence between variables using the mediating variable (M) was carried out in the final test. The significance test was carried out with the help of the SmartPLS 3.0 program. Based on the analysis of the SmartPLS 3.0 program, the indirect effect coefficient and the significance value of the indirect effect will be obtained (with a significant level of 5%).

3.3. Machine Learning

Machine learning is a method to make machines (computers) have the ability to learn and do a job automatically. The machine learning process is carried out through certain algorithms so the work order to the computer can be done automatically.

Machine learning is carried out through 2 phases, namely the training phase and the application phase. The training phase is the modeling process of the algorithm that will be learned by the system through training data. Meanwhile, the application phase is the process of using the output model of the training phase, to produce a certain decision using testing data.

Machine learning can be done in two ways, namely supervised learning and unsupervised learning. Unsupervised learning is the processing of sample data without requiring the final result to conform with a certain form, by using several data samples at once. The application of unsupervised learning can be found in the visualization process, or data exploration. Supervised learning is processing sample data x will be processed in such a way, to produce output that matches the final result y. Supervised learning can be applied to the classification process. One of the supervised processes is Random Forest.

3.3.1. Random Forest

Random Forest is a bagging method, which is a method that generates a number of trees from sample data where the creation of one tree during training does not depend on the previous tree, then decisions are taken based on the most votes. The Random Forest algorithm was first introduced by Breiman (2001). Random Forest has two problem-solving functions, namely classification, and regression (Svetnik et al., 2003). Random Forest can be used on several types of data such as discrete, continuous, multivariate combinations, and survival data (Wongvibulsin et al., 2020). Random Forest can detect interactions between dependent and independent variables, and is able to explore data with good flexibility (Auret and Aldrich, 2012). Analysis using Random Forest does not use certain assumptions that must be met.

The Random Forest method is used to build a decision tree consisting of root nodes, internal nodes, and leaf nodes by taking attributes and data randomly according to the applicable provisions (Prajwala, 2015). The root node is used to collect data, an inner node located at the root node contains questions about the data, and a leaf node is used to solve problems and make decisions. The decision tree begins by calculating the entropy value as a determinant of the level of attribute impurity and the value of information gain as in equations (3) and (4) (Reinders et al., 2019).

$$Entropy\ (Y) = -\sum_i p(c|Y) log_2\, p(c|Y) \tag{3}$$

where Y is set of cases and $p(c|Y)$ is value proportion of Y to class of c.

$$Information\ Gain\ (Y,a) = Entropy\ (Y) - \sum_{v \in\ Values\ (a)} \frac{|Y_v|}{|Y_a|} Entropy\ (Y_v) \tag{4}$$

with $Values(a)$ is all possible values in the case set a, Y_v is sub-class of Y with class of v have relationship with a, and Y_a is all values corresponding to a.

The random forest has two main parameters, namely: m is the number of trees to be used and k is the maximum number of features that are considered when branching (Kulkarni and Sinha). The more m values, then the better of classification results, while the recommended k value is the square root or logarithm of the total number of features. The following is an example of a plot tree from Random Forest in Figure 2.

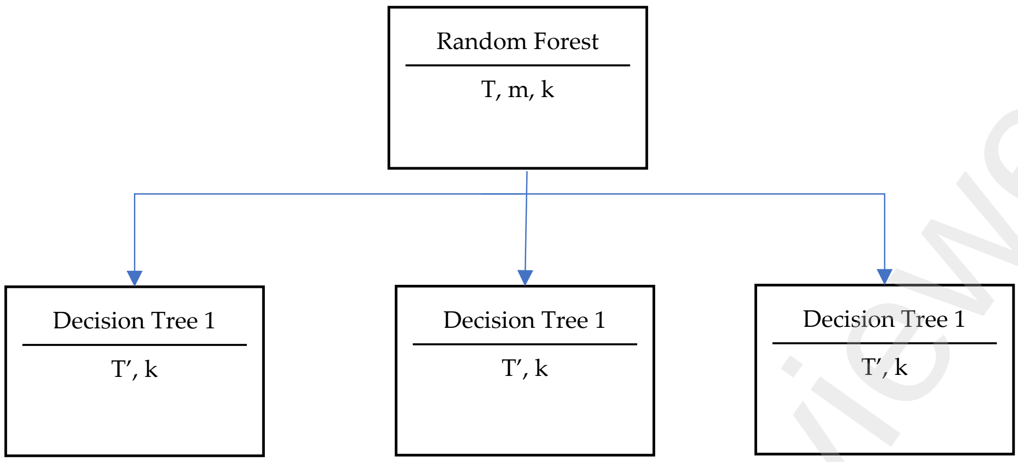

Figure 2. Illustration of Random Forest Plot Tree

Figure 2 shows the training process for a random forest using a dataset T with a number of m trees as the basic learner and k features chosen at random from the total features for branching in each tree. The training process on each tree uses the T' dataset which is the result of bootstrapping the dataset which is used as a parameter for a random forest.

3.3.2. Performance Evaluation of Random Forest Model

The evaluation was carried out for the selection of the best dataset distribution method and classification method seen through classification performance measures. The classification performance measure used in this study takes into account the confusion matrix. The confusion matrix is a useful tool to analyze how well or how accurately the classification method can recognize objects of observation from different classes. The following is an example of a confusion matrix for binary decision variables, which can be seen in Table 4 (Singh et al., 2021).

Table 4. Confusion Matrix

Confusion Matrix		Actual Class	
		True	False
Prediction Class	True	True Positive (TP) Correct Result	False Positive (FP) Unexpected Result
	False	False Negative (FN) Missing Result	True Negative (TN) Correct Result of "False"

Based on the prediction results from Table 4, then an evaluation of the performance of the classification model is carried out using the levels of accuracy, sensitivity, and specificity. The formula for measuring the performance evaluation is written as follows (Singh et al., 2021).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

Accuracy is the most commonly used method to assess classification performance. However, for the case of imbalanced data, accuracy places more weight on the majority class, so it is necessary to also pay attention to sensitivity and specificity. The analysis of the accuracy results can also be formed in the performance classification table as follows:

Table 5. Classification of Accuracy Level

Accuracy (%)	Classification
90 – 100	Very Good
80 – 90	Good
70 – 80	Enough
60 – 70	Bad
0 – 50	Very Bad

Based on Table 5, it is found that to achieve the model that is needed, the accuracy rate is more than 90%. The calculation of the accuracy, as well as the Random Forest algorithm process in this study, uses R software. The R packages used are *caret*, *e1071*, *randomForest*, *readxl*, *reprtree*, dan *ROSE*.

4. Results

This section describes the results of data processing and the interpretation of the results of data processing.

4.1. SEM-PLS Model Measurement

The model is measured by examining the validity and reliability of the indicators that compose the model's construct. Convergent validity with loading factor and AVE parameters is followed by discriminant validity with cross-loading parameters. In addition to Cronbach alpha and composite reliability metrics, reliability testing is performed too.

4.1.1. Convergent Validity

The process of evaluating the relationship between indicators used in a construct is called a convergent validity test. The requirements for convergent validity are met when the indicators used have a correlation in the same construct. The data used in this study came from a questionnaire filled out by 500 people. The data is used as input for the indicator value in testing the convergent validity. The convergent validity test was conducted based on the SEM model using SmartPLS 3.0.

The outer model is evaluated using three measurement criteria: convergent validity, discriminant validity, and composite validity. Convergent validity can be assessed using reflective indicators, as evidenced by the correlation between the indicators and their concept values. A loading factor value is stated to be valid/reliable if it has a correlation value more than 0.7; however, for research in the early phases of building a measurement scale, a loading value of 0.5 to 0.6 is regarded enough. However, if the final value is less than 0.5, the indicator is ruled incorrect and must be deleted from the model, requiring that data processing (running data) be repeated. Table 4 displays the results of assessing the convergent validity of this study's SEM model.

Table 4. Parameter Value of the First Iteration PLS SEM Loading Factor

	Organizational Commitment (OC)	Employee Performance (EP)	Job Satisfaction (JS)	Turnover Intention (TI)
OC1	0.800	-	-	-
OC2	0.784	-	-	-
OC3	0.880	-	-	-
OC4	0.268	-	-	-
OC5	0.757	-	-	-
EP1	-	0.868	-	-
EP2	-	0.788	-	-
EP3	-	0.868	-	-
EP4	-	0.937	-	-
JS1	-	-	0.946	-
JS2	-	-	0.890	-
JS3	-	-	0.920	-
JS4	-	-	0.767	-
TI1	-	-	-	0.948
TI2	-	-	-	0.932

Based on the results of the SEM-PLS Phase 1 study in Table 4, it is clear that several indications, such as OC04 with a loading factor value of 0.268, are not yet valid. An indication with a low loading factor value, such as OC04, indicates that the indicator's contribution to the construct is minor, and hence the indicator should be eliminated and re-analyzed. After removing OC04 from the model, the SEM-PLS analysis stage 2 was performed, based on the results of the stage 2 analysis, the loading factor value was generated as shown in Table 5.

According to the results of stage 2 SEM-PLS data processing in Table 5, all indicators were valid/had a loading factor value greater than 0.6, implying that all indicators met one of the criteria of convergent validity. Table 5 lists the prerequisites for loading factors greater than 0.6 that match the criteria for exploratory research.

Table 5. Parameter Value of Second Iteration PLS SEM Loading Factor

	Organizational Commitment (OC)	Employee Performance (EP)	Job Satisfaction (JS)	Turnover Intention (TI)
OC1	0.814	-	-	-
OC2	0.791	-	-	-
OC3	0.883	-	-	-
OC5	0.750	-	-	-
EP1	-	0.863	-	-
EP2	-	0.794	-	-
EP3	-	0.870	-	-

EP4	-	0.935	-	-
JS1	-	-	0.946	-
JS2	-	-	0.890	-
JS3	-	-	0.920	-
JS4	-	-	0.766	-
TI1	-	-	-	0.948
TI2	-	-	-	0.932

In addition to analysis the loading factor value, construct validity can be evaluated by testing the Average Variance Extracted (AVE) value. The loading factor value represents the original data score and shows the capacity of the latent variable value. The bigger the AVE value, the greater the ability to explain the value of the hidden variable's indicators. The AVE cut-off value used is 0.50, where an AVE value of at least 0.50 indicates a good measure of convergent validity, which means that the probability of an indicator in a construct entering another variable is lower (<0.50). So that the probability of the indicator converges and enters in constructs whose value in the block is greater than 50% of the value of convergent validity is greater than 50%.

Table 6. Parameter Values of AVE

	Average Variance Extracted (AVE)
Employee Performance (EP)	0.752
Job Satisfaction (JS)	0.780
Organizational Commitment (OC)	0.658
Turnover Intention (TI)	0.884

According to table 6, SEM-PLS data processing in the second stage of the test resulted in the AVE value of each variable being declared good because it exceeded the standards with a value greater than 0.5. This demonstrates that the latent variable can explain more than half of the variance in the indicators. As a result of Tables 5 and 6, it is possible to conclude that all indicators and constructs in the model have passed the convergent validity test. Following that, a discriminant validity test is performed to see whether the indicators of one concept are not highly associated with indications of other constructs.

4.1.2. Validity of Discriminant

The validity of discriminant test examines the relationship between indicators in one construct and indicators in other constructs. If the indications from distinct constructs are not associated, the discriminant validity conditions are met. This is tested using the criterion of the measured construct indicator loading being larger than or equal to the load-ing to other constructs or having a low cross-loading value. Table 7 displays the cross-loading parameter values of the indicators utilized in this investigation using the SmartPLS 3.0 software.

Table 7. Values of Cross Loading Parameter

	Organizational Commitment (OC)	Employee Performance (EP)	Job Satisfaction (JS)	Turnover Intention (TI)
OC1	0.814	0.142	0.336	-0.635
OC2	0.791	0.141	0.250	-0.521
OC3	0.883	0.360	0.522	-0.701
OC5	0.750	0.483	0.278	-0.546
EP1	0.345	0.863	0.321	-0.328
EP2	0.290	0.794	0.010	-0.333
EP3	0.268	0.870	0.140	-0.287
EP4	0.352	0.935	0.212	-0.354
JS1	0.450	0.157	0.946	-0.466
JS2	0.429	0.190	0.890	-0.509
JS3	0.382	0.181	0.920	-0.424
JS4	0.262	0.201	0.766	-0.271
TI1	-0.727	-0.419	-0.466	0.948
TI2	-0.676	-0.282	-0.448	0.932

An indication is also considered valid if its loading factor is greater than its cross-loading value. Table 7 displays the construct correlation of all loading values greater than cross-loading; in this example, the loading value is labeled yellow, indicating test passed with discriminant validity standards.

Another way to assess discriminant validity is to compare the square root value of each construct's AVE with the correlation value between the constructs and other constructs (latent variable correlation). If the AVE root for each construct is bigger than the correlation between the constructs and other constructs, as shown in Table 8, the model has acceptable for discriminant validity.

Table 8. Square Root of AVE

	Organizational Commitment (OC)	Employee Performance (EP)	Job Satisfaction (JS)	Turnover Intention (TI)
Employee Performance (EP)	0.867	-	-	-
Job Satisfaction (JS)	0.203	0.883	-	-
Organizational Commitment (OC)	0.365	0.441	0.811	-
Turnover Intention (TI)	-0.378	-0.487	-0.748	0.940

Table 8 demonstrates that the AVE root values of each construct are all bigger than the correlation between constructs. As a result of Tables 7 and 8, it is possible to conclude that all constructs in the estimated model and met the criteria for the discriminant validity test.

4.1.3. Reliability

A reliability test is performed to determine the consistency and reliability of product indicators when measuring a construct. Cronbach's alpha and composite reliability techniques were used in this work to examine the reliability, with Cronbach's alpha and composite reliability criteria > 0.6 for exploratory research. The SmartPLS 3.0 software is used to test the dependability of the SEM model under consideration; Table 9 displays the model's composite reliability parameter value.

Table 9. Composite Reliability Parameter Values

	Composite Reliability
Employee Performance (EP)	0.923
Job Satisfaction (JS)	0.934
Organizational Commitment (OC)	0.885
Turnover Intention (TI)	0.938

According to Table 9, all construct variables in this investigation showed a composite reliability value greater than the needed value minimum (0.6). Then, this demonstrates the Cronbach Alpha parameter value's reliability requirements in Table 10 were met by the variables in this study.

Table 10. Parameter Value of Cronbach's Alpha

	Cronbach's Alpha
Employee Performance (EP)	0.889
Job Satisfaction (JS)	0.906
Organizational Commitment (OC)	0.826
Turnover Intention (TI)	0.869

Based on the data in table 10, it is found that all construct variables have a Cronbach alpha value greater than 0.7. The minimum requirement of the Cronbach alpha parameter is 0.6 so that the results in Table 10 already meet the criteria of Cronbach alpha. In addition, this shows that the variables in this study meet both reliability standards. Then the measurement results of the three test parameters have also met the criteria. The fulfillment of the model's measurement criteria is based on the characteristics of validity and reliability. The validity test of this study resulted in a good level of accuracy and consistency for each indicator variable. In addition, the indicator variables also have good accuracy in assessing each construct.

4.2. Structural Model Evaluation

The structural model is evaluated by looking at the R-Square value for each dependent variable as the structural model's predictive power. Changes in the value of R-Square, in this case, can suggest some latent variables that have a significant effect on the dependent variable. Then to determine the level of significance use the bootstrapping process.

4.2.1. R-Square

R-square of the latent variable is one of the measurements used to assess the structural model's predictive potential. The R-square value states the percentage of variance of a variable that can be explained by other variables in the same model. R-square values of 0.67, 0.33, and 0.19 can be used to assess the model is strong, moderate, or weak. The R-Square value of the latent variable in the SEM model is shown in Table 11.

Table 11. R-Square Value

	R Square	Details
Employee Performance (EP)	0.158	Weak
Turnover Intention (TI)	0.590	Strong

According to the results of Table 11, there are two R-Square values in this study. Each indicator variable can be represented by an R-Square value. In addition, each indicator variable has unique characteristics. The characteristics of the variables in this study can be described as follows:

1) Employee Performance (EP) has an R-square value of 0.158, indicating the variable is weak. This demonstrates the factors in the model explained 15.8% of the variance in the EP variable, while the remainder is explained by variables outside the model.

2) The Turnover Intentions (TI) variable has an R-square value of 0.59, which is strong. This demonstrates that 59% of the variance in the TI variable can be explained by variables in the model, with the remainder explained by variables outside the model.

4.2.2. Hypothesis Testing

The final test of the structural model evaluation is to look at the level of significance using the bootstrapping process. The significance value of 5% was utilized as the requirement. Before understanding the significance test findings, the output of the PLS method and bootstrapping should be shown. The magnitude of the route coefficient values between latent variables is determined by the PLS algorithm output, but the size of the t-statistical significance value is determined by the bootstrapping output. The outcome of the PLS method and SEM bootstrapping using SmartPLS 3.0 is shown in Figure 2.

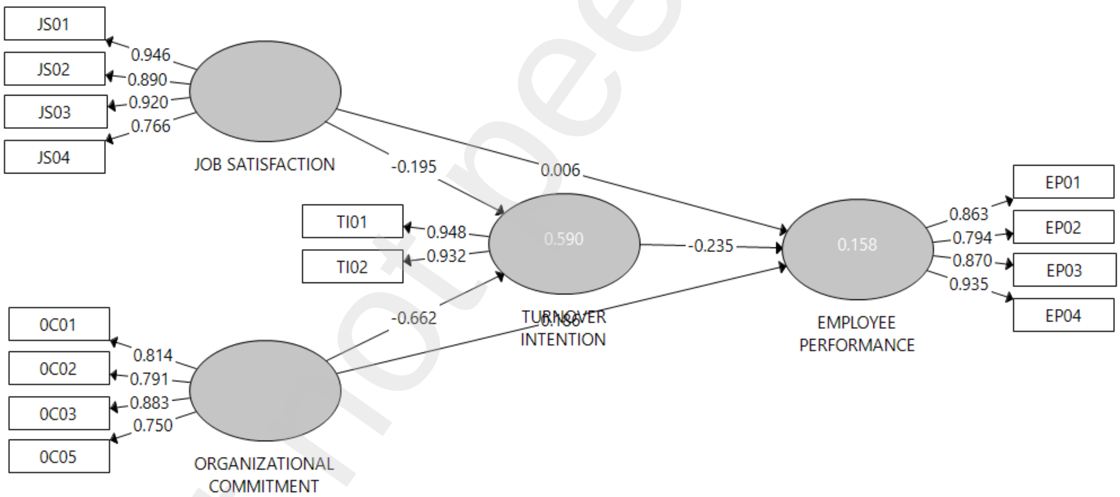

Figure 2. PLS Output

The magnitude of the route coefficient for each latent variable is observed in the results of Figure 2. Furthermore, the bootstrap results are shown in Figure 3 based on the significance value.

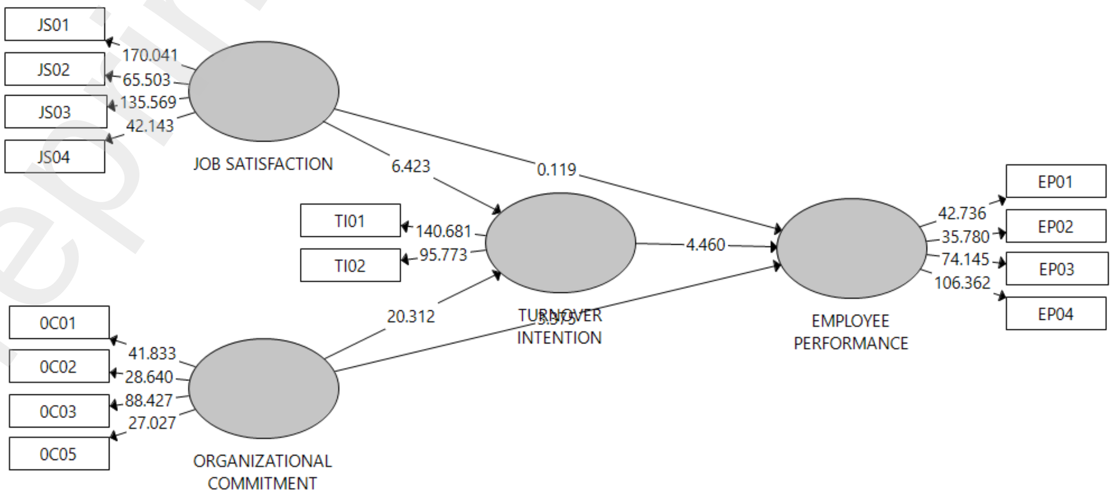

Figure 3. Bootstrap Output

The magnitude of the t-statistic value of the relationship of each latent variable is observed in Figure 3. The magnitudes of the path coefficient values and the t-statistical values can be seen in Figures 2 and 3. Then for a clearer analysis, it is shown in Table 12.

Table 12. Path Coefficient and P-Value

	Original Sample (O)	**T Statistics (|O/STDEV|)**	**P Values**
Job Satisfaction -> Employee Performance	0.006	0.117	0.907
Job Satisfaction -> Turnover Intention	-0.195	6.934	0.000
Organizational Commitment -> Employee Performance	0.186	3.375	0.001
Organizational Commitment -> Turnover Intention	-0.662	22.090	0.000
Turnover Intention ->Employee Performance	-0.235	4.544	0.000

Based on Table 12, it can be explained:

H_1: Job Satisfaction (JS) has a positive effect on Employee Performance (EP).

The results of the analysis in Table 12 shows the parameter coefficient of the influence of job satisfaction on employee performance is 0.006. The coefficient indicates a positive influence between the two variables. While the t-statistic value generated is 0.117. The t-statistic value is smaller than the t-table (0.117 < 1.96). Based on the t-statistic which is smaller than the t-table, it can be concluded that the hypothesis is rejected. In other words, job satisfaction has no significant effect on employee performance.

This can be interpreted as the greater level of job satisfaction can trigger an increase in employee performance. So, if an employee is satisfied with the job, then the results of the employee's performance will be good. But on the other hand, when the employee has a good level of job satisfaction, it can make employees feel that they can do work according to just their convenience. This can lead to stagnant employee performance because there is no demand for greater performance. Therefore, the effect of job satisfaction is not significant because other factors affect employee performance.

H_2: Job Satisfaction (JS) has a negative effect on Turnover Intention (TI)

The results of the analysis in Table 12 shows the parameter coefficient of the influence of job satisfaction on turnover intention is -0.195. The coefficient indicates a negative influence between the two variables. While the t-statistic value generated is 6.934. The t-statistic value is greater than the t-table (6.934 > 1.96). Based on the t-statistic which is greater than the t-table, it can be concluded that the hypothesis is accepted. In other words, job satisfaction has a significant effect on turnover intention.

This can be interpreted as the greater level of job satisfaction can trigger a decrease in turnover intention. So, if an employee is satisfied with his job, then the employee does not want to move or leave the company. Thus, the effect of job satisfaction is significant to suppress or reduce the level of the turnover intention of employees.

H_3: Organizational Commitment (OC) has a positive effect on Employee Performance (EP)

The results of the analysis in Table 12 shows the parameter coefficient of the influence of organizational commitment variable on employee performance is 0.186. The coefficient indicates a positive influence between the two variables. While the t-statistic value generated is 3.375. The t-statistic value is greater than the t-table (3.375 > 1.96). Based on the t-statistic which is greater than the t-table, it can be concluded that the hypothesis is accepted. In other words, organizational commitment has a significant influence on employee performance.

This can be interpreted as the greater level of organizational commitment of an employee can trigger an increase in employee performance. So, if an employee is fully committed to the work and organization (company), then the results of the employee's performance will be great. Thus, the effect of organizational commitment is significant to increase employee performance.

H_4: Organizational Commitment (OC) has a negative effect on Turnover Intention (TI)

The results of the analysis in Table 12 shows the parameter coefficient of the influence of organizational commitment variable on turnover intention is -0.662. The coefficient indicates a negative influence between the two variables. While the t-statistic value generated is 22.090. The t-statistic value is greater than the t-table (22.090 > 1.96). Based on the t-statistic which is greater than the t-table, it can be concluded that the hypothesis is accepted. In other words, organizational commitment has a significant effect on turnover intention.

This can be interpreted as the greater level of organizational commitment of an employee can trigger a decrease in turnover intention. So, if an employee is fully committed to the work and organization (company), then the employee does not want to move or leave the company. Thus, the effect of organizational commitment is significant to suppress or reduce the level of the turnover intention of employees.

H_5: Turnover Intention (TI) has a negative effect on Employee Performance (EP)

The results of the analysis in Table 12 shows the parameter coefficient of the influence of the turnover intention variable on employee performance (original sample) is -0.235, which means that there is a negative influence between the two variables. The t-statistic value generated is 4.544, which means that the result is greater than the t-table (4.544 > 1.96) so it can be concluded that the hypothesis is accepted, or Turnover Intention has a significant effect on Employee Performance.

This can be interpreted as the greater level of the turnover intention of an employee can trigger a decrease in employee performance. So, if an employee wants to move or leave the company, then the employee performance will decrease or stagnate because there is no passion for their work. Thus, the effect of turnover intention is significant to reduce the level of employee performance.

4.3. Testing of Mediation Effect

Testing of the mediation effect is a process of analyzing the connecting variables. Based on the connecting variable, it can be determined the relationship between the independent variable and the dependent variable. This indicates that the independent variable's influence on the dependent variable can occur directly or indirectly through a connecting variable or the media. Table 13 shows the results of the mediation effect test, including the size of the p-value, t-statistics, and the indirect effect value.

Table 13. Path Coefficient and P-Value

	Original Sample (O)	T Statistics (|O/STDEV|)	P Values
Organizational Commitment -> Turnover Intention -> Employee Performance	0.156	4.080	0.000
Job Satisfaction -> Turnover Intention -> Employee Performance	0.046	3.701	0.000

Based on Table 13, it can be explained:

H_6: Organizational Commitment (OC) affects Employee Performance (EP) through Turnover Intention (TI).

The results of the analysis in Table 13 shows the t-statistic value is 4.080, which means that the result is greater than the t-table (4.080 > 1.96) and the indirect effect is 0.156. This means that H_6 is accepted or the mediating model of the influence of organizational commitment on employee performance through turnover intention is accepted.

H_7: Job Satisfaction (JS) affects Employee Performance (EP) through Turnover Intention (TI).

The results of the analysis in Table 13 shows the t-statistic value is 3.701 which means that the result is greater than the t-table (3.701 > 1.96) and the indirect effect is 0.046. This means that H_7 is accepted or the mediation model of job satisfaction's influence on employee performance through turnover intention is accepted.

4.4. Analysis Factor using Machine Learning

The outputs of subsections 4.2 and 4.3 are the results of SEM analysis related to the relationship of the Turnover Intention variable with other independent variables. It was found that Turnover Intention was significantly influenced by Job Satisfaction and Organizational Commitment factors. Furthermore, the classification analysis of the Turnover Intention variable was carried out using the Random Forest algorithm. Turnover Intention classification analysis is based on Job Satisfaction and Organizational Commitment factors. Meanwhile, the classification of Turnover Intention is divided into 2 classes, namely: (1) “Stay”; employees who choose to stay working at the company and (2) “Leave”; employees who choose to leave/move from their current company.

The predictor variables used are indicators of Organizational Commitment and Job Satisfaction factors, namely OC01, OC02, OC03, OC05, JS01, JS02, JS03, and JS04. Based on these predictor variables, a Random Forest model will be obtained which can predict the classification of an employee's Turnover Intention. Then each predictor variable will be sorted based on the most important variable in predicting Turnover Intention.

4.4.1. Preparation of Training Data and Testing data

This study uses the percentage composition to divide between training data and testing data. The composition used is 80% for training data and 20% for data testing. The original data of this study was 500, so a lot of training data and data testing can be seen in Table 14 below.

Table 14. Summary of Training Data and Testing data

	Training	Test	Total
Data	400	100	500
Percentage	80%	20%	100%

Based on Table 14, it is found that the number of training data is 400. While the testing data is 100. Furthermore, the distribution of the training and testing data is carried out randomly using the RStudio software with the function of *sample(2,nrow(data),replace=TRUE,prob=c(0.8,0.2))*. The results of the random division produce training data and data testing containing each class "Stay" and "Leave" which can be seen in Table 15.

Table 15. Summary of Training Data and Testing data

	Classification of Turnover Intention		Total
	Stay	Leave	
Training Data	180	220	400
Testing data	47	53	100

4.4.2. Result of Random Forest Algorithm

Training data were obtained based on Table 15, then entered for the Random Forest algorithm process. The learning process carried out by machine learning based on the Random Forest algorithm will produce a classification model. This Random Forest classification model can be used as a predictor of whether an employee's Turnover Intention level belongs to the "Stay" or "Leave" class. The Random Forest Algorithm was performed with the help of R software using the function of *randomForest(Turnover~.,data=train.datarf)*. The results of the Random Forest algorithm using the R software can be seen in Figure 4.

```
Call:
 randomForest(formula = Turnover ~ ., data = train.datarf)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 3

        OOB estimate of  error rate: 5.74%
Confusion matrix:
    1   2 class.error
1 170  11  0.06077348
2  12 208  0.05454545
```

Figure 4. Random Forest Output

Based on Figure 4, it is found that the output of the Random Forest model type is classification. The number of trees grown is 500. Then the number of variables that are tested for each branch is 3. In addition, the OOB (Out Of Box) error rate is the estimated error that can be obtained if the data used is outside of the test data. The OOB obtained is 5.74%, this value is considered very good because the estimated error is less than 10%.

4.4.3. Plot of Random Forest Model

The Random Forest model obtained can be projected into a graph plot of each branching variable. The function used in R Software is *plot()*. The results of the Random Forest model graph projection can be seen in Figure 5.

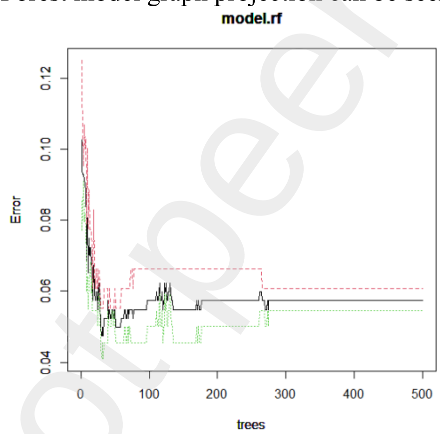

Figure 5. Plot of Random Forest Model

Based on Figure 5, it is found that the form of the Random Forest model has started to stabilize at the 300th tree. This indicates that the machine learning process for studying classification data training using the Random Forest model is stable or good. So that the Random Forest model can be used for general classification of other data.

4.4.4. Tree Plot of Random Forest Model

In addition to plotting the model graph against error, Random Forest can also be analyzed through tree plots. The function used in R Software is reprtree.:::plot.getTree(). The tree plot projection results from the Random Forest model can be seen in Figure 6.

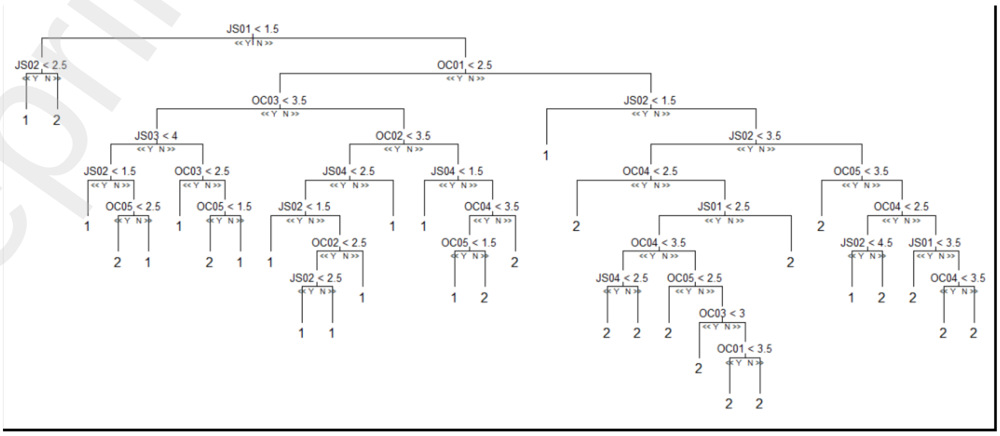

Figure 6. Tree Plot of Random Forest Model

Based on Figure 6, the tree plot form of the Random Forest model is stable. This tree plot shows the decision tree for each branching variable. The tree plot shows classifications 1 and 2, each representing the level of Turnover Intention for “Stay” and “Leave”. This indicates that the machine learning process for studying classification data training using the Random Forest model is stable or good. So that the Random Forest model can be used for the general classification of other data.

4.4.5. Result Evaluation of Random Forest Model

The result of applying the Random Forest model is to predict the classification based on the predictor variables. Prediction of the classification of the Random Forest model is carried out concerning the testing data. The function used in R software is *predict(model, test.data)*. The classification prediction results from the Random Forest model are presented in the form of a confusion matrix which can be seen in Table 16.

Table 16. Confusion Matrix Output

Confusion Matrix		Actual Class	
		Stay	Leave
Prediction Class	Stay	46	1
	Leave	1	52

Based on Table 16, it was found that the number of employees who belonged to the "Stay" class from the testing data was 47. The correct prediction results for the "Stay" class were 46 and the wrong predictions were 1. As for the employees belonging to the "Leave" class from the testing data is 53. The correct prediction results for the “Leave” class are 53 and the wrong prediction is 1. Furthermore, the measurement of the performance evaluation of the Random Forest model can be determined based on the level of accuracy, sensitivity, and specificity as follows.

$$\text{Accuracy } = \frac{46+52}{46+1+1+52} = \frac{98}{100} = 98\%$$

$$\text{Sensitivity } = \frac{46}{46+1} = \frac{47}{46} = 97.87\%$$

$$\text{Specificity } = \frac{52}{52+1} = \frac{52}{53} = 98.11\%$$

The results of the evaluation of the performance of the Random Forest model, each measure has exceeded 90%. The accuracy level is 98%, the sensitivity level is 97.87%, and the specificity level is 98.11%. Based on the classification of the accuracy level in Table 5, this Random Forest model is included in a very good value, which is more than 90%. Therefore, this model can already be used to determine a person's level of Turnover Intention in general on other real field data..

4.4.6. Importance Variabel Output

The Random Forest model produces an analysis of the importance of variables as the best predictor results that support the decision variables. The function used in R software is *View(importance(model))*. The results of the importance variable analysis to rank the best predictors for determining the level of Turnover Intention can be seen in Table 17.

Table 17. Importance Variable Output

Ranking	Predictor Variable	Level of Importance Variable (Mean Decrease Accuracy)
1	OC01	39.518173
2	JS02	33.334789
3	JS04	25.166420
4	OC02	23.092035
5	OC05	18.153701
6	JS01	15.089174
7	JS03	10.595010
8	OC03	6.644303

Based on Table 17, it is found that the best predictor to determine the level of Turnover Intention of an employee is OC01. OC01 has an importance variable level of 39.518173. The predictor variable OC01 means that employees have a high level of "strong belief in a career in the company" or believe that future career projections at the company will be better.

This can be said as a cause and effect, namely if the employee has confidence in the better career projections, then the employee will not have the desire to move or resign from the company.

5. Discussion

The analysis results can be used as a basis for discussion as follows:

Based on the results of the Random Forest model in Figure 6, it is found that the plot tree can be used as a way of classifying the level of Turnover Intention of an employee. The workings of the decision selection based on the plot tree is, following each branch of the variables that represent the decision tree. Each branch of the decision tree has a certain criterion value so the independent variable value affects the decision branch. The plot tree algorithm process is simplified with the help of R software using the *predict()* package. The *predict()* package can provide the final decision on the classification of the dependent variable, in this case, Turnover Intention. The classification of decisions taken is "Stay" or "Leave". The two classifications describe the condition of an employee's expectations regarding his future. "Stay" has an interpretation that employees will stay at the company while showing loyalty, on the contrary, "Leave" means that employees are not comfortable with their current job.

The process of predicting the Turnover Intention level will be very useful for human resource management in a company. HR management can get an overview of the loyalty of its employees, thus helping to evaluate employee performance. In addition, HR management has a reference for making company policies that are following conditions that affect employee turnover decisions. Based on practical theory in HR management, a high level of Turnover Intention will disrupt the stability of the company's performance. This is due to the changing human resources and the teamwork process is disrupted. Therefore, HR management needs to reduce the level of Turnover Intention to maintain company sustainability.

In addition, the results of this study can also be used as a consideration for the process of recruiting new employees. Based on Table 17, the best predictor variable was obtained to determine the level of Turnover Intention. Therefore, the employee recruitment team can have a reference regarding the measurement of the loyalty of prospective employees. Implementation in the recruitment process can be done by giving more points to the OC01 indicator because it is the best predictor. Thus, if the recruiting team knows this measure, it is expected that the employees who are accepted have a good level of loyalty to the company. As a result, the human resources working in the company will have good performance, while at the same time maintaining the company's sustainability.

6. Conclusions

In this study, research was done to analyze the prediction model of employee turnover intention based on the factors that influence it, using a combination of SEM-PLS and machine learning methods. Based on SEM-PLS results were found that Turnover Intention was significantly influenced by Job Satisfaction (JS) and Organizational Commitment (OC) factors. Thus, each indicator from JS and OC is used to make a predictive model for the classification of turnover intention based on the Random Forest algorithm. The results of the Random Forest model have an accuracy of 98%, so it is classified as very good to be applied to other general data. Based on the analysis of the importance of variables, it was found that OC01 was the most influential variable on the level of turnover intention, with a mean decrease accuracy of 39.51. The predictor variable OC01 means that employees have a high level of "strong belief in a career in the company" or believe that future career projections at the company will be better. Therefore, human resource managers must pay attention to career projections and the suitability of employees' positions better, so that employees will give loyalty to the company. Then, it can better maintain the company's sustainability.

Acknowledgments

We would like to thank to PT. TSM Kalimantan and Sekolah Tinggi Ilmu Ekonomi Balikpapan, Indonesia, for supporting this research collaboration.

Informed Consent Statement

All subjects gave their informed consent for inclusion before they participated in the study.

Ethics Committee Statement

This study was approved by the Ethics Committee of Tani Sejahtera Mandiri Ltd. (Date of approval: January 10th, 2021).

References

Ahmad, N., Iqbal, N., Javed, K., & Hamad, N. (2014). Impact of organizational commitment and employee performance on the employee satisfaction. *International Journal of Learning, Teaching and Educational Research, 1*(1), 84-92.

Ahmad, S., Zulkurnain, N. N. A., & Khairushalimi, F. I. (2016). Assessing the validity and reliability of a measurement model in Structural Equation Modeling (SEM). *Journal of Advances in Mathematics and Computer Science*, 1-8.

Alsaraireh, F., Quinn Griffin, M. T., Ziehm, S. R., & Fitzpatrick, J. J. (2014). Job satisfaction and turnover intention among J ordanian nurses in psychiatric units. *International journal of mental health nursing*, *23*(5), 460-467.

Auret, L. and Aldrich, C. (2012). Interpretation of nonlinear relationships between process variables by use of random forests. *Minerals Engineering, 35,* 27-42.

Aydogdu, S., & Asikgil, B. (2011). An empirical study of the relationship among job satisfaction, organizational commitment and turnover intention. *International review of management and marketing*, *1*(3), 43-53.

Aziri, B. (2011). Job satisfaction: a literature review. *Management Research & Practice*, *3*(4).

Badrianto, Y., & Ekhsan, M. (2020). Effect of work environment and job satisfaction on employee performance in PT. Nesinak industries. *Journal of Business, Management, & Accounting*, *2*(1).

Balouch, R., & Hassan, F. (2014). Determinants of job satisfaction and its impact on employee performance and turnover intentions. *International Journal of Learning & Development*, *4*(2), 120-140.

Bonenberger, M., Aikins, M., Akweongo, P., & Wyss, K. (2014). The effects of health worker motivation and job satisfaction on turnover intention in Ghana: a cross-sectional study. *Human resources for health*, *12*(1), 1-12.

Bothma, C. F., & Roodt, G. (2013). The validation of the turnover intention scale. *SA Journal of Human Resource Management*, *11*(1), 1-12.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5–32.

Carmeli, A., & Weisberg, J. (2006). Exploring turnover intentions among three professional groups of employees. *Human Resource Development International*, *9*(2), 191-206.

Cho, Y. J., & Lewis, G. B. (2012). Turnover intention and turnover behavior: Implications for retaining federal employees. *Review of public personnel administration*, *32*(1), 4-23.

Cicek, I. (2011). The Effect of Outsourcing Human Resource on Organizational Performance: The Role of Organizational Culture. *International Journal of Business and Management Studies*, *3*(2), 131-143.

Cohen, A. (2007). Commitment before and after: An evaluation and reconceptualization of organizational commitment. *Human resource management review*, *17*(3), 336-354.

Cohen, A. (2017). Organizational Commitment and Turnover: A Met A-Analysis. *Academy of management journal, 36,* 1140–1157.

Colquitt, J., Lepine, J., & Wesson, M. (2014). *Organizational Behavior: Improving Performance and Commitment in the Workplace (4e)*. New York, NY, USA: McGraw-Hill.

Dessler, G. (2013). *Fundamentals of human resource management*. Pearson.

Ertas, N. (2015). Turnover intentions and work motivations of millennial employees in federal service. *Public personnel management*, *44*(3), 401-423.

Fu, W., & Deshpande, S. P. (2014). The impact of caring climate, job satisfaction, and organizational commitment on job performance of employees in a China's insurance company. *Journal of business ethics*, *124*(2), 339-349.

Gocheva-Ilieva, S., Ivanov, A., and Stoimenova-Minova, M. (2022). Prediction of Daily Mean PM10 Concentrations Using Random Forest, CART Ensemble and Bagging Stacked by MARS. *Sustainability, 14*(2), 798.

Hair Jr, J. F., Matthews, L. M., Matthews, R. L., & Sarstedt, M. (2017). PLS-SEM or CB-SEM: updated guidelines on which method to use. *International Journal of Multivariate Data Analysis*, *1*(2), 107-123.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European business review, 31*(1), 2-24.

Inuwa, M. (2016). Job satisfaction and employee performance: An empirical approach. *The Millennium University Journal*, *1*(1), 90-103.

Islam, J., Mohajan, H., & Datta, R. (2011). A study on job satisfaction and morale of commercial banks in Bangladesh.

Khatimah, H., Susanto, P., & Abdullah, N. L. (2019). Hedonic motivation and social influence on behavioral intention of e-money: The role of payment habit as a mediator. *International Journal of Entrepreneurship*, *23*(1), 1-9.

Kremic, T., Tukel, O. I., & Rom, W. O. (2006). Outsourcing decision support: a survey of benefits, risks, and decision factors. *Supply Chain Management, 11*, 467-482.

Kulkarni V. Y. and Sinha, P. K. (2013). Random Forest Classifiers: A Survey and Future Research Directions. *International Journal of Advanced Computing, 36*(1), 144-153.

Kuo, H. T., Lin, K. C., & Li, I. C. (2014). The mediating effects of job satisfaction on turnover intention for long-term care nurses in T aiwan. *Journal of nursing management*, *22*(2), 225-233.

Lai, M.-C., & Chen, Y.-C. (2012). Self-efficacy, effort, job performance, job satisfaction, and turnover intention: The effect of personalcharacteristics on organization performance. *International Journal of Innovation, Management and Technology*, *3*(4), 387.

Liu, B., Liu, J., & Hu, J. (2010). Person-organization fit, job satisfaction, and turnover intention: An empirical study in the Chinese public sector. *Social Behavior and Personality: an international journal*, *38*(5), 615-625.

Mangkunegara, A. P., & Octorend, T. R. (2015). Effect of work discipline, work motivation and job satisfaction on employee organizational commitment in the company (Case study in PT. Dada Indonesia). *Marketing*, *293*, 31-36.

Mathis, R. L., Jackson, J. H., & Valentine, S. R. (2015). *Human resource management: Essential perspectives*. Cengage Learning.

Meyer, J. P., & Parfyonova, N. M. (2010). Normative commitment in the workplace: A theoretical analysis and re-conceptualization. *Human resource management review*, *20*(4), 283-294.

Mondego, D. Y., & Gide, E. (2020). Exploring The Factors That Have Impact On Consumers'trust In Mobile Payment Systems In Australia. *JISTEM-Journal of Information Systems and Technology Management*, *17,* 1-19.

Nazir, S., Shafi, A.L., Qun, W., Nazir, N., & Tran, Q. (2016). Influence of organizational rewards on organizational commitment and turnover intentions. *Employee Relations, 38*, 596-619.

Podsakoff, N. P., LePine, J. A., & LePine, M. A. (2007). Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: a meta-analysis. *Journal of applied psychology*, *92*(2), 438.

Prajwala, T. R. (2015). A Comparative Study on Decision Tree and Random Forest Using R Tool. *International Journal of Advanced Research in Computer and Communication Engineering*, *4*(1), 196-199.

Pratama, S. (2019). Effect of Organizational Communication and Job Satisfaction on Employee Achievement at Central Bureau of Statistics (BPS) Binjai City. *International Journal of Research and Review, 7*(11), 547-550.

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011). Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling*, *18*(2), 161-182.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological methods*, *15*(3), 209.

Qureshi, M. I., Iftikhar, M., Abbas, S. G., Hassan, U., Khan, K., & Zaman, K. (2013). Relationship between Job Stress, Workload, Environment and Employees Turnover Intentions: What We Know, What Should We Know. *World Applied Sciences Journal*, *23*(6), 764-770.

Reinders, C., Ackermann, H., Yang, M. Y., Rosenhahn, B. (2019). *Chapter 4 - Learning Convolutional Neural Networks for Object Detection with Very Little Training Data*. Multimodal Scene Understanding: Academic Press. ISBN 9780128173589, 65-100,

Richard, O. C., & Johnson, N. B. (2001). Strategic human resource management effectiveness and firm performance. *International Journal of Human Resource Management*, *12*(2), 299-310.

Shahzadi, I., Javed, A., Pirzada, S. S., Nasreen, S., & Khanam, F. (2014). Impact of employee motivation on employee performance. *European Journal of Business and Management*, *6*(23), 159-166.

Siengthai, S., & Pila-Ngarm, P. (2016). *The interaction effect of job redesign and job satisfaction on employee performance*. Evidence-based HRM: a Global Forum for Empirical Scholarship.

Singh, P., Singh, N., Singh, K. K., & Singh, A. (2021). *Chapter 5 - Diagnosing of disease using machine learning*. Machine Learning and the Internet of Medical Things in Healthcare: Academic Press. ISBN 9780128212295, 89-111.

Son, J.-M. (2012). Job embeddedness and turnover intentions: an empirical investigation of construction IT industries. *International Journal of Advanced Science and Technology*, *40*, 101-110.

Svetnik, V., Liaw, A., Tong, C., Culberson, J., Sheridan, R., & Feuston, B. (2003). Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *Journal of chemical information and computer sciences, 43*, 47-58.

Tarigan, V., & Ariani, D. W. (2015). Empirical study relations job satisfaction, organizational commitment, and turnover intention. *Advances in Management and Applied Economics*, *5*(2), 21.

Valentine, S., Godkin, L., Fleischman, G. M., & Kidwell, R. (2011). Corporate ethical values, group creativity, job satisfaction and turnover intention: The impact of work context on work response. *Journal of business ethics*, *98*(3), 353-372.

Wongvibulsin, S., Wu, K.C. & Zeger, S.L. (2020). Clinical risk prediction with random forests for survival, longitudinal, and multivariate (RF-SLAM) data analysis. *BMC Med Res Methodol, 20*(1), 1-14.

Yahaya, R., & Ebrahim, F. (2016). Leadership styles and organizational commitment: literature review. *Journal of Management Development*, *35*(2), 190-216.

Yang, M., Mamun, A. A., Mohiuddin, M., Nawi, N. C., & Zainol, N. R. (2021). Cashless transactions: A study on intention and adoption of e-wallets. *Sustainability*, *13*(2), 831.

Yi, Y., Nataraajan, R., & Gong, T. (2011). Customer participation and citizenship behavioral influences on employee performance, satisfaction, commitment, and turnover intention. *Journal of Business Research*, *64*(1), 87-95.

Yücel, İ. (2012). Examining the relationships among job satisfaction, organizational commitment, and turnover intention: An empirical study. *International Journal of Business and Management, 7*(20), 45-58.