

## COMPARING REGRESSION METHODS FOR ASSESSING AND PREDICTION THE IMPACT OF SALARY INCREASES ON EMPLOYEE PERFORMANCE

Palma Juanta<sup>1</sup>, Zachary Djuli<sup>2</sup>, Tifanny<sup>3</sup>, Delima Sitanggang<sup>4</sup>, Anita<sup>5</sup>

<sup>1,2,3,4,5</sup>Universitas Prima Indonesia

\*Email: <sup>1</sup>palmajuanta@unprimdn.ac.id, <sup>2</sup>zachdjuli@gmail.com, <sup>3</sup>tiffanykie03@gmail.com,

<sup>4</sup>delimasitanggang@unprimdn.ac.id, <sup>5</sup>anita@unprimdn.ac.id

(Received: 16 June 2025, Revised: 28 June 2025, Accepted: 30 July 2025)

### Abstract

In today's competitive digital era, data-driven decision-making is key to enhancing the efficiency of human resource management. One of the main challenges is objectively assessing the impact of salary increases on employee performance, which is often assumed to be a primary motivator but rarely proven quantitatively. This study conducts a comparative analysis of two data mining methods, Linear Regression and Decision Tree Regression, to assessing and predicting the impact of salary increases on employee performance. A case study was conducted at PT. Taipan Agro Mulia using the company's internal historical data. The analysis shows that Linear Regression performed better with an R-Square value of 0.731 or 73.1%, indicating that 73.1% of the variation in employee performance can be explained by salary increases. In comparison, Decision Tree Regression achieved an R-Square value of 0.700 or 70.0%. Additionally, Linear Regression recorded lower prediction errors (MAE = 4.78; MSE = 38.60; RMSE = 6.21) than Decision Tree (MAE = 5.61; MSE = 66.41; RMSE = 8.15). These findings demonstrate that data analysis approaches can serve as a strong foundation for formulating strategic salary policies aimed at improving employee performance.

**Keywords:** *Linear Regression, Decision Tree Regression, Salary Increase, Employee Performance*

*This is an open access article under the [CC BY](#) license.*



\*Corresponding Author: Palma Juanta

### 1. INTRODUCTION

One of the fundamental aspects of human resource management is employee performance, which is systematically managed through what is commonly referred to as performance management. Various concepts and frameworks regarding performance have been widely discussed in the management literature [1]. Performance is an outcome achieved as a result of motivation and job satisfaction [2]. Companies employ various strategies to enhance employee performance, such as providing education, training, salaries, benefits, and adequate workplace facilities—all of which must be given proper attention [3].

Salary is considered one of the most influential factors affecting employee performance. [4]. This is supported by data analysis from previous research, where the salary variable showed a significant distinction in the classification of performance categories. Employees earning Rp7,500,000 tended to fall into the "excellent performance" category, while

those with lower salaries were generally classified as having "good" or "adequate" performance. These findings suggest that salary increases can serve as a strong indicator contributing to improved performance, as salary functions not only as financial compensation but also as a form of appreciation that motivates employees to perform more optimally [5]. Salary also serves as a motivator for employees to develop their talents and skills [6]. Increases in base salary within a company have become a key focal point in the development of employees and the overall progress of the organization [7].

Raising employee salaries is not only a form of appreciation for their performance and loyalty, but also helps improve their well-being and motivates them to remain productive at work. A salary increase refers to the rise in an employee's pay over a certain period. The percentage of salary increases may vary, depending on the company's policies and available budget [8].

PT. Taipan Agro Mulia, a company operating in the agro-industrial sector, serves as the case study in

this research. The company has implemented a salary increase policy as a form of incentive but has not yet adopted a data-driven analysis system to evaluate the effectiveness of this policy on employee performance. Therefore, this study aims to apply data mining methods, namely Linear Regression and Decision Tree Regression, to compare the effectiveness of each in assessing and predicting the impact of salary increases on employee performance based on the available historical data.

Linear Regression is used because it is a powerful method for identifying and quantifying linear relationships between variables, allowing the researcher to measure how changes in salary affect changes in performance levels in a continuous scale [9]. Decision Tree Regression, on the other hand, is useful for capturing nonlinear patterns and providing interpretable decision rules that are easier for management to understand and apply in real-world compensation strategies. The combination of these two methods enables a comprehensive prediction approach that accommodates both linear and complex interactions [10].

Previous research by Prastika Buya Hakim and Zaehol Fatah (2024) titled “Analysis of the Effect of Periodic Salary Increases on Employee Performance Using the Naïve Bayes Algorithm” has explored the relationship between periodic salary increases and employee performance using the Naïve Bayes classification algorithm. The study confirmed a positive correlation, showing that employees who received regular salary increases tended to perform better in terms of attendance and project completion [11].

However, this study presents several limitations that form the research gap for the present study. First, the dataset was not sourced from a specific company such as PT. Taipan Agro Mulia, making it less contextually relevant for organizational-specific analysis. Second, the research relied solely on the Naïve Bayes algorithm, which, while effective for classification, does not provide regression analysis or predictive modeling on continuous performance metrics. Third, the absence of a comparative approach using different data mining techniques leaves room for further exploration regarding which methods are most suitable for performance prediction.

This approach is expected to provide a stronger foundation for management in making decisions related to compensation policies, by taking into account real data and objective analysis results. Through Linear Regression and Decision Tree Regression methods, the company can identify patterns in the relationship between the amount of salary increase and changes in employee performance levels. In this way, PT. Taipan Agro Mulia can design more targeted and evidence-based compensation strategies to enhance employee productivity while supporting the overall

achievement of organizational goals. Furthermore, the implementation of data-driven analysis enables the company to map out employee groups most affected by the salary increase policy and anticipate its long-term impact on employee retention and loyalty. This becomes increasingly important in the modern business era, which demands high efficiency and effectiveness in human resource management.

## 2. RESEARCH METHOD

In this research, the research method that will be carried out can be seen in Figure 1.

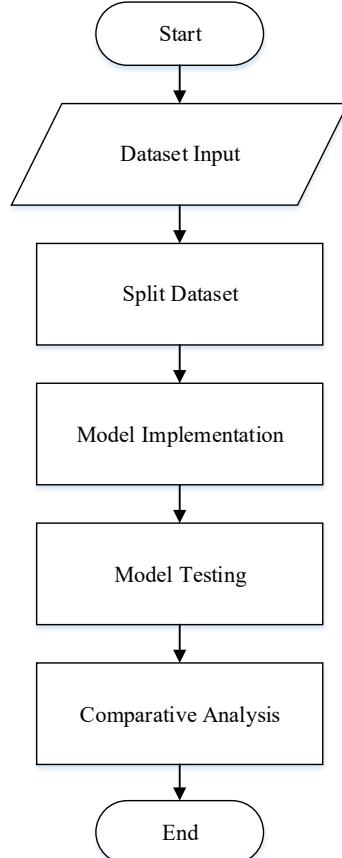


Figure 1. Stages of Research Methods

### 2.1 Dataset Input

At the initial stage of this study, a dataset was entered as the basis for analysis. The dataset was obtained from the employee performance data of PT. Taipan Agro Mulia, a company operating in the agro-industrial sector. It includes information such as employee codes, employee names, salary increase status, and two performance evaluation indicators: KPI1 and KPI2 scores. In total, the dataset consists of 119 employee records, each containing an input variable in the form of salary increase value and two target variables representing performance scores (KPI1 and KPI2). An example of the dataset can be seen in Table 1.

Table 1. Research Dataset

Employee Code	Alias	Salary Increases (x)	Score KPI 1	Score KPI 2
EMKL/ACC/00...	M	0	86	72

EMKL/ACC/00...	MB	0	84	60
EMKL/AST/00...	RT	0	76	80
EMKL/AST/00...	A	0	65	71
EMKL/AST/00...	HT	0	87	79
EMKL/BLW/00	AC	0	79	75
...				
EMKL/BLW/00	SU	0	60	77
...		.... etc		
RB/TAX/00...	I	0	76	86
RB/TAX/00...	A	0	69	86
RB/TAX/00...	K	0	76	84
RB/TAX/00...	RL	0	80	79
RB/TEK/00...	AE	0	77	77

## 2.2 Split Dataset

After the data is entered and the average of the two performance evaluation indicators (KPI1 and KPI2 scores) is calculated into a new column called average KPI, the next step is to prepare the data for the model training and testing process. At this stage, the input variable or feature used is the salary increase (X), while the target or predicted variable is the average KPI (Y). The dataset is divided using the train-test split method, where 80% of the data is allocated as training data to train the model, and the remaining 20% is used as testing data to evaluate the model's performance on previously unseen data. This division is crucial to test the model's ability to generalize and make accurate predictions on new data in an objective manner.

## 2.3 Model Implementation

In this study, two data analysis methods are applied to examine the relationship between salary increases and employee performance, namely Linear Regression and Decision Tree Regression. These methods were selected because they employ different approaches to understanding the relationship patterns between the independent variable (salary increase) and the dependent variable (average performance).

### 1. Linear Regression

Linear Regression is a statistical method used to model the relationship between one or more independent variables (predictors) and a dependent variable (response) in a linear manner [12]. In the case of simple linear regression, this relationship can be represented by a straight-line equation:

$$Y = a + bX + \epsilon \quad (1)$$

Where:

Y is the dependent variable (average KPI score),  
 X is the independent variable (salary increase),  
 a is the intercept (the point where the line crosses the Y-axis),  
 b is the regression coefficient (the slope of the regression line),  
 $\epsilon$  is the error term or the difference between the predicted and actual values.

Linear Regression is used to determine the extent to which changes in variable X (salary increase) can influence variable Y (employee performance).

This method is particularly useful when the relationship between variables is linear and easy to interpret [13].

### 2. Decision Tree Regression

Decision Tree Regression is a tree-based predictive method that splits data into groups based on specific rules to minimize prediction errors. This method works by selecting features (variables) and specific threshold values that optimally divide the data into subsets that are more homogeneous with respect to the target (output) values [14].

Each node in the decision tree represents a condition based on a feature, and each branch indicates the outcome of that condition. This process continues until certain criteria are met, such as the maximum depth of the tree or the minimum number of data points in each node. The leaf nodes of the tree represent the predicted output values [14].

One of the advantages of Decision Tree Regression is its ability to handle non-linear relationships and complex interactions between variables, as well as its capacity to produce models that can be easily interpreted in the form of decision rules (if-then rules) [14].

## 2.4 Model Testing

After building the Linear Regression and Decision Tree Regression models using the training data, the next step is to test the models using the testing data. This testing aims to evaluate the performance of each model in predicting the average employee performance based on the salary increase variable. The evaluation is carried out using several regression metrics, starting with R-Square ( $R^2$ ) to measure the strength of the relationship, followed by error metrics to assess prediction accuracy.

### 1. R-Square ( $R^2$ )-Coefficient of Determination

R-Square is a statistical measure that indicates how much of the variation in the dependent variable (Y) can be explained by the independent variable (X) in the regression model. The  $R^2$  value ranges from 0 to 1 (or 0% to 100%).

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (2)$$

Dimana:

$SS_{res} = \sum(y_i - \hat{y}_i)^2$  is the sum of squared residuals.

$SS_{tot} = \sum(y_i - \bar{y})^2$  is the total variation of the data.

The higher the  $R^2$  value (closer to 1), the better the model explains the variation in the data, meaning that the percentage influence of variable X on Y is greater. In this context,  $R^2$  indicates the extent to which salary increases influence the average employee performance.

### 2. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) measures the average absolute error between the actual values

and the predicted values, without considering the direction of the error. This metric provides an indication of how large the model's prediction errors are, expressed in the same units as the original data [15].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

The MAE value obtained from the calculations can be analyzed to determine whether a prediction has good performance. The performance categories of MAE values are described in Table 2 [16].

Table 2. MAPE Value Performance

Criteria	Interpretation
MAE ≈ 0	Prediction is very close to the actual value (ideal)
MAE < 5% of the average Y	Highly accurate
MAE between 5% – 10% of the Y	Acceptable accuracy
MAE > 10% of the average value of Y	Needs evaluation, the model may be inaccurate

### 3. Mean Squared Error (MSE)

Mean Squared Error (MSE) measures the average of the squared errors between the actual values and the predicted values. Because it squares the differences, MSE is more sensitive to large errors [17].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

MSE squares the error differences, making it highly sensitive to outliers. The smaller the value, the better. Since MSE is expressed in the squared units of Y, its interpretation is less intuitive, however [17]:

1. Low and stable MSE across experiments → the model is fairly accurate and not overfitting.
2. Very high MSE → indicates the presence of large outliers or that the model fails to capture the data pattern.
4. Root Mean Squared Error (RMSE)

Root Mean Squared Error (RMSE) is the square root of MSE, used to return the error unit to the same scale as the target value, making it easier to interpret [18].

$$RMSE = \sqrt{MSE} \quad (5)$$

The smaller the RMSE value, the more accurate the prediction model. The classification is as follows [18]:

Tabel 3. RMSE Value Performance

Criteria	Interpretation
RMSE < 5% of the average Y	High Accuracy
RMSE between 5% – 10% of the average Y	Moderate Accuracy
RMSE > 10% of the average Y	Model evaluation or improvement needed (low accuracy)

### 2.5 Comparative Analysis

At this stage, a comparative analysis was conducted on the two data mining models applied in

the study, namely Linear Regression and Decision Tree Regression. The purpose of this analysis is to compare the effectiveness of each model in assessing and predicting the impact of salary increases on employee performance based on historical data.

## 3. RESULT AND DISCUSSION

The test results obtained in this study were processed using the Python programming language with Google Colab and C# programming language with Visual Studio 2022. This study aims to compare the effectiveness of the Linear Regression and Decision Tree Regression models in assessing and predicting the impact of salary increases on employee performance based on historical data.

### 3.1 Result Dataset Input

The dataset used in this study was input and processed using the Google Colab. The dataset contains 119 employee records, each representing a unique entry related to performance evaluation and salary increases.

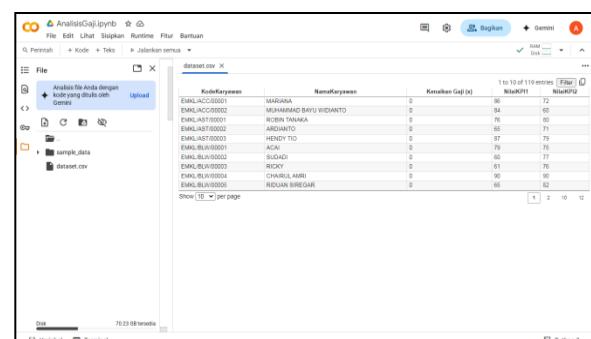


Figure 2. Result Dataset Input

The attributes or columns contained in the dataset are described in Table 3.

Table 3. Dataset Attributes

No	Attribute Name	Description
1	Employee Code	A unique code representing each employee
2	Employee Name	The employee's full name
3	Salary Increases (x)	The percentage of salary increase granted to the employee
4	Score KPI 1	The individual performance evaluation score based on the first indicator
5	Score KPI 2	The individual performance evaluation score based on the second indicator

### 3.2 Result Split Dataset

The dataset was split using the train-test split method with an 80:20 ratio, meaning that 80% of the data was used as training data and the remaining 20% as testing data. Out of the total 119 employee records available, 95 were used to train the model, while the remaining 24 were used to test the model's performance. The split was performed randomly but remained consistent by setting the 'random\_state' parameter in the splitting function. This is important

to ensure that the split can be reproduced if the process is repeated in the future. This step serves as a crucial foundation in the modeling process, as it enables a more realistic evaluation of prediction accuracy and helps prevent overfitting, which often occurs when a model adapts too closely to the training data.

### 3.3 Result Model Implementation

The implementation results are divided into two parts, namely the application of the Linear Regression model and the Decision Tree Regression model. First, the Linear Regression model was implemented through the developed application, as shown in Figure 3.

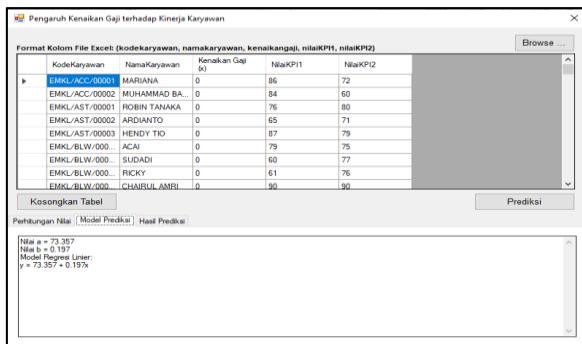


Figure 3. Linear Regression Model Implementation Results

The following section outlines the calculation of the Linear Regression model in this study based on the following formula.

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$= \frac{8751 * 910.8679 - 109.37 * 8202.63}{119 * 910.8679 - (109.37)^2} = 73.357$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$= \frac{119 * 8202.63 - 109.37 * 8751}{119 * 910.8679 - (109.37)^2} = 0.197$$

Thus, the resulting Linear Regression equation model is:

$$y = a + bx = 73.357 + 0.197x$$

The next step involves applying the Decision Tree Regression model, which generates data-splitting rules based on salary increase values automatically determined by the algorithm. These rules divide the data into smaller branches of the decision tree until reaching the terminal nodes (leaf nodes), where each node represents the predicted average KPI value for a group of data with similar characteristics. This process enables the model to capture non-linear relationships between salary increases and employee performance, while also providing deeper insights into how variations in salary increments can affect KPI achievement. Figure 4 illustrates the visualization of the resulting Decision Tree Regression model.

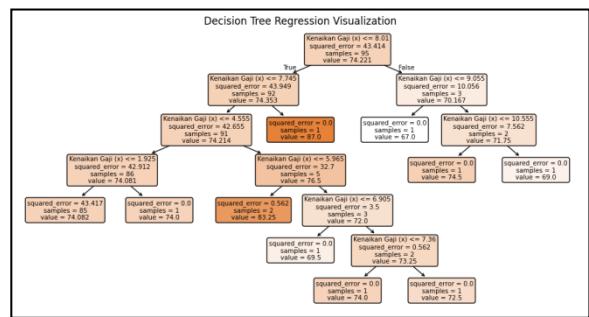


Figure 4. Decision Tree Regression Model Implementation Results

### 3.4 Result Model Testing

After constructing the two regression models, namely Linear Regression and Decision Tree Regression, a testing process was carried out using the test data to assess the predictive capabilities of each model in estimating the average employee performance based on the salary increase variable. The evaluation was conducted using four key performance metrics: R-Square ( $R^2$ ), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).

#### 1. R-Square ( $R^2$ ) Test Results-Coefficient of Determination

R-Square ( $R^2$ ) is used to measure the proportion of variation in the target variable (average performance) that can be explained by the predictor variable (salary increase). The higher the  $R^2$  value, the greater the influence of the input variable on the output.

1. The Linear Regression model produced an  $R^2$  value of 0.731, which means that 73.1% of the variation in employees' average performance can be explained by the salary increase.

2. The Decision Tree Regression model yielded an  $R^2$  value of 0.700, indicating that 70% of the variation can be explained by the model.

#### 2. Mean Absolute Error (MAE) Test Results

MAE measures the average of the absolute differences between the actual and predicted values. This metric is used to determine the average error magnitude of the model in the original units of the data.

1. The MAE of the Linear Regression model is 4.78, indicating that the model's average prediction error is approximately 4.78 points from the actual values.

2. The MAE of the Decision Tree Regression model is 5.61, indicating that the model's average prediction error is approximately 5.61 points from the actual values.

#### 3. Mean Squared Error (MSE) Test Results

MSE measures the average of the squared differences between actual and predicted values. It penalizes larger errors more heavily due to the squaring of the differences.

1. The MSE value of the Linear Regression model is 38.60, indicating that the model has a relatively low squared error.
2. The MSE value of the Decision Tree Regression model is 66.41, indicating that the model has a relatively high squared error.
4. Root Mean Squared Error (RMSE) Test Results  
RMSE is the square root of MSE and returns the error to the original unit of the target. RMSE is easier to interpret because it shares the same scale as the data.
  1. The RMSE of the Linear Regression model is 6.21, which falls into the category of fairly good accuracy.
  2. The RMSE of the Decision Tree model is 8.15, which falls into the low accuracy category because it exceeds 10% of the average target value.

### 3.5 Comparative Analysis Results

After testing both regression models, namely Decision Tree Regression and Linear Regression, the next step is to conduct a comparative analysis based on the evaluation results of each model. This analysis aims to compare how well each model can predict the average employee performance based on salary increases, by referring to four main evaluation metrics: R-Square ( $R^2$ ), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). Table 4 summarizes the comparative analysis results of both models.

Table 4. Comparative Analysis Results

Matrix	R-Square ( $R^2$ )	MAE	MSE	RMSE
Linear Regression	0,731	4,78	38,60	6,21
Decision Tree Regression	0,700	5,61	66,41	8,15

Next, the results of the model comparison analysis are presented in a bar chart visualization as shown in Figure 5.

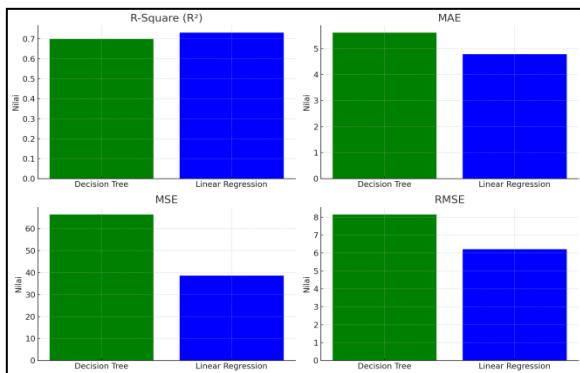


Figure 5. Bar Chart of Comparative Analysis Results

Based on the evaluation results presented in Table 4 and Figure 5, a comparison of the performance of the two regression models—Decision Tree Regression and Linear Regression in predicting

average employee performance based on salary increases is obtained.

In terms of R-Square ( $R^2$ ), the Linear Regression model shows a value of 0.731, slightly higher than the Decision Tree Regression model, which has a value of 0.700. A higher  $R^2$  value indicates that the Linear Regression model is able to explain 73.1% of the variability in the target data (average employee performance), while the Decision Tree model explains 70.0%. This shows that overall, the linear model performs slightly better in capturing the relationship between salary increases and employee performance.

For the Mean Absolute Error (MAE), the Linear Regression model also shows better performance with a value of 4.78, compared to the Decision Tree Regression which records an MAE of 5.61. This means that the average absolute difference between predicted and actual values in the linear model is smaller, indicating a lower and more consistent prediction error.

In the Mean Squared Error (MSE) metric, the results again demonstrate the superiority of the Linear Regression model, with a value of 38.60, which is much lower compared to the MSE value of 66.41 in the Decision Tree model. Since MSE places greater penalties on large prediction errors, the lower value in the linear model indicates not only more accurate predictions but also greater stability against potential outliers.

Furthermore, the Root Mean Squared Error (RMSE), which is the square root of MSE, also shows a similar performance. The RMSE for the linear model is 6.21, lower than the 8.15 recorded by the Decision Tree Regression. Because RMSE returns the error to the same scale as the target, it is easier to interpret: on average, the linear model has an error of around 6.21 points, while the Decision Tree Regression model has an error of around 8.15 points from the actual values.

The superior performance of the Linear Regression model can be attributed to the nature of the relationship between salary increases and employee performance, which appears to follow a relatively linear trend. Linear Regression is particularly effective when the relationship between independent and dependent variables is approximately linear and when the dataset is not too complex or noisy. In contrast, Decision Tree Regression tends to overfit small datasets and may struggle to capture smooth linear trends, especially if the tree is not sufficiently pruned. Additionally, Decision Tree models can be sensitive to slight variations in the data, leading to less consistent results. Therefore, in this context, Linear Regression provides more generalizable and stable predictions [19][20].

In conclusion, based on these four evaluation metrics, the Linear Regression model delivers more accurate and stable predictions than the Decision

Tree Regression model. Although the difference in  $R^2$  values is not very large, the lower prediction errors in the linear model make it a more reliable choice for predicting employee performance based on salary increases.

#### 4. CONCLUSION

Based on the results of this comparative study, it can be concluded that both Linear Regression and Decision Tree Regression models are capable of assessing and predicting the impact of salary increases on employee performance, but with varying levels of accuracy. The Linear Regression model demonstrated slightly better performance, as indicated by a higher R-Square value of 0.731, suggesting that 73.1% of the variance in employee performance could be explained by salary increases. In addition, it achieved lower prediction errors, with MAE of 4.78, MSE of 38.60, and RMSE of 6.21. In contrast, the Decision Tree model yielded an R-Square of 0.700 with higher prediction errors (MAE = 5.61, MSE = 66.41, RMSE = 8.15). These results indicate that while both models can be applied for predictive purposes, the Linear Regression model offers better generalization and more accurate performance prediction in this context. Therefore, it is recommended to use the Linear Regression approach when the relationship between salary increases and employee performance tends to be linear and consistent across data.

#### 5. REFERENCES

- [1] M. Pusparani, "Faktor Yang Mempengaruhi Kinerja Pegawai (Suatu Kajian Studi Literatur Manajemen Sumber Daya Manusia)," *J. Ilmu Manaj. Terap.*, vol. 2, no. 4, pp. 534–543, 2021, doi: 10.31933/jimt.v2i4.466.
- [2] A. Ariyanto and M. Ryansyah, "Decision Support System for Selecting the Best Employee At Pt Bank Digital Bca Using Saw Method," *JKO (Jurnal Inform. dan Komputer)*, vol. 6, no. 1, pp. 72–78, 2023, doi: 10.33387/jko.v6i1.5913.
- [3] R. Djutalov, P. R. S, and N. Sahirah, "Sistem Penilaian Kinerja Teknisi Berbasis Web Dengan Metode TOPSIS Pada IT Cyber Community," *BIIKMA Bul. Ilm. Ilmu Komput. dan Multimed.*, vol. 2, no. 4, pp. 759–765, 2024.
- [4] K. P. Ulandari, N. Chamidah, and A. Kurniawan, "Prediksi Risiko Gagal Bayar Kredit Kepemilikan Rumah Dengan Pendekatan Metode Random Forest," *SAINSMAT J. Ilm. Ilmu Pengetah. Alam*, vol. 13, no. 2, pp. 162–170, 2024.
- [5] Muhamni and W. Awaliah, "Penentuan Evaluasi Kinerja Karyawan Berbasis Logika Fuzzy Mamdani," *J. Secur. Comput. Information, Embed. Network, Intell. Syst.*, vol. 2, no. 2, pp. 109–117, 2024.
- [6] A. S. Riyadi, Suroso2, and M. M. Karnama, "Pengaruh Gaji dan Insentif Terhadap Kinerja Karyawan PT. Tenma Indonesia Karawang Factory," *J. Manag. Bussines*, vol. 6, no. 1, pp. 20–29, 2024.
- [7] A. Septiani, K. B. Riyanto, and E. Aprinato, "Pengaruh Kompensasi Finansial Dan Kompensasi Nonfinansial Terhadap Kinerja Pegawai Samsat Pesisir Barat Lampung," *J. Manaj. Divers.*, vol. 4, no. 3, pp. 729–739, 2024, doi: 10.25130/sc.24.1.6.
- [8] I. G. A. Wirajaya *et al.*, *Pengantar Perpajakan*. Bandung: CV. Media Sains Indonesia, 2023.
- [9] K. A. F. A. Samah, N. S. D. Wirakarnain, R. Hamzah, N. A. Moketar, L. S. Riza, and Z. Othman, "A linear regression approach to predicting salaries with visualizations of job vacancies: a case study of Jobstreet Malaysia," *IAES Int. J. Artif. Intell.*, vol. 11, no. 3, pp. 1130–1142, 2022, doi: 10.11591/ijai.v11.i3.pp1130-1142.
- [10] P. L. Jatika and K. R. Putra, "Comparison of Prediction Models: Decision Tree , Random Forest , and Support Vector Regression," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 39–49, 2025.
- [11] P. B. Hakim and Z. Fatah, "Analisis Pengaruh KGB Terhadap Kinerja Karyawan Menggunakan Algoritma Naïve Bayes," *Gudang J. Multidisiplin Ilmu*, vol. 2, no. 10, pp. 123–128, 2024.
- [12] M. A. Fikri and D. P. Prayogo, "Penerapan Metode Regresi Linear Untuk Prediksi Harga Layerzero Saat Listing," *J. Inform.*, vol. 1, no. 2, pp. 7–11, 2024.
- [13] W. S. Rudi, Y. A. Pranoto, and F. X. Ariwibisono, "Penerapan Metode Regresi Linier Dalam Peramalan Penjualan Kue Di Toko Karya Bahari Samarinda Berbasis Website," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 4, pp. 2451–2457, 2023, doi: 10.36040/jati.v7i4.7547.
- [14] A. T. Nurani, A. Setiawan, and B. Susanto, "Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," *J. Sains dan Edukasi Sains*, vol. 6, no. 1, pp. 34–43, 2023, doi: 10.24246/juses.v6i1p34-43.
- [15] F. R. Hariri and L. W. Rochim, "Sistem Rekomendasi Produk Aplikasi Marketplace Berdasarkan Karakteristik Pembeli Menggunakan Metode User Based Collaborative Filtering," *Teknika*, vol. 11, no. 3, pp. 208–217, 2022, doi: 10.34148/teknika.v11i3.538.
- [16] T. A. Prasetyo *et al.*, "Sales forecasting of marketing using adaptive response rate single exponential smoothing algorithm," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 31, no. 1, pp. 423–432, 2023, doi: 10.11591/ijeecs.v31.i1.pp423-432.
- [17] Andi, "Analisis Komparasi Algoritma Fuzzy Dalam Sistem Pendukung Keputusan Pemilihan

Sepeda Motor Bekas,” *J. TIMES*, vol. 12, no. 2, pp. 71–78, 2023, doi: 10.51351/jtm.12.2.2023711.

[18] L. Wiranda and M. Sadikin, “Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma,” *J. Nas. Pendidik. Tek. Inform.*, vol. 8, no. 3, pp. 184–196, 2019.

[19] X. Li, “Comparing Linear Regression and Decision Trees for Housing Price Prediction,” in *International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI)*, 2023, pp. 77–84, doi: 10.2991/978-94-6463-370-2\_9.

[20] I. Azure, “Predictive modeling for industrial productivity: Evaluating linear regression and decision tree regressor approaches,” *J. AppliedMath*, vol. 2, no. 4, p. 1435, 2024, doi: 10.59400/jam.v2i4.1435.