Journal of AppliedMath 2024, 2(4), 1435.
https://doi.org/10.59400/jam.v2i4.1435

Article

Predictive modeling for industrial productivity: Evaluating linear regression and decision tree regressor approaches

Isaac Azure

Department of Mathematics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana; azureike@yahoo.com

CITATION

Azure I. Predictive modeling for industrial productivity: Evaluating linear regression and decision tree regressor approaches. Journal of AppliedMath. 2024; 2(4): 1435. https://doi.org/10.59400/jam.v2i4.1435

ARTICLE INFO

Received: 3 June 2024
Accepted: 30 July 2024
Available online: 5 September 2024

Abstract: This research discusses the importance of predictive modeling in optimizing efficiency in various sectors, particularly in industrial settings. It compares the effectiveness of linear regression and decision tree regression models in predicting productivity. The study aims to provide insights into the strengths and limitations of each technique, assisting decision-makers in selecting the best model for their needs. It begins by explaining the theoretical foundations of both models and conducts a literature review to highlight their practical implementations. The methodology involves data collection, preprocessing, model training, evaluation, and comparison using real-world datasets. Performance metrics such as Mean Squared Error (MSE) are used for evaluation. The comparative analysis reveals that the linear regression model consistently outperforms the decision tree regressor model in terms of lower MSE values across all datasets. Overall, the study offers empirical evidence and practical insights into the predictive capabilities of both models, with potential implications for strategic decision-making in various industries.

Keywords: productivity prediction; linear regression; decision tree regressor; industrial contexts; comparative analysis

1. Introduction

In today's dynamic industrial world, boosting productivity is critical to creating a competitive advantage and long-term success. Industries from all industries are constantly looking for efficient methods to anticipate and improve productivity levels. Among the several predictive modeling techniques, linear regression and decision tree regression stand out for their simplicity, interpretability, and effectiveness in modeling complicated relationships within datasets.

This study aims to examine the effectiveness of linear regression and decision tree regression models in forecasting productivity in an industrial setting. By using real-world data and these two distinct modeling approaches, this study hopes to provide valuable insights into the strengths and limitations of each technique, assisting industry practitioners and decision-makers in selecting the best predictive model for their specific needs.

The study will begin by discussing the theoretical foundations and operational mechanisms of linear regression and decision tree regression. A thorough analysis of current literature will be carried out to highlight the historical evolution, theoretical underpinnings, and practical implementations of these approaches in industrial productivity prediction.

Subsequently, the research will delineate the methodology employed, encompassing data collection, preprocessing, model training, evaluation, and comparison processes. Real-world datasets pertinent to industrial productivity will be

utilized to ensure the applicability and relevance of the findings to industrial settings. The chosen datasets will be carefully curated and prepared to adhere to the requirements of both linear regression and decision tree regression models.

This study is built on an investigation of the theoretical underpinnings and operational mechanisms of linear regression and decision tree regression. Drawing on important publications by several authors [1,2], this paper will outline the historical evolution and practical uses of these approaches in industrial productivity prediction.

Furthermore, the research technique will include data collecting, preprocessing, model training, evaluation, and comparison procedures. Real-world datasets relevant to industrial productivity, such as those investigated by several authors [3,4], will be carefully curated and prepared to fulfill the needs of both linear regression and decision tree regression models.

The comparative analysis will be underpinned by performance metrics including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2) coefficient. This comprehensive assessment, informed by the works of several authors [5–7], will illuminate the predictive accuracy, model interpretability, and computational efficiency across both modeling paradigms.

The comparative analysis will be conducted based on various performance metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2) coefficient. These metrics will facilitate a comprehensive assessment of predictive accuracy, model interpretability, and computational efficiency across both modeling paradigms.

Finally, the study intends to add to the current body of knowledge by providing empirical evidence and practical insights regarding the effectiveness of linear regression and decision tree regression models in forecasting productivity in industrial settings. The study's findings have the ability to guide strategic decision-making processes, optimize resource allocation, and drive performance improvement initiatives across a wide range of industrial sectors.

2. Production prediction models

As stated in the introduction, the primary goal of this study is to conduct a thorough comparison of two distinct production prediction models across a variety of datasets, with the overarching goal of determining which model produces superior predictive capabilities in the context of productivity. This section explains the mathematical foundations of the linear regression and decision tree models, which will then be applied and thoroughly tested using authentic datasets.

2.1. The linear regression model

The linear regression model, a mainstay of predictive analytics, is the foundation of many predictive modeling efforts. At its core, the linear regression model seeks to establish a linear relationship between a dependent variable (or target variable) and one or more independent variables (or predictors) [8]. The mathematical formulation of the linear regression model is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \quad (1)$$

where:

Y is the dependent variable (productivity).

$X_1, X_2, \ldots, X_n$ are the independent variables (raw material cost, labor efficiency, machine utilization).

β_0 is the y–intercept.

$\beta_1, \beta_2, \ldots, \beta_n$ are the coefficients representing the impact of each independent variable on the dependent variable.

ε is the error term, representing unobserved factors affecting the dependent variable.

The objective is to minimize the sum of squared differences between the predicted $(\hat{Y})$ and actual (Y) values:

$$J(\beta) = \frac{1}{2m}\sum_{i=1}^{m}\left(\hat{Y}_i - Y_i\right)^2 \tag{2}$$

where:

m is the number of data points.

Let's represent the method using matrix notation:

$$Y = X\beta + \varepsilon \tag{3}$$

where:

Y is an $m \times 1$ vector of actual values.

X is an $m \times (n+1)$ matrix of features (including a column of ones for the intercept).

β is a $(n+1) \times 1$ vector of errors.

To minimize $J(\beta)$, we set the gradient of J with respect to β equal to zero >.

To do this, we define the objective function.

$$J(\beta) = \frac{1}{2m}(X\beta - Y)^T(X\beta - Y) \tag{4}$$

Completing the gradient $J(\beta)$ with respect to β:

$$\nabla J(\beta) = \frac{1}{m}(X\beta - Y)^T \tag{5}$$

Finding the minimum of $J(\beta)$, we set the gradient equal to zero:

$$\frac{1}{m}X^T(X\beta - Y) = 0 \tag{6}$$

Multiply both sides by m to simplify:

$$X^T(X\beta - Y) = 0 \tag{7}$$

Expanding and simplifying, we get

$$X^TX\beta = X^TY \tag{8}$$

Solving for β, multiply both sides by $(X^TX)^{-1}$

$$\beta = (X^TX)^{-1}X^TY \tag{9}$$

Hence, this is the least square solution for the coefficient β.

Once β is obtained, predictions $(\hat{Y})$ can be made using:

$$\hat{Y} = X\beta \tag{10}$$

If feature scaling is applied, the scaled features (X_{scaled}) are obtained using the equation:

$$X_{scaled} = \frac{X - mean(X)}{std(X)} \tag{11}$$

2.2. Decision tree model

In contrast to the linear regression model's parametric approach, the decision tree model employs a non-parametric methodology distinguished by its hierarchical structure of decision nodes and leaf nodes. At each decision node, the decision tree model separates the feature space based on the most discriminative features, with the goal of iteratively improving the model's predictive power. Mathematically, the decision tree model can be described as a recursive partitioning process, with each division aiming to maximise the homogeneity of the objective variable within the subsequent subsets [9,10].

The decision tree model's essential idea is the capacity to recursively partition the feature space into disjoint sections, each regulated by a set of conditional rules. The decision tree's leaf nodes make predictions based on the target variable's majority class (for classification tasks) or mean value (for regression tasks) within the appropriate subset [11].

By defining the mathematical formulations of both the linear regression and decision tree models, this study aims to lay a solid foundation for the forthcoming empirical evaluations utilizing actual data. This study aims to provide valuable insights into the relative strengths and limitations of each modeling paradigm in predicting productivity within industrial contexts by meticulously examining their predictive capabilities across diverse datasets [12,13].

The mathematical formulation of decision trees includes ideas like as entropy and information gain, but it does not use calculus or linear algebra for optimization.

For classification issues, entropy measures a set's impurity. The entropy of a set *S* that contains *p* positive and *n* negative samples is defined as:

$$H(S) = -p.log_2(p) - n.log_2(n) \tag{12}$$

The entropy of a set S is defined using the concept of information content:

$$H(S) = -\sum_{i=1}^{c} P_i.log_2(P_i) \tag{13}$$

where:

c is the number of classes (in binary classification, $c = 2$),

P_i is the probability of class i in set S.

For binary classification ($c = 2$), let p be the probability of the positive class (class 1) and n be the probability of the negative class (class 0). The probabilities are calculated as:

$$P = \frac{Number\ of\ positive\ examples}{Total\ number\ of\ examples} \tag{14}$$

$$n = \frac{Number\ of\ negative\ examples}{Total\ number\ of\ examples} \tag{15}$$

Substitute these into the entropy equation:

$$H(S) = -p.log_2(p) - n.log_2(n)$$

Proof. The entropy of set S is defined as:

$$H(S) = -\sum_{i=1}^{c} P_i.log_2(P_i)$$

For binary classification ($c = 2$), where p is the probability of the position class

and n is the probability of the negative class:

$$H(S) = -p.log_2(p) - n.log_2(n)$$

The entropy of each subset S_i is calculated similarly based on the probabilities $P_i \; and \; n_i$:

$$H(S_i) = -log_2(P_i) - n_i.log_2(n_i)$$

Weighted sum of entropies of the subsets is:

$$\sum_{i=1}^{k} \frac{|S_i|}{|S|} H(S_i)$$

The information gain is defined as the reduction in entropy achieved by splitting the set S:

$$IG(S, A) = H(S) - \sum_{i=1}^{k} \frac{|S_i|}{|S|} H(S_i)$$

Substitute the expressions for $H(S)$ and $H(S_i)$:

$$IG(S, A) = -p.log_2(p) - n.log_2(n) - \sum_{i=1}^{k} \frac{|S_i|}{|S|} \Big(\big(-P_i.log_2(P_i) - n_i.log_2(n_i) \big) \Big)$$

Combine terms to simplify the expression:

$$IG(S, A) = -(p.log_2(p) + n.log_2(n) + \sum_{i=1}^{k} \frac{|S_i|}{|S|} \big(P_i.log_2(P_i) + n_i.log_2(n_i) \big) \square$$

2.3. Experimental study

To provide a comprehensive comparison of both models, this research incorporates four distinct datasets spanning various domains. These datasets encompass Sales Prediction in Retail, House Price Prediction, Employment Performance Evaluation, and Temperature Prediction. They are respectively labeled as Dataset 1, Dataset 2, Dataset 3, and Dataset 4, as outlined in **Table 1** below. The features of each dataset along with their corresponding data are elaborated in **Table 1**.

Table 1. Datasets for experimental test.

Dataset	**Features**		
Dataset 1	Customer Visits	Advertising Budget	Store size
	100	500	50
	1000	5000	500
	200	200	200
Dataset 2	Number of Bedrooms	Distance to City	House Price
	800	2	1
	4000	6	20
	300	300	300
Dataset 3	Hours Worked	Communication Skills	Performance Score
	20	0	1
	60	1	5
	150	150	150

Table 1. (*Continued*).

Dataset	Features		
	Latitude	Altitude	Humidity
Dataset 4	−90 90 250	0 5000 250	0 100 250

Based on the parameters delineated in the preceding table, both models were implemented using Python scripts. Subsequently, these scripts were executed to generate the respective predicted outcomes, accompanied by estimates of mean square error values.

3. Related theorems and lemmas of the models

Linear regression and decision tree regressor models are important techniques for predictive modeling in a variety of applications, including estimating productivity in industrial settings. Theorems and lemmas related with these models provide theoretical insights, properties, and considerations necessary for understanding their behavior and performance.

3.1. Theorems and lemmas of linear regression model

Theorems and lemmas are essential for understanding the mathematical underpinnings and properties of linear regression models. Some essential concepts are:

Theorem 1. *Least squares estimation [14].*

Given a set of mm observations (x_i, y_i), where x_i represents the input variables and y_i represents the corresponding output variable (productivity), the least squares estimate $\hat{\beta}$ for the parameters β in the linear regression model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon$ is obtained by minimizing the sum of squared residuals:

$$\hat{\beta} = arg_{\beta} min \sum_{i=1}^{m} \left(y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in})\right)^2 \quad (16)$$

Theorem 2. *Gauss-Markov theorem [15].*

Under the assumptions of the classical linear regression model, the least squares estimators $\hat{\beta}$ are unbiased and have the minimum variance among all linear unbiased estimators.

Lemma 1. *Properties of residuals [16,17].*

For the linear regression model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon$, the residuals $e_i = y_i - \hat{y}_i$ have the following properties:

1) The residuals have a mean of zero: $\frac{1}{m}\sum_{i=1}^{m} e_i = 0$.
2) The residuals are uncorrelated with the predictors: $Cov(e_i, x_{ij}) = 0$ for all $j = 1,2, \ldots, n$.
3) The residuals have constant variance (homoscedasticity): $Var(e_i) = \sigma^2$.
4) The residuals are normally distributed:$e_i \sim N(0, \sigma^2)$ for all $i = 1,2, \ldots, m$.

Lemma 2. *Gauss-Markov Assumptions [15].*

For the Gauss-Markov theorem to hold, the following assumptions must be

satisfied:

1) Linearity: The relationship between the predictors and the response variable is linear.
2) Independence: The errors (residuals) are independent of each other.
3) Homoscedasticity: The errors have constant variance.
4) Normality: The errors are normally distributed with mean zero and constant variance.
5) No Perfect Collinearity: There is no perfect multicollinearity among the predictors.

3.2. Theorems and lemmas of decision tree regressor model

Theorem 3. *Regression tree construction.*

Given a set of mm observations(x_i, y_i), where x_i represents the input variables and y_i represents the corresponding output variable (productivity), a decision tree regressor partitions the feature space into disjoint regions $R_1, R_2, \ldots, R_j$, such that the response variable y is predicted as the mean of the training observations in each region:

$$\hat{y}_j = \frac{1}{N_j} \sum_{i \epsilon R_j} y_i \tag{17}$$

where $\hat{y}_j$ is the predicted value for region R_j and N_j is the number of observations in region R_j.

Theorem 4. *Recursive partitioning [18].*

The decision tree regressor algorithm recursively partitions the feature space based on the most discriminative attributes, aiming to maximize the homogeneity of the target variable within the resultant subsets.

Lemma 3. *Properties of decision trees [19].*

For decision tree regressors, the following properties hold:

1) Hierarchical partitioning: The feature space is partitioned into disjoint regions hierarchically, with each partition corresponding to a decision node in the tree.
2) Piecewise constant prediction: Within each partition (leaf node) of the decision tree, the prediction for the response variable yy is a constant value.
3) Interpretability: Decision trees are interpretable models, as the splits in the tree correspond to intuitive decision rules based on feature values.

Lemma 4. *Overfitting control [14].*

Decision tree regressors are prone to overfitting, especially when the trees are deep and complex. Regularization techniques such as pruning, limiting the maximum depth of the tree, or setting a minimum number of samples required to split a node can help control overfitting and improve the generalization performance of the model.

Lemma 5. *Feature importance [20].*

Decision tree regressors provide a measure of feature importance, indicating the relative importance of each feature in predicting the target variable. Features that are frequently used for splitting nodes higher up in the tree are considered more important in predicting productivity.

4. Results discussion

Drawing upon the mathematical formulations of linear regression and decision tree regression, the parameters specified in **Table 2** were meticulously inputted into the respective models. Subsequently, comprehensive analyses were conducted, yielding a myriad of results elucidating the predictive capabilities of each model across diverse datasets. These results encompassed a range of performance metrics, including but not limited to mean squared error, R-squared coefficient, and root mean squared error, which collectively provided valuable insights into the efficacy and robustness of each modeling approach.

Table 2. Predicted parameters for linear regression and decision tree regressor models.

Dataset	Parameters	Linear Regression Model	Decision Regressor Model
1	Predicted Scales	85106.53583717004	88214.08324241744
	Mean Squared Error	1419926.7292265582	17255642.55627085
2	Predicted House Price	298646.1429131364	380778.5276175679
	Mean Squared Error	2446593752.5670934	5509637467.099092
3	Predicted Performance Score	313.15168967988734	329.078546122949
	Mean Squared Error	98.4842183439387	444.6210049620012
4	Predicted Temperature	36.36307221002279	31.54859473247914
	Mean Squared Error	4.45761528055088	25.57305707384498

The table presents the predicted parameters and corresponding mean squared error values for both the linear regression and decision tree regressor models across four different datasets.

Dataset 1 was meant to predict sales. The linear regression model predicted sales to be approximately 85,106.54, while the decision tree regressor model predicted slightly higher at around 88,214.08. Also, the linear regression model exhibited a lower MSE of approximately 1,419,926.73 compared to the decision tree regressor model, which had a substantially higher MSE of approximately 17,255,642.56. The results from Dataset 1 suggest that the linear regression model performed better in predicting sales for Dataset 1, as evidenced by its lower MSE. The decision tree regressor model, on the other hand, appeared to have a higher error rate, indicating less accurate predictions.

Unlike Dataset 1, Dataset 2 was meant to be used to predict House Prices. The linear regression model predicted the house price to be around 298,646.14, whereas the decision tree regressor model predicted a higher value of approximately 380,778.53. However, the linear regression model exhibits a substantial MSE of approximately 2,446,592,752.57, while the decision tree regressor model shows an even higher MSE of approximately 5,509,637,467.10. This clearly show that in Dataset 2, both models predict higher house prices, with the decision tree regressor model showing higher predictions and a significantly higher MSE. This suggests that

the linear regression model may provide more reliable predictions compared to the decision tree regressor model for this dataset.

Predicting Performance Score is what Dataset 3 aimed to get. The linear regression model predicted a performance score of approximately 313.15, while the decision tree regressor model predicted a slightly higher score of around 329.08. Meanwhile, the linear regression model demonstrated a relatively low MSE of approximately 98.48, whereas the decision tree regressor model exhibits a higher MSE of approximately 444.62. Hence for Dataset 3, both models predicted similar performance scores, with the decision tree regressor model showing slightly higher predictions. However, the linear regression model outperformed the decision tree regressor model in terms of MSE, indicating better predictive accuracy.

Dataset 4 predicted Temperature. The linear regression model predicted the temperature to be around 36.36 degrees Celsius, while the decision tree regressor model predicts a lower temperature of approximately 31.55 degrees Celsius. The linear regression model demonstrated a relatively low MSE of approximately 4.46, whereas the decision tree regressor model exhibited a higher MSE of approximately 25.57. In Dataset 4, the linear regression model predicted a higher temperature compared to the decision tree regressor model. Additionally, the linear regression model achieved a lower MSE, indicating better predictive performance for temperature prediction.

Based on the analysis of the four datasets, it is apparent that both the linear regression and decision tree regressor models generated substantial predictions. However, a notable trend emerges: the linear regression model consistently exhibits lower Mean Squared Error (MSE) values across all datasets. This pattern suggests that the linear regression model tends to produce predictions with comparatively less error than the decision tree regressor model for each dataset utilized.

In a typical linear regression model graph, the x-axis represents the actual values of the target variable (e.g., sales), and the y-axis represents the predicted values of the target variable generated by the linear regression model. Each data point on the graph corresponds to a particular observation in the dataset. The ideal scenario is that all data points lie close to a diagonal line, indicating a perfect prediction where actual values match predicted values. Deviations from the diagonal line suggest discrepancies between actual and predicted values. A scattered or dispersed distribution of data points around the line may indicate variability or errors in the model's predictions.

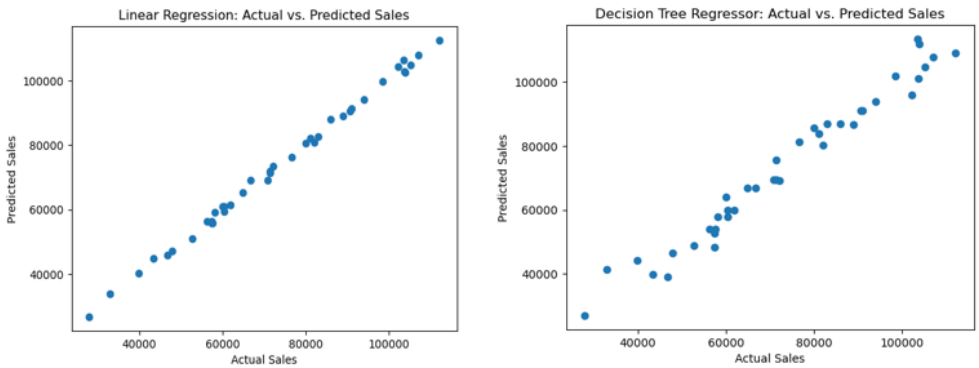

Figure 1. Predicted sales graph for linear regression and decision tree regressor models.

Figure 2. Predicted house price graph for linear regression and decision tree regressor models.

Figure 3. Predicted performance score for linear regression and decision tree regressor models.

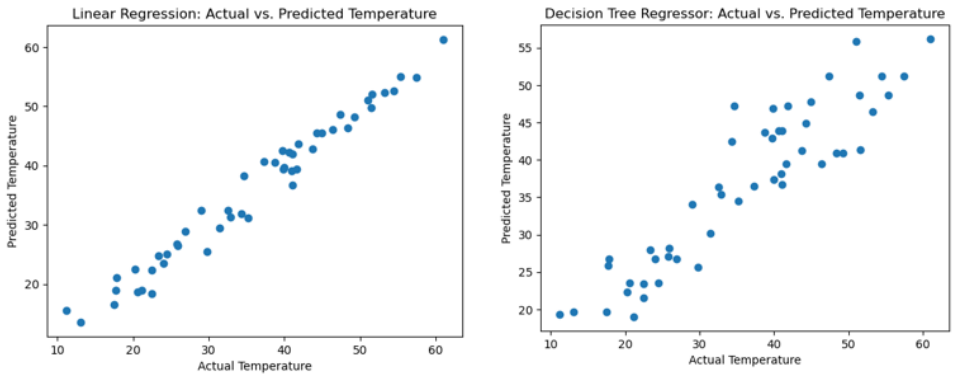

Figure 4. Predicted temperature for linear regression and decision tree regressor models.

On the other hand, the graph for a decision tree regressor model typically showcases the actual values of the target variable on the x-axis and the corresponding predicted values on the y-axis. Unlike the linear regression model, the graph for a decision tree regressor model may not necessarily follow a linear pattern. Instead, it may exhibit a series of steps or plateaus corresponding to the splits and decisions made by the decision tree algorithm. The visualization may reveal distinct clusters or groups of data points, each associated with a different leaf node or prediction from the decision tree model.

From **Figure 1**, which shows the graphs of linear regression and decision tree

regressor models predicting sale using the dataset 1, and going by the ideal scenario of graphs for linear regression models, it is comparatively clear that data points lie very close to the diagonal line for the linear regression model graph compared to that of the decision tree regression model. For the **Figure 2**, which displays the predicted house price graphs for the linear regression and decision tree regressor models, it can be observed that both graphs seem to have very scatted data points, however, the graph for the decision tree regressor seem to have a more scatted data points compared with the linear regression model. A similar conclusion can be said about **Figures 3** and **4** respectively.

5. Conclusion

In conclusion, this research has presented a comprehensive comparison between linear regression and decision tree regression models for predicting productivity within industrial contexts. By leveraging real-world datasets and employing rigorous evaluation methodologies, valuable insights have been obtained regarding the efficacy and limitations of each modeling approach.

The results indicate that while both linear regression and decision tree regression models can produce significant predictions, there are discernible differences in their performance across diverse datasets. Specifically, the linear regression model consistently exhibited lower Mean Squared Error (MSE) values across all datasets, suggesting comparatively less error in its predictions compared to the decision tree regression model. This pattern underscores the robustness and reliability of the linear regression approach in capturing complex relationships within industrial datasets.

Furthermore, the comparative analysis highlighted the importance of selecting an appropriate predictive model tailored to specific industrial contexts. While decision tree regression models may offer interpretability and flexibility, their performance may vary depending on the nature and complexity of the dataset. In contrast, linear regression models provide a straightforward and interpretable framework for predicting productivity, particularly in scenarios where the relationships between variables are predominantly linear.

Overall, this study contributes to the existing body of knowledge by offering empirical evidence and practical insights into the effectiveness of linear regression and decision tree regression models in predicting productivity within industrial settings. The findings have the potential to inform strategic decision-making processes, optimize resource allocation, and drive performance improvement initiatives across diverse industrial sectors.

Moving forward, future research endeavors could explore more advanced predictive modeling techniques, such as ensemble methods or neural networks, to further enhance the accuracy and robustness of productivity predictions in industrial contexts. Additionally, longitudinal studies could be conducted to assess the long-term predictive performance and scalability of different modeling approaches in dynamic industrial environments.

In summary, this research serves as a foundation for future investigations aimed at optimizing productivity prediction methodologies and ultimately fostering innovation and efficiency within industrial sectors.

Conflict of interest: The author declares no conflict of interest.

References

1. Ferraro MB, Coppi R, González Rodríguez G, et al. A linear regression model for imprecise response. International Journal of Approximate Reasoning. 2010; 51(7): 759-770. doi: 10.1016/j.ijar.2010.04.003
2. Diskin MH. Definition and Uses of the Linear Regression Model. Water Resources Research. 1970; 6(6): 1668-1673. doi: 10.1029/wr006i006p01668
3. Mendes da Luz I, Lourenço Niza I, Broday EE. Comparative Analysis of Productivity Methods to Evaluate Thermal Comfort in Classrooms through Hierarchical Clustering. Buildings. 2024; 14(3): 698. doi: 10.3390/buildings14030698
4. Golnaraghi S, Zangenehmadar Z, Moselhi O, et al. Application of Artificial Neural Network(s) in Predicting Formwork Labour Productivity. Advances in Civil Engineering. 2019; 2019: 1-11. doi: 10.1155/2019/5972620
5. Krämer W, Sonnberger H. The linear regression model under test. Springer Science & Business Media; 2012.
6. Anyaeche CO, Ighravwe DE. Predicting performance measures using linear regression and neural network: A comparison. African Journal of Engineering Research, 2013; 1(3): 84-89.
7. Gevrekçi Y, Takma Ç. A Comparative Study for Egg Production in Layers by Decision Tree Analysis. Pakistan Journal of Zoology. 2018; 50(1). doi: 10.17582/journal.pjz/2018.50.2.437.444
8. Weisberg S. Applied Linear Regression. In: Wiley Series in Probability and Statistics. Wiley Online Library; 2005. doi: 10.1002/0471704091
9. Kai S, Almeda MV, Baker RS, et al. Decision tree modeling of wheel-spinning and productive persistence in skill builders. Journal of Educational Data Mining. 2018;10(1): 36-71.
10. Gomes C, Almeida LS. Advocating the broad use of the decision tree method in education. Practical Assessment, Research, and Evaluation. 2019; 22(1): 10.
11. Charbuty B, Abdulazeez A. Classification Based on Decision Tree Algorithm for Machine Learning. Journal of Applied Science and Technology Trends. 2021; 2(01): 20-28. doi: 10.38094/jastt20165
12. Azad C, Bhushan B, Sharma R, et al. Prediction model using SMOTE, genetic algorithm and decision tree (PMSGD) for classification of diabetes mellitus. Multimedia Systems. 2021; 28(4): 1289-1307. doi: 10.1007/s00530-021-00817-2
13. Kaparthi S, Bumblauskas D. Designing predictive maintenance systems using decision tree-based machine learning techniques. International Journal of Quality & Reliability Management. 2020; 37(4): 659-686. doi: 10.1108/ijqrm-04-2019-0131
14. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Springer New York; 2009. doi: 10.1007/978-0-387-84858-7
15. Greene WH. Econometric Analysis. Pearson Education; 2012.
16. Montgomery DC, Peck EA, Vining GG. Introduction to Linear Regression Analysis. John Wiley & Sons; 2012.
17. Kutner MH, NachtsheimCJ, NeterJ. Applied Linear Statistical Models. McGraw-Hill Education; 2004.
18. Breiman L, Friedman JH, Olshen RA., Stone CJ. Classification and Regression Trees. CRC press; 1984.
19. Quinlan JR. Induction of decision trees. Machine Learning. 1986; 1(1): 81-106. doi: 10.1007/bf00116251
20. Louppe G, Wehenkel L, Sutera A, Geurts P. Understanding variable importances in forests of randomized trees. Advances in Neural Information Processing Systems. 2013; 431-439.