Journal of Advances in Manufacturing Engineering
Web page info: https://jame.yildiz.edu.tr
DOI: 10.14744/ytu.jame.2024.00004

Original Article

The application of machine learning algorithms in the estimation of production lead times: A case study of a steel construction manufacturing company

Şeyma DUYMAZ*[1], Ali Fuat GÜNERİ[2]

[1]*Department of Industrial Engineering, Faculty of Engineering and Natural Sciences, İstanbul Rumeli University, İstanbul, Türkiye*
[2]*Department of Industrial Engineering, Faculty of Mechanical Engineering, Yıldız Technical University, İstanbul, Türkiye*

ARTICLE INFO

Article history
Received: 24 May 2024
Revised: 16 June 2024
Accepted: 28 June 2024

Key words:
Machine learning, production time, random forest regressor, regression, steel construction.

ABSTRACT

In companies with a diverse range of products, it can be challenging to regulate production times. Having an understanding of production lead times is crucial for addressing issues such as deadlines, cost, production scheduling, and customer satisfaction. It is challenging for the company to provide the customer with an accurate estimation of the time required to produce and deliver a product that has not been produced before. One of the advantages of knowing the production times is to be able to adjust the machines to use them more efficiently when preparing the production plan. In this study, input data were obtained from a steel construction company, including dimensions such as size, diameter, and weight. Additionally, the times required for the production of different products were measured. Based on these times, production times were estimated using machine learning algorithms, including Decision Tree, Random Forest, and Gradient Boosting. Consequently, precise predictions were generated with an accuracy rate of approximately 96.9%. A test data set was then created with the objective of estimating the time required to produce a product that has never been produced. Additionally, the times of products that have not yet been produced were estimated. For each new product ordered, the machine must be adjusted and calibrated separately, which represents a significant loss of time and cost for the company. The objective of this research is to develop a model that can predict the time required to deliver a new product once it has been ordered. Furthermore, the aim is to enhance the efficiency of machine utilization.

Cite this article as: Duymaz, Ş., & Güneri, AF. (2024). The application of machine learning algorithms in the estimation of production lead times: A case study of a steel construction manufacturing company. *J Adv Manuf Eng*, 5(1), 21–28.

INTRODUCTION

In the contemporary business environment, the efficiency of production processes is a crucial factor in achieving success in a competitive market. Accurate forecasting of production lead times offers numerous advantages, including the efficient use of resources, cost reduction, and increased customer satisfaction. It also prevents delays in order deliveries. Therefore, estimating production lead times is of great importance for businesses [1]. All planning in the process, including cost, delivery, and other factors, is based on the realization time of the work steps [2]. In a manufacturing sector comprising a vast array of products, time studies are conducted on existing products to ascertain the time re-

***Corresponding author.**
*E-mail address: seyma.duymaz@rumeli.edu.tr

quired for production. These times are incorporated into the calculation of the delivery date or the production planning program. One of the methodologies employed to estimate production times is machine learning algorithms. Machine learning algorithms are employed in a multitude of fields. Their applications extend to diverse domains, including forestry studies [3], the retail sector [4], land cover mapping [5], the clustering of court decisions [6], and the detection and diagnosis of chronic diseases [7]. The deployment of machine learning algorithms facilitates the identification of production failures, the estimation of remaining useful life, and the estimation of production times. This paper presents a review of selected studies in the literature.

In a study by Lim et al. [8], support vector machines were employed to forecast production time based on work order and production data. The accuracy of the predictions was quantified at 84.62%. In a study by Schneckenreither et al. [9], an artificial neural network was employed to construct a flow time prediction model that could dynamically adjust lead times. In their study, Yüce et al. [10] employed artificial neural network, support vector regression, and gradient boosting algorithms from the field of machine learning to estimate production times in specific production areas of a production facility. In their study, Haeussler and colleagues [11] examined workload control models with fixed and dynamic lead times. In a related study, El Mekkaoui et al. [12] employed artificial neural networks and a random forest algorithm to predict the arrival time of ships. In a previous study, Agwu et al. [13] employed an artificial neural network to predict oil flow rates. In a study on production time estimation in a construction factory, Alsakka et al. [14] employed a methodology based on machine learning. Gyulai et al. [15] utilized machine learning methodologies to estimate production time, thereby facilitating informed decision-making in a manufacturing context. Chen et al. [16] employed machine learning in manufacturing studies. Dehghani et al. [17] employed a random forest algorithm to predict shear wave velocity.

In this study, for different product types of a manufacturing company producing in the field of steel construction;

1 The objective is to determine the time required to produce 99 distinct products on a single machine, each with a defined production time,
2 The objective is to calculate production times for products with varying length, volume, and weight values that have not yet been produced. This approach enables the provision of more precise delivery time estimates to customers without the necessity of manufacturing the products.

This study makes a significant contribution to the existing literature by providing insight into the time required to produce a diverse range of products by a company engaged in steel construction. The study also identifies the time required to deliver these products to the customer. Furthermore, a dynamic production planning schedule can be prepared by enhancing the efficiency of machine utilization. The utilization of machine learning algorithms will facilitate the estimation of production times. This study will enable the analysis of higher-dimensional data from production and the automatic calculation of the duration of instant orders. In contrast to previous studies, a data set comprising products that have never been manufactured was subjected to analysis. The model offers the company a number of benefits. Firstly, it will enable the company to ascertain the delivery times for instant orders. Secondly, it will allow the company to assign parts to machines at the optimal level. This will facilitate the creation of production planning schedules. The calibration time of the machines will be reduced.

In the Materials and Methods section of the study, the algorithms and performance metrics utilized are delineated. In the Results and Discussion section, the data is described and the application is made to a data set derived from a steel construction company. In the final section, the results of the study are interpreted and practical benefits are enumerated.

MATERIALS AND METHODS

Machine learning algorithms are employed in a multitude of sectors. A number of algorithms exist, including those based on artificial neural networks, support vector machines, random forests, decision trees, and gradient boosting. In this study, the random forest, decision tree, and gradient boosting algorithms are employed. As the problem addressed in this study is a forecasting problem, regression techniques are utilized.

Random Forest Algorithm

The random forest algorithm is based on tree-based models, which are a type of supervised machine learning algorithm. Tree-based models involve recursively partitioning the dataset into two groups, depending on a stopping condition. Each node depends on the previous node. This algorithm is useful for both classification and regression problems. In classification problems, the splitting criterion is determined by entropy calculation. In regression problems, the most commonly used splitting criterion is the mean squared error at each internal node, as described in [18]. The Random Forest Regressor is an ensemble model comprising a multitude of decision trees. Its function is to reduce variance by averaging the results. When forecasting for a given data set, it utilises the prediction of each decision tree and averages these predictions to create the final prediction. The tree structure grows incrementally and is averaged at each step. It is an algorithm that is sensitive to overlearning. Different examples can be generated from one data set [19]. Figure 1 illustrates the structure of the decision tree, which is a self-iterative structure.

The random forest algorithm is more accurate than decision trees in estimating the error rate.

Algorithm steps;

- The data set to be used in the study is prepared,
- Trees are created for the samples and the results are predicted,
- Averaging for the regression problem
- The process continues recursively according to the stopping criterion

as the most important factors.

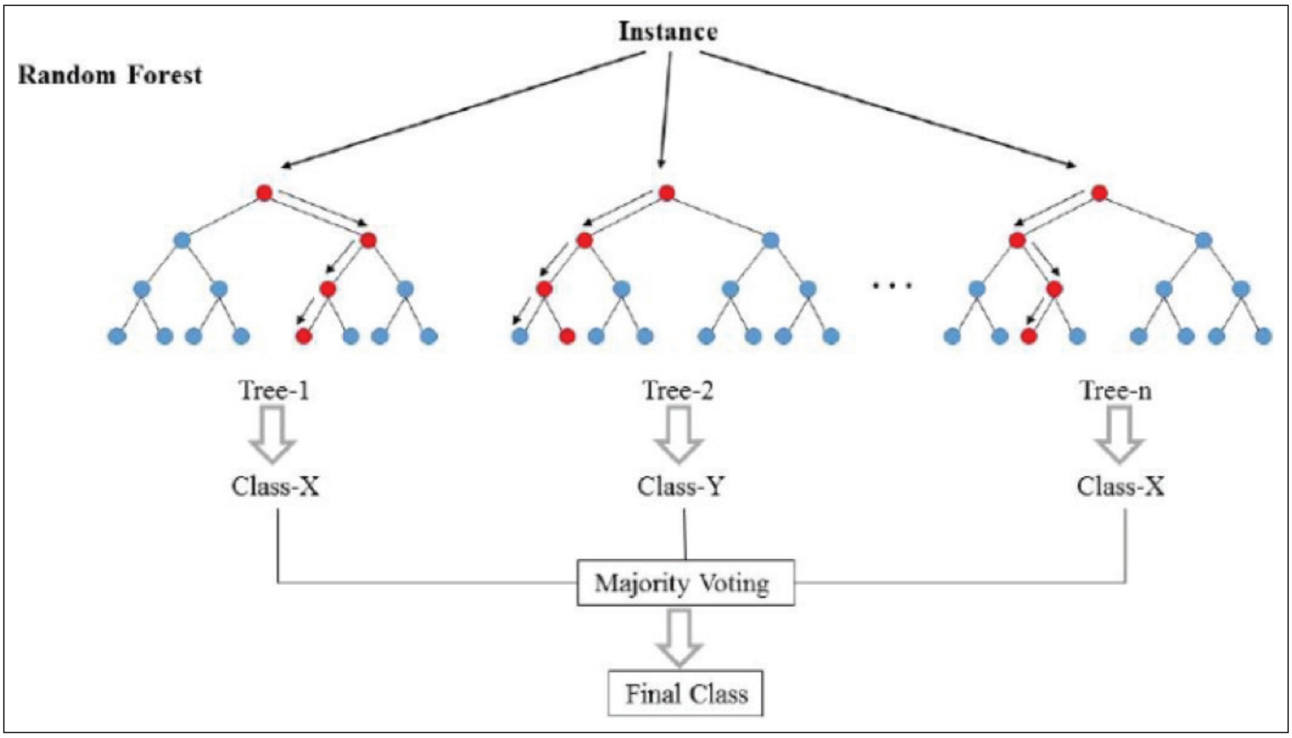

Figure 1. Structure of decision tree [20].

In this study, the Anaconda Jupyter interface was employed. The random forest regressor algorithm, one of the machine learning algorithms in scikit-learn, one of Python's open source libraries, was utilized because the problem at hand is a prediction problem.

Decision Tree Algorithm

Decision trees represent a supervised learning algorithm that automates the decision-making process by identifying the optimal solution from a set of alternatives. The decision tree algorithm is a nonparametric prediction model that can be used to address both regression and classification problems [21]. In the case of categorical final values, the problem is classified, whereas in the case of continuous values, it is a regression problem [22]. The algorithm begins by extracting examples from the data set and then subdividing them into subclasses. In essence, decision trees make predictions by dividing the data into branches according to their characteristics. This approach entails dividing the data set into a tree structure, with each node representing a specific value of a feature [23]. Leaf nodes contain continuous values, and the tree makes predictions based on the average of these values. The structure of the decision tree consists of a root node, an internal node, and leaves. The root node is the first node and contains the entire data set. The internal node represents the data set partitioned into multiple subsets [24]. The leaves are the final predictions. In this study, the decision tree algorithm is employed to estimate the production time.

Gradient Boosting Algorithm

Gradient Boosting Regressor is a powerful and flexible machine learning algorithm for regression problems. It is a decision tree-based algorithm. It is an iterative variant of sequentially organized tree models. It first runs the model, detects errors, then runs it again and so on iteratively. It continues in a stronger way by learning from the previous step [25]. Each new decision tree generated by the Gradient Boosting algorithm is based on the principle of minimizing the errors calculated in the previous tree [26]. This results in a more accurate outcome by correcting the errors in a sequential manner. Gradient Boosting employs gradient descent at each step to minimize the errors.

The algorithm functions according to the principle of transforming learners with limited abilities into those with enhanced capabilities. It can be described as an ensemble algorithm.

Steps of the algorithm;

- Examine the compatibility of the equation coefficients with the data by defining the loss function,
- Determining the state where the loss function is minimum by determining the fixed variable,
- Calculation of errors,
- Estimation for each observation,

it is listed in four articles [27].

Performance Metrics

Mean square error (MSE): The metric is employed in the context of regression problems and represents the average of the squares of the differences between actual and predicted values. As this value approaches zero, the performance of the model improves.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \overline{y}_i)^2 \quad (1)$$

Root mean square error (RMSE): It is obtained by taking the square root of the root mean square error. The RMSE value shows the closeness of the estimates to the ac-

Table 1. Data set

	Material name	Size (mm)	Diameter (mm)	Weight (kg)	Production time
1	Stud, scissors	166.5	37.0	0.721	10.0
2	Stud, scissors	163.0	36.0	0.730	8.0
3	Stud, scissors	163.8	35.0	0.728	7.9
4	Stud, scissors	163.0	35.0	0.730	10.0
5	Stud, scissors	164.2	37.0	0.732	10.2
...	...	...	...	...	...
95	Brake caliper pin	110.0	37.0	0.540	10.0
96	Brake caliper pin	111.0	36.0	0.530	8.0
97	Brake caliper pin	112.0	37.0	0.564	9.0
98	Brake caliper pin	111.0	37.0	0.558	9.0
99	Brake caliper pin	114.0	35.0	0.561	8.0

Table 2. Data types

RangeIndex: 99 entries, 0 to 98				
Data columns (total 5 columns):				
# Column			Non-null count dtype	
--- ------			-------------- -----	
0	Material name	99	Non-null	Object
1	Size (mm)	99	Non-null	Float64
2	Diameter (mm)	99	Non-null	Float64
3	Weight (kg)	99	Non-null	Float64
4	Production time	99	Non-null	Float64
Dtypes: Float64(4), Object(1).				

Table 3. Statistical information of the data set

	Size (mm)	Diameter (mm)	Weight (kg)	Production time
Count	99.000.000	99.000.000	99.000.000	99.000.000
Mean	124.601.515	43.448.990	0.856357	40.154545
Standart	39.386.151	10.024.778	0.549796	47.656259
Min	43.000.000	20.000.000	0.020000	7.000000
25%	104.500.000	37.000.000	0.429000	9.000000
50%	113.500.000	46.500.000	0.587000	16.000000
75%	163.400.000	49.300.000	1.377.000	45.000000
Max	181.000.000	60.100.000	1.860.000	153.000l000

tual values. It is a second-order error metric that measures the magnitude of the error. RMSE is the standard deviation of the difference between actual and predicted values. An RMSE value close to zero means good performance [28].

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \overline{y}_i)^2} \ (2)$$

Mean absolute error (MAE): The sum of the absolute values of the differences between the actual value and the predicted value.

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \overline{y}_i| \ (3)$$

Mean absolute percentage error (MAPE): It expresses the error between actual values and predicted values expressed as a percentage. The closer it is to 0%, the more meaningful the results [29].

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} \left|\frac{y_i - \overline{y}_i}{y_i}\right| \ (4)$$

R square (R^2): A performance metric describing the accuracy of the model.

$$R^2 = 1 - \frac{\sum_i (y_i - \overline{y}_i)^2}{\sum_i (y_i - \overline{y})^2} \ (5)$$

yi=Actual values
$\overline{y}$i=Estimated values
N=Number of observations

RESULTS AND DISCUSSION

Identification of Data

This study was conducted using a data set obtained from a company engaged in the production of steel structural products. The company provided input values for 99 finished products of varying sizes, diameters, and weights, as well as the time required to produce them on one machine. The data set is presented in Table 1. The size, diameter, and weight were used as input data, while the production time was used as an output value for estimation purposes.

Table 2 illustrates the data types and indicates that the data set does not contain null values. The analysis of data sets with empty values is a challenging endeavor.

Table 3 presents the results of the statistical calculations. It was observed that the products with the lowest standard deviation values were weight values. The highest standard deviation was observed between production times. In this case, it can be concluded that there is not much difference between weight values, but there is more difference between production times.

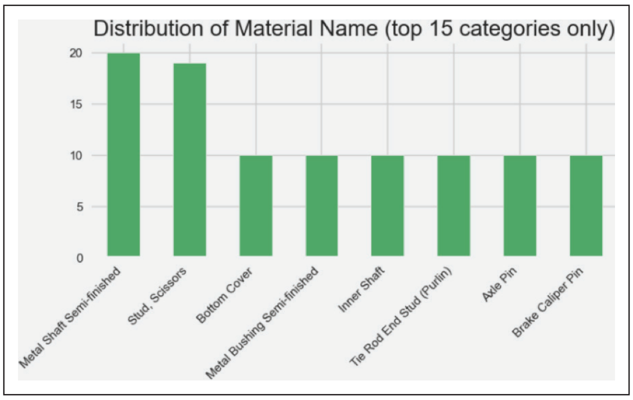

Figure 2. Distribution of material name.

Visualization of the Data Set

The Autoviz library, one of the open-source libraries in Python, was employed for the purpose of data visualization. Figure 2 illustrates the frequency graph of eight distinct product types. A total of 99 unique products belonging to eight different product types were produced. The most prevalent product type in this instance is Metal Shaft Semi-Finished Product.

Figure 3 illustrates the distribution of size, diameter, and weight values according to production times. It can be observed that in certain instances, the discrepancy between the production times of products with comparable lengths is considerable. This is due to the influence of additional variables on the time required.

Figure 4 illustrates the binary relationships between the variables.

Figure 5 shows the distribution of the products produced in terms of size, diameter, weight and production times.

Case Study

In this study, the Random Forest Regressor, Decision Tree Regressor, and Gradient Boosting Regressor algorithms were employed. The results are presented in Table 4.

In addition to the three distinct algorithms employed, the gradient boosting regressor algorithm, which yields the optimal outcome, is utilized once more through cross-validation (CV). The objective here is to enhance performance. In CV, the data set is partitioned into multiple training and test sets, the model is trained and tested on these sets, and the results are averaged to obtain more reliable performance metrics. This process allows for the identification and rectification of issues such as overfitting and underfitting.

The Gradient Boosting Regressor (CV) model demonstrated superior performance relative to the other models. In particular, the mean absolute percentage error (MAPE)

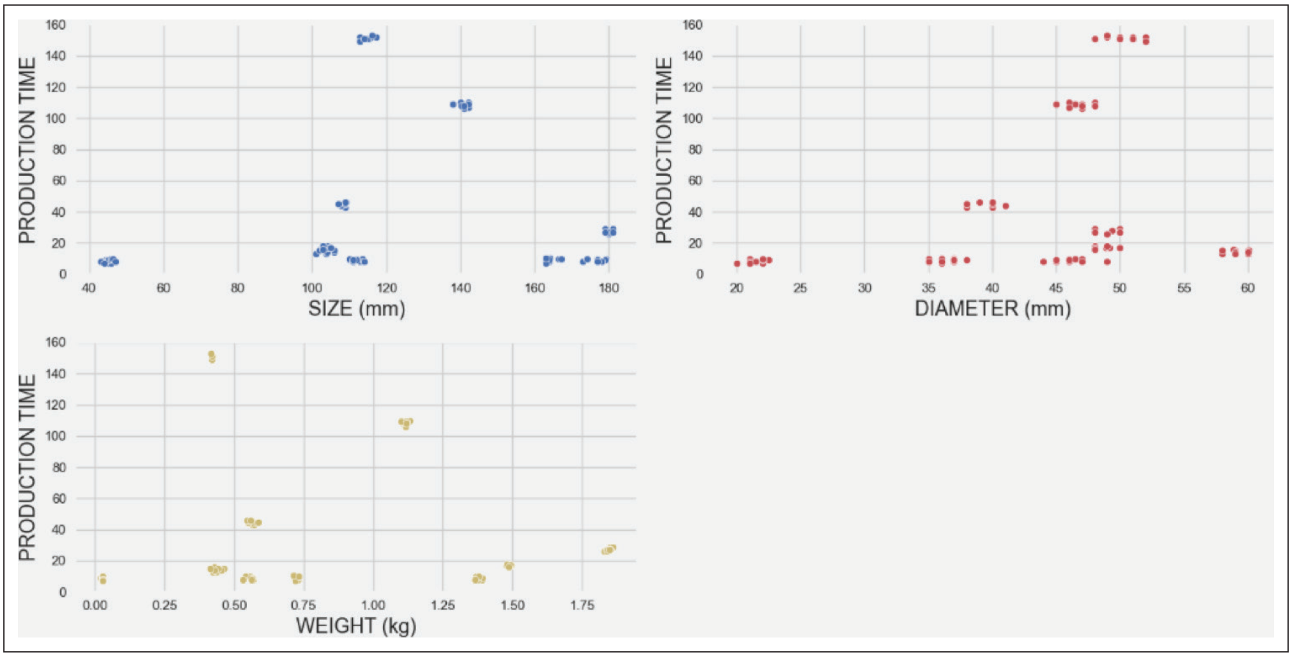

Figure 3. Scatter plot of each continuous variable vs target.

Table 4. Results

Algorithm	R2	MAE	MSE	RMSE	MAPE
Random forest regressor	0,924053709	8	192,32	13,87	198,90%
Gradient boosting regressor	0,968616182	4,6	79,48	8,91	214,46%
Decision tree regressor	0,962118471	4	95,93	9,79	213,13%
Gradient boosting regressor (CV)	0.969727010	4.3	76.66	8.75	19.25%

MAE: Mean absolute error, MSE: Mean square error, RMSE: Root mean square error, MAPE: Mean absolute percentage error.

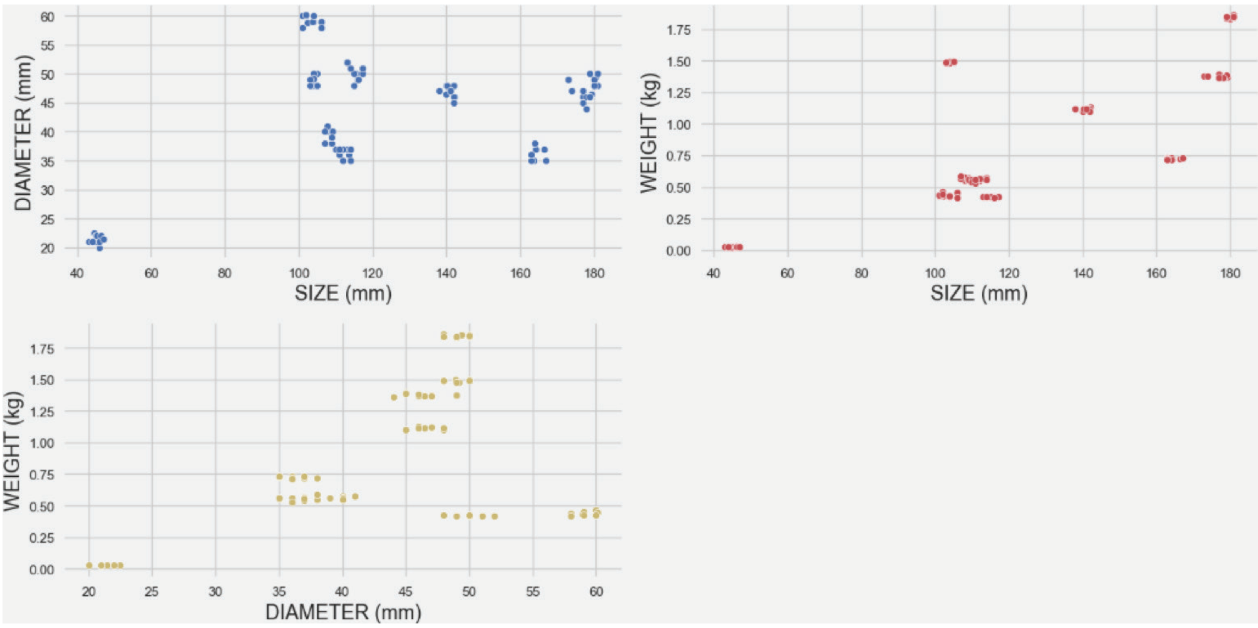

Figure 4. Pair-wise scatter plot of all continuous variables.

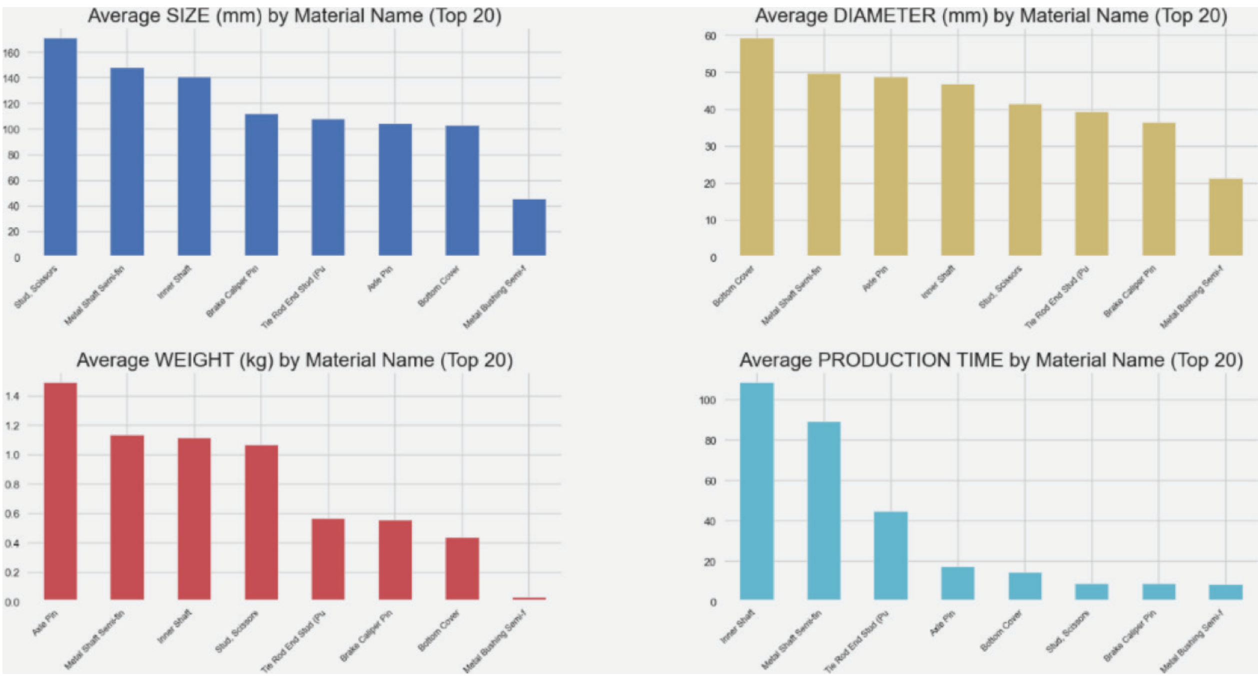

Figure 5. Distribution of products according to variables.

value of 19.25% was notably lower than the other models, indicating that the model's predictions were highly accurate.

The Random Forest Regressor and Decision Tree Regressor models exhibit high R2 values, yet simultaneously display high MAPE values. This indicates that the model is prone to making significant errors in certain instances.

The Gradient Boosting Regressor model also demonstrates satisfactory performance, with a high R2 value and low MAE, MSE, and RMSE. However, the MAPE value is relatively high, indicating that the model may be prone to significant errors.

In conclusion, the Gradient Boosting Regressor (CV) model is the most effective in terms of overall performance. In the future, it would be beneficial to investigate and implement methods to reduce the MAPE values. The algorithm's accuracy, estimated at approximately 96.9%, was utilized to create a dataset comprising product features that had never been produced. The dataset was subjected to testing and prediction in order to ascertain the estimated time required for the production of the products in question. The results are presented in Table 5.

Table 5. Test set

	Size (mm)	Diameter (mm)	Weight (kg)	Production time
0	156	50	1.5	108
1	156	43	1.5	110
2	140	50	0.45	110
3	168	47	1.3	10
4	145	52	1.6	108
5	160	35	2	10
6	143	44	0.45	109.2
7	200	45	0.3	9
8	177	56	0.89	10
9	145	64	0.96	110
10	113	52	0.4198	149

CONCLUSION

The estimation of production lead times is of significant importance for the purposes of planning and scheduling, cost control, resource management, inventory management, and customer satisfaction. In companies with a high product variety, the only way to determine the time required to produce each new order is to produce the product once. This necessitates the readjustment and recalibration of machines to facilitate the production of the new order. Consequently, communicating lead times to customers represents a significant challenge. It is of great importance to provide accurate lead times in order to ensure customer satisfaction. In this study, the estimated accuracy of production lead times for existing products was found to be approximately 96.9%. Nevertheless, forecasting the present circumstances will not assist the company. The company has already attempted this and has gained insight into the time required to produce existing products. The primary objective of this study is to predict the time required to produce a previously unproduced product order. The dataset provided by the company was subjected to analysis using machine learning algorithms, resulting in a highly accurate prediction. The study will enable the company to identify the production times of products with varying parameters. This will facilitate the estimation of production times.

Upon analysis of the algorithms, it becomes evident that reliable predictions can be made. These predictions can be utilized by the company in a number of ways, including predicting delivery times for products ordered, notifying customers of delivery times, and creating production planning schedules.

This study contributes to the use of machine learning algorithms in the estimation of production times. A dynamic scheduling can be created by optimizing the use of machines. These studies are of great importance to companies that wish to minimize costs and maximize customer satisfaction.

Data Availability Statement

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

Author's Contributions

Şeyma Duymaz: Modeling, Analyzing, Writing, Interpreting Results.

Ali Fuat Güneri: Obtaining the data set, Review-editing.

Conflict of Interest

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

Use of AI for Writing Assistance

Not declared.

Ethics

There are no ethical issues with the publication of this manuscript.

REFERENCES

[1] Işık, K., & Ulusoy, S. K. (2021). Determination of factors affecting production times in the metal sector by data mining methods. *Journal of Gazi University Faculty of Engineering and Architecture, 36*(4), 1949–1962. [CrossRef]

[2] Gökler, S. H., Ulus, İ., Cihat, F., & Boran, S. (2015). Estimation of machine part replacement times based on a cost model using two-parameter Weibull distribution. *Beykent University Journal of Science and Engineering, 8*(2), 61–84. [CrossRef]

[3] Eker, R., Alkiş, K. C., Uçar, Z., & Aydın, A. (2023). The use of machine learning in forestry. *Turkish Journal of Forestry, 24*(2), 150–177. [CrossRef]

[4] Sinap, V. (2024). Comparative performance analysis of machine learning algorithms in the retail sector: Black Friday sales forecasting. *Selçuk University Social Sciences Vocational School Journal, 27*(1), 65–90. [CrossRef]

[5] Danacıoğlu, Ş. (2023). Evaluation of different machine learning algorithms for land cover mapping: The case of Izmir Province. *Turkish Journal of Geography*, (84), 105–117. [CrossRef]

[6] Görentaş, M. B., & Uçkan, T. (2023). Clustering court decisions using machine learning methods. *Computer Science, 8*(2), 148–158.

[7] Melekoğlu, E. (2023). *Development of a new machine learning-based method for the diagnosis of chronic obstructive pulmonary disease* (Doctoral dissertation). Available from National Thesis Center. (Accession No. 789510.)

[8] Lim, Z. H., Yusof, U. K., & Shamsudin, H. (2019). Manufacturing lead time classification using support vector machine. In: Badioze Zaman, H., et al. Advances in Visual Informatics: 6th International Visual Informatics Conference, IVIC 2019, Bangi, Malaysia, November 19–21, 2019, Proceedings 6 (pp. 268–278). Springer International Publishing. [CrossRef]

[9] Schneckenreither, M., Haeussler, S., & Gerhold, C. (2021). Order release planning with predictive lead times: A machine learning approach. *International Journal of Production Research, 59*(11), 3285–3303. [CrossRef]

[10] Yüce, T., & Kabak, M. (2021). Predicting production times for detailed production areas at work center level using machine learning algorithms. *Erciyes University Journal of the Institute of Science and Technology, 37*(1), 47–60.

[11] Haeussler, S., & Netzer, P. (2020). Comparison between rule-and optimization-based workload control concepts: A simulation optimization approach. *International Journal of Production Research, 58*(12), 3724–3743. [CrossRef]

[12] El Mekkaoui, S., Benabbou, L., & Berrado, A. (2022). Machine learning models for efficient port terminal operations: Case of vessels' arrival times prediction. *IFAC-PapersOnLine, 55*(10), 3172–3177. [CrossRef]

[13] Agwu, O. E., Alkouh, A., Alatefi, S., Azim, R. A., & Ferhadi, R. (2024). Utilization of machine learning for the estimation of production rates in wells operated by electrical submersible pumps. *Journal of Petroleum Exploration and Production Technology*, (14), 1205–1233. [CrossRef]

[14] Alsakka, F., Yu, H., El-Chami, I., Hamzeh, F., & Al-Hussein, M. (2024). Digital twin for production estimation, scheduling and real-time monitoring in offsite construction. *Computers & Industrial Engineering, 191*, Article 110173. [CrossRef]

[15] Gyulai, D., Pfeiffer, A., Bergmann, J., & Gallina, V. (2018). Online lead time prediction supporting situation-aware production control. *Procedia CIRP, 78*, 190–195. [CrossRef]

[16] Chen, S., Li, X., Liu, R., & Zeng, S. (2019). Extension data mining method for improving product manufacturing quality. *Procedia Computer Science, 162*, 146–155. [CrossRef]

[17] Dehghani, M., Jahani, S., & Ranjbar, A. (2024). Comparing the performance of machine learning methods in estimating the shear wave transit time in one of the reservoirs in southwest of Iran. *Scientific Reports, 14*(1), Article 4744. [CrossRef]

[18] Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal, 20*(1), 3–29. [CrossRef]

[19] Das, O., & Bagci Das, D. (2022). Free vibration analysis of isotropic plates using regressive ensemble learning. *European Journal of Science and Technology*, (38), 428–434.

[20] Çelik, S., & Özdemir, D. (2023). Bitcoin price prediction using random forest regression algorithm. *Journal of Scientific Reports-B*, (008), 55–64.

[21] Rokach, L., & Maimon, O. (2014). *Data mining with decision trees: Theory and applications* (2nd ed.). World Scientific Publishing Co. Pte. Ltd. [CrossRef]

[22] Doğruel, M., & Ümit Fırat, S. (2021). Estimation of countries' innovation values using data mining decision trees and a comparative application with linear regression model. *Istanbul Business Research, 50*(2), 465–493.

[23] Özkan, B., Parim, C., & Çene, E. (2023). Prediction of countries' development levels using decision tree and random forest methods. *EKOIST Journal of Econometrics and Statistics*, (38), 87–104. [CrossRef]

[24] Çelik, S., Çetinkaya Bozkurt, Ö., & Ekşili, N. (2022). Determination of expressions in employee performance scale with decision tree algorithm. *Journal of Mehmet Akif Ersoy University Economics and Administrative Sciences Faculty, 9*(1), 561–584. [CrossRef]

[25] Rao, H., Shi, X., Rodrigue, A. K., Feng, J., Xia, Y., Elhoseny, M., & Yuan, X. (2019). Feature selection based on artificial bee colony and gradient boosting decision tree. *Applied Soft Computing Journal, 74*, 634–642. [CrossRef]

[26] Kelle, A. C., & Yüce, H. (2022). Classification of DoS attacks on MQTT traffic using machine learning and interpretation of the model with SHAP. *Journal of Materials and Mechatronics: A, 3*(1), 50–62. [CrossRef]

[27] Yeşilyurt, S., & Dalkılıç, H. (2021). *Daily river flow prediction with Xgboost and gradient boost machine.* In 3rd International Symposium of III Engineering Applications on Civil Engineering and Earth Sciences, Karabük, Turkey.

[28] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? *Geoscientific Model Development, 7*, 1247–1250. [CrossRef]

[29] Özen, N. S., Saraç, S., & Koyuncu, M. (2021). Prediction of COVID-19 cases with machine learning algorithms: The case of the United States. *European Journal of Science and Technology*, (22), 134–139.