



JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

RAW MATERIAL DEMAND FORECASTING IN AUTOMOBILE INDUSTRY

¹Abhishek Jadhav, ¹Nitesh Pandit, ¹Pratik Thorat, ¹Pratik Sonawane, ²Swati R. Khokale

¹Student, Sandip Institute of Technology and Research Centre, Nashik, India

²Assistant Professor, Department of Information Technology, Sandip Institute of Technology and Research Centre, Nashik, India

Abstract : Raw material procurement planning is crucial for automobile manufacturers to stay competitive in the global market. Forecasting demand for raw materials is very helpful for production planning. We implemented machine learning algorithms to forecast demand for raw material quantities needed in the future using past data. Machine learning algorithms can provide better accuracy than traditional forecasting techniques and can be flexible to changes. In this paper we implemented linear regression, Multiple output linear regression, random forest and k-nearest neighbour using the scikit-learn package to train ML models and evaluated their results using techniques such as RMSE, Rsquare.

IndexTerms - Machine Learning, Automobile Industry, Raw Materials, Demand Forecasting, Linear Regression.

INTRODUCTION

The Indian Automobile Industry is one of the major contributing factors in the growth of the country's economy. With ever-increasing demand, the Indian automotive industry is expected to surpass Rs. 16 trillion by 2026. With such growth, it is important to improve production planning processes.

Planning is regarded as a vital procedure in the manufacturing of products. The demand and sales of automobile vehicles are increasing. However, the business strategy requires aims based on production targets that require the need for raw material demand forecasting.

The manufacturing industry, especially for the automotive sector, has a great amount of production to fulfill the customer's demand. Production planning is one of the most crucial processes in the automobile industry. One of the most important activities in production planning is predicting the amount of raw materials required for production.

Continuous supply of raw material is essential for smooth production processes. Due to various factors there can be delays in raw material supply which can halt for months, which can disrupt production cycle. In order to overcome this situation and to create balance between production and supply of raw material. Experts have, for long, relied on traditional forecasting models to predict demand for raw materials.

But this can be improved with the help of machine learning algorithms, which can produce better results with increased accuracy, which will help increase revenue gains and aid in inventory cost optimization. Forecasting raw materials using machine learning will help the manufacturer to predict the quantity of raw materials in advance, which will indirectly contribute to company growth.

LITERATURE SURVEY

This paper describes a machine learning model which can be used in the manufacturing industry to maximize efficiency of production processes. The model trained on previously collected sample data is used to predict output for a given input. Efficiency of this model improves with experience. Thus, ML models can be used to solve various challenges faced by a manufacturing industry [1].

This paper presents information about materials used in the manufacturing of automobiles and changes in their relative proportions over the years. According to this study, A modern car consists of cast iron and steel - 55 %, plastics - 11%, aluminum - 9 %, rubber 7, glass 3 % and other non-ferrous alloys (magnesium, titanium, copper and zinc) are less than 1 %; other materials (varnishes, paints, electric wires, facing materials, etc.) - 13.5 %" [2].

This paper explains different machine learning techniques used for demand forecasting using historical data and analyses performance of different models. Different techniques have been analyzed, such as machine learning, time series regression, and deep learning models to develop strategies for demand forecasting [3].

This paper suggests forecasting raw material demand based on demand of finished goods will generate high cumulative error when the same material is used in production of multiple goods, another method of forecasting raw materials using historical data of raw material consumption might not represent the actual customer demand. These problems can be solved by using combined historical data of raw material quantities as well as finished goods [4].

This study uses the K-Nearest Neighbor algorithm to predict the accuracy of product delivery using the Administration of Raw Material model. The model developed in this study based on a KNN algorithm to predict delivery for products has high accuracy. In this study, the KNN Algorithm was found to be effective for predicting the accuracy of product delivery. This paper also shows that the KNN Algorithm is stable with a small error ratio, so the results are closer to the actual value of the target and can be used in decision-making for the company [5].

In this paper, a random forest regressor has been used to predict agricultural commodities prices and demand. Different models have been investigated for their performance and compared. The results show that the Random Forest Regression has higher accuracy compared to other models. Random Forests are an effective tool for this prediction. This model can assist decision-making and thereby achieve their profits goals and maintain a balance between deficit and surplus production [6].

This research compares 21 statistical and machine learning algorithms for demand forecasting in automotive OEM manufacturing and evaluates their performance. Various well-known machine learning algorithms were analyzed such as Multiple Linear Regression (MLR), Support Vector Machine (SVM), Linear Regression etc. and performance of these models were evaluated using mean absolute scaled error (MASE) and the R2 -adjusted (R2) metrics. Results showed that the models based on past demand had the best results in most aspects. In the experiments, the ML models outperformed the statistical ones [7].

This paper discusses various raw material forecasting techniques and evaluates them with different methods of evaluation. The aim of research is to calculate minimum requirements for each item and optimum stock quantities to be maintained. From the results it was concluded that linear regression had best forecasting accuracy among other techniques, and can be used to develop a planning strategy for procurement of raw materials [8].

This paper reviews performance metrics such as RMSE, MAE, MAPE and their use in machine learning regression and forecasting. Performance metrics are used to measure deviations of forecasts from observations in order to assess quality and choose forecasting methods [9].

This paper reviews Scikit-learn, a machine learning package in the Python programming language and shows how to use the machine learning methods included in the package. The Scikit-learn package is used to implement machine learning methods under unified data and modeling procedure conventions. It is a convenient tool for educational purposes[10].

SYSTEM METHODOLOGY

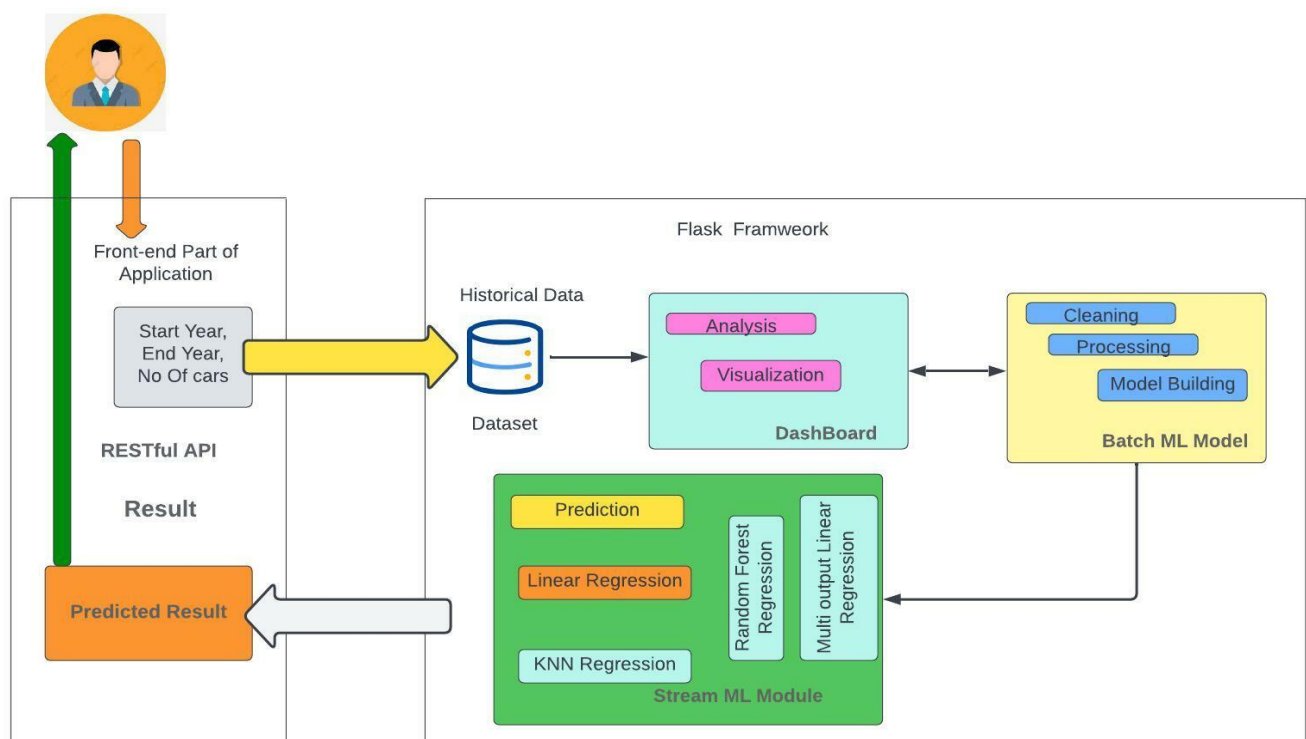


fig. System Architecture

The main objective of these proposed systems is to predict demand for raw materials using machine learning algorithms. Starting with the UI part, the application manufacturer first needs to specify starting year, ending year and number of cars to be manufactured in each year and pass a request for processing which will be sent to the backend where the result will be calculated through machine learning algorithms. For that firstly, we collect previous years raw materials data from various sources and convert it into machine understandable format.

Initially, when the data was gathered from various sites, the data was in an unstructured manner, so we preprocessed the dataset. After preprocessing the data, the data is ready for model building. There were nine columns in a dataset which are years, no. of cars, steels, plastics, aluminum, iron, rubber, copper, steel. We normalize the dataset so that values of material quantities should be at the common scale. After that, data is split into 75 and 25 percent ratios, 75 percent as training data and 25 percent as testing data. We used this dataset to train algorithms and find the model with best accuracy and least error.

Machine learning algorithms used are Linear Regression, Random Forest, Multi output Linear Regression, KNN Regression. After training the model it was tested on different values and cross validated their result. The calculated result from given input will be sent through the flask web server as response to the front-end where it shows the result in the tabular form and results are visualized using the highcharts library in the form of a bar chart.

ALGORITHMS USED

Linear Regression

It is a supervised machine learning algorithm, used to predict values which are continuous in nature. This algorithm finds the underlying relationship between a dependent variable and an independent variable. The equation of linear regression is:

$$y = b_0 + b_1 x$$

Where y is output, x is a dependent variable and b_0 and b_1 are constants. The aim of this algorithm is to find the best linear line with minimized error.

Multi-output Regression

Multi Output Regression is used to deal with problems that involve predicting two or more numerical values based on given input. Multiple Linear Regression accepts multiple inputs simultaneously and trains models with given data. It can also give multiple values as output.

Random Forest

Random forest is a supervised machine learning algorithm which is mainly used to solve classification and linear regression problems. Random Forest also handles the missing value and maintains accuracy. In these algorithms models create different trees while running. Random forest builds decision trees on different samples and take their majority vote in case of regression.

KNN Regression

KNN regression is a non-parametric method that, in an intuitive manner, approximates the association between independent variables and the continuous outcome by averaging the observations in the same neighborhood. The general concept of KNN for regression is the same as for classification: first, we find the K-nearest neighbors in the dataset; second, we make a prediction based on the labels of the k nearest neighbors. However, in regression, the target function is a real- instead of discrete-valued function.

RESULTS AND DISCUSSION

we compared the trained ML models, we are using RMSE and Rsquare. The Rsquare finds the accuracy of the model. Higher the accuracy, the better the model fits on the dataset and RMSE is used to find the difference between actual values in the dataset and predicted values.

Following are results of our tests,

Algorithms	Root mean square error (RMSE)	Accuracy
Linear Regression	0.164	96.57
multi output Regression	0.168	91
Random forest	0.154	97.14
Knn	0.292	87.42

RMSE and R-Square value of different algorithms

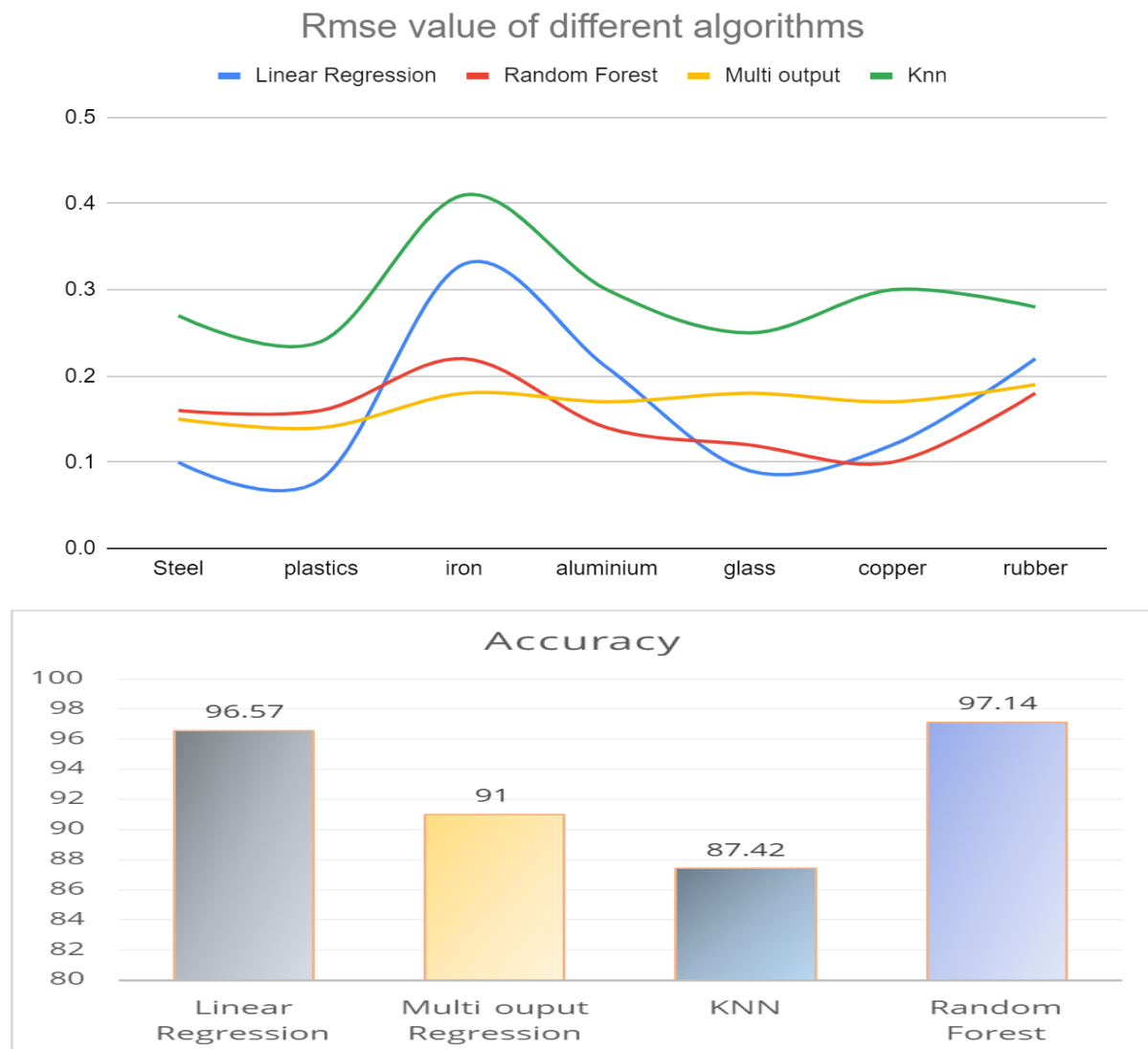


Fig. Accuracy value of different algorithms

After performing the experiments with four different algorithms, in terms of error, all other algorithms show lesser error compared to knn. From the observation, in terms of accuracy, random forest shows the best accuracy and least error, but linear regression is not far behind. Linear regression and random forest shows accuracy greater than 95%.

CONCLUSION

In this system the dataset has been tested on four widely used machine learning methods in regression and their comparative study is presented using standard performance metrics. From the experimental results we conclude that Random Forest Regression outperforms the other algorithm in terms of accuracy and error. These algorithms can generate better results if used in place of Traditional Forecasting Methods for Raw Material Demand Forecasting. This can help manufacturers in production planning to a greater extent.

REFERENCES

- [1] Vignesh Manoj Varier, Dhruv Sharma , Srinath Karmungi, "Impact of Machine Learning on Manufacturing Industries" October–2017 International Journal of Innovative Science and Research Technology , Volume 2, Issue 10, <https://www.scribd.com/document/364059280/Impact-of-Machine-Learning-on-Manufacturing-Industries>
- [2] T Hovorun, K.V. Berladir, V.Pererva, S.G.Rudenko, A.Martynov, "Modern materials for automotive industry ", December 2017, Volume 4; Issue 2, Journal of Engineering Sciences. doi:[10.21272/JES.2017.4\(2\).F8](https://doi.org/10.21272/JES.2017.4(2).F8)
- [3] Zeynep Hilal Kilimci, A.Okay Akyuz,Mitat Uysal,Selim Akyokus, "An Improved Demand Forecasting Model Using Deep Learning Approach and Proposed Decision Integration Strategy for Supply Chain", February 2019 ,Hindawi Complexity Volume 2019, Article ID 9067367, <https://doi.org/10.1155/2019/9067367>

[4] Adnan Mukattash and Murad Samhour, 2011. "Supply Planning Improvement: A Causal Forecasting Approach". Journal of Applied Sciences, 11: 2207-2213.

DOI: [10.3923/jas.2011.2207.2213](https://doi.org/10.3923/jas.2011.2207.2213)

[5] H B Novitasari, N Hadiano, W Gata, J Miharja, "K-nearest neighbor analysis to predict the accuracy of product delivery using administration of raw material model in the cosmetic industry (PT Cedefindo)", et al 2019 J. Phys.: Conf. Ser. 1367 012008, doi:[10.1088/1742-6596/1367/1/012008](https://doi.org/10.1088/1742-6596/1367/1/012008)

[6] S Brunda, Nimish L, Chiranthan S, Arbaaz Khan, "Crop Price prediction using Random Forest and Decision Tree Regression", Sep 2020, International Research Journal of Engineering and Technology, Volume: 07 Issue: 09, <https://mail.irjet.net/archives/V7/i9/IRJET-V7I938.pdf>

[7] Rožanec, J.M.; Kažič, B.; Škerjanc, M.; Fortuna, B.; Mladenčić, D. "Automotive OEM Demand Forecasting: A Comparative Study of Forecasting Algorithms and Strategies". Appl. Sci. 2021, 11, 6787. <https://doi.org/10.3390/app11156787>

[8] Salam Imam Taifur, Tukhas Shilul Imaroh, "Forecasting Planning And Procurement Strategy Of Raw Material Using Material Requirements Planning Method", June 2020, Dinasti International Journal of Digital Business Management Volume 1, Issue 4, dinastipub.org > DIJDBM > article

[9] Alexei Botchkarev, "Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology", Interdisciplinary Journal of Information, Knowledge, and Management, 2019, 14, 45-79. <https://doi.org/10.48550/arXiv.1809.03006>

[10] Tin Kam Ho, Jiangang Hao, "Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language", February 2019, Journal of Educational and Behavioural Statistics Vol. XX, No. X, pp. 1-14, <https://doi.org/10.3102/1076998619832248>

