iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: https://doi.org/10.22214/ijraset.2023.57210

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

Material Property Prediction Using Machine Learning

Parag Prasad Kshirsagar[1], Shruti Sunil Kulat[2], Bhargav Atul Kulkarni[3], Vedant Kulkarni[4], Sanket More[5]
Department of Mechanical Engineering, Vishwakarma Institute Of Technology, Pune, India

Abstract: Material research is an ever-growing field with the invention of new materials almost every day. It is not feasible for everybody working in the field to test each and every material on the Universal Testing Machine (UTM). This poses a limitation that has been sought to be tackled for quite some time now. One approach to tackle this is to use machine learning to predict the properties of the materials, thus eliminating the need of extensive time requirement and cost of performing experiments on the UTM. In this paper, the authors have trained and tested a machine learning based model that can predict material properties on the basis of composition and given temperature.

Keywords: Material Properties, Prediction, Random Forest, Low Alloy Steel, Tensile Strength, Linear Regression

I. INTRODUCTION

This manuscript explores the application of machine learning in material property prediction. Traditional strategies like trial and error, experiences of domain experts consume a lot of time and cost. In order to overcome these drawbacks, we propose a machine learning design system which involves three features of Machine Learning which are Modelling, Compositional Design and property prediction of the materials which can help us discover new alloys based on low alloy steels and carbon steels. We investigated 13 elements, using temperature as the 14th, to predict the tensile strength of the materials. We selected low alloy steels, which are used in industry because they are affordable, readily available, highly durable, and machinable. Low alloy steels have precise compositions and offer better mechanical qualities than many common carbon steels due to the addition of some specific alloys.

II. LITERATURE SURVEY

A. Property Oriented Design Strategy for high Performance Copper Alloys.

In this study, machine learning has been utilised to design new content in place of more conventional methods like trial and error and the expertise of subject matter experts. The system is intended to develop high performance copper alloys with a targeted Ultimate tensile strength of 600-950 MPa and an electrical conductivity of 50% international annealed copper standard. This study only focuses on copper alloys, utilising copper as the basic material. They were remarkably successful, with only a 10% difference between the actual and projected results.

B. Machine Learning Elastic Constants of multi-component Alloys

This study investigates the use of machine learning techniques for working out the elastic constants and other mechanical properties generated from multi-component alloys. On a dataset of binary alloys produced using density functional theory (DFT) calculations and spanning over a large number of elemental species in the periodic table, a variety of machine learning models, including linear regression, neural network, and random forest-based models, are trained and tested. A correlation-based feature selection strategy was used to systematically down-select a set of most pertinent features towards the prediction of the elasticity tensor components from a large variety of simple and easily accessible compositionally-averaged elemental features. Through testing on hypothetical data and bootstrapping, the models' actual predictive performance and related uncertainties were determined.

C. Machine-learning-assisted prediction of the Mechanical Properties of Cu–Al alloy

To speed up the production of new materials, describe them more quickly, and get physical insights into their properties, the mechanical properties of Cu-Al alloys synthesized using the powder metallurgy technique were predicted using a machine-learning approach.

The prediction models were developed using six algorithms, and the compacts' chemical composition and porosity were chosen as the descriptors.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

III. METHODOLOGY

A. DataSet

This study requires Data gathering which was necessary for this project. A data set was created by collecting data from numerous websites. It was challenging to collect all the data because several factors were considered. The data for this study were gathered using a Japanese website as a reference. Following data collection, a data set was constructed that contained all 13 elements in addition to temperature as the fourteenth factor.

The collected data was not organized in a way that we could use it or that the code could simply comprehend. In order to make the data helpful to the model, it was organized in that manner. A spreadsheet was made, and specific parameters were written under specific elements. The order of the data was established such that the code could easily interpret it.

B. Model Training

The model was trained after data collection and preparation. The model in this study was trained using a variety of techniques. Python is primarily used for the machine learning portion. To train the data, three packages and libraries—Pandas, NumPy, and Matplotlib—were utilized. The supervised machine learning algorithms Random Forest Regressor and Linear Regression were used to estimate the values of tensile stress based on temperature values and element composition. Both algorithms were applied during this study.

C. Linear Regression

Multiple linear regression refers to a statistical technique that is used to predict the outcome of a variable based on the value of two or more variables. It is sometimes known simply as multiple regression, and it is an extension of linear regression. The variable that we want to predict is known as the dependent variable, while the variables we predict is known as independent or explanatory variables.

$$yi = \beta0 + \beta1xi1 + \beta2xi2 + \ldots + \beta pxip + \epsilon$$

Where:

- I is number of observations
- yi is the dependent or predicted variable
- β0 is the y-intercept, i.e., the value of y when both xi and x2 are 0.
- β1 and β2 are the regression coefficients representing the change in y relative to a one-unit change in xi1 and xi2, respectively.
- βp is the slope coefficient for each independent variable
- ϵ is the model's random error term.

So, we trained our model on this algorithm as it takes multiple inputs and predicts the required result. This cannot be done manually as you need a to compare a lot of data to get a semi-accurate result. Using an AI model also gives us an output on the mean deviation and highest deviation.

D. Random Forest

Random Forest is a classifier, it contains several decision trees on various subsets of the given dataset and using this the average id taken to improve the predictive accuracy of that dataset. It is based on the concept of ensemble learning which is a process of combining multiple classifiers to solve a complex problem and improve the performance of the model.

Random Forest algorithm works with selection of random Samples from given data or training data.it will construct a decision tree for every training data. After averaging the values from data, voting is done. Finally, selection of most voted prediction result is decided as final prediction result.

So, we trained our model using this algorithm.

E. Prediction

The trained model was then used to try and predict the values for a known set of values. Some entries had been kept out of the dataset for testing purposes. These values were then given to the model and the results acquired were compared to the actual values.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV. RESULTS

The model was trained on the basis of 75% of the data and the remaining 25% was used to test the accuracy and prediction power of the project. Two datasets were used for the process and two algorithms were applied to each of the models. Results of both the algorithms were compared. Three parameters were decided upon to judge the output of the models. The parameters being, Root Mean Square Error, Mean Absolute Error and Maximum encountered error. Plots of the results were also made to judge the results better.

A. Carbon Steel Dataset

This was the smaller dataset that we first worked on to test the grounds. Around 400 entries were present in the dataset.

1) Linear Regression

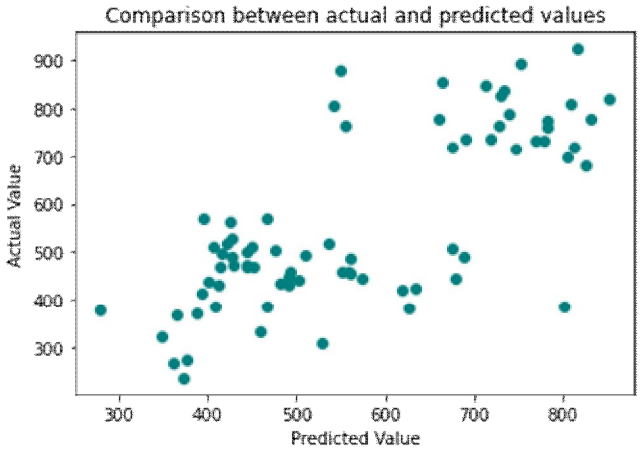
This is one of the most basic algorithms out there. This was used first as it is easy to use and is many times considered fit for such applications.

Root Mean Square Error: 123.54

Maximum Error: 415.63

Mean Absolute Error: 94.52

The maximum encountered error was 415 MPa. The Mean absolute error value was quite high.



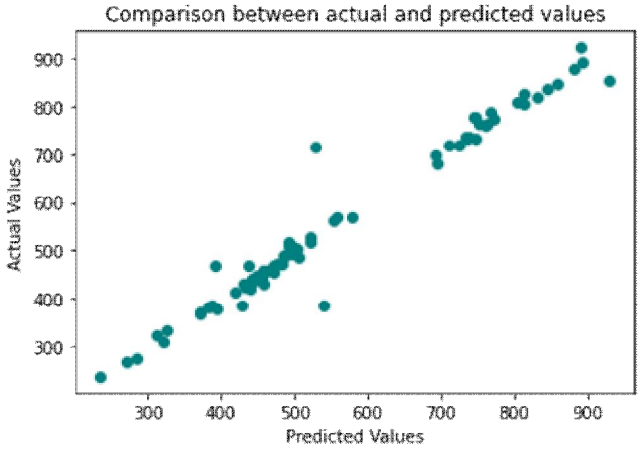
2) Random Forest

Random forest is a popular algorithm known for it's high accuracy scores. It works quite differently compared to LR.

Root Mean Square Error: 33.34

Maximum Error: 188.08

Mean absolute Error: 15.42

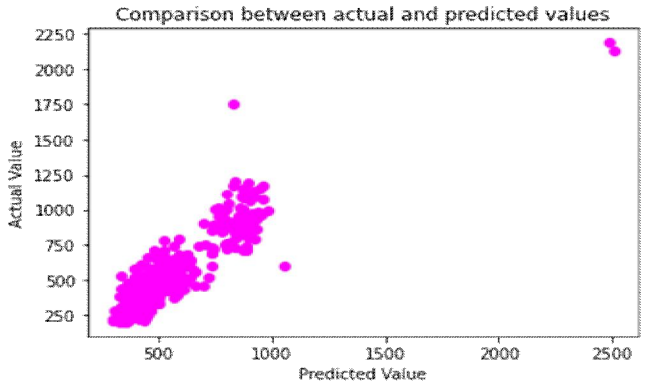


B. Low Alloy Steels

This was the bigger and primary dataset used for the project. This included data of 13 elements and temperature values. Two algorithms were applied just as was done for Carbon steels. Following are the results

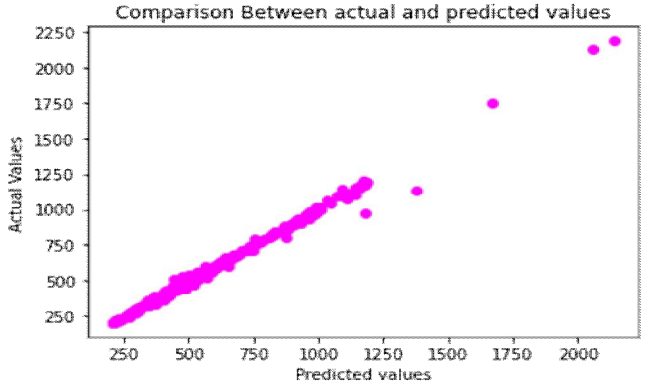
1) Linear Regression

Root Mean Square Error: 107.2
Maximum Error: 916.6
Mean absolute Error: 78.11



2) Random Forest

Root Mean Square Error: 18.2
Maximum Error: 242.55
Mean absolute Error: 7.69



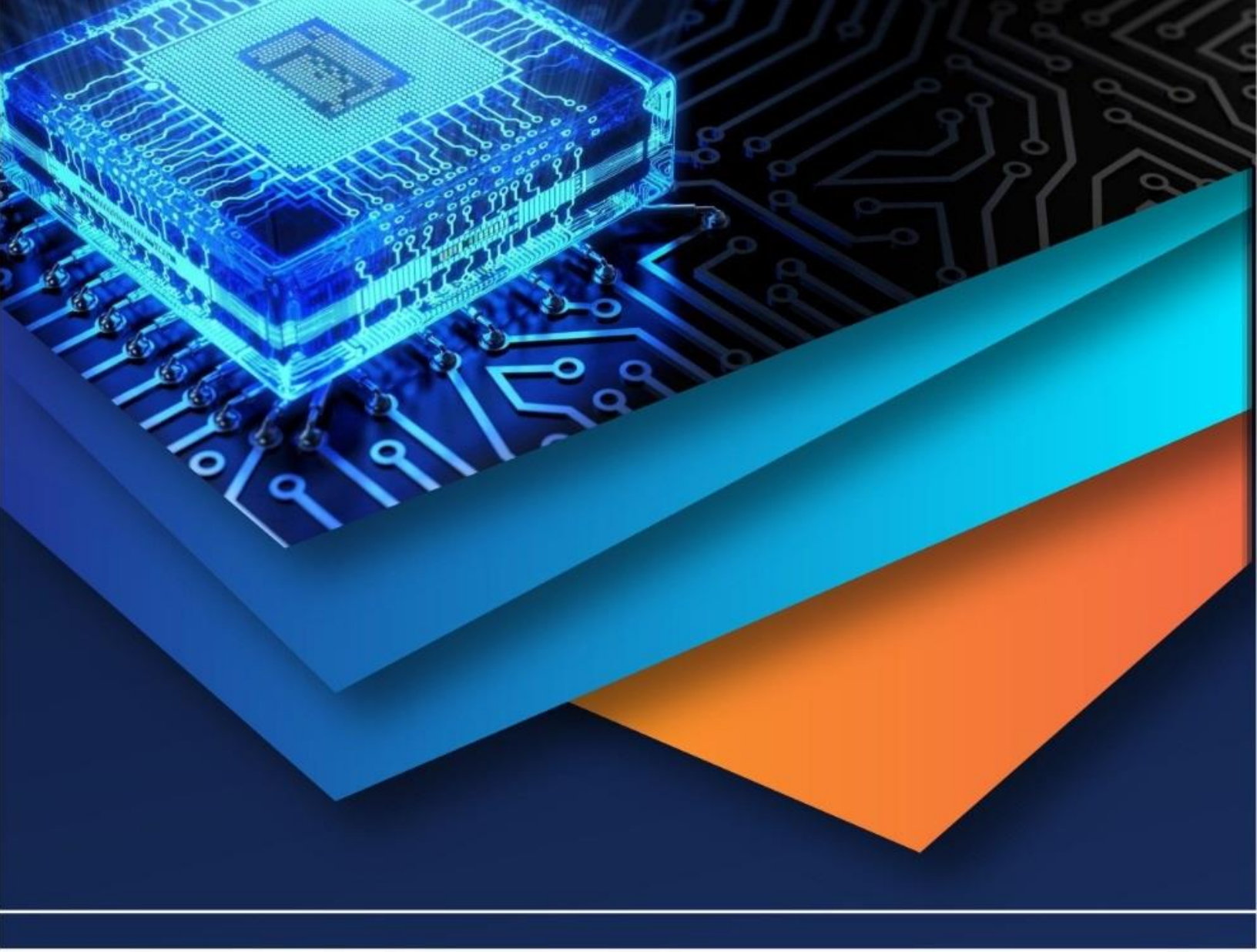
V. CONCLUSION

A machine learning model was successfully designed to predict the tensile strength of any given metal. The results of Random Forest algorithm were far superior when compared to the results of linear regression. Linear regression is quite a basic algorithm whereas Random Forest algorithm is quite advanced. Results of the RF algorithm showed around 10 times as much as improvement over the linear regression algorithm for both the datasets. The average error in the values given by both the models (for both datasets) was 7-15Mpa. Thus, the model has a high accuracy rate.

REFERENCES

[1] Wang, Juan, et al. "New methods for prediction of elastic constants based on density functional theory combined with machine learning." Computational Materials Science 138 (2017): 135-148.
[2] Revi, Vivek, et al. "Machine learning elastic constants of multi-component alloys." Computational Materials Science 198 (2021): 110671.
[3] Deng, Zheng-hua, et al. "Machine-learning-assisted prediction of the mechanical properties of Cu-Al alloy." International Journal of Minerals, Metallurgy and Materials 27.3 (2020): 362-373.
[4] National Institute for Materials Science Japan, Website

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)