Proceedings of the International Conference on Smart Electronics and Communication (ICOSEC 2020)
IEEE Xplore Part Number: CFP20V90-ART; ISBN: 978-1-7281-5461-9

# *An Informatic Approach to Predict the Mechanical Properties of Aluminum Alloys using Machine Learning Techniques*

M Aruna Devi
*Department of Mechanical Engineering dayananda sagar college of Engineering*
Bangalore, India
arunadevi.dsce@gmail.com

C P S Prakash
*Department of Mechanical Engineering dayananda sagar college of Engineering*
Bangalore, India
drcpsprakash@gmail.com

Rahul Pandappa Chinnannavar
*Department of Mechanical Engineering dayananda sagar college of Engineering*
Bangalore, India
rahulchinnannavar@gmail.com

Venugopal Prasanna Joshi
*Department of Mechanical Engineering dayananda sagar college of Engineering*
Bangalore, India
venugopal.joshi@yahoo.com

Rohit Shankar Palada
*Department of Mechanical Engineering dayananda sagar college of Engineering*
Bangalore, India
paladarohit@gmail.com

Ravut Dixit
*Department of Mechanical Engineering dayananda sagar college of Engineering*
Bangalore, India
mailravut@gmail.com

***Abstract*— One of the major problems faced by the industries during the manufacturing of aluminum components is by achieving the required properties of aluminum alloys. Lot of time and energy is involved in the experimentation and testing of properties of new aluminum alloys. It leads to wastage of many resources and even sometimes ends up with no results. This paper presents an algorithm for prediction of mechanical properties of aluminum alloys using different machine learning techniques such as linear regression (LR), artificial neural network (ANN), and k-nearest neighbor (KNN) algorithm. In this paper, KNN algorithm gives a better prediction of tensile strength and hardness values. In yield strength predictions, ANN gives the better and accurate results compared to other two algorithms. More amount of energy and time is saved using machine learning techniques.**

***Keywords*— *Aluminum alloys, neural network, machine learning, linear regression, K-nearest algorithm, tensile strength.***

## I. INTRODUCTION

In recent years, machine learning entered in the field of material science for the discovery of new material. Discovery of new material needs a lot of experimentation, testing and it is also remaining as a time consuming process along with uncertainty [1]. Traditional procedure of finding a new material like trial &error method and DFT (Density Functional Theory) - based methods have long development cycles, higher cost and lower efficiency. Due to its low computational cost and better prediction performance, machine learning techniques plays important role in discovery of new material, material analysis and material design [2]. In [3], Lasso regression is used for classification and prediction of experimental tests on ductile or non-ductile joint shear failure prior to or after beam yielding. In [4], an informatics approach to transformation temperatures of NiTi based shape memory alloys is explained. A quantitative statistical learning model is built to predict and estimate the targeted property of set of unexplored or possible materials. In [5], single layer and multi-layer feed forward back propagation of an artificial neural network is explained to predict the mechanical properties of hydrogen charged metallic materials. In [6], the predictions of cetane number for furanic biofuel additives are obtained using an artificial neural network with high accuracy. In [7], the efficient ANN model is developed to predict the weight loss values after the wear test to find of the effect of aging parameters on wear behavior of PM Inconel 706. In this paper[8], different machine learning techniques such as linear and quadratic interpolation, neural network, Gaussian processes (GP),dynamic trees and the scalable Gaussian process regression (SGP) are used for the prediction of physical properties of various compositions of steel. In [9], to find the influence of alloying elements on mechanical properties of steels, the artificial neural network model was used. The ANN used in many fields of materials science [10,11,12,14,15,16,17], because of its self-learning ability, enough storage function, and optimal solutions at a high speed. In this paper [13], the support vector machine is proposed to achieve improved prediction performance of lattice constant of structurally known perovskites and results are compared with ANN model.

In aluminum alloys, aluminum is the predominant one; others are manganese, copper, tin, zinc, magnesium, etc. Aluminum alloys are light weight material and used as corrosion resistant in engineering structures and components. Amount of silicon in alloy gives good casting properties. Cast aluminum alloys are less in cost because of its low melting point and having low tensile strength. Increasing magnesium in aluminum alloy leads to less flammable property and less weight, which is suitable for aerospace manufacturing. So, in this study different machine learning techniques are proposed to predict the tensile strength, yield strength and hardness using composition of aluminum alloys which is shown in figure.1.

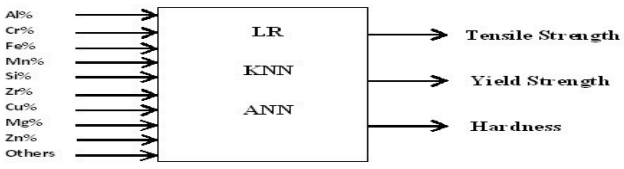


Fig. 1. Prediction of mechanical properties of Aluminum alloy

TABLE I.

ALUMINUM ALLOYS AND ITS MECHANICAL PROPERTIES

| Sl.No | Al% | Cr% | Fe% | Mn% | Si% | Zr% | Cu% | Mg% | Zn% | others | Tensile Strength | Yield Strength | Hardness |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 11 | 0 | 4.7 | 0.3 | 0 | 0 | 79.1 | 0 | 0 | 0 | 850 | 185 | -70 |
| 2 | 96.3 | 0 | 0.6 | 0 | 0.6 | 0.1 | 0 | 1.5 | 0 | 0 | 135 | 60 | -44 |
| 3 | 95 | 0 | 0.7 | 1.2 | 2.2 | 0 | 0.2 | 0.5 | 0.2 | 0 | 200 | 170 | -44 |
| . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| 128 | 86.05 | 0 | 0.65 | 0.55 | 11 | 0 | 0.35 | 0.45 | 0.35 | 0.2 | 180 | 380 | -80 |
| 129 | 90.5 | 0 | 0.55 | 0.45 | 1.5 | 0 | 0.05 | 6.5 | 0.1 | 0.2 | 219 | 250 | -65 |
| 130 | 87.8 | 0.1 | 0.8 | 1 | 0.8 | 0 | 4.6 | 1.8 | 0.8 | 0.2 | 370 | 250 | -100 |

## II. PROBLEM DESCRIPTION

One of the major problems faced by the industries during manufacturing of aluminum components is achieving the required properties of aluminum alloys. Lot of time and energy is involved in the experimentation and testing of properties of new aluminum alloys. It leads to wastage of many resources and even sometimes ends up with no results. when large amount of data is available with more number of variables and no mathematical equation is available to derive the predictions, machine learning is the only solution to solve the problem.

### A. Objective

To overcome the above issue, the objectives of this study is

- To predict the properties of aluminum alloy composites using machine learning algorithm.
- To reduce the experimental cost using machine learning algorithm.
- To save the time of experimentation using machine learning algorithm.

## III. METHODOLOGY

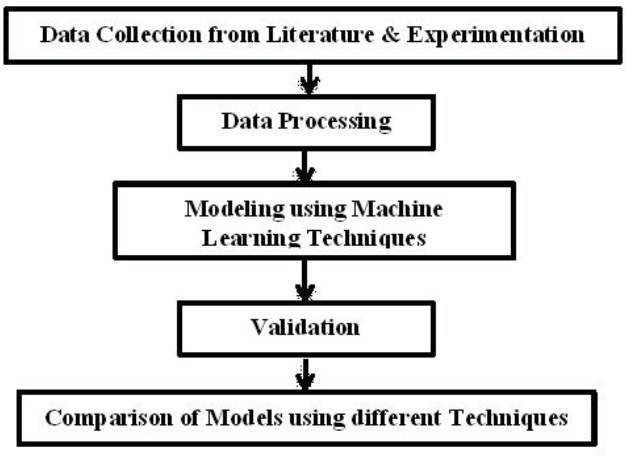


Fig. 2. Methodology

The methodology of this study is explained using flowchart which is shown in figure.2. First step is data collection from different literature and experimentation. Then the machine learning algorithm involves the following

- Step1: Data Processing
- Step2: Modeling
- Step3: Validation

### A. Data Processing

Application of data processing in material science is the important step of machine learning, because the performance of machine learning model mainly depends on data processing.

Data processing consists of two parts:

- Data selection
- Feature engineering.

*1) Data Selection:* In this study, A 130 datasets were collected from matmatch[18], which includes different composition of aluminum alloys, its tensile strength , yield strength and hardness values of Aluminum composites is shown in table 1.

*2) Feature Engineering:* Feature engineering means extracting the suitable characteristics for prediction of target. In this study different composition of aluminum alloys are considered as features for the prediction of Tensile strength, Yield strength and Hardness value.

### B. Modeling

After data processing, a model is built to analyze the data. The steps involved in the modeling of algorithm are selection of algorithms, training and making predictions. Supervised learning models such as LR, KNN and ANN are selected for modeling, because data collected are the labeled one. As our output is continuous, the regressor of all above algorithms is selected to build model, train the model and predictions are made using python.

#### 1) Linear regression

Linear regression builds a model which explains the relationship between independent variable and dependent variable by fitting a linear equation for the observed data.

2) *K-Nearest Neighbors (KNN)*

It is a supervised learning which can be used to solve both classification and regression problems. In this algorithm, the data points similar to the testing data point are identified as neighbors and output is derived. The predictions are made by observing the k neighbors. The steps to be followed in this algorithm is

Step 1: Loading the data and initialize the K value based on your data size

Step 2: Calculating the distance between test data and the current data for each example.

Step 3: Sorting the ordered collection of distances and indices in ascending order.

Step 4: Picking the first K entry and get the labels

Step 5: If regression, return the mean of the K labels or mode of the K labels.

*3) Artificial Neural Network*

In neural network [6], neurons are combined according to different architectures. Basically neurons are arranged in a layers consist of input layer, hidden layer and output layer. Each neuron receives input signals and connected with weightage which sum to an activation function. A suitable activation function transforms it into the output. The accuracy of network depends only on connected weightage values.

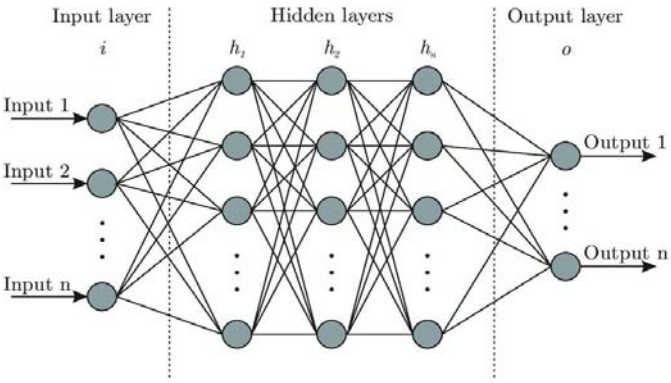


Fig. 3. Artificial Neural Network

## C. Validation of model

Validation of a model is conducted to assess performance and the accuracy of developed model. In every machine learning technique, the original data is divided in to a training set and test set. The training set of data is used for training of model and the validation is conducted by test set of data.

## IV. RESULTS AND DISCUSSIONS

In this work, the percentage of different metal in the aluminum alloy is considered as the input and the targets are tensile strength, yield strength and hardness. LR, KNN and ANN algorithms are used for training the network. The above work is performed in Python3.7.

RMSE (Root Mean Square Error) plots are used for analysis in this study. Residuals are a measure of how the data points are far from regression line. RMSE is a measure of how residuals are spreaded and how the data is concentrated around the line of best fit. The formula to find the RMSE value is:

$$RMSE = \sqrt{1/n \sum_{i=1}^{n} (f_i - o_i)^2}$$

$f_i$ - Predicted value

$o_i$ -Actual value

$n$ - Sample size

R- Squared value is used to predict the model accuracy.

R-squared = Explained variation / Total variation

It ranges from zero to one. If it is zero, the model poor prediction and one indicates perfect prediction.

## A. Prediction of Tensile Strength

All three techniques are performed well by maintaining an average R squared value above 85%. However, the KNN algorithm has lowest RMSE value 63.84 and highest R squared value 0.947 by comparing with other two algorithms which is shown in Table 2. RMSE values for the different percentage of training dataset (50% to 90%) are plotted in Figure 4, which shows the KNN algorithm gave the lowest RMSE value for all the different percentage of training datasets.

TABLE II.

RMSE AND $R^2$ VALUE FOR TENSILE STRENGTH

| Technique | RMSE | R Squared value |
|---|---|---|
| LR | 79.69 | 0.914 |
| KNN | 63.84 | .947 |
| ANN | 94.96 | 0.88 |

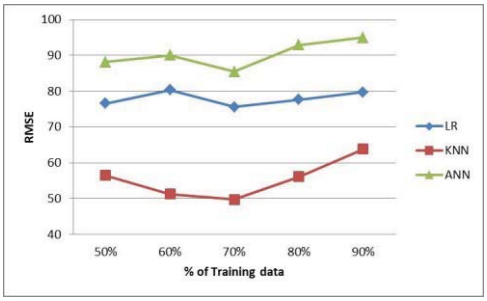


Fig. 4. Comparison of RMSE for Tensile Strength Predictions

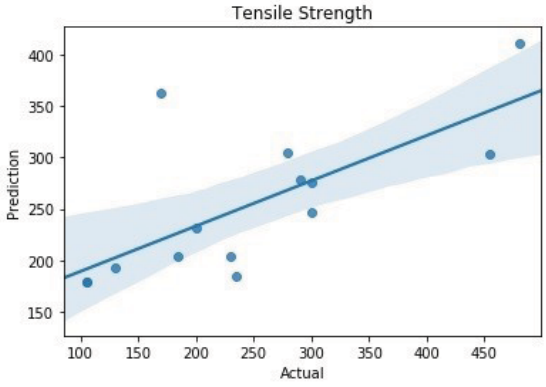


Fig. 5. Actual vs Predicted Tensile Strength using Linear Regression

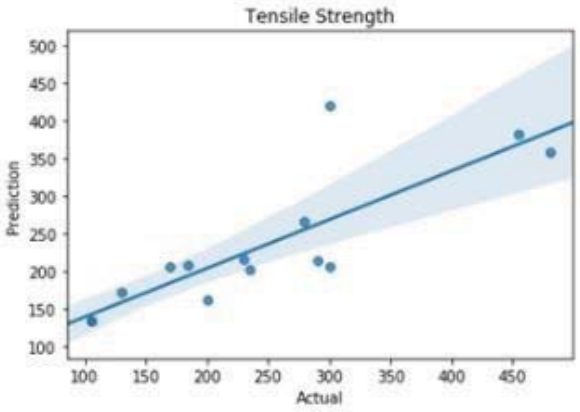


Fig. 6. Actual vs Predicted Tensile Strength using KNN

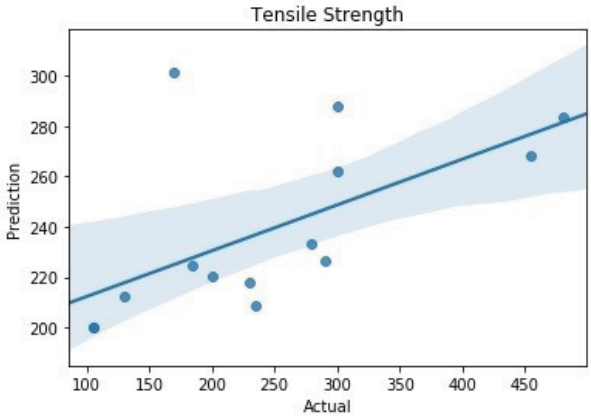


Fig. 7. Actual vs Predicted Tensile Strength using ANN

Actual values versus Prediction plots for tensile strength for three algorithms (LR, KNN &ANN) are shown in Figure 5-7. In Figure 5, around six data points are out of the fitting region and its R square value is 0.914 which is obtained using Linear Regression. In Figure 6, four data points are out of best fit and R square value is 0.947 obtained using KNN. In Figure 7, five data points are out of the fitting region defined by ANN model. From the above graph, it is observed that KNN plot has the better fit compared to other two algorithms (LR& ANN). For tensile strength, it clearly shows the KNN algorithm has better fit of points with regression line and only one point out of 13 points does not fit with rest of data points and far from regression line which is called as outlier. So, in prediction of tensile strength of aluminum alloys, KNN algorithm gives better and accurate predictions compared to other two algorithms.

### *B. Prediction of Yield Strength*

The ANN algorithm has lowest RMSE value 72.7 and highest R squared value 0.92 by comparing with other two algorithms which is shown in Table 3. RMSE values for the different percentage of training dataset (50% to 90%) are plotted in Figure 8, which shows the ANN algorithm gave the lowest RMSE value for all the percentage of training datasets.

TABLE III.

RMSE AND $R^2$ VALUE FOR YIELD STRENGTH

| Technique | RMSE | R Squared value |
|---|---|---|
| | | |
| LR | 350 | 0.71 |
| KNN | 135 | 0.723 |
| ANN | 72.7 | 0.92 |

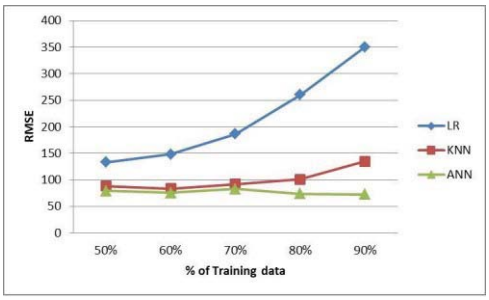


Fig. 8. Comparison of RMSE for Yield Strength Predictions

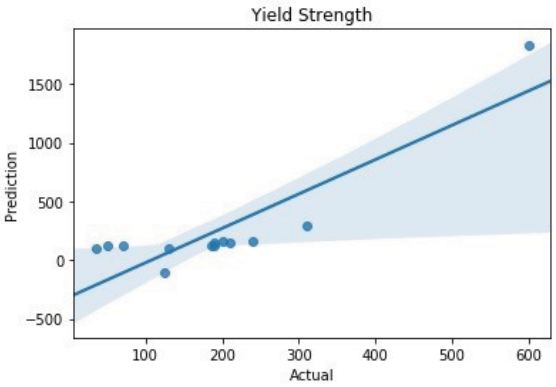


Fig. 9. Actual vs Predicted Yield Strength using Linear Regression

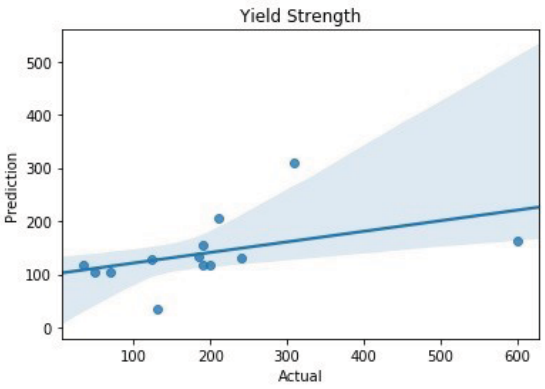


Fig. 10. Actual vs Predicted Yield Strength using KNN

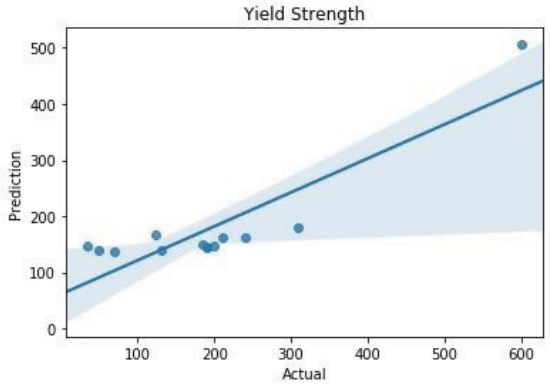


Fig. 11. Actual vs Predicted Yield Strength using ANN

Actual values versus Prediction for yield strength for three algorithms (LR, KNN &ANN) are shown in Figure 9-11. In Figure 9, from the graph, it is observed that ANN residuals have the better fit compared to other two algorithms (LR& KNN). For yield strength, it clearly shows the ANN algorithm has better fit of points with regression line. So, in prediction of yield strength of aluminum alloys, ANN algorithm gives better and accurate predictions compared to other two algorithms.

### *C. Prediction of Hardness*

All three techniques are performed well by maintaining an average R squared value above 90%. However, the KNN algorithm has lowest RMSE value 15.86 and highest R squared value 0.96 by comparing with other two algorithms which is shown in Table 4. RMSE values for the different percentage of training dataset (50% to 90%) are plotted in Figure 12, which shows the KNN algorithm gave the lowest RMSE value for all the different percentage of training datasets.

TABLE IV.

RMSE AND $R^2$ VALUE FOR HARDNESS

| Technique | RMSE | R Squared value |
|---|---|---|
| LR | 18.81 | 0.958 |
| KNN | 15.86 | 0.96 |
| ANN | 21.5 | 0.935 |

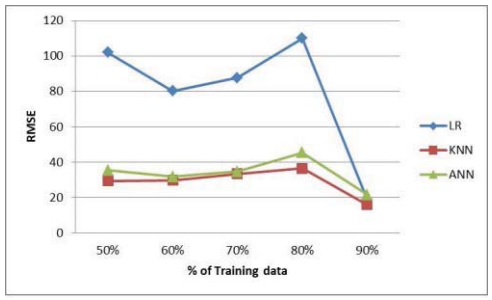


Fig. 12. Comparison of RMSE for Hardness Predictions

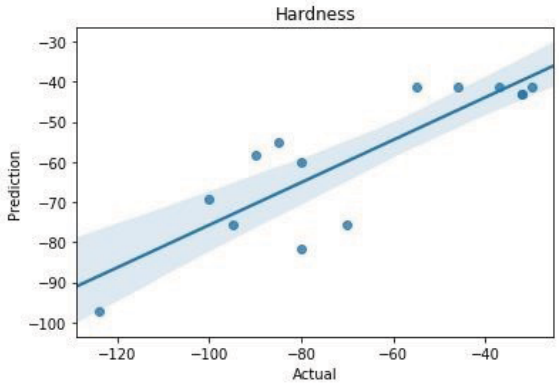


Fig. 13. Actual vs Predicted Hardness using Linear Regression

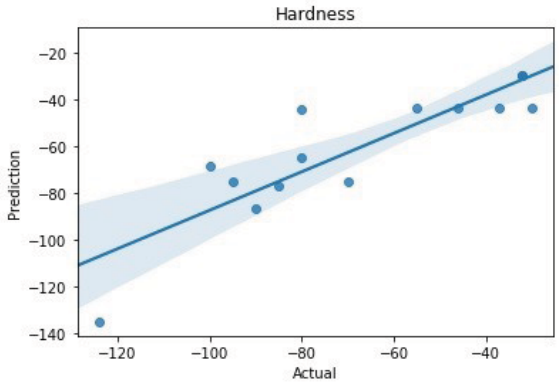


Fig. 14. Actual vs Predicted Hardness using KNN

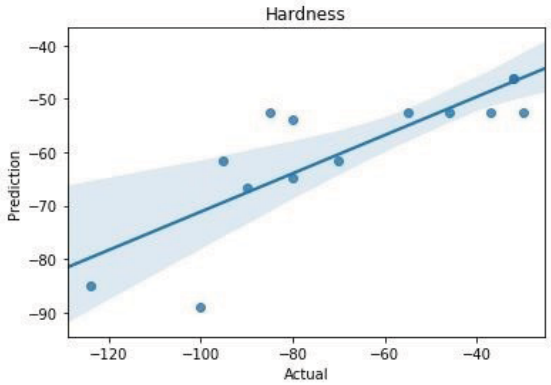


Fig. 15. Actual vs Predicted Hardness using ANN

Actual values versus Prediction plots for the hardness for three algorithms (LR, KNN &ANN) are shown in Figure 13-15.In Figure 13, around five data points are out of the fitting region and its R square value is 0.958 which is obtained using Linear Regression. In Figure 14, three data points are out of best fit but near to the region and R square value is 0.96 obtained using KNN. In Figure 15, five data points are out of the fitting region defined by ANN model. From the above graph, it is observed that KNN plot have the better fit compared to other two algorithms (LR& ANN). For hardness, it clearly shows the KNN algorithm has better fit of points with regression line and only 3 points out of 13 points does not fit with rest of data points and far from regression line which is called as outlier. So, in prediction of t hardness of aluminum alloys, KNN algorithm gives better and accurate predictions compared to other two algorithms.

TABLE V.

PERFORMANCE COMPARISON METRICS TABLE

| Performance metric | Tensile Strength | | | Yield Strength | | | Hardness | | |
|---|---|---|---|---|---|---|---|---|---|
| | LR | KNN | ANN | LR | KNN | ANN | LR | KNN | ANN |
| R-square | 91.4 | 94.7 | 88 | 71 | 72.3 | 92 | 95.8 | 96 | 93.5 |
| RMSE | 79.69 | 63.84 | 94.96 | 350 | 135 | 72.7 | 18.81 | 18.81 | 18.81 |

Performance comparison metrics of all the parameters are tabulated in Table 5.

## V. CONCLUSION

In this paper, three machine learning algorithms such as LR, KNN and ANN used to predict the mechanical properties of aluminum alloys. The training and testing data set for the machine learning algorithms is obtained from the mat match. KNN algorithm gives better prediction of tensile strength and hardness values. In yield strength predictions, ANN gives the better and accurate results compared to other two algorithms. Good results obtained using Linear regression is limited only to linear relationship and sensitive to outliers. KNN algorithm does not give good results with high dimensional data, sensitive to noises in data, missing values and outliers. But ANN gives better and accurate results even for nonlinear relationships, while predicting the response in ANN model, unseen relationship on missed data will be considered. Future research will be, the same methodology can be used for discovery of new materials in other composites by using different machine learning techniques. Also machine learning can be applied to other areas of material science

## REFERENCES

[1] Rajiv Kumar , Amrita , Pradeep Kumar Mishra, "An intelligent computing system to detect material", Materials Today proceedings(2020)

[2] Jing Wei, Xuan Chu, Xiang-Yu Sun, Kun Xu1, Hui-Xiong Deng, Jigen Chen, Zhongming Wei, Ming Lei, "Machine learning in materials science", Wiley journal (2019) InfoMat. 2019;1:338–358.

[3] Sujith Mangalathua, Jong-Su Jeonb, "Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques" Engineering Structures 160 (2018) 85–94.

[4] Dezhen Xue , Deqing Xue , Ruihao Yuan, Yumei Zhou, Prasanna V. Balachandran ,Xiangdong Ding, Jun Sun, Turab Lookman, "An informatics approach to transformation temperatures of NiTibased shape memory alloys", Acta Materialia 125 (2017) 532-541.

[5] Titus Thankachan, K. Soorya Prakash, Christopher David Pleass, Devaraj Rammasamy, Balasubramanian Prabakaran, Sathiskumar Jothi, "Artificial neural network to predict

[6] the degraded mechanical properties of metallic materials due to the presence of hydrogen", International Journal of hydrogen energy(2017) pages 1-10

[7] Travis Kessler, Eric R. Sacia, Alexis T. Bell, J. Hunter Mack, "Artificial neural network based predictions of cetane number for furanic biofuel additives", Fuel 206 (2017) 171–179

[8] Osman Palavar, Dursun Ozyurek, Ali Kalyon, "Artificial neural network prediction of aging effects on the wear behavior of IN706 super alloy", Materials & Design 82(2015) 164-172

[9] Eve Bélisle , Zi Huang , Sébastien Le Digabel , Aïmen E. Gheribi , "Evaluation of machine learning interpolation techniques for prediction of physical properties", Computational Materials Science 98 (2015) 170–177

[10] S. Krajewski,J.Nowackin , "Dual-phase steels microstructure and properties consideration based on artificial intelligence techniques",Archives of civil and mechanical engineering 14(2014)276-286

[11] Ravindranadh bobbili, b. Ramakrishna, v. Madhu, a.k. Gogia, "Prediction of flow stress of 7017 aluminium alloy under high strain rate compression at elevated temperatures", Defence Technology (2014) 1-6

[12] Mohsen Ostad Shabani, Ali Mazahery, "Optimization of process conditions in casting aluminum matrix composites via interconnection of artificial neurons and progressive solutions" Ceramics International 38 (2012) 4541–4547

[13] Y.C. Lin a, Jun Zhang, Jue Zhong, "Application of neural networks to predict the elevated temperature flow behavior of a low alloy steel", Computational Materials Science 43 (2008) 752–758

[14] R. Koker, N. Altinkok, A. Demir, "Neural network based prediction of mechanical properties of particulate reinforced metal matrix composites using various training algorithms", Mater. Design 28 (2007) 616–627.

[15] Z. Sterjovski, D. Nolan, K.R. Carpenter, D.P. Dunne, J. Norrish, "Artificial neural networks for modelling the mechanical properties of steels in various applications", J. Mater. Process. Technol. 170(2005) 536–544.

[16] M.O. Shabani, A. Mazahery, "Artificial Intelligence in numerical modeling of nano sized ceramic particulates reinforced metal matrix composites", Appl. Math. Model. (2012),

[17] L. Fratini, G. Buffa, D. Palmeri, "Using a neural network for predicting the average grain size in friction stir welding processes", Comput. Struct. 87 (2009) 1166–1174.

[18] Matmatch.2017.Accessed on: jan 30 , 2020 [online]vailable:https://matmatch.com/advancedsearch?categories=aluminium