Analysis of Performance of Magnesium Alloys using Machine Learning Techniques

COMP 4560 (Advanced Computing Research)

Prateek Arora (u6742441)
Supervisor – Professor Nick Birbilis
Australian National University, Canberra
(Semester 2, 2019 - Semester 1, 2020)
Acknowledgement

The project titled “Analysis of Performance of Magnesium Alloys using Machine Learning Techniques” has been completed under the dynamic and unfailing supervision of respected Professor, Nick Birbilis. I want to extend my humble thanks to him for allowing me to showcase my skills through this project. He provided incessant and immediate motivation from time to time as well as the feedback on the project thereby helping us in cleaning the database as and when required. Without his expert-assistance and immaculate guidance, I would not have been able to finish my project.

At this moment, I would also like to express my gratitude to Professor Zhouran Zheng for providing the research papers based on which most of database was collected. It would have been very difficult and time consuming for me in the absence of the requisite research papers for the database.

I would also like to thank my team consisting of Samyak Jain(u6734495), Bhumipat(u6069393) and Jia Ye(u5879731) for collecting the dataset together as a team.

At this time, my ‘Note of Thanks’ is incomplete without mentioning the name of Professor Liang Zheng for devoting his valuable time to go through the project and awarding the due credit to me whatever I deserve. I deeply appreciate the acumen of his stream and I am thankful to him for stamping the project with the mettle of his name.

Above all I am very thankful to Lord Almighty who gave me the perseverance to go through the bumps and smooth flows that I face and with His grace, I was able to complete this project.
Abstract

Alloys play an important role in our daily lifestyle. So, forming a good alloy is the need of the hour for the industries. It is a well-known fact that the Magnesium is the lightest element in the Periodic Table. Therefore, alloys formed using Magnesium will be much lighter. These alloys can be made in the furnace using different conditions and the temperature, but alloys formation can be made easier using the Machine Learning Techniques. The motive of our project is to form a strong Mg alloy with 350 MPA as the Ultimate Tensile Strength, 300 MPA as the Yield Strength and 20 as the Ductility.

Further in the report, different types of Machine Learning Techniques such as Neural Network, Elastic Net Regression and SVR (Support Vector Regression) are used to build the model respectively on the database provided. Among these three Machine Learning Techniques, the Neural Network Technique outstands with the better R2 score for all the parameters. So, Neural Network is further used as a model and on that basis, objective function, as well as the data generator, is formed. Data generator helps in generating 2.5 million datasets and the objective function will use the solubility table which will help to set the element’s composition range in the alloy and will try to get the results that are very close to the desired output. This project, in the end, is successful in achieving the desired mechanical properties by forming strong Magnesium alloys using the Machine Learning Techniques.

Keywords:
Magnesium alloys, Elastic Net regression, Neural Network, SVR, Objective function, Data generator, Alloy conditions, Solubility, Overfitting
## Contents

Acknowledgement .......................................................................................................................... 2  
Abstract ........................................................................................................................................ 3  
Introduction .................................................................................................................................. 6  
  1.1 Objective ................................................................................................................................. 6  
  1.2 Scope ...................................................................................................................................... 6  
  1.3 Contribution ............................................................................................................................ 7  
  1.4 Report Outline ......................................................................................................................... 7  
Background work ........................................................................................................................... 8  
  2.1 Alloys’ Conditions .................................................................................................................. 8  
    2.1.1 ECAE or ECAP .................................................................................................................. 8  
    2.1.2 Extrusion .......................................................................................................................... 8  
    2.1.3 Sand or Cool Cast .............................................................................................................. 9  
    2.1.4 Die Cast, HPDC, fast cool Casting .................................................................................... 10  
    2.1.5 Casting and Heat Treatment ............................................................................................ 11  
  2.2 Solubility .................................................................................................................................. 12  
  2.3 Machine Learning Technique ............................................................................................... 13  
    2.3.1 Machine Learning ............................................................................................................ 13  
    2.3.2 Neural Network ............................................................................................................... 13  
    2.3.3 Elastic Net Regression ................................................................................................... 14  
    2.3.4 Support Vector Regression ........................................................................................... 14  
Methodology .................................................................................................................................... 16  
  3.1 Dataset Collection ................................................................................................................... 16  
  3.2 Cleaning Dataset and Pre-Processing .................................................................................... 16  
  3.3 Train Test Split and Relationship between Inputs and Outputs ........................................... 17  
  3.4 Model ....................................................................................................................................... 18  
    3.4.1 Elastic Net Regression Model ......................................................................................... 18  
    3.4.2 SVR (Support Vector Regression) Model ....................................................................... 19  
    3.4.3 Neural Network Model .................................................................................................... 19  
  3.5 Objective Function .................................................................................................................. 22  
Results and Analysis ...................................................................................................................... 25  
Conclusion and Future work ......................................................................................................... 29  
Bibliography .................................................................................................................................... 30  
Appendix A ....................................................................................................................................... 31
Introduction

Alloys are made from combination of two or more metals under certain prescribed condition. Forming of good magnesium alloys is the biggest concern for this project. So, forming an alloy with high mechanical properties will fulfil our goal of achieving a good magnesium alloy.

Many types of researches have been made on the Magnesium alloy’s different properties but in this particular project, we are more concerned about the three major properties of alloy which are Ultimate Tensile Strength, Ductility and Yield Strength. In addition to it, processes such as Extrusion, Cast, RHT and many more are used for the extraction of the great alloy which can be seen from the database used. Overall, there are so many parameters which are to be dealt with and which will lead to the making of the alloy. Using hit and trial method will take a lot of time, therefore, to reduce the time and to increase the efficiency of making a great alloy, Machine Learning Techniques are used which will provide a greater edge over the usual method.

Initially, the main two machine learning algorithms can be applied to the data. These two algorithms are Regression and Neural Networks. Both of these machine learning algorithms are tried to create a model and the best one is chosen among them. Under Regression, two algorithms are used which are Support Vector Regression and Elastic Net Regression. The Third algorithm of Machine Learning is Neural Networks which is built in Keras. Nonetheless, all these algorithms give quite decent results but the best algorithm, i.e., the neural network is chosen for the model as it performs the best.

1.1 Objective

The objective of this research project is to produce a Magnesium alloy of the properties which are 350 MPA as the Ultimate Tensile Strength, 300 MPA as the Yield Strength and Ductility as 20 using the Machine Learning Techniques.

1.2 Scope

The scope of the project helps us to know the pace as well as the tasks that have been applied in the research project. So, the research project is divided into four tasks:

1. Collecting the dataset from the Google scholar papers and also the papers provided by Professor Zhouran Zheng, Professor Nick Birbilis as well as the form the Mat Web. Combing and cleaning the database for the application of machine learning comes under this.

2. Application of the different machine learning techniques such as Elastic Net Regression, Neural Network and SVR (Support Vector Machines) on the database extracted

3. Selecting the best algorithm among these three algorithms for further development of the project. Using the best algorithm, generate the data using the data generator by following the solubility table provided
4. Then, make the objective function which will help in extracting the results near to the desired results. Two GUI have been added in the objective function so as to make the objective function more user friendly.

1.3 Contribution
The contribution in this project is to use the machine learning techniques, form a data generator and objective function which will help in making good Magnesium alloys. This will help the alloy industry to use this efficient technique of making alloy by using the Machine Learning methods that have been used in this project to ease their work.

1.4 Report Outline
Report outline will guide all the sections of the report. The first section will be the introduction of the report. Then, the second section will be there for the background or the literature of the project. The third section will explain the different Machine Learning Techniques made and their parameters. Fourth section will explain in detail about choosing the algorithm, result of the model as well as the conclusion and the future work.
Background work

This chapter is divided into three main parts, one is where we explain about the different conditions in which alloys will be made, solubility of an alloy and other would be the explanation about the different Machine Learning Techniques used in this project. Learning about the conditions at which alloys will be extracted is crucial for this project. So, giving a background on these conditions will provide more depth. Knowing about the solubility of an alloy will also help in understanding whether that composition made by the model is feasible or not.

2.1 Alloys’ Conditions
In this section, five main types of conditions used in the database that are given below will be discussed.

2.1.1 ECAE or ECAP
ECAE or ECAP is known as Equal Channel Angular Extrusion or Pressing. This process is used to introduce severe plastic deformations to processed materials to improve their mechanical properties by reducing the grain size. The ECAE die is manufactured with two channels that usually intersect at an angle with the same cross-section. The material is extruded through the die and it is mainly deformed by a shear mechanism combined with a high hydrostatic pressure which exists within the die channels [2].

![ECAE or ECAP Image](image)

Fig
(These optical photographs show the AZ31 Mg alloy by two-step ECAE process at different temperatures at different steps)

2.1.2 Extrusion
Extrusion is a manufacturing process used to create objects of a fixed cross-sectional area. Extrusion is a metal forming process which is widely used in industry and daily life equipment. It is another kind of manufacturing process which involves shearing and compression [3]. The advantage of extrusion is that it can help in developing or building difficult shape alloys as it is known to create the object of fixed cross-sectional area.
Extrusion is generally classified into four types. They are: Direct Extrusion, Indirect Extrusion, Impact Extrusion and Hydrostatic Extrusion. In Direct Extrusion, a solid ram drives the entire billet through a stationary die and provides additional power to overcome the frictional resistance between the surface of the moving billet and confining chamber. In Hydrostatic Extrusion, high pressure fluid applies the force to the work piece through a die. Billet-chamber friction is eliminated, and the pressurised fluid acts as a lubricant between the billet and the die [3].

![Grain structure micrograph of AZ31 alloy after extrusion at 300°C](image)

Fig

(Grain structure micrograph of AZ31 alloy after extrusion at 300°C)

[9]

2.1.3 Sand or Cool Cast
The process is referred to as sand casting because the mold that contains the cavity into which metal is poured is made of compressed or compacted sand. The sand contains some other material that encourages it to hold its shape [4].

The sand-casting process employs techniques that produce shaped parts of nearly any design, including very large parts and those with internal passageways. There may be more optimal casting or metalworking processes for any specific product, based on needed tolerance, design intricacy, volume, tooling availability, or lead time, but a casting in the needed configuration could like be made using the sand process. That is a decision left to the design engineer [4].
Die casting is a process in which the molten metal is forced under enormous pressure into the mold cavity. These are: (1) permanent mold casting, also called gravity die casting, (2) low-pressure die casting, and (3) high-pressure die casting. The three processes differ mainly in the amount of pressure that is used to force the molten metal into the die. The advantages of such a process include high production rates, exceptional dimensional repeatability, low part costs, and less machining due to reduced casting finish stock. The great heat transfer rate obtained through metal molds can further refine and improve the final cast structure, and therefore the mechanical properties of the castings. The reason for using this type of condition or processing is required because it is done on the nonferrous metals and in that list comes the Mg alloys which is our main focus. The disadvantages are design limitations due to metal dies, higher initial die cost, and longer lead time for die construction and for changes to the die caused by a casting design change [6].
2.1.5 Casting and Heat Treatment

There are four main types of casting heat treatments which are listed below:

Annealing

Annealing is generally used for increasing the ductility of the metal. In this process, decreasing the grain size of the metal will help in increasing the ductility. In annealing, metal is beaten at certain temperature and is kept constant for a while and thereafter it is cooled down. Sometimes, metals require cool furnace for the cooling process.

Normalizing

Normalizing process is quite similar to that of annealing but the difference is that, in Normalizing the cooling of metal is always done at room temperature and this process is only applicable to ferrous metals.

Hardening

In this process, the metal is heated at a certain temperature and then the metal is cooled immediately using water or oil. This makes the metal hard. Due to this brittleness in the metal, the ductility of the alloy or metal decreases. Thus, the metal is tempered so as to increase the ductility.

Tempering

It is a process which is always followed by hardening. In this, hardness of the metal is removed by heating it up to a certain temperature and then cooling it in the air. This process helps in deciding the hardness as well as strength of the alloy which is essential for this project.
2.2 Solubility

Solubility of an element is the ability of a particular element to get dissolved in the solution. Accordingly, in the table given below, it can be clearly seen, which elements are soluble up to what extent and which elements are not. The reason for including the solubility is, because it will help in knowing that in how much concentration, the element can be easily used to form an alloy. For this, the list given below is of the elements which are used in the database. So, to form an alloy of the desired mechanical properties, the last column explains the maximum concentration of an element in the alloy. If the alloy is soluble then, 1.5 times its solubility; that much amount of concentration is allowed in alloy. As in this project, formation of Magnesium alloys is the main focus, the concentration of Magnesium should be from 80 to 100 percent. The reason for selection of 1.5 as the number is to achieve the goal of certain concentration and it is only possible in this particular way.

<table>
<thead>
<tr>
<th>Name of the element</th>
<th>Soluble / Insoluble</th>
<th>Maximum Solubility</th>
<th>Maximum concentration of an element in Mg alloy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yttrium (Y)</td>
<td>Soluble</td>
<td>12.47</td>
<td>18.705</td>
</tr>
<tr>
<td>Zirconium (Zr)</td>
<td>Soluble</td>
<td>2.69</td>
<td>4.035</td>
</tr>
<tr>
<td>Neodymium (Nd)</td>
<td>Soluble</td>
<td>3.6</td>
<td>5.4</td>
</tr>
<tr>
<td>Zinc (Zn)</td>
<td>Soluble</td>
<td>6.21</td>
<td>9.315</td>
</tr>
<tr>
<td>Gadolinium (Gd)</td>
<td>Soluble</td>
<td>23.49</td>
<td>35.235</td>
</tr>
<tr>
<td>Manganese (Mn)</td>
<td>Soluble</td>
<td>2.2</td>
<td>3.3</td>
</tr>
<tr>
<td>Aluminium (Al)</td>
<td>Soluble</td>
<td>12.61</td>
<td>18.915</td>
</tr>
<tr>
<td>Silicon (Si)</td>
<td>Insoluble</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Copper (Cu)</td>
<td>Insoluble</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Cerium (Ce)</td>
<td>Soluble</td>
<td>0.23</td>
<td>0.345</td>
</tr>
<tr>
<td>Calcium (Ca)</td>
<td>Soluble</td>
<td>1.35</td>
<td>2.025</td>
</tr>
<tr>
<td>Lithium (Li)</td>
<td>Soluble</td>
<td>5.53</td>
<td>15</td>
</tr>
<tr>
<td>Bismuth (Bi)</td>
<td>Soluble</td>
<td>7.99</td>
<td>11.985</td>
</tr>
<tr>
<td>Antimony (Sb)</td>
<td>Insoluble</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Lanthanum (La)</td>
<td>Soluble</td>
<td>0.74</td>
<td>1.11</td>
</tr>
<tr>
<td>Nickel</td>
<td>Insoluble</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Praseodymium (Pr)</td>
<td>Soluble</td>
<td>0.52</td>
<td>0.78</td>
</tr>
<tr>
<td>Tin (Sn)</td>
<td>Soluble</td>
<td>14.5</td>
<td>21.75</td>
</tr>
<tr>
<td>Silver (Ag)</td>
<td>Soluble</td>
<td>15</td>
<td>22.5</td>
</tr>
<tr>
<td>Ytterbium (Yb)</td>
<td>Soluble</td>
<td>7.96</td>
<td>11.94</td>
</tr>
<tr>
<td>Thorium (Th)</td>
<td>Soluble</td>
<td>4.75</td>
<td>7.125</td>
</tr>
<tr>
<td>Beryllium (Be)</td>
<td>Insoluble</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Iron (Fe)</td>
<td>Insoluble</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Gallium (Ga)</td>
<td>Soluble</td>
<td>8.4</td>
<td>12.6</td>
</tr>
</tbody>
</table>

Fig 2.2
The concentration for the insoluble elements lies in the range of [0, 0.1]. However, there is an exception in this table. Lithium element does not follow the 1.5 times solubility rule. So, it can be said as an exception in the table. By using all these constraints, the alloy with the desired mechanical properties is made.

2.3 Machine Learning Technique

Before diving into the Machine Learning Techniques that have been used for this project, explaining the overview of machine learning will give an insight into the topic that is to be covered in the following paragraphs.

2.3.1 Machine Learning

According to the Tom Mitchell’s definition of Machine Learning, it is a computer program which is said to learn from the experience E with respect to some class of task T and performance measure P, if its performance in task T, as measured by P, improves with the experience E (Chinmay Das, 2017, [1]). Machine Learning will not only help in saving time but it is also helpful in understanding relationship between the inputs and outputs. Sometimes, it is also called a black box in a naive term as we just put the data and expect the output from the algorithm but it is always more than that. It is a learning of the machine which learns from the data without explicitly coding for that learning. So, there are types of Machine Learning Techniques that could be used to get the desired results but finding an optimum algorithm will ensure an upper hand of providing better results as well as analysis.

In this project, three different types of Machine Learning models have been tried and the best one is chosen for the model. So, before diving more into the model, getting a background brief for that type of Machine Learning will help in understanding the deeper concepts.

2.3.2 Neural Network

Neural network is a type of Machine Learning algorithm which resembles the multi-layer brain neurons and helps in learning the patterns and the relationship from the dataset provided in the project.
The above graph shows one of the neural network structures that could have been formed. These connections show how the inputs are connected to the hidden neurons and how hidden neurons are connected to the output layer. The connections between the neurons or the nodes in the neural network are known as the edges. These edges as well as the neurons contain the weights which change as the neural network learns the patterns as well as the relationship between the inputs. These weights and the inputs usually pass through a nonlinear activation function which gives the output accordingly. Neural networks have been used for this project so that the algorithm can understand the patterns between the inputs and can make prediction more accurate. In this project, “Keras” is used as platform for implementing the neural networks.

2.3.3 Elastic Net Regression
It is a form of regularized optimization for linear regression that provides a bridge between the ridge regression and lasso. The estimate that it produces is they can be viewed as a Bayesian Posterior Mode under a prior distribution implied by a form of penalty. Regularization performed in the elastic net is used to improve the predictions based on the ordinary least square error by shrinking the parameter estimator to zero [12]. Elastic Net Regression has been used for this project as one of the models, though it does not give really good results but helps us in choosing the best model. The result and analysis section will help in explaining how the results lead to not choosing this algorithm.

2.3.4 Support Vector Regression
Support Vector Regression is somewhat similar to the Support Vector Machines but the previous one is used for the regression while the latter one is used for binary classification. Now let us dive into the algorithm.

The continuous-valued function being approximated can be written as in equation (2.9.1). For multidimensional data, you augment x by one and include b in the w vector to the mathematical notation to obtain the multivariate regression in the SVR formulates. This functions the approximation problem as an optimization problem that attempts to find the narrowest tube centred around the surface, while
minimizing the prediction error, that is, the distance between the predicted and the desired outputs. The
framer condition produces the objective function in the equation (2.9.3), where $\|w\|$ is the magnitude of
the normal vector to the surface that is being approximated [13]:

$$y = f(x) = \langle w, x \rangle + b = \sum_{j=1}^{M} w_j x_j + b, \quad y, b, \in \mathbb{R}, \quad x, w, \in \mathbb{R}^{M+1} \quad (2.9.1)$$

$$f(x) = \begin{bmatrix} w^T \\ b \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = w^T x + b, \quad x, \quad w, \in \mathbb{R}^{M+1} \quad (2.9.2)$$

$$\min_w \frac{1}{2} \|w\|^2 \quad (2.9.3)$$
Methodology

This section of the report covers the body of the project which is the main focus. It includes the dataset collection, cleaning, encoding, keras model as well as the regularizers used which made the algorithm more generalized and better at prediction.

3.1 Dataset Collection
The dataset plays a vital role in machine learning algorithms. Bad dataset lead to bad results and sometimes, the dataset has to be changed if the results are not good. So, good collection of data will provide not only good results but also provides better generalization of the model. Thus, for collecting the dataset for this project, our group approach was to extract data from different platforms so as to make a huge database for the neural network. 746 data points were collected for the database. Most of the papers were provided by Professor Zhouran Zheng for the extraction of the dataset. From those as well as some additional papers from the Google Scholar, the papers were divided into decades starting from 1950 till now. So, each of the group member took papers from 2 decades respectively, then joined the database. Data was collected from the mat web database also.

3.2 Cleaning Dataset and Pre-Processing
Dataset extracted initially from the papers as well as the mat web was not cleaned. There were some of the data points which were removed from the dataset because they had incomplete columns which will in fact make the learning harder for machine. Moreover, the condition used in the metallurgy industry is a bit different because some conditions come under a certain label. So, using label encoder for those conditions would not have worked. Our esteemed and cautious Supervisor did the labelling of the conditions manually as he is an expert in the field of metallurgy. After the manual labelling, around 30 inputs were there in the database. Some of the columns such as Sc, Sr, Er, Tb, Ty had very few entries of the data due to which those entries had to be deleted because the learning as well as prediction will be difficult because of insufficient data. Professor Nick also made some changes in the dataset so as to make the dataset easy for machine learning. Professor Nick Birbilis also provided about 80 data entries for the database.

Alloys which had rare metal alloys, the concentration of that rare metal was divided into Ce, La, Nd in the proportion of 50 percent, 25 percent and 25 percent respectively. There were some redundant elements’ alloys which were removed by Professor Nick Birbilis. Entries with Sic and alloys which were made by powder metallurgy were also removed from the dataset. After this, the dataset was almost ready for the machine learning. The dataset that is being used is almost same for all the group members in the project. Dataset was loaded in the jupyter notebook for different machine learning techniques that have been applied. It was found that 14 entries were duplicate in the dataset which were further removed from the data frame using inbuilt function called drop duplicates. Duplicate entries were removed from the dataset so as to get good accuracy for the model. The names of the alloys were removed from the dataset as they were unique and we cannot get any useful information from that column for machine learning. Condition was also removed from the dataset as it is redundant because we already have the manual labels which had been assigned to those conditions. The figure 3.2 shows a glimpse of the database made.
3.3 Train Test Split and Relationship between Inputs and Outputs

From the dataset, 80 percent of the dataset was used for training data and the rest of the dataset which consists of 20 percent of data is used as testing data. The dataset for training and testing set is selected randomly for the neural network so that the training can be done on a balanced dataset. If we do not select the dataset randomly, then there is a fair chance of getting poor results on the testing data. Three columns which namely are Yield Strength, Ductility and Ultimate Tensile Strength are set as the target for the neural network as our motive is to predict the properties of alloy so as to complete our 2\textsuperscript{nd} task of the project according to the scope.
Fig 3.3

(Heat map of the inputs and the outputs of the database formed)

This heat map explains the relationship between the inputs and the outputs. It also shows that there is not much inference that can be drawn as there is not much relationship with the inputs as well as with the outputs. So, we can deduce that it will be a hard time for the Neural Network Algorithm to learn from this dataset as there is not much relationship that can be seen.

3.4 Model

Machine Learning Model helps in the prediction of the output. Better model provides better outputs and better generalization. So, there are three models used for the database and the best model will be used for the objective function and data generation.

3.4.1 Elastic Net Regression Model

Inbuilt functions such as Elastic Net has been used for building the model. This inbuilt function helps the model to fit to the database made and gives the results accordingly. Before applying the algorithm, it was made sure that the data is cleaned and pre-processing is done. In this model, normalize is set to true so that the dataset extracted can be normalized using the L2 norm. By normalizing the data, the results improve. Tolerance in the model is set to 1e-3. It is updated at every iteration and it checks whether the model’s tolerance is lower than the previous time or not. If the tolerance becomes lower...
than the previous one, the model stops optimising. Alpha in this model is the penalty terms. So, if alpha is 0, then the model is using the L2 norm and if alpha is 1, then the model is using the L1 norm. Therefore, maintaining a balance is needed in this model. So, alpha is kept to 0.001. The results obtained using the model have been discussed in the results and analysis sections.

3.4.2 SVR (Support Vector Regression) Model

Inbuilt functions such as Linear SVR has been used for building the model. This inbuilt function helps the model to learn the relationship between the input and the output in the database. In this model, variable C is set to 0.8. It is a parameter which helps in defining the regularization in the model. C is inversely proportional to the regularization. So, as the default value of C is one, decreasing the value by too much will lead to highly regularized model due to which there will be large penalty, if the prediction is not right. Therefore, using regularization properly is an apt way of making the model better. C is set to 0.8 so as to regularize the model as well as to get a good R2 score. Tolerance is also set to 1e-3. One major problem with linear SVR is that it only takes one output. But, in this database, there are 3 outputs. So, to solve this problem, “Multiple Output Regressor” is used. This is an inbuilt function which will take regression model as an argument. It will then create one instance of the provided model for each output [14]. The results obtained using the model has been discussed in the result and analysis section.

3.4.3 Neural Network Model

Neural Network Model for this project is made from scratch. So, there are lots of functions and parameters which need to be discussed to understand the model made for the database.

3.4.3.1 Topology of the model

There are 25 columns as the inputs and 3 columns are named “Ductility”, “Tensile Strength” and “Yield Strength”. In the model, there are 4 dense hidden layers with the 512 nodes or neurons in first hidden layer, 256 nodes or neurons in second hidden layer, 128 nodes or neurons in third hidden layer and 64 nodes or neurons in the fourth neuron layer. Large number of hidden layers are used in this model because firstly, there are 26 inputs, so more neurons will help in training of the model properly as well as there is not much relationship between the inputs and the outputs which can be derived from the heat map shown in the fig 3.3 due to which addition of more neurons will help in understanding the complexity of the dataset and will produce really good results. Though, increasing the layers increases the risk of overfitting, some regularizers should be added to avoid that. By adding regularizers, it will help in making model more generalized which will help in predicting good results. For better optimisation of the model, Adam optimiser is used because it can handle the gradients very well as well as it is one of the best optimisers till date. Graph displayed below shows the summary of the model-
Fig 3.4.1

(Summary of the keras feed forward neural network model made)

3.4.3.2 R2 Score

R2 Score is the regression scoring coefficient. This is used as the measure of accuracy because in regression, there is not any parameter of “accuracy” or “precision” as it is there in the classification. The range of R2 score lies between the range of 0 and 1. More the R2 score is close to 1, better the model is. R2score is a sklearn library which is imported in this project for the usage. Though we have the training error as well as the test error, but RMSE error is not used as the parameter for the distinguishing whether the model performs good or bad. The reason for not choosing the RMSE error is because there is no range for the RMSE error. There is no set bar which tells us that if the RMSE error reaches that level, then our model is optimum. So, R2 score is apt for calculating the score as it has the range.

3.4.3.3 Kernel Initializer

Initial weights are required in the neural network because it helps the underlying function to learn and make the prediction closer to the desired output. So, it is one of the key components in the neural network. Kernel Initializer is another name for weight initialization. In this, if the Kernel Initializer is set as “normal”, then the weights in that particular layer is initialized to be normally distributed or Gaussian normally distributed. In the same way, the model made in this project has used glorot normal as well as normal as the Kernel Initializer in different respective layers. The reason for using glorot normal and normal Kernel Initializer is because both of the initializers, initialize the weights at the mean 0 and standard deviation 1 due to which weights are really small which helps in the regularization of the model and it gives good results. These initialized weights are used for the back propagation in the model. As, the platform used for this project is keras, we do not have to specify the back propagation because keras does the back propagation automatically on its own.
3.4.3.4 Overfitting
Overfitting in the neural network is one of the common things to occur. Neural networks with lots of layers are susceptible to overfitting. The reason is that the dataset for analysis is less and in addition to it, there is not much relationship between the inputs and the outputs. So, trying the model to fit correctly so as to get better results could lead us to the overfitting of the model. Thus, many of the techniques are used for prevention.

3.4.3.4.1 Dropout
Dropout is a regularization technique which prevents overfitting of the model by removing some connections between the nodes in the model. Dropout function has a parameter which decides at what probability rate, the connections from the nodes should be removed. If the parameter is set to zero, then all the nodes are removed and if the value is set to 1, then none of nodes are removed. In this project, this parameter has been set to 0.1. The reason for doing this is because the model was overfitting as well as dropout works best if the dense layers are there in the model. The R2 score value after implementing the dropout increased from the 0.59 to 0.65 for Ultimate Tensile Strength, 0.50 to 0.63 for Yield Strength and 0.25 to 0.35 which is tremendous change in the accuracy of the model. Dropout works for the dense models because it makes the training noisy which forces the layer that is left to take the responsibility of the inputs for making the prediction which makes the model more robust.

![Image](https://via.placeholder.com/150)

(a) Standard Neural Net  
(b) After applying dropout.

[7]

Fig 3.4.2

(This the graph comparing between the before dropout and after dropout)

3.4.3.4.2 Kernel Regularizers
Kernel Regularizers are applied to the weight matrix of the neural network model formed. It is a type of regularization on the Kernel which adds penalty to loss function. We can control that in which layer do we want to add regularization. In this model, at every layer, regularizers have been added because layers used in the model are dense due to which overfitting can happen. So, regularization will prevent overfitting. Similarly, Bias regularizers are also added in the model for the same context, i.e., to avoid overfitting.
3.4.3.3 Call Backs

Different types of Call Backs that can be used in the model, but the Call back named “Reduce on Plateau” has been used in the neural network model. Further, it reduces the learning when the model stops improving. While using this, patience variable is kept 2. This means that if model’s learning rate has been constant since the last two epochs, then the learning rate will be decreased. By doing this, the model keeps on learning the relationships in the dataset which helps in prediction, closer to the target.

After initializing all the parameters and making the model for training, we try to fit the model on the training data by running the model for 300 epochs containing the batch size of 13. After training the model, when the testing dataset is used for testing of the model, it gives the prediction for the comparison of the results. After the completion of making the model and tweaking the parameters so as to get the optimal results, now proceeding on the objective function will lead us to the completion of the project.

3.5 Objective Function

Now completion of task 4, according to the scope, is important because it will not only complete the project but will serve the purpose of the project that we are heading to. To make solver which will solve this problem, first data has to be generated using the neural network. In this project, 2.5 million data points have been generated using the neural network model.

Fig 3.5.1

(Data generator explained in the form of flowchart)

The above figure explains about the data generator formed in the model. For generating the random inputs, inbuilt function known as Faker is used to generate random numbers. It is made sure that the solubility table as shown in Fig 2.2 is followed strictly. Then, those numbers are restricted in a particular range as well as normalized so that the overall concentration does not increase by 100 percent. After the data has been generated, the whole generated dataset runs through the neural network model and makes the predictions accordingly. As, there are 2.5 million data points, the time complexity of running the program will be higher.

After the data is generated and the predictions are made, we want to again follow the solubility table because when the data is normalized, then there is a possibility that the elements might have exceeded their range provided to it. So, again the range is set accordingly. Alloys concentration should be restricted accordingly because by doing that, alloy can be made which has 350Mpa as the Ultimate Tensile Strength, 300Mpa as the Yield Strength and 20 as the Ductility.

Then we try to enter the desired outputs in the GUI which is shown below:
Our objective function then tries to search through all the database and tries to find all the values which are close to these mechanical properties entered in the GUI above.

Now, concentration range is set again because when the data was normalized, it could have happened that the range of the element that have been set, would have exceeded. So, to keep a check, again the conditions are run to get the fine results that we desire. However, narrowing the range of the concentration is also possible. To set the range of the elements, GUI has been used which is given below:
When the values are entered and all the conditions are satisfied, then program runs which will use norm so as to find the output which is closer to the desired result considering the solubility table is satisfied.
Results and Analysis

Analysing the results obtained from the model is a crucial part of the report because it helps to know how the project results are and how well the project is done.

Our initial goal for this project is to find the best machine learning technique for the database extracted. So, after the algorithm has generated the results, comparing the R2 score will give an insight about which model performs the best. So, the results obtained after running the program has been shown below:

<table>
<thead>
<tr>
<th>Model</th>
<th>Ultimate tensile strength</th>
<th>Yield Strength</th>
<th>Ductility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Network</td>
<td>0.7379</td>
<td>0.7059</td>
<td>0.4070</td>
</tr>
<tr>
<td>SVR</td>
<td>0.2642</td>
<td>0.3637</td>
<td>0.3035</td>
</tr>
<tr>
<td>Elastic Net regression</td>
<td>0.5040</td>
<td>0.5119</td>
<td>0.3388</td>
</tr>
</tbody>
</table>

The above table shows different R2 score values for different machine learning techniques followed. It can be deduced from the above table that the neural network outperforms other two Machine Learning Techniques because the neural network’s R2 score is much higher for all three parameters which are the mechanical properties of the alloy. After the Neural Network model is selected, knowing about how well model predicts is the key.

![Fig 4.1](image)

(Prediction vs Actual graph for Ultimate Tensile Strength)
These above graphs show that prediction made by neural network regression model is good. So, after the selection of the model, create a data generator and the objective function. These two functions are the main pillars of the project. After the data has been generated using the data generator, we try to get the nearest possible results using the objective function. The results that have been got using both the functions are given below:
<table>
<thead>
<tr>
<th>Element Name</th>
<th>Alloy 1</th>
<th>Alloy 2</th>
<th>Alloy 3</th>
<th>Alloy 4</th>
<th>Alloy 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mg</td>
<td>88.76464</td>
<td>93.58974</td>
<td>89.20735</td>
<td>87.15379</td>
<td>87.28376</td>
</tr>
<tr>
<td>Y</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.278079</td>
</tr>
<tr>
<td>Zr</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.1748</td>
</tr>
<tr>
<td>Nd</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.11025</td>
</tr>
<tr>
<td>Zn</td>
<td>0</td>
<td>1.893939</td>
<td>0</td>
<td>1.332102</td>
<td>0</td>
</tr>
<tr>
<td>Gd</td>
<td>4.144192</td>
<td>0</td>
<td>0</td>
<td>2.334476</td>
<td>1.859024</td>
</tr>
<tr>
<td>Mn</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.929512</td>
</tr>
<tr>
<td>Al</td>
<td>0</td>
<td>1.369464</td>
<td>1.014811</td>
<td>1.622263</td>
<td>0</td>
</tr>
<tr>
<td>Si</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cu</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ce</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ca</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.916602</td>
</tr>
<tr>
<td>Li</td>
<td>1.039337</td>
<td>0</td>
<td>3.250137</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bi</td>
<td>1.828707</td>
<td>0</td>
<td>1.892485</td>
<td>1.358481</td>
<td>0.916002</td>
</tr>
<tr>
<td>Sb</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>La</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.002374</td>
<td>0</td>
</tr>
<tr>
<td>Ni</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pr</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sn</td>
<td>2.249704</td>
<td>0</td>
<td>3.231337</td>
<td>1.988123</td>
<td>0</td>
</tr>
<tr>
<td>Ag</td>
<td>1.973425</td>
<td>1.384033</td>
<td>1.467361</td>
<td>1.965181</td>
<td>2.543248</td>
</tr>
<tr>
<td>Yb</td>
<td>0</td>
<td>1.762821</td>
<td>1.919912</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Th</td>
<td>0</td>
<td>0</td>
<td>1.247943</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Be</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fe</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ga</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Conditions under which elements go through to form an alloy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**Mechanical Properties of the alloy formed**

<table>
<thead>
<tr>
<th></th>
<th>Alloy 1</th>
<th>Alloy 2</th>
<th>Alloy 3</th>
<th>Alloy 4</th>
<th>Alloy 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTS (Ultimate Tensile Strength)</td>
<td>349.8074</td>
<td>350.2972</td>
<td>350.1059</td>
<td>349.1685</td>
<td>349.3119</td>
</tr>
<tr>
<td>YS (Yield Strength)</td>
<td>300.2112</td>
<td>299.929</td>
<td>299.3131</td>
<td>300.2816</td>
<td>299.8255</td>
</tr>
</tbody>
</table>

Fig 4.4

(Final result table obtained)

As our goal was to form an alloy with the mechanical properties like, 350Mpa as the Ultimate Tensile Strength, 300Mpa as the Yield Strength and 20 as Ductility. From the above, it can be deduced that there 5 different composition of Mg alloys which can be made and have the desired properties. So, we can conclude that the we have achieved our goal for the project. In addition to it, in the above given under the Condition section, numeric have been assigned to the alloy. Those numeric indicate different conditions through which alloy can be made. Those alloy conditions can be seen below.
<table>
<thead>
<tr>
<th>Condition number</th>
<th>Condition Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Extrusion</td>
</tr>
<tr>
<td>2</td>
<td>ECAE/ECAP</td>
</tr>
<tr>
<td>3</td>
<td>Cast (Sand/Slow/Cool)</td>
</tr>
<tr>
<td>4</td>
<td>Cast (Die cast/HDPC/Fast cool)</td>
</tr>
<tr>
<td>5</td>
<td>Cast (Any type, then heat treatment)</td>
</tr>
</tbody>
</table>

Fig 4.5

(Condition table)
Conclusion and Future work

Choosing the best machine learning algorithm is one of the best ways of building the model for the database because it will allow us in making great predictions. Over here, Neural Network Technique outperformed the Elastic Net Regression as well as the SVR algorithm because the R2 score for that algorithm is closer to 1 and it is well known that Regression Model performs the best when R2 score of the model is closer to 1. Using the best technique with the combination of the objective function and the data generator produces the results for the model which are required. From the above table, we can conclude that the mechanical properties that we are looking for, have been found satisfying the solubility table. So, we can say that our project was successful in getting the desired results and forming an alloy using the machine learning techniques.

However, future work will help in continuing with this project. Increasing the database will give an aspect of generalizing the model more which will help to make the model better. The R2 score for the Ductility can be improved in the further version of the model. In addition to it, a few other physical properties such as atomic radii, electronegativity, temperature at which alloy can be made can be added into the database which will make the database denser and by adding all these entities, we can make an alloy in the furnace for the usage and it can be used in automobiles in the years to come.
Bibliography

[9] H.K.Lin, J.C.Huang, T.G.Langdon, “Relationship between Texture and Low Temperature Super Plasticity in an extruded AZ31 Mg alloy processed by ECAP.”
Appendix A: Independent Study Contract Projects

INDEPENDENT STUDY CONTRACT PROJECTS

Note: Enrolment is subject to approval by the course convenor.

SECTION A (Students and Supervisors)

| Unit ID:  | U6742441                  |
| Surname:  | Pawar                      |
| First Names: | Prateek                |
| Formal Supervisor: | Professor Nick Birbilis  |
| Course Code, Title and Units: | Comp 4560, Advanced Computing Research |

COMMENCING SEMESTER: ☐ S1 ☒ S2 YEAR: 2019 Two-semester project (12u courses only): ☒

PROJECT TITLE:
Analysis of performance of magnesium alloys using Machine learning techniques

LEARNING OBJECTIVES:
The project will explore the use of machine learning to rationalise the performance of magnesium alloys fabricated by extrusion. The variables in alloy production and fabrication are large (composition of alloy, extrusion temperature, speed, reduction of area, etc.). At present, human experience is capable of wholly capturing the role of inputs (compositions + processing) on the outputs (strength and ductility) for such alloys, warranting a machine learning approach.

PROJECT DESCRIPTION:
Light alloys with high strength are crucial to the future of transportation, including flight, automotive, etc. In addition, light alloys now populate smartphones, laptop chassis, etc.

The lightest of the 'engineering' or structural alloys are magnesium alloys.

The development of a magnesium alloy is carried out on the basis of selecting a composition (i.e., the alloy ingredients, which may include up to 10 alloying elements) and processing conditions (which in the case of extrusion includes temperature, extrusion speed, extrusion ration, quench rate, and the possibility of multistep processing).

In short, the number of input variables that contribute to the alloys properties (such as strength) are many. There are too many inputs for human level interpretation or visualization (by graphs / plots) of compositional and processing relationships.

This complexity (i.e., high dimensionality) warrants the exploration of machine learning as a means of learning the relationships between the many inputs, and the few outputs (such as strength and ductility of the final alloy).

ASSESSMENT (as per the project course’s rules web page, with any differences noted below).

<table>
<thead>
<tr>
<th>Assessed project components:</th>
<th>% of mark</th>
<th>Due date</th>
<th>Evaluated by:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report; style: (e.g. research report, software description...)</td>
<td>(min 45, max 60)</td>
<td>60</td>
<td>(examiner)</td>
</tr>
<tr>
<td>Artifact: wireframe (e.g. software, user interface, robot...)</td>
<td>(max 45, min 30)</td>
<td>30</td>
<td>(supervisor)</td>
</tr>
<tr>
<td>Presentation:</td>
<td>(10)</td>
<td></td>
<td>(course convener)</td>
</tr>
</tbody>
</table>

MEETING DATES (IF KNOWN):

STUDENT DECLARATION: I agree to fulfill the above defined contract:

.......................................................... .......................... 31st June 2019...
Signature Date

SECTION B (Supervisor):

I am willing to supervise and support this project. I have checked the student’s academic record and believe this student can complete the project. I nominate the following examiner, and have obtained their consent to review the report (via signature below or attached email):

.......................................................... .......................... 5 - August - 2019
Signature Date

Examiner:
Name: ... Dr. Liang Zheng.............................................. Signature .................
(Nominated examiners may be subject to change on request by the supervisor or course convener)

REQUIRED DEPARTMENT RESOURCES:


SECTION C (Course convener approval)

.......................................................... .......................... 6/9/119
Signature Date

Research School of Computer Science Form updated Jun 2018
Appendix B:

There are three files for the coding in this project which are explained below:

**Neural_Network_Comp4560_Project.ipynb**

Neural Network has been implemented in this file. Dataset has also been cleaned in this file. As this model produces the best R2 score, it is further used for data generation as well as the objective function which will help in generating an alloy with the desired mechanical properties. GUI has been added at different steps so as to make the project more user friendly.

**ElasticNet_Comp4560_Project.ipynb**

Elastic Net is a form of regularized Regression Model. This has been implemented in this file. Three models have been made in this project so as to select the best model for the objective function as well as the data generator. Although it performs better than SVR, it does not perform better than neural network.

**SVR_Comp4560_Project.ipynb**

This file contains the Machine Learning Algorithm named SVR, i.e., Support Vector Regression which is applied to the database formed. This algorithm gives the least result among the other two algorithms.
Appendix C:

Readme file

Background:

This project of machine learning helps in extracting the good Mg alloys of particular Tensile Strength, Yield Strength and Ductility. The model properties are fed so as to find the different composition and condition which could be used to make an alloy.

Getting Started

This project is made on windows environment. So, the installing command for the libraries such as keras would be different for different operating systems.

Required Libraries and installation:

The libraries used in this model are:

1. **Keras**
   
   Installation command – “pip install Keras”

2. **Faker**
   
   Installation command – “pip install Faker”

3. **NumPy**
   
   Installation command – “pip install numpy”

4. **Pandas**
   
   Installation command – “pip install pandas”

5. **Seaborn**
   
   Installation command – “pip install seaborn”

6. **Matplotlib**
   
   Installation command – “pip install matplotlib”

7. **TensorFlow**
   
   Installation command – “pip install tensorflow”

These installation commands work on the anaconda prompt. Moreover, these commands are for windows and ubuntu environment.

Prerequisites

1. Python and Anaconda (jupyter notebook) should be installed in the pc to run the code on pc
2. Running the following code on the GPU will help in running the code efficiently. Not using the GPU for code will increase the time complexity or the time taken for the code to run.
3. Alternatively, if Python is not installed, the project can be run on the Microsoft azure notebooks. It is an online local host on which jupyter notebook can run

Running the Code of the Project
1. For running the code on Microsoft azure notebooks, find the file from the pc, upload it on the azure and click on the ‘.ipynb’ file.

2. For anaconda prompt users, change the drive from the C:/Users to the drive where the file is located and then type “jupyter notebook”. It will open a local host on your web browser. Then, find the file and open it for the running.

3. Before running the code, please do not make changes in the numeric or any variables. The variables in the model are set accordingly so as to get optimum results.

4. So, for running the cell in the code, a shortcut key can be used which is “Shift + Enter” or just click on “Cell” column and click on “run all” to run the entire program.

5. So, for this project, three files have to be run so as to see the results accordingly. First file is named “Elasticnet_Comp4560_Project” in which elastic net regression is used for the database which is a type of machine learning algorithm. Similarly, there is another file named “SVR_Comp4560_project” in which SVR (Support vector regression) is implemented on the database. These two algorithm’s R2 scores, i.e the accuracy measure for the model which helps in deducing that the neural network which is implemented in “Neural_Network_Comp4560_Project” works the best out of all the models implemented. This can be seen when the model is run and the R2 score value is compared. For running the program, follow the above given steps.

6. In the GUI provided in the neural network model formed, enter the integer or float values in the blank rows provided and then click on the button to get the results which will be stored as a variable. Run all the cells continuously and then the Mg alloy formed will be displayed on the screen as well as the results will be saved in the csv named “Generated_alloy_properties.csv”.

NB: The code running for data generation or objective function may take 20 to 40 minutes. Comments have also been added in the code for understanding.

Author

Code and readme file is made by Prateek Arora, u6742441

Database is made with the help of Professor Nick Birbilis, Professor Zhouran Zeng, Prateek Arora and the group members (Samyak Jain(u6734495), Bhumipat(u6069393) and Jia Ye(u5879731)).