Project 2 Foundation of Intelligent System

MLP and Decision Tree

By – Sanyukta Sanjay Kate (ssk8153)

Amit Magar (ajm6745)

1. We have implemented the Multi layer preceptron and the decision tree.

   Differences: We see that in mlp we all the instances are used by the neural network/system. Whereas, in decision tree, the decision tree uses only those attributes which contribute significantly on the accuracy or are important. Whereas, in mlp, all instances are used.

Similarity: The neural network and decision tree are used for classifying instances.

The multi layer perceptron, a neural network is created with 1 hidden layer (with 5 hidden neurons) and 1 output layer with 4 output neurons. In this, for every instance we do forward propogation, which gives the predicted class for that instance. We then find the sum of square errors which determines how far are we from the actual class. We propogate back the error we have got an then update the weights and the bias node weights. We use the updated weights on the next instance (stochistic updation). This is gradient decsent method. In decision tree we have used the information gain as the threshold value, to split the dataset/instances into two parts. The similarity here, is that we have a set of instances which are classified into different classes but based on a different approach.

2. Data

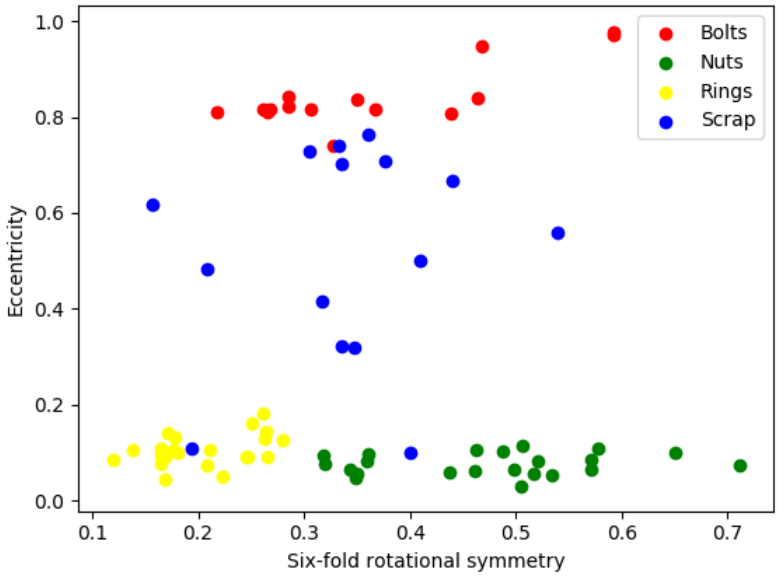


Figure 1. Classification of Data Points using MLP

This is a plot of training data points which either represent bolts, nuts, rings or scrap.

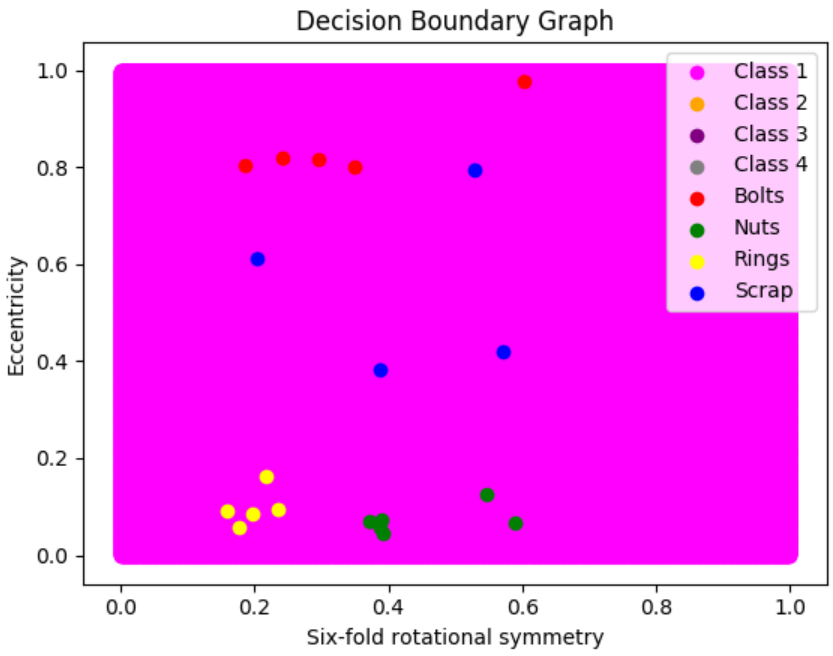
- From the above graph we observe that the bolts (in red) are have a very high eccentricity and are spread across the six fold rotational symmetry. One of the bolts have a very high eccentric value and a very high six fold rotational symmetry (around x-value = 0.6 and y_value=1.0).
- The nuts which are green in color, have a very low eccentricity but are spread across six fold rotational symmetry (0.3<x_value<0.8). The points are not very close to each other that is they are not clustered together in one space but also, do not have any outliers.
- The Rings which are yellow is color are clustered near to each other. They have a very low eccentric value (between 0.0 to 0.2), and also do not have a very high six fold rotational symmetry (between 0.1 to 0.3).
- The scrap are spread across the regions (space). One of them have a low eccentricity (0.1)and low six fold rotational symmetry (0.2), whereas the rest have different values. They are not clustered together and are highly spread out from each other.

This shows that the data points have different 2 features but with different feature values.

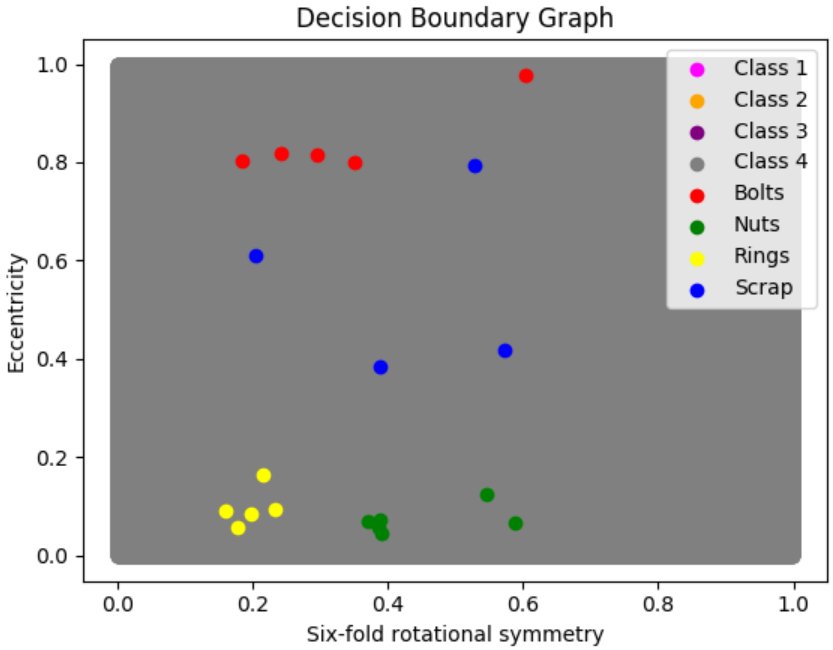
3. Results (MLP)

(I). Decision Boundary

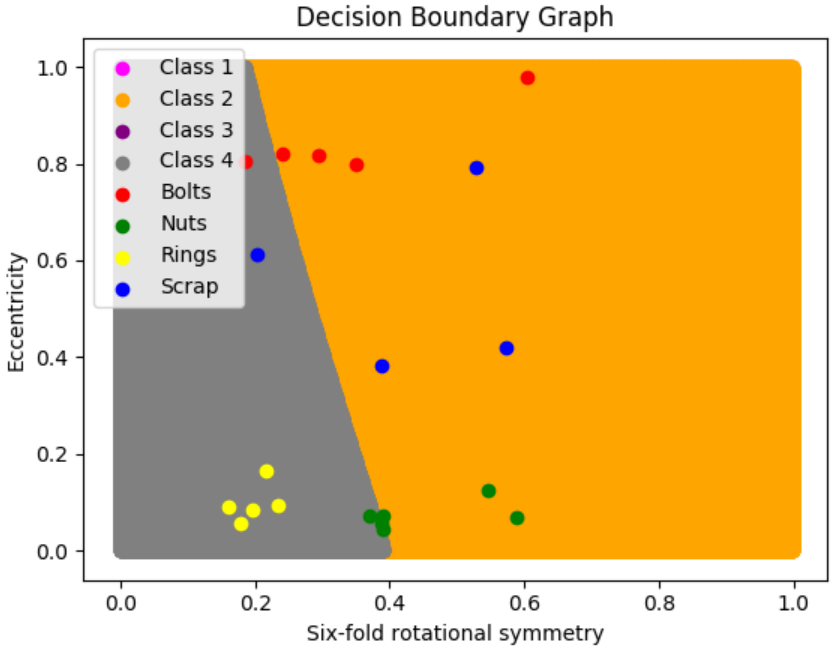
a. 0 Epochs



B). 10 epochs

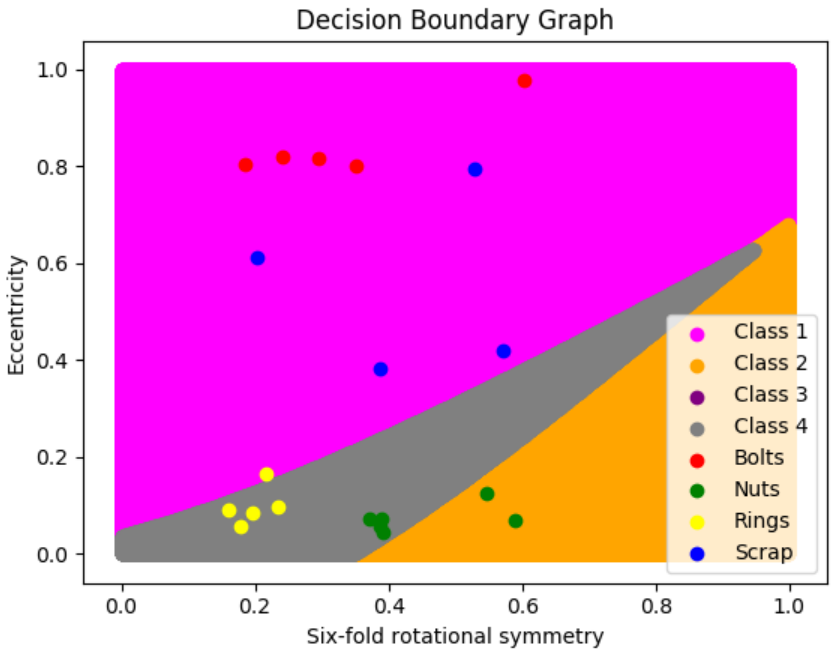


C). 100 Epochs

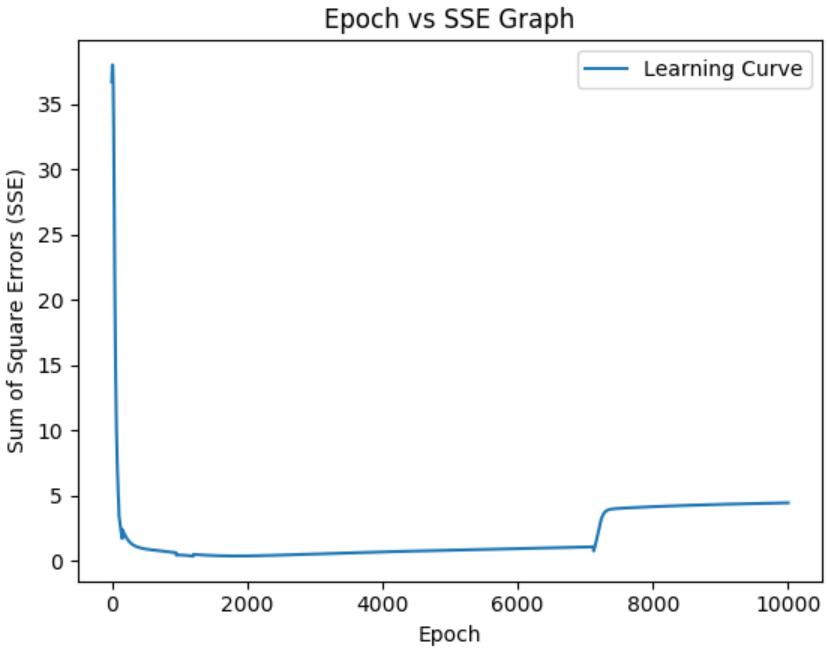


D) 1000 epochs

E). 10000 epochs



Ii)Learning curve for 10000 epoch

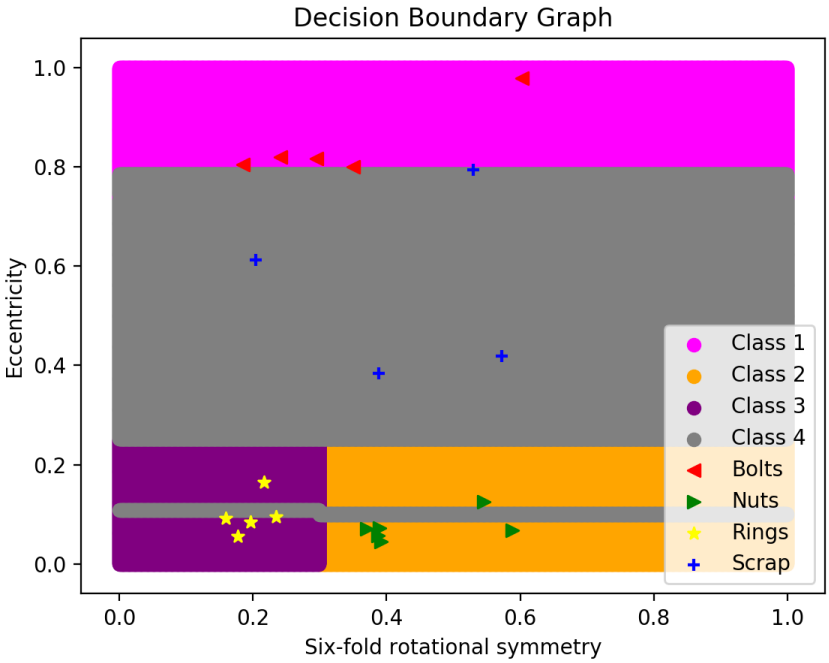


Iii).

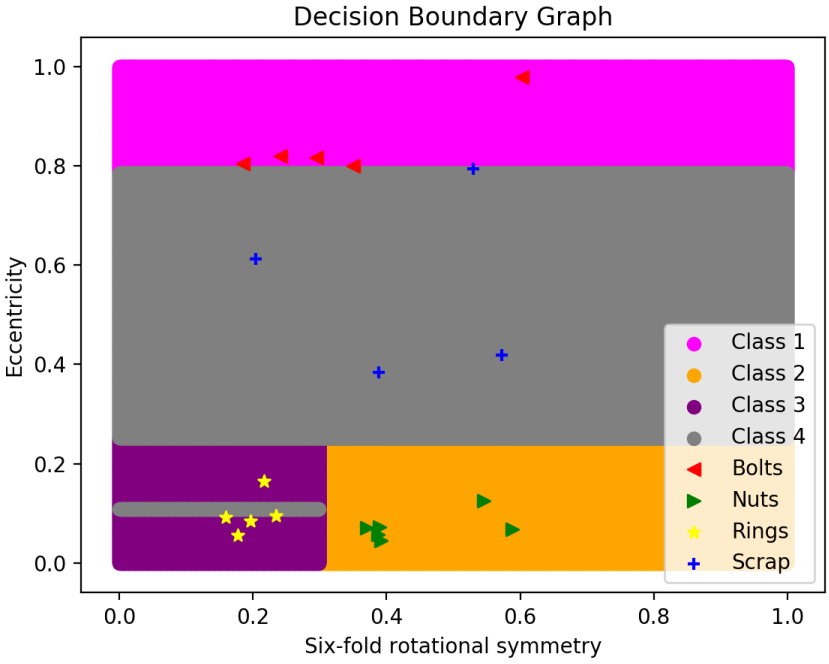
Epoch number	Recognition rate	profit
0	25%	32
10	35%	47
100	45%	20
1000	55%	127
10000	85%	191

Decision Tree

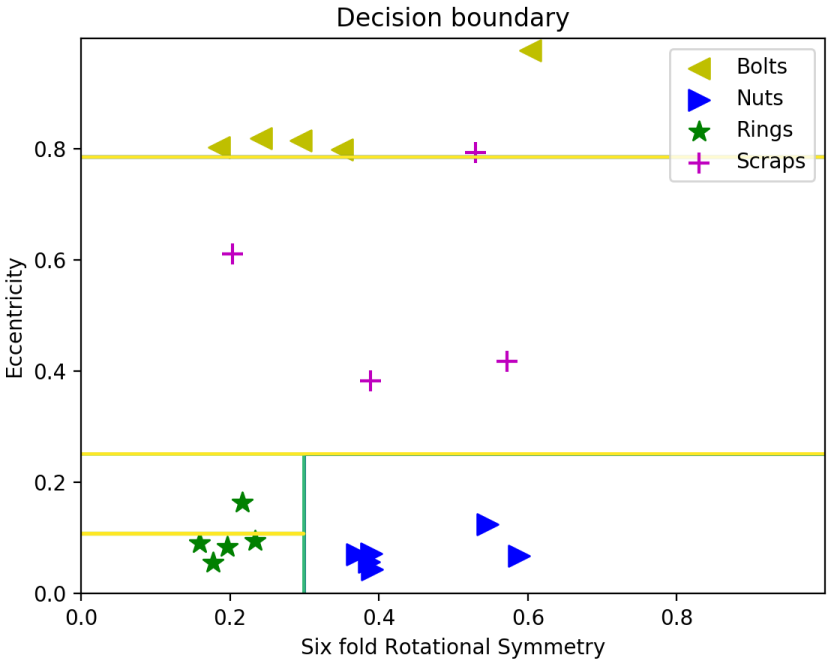
1.1) Decision Boundry with Inductive Tree (not Pruned Tree)



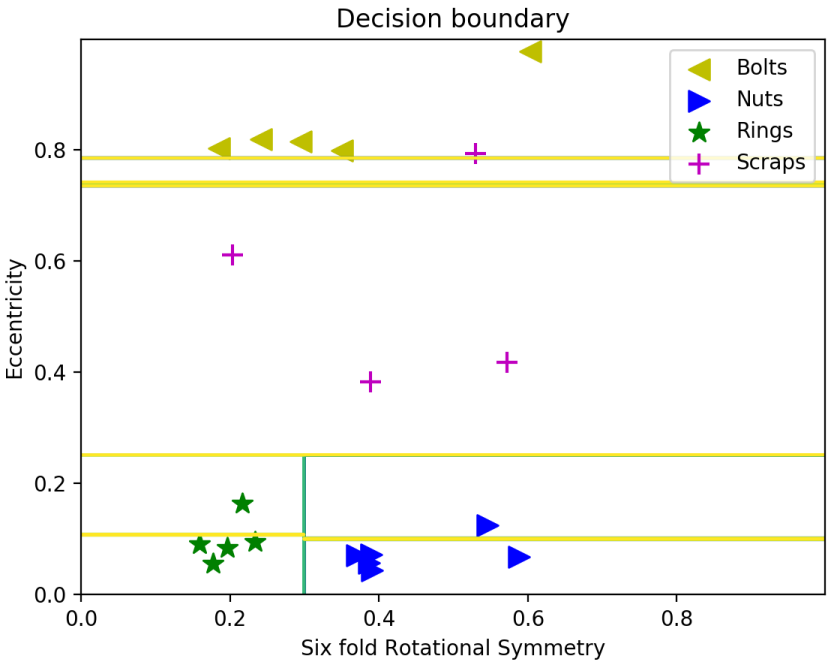
2.2) Decision Boundry with Pruned Tree



2.1) Feature Space of Pruned Tree



2.2) Feature Space with Inductive Tree



3.1) Recognition Rate and Profit earned

	Inductive Tree(Non-pruned)	Pruned Tree
Recognition Rate	95.00%	95.00%
Profit Earned	199	199

3.2)

Decision Tree Information of Inductive Tree

----------Summary----------

Maximum Height of Decision Tree is : 4

Minum Height of Decision Tree is : 2

Average Height of Decision Tree is : 2.083333333333333

Number of Nodes in Decision Tree : 19

Number of Leaf Nodes in Decision Tree : 10

Decision Tree information of Pruned Tree

----------Summary----------

Maximum Height of Prunned Decision Tree is : 4

Minimum Height of Prunned Decision Tree is : 2

Average Height of PrunnedDecision Tree is : 2.083333333333333

Number of Nodes in Prunned Decision Tree : 11

Number of Leaf Prunned Nodes in Decision Tree : 6

Discussion

1) Which of the different classifiers performed best in terms of 1) accuracy and 2) profit? Did this meet your expectations?

From Above results we can see that decision tree has performed better than Multi Level Perceptron. The Profits and Accuracy is higher for decision tree are higher than the MLP.

This was not as expected the MLP should have performed better, May be after more epoch the MLP will give us better results than the Decision Tree. Here we can see that data is linearly separable, if that was not case then decision tree would have not performed better than MLP

2)How do the hypotheses (i.e. class boundaries) and performance metrics differ between the different version of the MLP and decision trees, and why?

We can see that Profit and Accuracy achieved by inductive decision tree and pruned decision tree are same. But this will not be always case, Also the test data we are provided with is less, This might be reason of getting same performance. Generally the Inductive Tree are prone to be overfit for the train data, and they perform badly for test data.