Assignment-2

M.Sushanth-20CS10034
S.Bhanu Teja-20CS10059

2.Supervised Learning:

1.Standard Scalar Normalisation:

It is also Known as Z-score normalisation. It is the transformation of features by subtracting from mean and dividing by standard deviation. This is often called as Z-score.Standardization can be helpful in cases where the data follows a Gaussian distribution.

```
X_new = (X - mean)/Std
```

2.Binary SVM Classifier:

Support Vector Machines (SVMs) are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis.Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier.

SVM Accuracy:

- Linear function SVM accuracy : 0.9666666666666667
- Quadratic function SVM accuracy : 0.933333333333333
- RBF(radial basis function) SVM accuracy :0.9

3.MLP Classifier:

MLP Classifier stands for Multi-layer Perceptron classifier which in the name itself connects to a Neural Network. Unlike other classification algorithms such as Support Vectors or Naive Bayes Classifier, MLP Classifier relies on an underlying Neural Network to perform the task of classification.

Accuracy of MLP classifier for 1 hidden layer with 16 nodes: 0.9

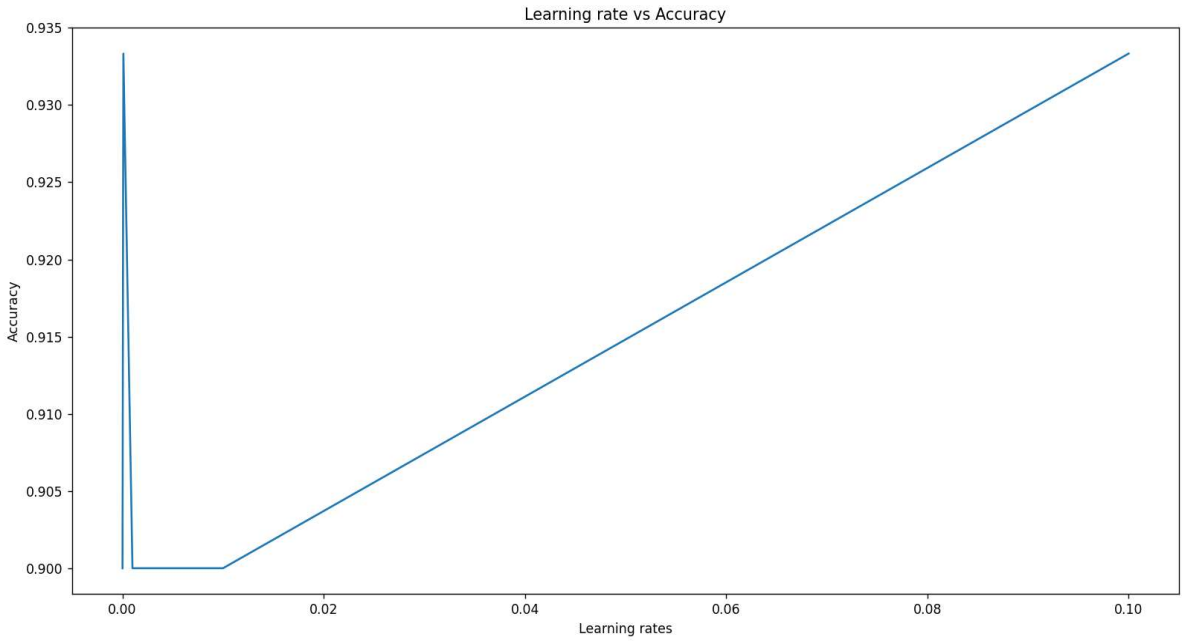
Accuracy of MLP classifier for 2 hidden layers with 256 and 16 nodes respectively:0.9

Using the best accuracy model from the above

Learning rates : [0.1,0.01,0.001,0.0001,0.00001]

Corresponding MLP classifier accuracy : [0.9333333333333,0.9,0.9,0.9333333333333,0.9]

Learning rate vs accuracy graph



4.Backward Elimination Method:

Backward elimination is a simple and effective way to select a subset of variables for a linear regression model. It is easy to implement and can be automated. The backward elimination process begins by fitting a multiple linear regression model with all the independent variables. The variable with the highest p-value is removed from the model, and a new model fits. This process is repeated until all variables in the model have a p-value below some threshold, typically 0.05.

Best Features :

PetalWidth, SepalWidth, PetalLength.

5.Max Voting Technique:

The max voting method is generally used for classification problems. In this technique, multiple models are used to make predictions for each data point. The predictions by each model are considered as a ‘vote’. The predictions which we get from the majority of the models are used as the final prediction.

Accuracy of Max-vote classifier : 0.9