**Muhammad Mubeen**

**SP18-BCS-034**

**B**

**ML Project**

**Task:**

The aim of this project is to predict the Type of IRIS flower, the task is to pass the data by splitting it into training and testing dataset, as 1 attribute has categorical data as wells so the data must be label encoded and then it would fit in 10 different model and on the basis of accuracy values, the best model will be selected and will predict the survival of the passengers after getting input from user.

**Dataset:**

The dataset is the Iris Flower Dataset which has 4 input attributes and have the Species as an output attribute, the input attributes are sepal\_length, sepal\_width, petal\_length, petal\_width species which contains numeric values and categorical values such as species needs to be label encoded. It is significantly small dataset to train the model.

**Evaluation Methodology:**

One evaluation methodologies has been used in this experimental task, one is Train-Test split ratio. In the Train Test split ratio, 80/20 split ratio has been used in which 80% dataset is Train data and 20 % data is Test data.

**Evaluation Measures:**

In this project 1 evaluation measures have been used which is accuracy to get the best and accurate model for our task to correctly predict the specie of iris flower.

**Machine Learning Algorithms:**

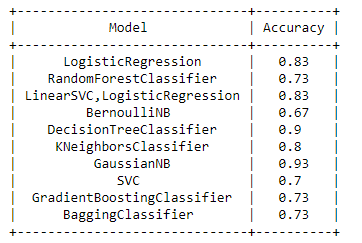
In this project I have used 10 classifiers algorithm of machine learning and have trained them with my data to see which model works best with the Iris Flower Dataset, the model which gives higher accuracy, precision, recall, f1 will be the best model. The applied algorithms include:

* Linear Regression
* Random Forest Classifier
* Linear SVC
* Bernoulli NB
* Kneighbors Classifier
* Decision Tree
* Gaussian NB
* SVC
* Gradient Boosting Classifier
* Bagging Classifier

**Results Summited in Table:**

In this project 2 evaluation methodologies have been used to find the best model, so the results that they have produced are:

Train Test Model Result:



**Observation:**

Different classifier algorithms have been used in this project including Instance Based Evaluation, Artificial Neural Network, Decision Tree but the best model that has the best evaluation measures for this data is GaussianNB for the train test validation, so I have save this model and I have used it in the Application phase and used to predict the output/survival of the user given inputs.