Eucalyptus

My data set has 19 features and one target variable. Out of which 11 are numerical variable and 8 are categorical variable.

As my dataset has some ‘?’ cleaning process is done before providing the data to the model. Question mark is filled with the median of the respective column.

Fixed outliers in few columns.

Categorical variables and the target variable are encoded using Label encoder

Numerical variable is pre-processed using scale

Then the data is given to logistic regression and the accuracy percentage is 75%

The accuracy level for Random forest is 87.5%

When I tried Neural Network with four hidden layers (600,500,400,300), I got 100% accuracy.