**CAPSTONE HEATHCARE PROJECT**

1. Devise strategies for model building. It is important to decide the right validation framework. Express your thought process. Would Cross validation be useful in this scenario?
2. Apply an appropriate classification algorithm to build a model. Compare various models with the results from KNN.
3. Create a classification report by analysing sensitivity, specificity, AUC(ROC curve) etc. Please try to be as descriptive as possible to explain what values of these parameter you settled for? any why?

***This week we selected couple of models to train and predict.***

***Logistic Regression and Decision Tree Classifier models were used to further analyze the dataset.***

***Here are the results of the same***

***Naive Bayes: 73.78106821787414***

***K-Nearest Neighbour: 67.42728714965628***

***Support Vector Machine-linear: 75.88577472236912***

***Support Vector Machine-rbf: 64.01639344262296***

***Ranom Forest: 72.47752511898466***

***eXtreme Gradient Boost: 74.75409836065575***

***Logistic Regression gave us a score of 82% accuracy***

***Where as Decision tree classifier yielded aroung 76% of accuracy***

***From the AUC Roc analysis we gathered that following results –***

***Threshold of 0.5 is used by default (for binary problems) to convert predicted probabilities into class predictions***

***Threshold can be adjusted to increase sensitivity or specificity***

***Sensitivity and specificity have an inverse relationship***

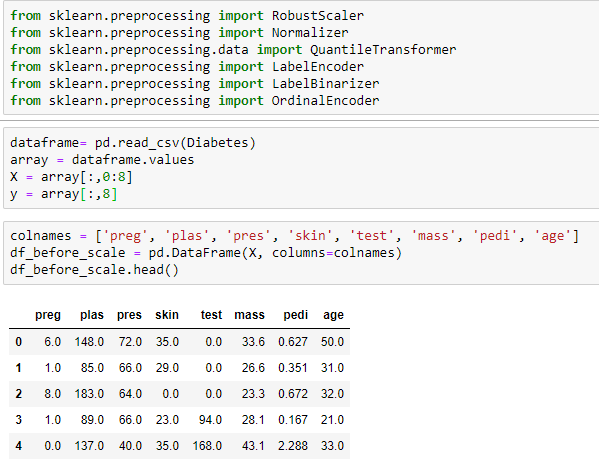
***Increasing one would always decrease the other***

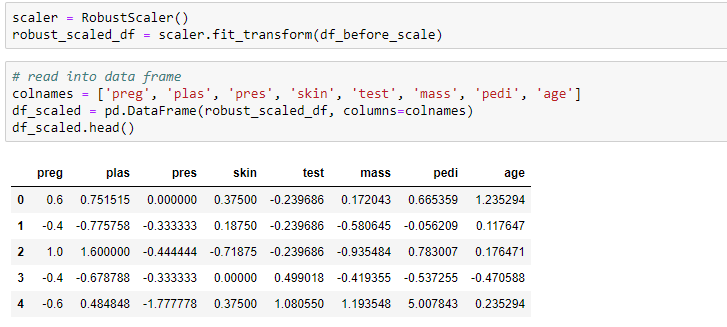
***Adjusting the threshold should be one of the last step you do in the model-building process***

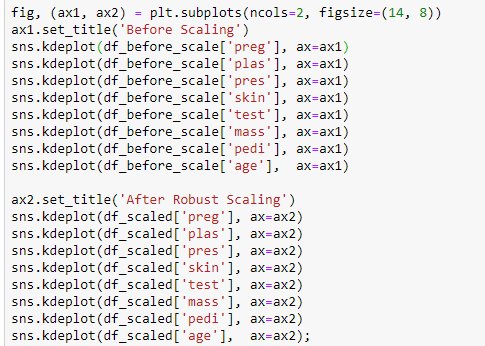
***The most important steps are***

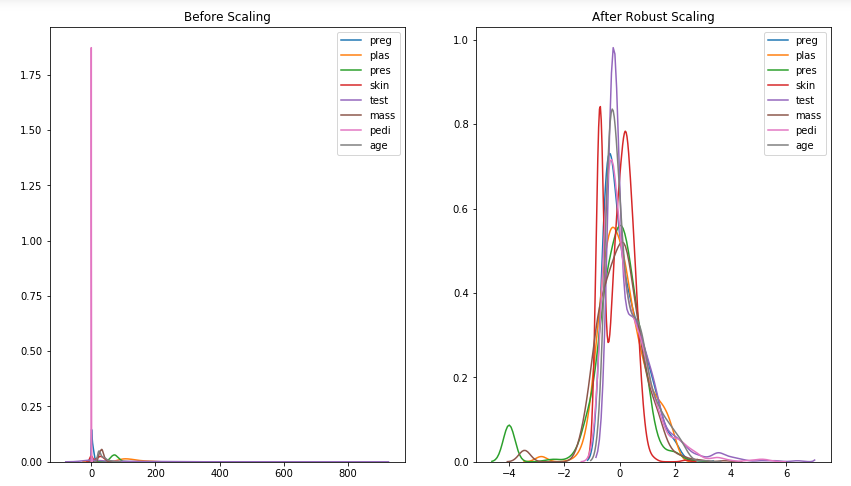
***Building the models***

***Selecting the best model***

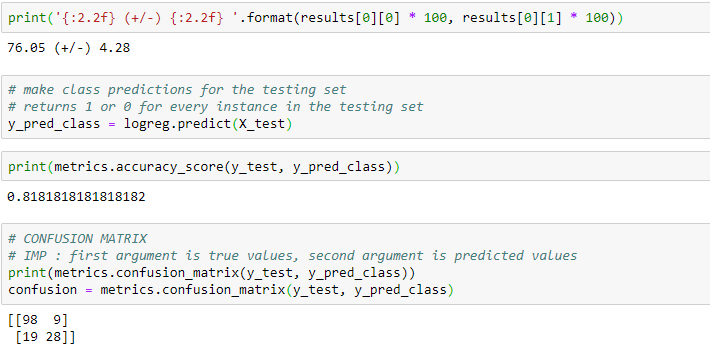


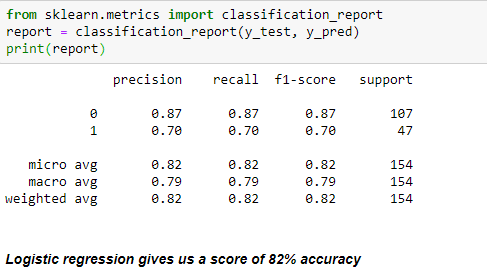


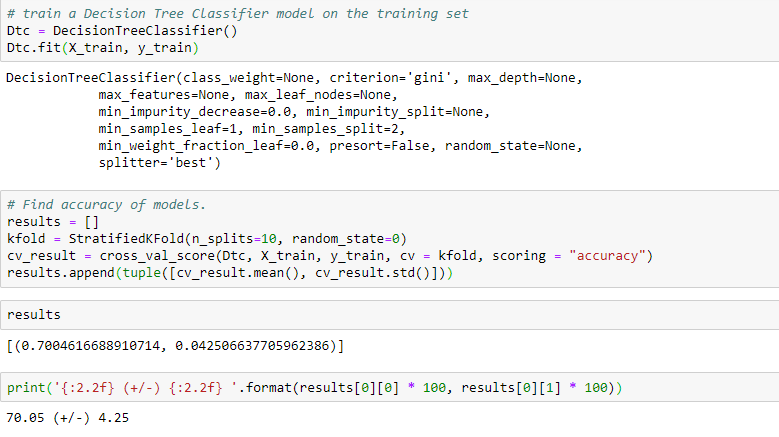


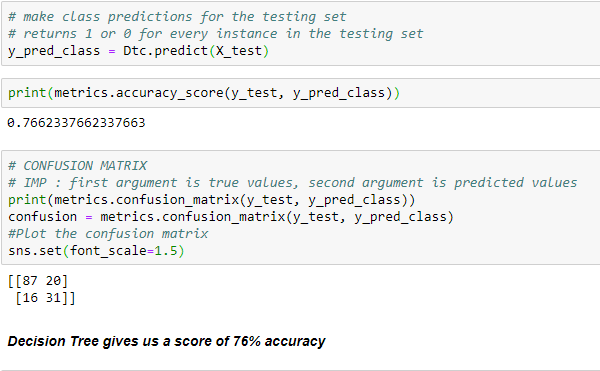




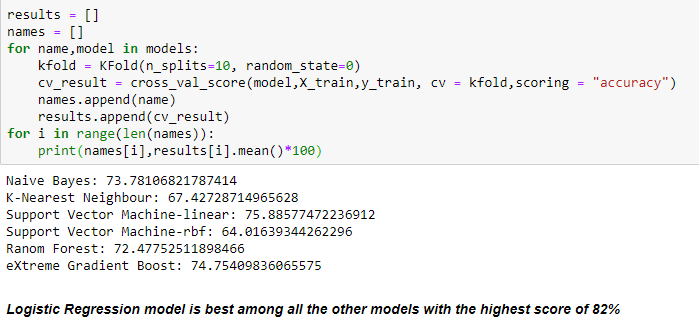












Analysing sensitivity, specificity AUC(ROC) curve

