Title: Implementation of different types of classifiers.

Objective/Aim: To implement

1. Regression classifier.
2. Naïve Bayesian Classifier.
3. k-NN classifier (Take k = 1,3,5,7)
4. Three layer Artificial Neural Network (ANN) classifier (use back propagation). Plot error graph (iteration vs error).

Evaluate performance of above classifiers using metrics

1. Recognition rate
2. Misclassification rate
3. Sensitivity
4. Specificity
5. Precision & Recall

Theory/Algorithm:

1. Regression classifier: Logistic Regression is a 'Statistical Learning' technique categorized in 'Supervised' Machine Learning (ML) methods dedicated to 'Classification' tasks. A contradiction appears when we declare a classifier whose name contains the term 'Regression' is being used for classification, but this is why Logistic Regression is magical: using a linear regression equation to produce discrete binary outputs
2. Naïve Bayesian Classifier: In statistics, naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong (naïve) independence assumptions between the features. Naïve Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem.
3. k-NN classifier: In statistics, the k-nearest neighbors algorithm (k-NN) is a non-parametric classification method first developed by Evelyn Fix and Joseph Hodges in 1951,[1] and later expanded by Thomas Cover.[2] It is used for classification and regression. In both cases, the input consists of the k closest training examples in a data set. The output depends on whether k-NN is used for classification or regression.
4. Three layer Artificial Neural Network (ANN) classifier (use back propagation). Plot error graph (iteration vs error): Artificial neural networks (ANNs), usually simply called neural networks (NNs), are computing systems inspired by the biological neural networks that constitute animal brains.
5.

Authors: 2018BTECS00050
2018BTECS00064

6. An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons.

Performance metrics: Once we build Confusion Matrix, we get True Positive(TP), False Positive(FP), False Negative(FN), True Negative(TN) Values.

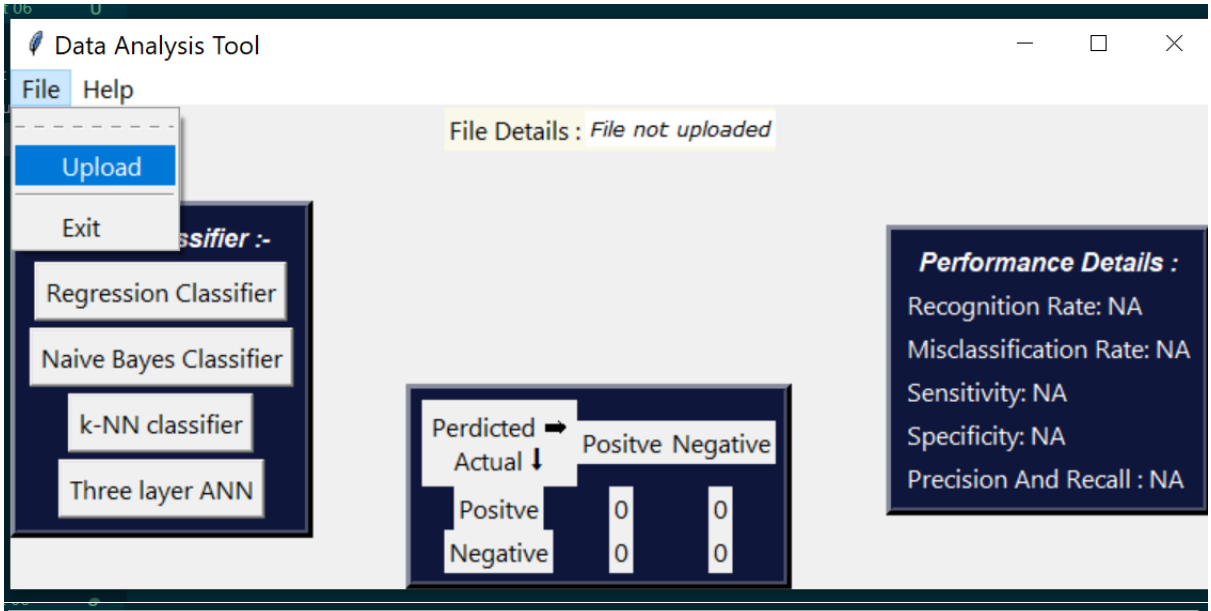
1. Recognition rate: TP + TN/ P+N
2. Misclassification rate: FP+FN/P+N
3. Sensitivity: TP/P
4. Specificity: TN/N
5. Precision: TP/TP+FP

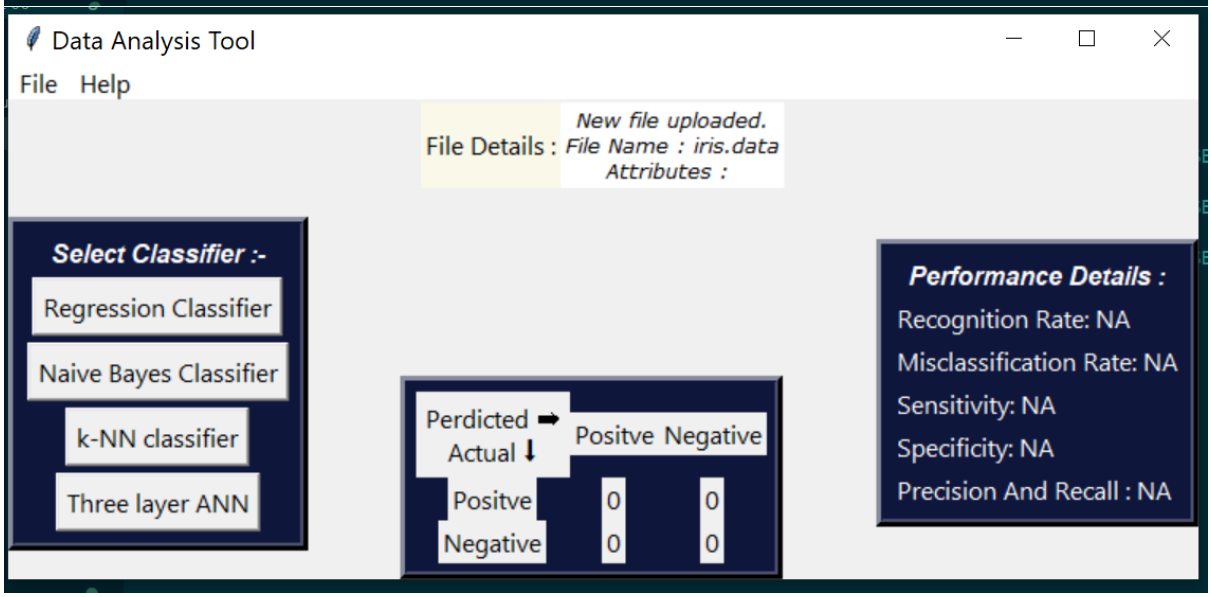
Actual Experimentation/ simulation/ result/ Observation:

Screenshots of the GUI:

1. 

Authors: 2018BTECS00050
2018BTECS00064

2. 

3. 

Conclusion: In this assignment we learnt:

1. How to build a dynamic gui using tkinter
2. For to parse .xcel, .csv files
3. Building Classifiers of given type.

References:

1. https://en.wikipedia.org/wiki/Artificial_neural_network
2. https://towardsdatascience.com/
3. https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm
4. https://en.wikipedia.org/wiki/Naive_Bayes_classifier

Authors: 2018BTECS00050
2018BTECS00064

Authors: 2018BTECS00050
2018BTECS00064