IT1244 Project Instructions

1. There are two colab files: 'IT1244 Team 5 Project Code Using Selected Features.ipynb' and 'IT1244 Team 5 Project Code Using All Features.ipynb'.
2. Open 'IT1244 Team 5 Project Code Using Selected Features.ipynb'.
3. Upload the dataset, 'data.csv' into the colab file.
4. There are 8 main sections: Load Dataset, Filtering Out Relevant Data from the Dataset, Standardising Data, Confusion Matrix Visualizer, Split Dataset into Training and Testing Sets, Logistic Regression Model, k-nearest Neighbors Model and Neural Network Model.
5. Run the code for each section in order.
6. Each model will be trained using the training set and subsequently tested on the testing set.
7. For each model evaluation, a visual representation of the confusion matrix after 10-fold cross-validation as well as the average accuracy, precision, and recall scores will be printed.
8. Repeat Steps 3 to 7 on 'IT1244 Team 5 Project Code Using All Features.ipynb'.
9. For this colab file, there is one extra section called Varying Threshold to Optimise Recall Score. Run the code under this section.
10. A graph will be printed showing the precision, recall and accuracy scores for each of the three models.