**CI213 - Development of a classification algorithm**

**Task 1: Importing the modules**In order to develop a classification algorithm, I imported from the Keras and Sklearn libraries. Then I loaded in the dataset and split the CSV data into input (X) values and the output (Y) value. Essentially, the X values are the data that will be trained on to determine whether the output Y is a spam email or not.

**Task 2: Set up cross fold validation and Neural Network**Afterwards, the data is split into 5 separate ‘folds’ for training and testing. An array of emails is created to hold the accuracy scores generated by these folds. I then created a sequential Neural Network model, with 3 layers. An input layer with neurons equal to the number of features (57), an output layer with only one neuron and a single hidden layer with neurons equal to the mean of the input and output layers. To prevent overfitting of the model, I also installed a dropout rate of 0.2 to randomly remove 20% of the inputs as they traverse between the input and hidden layer.

**Task 3: Compile Neural Network**I then compiled the network, specifying *binary\_crossentropy* as the loss function to minimize, as it is designed for binary outputs. I set the optimizer to the default ‘adam’ and specified the accuracy metric as the visual data between epochs, to check the progress of the training.

**Task 4: Train Neural Network and print accuracy scores**The training data is then loaded into the model, with the number of epochs and batch size set. The model is trained on the 5 folds of data, producing accuracy scores that will then be averaged to give a more robust accuracy score that is then printed to the console.

**Side Note:** Because of the use of K-fold cross validation on top of a neural network, it will likely take several minutes to complete running. Your mileage may vary, but just in case here is a list of my outputs from my last execution of the program.

K-fold 1  
  
K-fold 2  
  
K-fold 3  
  
K-fold 4  
  
K-fold 5  
  
  
Average accuracy  
