Assignment 1A

CAB420 – Machine Learning

By Duane Johnston & Nihar

Contents

[Question 1: Regression 3](#_Toc68532645)

[Question 2: Classification 4](#_Toc68532646)

[Splitting Data 4](#_Toc68532647)

[Part 1: K-Nearest Neighbours Classifier 4](#_Toc68532648)

[Part 2: Random Forrest Classifier 4](#_Toc68532649)

[Part 3: Support Vector Machine Ensemble (SVM) 4](#_Toc68532650)

# Question 1: Regression

# Question 2: Classification

## Splitting Data

To split the testing data, the train\_test\_split method from SK Learn was used to split the test data in half to create a validation set for finding the optimal parameters. This splits our training data so that we have 198 samples for training, 162 samples for validation and 163 samples for testing. This is ok because the training set does contain more samples than the test and validation but not ideal. The set should contain should be more values

## Part 1: K-Nearest Neighbours Classifier

## Part 2: Random Forrest Classifier

## Part 3: Support Vector Machine Ensemble (SVM)