

## **ITCS 6156: Machine Learning**

### **Homework 1**

Student Name: Rajdeep R Rao

Student ID: 800972470

a) Explanation of status and stopping point, if incomplete.

**Answer:**

Naïve Bayes: I have adhered to the given Instructions regarding the format of code, but however, I have only implemented it for a dataset of the first 500 documents since computing it for the entire dataset was a long time.

Decision Tree: I have programmed decision tree train and predict functions as per the given template, but again, for the first 500 documents.

Cross Validation: I have programmed cross validation for both decision tree and naïve Bayes again, for the first 500 documents.

Information Gain: I have partially computed the information gain for decision tree in its entirety. I could finish it for 500 documents, but since I ran out of time when I was converting it for all the documents, It's kind of broken. Since I had no backup of the code for 500 documents, I've had to upload it like that. I would recommend you run the code partially, all the way up till the infoGain part, since that produces the output correctly.

Plotting of Cross Validation: This has not been implemented.

Report has been written and submitted with the assignment.

b) Explanation of additional functions and analysis, if any.

**Answer:**

None

c) Discuss the easy and challenging parts of the assignment. How did you overcome all or some of the challenges?

**Answer:**

The biggest challenge, I faced was to understand what exactly needs to be done. I faced lot of difficulty in developing the flow to adopted, since this was my first project involving so much of data and also converting the data into suitable formats to work with.