Analysis Report

1 Using Navie bayes predict whether a given email is spam/ham.

Ans. First we split the data into 30:70 ratio for test and train respectively. Then calculated the total number of spam/ham emails to find out probability of spam and ham in the entire training dataset.Bag of words of email was also made. Using bayes rule probabilty of the given mail to be spam/ham was calculated. Because of Navie baye's we were able to apply chain rule of differentiation. Naive baye's for this case means that every occurence of word in the email is independent of the other. Stop words were removed form the mails for better effeciency of the code. Logarithm was used since various multiplication of less than one probabilities could lead the computer to ignore significant figures. Accuracy of 86 percent was achieved on the test data.

2 Using Baye's Decision Rule classify whether a given pixel location belongs to river or non river class and output black and white pixel corresponding to river and non-river area.

Ans. First two text files were made in which the random 50 and 100 pixel location of river and non-river were saved with x,y co-ordinates. Using these pixel location we calculated whether a pixel location belongs to river or non-river class. Calculating baye's descision rule was done in two parts, fisrt we calculated the class conditional density function which was broken into several steps and in second part we prior probailities of river and non-river class were taken. Using the prior probability and class coditional density function whether a pixel belong to river/non-river were calculated by comparing two probabilities. White pixel was printed corresponding to river class and black pixel corresponding to non-river class.We repeated the above with three prior probabilities value.