**NAIVE-BAYES**

1. The first data set contains the text data .It contains all the Ham, Spam e-mails . So as it is a text data ,I have used tokenization, stemming, lemmatization, stop words, removing white spaces, removing punctuations, removing numbers . AS the data is cleaned I have applied the naive-bayes classifier on the structured data. As there are no graphs regarding this Module ,So Iam just explaning how I have done My assignment in brief.
2. This salary prediction data consists of around 14 columns , Based on the persons Qualification, Job,city, Relationship status, race, native , upskill ,down skill .We need to predict wether he will get greater than 50k or less than 50k . So I have done the preprocessing like creating dummy variables , Cateogary encoder, checking the nulls . So in this data set I have removed the race and relationship status for predicting the income of a person .I personally didn’t fell these 2 features are important. As there are no graphs regarding this Module ,So Iam just explaning how I have done My assignment in brief.