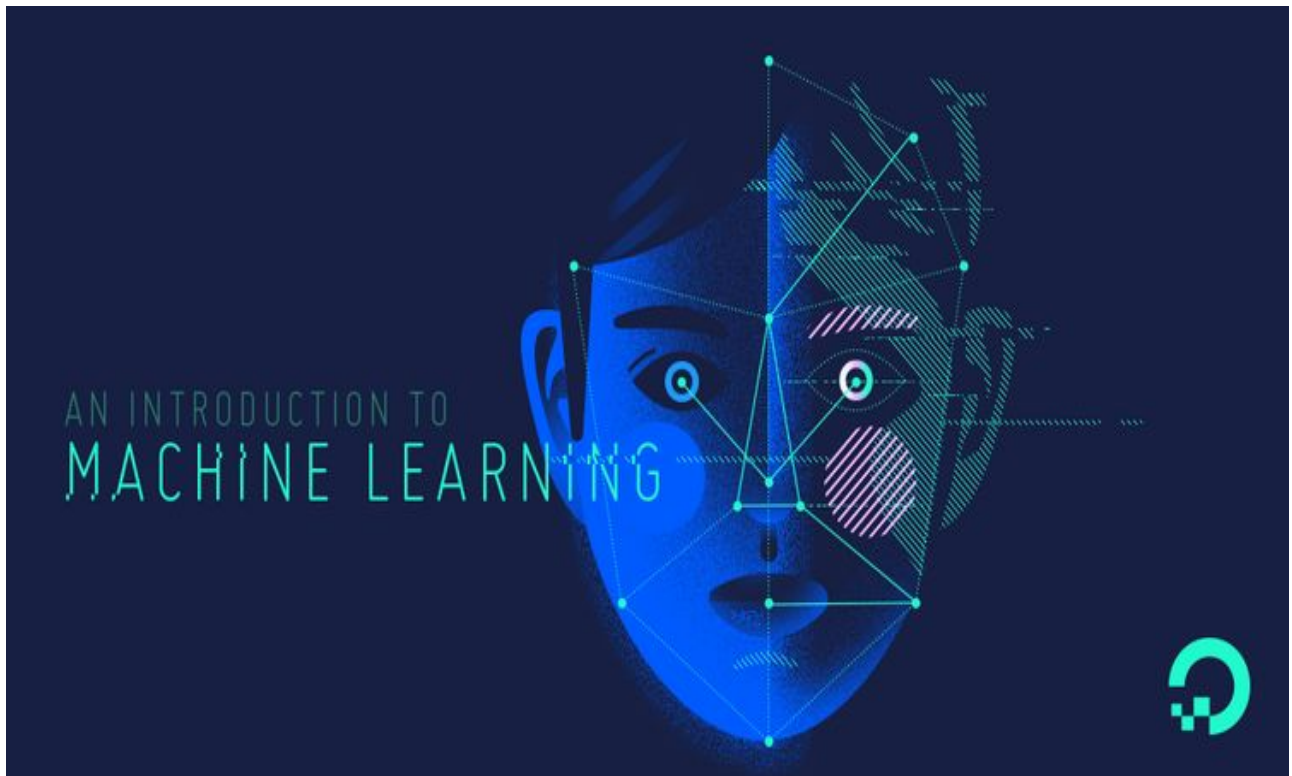# Email Classification, River & Non-River Classification (Using Naive Bayesian Classifier, Assignment # 5)

Amarjeet Yadav

MIT2018053

1. Email Classification using Naive Bayesian Classifier

1. I . 70 % Training and 30 % Testing


```
0     go jurong point, crazy.. avail bugi n great wo...
1                          ok lar... joke wif u oni...
2     free entri 2 wkli comp win fa cup final tkt 21...
3            u dun say earli hor... u c alreadi say...
4              nah think goe usf, live around though
5     freemsg hey darl 3 week' word back! i'd like f...
6     even brother like speak me. treat like aid pat...
7     per request 'mell mell (oru minnaminungint nur...
8     winner!! valu network custom select receivea å...
9     mobil 11 month more? u r entitl updat latest c...
10    i'm gonna home soon want talk stuff anymor ton...
11    six chanc win cash! 100 20,000 pound txt> csh1...
12    urgent! 1 week free membership å£100,000 prize...
13    i'v search right word thank breather. promis w...
14                           date sunday will!!
15    xxxmobilemovieclub: use credit, click wap link...
16                         oh k...i'm watch here:)
17    eh u rememb 2 spell name... ye did. v naughti ...
18          fine thatåõ way u feel. thatåõ way gota b
19    england v macedonia - dont miss goals/team new...
20                          serious spell name?
```

Emails after preprocessing

Correct predictions out of 1672 is : 1614
Accuracy achieved : 96.5311004784689

2. River and Non-River classification using Naive Bayesian Classifier

1. I . 50 River points and 100 Non-River points taken

Accuracy achieved : 90.98625183105469