Data Science  
Lab Exercise (Naïve Bayes)

1. Initially perform the following three exercises on the following Links:

<https://www.statology.org/k-fold-cross-validation-in-python/>

<https://www.analyticsvidhya.com/blog/2021/01/a-guide-to-the-naive-bayes-algorithm/>

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/>

<https://www.datacamp.com/community/tutorials/naive-bayes-scikit-learn>

1. In the following step, you will be working on building your own Naïve Bayes Classifier.
2. Create a text file from the following data or you may also use any of the datasets used in the above exercises.
3. Load data through Pandas libraries
4. Build your own Naïve Bayes classifier model using steps below
5. Class: P(C) = Nc/N
6. e.g., P(No) = 7/10,   
    P(Yes) = 3/10
7. For discrete attributes:  
      
    P(Ai | Ck) = |Aik|/ Nck

where |Aik| is number of instances having attribute Ai and belongs to class Ck

Examples:

P(Status=Married|No) = 4/7P(Refund=Yes|Yes)=0

1. Normal distribution:

One for each (Ai,ci) pair





For (Income, Class=No):

If Class=No

sample mean = 110

sample variance = 2975

1. Once Trained, test your model for the cases below

X1 = {Refund = Yes, Status = Divorced, Income = 90K, Evade = ?}

X2 = {Refund = No, Status = Married, Income = 60K, Evade = ?}