Write Pseudo Code of Naive Bayes Classifier

Ans:

Input : $D = \{x_1, x_2, \ldots, x_n\}$ // Training Data

Output : A naive bayes model

Method:

```
for each class, Ci ∈ D, do
    find the Prior Probabilities, P(Ci)
end for

for each attribute, Ai ∈ D, do
    for each attribute Value, Aij ∈ Ai, do
        find the class Conditional Probabilities, P(Aij | Ci)
    end for
end for

for each instance, Xi ∈ D, do
    find the Posterior Probability, P(Ci | Xi);
end for
```

Q Write the Major steps of Naive Bayes Classifier

Ans:

The major steps are

a. Find the Prior Probabilities, $P(C_i)$
b. Find the class conditional Probabilities, $P(A_{ij}|C_i)$
c. Find the Posterior Probability, $P(C_i|X_i)$

Q What is Laplace Correction? Why we need Laplace Correction

Ans:

A zero Probability cancels the effect of all other (Posteriori) Probabilities (on $C_i$) involved in the Product. This is known as laplace Correction.

We need Laplace Correction" to conveniently avoid the case of Probability values of zero.

Q Write the advantages & disadvantages of KNN & Naive Bayes Classifier.

Ans:

Advantages:

KNN:

a. No assumptions
b. No training step
c. Easy to implement for multi-class Problem.
d. Only hyper Parameter
e. Variety of distance Criteria.

## Naive Bayes:

a. Easy to use
b. Only one scan of the training data required
c. Handling missing attribute values.
d. Continuous data
e. High classification Performance

## Disadvantages

### KNN:

a. Slow algorithm
b. Dimensonality Constraint
c. Very high run time
d. Missing Value Problem.

### Naive Bayes:

a. Zero frequency Problem
b. Assumption of independent Predictors.
c. Data Scarcity
d. Continuous features.
e. Incomplete training data

Q How to handle big data. Write the advantage of handling big data by Naive Bayes.

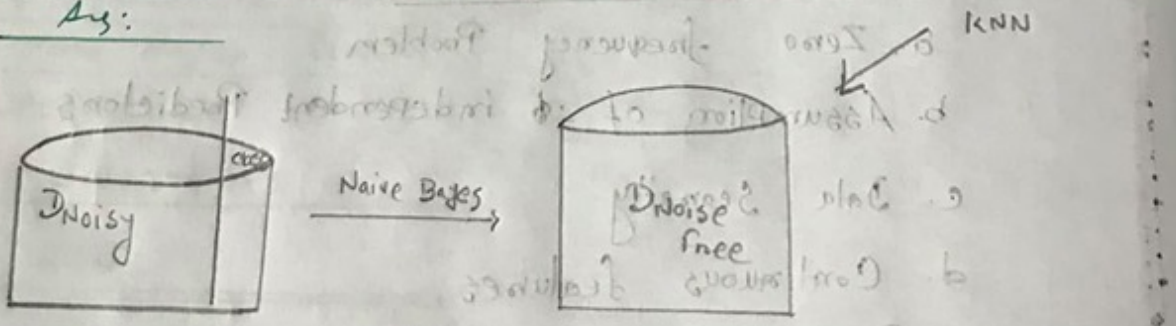
Ans: We can divide a bigset into small Parts and load them.

Naive bayes only requires a small amount of training data to estimate Parameters necessary for classification & the classifier can be trained incrementally.

So for that we can use Gaussian distribution to handle big data.

Q Suppose You have some noisy data. Now write a ~~techin~~ techniaue to make data noise free

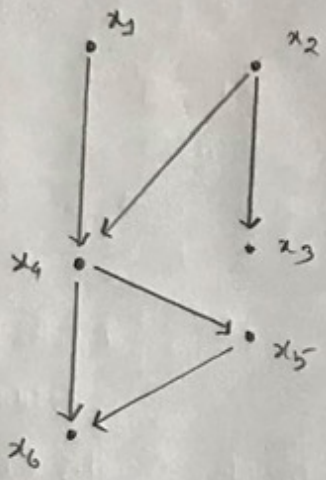
Ans:




Naive Bayes can remove noisy vectors Then with the noise free data KNN would classify.

## From the figure Construct Bayes Net

Ans:




Ans:

$$P(X_6 \mid X_5, X_4, X_3, X_2, X_1) = P(X_6 \mid X_5, X_4)$$

$$P(X_5 \mid X_4, X_3, X_2, X_1) = P(X_5 \mid X_4)$$

$$P(X_4 \mid X_3, X_2, X_1) = P(X_4 \mid X_1, X_2)$$

$$P(X_3 \mid X_2, X_1) = P(X_3 \mid X_2)$$

$$P(X_2 \mid X_1) = P(X_2)$$