BAYES THEOREM

Let X → data sample whose class label is unknown

H → hypothesis is such as that the data sample X belongs to a specified class C

P(H|X) → probability that the hypothesis H holds the given observed data sample X.

It is the posterior prob of H conditioned on X.

Ex. Suppose the ~~void~~ world of data sample consists of fruits described by their color and shape.

Suppose that X is Red and Brown and that H is the hypothesis that X is an apple. Then P(H/X) reflects our confidence that X is an apple given that we have seen that X is Red and Brown.

P(H) - This is the prior probability and it is independent of X.

P(X/H) - This is the posterior probability of X conditioned on H.

Bayes theorem is

$$\boxed{P(H/X) = \frac{P(X/H)\,P(H)}{P(X)}}$$

Q1.

Rid	age	income	student	credit rating	class: buys computer
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	mid	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	mid	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes
12	mid	medium	no	excellent	yes
13	mid	high	yes	fair	yes
14	senior	medium	no	excellent	no

Q10 predict the class label of an unknown sample using Naive Bayesian Classification on the following training dataset from all electronics customer database.
The unknown sample to classify is
X = (age = "youth", income = "medium", student = "yes", credit_rating = "fair")

Ans) Maximise $P(X|C_i)\ P(C_i)$ for $i = 1,2$
$P(C_i)$ is the prior probability of each class.

P (buys_computer = "yes") = 9/14 = 0.643
P (buys_computer = "No") = 5/14 = 0.357.

To Compute $P(X|C_i)$ for $i = 1,2$ we compute following conditional probabilities.

P(age = "youth" | buys_computer = "Yes") = 2/9 = 0.222
P(| buys_computer = "No") = 3/5 = 0.6

P(income = "medium" | buys_computer = "Yes") = 4/9 = 0.444
P(| buys_computer = "No") = 2/5 = 0.4

P(student = "Yes" | buys_computer = "Yes") = 6/9 = 0.667
P(| buys_computer = "No") = 1/5 = 0.2

P(credit_rating = "Yes" | buys_computer = "Yes") = 6/9 = 0.667
P(| buys_computer = "No") = 2/5 = 0.4

using the above probabilities we obtained.

$P(X \mid \text{buys\_computer} = \text{"Yes"})$

$= 0.222 \times 0.444 \times 0.667 \times 0.667$

$= 0.044$

$P(X \mid \text{buys\_computer} = \text{"No"})$

$= 0.6 \times 0.4 \times 0.2 \times 0.4$

$= 0.019$

$P(X \mid \text{buys\_computer} = \text{"Yes"})\, P(\text{buys\_computer} = \text{"Yes"})$

$= 0.044 \times 0.643$

$= 0.028$

$P(X \mid \text{buys\_computer} = \text{"No"})\, P(\text{buys\_computer} = \text{"No"})$

$= 0.019 \times 0.357$

$= 0.007$

Since, 0.028 is greater than 0.007 i.e $0.028 > 0.007$

Therofor, the Naive Bayesian Classifier predicts buys_computer = "Yes" for the sample X.

$\therefore$ X = (age = "youth", income = "medium", student = "Yes", credit-rating = "fair", buys_computer = "Yes")

Q20 predict a class label of an unknown sample X

X = (outlook = "Rain", temp = "cool", humidity = "Normal", wind = "strong", play-football = ?)

Maximise

~~Minimise~~ P(X | Ci) P(Ci) for i = 1,2 P(Ci) is the prior probability of each class.

P(plays-football = "Yes") = 9/14 = 0.643

P(plays-football = "No") = 5/14 = 0.357

To compute P(X|Ci) for i = 1,2 we compute following conditional probability.

P(outlook = "Rain" | play-football = "Yes") = 3/9 = 0.333

P(= "No") = 2/5 = 0.4

P(temp = "cool" | play-football = "Yes") = 3/9 = 0.333

P(= "No") = 1/5 = 0.2

P(humidity = "Normal", play-football = "Yes") = 6/9 = 0.667

P(= "No") = 1/5 = 0.2

P(wind = "strong", play-football = "Yes") = 3/9 = 0.333

P(= "No") = 3/5 = 0.6

using the above probabilities we obtained

P(X | play-football = "Yes") = 0.333 × 0.333 × 0.667 × 0.333

= 0.025

P(X | play-football = "No") = 0.4 × 0.2 × 0.2 × 0.6

$= 9.6 \times 10^{-3}$ = 0.0096

$P(X \mid play\text{-}football = \text{"yes"})\, P(play\text{-}football = \text{"yes"})$

$= 0.025 \times 0.643$

$= 0.016$

$P(X \mid play\text{-}football = \text{"}\overset{No}{\cancel{yes}}\text{"})\, P(play\text{-}football = \text{"No"})$

$= 0.357 \times 0.0096$

$\cancel{0 =}\ 3.4272 \times 10^{-3}$

$= 0.00343.$