Assignment 1

Ans1) - Bayes' theorem states that the conditional probability of an event, based on the occurrence of another event, is equal to the likelihood of the second event given the first event multiplied by the probability of the first event.

$$P(A|B) = \frac{P(A)\,P(B|A)}{P(B)}$$

$P(A|B)$ = probability of how often A happens given that B happens.

$P(B|A)$ = probability of how often B happens given that A happens

$P(A)$ = likelihood of A

$P(B)$ = likelihood of B

Maximum a Posteriori hypothesis is a probabilistic framework for solving the problem of density estimation. It involves calculating a conditional probability of observing the data given a model weighted by a prior probability or belief about the model. It provides an alternate probability framework to maximum likelihood estimation for machine learning. For example a learner considers some set of candidate hypothesis H and is interested in finding the most probable hypothesis $h \in H$ given the observed data D (or at least one of the maximally probable if there are several). Any such maximally probable hypothesis is called a maximum a Posteriori hypothesis

$$h_{MAP} = \underset{h \in H}{\operatorname{argmax}}\, P(h|D)$$

$$= \underset{h \in H}{\operatorname{argmax}}\, \frac{P(D|h)\,P(h)}{P(D)}$$

$$= \underset{h \in H}{\operatorname{argmax}}\, P(D|h)\,P(h)$$

Ans2) - Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle i.e every pair of features being classified is independent of each other. In simple terms, it assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature

Assignment 1

Ans1) - Bayes' theorem states that the conditional probability of an event, based on the occurrence of another event, is equal to the likelihood of the second event given the first event multiplied by the probability of the first event.

$$P(A|B) = \frac{P(A)\,P(B|A)}{P(B)}$$

$P(A|B)$ = probability of how often A happens given that B happens.

$P(B|A)$ = probability of how often B happens given that A happens.

$P(A)$ = likelihood of A

$P(B)$ = likelihood of B.

Maximum a Posteriori hypothesis is a probabilistic framework for solving the problem of density estimation. It involves calculating a conditional probability of observing the data given a model weighted by a prior probability or belief about the model. It provides an alternate probability framework to maximum likelihood estimation for machine learning. For example a learner considers some set of candidate hypothesis H and is interested in finding the most probable hypothesis $h \in H$ given the observed data D (or at least one of the maximally probable if there are several). Any such maximally probable hypothesis is called a maximum a Posteriori hypothesis

$$h_{MAP} = \underset{h \in H}{\operatorname{argmax}}\, P(h|D)$$

$$= \underset{h \in H}{\operatorname{argmax}}\, \frac{P(D|h)P(h)}{P(D)}$$

$$= \underset{h \in H}{\operatorname{argmax}}\, P(D|h)\, P(h)$$

Ans3) - Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle i.e every pair of features being classified is independent of each other. In simple terms, it assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

Ans 4)

Probability of Team 0 wins is $P(Y_0) = 0.95$

Probability of Team 1 wins is $P(Y_1) = 1 - P(Y_0)$
$= 1 - 0.95$
$= 0.05$

Probability that Team 1 hosted the match, it had won is
$P(X_1|Y_1) = 0.75$

Probability that Team 1 hosted the match won by Team 0 is
$P(X_1|Y_0) = 0.30$

We have to calculate the probability of Team 1 winning the ne match if it hosts i.e $P(Y_1|X_1)$

using Bayes Theorem

$$P(Y_1|X_1) = \frac{P(X_1|Y_1) \times P(Y_1)}{P(X_1)}$$

$$= \frac{P(X_1|Y_1) \times P(Y_1)}{P(X_1|Y_1) \times P(Y_1) + P(X_1|Y_0)\, P(Y_0)}$$

$$= \frac{0.75 \times 0.05}{(0.75 \times 0.05) + (0.30 \times 0.95)}$$

$$P(Y_1|X_1) = 0.1162$$

$$\therefore \quad P(Y_0|X_1) = 1 - P(Y_1|X_1) = 0.8838$$

Since $P(Y_0|X_1) > P(Y_1|X_1)$

Team 0 has a better probability of winning match than Team 1

Ans 5) Machine learning, a branch of artificial intelligence, concerns t construction and study of systems that can learn from data. I to develop methods that can automatically detect patterns and then to use the uncovered patterns to predict futur or other outcomes of interests.

Steps for designing learning Systems are :-

① Choosing the training Experience : The very important and firs is to choose the training data or training experience which w fed to the Machine learning Algorithm.

$$P(X_1|Y_1) * P(Y_1) + P(X_1|Y_0)\, P(Y_0)$$

$$= \frac{0.75 \times 0.05}{(0.75 \times 0.05) + (0.30 \times 0.95)}$$

$$P(Y_1|X_1) = 0.1162.$$

$$\therefore \quad P(Y_0|X_1) = 1 - P(Y_1|X_1) = 0.8838$$

Since $P(Y_0|X_1) > P(Y_1|X_1)$

Team 0 has a better probability of wining next match than Team 1

Ans5) Machine learning, a branch of artificial intelligence, concerns the construction and study of systems that can learn from data. It is to develop methods that can automatically detect patterns in data and then to use the uncovered patterns to predict future data or other outcomes of interests.

Steps for designing learning Systems are :-

① Choosing the training Experience : The very important and first task is to choose the training data or training experience which will be fed to the Machine learning Algorithm.

② Choosing ~~Representation~~ ~~for~~ Target function :-
It means according to the knowledge fed to the algorithm the machine learning will choose NextMove function which will describe what type of legal moves should be taken.

③ Choosing Representation for Target function :-
when the machine algorithm will know all possible legal moves the next step is to choose the optimized move using any representation i.e using linear equations, Hierarchical graph etc

④ choosing Function Approximation Algorithm :-
An optimized move cannot be choosen just with the training data. The training data had to go through with set of example and through these ~~exp~~ examples the training data will approximates which steps are chosen after the ML provides feedback on it.

⑤ The final design is created at last when system goes from number of examples, failures and success, correct & incorrect division and what will be the next step.

Ans 6) -
Hypothesis space is described by a conjunction of constraints on the attribute, the constraints may General hypothesis "?"; specific "φ"

what will be the next step.

Ans 6) -

Hypothesis space is described by a conjunction of constraints on the attribute, the constraints may general hypothesis "?", specific hypothesis "φ"

Instance space is a subset of all possible example or instance

Version space is the subset of hypothesis from the consistent - with training example in D.

Ans 8) -

Colour	Type	Origin	Stolen
Red	Sports	Domestic	Yes
Red	sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	sports.	Imported	Yes
Red		Imported	No
			No
			Yes

$$y' = \underset{y \in Y}{\text{argmax}}\, P(Y) \prod_{i=1}^{m} P(x_i \mid Y)$$

(colour | Red, Yellow)

(Type | Sports, SUV)

(Origin | Domestic, Imported)

Target

Values

colour	Yes	No
R	3	2
Y	2	3

$P(\text{Red} \mid \text{stolen}) = \frac{3}{5} = 0.6$

$P(\text{Red} \mid \neg\,\text{stolen}) = \frac{2}{5} = 0.4$

$P(\text{Yellow} \mid \text{stolen}) = \frac{2}{5} = 0.4$

$P(\text{Yellow} \mid \neg\,\text{stolen}) = \frac{3}{5} = 0.6$

Type	Yes	No
Sports	4	2
SUV	1	3

$P(\text{Sports} \mid \text{Yes}) = 4/5 = 0.8$

$P(\text{Sports} \mid \text{No}) = 2/5 = 0.4$

$P(\text{SUV} \mid \text{Yes}) = 1/5 = 0.2$

$P(\text{SUV} \mid \text{No}) = 3/5 = 0.6$

Origin	Yes	No
D	2	3
I	3	2

$P(D \mid \text{Yes}) = 2/5 = 0.5$

$P(D \mid \text{No}) = 3/5 = 0.6$

$P(I \mid \text{Yes}) = 3/5 = 0.6$

$P(I \mid \text{No}) = 2/5 = 0.4$

classify the new data = (Red, SUV, Domestic)

for stolen = Yes :

$P(\text{Yes}) * P(R \mid \text{Yes}) * P(\text{SUV} \mid \text{Yes}) * P(D \mid \text{Yes})$

$= 0.6 * 0.2 * 0.4 * 0.5 = 0.024$

for stolen = NO :

$= P(\text{NO}) * P(R \mid \text{NO}) * P(\text{SUV} \mid \text{NO}) * P(D \mid \text{NO})$

$= 0.4 * 0.6 \times 0.6 \times 0.5 = 0.072.$

Therefore we would classify the data as not stolen.