Chapter -6

1. Three companies X, Y, and Z supply 40%, 45%, and 15% of the uniforms to a school. Past experience shows that 2%, 3%, and 4% of the uniforms supplied by these companies are defective. If a uniform was found to be defective, what is the probability that the uniform was supplied by Company X?

   1. $^{34}/_{69}$

   2. $^{25}/_{55}$

   3. 16/55

   4. $^{16}/_{12}$

   Hint: P(X|D) = P(D|X) P(X) / P(D)

   P(D|X) = 2/100, P(X) = 40/100

   P(D) = P(D|X) * P(X) + P(D|Y) * P(Y) + P(D|Z) * P(Z) = 2/100*4/10 + 3*45/1000+ 4*15/1000

   Ans = 2*4 / (8+13.5 +6) = 16 / (16+26+12) =16/54

2. A box of apples contains 10 apples, of which 6 are defective. If 3 of the apples are removed from the box in succession without replacement,

what is the probability that all the 3 apples are defective?

  1. (6*5*4)/ (10*10*10)

  2. (6*5*4)/ (10*9*8)

  3. (3*3*3)/ (10*10*10)

  4. (3*2*1)/ (10*9*8)

  Hint : P(D1 and D2 and d3) = P(D1) . P(D2).P(D3)

P(D1) = 6/10 , P(D2) = 5/9 , P(D3) = 4/8

3. Two boxes containing chocolates are placed on a table. The boxes are labelled B_1 and B_2. Box B_1 contains 6 dark chocolates and 5 white chocolates. Box B_2 contains 3 dark chocolates and 8 orange chocolates. The boxes are arranged so that the probability of selecting box B_1 is $^1/_3$ and the probability of selecting box B_2 is $^2/_3$. Sneha is blindfolded and asked to select a chocolate. She will win Rs. 10,000 if she selects a dark chocolate. What is the probability that Sneha will win Rs. 10,000 (that is, she will select a dark chocolate)?

  1. $^7/_{33}$

  2. $^6/_{33}$

  3. $^{11}/_{33}$

  4. $^{10}/_{33}$

  Hint : P(B1) = 1/3 , P(B2) =2/3

  P(D|B1) = 6/11, P(D|B2) = 3/11

  P(D) = P(D|B1) * P(B1) + P(D|B2) * P(B2)

  = 6/11 * 1/3 + 3/11 * 2/3 = 6/33 + 6/33 = 12/ 33

  Ans : 12/33

4. Two boxes containing chocolates are placed on a table. The boxes are labelled B_1 and B_2. Box B_1 contains 6 Cadbury chocolates and 5 Amul chocolates. Box B_2 contains 3 Cadbury chocolates and 8 Nestle chocolates. The boxes are arranged so that the probability of selecting box B_1 is $^1/_3$ and the probability of selecting box B_2 is $^2/_3$. Sneha is blindfolded and asked to select a chocolate. She will win Rs. 10,000 if she selects a Cadbury chocolate. If she win Rs 10,000, what is the probability that she selected a Cadbury chocolate from the first box?

  1. $^7/_{12}$

2. $^6/_{12}$

3. $^{11}/_{12}$

4. $^{10}/_{33}$

Hint : P(B1) = 1/3 , P(B2) =2/3

P(C|B1) = 6/11, P(C|B2) = 3/11

P(C) = P(C|B1) * P(B1) + P(C|B2) * P(B2)

= 6/11 * 1/3 + 3/11 * 2/3 = 2/11 + 1/22 = 5/ 22

P(B1|C) = P(C|B1) * P(B1) / P(C) = (6/11 * 1/3) / (12/33) = 6/12

5. In a certain basketball club, there are 4% of male players who are over 6 feet tall and 1% of female players who are over 6 feet tall. The ratio of male to female players in the total player population is male:female =

2:3. A player is selected at random from among all those who are over 6 feet tall. What is the probability that the player is a female?

1. 3/11

2. 2/5

3. 2/11

4. 1/11

Hint: P(6|M) = 4/100, P(6|F) = 1/100   P(M) = 2/5 , P(F) = 3/5

P(6) = P(F)* P(6|F) + P(M)* P(6|M) = 3/5 * 1/100 + 2/5 * 4/100 = 11/500

P(F|6) = P(6|F) * P(F) / P(6) = 3/11

6. The probability that a particular hypothesis holds for a data set based on the Prior is called
   1. Independent probabilities
   2. Posterior probabilities
   3. Interior probabilities
   4. Dependent probabilities
      Hint: 2 Posetrior

7. One main disadvantage of Bayesian classifiers is that they utilize all available parameters to subtly change the predictions.
   1. True
   2. False

8. In a bolt factory, machines A1, A2, and A3 manufacture respectively 25%, 35%, and 40% of the total output. Of these 5%, 4%, and 2% are defective bolts. A bolt is drawn at random from the product and is found to be defective. What is the probability that it was manufactured by machine A2?

1. 0.0952

2. 0.452

3. 0.952

4. 0.125

Hint :P(A1) = 25/100, P(A2) = 35 /100, P(A3) = 40/100

P(D|A1) = 5/100, P(D|A2) = 4/100, P(D|A3) = 2/100

P(A2|D) = P(D|A2) * P(A2) / P(D) = 4*35/345=0.405

P(D)= 5/100 * 25/100 + 4/100 *35/100 + 2/100 *40/100 = 125+140+80/10000

9. Bayesian methods can perform better than the other methods while validating the hypotheses that make probabilistic predictions.
   1. True
   2. False

10. Naïve Bayes classifier makes the naïve assumption that the attribute values are conditionally dependent given the classification of the instance.
    1. True
    2. False

Chapter -7

1. Predicting whether a tumour is malignant or benign is an example of?
   1. Unsupervised Learning
   2. Supervised Regression Problem
   3. Supervised Classification Problem
   4. Categorical Attribute
   Hint: We need previous records for past patient data and categorize them into different classes
2. Price prediction in the domain of real estate is an example of?
   1. Unsupervised Learning
   2. Supervised Regression Problem
   3. Supervised Classification Problem
   4. Categorical Attribute
Hint: We need past records of prices and a prediction model

17. The distance between hyperplane and data points is called as:
    1. Hyper Plan

2. Margins
3. Error
4. Support Vectors

Hint : Margins

25. Which of the following will be Euclidean distance between the two datapoints A(4,3) and B(2,3)?

1. 1

2. 2

3. 4

4. 8

Hint: P1(x1,y1) and P2(x2,y2) is square root of (x2-x1)2 + (y2-y1)2

18. Which of the following will be Manhattan distance between the two data points A(8,3) and B(4,3)?

1. 1

2. 2

3. <mark>4</mark>

4. 8

Hint : |x2-x1| + |y2-y1|

19. When you find many noises in data, which of the following options would you consider in kNN?
    1. Increase the value of k
    2. Decrease the value of k
    3. Noise does not depend on k
    4. $K = 0$
       Hint When K is less the class value of noisy data or outlier in the training data set which is nearest neighbour to the test data will be assigned to test data

20. What would be the relationship between the training time taken by 1-NN, 2-NN, and 3-NN?
    1. 1-NN > 2-NN > 3-NN

    2. <mark>1-NN < 2-NN < 3-NN</mark>

    3. 1-NN ~ 2-NN ~ 3-NN

    4. None of these

    B.

21. Which of the following algorithms is an example of the ensemble learning algorithm?
    <mark>1. Random Forest</mark>
    2. Decision Tree
    3. kNN
    4. SVM
       Hint.1

22. Which of the following is not an inductive bias in a decision tree?
    1. It prefers longer tree over shorter tree
    2. Trees that place nodes near the root with high information gain are preferred
    3. Overfitting is a natural phenomenon in a decision tree
    4. Prefer the shortest hypothesis that fits the data

Chapter – 8

1. Price prediction of crude oil is an example of
   1. Unsupervised learning

2. Supervised regression problem
3. Supervised classification problem
4. Categorical attribute

5. If the regression involves only one independent variable, it is called as
    1. Multiple regression
    2. One regression
    3. Simple regression
    4. Independent regression
        Hint : Y = aX + b
6. Which equation represents simple imperfect relationship?
    1. $Y = (a + bx) + \varepsilon$
    2. $Y = (a + bx)$
    3. DY = Change in Y / Change in X
    4. $Y = a + b1X1 + b2X2$

7. What is the formula for slope in a simple linear equation?
    1. $Y = (a + bx) + \varepsilon$
    2. $Y = (a + bx)$
    3. DY = Change in Y / Change in X
    4. $Y = a + b1X1 + b2X2$

8. This slope always moves upward on a graph from left to right.
    1. Multilinear slope
    2. No relationship slope
    3. Negative slope
    4. Positive slope
        Hint : y2-y1 / x2-x1   here y2-y1 and x2-x1 are both postive value

9. This slope always moves downwards on a graph from left to right.

1. Multilinear slope
2. No relationship slope
3. Negative slope
4. Positive slope
   Hint : y2-y1 / x2-x1   here y2-y1 is negative value

10. Maximum and minimum points on a graph are found at points where the slope of the curve is
    1. Zero
    2. One
    3. 0.5
    4. Random number

Chapter -9

1. k-means clustering algorithm is an example of which type of clustering method?
   1. Hierarchical
   2. Partitioning
   3. Density based
   4. Random

2. Which of the below statement describes the difference between agglomerative and divisive clustering techniques correctly?
   1. Agglomerative is a bottom-up technique, but divisive is a top-down technique
   2. Agglomerative is a top-down technique, but divisive is a bottom-up technique
   3. Agglomerative technique can start with a single cluster
   4. Divisive technique can end with a single cluster

3. Which of the below is an advantage of k-medoids algorithm over k-means algorithm?
   1. both are equally error prone
   2. k-medoids can handle larger data set than k-means
   3. k-medoids helps in reducing the effect of the outliers in the objects
   4. k-medoids needs less computation to arrive at the final clustering

4. The principle underlying the Market Basket Analysis is known as
   1. Association rule
   2. Bisecting rule
   3. *k*-means
   4. Bayes' theorem
5. A Voronoi diagram is used in which type of clustering?
   1. Hierarchical
   2. Partitioning
   3. Density based
   4. Intuition based

Chapter -10

1. A binary sigmoid function has range of____.
   1. $(-1, +1)$
   2. $(0, -1)$

   3. $(0, 1)$

   4. $(1, 0)$

   Hint: y out = $1/1+e^{-kx}$

2. A bipolar sigmoid function has range of____.
   1. $(-1, +1)$
   2. $(0, -1)$

   3. $(0, 1)$

   4. $(1, 0)$

   Hint: y out = $1-e^{-kx}/1+e^{-kx}$

3. The inputs of the McCulloch–Pitts neuron could be____.
   1. either -1 or 1
   2. either 0 or 1
   3. either 0 or -1
   4. none of the above
4. Single-layer perceptron is able to deal with
   1. linearly separable data
   2. non-linearly separable data
   3. linearly inseparable data
   4. none of the above
5. Single-layer feed forward network consists of____layers.
   1. two
   2. one
   3. three
   4. many

6. Multi-layer feed forward network consists of____layers.
   1. two
   2. one
   3. three
   4. many
7. In the backpropagation algorithm, multiple iterations are known as
   1. degree
   2. epoch
   3. cardinality
   4. none of the above
8. Each epoch consists of phases –
   1. forward
   2. backward
   3. both a and b
   4. none of the above
9. Deep neural networks generally have more than____hidden layers.
   1. 1
   2. 2

   3. 3

   4. none of the above

10. To handle intense computation of deep learning,____is needed.
    1. parallel computing
    2. CPU-based traditional computing
    3. GPU computing
    4. none of the above

Chapter-11

1. In representation learning, when unknown causes (H) are combined, it can also generate the data (V), it is called as ____

    1. Recognition
    2. Generation
    3. Representative
    4. Combinative

2. When the causes (H) explains the data (V) we observe, it is called as ___

    1. Recognition
    2. Generation
    3. Representative
    4. Combinative

3. When unknown causes (H) are combined, it can also generate the data (V), it is called as ____

    1. Recognition
    2. Generation
    3. Representative
    4. Combinative

5. Independent component analysis is similar to

    1. Dependent analysis
    2. Sub Component analysis
    3. Data analysis
    4. Principal component analysis