Note: Question No One is Compulsory. Attempt any three from remaining questions.

Que 1 a) Define Machine Learning and Explain with example importance of Machine Learning.(5)

Ans 1 definition 2 marks

Importance with examples 3 marks

b) Explain Multilayer perceptron with a neat diagram. (5)

Answer : MLP explanation 3 marks

Labelled diagram 2 marks

c) Why is SVM more accurate than logistic regression?(5)

Answer: SVM try to maximize the margin between the closest support vectors while LR the posterior class probability. Thus, SVM find a solution which is as fare as possible for the two categories while LR has not this property.

LR is more sensitive to outliers than SVM because the cost function of LR diverges faster than those of SVM.

Logistic Regression produces probabilistic values while SVM produces 1 or 0. So in a few words LR makes not absolute prediction and it does not assume data is enough to give a final decision. This maybe be good property when what we want is an estimation or we do not have high confidence into data.

d) Explain Radial Basis Function with example.(5)

Answer: In Single Perceptron / Multi-layer Perceptron(MLP), we only have linear separability because they are composed of input and output layers(some hidden layers in MLP)

☐   RBNN what it does is, it transforms the input signal into another form, which can be then feed into the network to get linear separability. ☐ RBNN is structurally same as perceptron(MLP). RBNN is composed of input, hidden, and output layer. RBNN is strictly limited to have exactly one hidden layer. We call this hidden layer as feature vector. ☐ RBNN increases dimension of feature vector.

Gaussian Radial Function := $\varphi(r) = \exp(-r^2/2\sigma^2)$ where $\sigma > 0$

Example. XOR function

Que 2 a) What is Dimensionality reduction? Describe how Principal Component Analysis is carried out to reduce dimensionality of data sets.   **10**

Answer: Dimensionality Reduction explanation 5 marks

PCA steps 5 marks

b) Find the singular value decomposition of

$$A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$$ **10**

Answer: $U = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ $V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ $S = \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$

Que 3 a) For a unknown tupke t =<Outllook =Sunny,Temperature =Cool, Wind= Strong> use naïve Bayes classifier to find whether the class for PlayTennis is yes or no. The dataset is given below   10

Answer : P(Outlook=*Sunny*|Play=*Yes*) = 2/9

P(Temperature=*Cool*|Play=*Yes*) = 3/9

P(Wind=*Strong*|Play=*Yes*) = 3/9

P(Play=*Yes*) = 9/14

P(Outlook=*Sunny*|Play=*No*) = 3/5

P(Temperature=*Cool*|Play==*No*) = 1/5

P(Wind=*Strong*|Play=*No*) = 3/5
P(Play=*No*) = 5/14
MAP Rule
P(*Yes*|x'): [P(*Sunny*|*Yes*)P(*Cool*|*Yes*) P(*Strong*|*Yes*)]P(Play=*Yes*) = 0.0158
P(*No*|x'): [P(*Sunny*|*No*) P(*Cool*|*No*) P(*Strong*|*No*)]P(Play=*No*) = 0.0257
Given the fact P(*Yes*|x') < P(*No*|x'), we label x' to be "*No*".

b)List some advantages of derivative-based optimization techniques. Explain Steepest   **10**
Descent method for optimization.
                                                              5 marks
Answer: Advantages   5 marks
        Steepest Descent method

Que 4 a) Given the following data for the sales of car of an automobile company for six   **10**
consecutive years. Predict the sales for next two consecutive years.

Years	2013	2014	2015	2016	2017	2018
Sales	110	100	250	275	230	300

Answer: b_0 = -78393.67, b_1= 39 . when x=2019 sales 347 and when x=2020 sales = 386
b) Explain various basic evaluation measures of supervised learning Algorithm for   **10**
Classification.
Answer: Expected Definition of Confusion matrix, accuracy,Precision,Recall/sensitivity, F score each
2 marks
Que 5 a) Consider following table for binary classification. Calculate the root of the
decision tree using gini index.   **10**

Answer: Gini Index for overall example is 0.460 (2marks)

The Gini index for the Customer Income is 0.449 (2marks)
The Gini index for gender is 0.3675 (2marks)
The Gini index for car type is 0.4284 (2marks)

Gender will be the root. As it has minimum Gini index (2marks)

b) Define Support Vector Machine. Explain how margin is computed and optimal hyperplane
is decided. 10

Answer: SVM definition with goals 2 marks
        Margin calculation with plane $g(x) = W^T x+b$ equation 6 marks
        Optimal Hyperplane 2 marks

                                                                        20

Que 6 Write Short notes on any two
    a) Hidden Markov Model
    b) EM Algorithm
    c) Logistic Regression
    d) McCulloch-Pitts Neuron Model