Total Marks: 80

(3 Hours)

N.B. (1) Question No. 1 is compulsory.
(2) Attempt any three from the remaining five questions.
(3) Answers to questions should be grouped and written together.

Q.1 (a) Explain classification with logistic regression and sigmoid function. [10]

(b) Explain implementation of classification can be improved with AdaBoost algorithm. [10]

Q.2 (a) Consider following data of **buying computer** and classify a tuple [10]
X= (age <=30, Income = medium,Student = yes, Credit_rating = Fair)
Using Bayesian classifier.

Age	Income	Student	Credit-rating	buy-
<30	high	no	fair	no
<=30	high	no	excellent	no
31….40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31-40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31-40	high	yes	fair	yes
31-40	medium	no	excellent	yes
>40	medium	no	excellent	no

(b) Explain the various steps in developing Machine learning application [10]

Q.3 (a) Construct Decision tree based on ID3 for the following training data [10]

Day	Outlook	Temp.	Humidity	Wind	Decision
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

(b) Describe working of support vector machine and the calculation of maximum margin [10]

Q.4 (a) Explain agglomerative clustering with suitable algorithm. [10]

Explain the following terms

(b) [10]

i. Bias
ii. Error
iii. Accuracy
iv. Variance
v. Dimensions

Q.5 (a) Describe principal component analysis and its importance. [10]

(b) What is clustering? Explain K means clustering algorithm. Explain K-Means clustering algorithm. Using K-means clustering, cluster the following data into two clusters and show each step. {2, 4, 10, 12, 3, 20, 30, 11, 25} [10]

Q.6 Write short note on (**Attempt any Four)** [20]

(a) Recommender systems
(b) Applications of machine learning
(c) Anomaly detection
(d) Big data analysis
(e) Supervised VS unsupervised learning

****