

(Time: 3 Hours)

(Total Marks: 80)

N. B:

1. Question No. 1 is Compulsory.
2. Solve any **THREE** from Question No. 2 to 6.
3. Draw neat well labeled diagram wherever necessary

- Q. 1 a)** What is the significance of machine learning techniques in pattern recognition? (05)
- b)** How maximum likelihood Hypothesis can be used for predicting probabilities? (05)
- c)** Differentiate between training and testing data. (05)
- d)** State and explain similarity and distance measures. (05)
- Q. 2 a)** What do you understand by a well posed learning problem? Explain the important features that are required to well define a learning problem. (10)
- b)** What are the issues in Decision Tree Learning? Demonstrate overfitting the data in Decision Tree with suitable example. (10)
- Q. 3 a)** How PCA is used to arrive at the transformed matrix for given matrix? Illustrate with suitable example. (10)
- b)** The Support Vector Machine (SVM) is a highly accurate classification method. However, SVM classifier suffer from slow processing when training with large set of types. Discuss how to overcome this difficulty and develop a scalable SVM algorithm for efficient SVM classification inn large datasets. (10)
- Q. 4 a)** For the following dataset create Decision Tree using ID3 algorithm. (10)

age	income	student	credit_rating	buys_computer
<=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	medium	no	excellent	yes
31...40	high	yes	fair	yes
>40	medium	no	excellent	no

- b)** What is the role of dimension reduction in Machine Learning? State and explain different methods. (10)

- Q. 5 a)** Ensemble of classifiers (EOC) combines a set of diverse classifiers to solve a classification problem. What are the consequences of combination of multiple classifiers? **(10)**
- b)** What are the Performance Evaluation metrics for Classification and clustering problems? **(10)**

Q. 6 Write short notes on any **FOUR**:

1. Temporal difference learning **(20)**
2. Hierarchical clustering
3. Linear Regression
4. Q Learning algorithm
5. Parameter tuning
