TE (ELEX) ML R'19 C-scheme sem VI
26.05.2025

Duration 3 Hours [Maximum Marks 80]

NOTE: -1) Question 1 is compulsory
2) Solve any three from the remaining five questions
3) Assume suitable data if necessary.
4) Figures to the right indicate full marks

Q.1. Attempt **any four** 20
a) What key factors should be taken into account when choosing the most suitable algorithm for a machine learning application?
b) Justify there is dilemma in Bias variance tradeoff for model selection.
c) How do supervised learning and unsupervised learning differ from each other?
d) Explain Logistic model based on feature construction, transformation, and feature selection.
e) What role does a loss function play in machine learning, and why is it crucial for both model training and performance evaluation?

Q.2.a) Explain the Principal Component Analysis (PCA) technique. Discuss its purpose, how it works, and provide an example to illustrate its application in Machine learning. 10
b) Explain the concept of kernels in machine learning. How do kernels help in learning non-linear functions? Discuss the types of kernels commonly used and provide examples to illustrate their application. 10

Q.3 a) How is machine learning applied in credit card fraud detection? Please explain the techniques used, the types of data involved, and the benefits of using machine learning in this context.
b) Explain the different error measures used to evaluate the performance of regression models. Additionally, calculate the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) for the following dataset: 10

Actual Value	Predicted value
4.0	3.5
4.5	4
3.0	3.5
5.0	5.5
3.5	3.0

QP Code
80967

Prog. code
IT01136

Q.4. a) Explain the Expectation Maximization (EM) algorithm in detail. How does it work, and what are its key applications? Additionally, discuss the role of supervised learning after clustering, and how clustering results can be used to improve the performance of a supervised learning model. 10

b) Given the following 3x3 confusion matrix for a classification model with three classes (A, B, and C), calculate the accuracy, precision, recall, and F1 score for Class A and B 10

		Predicted		
		Predicted Class A	Predicted Class B	Predicted Class C
Actual	Actual Class A	30	5	10
	Actual Class B	3	40	7
	Actual Class C	4	8	25

Q.5. a) Short note on Bayesian belief networks 10

b) Explain the Support Vector Machine (SVM) algorithm in detail. How does SVM work for classification and regression tasks? 10

Q.6. a) How are decision trees constructed. construct decision tree for following dataset. The target variable is play tennis (Yes/No). 10

Outlook	Temperature	Humidity	Wind	Play Tennis
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rainy	Mild	High	Weak	Yes
Rainy	Cool	Normal	Weak	Yes
Rainy	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rainy	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rainy	Mild	High	Strong	No

b) What is the ROC curve, and how does it relate to model performance evaluation? Explain the concept of "one-versus-many symmetric" in the context of classification problems. 10

___