

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

School of Computing Science and Engineering

Master of Technology in Computer Science and Engineering
Semester End Examination - Jun 2024

Duration : 180 Minutes

Max Marks : 100

Sem II - R1PV202B - MachineLearning*General Instructions*

Answer to the specific question asked

Draw neat, labelled diagrams wherever necessary

Approved data hand books are allowed subject to verification by the Invigilator

- 1) List the advantages of bagging over boosting. K1 (2)
- 2) Describe the ROC curve and AUC score. How are they used to assess the performance of binary classification algorithms? K2 (4)
- 3) Describe precision and recall. How are they calculated from a confusion matrix. K2 (6)
- 4) Evaluate the performance of a logistic regression model using receiver operating characteristic (ROC) curve and area under the curve (AUC) metrics. K3 (9)
- 5) Use logistic regression to predict the probability of an event occurring given certain independent variables, and determine the optimal threshold for classification. K3 (9)
- 6) Discuss strategies for dealing with class imbalance in logistic regression, such as oversampling, undersampling, or using class weights. K5 (10)
- 7) Discuss strategies for tuning hyperparameters in a Random Forest model, such as the number of trees, maximum depth, or minimum samples per leaf. K4 (12)
- 8) Using the concept of conditional independence, simplify a given Bayesian belief network by removing redundant nodes and edges. K5 (15)
- 9) Construct a Bayesian belief network to model the relationship between symptoms and diseases, given a set of conditional probability tables (CPTs). K5 (15)
- 10) Construct a Bayesian belief network to model the relationship between various features of a financial dataset and predict the likelihood of a stock price increase or decrease. K6 (18)