CVM UNIVERSITY
M.Sc. (INFORMATION TECHNOLOGY) Semester-I Examination-2021
Monday, 1st March – 2021
2:00 PM to 4:00 PM
101410108: MACHINE LEARNING TECHNIQUES

Total Marks: 60

Note: (1) Attempt all questions.
(2) Figures to the right indicate marks.

Q. 1   (a)   Answer the following multiple-choice questions.   (08)

(1) A computer program is said to learn from experience E with respect to some task T and some performance measure P if its performance on T, as measured by P, improves with experience E. Suppose we feed a learning algorithm a lot of historical weather data, and have it learn to predict weather. What would be a reasonable choice for P?
(A) The probability of it correctly predicting a future date's weather.
(B) The weather prediction task.
(C) The process of the algorithm examining a large amount of historical weather data.
(D) None of these.

(2) Suppose you are working on stock market prediction, and you would like to predict the price of a stock tomorrow (measured in dollars). You want to use a learning algorithm for this. Would you treat this as a classification or a regression problem?
(A) Regression   (B) Classification

(3) Suppose you are working on stock market prediction. You would like to predict whether or not a certain company will declare bankruptcy within the next 7 days. Would you treat this as a classification or a regression problem?
(A) Regression   (B) Classification

(4) Suppose that there are a total of 50 data mining related documents in a library of 200 documents. Suppose that a search engine retrieves 10 documents after a user enters 'data mining' as a query, of which 5 are data mining related documents. What are the precision and recall?
(1) 70% & 10%   (2) 45% & 20%   (3) 50% & 10%   (4) 60% & 5%

(5) High entropy means that the partitions in classification are
(A) Pure   (B) Not pure   (C) Useful   (D) Useless   (E) None of the above

(6) Which of the following is **NOT** supervised learning?
(A) PCA & Clustering   (B) Decision Tree   (C) Linear Regression   (D) Naive Bayesian

(7) Suppose you are working on weather prediction and use a learning algorithm to predict tomorrow's temperature (in degrees Centigrade/Fahrenheit). Would you treat this as a classification or a regression problem?
(A) Regression   (B) Classification

(8) Suppose you are working on weather prediction, and your weather station makes one of three predictions for each day's weather: Sunny, Cloudy or Rainy. You'd like to use a learning algorithm to predict tomorrow's weather. Would you treat this as a classification or a regression problem?
(A) Regression   (B) Classification

(b) Answer the following (08)

(1) Suppose we would like to convert a nominal attribute X with 4 values to a data table with only binary variables. How many ---- attributes are needed?

(2) Attributes are statistically dependent of one another given the class value. (TRUE/FALSE)

(3) High training accuracy doesn't necessarily imply high test accuracy. (TRUE/FALSE)

(4) Formula for Recall is------

(5) In Gradient Descent, you either use the entire data or a subset of training data to update a parameter in each iteration. (TRUE/FALSE)

(6) Adding a non-important feature to a linear regression model may result in----R SQUARE

(7) MSE stands for ---------

(8) ANN stands for ---------

Q.2 Attempt **any six** of the following. (12)

(1) When Will You Use Classification over Regression?

(2) Define and explain any one application of supervised learning.

(3) Define accuracy and error rate.

(4) Briefly Explain Logistic Regression.

(5) What is Decision Tree Classification?

(6) What is Pruning in Decision Trees, and How Is It Done?

(7) What Is a False Positive and False Negative and How Are They Significant?

(8) What are the Three Stages of Building a Model in Machine Learning?

Q. 3 Why do we need back propagation method? (08)

OR

Q. 3 What is Bias and Variance in a Machine Learning Model? (08)

Q. 4 What do you mean by Gradient Descent? (08)

OR

Q. 4 What are the conditions in which Gradient Descent is applied ? (08)

Q. 5 Explain Naïve Bayes Classifier with an Example. (08)

OR

Q. 5 Calculate based on below table (08)

(A) IG (S, outlook)   (B) IG (S, temperature)

In each case please do draw the decision tree in proper and desired format

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Hot	High	Strong	No
D5	Sunny	Mild	Normal	Weak	Yes
D6	Sunny	Mild	Normal	Strong	Yes
D7	Overcast	Mild	Normal	Strong	Yes
D8	Sunny	Cool	Normal	Weak	Yes
D9	Sunny	Mild	Normal	Weak	Yes
D10	Rain	Hot	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Weak	Yes
D12	Overcast	Mild	Normal	Weak	Yes
D13	Overcast	Cool	High	Weak	No
D14	Rain	Cool	High	Strong	No

Q. 6   What is 'Training Set' and 'Test Set' in a Machine Learning Model?   (08)
                                        OR
Q. 6   Enlist strength and weakness of SVM with proper exemplification   (08)

****