

CIA-1 Machine Learning Techniques

*Required

CIA-1 MLT

Morning slot: (09:30 AM to 10:30 AM) /Evening slot: (01:00 PM to 02:00 PM)

Course Outcomes:

C337.1 To understand the need for machine learning for various problem solving.(K1,K2)

C337.2 To understand a wide variety of learning algorithms and how to evaluate models generated from data.(K1,K3)

C337.3 To understand the latest trends in machine learning.(K2,K3)

C337.4 To design appropriate machine learning algorithms and apply the algorithms to a real-world problems.(K4,K6)

C337.5 To optimize the models learned and report on the expected accuracy that can be achieved by applying the models.(K4,K5)

Attempt All Questions

30*1 Marks =30 Marks

Q.1 Which of the following is true about Naive Bayes ? *

- ☐ A. Assumes that all the features in a dataset are equally important
- ☐ B. Assumes that all the features in a dataset are independent
- ☒ C. Both A and B
- ☐ D. None of these

Q.2 Is Logistic regression a supervised machine learning algorithm? *

☒ A TRUE

☒ A TRUE☐ B FALSE

Q.3 Can a Logistic Regression classifier do a perfect classification on the below data? Note: You can use only X1 and X2 variables where X1 and X2 can take only two binary values(0,1). *

☐ A TRUE☒ B FALSE☐ C Can't say☐ D None of these

Q.4 Standardisation of features is required before training a Logistic Regression. *

☐ A TRUE☒ B FALSE

Q.5 Is it possible to apply a logistic regression algorithm on a 3-class Classification problem? *

☒ A TRUE☐ B FALSE

Q.6 Which of the following is an example of feature extraction? *

☐ A Constructing bag of words vector from an email☐ B Applying PCA projects to a large high-dimensional data

- ☐ C Removing stopwords in a sentence
- ☒ D All of the above

Q.7 How do you handle missing or corrupted data in a dataset? *

- ☐ A Drop missing rows or columns
- ☐ B Replace missing values with mean/median/mode
- ☐ C Assign a unique category to missing values
- ☒ D All of the above

Q.8 When performing regression or classification, which of the following is the correct way to preprocess the data? *

- ☒ A Normalize the data → PCA → training
- ☐ B PCA → normalize PCA output → training
- ☐ C Normalize the data → PCA → normalize PCA output → training
- ☐ D None of the above

Q.9 In modeling, an optimal solution is understood to be *

- ☐ A a solution that can only be determined by an exhaustive enumeration and testing of alternatives.
- ☐ B a solution found in the least possible time and using the least possible computing resources
- ☒ C a solution that is the best based on criteria defined in the design phase
- ☐ D a solution that requires an algorithm for determination



Q.10 Multiple linear regression (MLR) is a _____ type of statistical analysis. *

- ☐ A univariate
- ☐ B bivariate
- ☒ C multivariate
- ☐ D none of these

Q.11 In practice, Line of best fit or regression line is found when *

- ☐ A Sum of residuals ($\sum(Y - h(X))$) is minimum
- ☐ B Sum of the absolute value of residuals ($\sum|Y-h(X)|$) is maximum
- ☒ C Sum of the square of residuals ($\sum (Y-h(X))^2$) is minimum
- ☐ D Sum of the square of residuals ($\sum (Y-h(X))^2$) is maximum

Q.12 If Linear regression model perfectly first i.e., train error is zero, then _____ *

- ☐ A Test error is also always zero
- ☐ B Test error is non zero
- ☒ C Couldn't comment on Test error
- ☐ D Test error is equal to Train error



Q.13 Which of the following metrics can be used for evaluating regression models? i) R Squared ii) Adjusted R Squared iii) F Statistics iv) RMSE / MSE / MAE *

- ☐ A ii and iv
- ☐ B i and ii
- ☐ C ii, iii and iv
- ☒ D i, ii, iii and iv

Q.14 How many coefficients do you need to estimate in a simple linear regression model (One independent variable)? *

- ☐ A 1
- ☒ B 2
- ☐ C 3
- ☐ D 4

Q.15 In the mathematical Equation of Linear Regression $Y = \beta_1 + \beta_2 X + \epsilon$, (β_1, β_2) refers to _____ *

- ☐ A (X-intercept, Slope)
- ☐ B (Slope, X-Intercept)
- ☒ C (Y-Intercept, Slope)
- ☐ D (slope, Y-Intercept)



Q.16 A multiple regression model has the form: $y = 2 + 3x_1 + 4x_2$. As x_1 increases by 1 unit (holding x_2 constant), y will *

- ☒ A increase by 3 units
- ☐ B decrease by 3 units
- ☐ C increase by 4 units
- ☐ D decrease by 4 units

Q.17 Which of the following methods do we use to find the best fit line for data in Linear Regression? *

- ☒ A Least Square Error
- ☐ B Maximum Likelihood
- ☐ C Logarithmic Loss
- ☐ D Both A and B

Q.18 In a simple linear regression model (One independent variable), If we change the input variable by 1 unit. How much output variable will change? *

- ☐ A by 1
- ☐ B no change
- ☐ C by intercept
- ☒ D by its slope



Q.19 The attribute x is the input variable and y is the output variable that we are trying to predict. Calculate the intercept value (B_0) for the given data as below. (Use linear regression algorithm) *

x	y
1	1
2	3
4	3
3	2
5	5

- ☒ A 0.4
- ☐ B 0.8
- ☐ C 0.6
- ☐ D 0.7

Q.20 Which of the factors affect the performance of learner system does not include? *

- ☐ A Representation scheme used
- ☐ B Training scenario
- ☐ C Type of feedback
- ☒ D Good data structures

Q.21 A machine learning problem involves four attributes plus a class. The attributes have 3, 2, 2, and 2 possible values each. The class has 3 possible values. How many maximum possible different examples are there? *

- ☐ A 12
- ☐ B 24



- ☐ B 24
- ☐ C 48
- ☒ D 72

Q.22 Which logistic function is used in Logistic Regression? *

- ☐ A Relu
- ☒ B Sigmoid
- ☐ C Linear
- ☐ D Tanh

Q.23 The most widely used metrics and tools to assess a classification model are: *

- ☐ A Confusion matrix
- ☐ B Cost-sensitive accuracy
- ☐ C Area under the ROC curve
- ☒ D All of the above

Q.24 Consider a problem with three random variables: A, B, and C. A is dependent upon B, and C is dependent upon B. The joint probability of P(A,B,C) is: *

- ☐ A $P(A, B, C) = P(B|A) * P(A|B) * P(B)$
- ☒ B $P(A, B, C) = P(A|B) * P(C|B) * P(B)$
- ☐ C $P(A, B, C) = P(C|B) * P(B|C) * P(B)$
- ☐ D $P(A, B, C) = P(A|B) * P(B|A) * P(C)$



Q.25 For the new instance (v_i), given training data (D) and the given a space of hypotheses (H). What is the conditional probability *

- ☒ A $P(v_j | D) = \sum \{h \in H\} P(v_j | h_i) * P(h_i | D)$
- ☐ B $(v_j | D) = \max(\sum(\{h \in H\} P(v_j | h_i) * P(h_i | D))$
- ☐ C $P(D | v_i) = \sum \{h \in H\} P(v_j | h_i) * P(h_i | D)$
- ☐ D None

Q.26 For Bayes theorem to be applied, the following relationship between hypothesis H and evidence E must hold. *

- ☒ A $P(H|E) + P(H| \sim E) = 1$
- ☐ B $P(H|E) + P(\sim H| E) = 1$
- ☐ C $P(H|E) + P(H| \sim E) = 0$
- ☐ D $P(H|E) + P(\sim H| E) = 0$

Q.27 Suppose you are given an EM algorithm that finds maximum likelihood estimates for a model with latent variables. You are asked to modify the algorithm so that it finds MAP estimates instead. Which step or steps do you need to modify? *

- ☐ A Expectation
- ☒ B Maximization
- ☐ C No modification necessary
- ☐ D Both A and B



Q.28 In Bayes Theorem, the unconditional probability is called as *

- ☐ A Evidence
- ☒ B Likelihood
- ☐ C Prior
- ☐ D Posterior

Q.29 Which of the following tends to work best on small data sets (few observations)? *

- ☐ (a) Logistic regression
- ☒ (b) Naive Bayes

Q.30 Algorithm X is better than Algorithm Y if the training error of algorithm X is better than that of Y *

- ☐ True
- ☒ False

Back

Submit

Never submit passwords through Google Forms.

This form was created inside JSS Academy of Technical Education, NOIDA. [Report Abuse](#)

Google Forms

