Assignment 3

Question 1) Among the following identify the one in which dimensionality reduction reduces.

Answer: (d) Collinearity

Question 2) Which of the following machine learning algorithm is based upon the idea of bagging?

Answer: (b) Random Forest

Question 3) Choose a disadvantage of decision trees among the following.

Answer: (c) Decision Tree are prone to overfit

Question 4) What is the term known as on which the machine learning algorithms build a model based on sample data?

Answer: (a) Data Training

Question 5) Which of the following machine learning techniques helps in detecting the outliers in data?

Answer: (c) Anamoly detection

Question 6) Identify the incorrect numerical functions in the various function representation of machine learning.

Answer: (c) Case based

Question 7) Analysis of ML algorithm needs.

Answer: (d) Both a and b

Question 8) Identify the difficulties with the k-nearest neighbor algorithm.

Answer: (c) Both a and b

Question 9) The total types of the layer in radial basis function neural networks is _____

Answer: (c) 3

Question 10) Which of the following is not a supervised learning

Answer: (a) PCA

Question 11) What is unsupervised learning?

Answer: (c) Neither feature nor number of groups is known

Question 12) Which of the following is not a machine learning algorithm?

Answer: (a) SVM

Question 13) _____ is the scenario when the model fails to decipher the underlying trend in the input data

Answer: (b) Underfitting

Question 14) Real-Time decisions, Game AI, Learning Tasks, Skill acquisition, and Robot Navigation are applications of .....

Answer: (a) Reinforcement learning

Question 15) What is called the average squared difference between classifier predicted output and actual output? 55) What is called the average squared difference between 55classifier

Answer: (b) Mean squared error

Question 16) Logistic regression is a ..... regression technique that is used to model data having a ..... outcome.

Answer: (a) Linear, binary

Question 17) You are given reviews of few netflix series marked as positive, negative and neutral. Classifying reviews of a new netflix series is an example of

Answer: (a) supervised learning

Question 18) Following is powerful distance metrics used by Geometric model

Answer: (c) both a and b

Question 19) Which of the following techniques would perform better for reducing dimensions of a data set?

Answer: (b) removing columns which have high variance in data

Question 20) Supervised learning and unsupervised clustering both require which is correct according to the statement.

Answer: (c) input attribute.

Question 21) What is the meaning of hard margin in SVM?

Answer: (a) SVM allows very low error in classification

Question 22) Increase in which of the following hyper parameter results into overfit in Random forest? (1). Number of Trees. (2). Depth of Tree, (3). Learning Rate

Answer: (b) Only 2

Question 23) Below are the 8 actual values of target variable in the train file: [0,0,0, 0, 1, 1,1,1,1,1], What is the entropy of the target variable?

Answer: (A) -(6/10 log(6/10) + 4/10 log(4/10))

Question 24) Lasso can be interpreted as least-squares linear regression where

Answer: (C) the solution algorithm is simpler

Question 25) Consider the problem of binary classification. Assume I trained a model on a linearly separable training set, and now I have a new labeled data point that the model properly categorized and is far away from the decision border. In which instances is the learnt decision boundary likely to change if I now add this additional point to my previous training set and re-train? When the training model is,

Answer: (D) Perceptron

Question 26) Assume you've discovered multi-collinear features. Which of the following actions do you intend to take next? (1). Both collinear variables should be removed. (2). Instead of deleting both variables, we can simply delete one. (3). Removing correlated variables may result in information loss. We may utilize penalized regression models such as ridge or lasso regression to keep such variables.

Answer: (D) Either 2 or 3

Question 27) A least squares regression study of weight (y) and height (x) yielded the following least squares line: y = 120 + 5x. This means that if the height is increased by one inch, the weight should increase by what amount?

Answer: (B) increase by 5 pound

Question 28) The line described by the linear regression equation (OLS) attempts to ____?

Answer: (D) Minimize the squared distance from the points

Question 29) For two real-valued attributes, the correlation coefficient is 0.85. What does this value indicate?

Answer: (B) As the value of one attribute increases the value of the second attribute also increases

Question 30) Which neural network architecture would be most suited to handle an image identification problem (recognizing a dog in a photo)?

Answer: (B) Convolutional Neural Network