20U5115 (Pages: 2) Name: …………………….

Reg. No: …………………..

FIFTH SEMESTER B.Voc. DEGREE EXAMINATION, NOVEMBER 2022
(Regular/Supplementary/Improvement)
CC18U SDC5 ML20 - MACHINE LEARNING USING PYTHON
(Information Technology)
(2018 Admission onwards)

Time: Three Hours   Maximum: 80 Marks

PART A
Answer *all* questions. Each question carries 1 mark.

1. A term used to describe the case when the independent variables in a multiple regression model are correlated is _____

2. Machine learning is a subset of _____

3. Decision tree is a _____ algorithm.

4. In Regression, a decision tree splits the dataset based on _____

5. For two events A and B, the Bayes theorem will be _____

6. _____ is a library consisting of multidimensional array objects and a collection of routines for processing those arrays.

7. _____ an information filtering system that predicts what a user might want to hear or see based on choice patterns provided by the user.

8. A _____ is a measurable property of the object you're trying to analyze.

9. Machine learning algorithm is said to have _____ when it cannot capture the underlying trend of the data.

10. _____ algorithm works with the categorical variable such as 0 or 1, Yes or No, True or False, Spam or not spam, etc.

(10 × 1 = 10 Marks)

PART B
Answer any *eight* questions. Each question carries 2 marks.

11. What is unsupervised learning?

12. What is ensemble learning?

13. Compare K-means and KNN algorithms.

14. How to create 1D and 2D array using Numpy?

15. What is Dimensionality Reduction?

16. Define feature selection.

17. What are the differences between classification and clustering?

18. What is classifier in machine learning?

19. Define binary and multiclass classification

20. Define KFold cross validation technique.

21. Define independent variable with example.

22. What is overfitting?

(8 × 2 = 16 Marks)

PART C
Answer any *six* questions. Each question carries 4 marks.

23. Explain Linear regression with example.

24. Explain bag of word approach. Implement feature extraction from text document.

25. Explain support vector regression.

26. Explain broadcasting in Numpy.

27. What are the advantages of using a Navie Bayes for classification?

28. What are the differences between Artificial intelligence and Machine learning?

29. What is Bias, Variance and what do you mean by Bias-Variance Tradeoff?

30. Explain Count Vectorizer with example.

31. What is a confusion matrix and why do you need it?

(6 × 4 = 24 Marks)

PART D
Answer any *two* questions. Each question carries 15 marks.

32. Explain Logistic Regression. Implement the Logistic regression algorithm.

33. What are the different types of ML Classification Algorithms?

34. What is Machine learning? Why do you need it? Explain Machine learning Life cycle.

35. Explain the K Nearest Neighbor Algorithm with example.

(2 × 15 = 30 Marks)

****