



M.C.A. III Semester Degree Examination, April/May - 2024

MACHINE LEARNING

(NEP)

Time : 3 Hours

Maximum Marks : 70

Note : Answer **any five** of the following questions with question **No.1 Compulsory**.

1. (a) Discuss different perspectives on prediction models in Machine Learning. **7**
(b) Design a visual representation, such as a histogram or box plot, to effectively communicate the descriptive statistics of a given continuous dataset. Justify your choice of visualization based on the characteristics of the data. **7**
2. (a) What are the primary objectives of predictive data analytics ? **7**
(b) Define Machine Learning. What can go wrong with Machine Learning ? **7**
3. (a) Consider a case where you're given that a fruit is long, sweet and yellow and you need to predict what type of fruit it is using Naive bayes. **7**

Type	Long	Not long	Sweet	Not sweet	Yellow	Not yellow	Total
Banana	400	100	350	150	450	50	500
Apple	0	300	150	150	300	0	300
Other	100	100	150	50	50	150	200
Total	500	500	650	350	800	200	1000

- (b) Develop a tutorial or guide for beginners on implementing SVM for classification tasks in machine learning . **7**
4. (a) Describe the process of splitting nodes in a decision tree and how it contributes to the tree's predictive accuracy. **7**
(b) Compare and contrast the properties of Gini index and entropy as measures of impurity. **7**
5. (a) Describe the difference between Euclidean distance and Cosine Similarity in measuring similarity between vectors. **7**
(b) Why do we need Data Normalization in Machine Learning ? **7**



6. (a) What are the primary objectives of each phase in the CRISP-DM methodology ? **7**
- (b) Verify two events A and B are conditionally independent given an event C. A box contains two coins : a regular coin and one fake two-headed coin ($P(H)=1$). I choose a coin at random and toss it twice. Define the following events. **7**
- A=First coin toss results in an H.
 - B=Second coin toss results in an H.
 - C=Coin 1 (regular) has been selected.
- Find $P(A|C), P(B|C), P(A \cap B|C), P(A), P(B)$ and $P(A \cap B)$
7. (a) How is the value predicted by a leaf node in a regression tree typically determined ? **7**
- (b) Why do we need KNN algorithm ? How does KNN work ? **7**
8. Write short notes on the following :
- (a) Efficient Memory search **5**
 - (b) Noisy data **5**
 - (c) Decision tree **4**

- o 0 o -

