Exam.Code:1000 Sub. Code: 7630

## 2072

## M.E. (Computer Science and Engineering) Second Semester Elective - III

CS-8205: Machine Learning

Time allowed: 3 Hours Max. Marks: 50

NOTE: Attempt five questions in all, including Question No. I which is compulsory and selecting two questions from each Part.

| Balleton and I have |  | AND DESCRIPTION OF |
|---------------------|--|--------------------|
|                     | a) What is the most commonly used cost function for classification model? Discuss in     |                    |
|                     | brief.   | (02)               |
|                     | b) What is cross-validation procedure? What is its use?                                  | (02)               |
|                     | c) Why do we need soft margin in SVM?  | (02)               |
|                     | d) Differentiate between core-points and border-points in DBSCAN clustering.             | (02)               |
|                     | e) Give the general form of probability density function of normal distribution.         | (02)               |
|                     | PART I   |                    |
|                     | a) If True Positive (TP) = 70, True Negative (TN) = 25, False Positive (FP) = 25 and     |                    |
|                     | False Negative = 20, then what is i) Recall ii) Specificity iii) False Positive Rate iv) |                    |
|                     | Precision and v) Error Rate  | (05)               |
|                     | b) Why are overfitting and underfitting situations bad? How can they be detected and     |                    |
|                     | avoided? Is overfitting more critical than underfitting? Discuss.                        | (05)               |
| III .               | a) Discusses any two strategies for reducing the problem of multiclass classification    |                    |
|                     | to multiple binary classification problems.  | (04)               |
|                     | b) What is multivariate linear regression? How parameter learning happens in             |                    |
|                     | multivariate linear regression? Discuss.   | (06)               |
| IV                  | a) What is class imbalance problem? What tactics can be employed to handle this          | ) (                |
|                     | problem? Discuss.  | (04)               |
|                     | b) What is the use of kernel function in SVM? Discuss any two commonly used              |                    |
|                     | kernel functions.  | (06)               |
|                     | PART II  |                    |
| , V                 | a) Discuss the working of k-Means clustering algorithm and its ability to handle         |                    |
|                     | outliers.  | (06)               |
|                     | b) Compare the characteristics of density-based methods of clustering with               | \ · · · .          |
|                     | hierarchical methods of clustering.  | (04)               |
| VI                  | a) What is Expectation Maximization Algorithm? What is the purpose of E-Step and         | 1                  |
|                     | M-Step? Discuss.   | (05)               |
|                     | b) What are the steps involved in building an anomaly detection system? What are the     | 1                  |
|                     | challenges involved? Discuss.  | (05)               |
| VII                 | Write notes on the following:  | 1                  |
|                     | a) Collaborative Filtering   | (05)               |
|                     | b) Density Estimation  | (05)               |