**MACHINE LEARNING**

**Assignment-1**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **SECTION – A** |  |
| **1.** | **All questions are compulsory** | |  |
|  | **a.** | What do you expect will happen with bias and variance as you increase the size of training data?  A: Bias increases and Variance increases B: Bias decreases and Variance increases C: Bias decreases and Variance decreases  D: Bias increases and Variance decreases | **CO3** |
|  | **b.** | Regression trees are often used to model…… data. A: Linear  B: Nonlinear C: Categorical  D: Symmetrical | **CO2** |
|  | **c.** | Which of the following is not a supervised learning? A: Naïve Bayesian  B: PCA  C: Linear Regression D: Decision Tree | **CO2** |
|  | **d.** | Naïve Bayes algorithm is based on and used for solving  classification problems. A: Bayes Theorem  B: Candidate Elimination Algorithm C: EM algorithm  D: None of the above | **CO3** |
| **2.** | **All questions are compulsory** | |  |
|  | **a.** | Define Overfitting and Underfitting. | **CO2** |
|  | **b.** | Compare Classification & Regression. | **CO3** |

|  |  |  |  |
| --- | --- | --- | --- |
| **SECTION – B** | | |  |
| **3.** | **Answer any two of the following-** | |  |
|  | **a.** | Explain the Naïve Bayes concept in detail with the help of an example  and the python code. | **CO3** |
|  | **b.** | Differentiate between Linear and Logistic regression. Explain both with  the help of an example and the Python code. | **CO2** |
|  | **c.** | Write the python code model data splitting and importing. Also, explain the concept of Confusion matrix and write the formulas for calculating  accuracy, recall and precision. | **CO3** |
| **SECTION – C** | | |  |
| **4** | **Answer any one of the following-** | |  |
|  | **a.** | Explain the procedure of building decision tree with Information Gain &  Entropy. Illustrate with example. | **CO3** |
|  | b. | Explain the concept of training and evaluation of the model using  Confusion Matrix. | **CO3** |
| **5.** | **Answer any one of the following-** | |  |
|  | **a.** | Define Instance-based learning. Explain the concept of KNN algorithm  with the help of an example. | **CO3** |
|  | **b.** | Elaborate the issues in decision tree learning. Explain briefly how can they be overcome? Discuss the following in detail.   1. Avoiding Overfitting in decision trees. 2. Handling training examples with missing attributes values. 3. Handling attributes with different costs. | **CO2** |