**Department of Computer Science & Engineering**

**MACHINE Learning LAB (R20)**

**LIST OF EXPERIMENTS**

**Size – (2\*3) Count- 1**

|  |  |
| --- | --- |
| Exercise-1 | Implement and demonstrate the FIND-S algorithm for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file. |
| Exercise-2 | For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples. |
| Exercise-3 | Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample. |
| Exercise-4 | Exercises to solve the real-world problems using the following machine learning methods: a) Linear Regression b) Logistic Regression c) Binary Classifier |
| Exercise-5 | Develop a program for Bias, Variance, Remove duplicates , Cross Validation |
| Exercise-6 | Write a program to implement Categorical Encoding, One-hot Encoding |
| Exercise-7 | Build an Artificial Neural Network by implementing the Back propagation algorithm and test the same using appropriate data sets. |
| Exercise-8 | Write a program to implement k-Nearest Neighbor algorithm to classify the iris data set. Print both correct and wrong predictions. |
| Exercise-9 | Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs |
| Exercise-10 | Assuming a set of documents that need to be classified, use the naïve Bayesian Classifier model to perform this task. Built-in Java classes/API can be used to write the program. Calculate the accuracy, precision, and recall for your data set. |
| Exercise-11 | Apply EM algorithm to cluster a Heart Disease Data Set. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering. You can add Java/Python ML library classes/API in the program. |
| Exercise-12 | Exploratory Data Analysis for Classification using Pandas or Matplotlib. |
| Exercise-13 | Write a Python program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set |
| Exercise-14 | Write a program to Implement Support Vector Machines and Principle Component Analysis |
| Exercise-15 | Write a program to Implement Principle Component Analysis |