Sreenidhi Institute of Science and Technology
Department of Computer Science and Engineering
B.Tech Computer Science and Engineering
Research Laboratory

Data science & Machine Learning Lab (Lab No:2414B):

Objective:

The purpose of the lab is to do experiments on the modern AI and ML Algorithms, to gain knowledge on how they actually work. And to some research on improving the models created or optimize them. Machine Learning is a set of methods and techniques for constructing software systems automatically by analyzing only examples of the desired behavior. We do not write a program. We only provide examples of what we want. The program will be synthesized automatically.

Activities carried out in the Machine Learning Lab:

- 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.
- Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets.
- Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.
- Write a 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. You can use Python ML library classes.
- Apply EM algorithm to cluster a set of data stored in a .CSV file. 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 Python ML library classes in the program.
- Write a program to implement k-Nearest Neighbor algorithm to classify the iris data set. Print both correct and wrong predictions. Python ML library classes can be used for this problem.
- Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs.

Activities carried out in the Data science Lab:

- To Learn R Basics and Statistical Modelling in R

- Testing the Regression algorithms like – Linear & Multiple Regression (case studies)
- Solving the case studies on Classification algorithms like Logistic regression, The Stock Market Data, K-Nearest Neighbors.
- To solve the case studies using Decision Trees concepts like Bagging, Boosting and Random Forests.
- To solve the case studies using SVM model.
- To understand and analyze the Unsupervised Learning based algorithms.

Hardware Configuration	
Company	HP
Make	Compaq 4000 series (64 Bit Edition)
Software Configuration	
For Python	Python 3.x, 2.x
For R	Rstudio
OS	Ubuntu 18.04
Platform for Python	Jupyter Notebook, Spyder