DATA SCIENCE LAB

B.Tech III Year I Semester CSE-DS

L	T	P	C
0	0	2	1

Course Outcomes:

1. To make students understand learn about a Data Science – Python Programming, way of solving problems.
2. To teach students to write programs in Python to solve problems.
3. Demonstrate the usage of built-in objects in Python.
4. Analyze the significance of python program development environment by working on real world examples
5. Implement numerical programming, data handling and visualization through NumPy, Pandas and MatplotLib modules.

1. INTRODUCTION TO PYTHON

Structure of Python Program-Underlying mechanism of Module Execution-Branching and Looping-Problem Solving Using Branches and Loops-Functions - Lists and Mutability-Problem Solving Using Lists and Functions

Week 1:

1. Demonstrate usage of branching and looping statements
2. Demonstrate Recursive functions
3. Demonstrate Lists

Week 2:

SEQUENCE DATATYPES AND OBJECT-ORIENTED PROGRAMMING

Sequences, Mapping and Sets- Dictionaries- -Classes: Classes and Instances-Inheritance-Exceptional Handling-Introduction to Regular Expressions using "re" module.

Lab Exercises

1. Demonstrate Tuples and Sets
2. Demonstrate Dictionaries
3. Demonstrate inheritance and exceptional handling
4. Demonstrate use of "re".

Week 3:

USING NUMPY

Basics of NumPy-Computation on NumPy-Aggregations-Computation on Arrays-Comparisons, Masks and Boolean Arrays-Fancy Indexing-Sorting Arrays-Structured Data: NumPy's Structured Array.

Lab Exercises

1. Demonstrate Aggregation
2. Demonstrate Indexing and Sorting

Week 4:
DATA MANIPULATION WITH PANDAS -I
Introduction to Pandas Objects-Data indexing and Selection-Operating on Data in Pandas- Handling Missing Data-Hierarchical Indexing - Combining Data Sets
Lab Exercises
1. Demonstrate handling of missing data
2. Demonstrate hierarchical indexing

Week 5:
DATA MANIPULATION WITH PANDAS -II
Aggregation and Grouping-Pivot Tables-Vectorized String Operations -Working with Time Series-High Performance Pandas- and query ()
Lab Exercises
1. Demonstrate usage of Pivot table
2. Demonstrate use of and query ()

Week 6:
VISUALIZATION AND MATPLOTLIB
Basic functions of matplotlib-Simple Line Plot, Scatter Plot-Density and Contour Plots-Histograms, Binnings and Density-Customizing Plot Legends, Colour Bars-Three-Dimensional Plotting in Matplotlib.
Lab Exercises
1. Demonstrate Scatter Plot
2. Demonstrate 3D plotting

Week 7:
Perform Data exploration and pre-processing in Python

Week 8:
Implement regularized linear regression

Week 9:
Implement Naive Bayes classifier for dataset stored as CSV file.

Week 10:
Implement regularized logistic regression
Week 11:
Build models using different Ensembling techniques

Week 12:
Build models using Decision trees

Week 13:
Build model using SVM with different kernels

Week 14:
Implement K-NN algorithm to classify a dataset.

Week 15:
Build model to perform Clustering using K-means after applying PCA and determining the value of K using Elbow method.