# 23AD21P1 - DATA SCIENCE LAB

(CSE (DS))

|  |  |  |  |
| --- | --- | --- | --- |
| Course Category: | Professional Core | Credits: | 1.5 |
| Course Type: | Practical | Lecture-Tutorial-Practical: | 0-0-3 |
| Prerequisite: | Familiarity with Jupyter Notebooks for interactive coding and data analysis | Sessional Evaluation:Univ. Exam Evaluation:Total Marks: | 3070100 |
| Objectives: | * The main objective of the course is to inculcate the basic understanding of Data Science and its practical implementation using Python.
 |

|  |  |
| --- | --- |
| Course Outcomes | Upon successful completion of the course, the students will be able to: |
| CO1 | Apply principles and techniques for optimizing the performance of Pythonapplications (L3) |
| CO2 | Implement parallel computing applications using Python (L5) |
| CO3 | Develop GPU accelerated Python applications (L6) |
| Course Content | **List of Experiments**1. Creating a NumPy Arraya. Basic ndarrayb. Array of zerosc. Array of onesd. Random numbers in ndarraye. An array of your choicef. Imatrix in NumPyg. Evenly spaced ndarray
2. The Shape and Reshaping of NumPy Arraya. Dimensions of NumPy arrayb. Shape of NumPy arrayc. Size of NumPy arrayd. Reshaping a NumPy arraye. Flattening a NumPy arrayf. Transpose of a NumPy array
3. Expanding and Squeezing a NumPy Arraya. Expanding a NumPy arrayb. Squeezing a NumPy arrayc. Sorting in NumPy Arrays
4. Indexing and Slicing of NumPy Arraya. Slicing 1-D NumPy arraysb. Slicing 2-D NumPy arraysc. Slicing 3-D NumPy arraysd. Negative slicing of NumPy arrays
5. Stacking and Concatenating Numpy Arraysa. Stacking ndarraysb. Concatenating ndarraysc. Broadcasting in Numpy Arrays
6. Perform following operations using pandasa. Creating dataframeb. concat()c. Setting conditionsd. Adding a new column
7. Perform following operations using pandasa. Filling NaN with string.

b. Sorting based on column valuesc. groupby()1. Read the following file formats using pandasa. Text filesb. CSV filesc. Excel filesd. JSON files
2. Read the following file formatsa. Pickle filesb. Image files using PILc. Multiple files using Globd. Importing data from database
3. Demonstrate web scraping using python
4. Perform following preprocessing techniques on loan prediction dataseta. Feature Scalingb. Feature Standardizationc. Label Encodingd. One Hot Encoding
5. Perform following visualizations using matplotliba. Bar Graphb. Pie Chartc. Box Plotd. Histograme. Line Chart and Subplotsf. Scatter Plot
6. Getting started with NLTK, install NLTK using PIP
7. Python program to implement with Python Sci Kit-Learn & NLTK
8. Python program to implement with Python NLTK/Spicy/Py NLPI.
 |
| E-Resources | 1. <https://www.analyticsvidhya.com/blog/2020/04/the-ultimate-numpy-tutorial-for-datascience-beginners/>
2. <https://www.analyticsvidhya.com/blog/2021/07/data-science-with-pandas-2-minutesguide-to-key-concepts/>
3. <https://www.analyticsvidhya.com/blog/2020/04/how-to-read-common-file-formatspython/>
4. <https://www.analyticsvidhya.com/blog/2016/07/practical-guide-data-preprocessingpython-scikit-learn/>
5. <https://www.analyticsvidhya.com/blog/2020/02/beginner-guide-matplotlib-datavisualization-exploration-python/6>.
6. <https://www.nltk.org/book/ch01.html>
 |