# 23AD22P2 - DATA SCIENCE USING PYTHON LAB

(AI&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: | 30  70  100 |
| Objectives: | **Students undergoing this course are expected:** | | |
| * 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 Python  applications |
| CO2 | Implement parallel computing applications using Python |
| CO3 | Develop GPU accelerated Python applications |
| Course Content | List of Experiments  1. **Creating a NumPy Array** a. Basic ndarray b. Array of zeros c. Array of ones d. Random numbers in ndarray e. An array of your choice f. Imatrix in NumPy g. Evenly spaced ndarray 2. **The Shape and Reshaping of NumPy Array** a. Dimensions of NumPy array b. Shape of NumPy array c. Size of NumPy array d. Reshaping a NumPy array e. Flattening a NumPy array f. Transpose of a NumPy array 3. **Expanding and Squeezing a NumPy Array** a. Expanding a NumPy array b. Squeezing a NumPy array c. Sorting in NumPy Arrays 4. **Indexing and Slicing of NumPy Array** a. Slicing 1-D NumPy arrays b. Slicing 2-D NumPy arrays c. Slicing 3-D NumPy arrays d. Negative slicing of NumPy arrays 5. **Stacking and Concatenating Numpy Arrays** a. Stacking ndarrays b. Concatenating ndarrays c. Broadcasting in Numpy Arrays 6. **Perform following operations using pandas** a. Creating dataframe b. concat() c. Setting conditions d. Adding a new column 7. **Perform following operations using pandas** a. Filling NaN with string.   b. Sorting based on column values c. groupby()   1. **Read the following file formats using pandas** a. Text files b. CSV files c. Excel files d. JSON files 2. **Read the following file formats** a. Pickle files b. Image files using PIL c. Multiple files using Glob d. Importing data from database 3. **Demonstrate web scraping using python** 4. **Perform following preprocessing techniques on loan prediction dataset** a. Feature Scaling b. Feature Standardization c. Label Encoding d. One Hot Encoding 5. **Perform following visualizations using matplotlib** a. Bar Graph b. Pie Chart c. Box Plot d. Histogram e. Line Chart and Subplots f. 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-pre processing python-scikit-learn/](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> | |