Antrix Academy of Data Science

Summer Online Internship Program
(in collaboration with Ajay Kumar Garg Engineering College)
Fees : INR : 3290/-
Payment Link : http://antrixacademy.com/Pay.aspx
Support No. : 9971-283-969/8130-371-253
Duration : 40 Hrs/ 6 Weeks

Recap-Python	Statistical Fundamentals	Machine Learning

Training: Implementing Machine Learning with Python

Recap Key Concept of Data Analytics with Python

Data Analysis with Pandas

- Using Series, DataFrame, Panels
- Data wrangling
- Sorting and filtering data
- Aggregate operations
- Analyzing time series
- Visualization with Pandas

Vectorizing Data in Numpy

- Creating Numpy arrays
- Common operations on matrices
- Using Analytics functions
- Views and broadcasting on Numpy arrays
- Optimizing performance by avoiding loops

Python: Data Manipulation – Cleansing

- Cleansing Data with Python
- Data Manipulation steps(Sorting, filtering, duplicates, merging, appending, subsetting, derived variables, sampling, Data type conversions, renaming, formatting etc)
- Data manipulation tools(Operators, Functions, Packages, control structures, Loops, arrays etc)
- Python Built-in Functions (Text, numeric, date, utility functions)

- Python User Defined Functions
- Stripping out extraneous information
- Normalizing data
- Formatting data
- Important Python Packages for data manipulation (Pandas, Numpy etc)

Python: Accessing/Importing and Exporting Data

- Importing Data from various sources (Csv, txt, excel, access etc)
- Database Input (Connecting to database – MySQL, MS SQL, Oracle, Teradata)
- Viewing Data objects - subsetting, methods
- Exporting Data to various formats

Python: Data Analysis – Visualization

- Introduction exploratory data analysis
- Descriptive statistics, Frequency Tables and summarization
- Univariate Analysis (Distribution of data & Graphical Analysis)
- Bivariate Analysis(Cross Tabs, Distributions & Relationships, Graphical Analysis)
- Creating Graphs- Bar/pie/line chart/histogram/boxplot/scatter/density etc)
- Important Packages for Exploratory Analysis(NumPy Arrays, Matplotlib, Pandas and scipy.stats etc)

Python: OpenCV

- Introduction to OpenCv Image processing Library
- Processing real-time images with OpenCv
- Industrial case study on use case of OpenCv

Python: Natural Language Processing with NLTK

- Introduction to NLP
- Text Tokenization, chunking, pos-tagging using NLTK.
- Syntactic Parsing in Python
- Entity Recognition from document
- Text Mining in Python
- Sentiment Analysis

Python: Basic statistics

- Basic Statistics - Measures of Central Tendencies and Variance
- Building blocks - Probability Distributions - Normal distribution - Central Limit Theorem
- Inferential Statistics -Sampling - Concept of Hypothesis Testing
- Statistical Methods - Z/t-tests (One sample, independent, paired), Anova, Correlation and Chi-square

Introduction to Machine Learning with Python

- Orientation of Machine Learning and Industry Use-cases.
- Different types of Machine Learning Techniques.

- Understanding the SKLEARN package for Machine Learning

Machine Learning (Supervised Learning) - I

- Generalised Linear Models
  - Linear Regression
  - Ridge and Lasso Regression
  - Logistic Regression
- Classification
  - Random Forest
  - Decision Trees
  - Support Vector Machines
  - KNN
  - Naïve Bayes
  - Usage

Machine Learning (Unsupervised Learning) - II

- Clustering
  - K-Means
  - K Nearest Neighbours
  - Association Rule Learning