



CERTIFICATE IN DATA SCIENCE

DURATION: 120 Hours

TOTAL CREDITS: 4

COURSE SYLLABUS

Objective

The Certificate Course in Data Science aims to equip learners with the essential skills to analyze, process, and visualize data using industry-standard tools and techniques. Participants will develop proficiency in programming, statistical analysis, and machine learning, enabling them to build data-driven solutions for real-world applications.

Exit Profile

- Complete Knowledge in Data Science
- Knowledge in data analysis
- Data Modeling
- Python, R, and SQL
- Data analysis
- Web mining

Career Path

- Data Analyst
- Data Scientist

Course Outline

Course Name:	Certificate in Data Science	Duration:	120 H
Module	Topic	Duration	Total Duration
Module 1	Data analysis	5 H	50 H
	Data mining concepts and application	5 H	
	Data pre-processing	15 H	
	Data wrangling	5 H	
	Data integration	5 H	
	Data reduction	10 H	
	Web mining (web scraping)	5 H	
Module 2	Data visualization	20 H	70 H
	Scikit learn	20 H	
	Implementation of kaggle data set projects	29 H	
	Who is data analyst	1 H	

Course In Detail

Module - 1

DATA ANALYSIS

- What is data science
- Applications
- Types of data analysis
- Data analytics tools

PYTHON LIBRARIES FOR DATA SCIENCE

- Introduction to series and dataframes
- Data using Numpy
- Pandas
- Matplotlib
- SciPy
- Data analysis process
- Data Cleaning Introduction, with text function, dates and times
- Summarizing the Data
- Statistical Inference, types, importance

DATA MINING CONCEPTS AND APPLICATIONS

- Introduction to data mining
- Data Trial Uses of Mining
- Application of Data Mining
- Major issues in Data Mining
- Stages of Data Mining
- Selection of target data: Introduce Dataset sources

DATA PREPROCESSING

- MAJOR TASK IN DATA PREPROCESSING
 - Data cleaning
 - Data integration
 - Data reduction
 - Data transformation
- DATA SCIENCE ALGORITHMS
 - Linear regression
 - Logistic regression

- Decision trees
- Naive bayes
- K nearest neighbors
- K means clustering
- Support vector machine
- Principal component analysis
- Neural networks
- Random forest

DATA WRANGLING

- Tasks
- Importance
- 6 basic steps of data wrangling
- Tools and techniques used for data wrangling
- Data wrangling in python
- Algorithm for wrangling

DATA INTEGRATION

- Importance
- Tools and techniques
- Process for Integrating Data

DATA REDUCTION

- Basics
- Tools and techniques
- Steps involved in reduction

WEB MINING (WEB SCRAPING)

- Web content mining
- Web usage mining – (log files)
- Web structure mining
- Working example of web scraping using scrapy

Module - 2

DATA VISUALIZATION

- Matplotlib
 - Basics of Matplotlib
 - Creating Line Charts
 - Creating Bar Charts and Histograms
 - Plot Customization: Titles, Labels, Legends, and Gridlines
 - Subplots and Layout Adjustments
 - Saving and Exporting Visualizations

- Seaborn
 - Creating Scatterplots, Lineplots, and Barplots
 - Using Pairplots for Data Relationships
 - Heatmaps: Correlation and Customization
- Plotly
 - Basics of interactive visualizations with Plotly
 - Create Line, Scatter, and Bar Charts
 - Create Pie Charts and Treemaps
 - Exporting Plotly Visualizations
- Power BI
 - Introduction to Power BI and Installation
 - Connecting to Datasets (Excel, CSV, SQL)
 - Creating Bar, Area, and Line Charts
 - Interactive Dashboards: Filters
 - Working with Calculated Columns and Measures
 - Understand Aggregation, Granularity, and Level of Detail
 - Best Practices for Visualizing Data
 - Practical Project: Build a Dashboard

SCIKIT-LEARN

- What is scikit-learn
- Installation
- Features
- Data representation
- Training and Testing Algorithm
- Label Encoder and One Hot Encoder
- Packages of Scikit Learn

IMPLEMENTATION OF KAGGLE DATASET PROJECTS

- IRIS Dataset
- City Accident Analysis Dataset
- Flipkart, Amazon Dataset using Scrapy
- Pima Diabetes Dataset
- Titanic Dataset

WHO IS A DATA ANALYST?

- Data Analyst Jobs and Description
- Data Scientist in a Hospital Scenario
- Data Analyst Resume Description
- Data Preparation Videos
- Kaggle Dataset Implementation of Data Analytics using Machine Learner
- Hacker Earth Exam Scenar