Data Science Training Program

Overview

This comprehensive Data Science Training Program is designed to transform beginners into industry-ready data scientists. The program covers the entire data science workflow — from understanding data and performing exploratory analysis to building machine learning models and deploying them in real-world applications. Learners will gain hands-on experience through mini-projects, real datasets, and a capstone project, ensuring a strong practical foundation.

Objectives

- Provide a solid understanding of data science fundamentals, tools, and workflows.
- Teach Python programming, data manipulation, and visualization techniques.
- Develop the ability to apply machine learning algorithms to solve real-world problems.
- Introduce deep learning and neural networks for advanced applications.
- Prepare learners for job opportunities, freelancing, and portfolio building.

Key Modules

1. Python for Data Science – NumPy, Pandas, Data Cleaning
2. Data Visualization – Matplotlib, Seaborn, Plotly, Dashboards
3. Machine Learning – Regression, Classification, Clustering
4. Deep Learning – Neural Networks, ANN using TensorFlow/PyTorch
5. Project Deployment

Outcomes

1. Ability to analyze, clean, and visualize data effectively.
2. Build and evaluate ML models for prediction and classification tasks.
3. Develop end-to-end projects and deploy them as web apps.
4. Gain confidence in handling real-world datasets and problem statements.
5. Build a strong portfolio for jobs, internships, or freelancing opportunities.

Curriculum:

1. **Introduction to Data Science**
    - Learn what Data Science is and its real-world applications
    - Explore career paths and tools (Python, Jupyter, GitHub)
    - Understand the data science workflow

2. **Python for Data Science**
    - Master Python basics: variables, loops, functions
    - Work with NumPy for numerical computing
    - Use Pandas for data manipulation & cleaning
    - Load and process CSV, Excel, and JSON files.
3. **Data Visualization**
    - Create plots with Matplotlib, Seaborn, and Plotly
    - Build line, bar, scatter, and pie charts
    - Visualize trends with heatmaps, pairplots
    - Intro to interactive dashboards

4. **Statistics & Probability**
    - Learn mean, median, mode, variance, and standard deviation
    - Study probability distributions
    - Perform hypothesis testing
    - Analyze correlation & covariance

5. **Exploratory Data Analysis (EDA)**
    - Clean data: handle missing values and outliers
    - Perform feature engineering & transformation
    - Normalize and scale datasets
    - Conduct a real-world EDA case study

6. **Machine Learning Basics**
    - Understand supervised vs. unsupervised learning
    - Implement Linear & Logistic Regression
    - Learn Decision Trees & Random Forests
    - Evaluate models with accuracy, precision, recall

7. **Advanced Machine Learning**
    - Learn Support Vector Machines (SVM)
    - Perform K-Means clustering
    - Apply dimensionality reduction (PCA)

Introduction to Deep Learning

- Understand the basics of neural networks
- Learn TensorFlow or PyTorch fundamentals
- Build a simple ANN model for classification

8. **Data Science Projects**