HTTPS://FIFTHFORCE.IN
PH: +917003720431

Fifth Force
Giving Wings to Your Dream

Course Title: Real World ML & AI: Scratch to Impact

Duration: 6 Week
Who it is for: Pursuing B.Tech/B.E./MCA/M.Tech/BCA/M.Sc(IT)/B.Sc(IT)
Mode of Training: Live, Online, Instructor Led
Access to Recording: Yes
Pre-Requisite: No Pre-requisite, Concepts would be built from Scratch
Teaching Methodology: No to Minimum theory, maximum Hands on. Object oriented; problem driven.
Doubt Session: Weekly Doubt Session to ensure no one is left behind
Certificate: Industry Endorsed Certificate (Jointly Signed By Our Primary Industry Partner)
Project:

1) Module Specific 3-4 Mini-Project.
2) Hand On Major Project
3) Sample Project
4) Capstone Project

Teaching Approach:

- **Interactive Lectures**: For concept clarity.
- **Hands-On Labs**: Real World Simulated datasets and exercises.
- **Capstone Project**: Integration of all tools and concepts.
- **Feedback Sessions**: Instructors review and guide project presentations.

Course Objectives:

- Develop Python programming skills for Machine Learning workflows.
- Build foundational knowledge of Statistics and SQL tailored for ML tasks.
- Master supervised and unsupervised ML algorithms.
- Gain hands-on experience through practical projects and real-world datasets.

Detailed Syllabus:

Module 1: Python Programming for Machine Learning (10 Hours)
Learning Objectives:

- Build foundational Python programming skills for ML tasks.
- Learn data manipulation, visualization, and cleaning with Python libraries.

Content:

1. **Python Basics for ML** (4 Hours)
    - Python Syntax, Variables, Data Types, and Control Structures.
    - Functions, Loops, and File Handling.
2. **Data Manipulation with Pandas** (4 Hours)
    - DataFrames: Creation, Indexing, and Selection.
    - Data Cleaning: Handling Missing Values and Duplicates.
    - Grouping, Aggregating, and Filtering Data.
3. **Data Visualization with Matplotlib and Seaborn** (2 Hours)
    - Line, Bar, Scatter, and Histogram Plots.
    - Advanced Visualizations: Heatmaps and Pairplots.

Mini-Project: Clean and visualize a dataset (e.g., sales or weather data) using Pandas and Seaborn.

Module 2: SQL for Machine Learning (8 Hours)
Learning Objectives:

- Use SQL to query, manipulate, and preprocess data for ML workflows.
- Integrate SQL with Python for seamless analysis.

Content:

1. **SQL Basics** (3 Hours)
    - Writing Queries: SELECT, WHERE, GROUP BY, and ORDER BY.
    - Understanding Relationships and JOINs (INNER, LEFT, RIGHT).
2. **Data Manipulation with SQL** (3 Hours)
    - Aggregation Functions: COUNT, SUM, AVG, MAX, MIN.
    - Writing Subqueries for Complex Data Analysis.

3. **SQL-Python Integration** (2 Hours)
    - Using Python's sqlite3 Library to Connect to Databases.
    - Importing SQL Query Results into Pandas DataFrames.

Mini-Project: Query and preprocess a database of customer transactions for ML workflows.

Module 3: Statistics for Machine Learning (6 Hours)
Learning Objectives:

- Understand the statistical concepts critical for ML models.
- Apply descriptive and inferential statistics in Python.

Content:

1. **Descriptive Statistics** (3 Hours)
    - Measures of Central Tendency: Mean, Median, Mode.
    - Measures of Spread: Range, Variance, and Standard Deviation.
    - Visualizing Distributions: Histograms and Boxplots.
2. **Inferential Statistics** (3 Hours)
    - Confidence Intervals and Sampling Techniques.
    - Hypothesis Testing: t-Tests and p-Values.
    - Correlation vs. Causation.

Mini-Project: Analyze student performance data to calculate summary statistics and test hypotheses using Python.

Module 4: Data Preprocessing and Feature Engineering (6 Hours)
Learning Objectives:

- Preprocess and prepare data for ML model training.
- Engineer features to improve model accuracy.

Content:

1. **Data Cleaning** (2 Hours)
    - Handling Missing, Outlier, and Inconsistent Data.
    - Encoding Categorical Variables.

2. **Feature Engineering and Scaling** (2 Hours)
    - Feature Extraction and Selection Techniques.
    - Scaling: Normalization and Standardization.
3. **Data Splitting and Cross-Validation** (2 Hours)
    - Train-Test Splits.
    - K-Fold Cross-Validation.

Mini-Project: Preprocess a house pricing dataset for ML model building.

Module 5: Supervised Learning Algorithms (12 Hours)
Learning Objectives:

- Understand supervised learning concepts and their applications.
- Build and evaluate regression and classification models.

Content:

1. **Regression Models** (6 Hours)
    - Linear and Polynomial Regression.
    - Regularization: Ridge and Lasso Regression.
    - Evaluation Metrics: MSE, RMSE, R-squared.
2. **Classification Models** (6 Hours)
    - Logistic Regression, Decision Trees, and Random Forests.
    - Gradient Boosting Algorithms (e.g., XGBoost).
    - Evaluation Metrics: Precision, Recall, F1-Score, ROC-AUC.

Mini-Project: Build a regression model for house price prediction and a classification model for customer churn analysis.

Module 6: Unsupervised Learning Algorithms (8 Hours)
Learning Objectives:

- Learn and implement unsupervised learning methods.
- Discover patterns and reduce dimensions in datasets.

Content:

1. **Clustering Techniques** (4 Hours)
    - k-Means and Hierarchical Clustering.
    - Density-Based Clustering with DBSCAN.

2. **Dimensionality Reduction** (4 Hours)
    - Principal Component Analysis (PCA).
    - t-SNE for Data Visualization.

Mini-Project: Cluster customers based on transaction data and visualize high-dimensional data using PCA.

Module 7: Capstone Project (10 Hours)
Objective:
Combine knowledge of Python, SQL, Statistics, and ML to solve a real-world data problem.
Structure:

1. **Dataset**: A comprehensive dataset (e.g., sales, healthcare, or finance).
2. **Workflow**:
    - **Preprocessing**: Clean and prepare the data using Python and SQL.
    - **EDA and Statistics**: Explore data trends and perform hypothesis testing.
    - **Modeling**: Apply both supervised and unsupervised ML models.
    - **Evaluation**: Optimize models using metrics and hyperparameter tuning.
3. **Deliverables**:
    - Python notebooks demonstrating the full ML pipeline.
    - A report summarizing findings, methodologies, and actionable insights.