Data Science for Fraud Detection and Risk Management

Duration: 05 days

Course Outcomes:

- Apply data science techniques to solve business problems.
- Process and analyze data for meaningful insights.
- Develop machine learning models for classification, regression, and clustering.
- Evaluate and communicate the results of data science projects, with a focus on fraud detection and risk analysis.

Course Outline

Lesson 1: Addressing Business Issues with Data Science

- **Topic A: Initiate a Data Science Project**
- **Topic B: Formulate a Data Science Strategy**
- **Topic C: Understanding Fraud and Risk in Business**
    - Impact of fraud and risk on organizations.
    - Case studies of data science applications in fraud detection and risk management.

Lesson 2: Extracting, Transforming, and Loading Data

- **Topic A: Extract Data**
- **Topic B: Transform Data**
- **Topic C: Load Data**
- **Topic D: Data Preprocessing for Fraud and Risk Analysis**
    - Techniques for preparing fraud and risk datasets.

Lesson 3: Analyzing Data

- **Topic A: Examine Data**
- **Topic B: Explore the Underlying Distribution of Data**
- **Topic C: Use Visualizations to Analyze Data**
- **Topic D: Analyzing Fraud and Risk Data**
    - Techniques for identifying anomalies and patterns.

Lesson 4: Designing a Machine Learning Approach

- **Topic A: Identify Machine Learning Concepts**

- **Topic B: Test a Hypothesis**
- **Topic C: Risk Factors in Data**
    - Understanding key variables impacting risk and fraud.

Lesson 5: Developing Classification Models

- **Topic A: Train and Tune Classification Models**
- **Topic B: Evaluate Classification Models**
- **Topic C: Fraud Detection with Classification Models**
    - Use case examples and evaluation metrics specific to fraud detection.

Lesson 6: Developing Regression Models

- **Topic A: Train and Tune Regression Models**
- **Topic B: Evaluate Regression Models**
- **Topic C: Risk Assessment with Regression Models**
    - Use case examples in assessing risk (e.g., credit scoring).

Lesson 7: Developing Clustering Models

- **Topic A: Train and Tune Clustering Models**
- **Topic B: Evaluate Clustering Models**
- **Topic C: Clustering for Anomaly Detection**
    - Techniques for identifying unusual patterns indicative of fraud.

Lesson 8: Finalizing a Data Science Project

- **Topic A: Communicate Results to Stakeholders**
- **Topic B: Demonstrate Models in a Web Application**
- **Topic C: Implement and Test Production Pipelines**
- **Topic D: Case Studies in Fraud and Risk Analysis**
    - Present case studies showcasing successful projects.