

Institute of Business and Economic Research

United **International** University



COURSE OUTLINE



Post Graduate Diploma in **Data Science (PGDDS)**

Institute of Business and Economic Research (IBER)

Course Outline

Post Graduate Diploma (PGD) in Data Science

2 Terms (6 Courses | 18 Credits)

Term I

(3 Courses | 9 Credits)

COURSE 1

PGDDS 101: PROGRAMMING ON CORE PYTHON

Module 01: Python Fundamentals

- Introduction to Python
- Basic Python Syntax
- Language Components and Loops
- Lists, Tuples, Sets, and Dictionaries

Module 02: Functions and Object-Oriented Programming

- Functions
- Classes, Objects, and Methods
- Inheritance
- Encapsulation

Module 03: Data Analysis with Python

- NumPy for Data Analysis
- Pandas – Part I
- Pandas – Part II
- Project using Pandas

Module 04: Problem Solving and Utilities

- Problem Solving
- Review Session
- Date and Time Module
- Math Module

COURSE 2

PGDDS 102: MACHINE LEARNING USING PYTHON

Module 01: Introduction to Machine Learning

- Introduction to Machine Learning

Module 02: Data Preprocessing

- Feature Scaling
- Principal Component Analysis (PCA)
- Label Encoding
- Ordinal Encoding
- One-Hot Encoding
- Outlier Removal

Module 03: Regression Models

- Simple Linear Regression
- Multiple Linear Regression
- Logistic Regression

Module 04: Classification Algorithms

- K-Nearest Neighbors (KNN)
- Naïve Bayes Classification
- Decision Tree

Module 05: Ensemble Learning

- Random Forest
- Bootstrap
- Bagging
- Entropy and Information Gain

Module 06: Advanced Machine Learning

- Support Vector Machines (SVM)
- Saving and Loading Trained Models
- Google Colab

Module 07: Unsupervised Learning

- K-Means Clustering

Module 08: Natural Language Processing

- Introduction to NLP

COURSE 3

PGDDS 103: DATA ANALYSIS & DATA VISUALIZATION

Module 01: Data Analysis Fundamentals and Process

- Problem framing and decision context
- Types of data analytics
- Data concepts and variables
- Data analysis process
- AI in data analysis
- Responsible analytical practices

Module 02: Data Acquisition, Inspection, and Quality Assessment

- Sources of data
- Importing data
- DataFrame basics
- Inspecting data
- Detecting anomalies
- Data quality assessment
- Data documentation

Module 03: Data Cleaning, Transformation, and Integration

- Prompt engineering for data wrangling
- Data correction and cleaning
- Handling missing values and anomalies
- Data manipulation
- Feature engineering
- Data aggregation
- Data reshaping
- Data integration
- Data validation

Module 04: Exploratory Data Analysis (EDA)

- Frequency distributions
- Summary statistics
- Outlier analysis
- Segmented summaries
- Cross-tabulations
- Data visualization
- Pattern identification

Module 05: Statistical Analysis and Modeling

- Statistical inference
- t-test and ANOVA
- Chi-square test
- Multiple Linear Regression
- Logistic Regression
- Model evaluation
- Causation vs. Association

Module 06: Data Visualization with Python

- Chart selection
- Basic and advanced visualizations
- Matplotlib
- Seaborn
- Visualization principles
- AI-assisted visualization

Module 07: Interactive Dashboards with Power BI

- Power Query
- Data Modeling
- DAX Fundamentals
- KPIs
- Dashboard Design
- Interactive Reports
- AI Features

Module 08: Data Storytelling and Reporting

- Audience-oriented reporting
- Business storytelling
- Communicating uncertainty
- Report writing
- Reproducible analysis
- AI-assisted documentation

Term I
(3 Courses | 9 Credits)

COURSE 4

PGDDS 201: DATABASE DESIGN AND DEVELOPMENT (SQL SERVER)

Module 01: Introduction to Database Systems

- Introduction
- RDBMS
- Normalization
- NoSQL

Module 02: SQL Fundamentals

- DDL, DML, DCL, TCL
- SQL Comments
- Database Keys
- Constraints
- Nested Queries
- Joins
- Set Operations
- Dates, Auto Increment
- Views
- Stored Procedures
- Triggers

Module 03: Constraints and Identity

- Primary Key, Foreign Key
- Unique, Not Null
- Default
- Check Constraint

Module 04: Querying Data

- SELECT Statement
- WHERE Clause
- ORDER BY, DISTINCT
- TOP, LIKE
- DELETE vs TRUNCATE

Module 05: SQL Operators

- Arithmetic Operators
- Comparison Operators

- Logical Operators
- GROUP BY
- HAVING

Module 06: SQL Joins

- INNER JOIN, LEFT JOIN
- RIGHT JOIN, FULL JOIN
- CROSS JOIN
- SELF JOIN
- MERGE
- PIVOT

Module 07: SQL Functions

- ISNULL
- COALESCE
- CASE
- UNION
- INTERSECT
- EXCEPT
- Execution Plan

Module 08: Built-in Functions

- String Functions
- Date Functions
- CAST
- CONVERT
- Wildcards

Module 09: Advanced SQL

- Temporary Tables
- Subqueries
- CTE
- Views
- Triggers

Module 10: Flow Control

- BETWEEN, ALIASES
- NULLIF, IF...ELSE
- WHILE LOOP
- INSERT INTO SELECT

Module 11: Stored Procedures

- Overview
- Parameters
- Variables
- Output Parameters
- Advantages

Module 12: Indexing and Transactions

- Clustered Index
- Non-Clustered Index
- SQL Transactions

COURSE 5

PGDDS 202: DEEP LEARNING USING PYTHON

Module 01: Orientation and ANN Fundamentals

- Data Science
- Artificial Intelligence
- Artificial Neural Networks
- Deep Learning Applications
- Single-Layer Perceptron

Module 02: Multi-Layer Perceptron

- Concept
- Algorithm
- Mathematical Example
- Project

Module 03: Convolutional Neural Network (CNN)

- Concept
- Algorithm
- Mathematical Example
- Project

Module 04: Self-Organizing Map (SOM)

- Concept
- Algorithm
- Mathematical Example
- Project

Module 05: Restricted Boltzmann Machine (RBM)

- Concept
- Algorithm
- Mathematical Example
- Project

Module 06: Autoencoder

- Concept

- Algorithm
- Mathematical Example
- Project

Module 07: Deep Belief Network (DBN)

- Concept
- Algorithm
- Mathematical Example
- Project

Module 08: Midterm Assessment

- Midterm Examination
- Two Class Tests

Module 09: Recurrent Neural Network (RNN)

- Concept
- Algorithm
- Mathematical Example
- Project

Module 10: Long Short-Term Memory (LSTM)

- Concept
- Algorithm
- Mathematical Example
- Project

Module 11: Generative Adversarial Network (GAN)

- Concept
- Algorithm
- Mathematical Example
- Project
- Review Class

COURSE 6

PGDDS 203: DATA SCIENCE PROJECT

Module 01: Project Orientation

- Data science project lifecycle
- Agile methodology
- Research questions
- Project proposal
- Team formation

Module 02: Python for Projects

- Python review
- NumPy
- Pandas
- Scikit-learn
- GitHub
- Documentation

Module 03: Data Collection and Preprocessing

- CSV and Excel
- Web scraping
- Missing values
- Outliers
- Encoding
- Feature extraction
- PCA

Module 04: Exploratory Data Analysis

- Descriptive statistics
- Correlation
- Matplotlib
- Seaborn
- Plotly
- Storytelling

Module 05: Project 1 – Classification

- Decision Tree

- SVM
- Naïve Bayes
- Model Evaluation
- Hyperparameter Tuning

Module 06: Project 2 – Clustering and Regression

- KNN
- K-Means
- Linear Regression
- Multiple Regression
- Logistic Regression

Module 07: Project 3 – Ensemble Learning

- Random Forest
- Bagging
- Boosting
- Voting
- Cross-validation
- Feature Importance

Module 08: Midterm Project Review

- Progress Presentation
- Peer Review
- Interim Report
- Feedback Session

Module 09: Real-Life Predictive Analytics

- Stock Price Prediction using LSTM
- Sales Forecasting
- Time-Series Analysis
- Waiter Tips Prediction

Module 10: NLP and Sentiment Analysis

- Online Review Analysis
- Tokenization
- Stopwords
- TF-IDF
- NLP Models

Module 11: Recommendation Systems

- Product Recommendation
- Collaborative Filtering
- Content-Based Filtering

- Final Report Writing
- Presentation Preparation

Module 12: Final Project Presentation

- Project Demonstration
- Presentation
- Peer Evaluation
- Viva and Q&A
- Course Reflection