DATA SCIENCE & ANALYTICS
COURSE SYLLABUS

Module 1 — Excel for Data Analysis

- Excel Basics (Functions, Shortcuts, Formatting)
- Data Cleaning & Handling Large Datasets
- Pivot Tables & Pivot Charts
- Lookup Functions (VLOOKUP, HLOOKUP, XLOOKUP)
- Conditional Formatting & Data Validation
- What-If Analysis & Solver
- Excel Dashboards & Reporting
- Introduction to Power Query

Module 2 — Python Core and Advanced

- Python basics → Variables, Data types, Operators
- Control structures → If-Else, Loops
- Functions & Scope
- Object-Oriented Programming (OOP) → Classes, Inheritance, Polymorphism
- Error handling & Exception management
- File handling (CSV, JSON, Excel)
- Advanced concepts → Iterators, Generators, Decorators, Lambda functions
- Data analysis libraries → NumPy (arrays, operations), Pandas (dataframes, merging, grouping), Matplotlib/Seaborn (visualization)

Module 3 — Exploratory Data Analysis (EDA) using Python

- Data import & preprocessing → Handling missing values, duplicates
- Outlier detection & treatment
- Feature engineering → Encoding, Scaling, Transformation
- Correlation & Covariance analysis
- Univariate, Bivariate & Multivariate analysis
- Visualization for EDA → Histograms, Heatmaps, Pairplots, Boxplots
- Case studies using real datasets

Module 4 — Advanced Statistics

- Descriptive statistics → Mean, Median, Mode, Variance, Std. Dev.
- Probability distributions → Normal, Binomial, Poisson, Uniform
- Inferential statistics → Sampling, Estimation, Confidence intervals
- Hypothesis testing → Z-test, T-test, Chi-square test, ANOVA
- Correlation & Regression analysis
- P-value interpretation & statistical significance
- Advanced regression diagnostics → Multicollinearity, Heteroscedasticity

Module 5 — SQL for Data Analysis

- SQL basics → SELECT, WHERE, ORDER BY
- Joins → Inner, Left, Right, Full
- Subqueries & Nested queries
- Common Table Expressions (CTEs)
- Window functions → RANK, DENSE_RANK, ROW_NUMBER, LEAD, LAG
- Aggregate functions → SUM, AVG, COUNT, GROUP BY, HAVING
- Stored procedures, Views, Indexing
- SQL case studies on business datasets

Module 6 — Machine Learning – Supervised Learning

- ML workflow → Data preparation, training, testing, evaluation
- Regression algorithms → Linear, Multiple, Polynomial, Ridge, Lasso
- Classification algorithms → Logistic Regression, Decision Trees, Random Forest, Naïve Bayes, SVM, Gradient Boosting, XGBoost
- Model evaluation metrics → Accuracy, Precision, Recall, F1-score, ROC-AUC
- Feature selection techniques
- Cross-validation, GridSearchCV & Hyperparameter tuning

Module 7 — Machine Learning – Unsupervised Learning

- Clustering → K-Means, Hierarchical Clustering, DBSCAN
- Dimensionality Reduction → PCA, t-SNE
- Association Rule Learning → Apriori, Market Basket Analysis
- Anomaly detection techniques
- Customer segmentation projects

Module 8 — Artificial Intelligence and Deep Learning

- Introduction to Neural Networks → Perceptron, Feed-forward, Backpropagation
- Activation functions → Sigmoid, ReLU, Tanh, Softmax
- Optimization techniques → Gradient Descent, Adam Optimizer
- Deep learning frameworks → TensorFlow, Keras, PyTorch basics
- Architectures → ANN, CNN, RNN, LSTM
- Regularization techniques → Dropout, Batch Normalization
- Transfer learning & Pre-trained models

Module 9 — Computer Vision (CV)

- Image preprocessing → Resizing, Normalization, Augmentation
- Convolutional Neural Networks (CNNs) → Filters, Pooling, Feature Maps
- Image classification & recognition projects
- Object detection → YOLO, Faster R-CNN basics
- OpenCV for image processing → Edge detection, Face detection, Image transformations
- Real-time applications → Face recognition, Image captioning

Module 10 — Natural Language Processing (NLP)

- Text preprocessing → Tokenization, Stopword removal, Stemming, Lemmatization
- Representations → Bag-of-Words, TF-IDF
- Word embeddings → Word2Vec, GloVe, FastText, BERT basics
- Sentiment analysis projects
- Topic modeling → LDA
- Sequence models → RNN, LSTM basics for NLP
- Chatbot development fundamentals

Module 11 — Power BI

- Introduction to BI & Data Visualization
- Data loading & transformation with Power Query
- Data modelling & relationships
- DAX functions & calculated columns
- Building interactive reports & dashboards
- Advanced charts & KPIs
- Drill-through, Filters, Bookmarks
- Publishing & Sharing dashboards in Power BI service