Basic Data Science Course

Duration: 4 Months

Sessions: 2 sessions/week

Module 1: Python for Data Science

- **Python Essentials:** Variables, Data Types, Loops, Functions
- **Advanced Python**: List Comprehension, Lambda, Map/Reduce, OOP Basics
- **Numpy & Pandas**: Arrays, Data Frames, Indexing, Filtering, Slicing
- **Data Grouping and Visualization**: Groupby, Aggregation, Visualization (Matplotlib & Seaborn)

Module 2: Statistics and Exploratory Data Analysis(EDA)

- **Descriptive Statistics**: Mean, Median, Mode, Variance, Standard Deviation
- **Inferential Statistics**: Hypothesis Testing, p-value, t-test, chi-square test, ANOVA test.
- **Probability & Distributions**: Normal, Binomial, Central Limit Theorem
- **EDA:** Data Cleaning, Outlier Detection, Transformation, Correlation Analysis, Univariate Analysis, Bi-Variate Analysis
- **Data Preprocessing:** Imputation, Encoding(One-Hot-Encoding, Label Encoding, etc.), feature engineering.
- **Feature Selection:** RFE (Recursive Feature Elimination), SelectKBest

Module 3: Machine Learning

- **ML Basics**: Supervised, Unsupervised learning, ML Pipeline, Train-Test Split.
- **Regression Models**: Linear, Multiple, Regularization (Ridge, Lasso), Polynomial Regression
- **Classification Models**: Logistic Regression, Decision Trees, KNN, SVM
- **Ensemble Learning:** Bagging, Boosting, Stacking, XGBM, LBBM, XGBOOST, Naive Bayes
- **Model Evaluation Metrics**: Accuracy, Confusion Matrix, ROC AUC, Cross-validation, R2, MSE, RMSE, MAE
- **Unsupervised Learning**: Hierarchical, K-Means, DBSCAN Clustering, PCA & t-SNE, Association Rule, Recommendation System
- **Model Validation and Optimization:** K-Fold Cross Validation, Overfitting/Underfitting handling (Regularization, Early Stopping), Bias-Variance Tradeoff, Hyperparameter Tuning

Module 4: Deep Learning

- **Neural Network:** Artificial Neural Network(ANN), Forward Propagation, Back Propagation, Perceptron, Activation Functions (Sigmoid, ReLU, Tanh)
- **Evaluation Metrics:** Loss Functions (MSE, Cross Entropy), Gradient Descent & Variants (SGD, Adam, RMSProp)
- **Keras and TensorFlow:** Basics (Sequential API), Epochs, Batches, Learning Rate, Dense Layers, Compile, Fit, Dropout
- **Forecasting:** Time Series Analysis, Moving Average Method, ARIMA, SARIMA, Prophet

Note: Weekly assignments on each topic. Capstone Project(End-to-End ML Project)

Course Fees: ₹80,000/ Per Candidate

Advanced Data Science Course

Duration: 2 Months

Sessions: 2 sessions/week

Module 1: Natural Language Processing(NLP)

- **Text preprocessing:** Tokenization, Stopwords, Lemmatization, Stemming
- **Text Normalization:** Lowercasing, Punctuation Removal, Spelling Correction
- **Feature Engineering for Text:** Bag of Words (BoW), TF-IDF
- **Word Embeddings:** Word2Vec, GloVe, FastText
- **NLP Tasks & Applications:** Text Classification, Named Entity Recognition (NER), Sentiment Analysis, n-grams

Module 2: Sequential Deep Learning

- **Sequence Data & RNNs:** Recurrent Neural Networks (RNN), Implementing RNN in Keras, Vanishing Gradient Problem
- **Long Short-Term Memory (LSTM):** LSTM internals: Gates and Memory Cells, • Bidirectional LSTM, Stacked LSTM, Keras Implementation
- **Gated Recurrent Unit (GRU):** GRU vs LSTM, Build GRU-based Text Classifier, Time Series Forecasting with LSTM/GRU

Module 3: Transformers & Attention Mechanisms

- **Attention Mechanism:** Attention is All You Need, Multi-Head Attention
- **Transformer Architecture:** Encoder-Decoder Structure, Positional Encoding, Transformer from Scratch in PyTorch/TensorFlow

- **Hugging Face Transformers:** Introduction to Hugging Face ecosystem, Using pre-trained models (BERT, RoBERTa, DistilBERT), Tokenizers and Pipelines

- **Applications of Transformers:** Text Classification using BERT, Question Answering, Summarization, Translation

Course Fees: ₹30,000/ Per Candidate