BIG DATA & VISUAL ANALYTICS

PROGRAM CURRICULUM

Quantitative Methods

Linear Algebra; Matrices, Calculus: Differentiation, Integration (single variate), Optimisation, Discrete Mathematics, Simplex Methods, Assignment & Transportation, Game Theory, Graph Theory, Brownian Motion, Lebesgue Sample Space, Measure Theory, Martingale

Statistics

R-studio Fundamentals, Programming in R, Data Handling, Transformations, Descriptive Statistics, Probability Theory; Central Tendencies, Distributions, Regression, Stochastic Processes, Hypothesis Testing, Time Series

Databases

ER Modelling, SQL, Indexes and Constraints, Relational Databases, MySQL Workbench, NoSQL Theory and Clustered Databases, Document Databases such as MongoDB, Graph Databases, Google Firebase / AWS Redshift

Data Visualisation

Visual Cognition, Perception, Analytical Design, Dashboard and Storytelling, Tableau, Work Sheet & Dashboard Actions, Guided Analytics, JavaScript, React-REDUX

Data Structures & Algorithms Using Python

Python Programming; Data Representation such as JSON and XML, Python Scripting, Objects & Data Types, Functions, Strings, Boolean Logic, Data Libraries; SciPy, NumPy, Pandas, Matplotlib Algorithmics: Searching & Sorting, Divide & Conquer, Dynamic Programming, Augmented Search Structure, Amortised Analysis, Monte Carlo Simulation

Machine Intelligence & Deep Learning

Neural Networks, Perceptron, Confusion Matrix, Kernel Trick, Supervised Learning: Linear and Logistic Regression, Decision Trees and Random Forests, Support Vector Machines, Unsupervised Learning: Hierarchical Clustering, K-means Gaussian Mixture Model, Dimension Reduction, Linear Discriminant Analysis, Ensemble Methods: Bias-variance Decomposition, Over-fitting, Random Forest, AdaBoost Algorithm, Gradient Boosting, Reinforcement Learning. Packages – Keras / Theano / TensorFlow

Natural Language Processing

Naïve Bayes Theorem, Markov Model, Support Vectors, Probabilistic Language Modelling, N-Grams, NLTK (Python Libraries), POS Tagging, Parsing, Semantics, Information Retrieval and Extraction, Sentiment Analysis

Recommender Systems

Collaborative Filtering, Content-based Filtering, Hybrid Models

Data Mining

Artificial Neural Nets, Data Exploration & Visualisation, Classification, Association Analysis, Clustering, Anomaly Detection

Functional Programming in Scala

Scala Basics, Types, Classes, Special Methods, Currying, Types, Implicits, Anonymous Classes, Var Args, Partial Functions, Recursion, Collections, For Loops

Data Engineering (Hadoop & Apache Spark)

Design, Hadoop: Introduction to Big Data, Hadoop Setup, Map Reduce, Yarn Architecture, Hive, Pig, Sqoop, Flume - Data Ingestion, SQL, NoSQL, H-base, Kafka, Zookeeper, Oozie Apache Spark: Stream Processing, Spark Streaming Library with PySpark, Using SQL & Dataframe

Cloud Computing & DevOps

AWS EC-2, S3, Virtualisation, CI/CD with Jenkins, Google Kubernetes

BIG DATA & VISUAL ANALYTICS

Banking & Financial Analytics

Stochastic Calculus, Value at Risk, CAPM, Black-Scholes Theorem, Volatility Estimation: Exotic Options, Stochastic and Local Volatility Models, RSI, Equivalent Martingale Measure Approach, Interest Rate Derivatives, Risk Analysis, Credit Research

Marketing Analytics

CRM with Salesforce, Customer Segmentation, HubSpot Digital Marketing Models, ML-based customer segmentation and prediction for cross-sell/up-sell, Review of Algorithms, Digital Banking and Channel Effectiveness, Digital Media, Game Theory, Competitive Marketing Strategy, Mining social data to segment, target, create campaigns, execute, visualize and predict customer churns, Analytics