PROFESSIONAL PROGRAM IN BIG DATA & VISUAL ANALYTICS
CURRICULUM

DATA SCIENCE FOUNDATION:

Mathematics for Data Science
- MATHS: Linear Algebra; Matrices, Calculus: Differentiation, Integration, Optimization, Discrete Mathematics
- STATS: Probability theory; Central Tendencies, Distributions, Hypothesis Testing, Time Series

R Programming with Advanced Statistics
- R-studio fundamentals, Programming, data handling, transformations, regression, Stochastic Processes, Structures

RDBMS and SQL
- RDBMS, data modelling, ER Modelling, SQL Queries, Indexes and Constraints, MySQL workbench, NoSQL theory and clustered databases, Document databases such as MongoDB, Column family stores such as Cassandra, Graph databases such as Neo4j

Data Visualization using Tableau
- Data Visualization: Visual Cognition, Perception, Analytical Design, Dashboard and storytelling, Tableau, Special chart types, Work Sheet & Dashboard Actions, Guided Analytics, R integrated with Tableau

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; SciPi, NumPi, Pandas, MatPotLib
- Algorithmics: Searching & Sorting, Divide & Conquer, Dynamic Programming, Augmented search structure, Amortized analysis, Monte Carlo Simulation

DATA SCIENCE ADVANCED:

Machine Learning
- Supervised Learning: Linear and Logistic Regression with and without Regularization (L1, L2, ElasticNet is optional), Decision Trees and Random Forests, optionally Gradient Boosting Trees, Support Vector Machines, Structural Risk Minimization, Detect Over-fitting, Tensor Flow

Neural Networks
- Perceptron: Multi-layer perception, Backpropagation, Perceptron, Multi-layer perception, Backpropagation, Duality, KKT Condition, Kernel Trick.

Deep Learning
- Unsupervised Learning: Hierarchical Clustering, K-means Gaussian mixture model, Dimension reduction, Linear Discriminant Analysis
- Ensemble Learning: Bias-variance decomposition, Over-fitting, Bagging, Random Forest, Adaboost algorithm, Gradient boosting

Data Mining using Excel Miner
- Data Mining, Data Exploration & Visualization, Classification, Association Analysis, Clustering, Anomaly Detection

Apache Hadoop & Spark
- Data Warehousing Architecture, 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

NLP & NN
- Predictive Analytics, Neural Networks, Text Mining and NLP, Latent Dirichlet Allocation, Techniques: N-Grams, POS Tagging, Parsing, Semantics, Information Retrieval and Extraction, Sentiment Components, Application

SPECIALISATIONS

Stream A: Banking & Financial Analytics
- Business Case Analysis and Modelling
- Quant for BFSI: Value at Risk. Principal component analysis. volatility estimation: ARCH, GARCH. Exotic Options. Stochastic and local volatility models. RSI, Equivalent martingale measure approach. Interest rate derivatives, standard market models, Algo Models and developments in Indian Equity Markets
- CRM and Analytics: CRM with Salesforce, 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, predict customer churns, Google Analytics
- Credit Research 1: CIBIL data, market risk, credit risk, liquidity risk, settlement risk, model risk, volatility risk, kurtosis risk and other types of financial risks. Risk measurement techniques, dynamic portfolio distribution analysis and extreme value analysis, Volatility, Monte Carlo. Financial Statement Analysis, Credit Specific Analysis – Banker's and Analyst Perspective, Understanding of Loans and Bonds
- Credit Research 2: Credit Rating; Credit Risk Management, 5C Credit Model, Analytical Framework for Credit Analysis, Ratings for Financial Institutions and Banks, Sovereign Ratings, Internal Rating models (Basel), and Case studies; Credit Risk Models and Regulations, Understanding and Analysis of Corporate, Banking Facilities and Other Financing Forms

* At the end of CR1 & 2, students will be able to appear for Certifications (like CCRA) and thereby, enhance their placement potential.
** Case studies will be provided through corporate associations with organisations like Salesforce, Crisil, SEBI, Bloomberg and Edelweiss.

Stream B: Marketing Analytics
- Business Case Analysis and Modelling
- Quant for Marketing: Summarize Marketing Data, PivotTables, Statistical Functions, Estimating Demand Curves and Using Solver to Optimize Price Estimating Linear and Power Demand Curves to Optimize, Pricing, Revenue Management, Estimating Demand, Regression Output, Modelling Interactions and Nonlinearities Testing, Neural Networks to Forecast Sales
- CRM and Analytics: CRM with Salesforce, 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, predict customer churns, Google Analytics
- Market Research 1: What do Customers Want? Conjoint Analysis Products, Attributes, and Levels, Price and Brand Equity into Discrete Choice Analysis, Independence of Irrelevant Alternatives (IIA), Customer Lifetime Value Model, Customer Acquisition and Retention Modelling, Market Segmentation, Cluster Analysis, Forecasting New Product Sales, Retailing: Market Basket Analysis and Lift Computing, RFM Analysis and Optimizing Direct Mail Campaigns
- Market Research 2: Model and Its Variants, Sales, Allocating Retail Space and Sales Resources, Allocation of Sales Advertising, Effectiveness, Media Selection Models, Google AdWords Auction, Bid Simulator. Marketing Research: Principal Components Analysis (PCA), Classification Algorithms, Discriminant Analysis, Internet and Social Marketing, Networks Measuring, Node Measuring, Complex Viral Marketing Model; Text Mining

*At the end of MR1 & 2, students will be able to appear for Industry Certifications and thereby, enhance their placement potential.
**Case studies will be provided through corporate associations with organisations like Salesforce, Nielson, InterSpace, Amazon, Future Group and BigBasket.