INNOVATICS
INNOVATION THROUGH ANALYTICS

CPD MEMBER
The CPD Certification Service

FROM ZERO TO HERO

ONLINE LIVE TRAINING PROGRAM

40 Weeks Program

BATCH 20

APRIL
20
2024

This Batch is Now Accredited

DATA SCIENCE & AI

www.innovatics.ai S.smaoui@innovatics.ai +971 58 971 9975

Overview

To maintain a competitive edge, today's organizations must gain practical knowledge of data science, machine learning and artificial intelligence. Careers especially in engineering, data science, and technical management depend on the immense value of those advanced analytical and programming skills.

In the Data Science & Machine Learning program from zero to hero program, an executive education program led by the InnovatiCS, you will dive deep into learning the applied mathematical structure of core machine learning (ML) and artificial intelligence (AI) methods. The course focuses on the skills needed to work efficiently with machine learning models, such classification models, regression models, clustering models, and optimization.

In this course, you will conduct your own analyses and apply the most effective methods to solve various problem types. This is what makes our program unique — we address both the how and the why of machine learning to provide you with the technical knowledge to thrive in any ML-essential field.

You will leave this program with.....

- The tools required to put analytics to practical use and solve specific business problems
- The language and intuition to work effectively with data scientists
- The necessary insights for leveraging analytics to accelerate growth and increase efficiency and productivity

Benefits of Developing a Data-Driven Organization

- Improve Marketing
- Create Operational Efficiencies
- Build New Business Models
- Disrupt the Industry Status Quo
- Master Insights to Leverage Data
- Boost Productivity

Key Benefits from the Program

Learn from basics of Data Science, R & Python Programming to Advanced ML

Build and Apply: Develop the ability to translate business problems to analytics & ML problems and develop impactful solutions.

Interact with Academic Experts and Industry Practitioners and get mentored - Train on detailed concepts and the nuances of solutioning

Reduce learning curve and be Industry Ready: The in-built Analytics and AI functions of the platform will reduce the learning curve and help to focus more on solutioning rather than just programming.

Dynamic Collaboration: The Collaboration features of the platform will enable to form teams, develop effective solutions and becomebetter team players.

Regular Hackathons: Compete & apply learnings under strict timelines and achieve tangible recognition/personal brand in the Data Science community.

Build and showcase cutting edge ML solutions using the INNOVATICS platform, with long term access to learning content/platform.

Industry standard assessment tests and long-term access to material will enable one's Data Science journey, even after the program completion.

Who is this program for ?

Engineers / Associates / IT Professionals: Software engineers in IT/ITES Startup teams building ML products/services

Data Analysts/Scientists & Business Analysts: Who want to transition to or progress into data science/ analytical roles and become more efficient and effective in data-driven decision-making.

Managers: Product managers, Program managers, General Managers, etc. interested in improving their analytical skills and effectively managing analytics, data science and machine learning projects

Consultants: Who are driving client projects and looking for acquiring and honing cutting edge analytical and data science skills for a career transition or progress

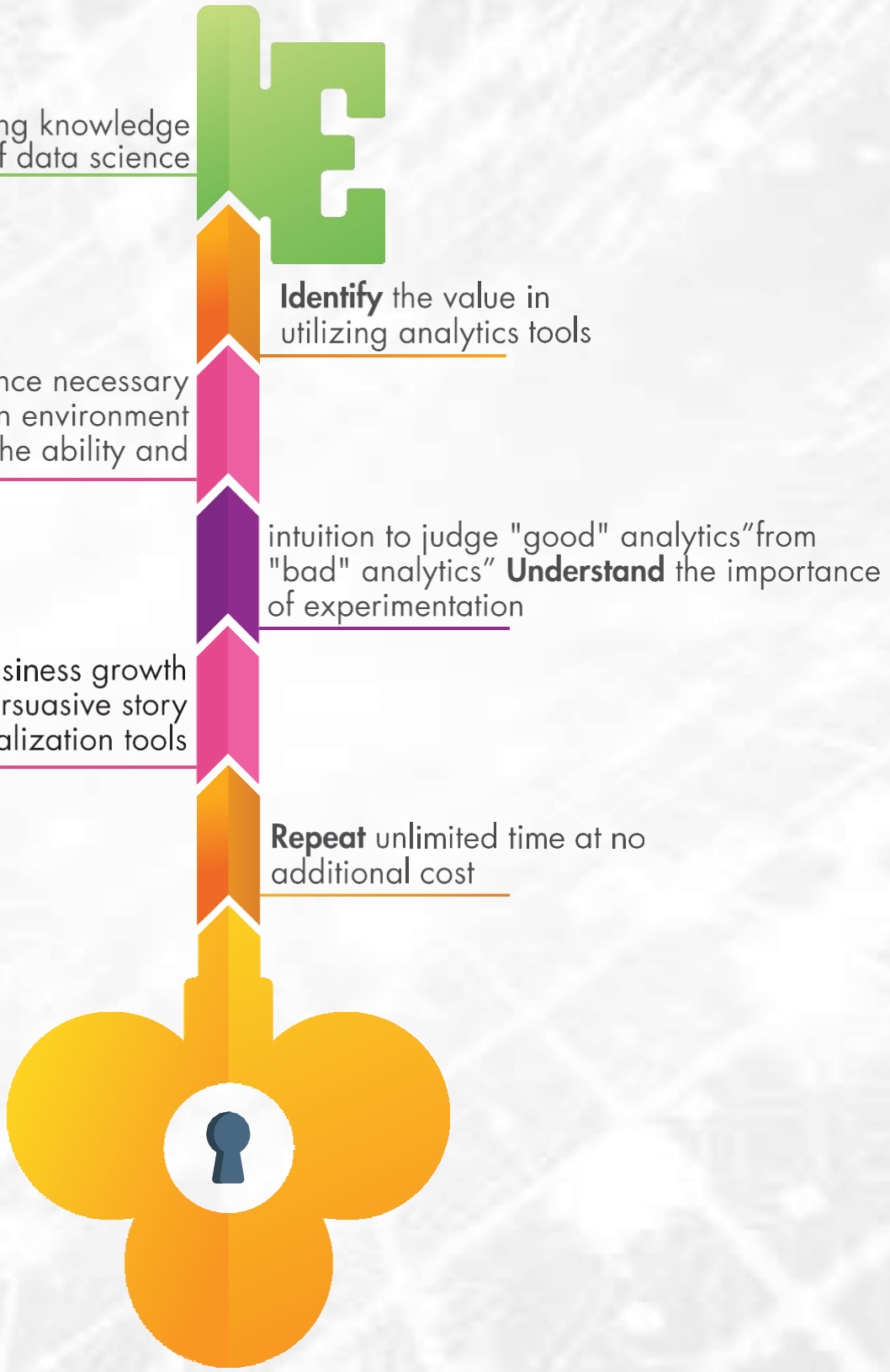
Requirements:

- 1-6 Years of work experience
- Bachelor's /Master's degrees.
- Commitment for Hands On/Lab effort of 8-12 hours per week throughout the program.

PREREQUISITES:

This course requires patience, passions, dedication and commitments and we will do our job

Key Takeaways



Program Experience

- Office Hours with Learning Facilitators
- Knowledge Checks
- Dedicated Program Support Team
- Insurance Data
- Python/R Coding Exercise in Each Module
- Bite-Sized Learning
- Peer Discussion
- Faculty Interaction

Regular office hours, coaching sessions, and live instructor sessions ensure that you 'get it'.

Graded assignments allow you to get the feedback you need to deepen your understanding of core concepts.

Having a portfolio will allow you to hit the ground running.

Program Features

Live : The program is delivered in an **Online** format with a highly effective blend of weekend live sessions followed 4 weekdays live sessions , then by self-paced practice sessions spanning across 8 months oriented learning.

Experiential Learning : More than 50 industry standard data cases & case studies across domains to be used for applied learning

Collaborative Learning : The INNOVATICS platform enables learners to collaborate on projects, helping them perform better in teams. Develop and showcase a portfolio of projects and solutions

Academic/Industry Leaders Interactions: 100 live sessions planned with Academic and Industry leaders for detailed and application oriented learning.

Milestone based Learning: Entire curriculum is divided into three manageable learning blocks to help the learner reap benefits in an incremental manner. Separate certificate for each block, followed by a certificate for the overall program

Capstone Project: The learner will get the opportunity to participate in live capstone project and develop cutting edge ML solutions under the guidance of experts and implement those, with the opportunity to develop and showcase a portfolio of projects and solutions

Sample Case Studies

Retail/Etail: Market Basket Analysis in Retail - Application of Association Rule | Mining Technique | Customer Sentiment analysis for an e-commerce retailer | Predict Holiday Sales for A Retail Client | Application of Linear Regression

Human Capital: Application of Non-Hierarchical Clustering in HR Analytics | Domain Recognizing human activity An application of supervised machine learning

Banking & Financial Services : Analyze Credit Card Spend Data | Application Of Descriptive Analysis Techniques | Identify risk class and eligibility of a customer | Application of Machine Learning | Credit Card Attrition

Manufacturing: Estimating Price for Diamonds | Supervised Learning | Hyper-parameter Optimization | Predictive maintenance of equipment data | Building predictive models using sensor data

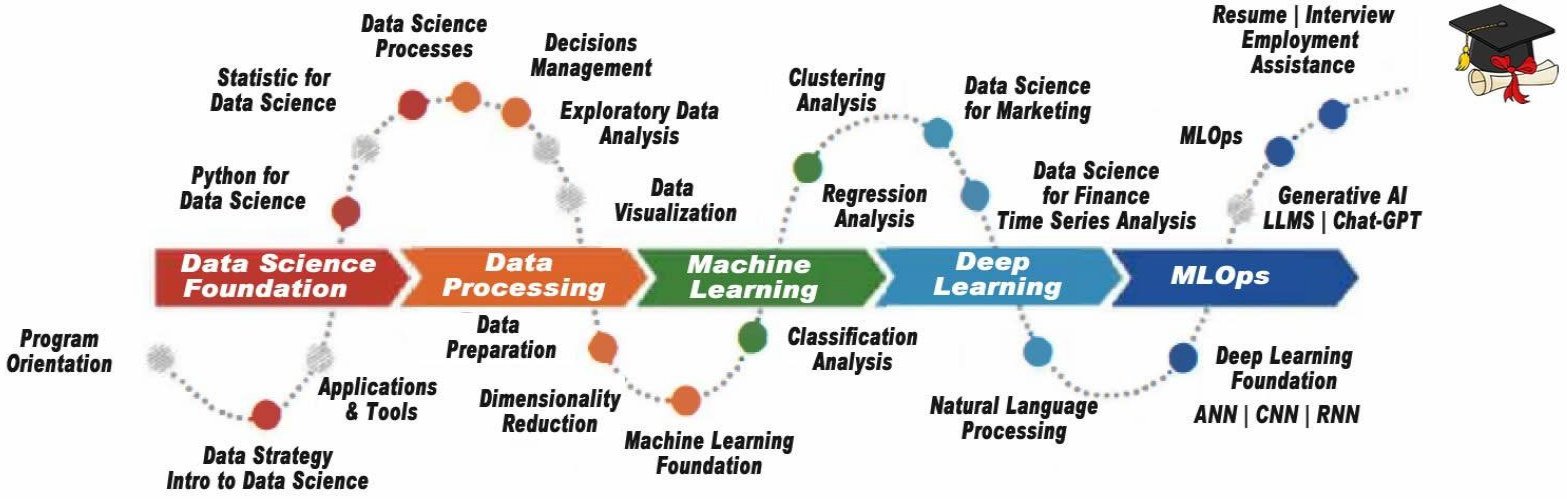
Healthcare: Detection of Breast Cancer in A Clinical Trial | Application Of SVM COVID-19 Data Exploration & Visualization

Sports: Identify the Top Performing Players in a Domestic Cricket League | Application of Descriptive Analysis Techniques | Fantasy Cricket Team Creation | Application of Linear Programming

Learning Journey

Over the course of eight months, you will develop a competency in data science fundamentals. By working directly with and receiving support from industry-expert mentors, you will be prepared for a shift into a career in data science.

Your Portfolio

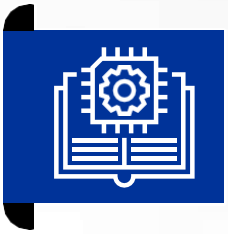


Program Highlights

This curriculum is sourced from 3 Master Degrees (MS-Analytics, MS-Data Science and MS-AI) and 10 years of industry experience. It is organized around the skills that technology giants such as Amazon, Google, Facebook, Apple, and McKinsey value in data science professionals. Numerous participants in the previous batches of the program have leveraged the skills acquired to obtain positions in data analytics oriented roles across a variety of industries.

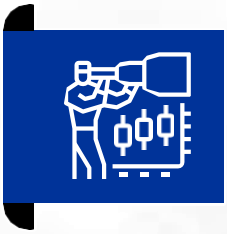
Target Job Functions Include

Data Science Foundations



R & Python Programming

Dedicated Program Support Team



Insurance Data

Tools in the Program

Schedule

Duration

40 Weeks Online (Live)
12 - 15 hours per week

Schedule: 3 Times a week

Sat 12:00pm to 05:00pm UAE Time
Sun 12:00pm to 05:00pm UAE Time
Thur 12:00pm to 03:00pm UAE Time

(Self-Paced)

Office Hourse: Every day per schedule

Why Now ?

Why INNOVATICS Data Science Solutions?

INNOVATICS**has a reputation for rigor, excellence, and cross-pollination across disciplines. What makes this program dierent?**

- ✓ One-to-one career coaching*
- ✓ Individual feedback
- ✓ Learning in real time
- ✓ Regular live webinars*
- ✓ Rigorous, graded assignments (Professional-level certificate)
- ✓ Assistance with career planning
- ✓ Real-world application of knowledge
- ✓ Small group mentoring sessions
- ✓ World-class faculty and thought leadership
- ✓ Learn from and network with your colleagues through peer discussions
- ✓ Ivy League education

*Services provided by Emeritus, a learning partner for this program.

Program Faculty

Dr. Mo Medwani
INNOVATICS CEO

Dr. Medwani, Founder of InnovatiCS (formerly, Big Data Science Solutions), has a PhD in Artificial Intelligence and an expert Data Scientist with a passion for transforming data into useful products. He has over 20 years of experience in Service Delivery Management; Four master's degrees in data science, IT, Machine Learning, and Business Administration; and over 9 years of experience working with Data Science. Mo's specialties include Data Science, Machine Learning, Big Data, Deep Learning, Data Analytics, Application Support and IT Service Delivery Management.

Prof. Ed Bujak
INNOVATICS Co-founder
Python Instructor

Prof. ED Bujak, Data science and Python expert withover 26 years of experience in IT and 20 years of education experience in Computer Science, Mathematics, and Engineering. He has two master's degrees in ElectricaEIngineering and Education, respectively. Ed is a public speak two master's everything related to programming, Eata Ecience, and Eachine Eearning.

Prof. Mohammed Mojeeb
INNOVATICS Co-founder
Data Engineering Instructor

Prof. Mohammad Mojeeb, PBS in Computer science and a Master of Business Adminis-tration (MBA)Microsoft Certified. Technology two master's Specialist in SQL Server Implementation and Maintenance Microsoft Certified. Azure cloud Fundamentals Certified. Computer Programming and Database Development

Over 24 years of experience in Software Development, Database Develop-ment, and Cyber Security, and a strong background in all varieties of Application Development and Application Security 10 years of training experience Currently working as Manager, Security Business Intelligence at LexisNexis.

Program Advisors

Dr. Kirk Borne
(Chief Science Officer)

Dr. Kirk Borne is the Chief Science Officer at DataPrime, Inc. He is a sought-after global speaker on topics ranginfrom data mining, data management, big data analytics, data science, machine learning, artificial intelligence, the internet of things, data-driven decision-making, modeling & simulation of dynamic systems to emerging technologies, the future of work, education, and science. Kirk provides leadership and mentoring to multi-disciplinary teams of scientists, modelers, and data scientists; and he consults with some external organizations, industries, agencies, and partners in the use of large data repositories and machine learning for discovery, decision support, and innovation. In addition to being a top big data influencer, he was a professor of astrophysics and computational science at George Mason University (GMU) for 12 years. He served as an undergraduate advisor for the GMU data science program and a graduate advisor in the computational science and informatics Ph.D. program.

Salwa Samaoui
(InnovatiCS MEA Director)

Salwa Smaoui is a digital transformation leader with over 25 years of experience in the world of technology, marketing and business development, with the majority of her career spent at some of the world's finest global companies. Since February 2016 Salwa has worked as the Microsoft Middle East & Africa Government Business leader. Before that she was the Business Group General Manager for Windows and Surface for Microsoft in Western Europe Region since 2014. Prior to that and in 2011, Mrs Salwa Smaoui was appointed the Director of Advertisement & Online for Middle East and Africa. In this role, Mrs. Smaoui was driving the strategy of Online and Advertisement Microsoft platform to increase its impact and usage by the consumers and businesses in the region. She helped build a strong ecosystem of IT partners, media and advertising partners in the region.

Program Modules

Week 1	Program Orientation

Learning Objectives: The program begins with a lecture on "Why Learning Data Science is an Absolute Must!" where you will be introduced to Data Science process, Data Science portfolio, analytics types, day to day activities,quantitative & statistical techniques required & the career path to start the journey.

Session 1	Program Orientation
	• Program Orientation (Agenda – Curriculum) • InnovatiCS Website | Canvas | Slack walk through
Session 2	**Why should you become a Data Scientist?**
	• Data Explosion • Why Data Science? | What is Data Science? | Type of Analytics • Data Science Portfolio | Data Science Process | Career in Data Science
Session 3	**Introduction to Data Science | CAP Certificate**
	• Introduction to Data Science • Introduction to Machine Learning • Introduction to Deep Learning overview • Certified Analytics Professional (CAP)
Session 4	**Projects, Teams & Team Leads**
	• Introduction to Data Science Project • Projects Discussion (Milestone Projects Assignment) • Projects List | Project Templates | Project Team
Session 5	**Program Tools & Installation 2**
	• **Jupyter Notebook** • **MySQL**: Installation • **Excel** • **Tableau**: Installation • **RapidMiner**: Installation

Week 2	Business Understanding (Strategy | Literacy | Problem Framing)

Learning Objectives: The basic workflow is now in place. You will dive into the first stage of the Cross -Industry Standard Process for Data Mining (CRISP-DM) process which deals with understanding analytical project objectives from a business perspective. Customers often have competing objectives and constraints that must be properly identified and balanced. The goal is to uncover important factors that could influence the outcome of the analytical projects. Neglecting thisstep can mean that a great deal of effort is put into producing the right answers to the wrong questions. In this section also we will provide an overview of the importance of data strategy as a tool that enables any organization to make better-informed decisions, improve business processes and gain new revenue streams. Data literacy involves articulating a problem that can potentially be solved using data.

Session 1	Introduction to Data Literacy and Data Strategy
	• Deciding Your Strategic Data Needs ○ Using Data to Improve Your Decisions | Using Data to Understand Your Customers and Markets | Using Data to Provide More Intelligent Services | Using Data to Make More Intelligent Products | Using Data to Improve Your Business Processes • Monetizing Your Data

	○ Defining Your Data Use Cases | Sourcing and Collecting the Data | Data Governance | Turning Data into Insights | Creating the Technology and Data Infrastructure | Building the Data Competencies in Your Organization | Executing and Revisiting Your Strategy • Introduction To Data Literacy ○ What Exactly is Data Literacy | Why do We Need Data Literacy | Data-driven Decision Making | Benefits of Data Literacy | How to Get Started
Session 2	**Introduction to Artificial Intelligence (ML & DL)**
	• Introduction to Machine Learning from different perspectives • Understand the Need of Machine Learning • Understand Machine Learning major elements. • Understand Machine Learning tasks & types of Machine Learning • Understand Types of Algorithms & types of Analysis • Understand Algorithms Performance • Deep Learning Explained • Biological Neural Network
Session 3	**Introduction to Big Data**
	• The job market around Big Data • What is Big Data? • Big Data use cases & Big Data Ecosystem • HDFS | What is MapReduce | Pig, Hive etc. • NoSQL Databases
Session 4	**Introduction to CRISP-DM**
	• Solutions Methodologies (Macro vs. Micro) ○ Scientific Research Method ○ Operations Research Method | Water Fall Method | CRISP-DM ○ Exploration & Discovery ○ Solutions are Dependent on Data | Solutions are Independent on Data • CRISP-DM Process • Business Understanding | Data Understanding | Data Preparation | Modeling |Evaluation | Deployment
Session 5	**Decisions Management & Problem Framing**
	• Quantifying Business Problem • Converting Business Problems into Analytics Solutions • Defining the objectives, assumptions, contains, analytical approaches. • Designing the Analytics Base Table & Implementing Features • Identifying Y variable & Measures of Success • Stakeholders & Analytical team
Week 3	**Data Science Foundations (Python Programming | GitHub)**
Learning Objectives: In this Week, students will be introduced Python programming language commonly used. Students will work with the tools that professional Data Scientists work with, like Jupyter Notebooks, IDE, and others. Students will learn about what each tool is used for, what codes they can execute, and their features and limitations	
Session 1	**Basic Python Programming 1**
	• Introduction to Python & Jupyter Notebook • Basic operators in Python • Data Structure: Sting | Lists

Session 2	Basic Python Programming 2
	• Data Structure ○ Tuples | Sets | Dictionaries
Session 3	**Intermediate Python Programming**

	• Python data types (basic and Boolean)
	• Conditional Statements.
	  ○ Python built-in data types
	  ○ Basic operators in Python
	  ○ Loop and control statements like break, if, for, continue, else, range () and more

Session 4	**Advanced Python Programming**
	• Functions, Assignment & Operations • Conditions & Loops
Session 5	**Introduction - Git and GitHub**
	• Introduction To Git • Installing Git | Configure Git • Basic commands • Exploring Git log | Git diff • Making corrections Git reset • Branching
Week 4	**Data Science Foundations (Statistics - Probability - SQL)**

Learning Objectives: In this Week, students will build a solid foundation of statistics for Data Science, learn probability, distributions, hypothesis testing and SQL. Statistics is the science of assigning a probability to an event based on experiments. It is the application of quantitative principles to the collection, analysis, and presentation of numerical data. Ace the fundamentals of Data Science, statistics, and Machine Learning with this course, it will enable students to define statistics and essential terms related to it, explain measures of central tendency and dispersion, and comprehend skewness, correlation, regression, distribution. Students will be able to make data-driven predictions through statistical inference. We will also learn about creating and working with tables, Databases Data Definition Language (DDL) and Data Manipulation Language (DML), Data Types SELECT, INSERT, UPDATE, AND DELETE Statements in SQL.

Session 1	Introduction to Statistics
	• Understanding Data • Levels of Measurement • Measures of Central Tendency • Population and Sample • Measures of Dispersion • Quartiles and IQR
Session 2	**Advanced Statistics / Probability & Information Theory**
	• Introduction to Distribution • Uniform Distribution | Binomial Distribution • Poisson Distribution | Normal Distribution • Skewness | Standardization and Z Score • Central Limit Theorem • Hypothesis Testing • ANOVA - Analysis of Variance • Chi-Square Analysis

Session 3	**Crash Course in SQL (Tables and Constraints)**
	• Creating, and Working with Tables • Default Constraints • Cascading referential integrity • Check Constraint | Identity column | Unique key constraint. • Get last generated id in SQL server.
Session 4	**SQL Statements Group by Clause**
	• SQL Statements (Select, Insert, Delete, Update) • Group by Clause • Difference between where and having in SQL server. • Basic Joins | Advance Joins | Self Joins • Different ways to replace NULL | Coalesce function. • Union and Union All
Session 5	**Team Projects 1st Presentation (Framing Business Problem)**
	• Team Projects (Business Understanding) • Team Leads' First Presentation
Week 5	**Data Science Foundations (Visual Data Sensemaking)**
Learning Objectives: Before you can present information to others, you must understand its story. In this Week students learn concepts, principles, and practices of visual data sense-making. The skills taught in this book rely primarily on something that most of us possess—vision—interactively using graphs to find and examine the meaningful patterns and relationships that reside in quantitative data.	
Session 1	**Building a framework for Visual Data Sensemaking**
	• History of Data Visualization • Prerequisites for Data Sensemaking • Thinking with our eyes • Visual Data Sensemaking Interaction and Navigation
Session 2	**Data fitness, Variation and Relationships (Categorical Variables)**
	• Variation withing Categorical Variables • Relationships among Categorical Variables • Variation across space | Variation through time
Session 3	**Data fitness, Variation and Relationships (Quantitative Variables)**
	• Variation within Quantitative Variables • Relationships among Quantitative Variables
Session 4	**Data fitness, Variation and Relationships (Multiple Variables)**
	• Relationships among multiple Variables and Perspectives • Multi-Perspective View • Multivariate Relationships
Session 5	**Interactive Dashboards (Power BI)**
	• Tableau Interactive dashboards • Power BI Interactive Dashboards

Week 6	Data Understanding | EDA & Statistical Analysis
Learning Objectives: In the Week, students will be introduced the second phase of the Cross- Industry Standard Process for Data Mining (CRISP-DM) process model. Students obtain data and verify that it is appropriate for the needs. Students learn to identify issues that cause the analyst to return to the 'Business Understanding' phase of theproject and revise the plan. Students may even discover flaws in the 'Business Understanding', another reason to rethink goals and plans. The Data Understanding phase includes four tasks. These are Gathering data – Describing data – Exploring data –Verifying data quality.	
Session 1	**Data Loading & Manipulation**
	• Read data from different sources. • Identifying Categorical Data: Nominal, Ordinal and Continuous • Univariate | Bivariate | Multivariate Analysis • Types of Data | Type of Central Tendency | Data distribution | Data Density
Session 2	**Data Types- Measure of Shape - Position - Dispersion**
	• Measures of the Spread: Range – IQR – Variance – Standard Deviation • Measures of Dispersion • Measures of Position: Statistical Analysis (John T. 5 Numbers Summary) • Measures of Relationships: Correlation | Variance & Covariance • Measures of Shapes: Skewness & Kurtosis
Session 3	**Data Visualization (Numerical Data & Graphical Descriptive Statistics)**
	• Data Visualization (Histograms, Bar Plot, Scatter Plot, Box Plot) • Data Quality Report | Develop the Code Book o Summary of Data | Type of variable | Ranges of variables | Missing fields | Identify the primary list of variables to solve the business problem • Detecting missing values & outliers | Duplicates & redundant records
Session 4	**NumPy Array | The Shape and Reshaping of NumPy Array**
	• NumPy ndarray Data Type, Features, and Array Dtype • Matrices, Dimensions, Axis, Shape, Reshape • Sequences (Index and Slice), Iterable
Session 5	**ML with RapidMiner | Tableau | SAP Analytics Cloud**
	• **RapidMiner** ○Installation | RapidMiner Basics | Studio - GUI Intro | Visualizing Data in RapidMiner Marketplace – Extensions • **SAC** ○Using in-story calculations | Add a measure input control | Conditional formatting | Using Conditional Formatting | Linked analysis | Data blending | Share a story | Leave comments | Scheduling publications | Use bookmarks | Smart Assist ○Smart Assist overview | Search to Insights | Indexing | Smart Discovery | Smart Insights | Using Smart Insights | Smart Discovery vs Smart Insights | Time Series Forecast | Integrating R visualization | Smart Grouping | Using Smart Grouping | Smart Predict | Datasets | Predict incident cost | • **Tableau:** ○ Introduction to Tableau | Your First Bar chart ○ Connecting Tableau to a Data File - CSV File ○ Navigating Tableau | Creating Calculated Fields ○ Adding Colors | Adding Labels and Formatting | Exporting Worksheets

Week 7	Data Preparation & Preprocessing
Learning Objectives: In the Data Preparation phase, students learn the process of cleaning and transforming raw data prior to processing and analysis. This is an important step prior to processing and often involves reformattingdata, making corrections to data and the combining of data sets to enrich data. Students also learn how to fix data quality issues discovered through EDA & Data visualization from the DataUnderstanding phase	
Session 1	**Data Preparation & Processing (Categorical Attributes)**

	• Encoding Categorical Data
	  ○ Replacing values | Creating Dummy Variables | Encoding labels | One-Hot encoding |Binary encoding | Backward difference encoding | Miscellaneous Features
	• Data Quality Report
	• Develop the Code Book
	  ○ Summary of Data | Type of variable | Ranges of variables | Missing fields | Identify the primary list of variables to solve the business problem.

Session 2	**Data Preparation & Processing (Numeric Attributes & Transformation)**
	• Statistical Analysis (John T. 5 Numbers Summary) • Variance & Covariance • Correlation Analysis
Session 3	**Data Preparation & Processing (Missing Values | Outliers | Duplicates)**

	• Dealing with Skewness & Kurtosis | Missing Values & Outliers | Duplicates & redundant
	• Variable Conversion
	  ○ Discretization | Binning
	• Variable Transformation
	  ○ Normalization | Standardization | Factorizations | Binarization
	• Box-Cox Transformation
	• Data Partitioning

Session 4	**Expanding, Squeezing, Indexing and Slicing of NumPy Array**
	• Filtering, Sorting, Statistics • Correlation, Regression, Outliers • Advanced Operations
Session 5	**Data Prep with RapidMiner | Excel**

	• **RapidMiner:**
	  ○ Turbo Prep – Introduction | Data Cleansing | Merging Data
	  ○ Data Pivoting | Data Preparation | Connecting to Databases
	• **Excel**
	  ○ Inserting a Line Break with Alt + Enter | Create Easily Printable Excel Documents
	  ○ Insert Hyperlinks into Excel Spreadsheets | Using Excel's Freeze Panes to Handle Large Datasets | Introduction to Excel's Pivot Tables
	  ○ Macros | Absolute and Relative Cell References
	  ○ Create Dynamic Names in Excel Spreadsheets
	  ○ Using Named Ranges to Make Formulas More Readable
	• **Tableau:**
	  ○ Time series, Aggregation, and Filters | Working with Data Extracts in Tableau
	  ○ Working with Time Series | Understanding Aggregation, Granularity, and Level of Detail | Creating an Area Chart & Learning About Highlighting | Adding a Filter and Quick Filter

Week 8	Factor Analysis & Dimensionality Reduction
Learning Objectives: Data preprocessing includes imputing missing values, dealing with outliers, duplicates, redundant data, skewness, kurtosis, distributions, correlation, feature selection, feature extraction, generic wrangling, data manipulation, scaling, factorization, binarization, transformation (Box-Cox), normalization & many more	
Session 1	**Principal Component Analysis**

	• Goals of Feature Selection
	• Classes of Feature Selection Methodologies
	  ○ Correlation Method
	  ○ Machine Learning Method | Feature Importance
	• Regularization & Embedded Feature Selection
	• Feature Subset Selection | Feature Creation
	• Factor Analysis: PCA | KPCA
	• Aggregation | Sampling | Dimensionality Reduction
	• Feature Subset Selection | Feature Creation
	• Discretization & Binarization | Variable Transformation | Data Partitioning

Session 2	**Singular Value Decomposition & Linear Discriminant Analysis**
	• Factor Analysis: LDA & SVD • Aggregation | Sampling | Dimensionality Reduction
Session 3	**ISOMAP | Locally Linear | Modified Locally Linear Embedding**
	• Introduction to Manifold Learning (t-SNE, LLE, Isomap) • Local Linear Embedding (LLE) ○ Intuition | Least squares problem • Eigenvalue problem
Session 4	**Data Prep with RapidMiner | Excel**

	• **RapidMiner:**
	  ○ Data Cleansing | Merging Data
	  ○ Data Pivoting | Data Preparation | Connecting to Databases
	• **Excel**
	  ○ Add a Drop-down List in Excel | Sort Multiple Columns Within a Table
	  ○ Saving Time in Excel and Doing Everything Faster by Using Excel Shortcuts
	  ○ Find and Replace – References | Find and Replace - Formatting
	  ○ Beauty Saving - The Professional Way of Saving Files
	  ○ The Power of F2 | Conditional Formatting
	  ○ Introduction to Custom Cell Formatting
	  ○ Custom Formatting - An example
	  ○ Beginner, Intermediate & Advanced Functions
	• **Tableau:**
	  ○ Maps, Scatterplots, and Your First Dashboard | Relationships vs Joins (v00update)
	  ○ Joining Data in Tableau | Creating a Map, Working with Hierarchies
	  ○ Creating a Scatter Plot, Applying Filters to Multiple Worksheets | Creating a Dashboard | Adding an Interactive Action - Filter | Adding an Interactive Action - Highlighting

Session 5	**Capstone Projects 2nd Presentation (Data Understanding & Data Viz)**
	• Team Projects (Data Understanding & Data Visualization) • Team Leads' Second Presentation

Week 9	Supervised Learning - Classification
Learning Objectives: This week broadens concepts learned in Data Understanding, Data Preparation & Exploratory Data Analysis (EDA) by extending to Machine Learning where you will learn about the models &methods used in machine learning & apply them to a real-world. The aim of supervised machine learning is to build a model that makes predictions based on evidence in the presenceof uncertainty. In this session, you will learn about the different algorithms of supervised learning such as Decisions Tree, Rule-Based classifier & Naive Bayes, Logistic Regression, Support Vector Machine, Nearest Neighbor Classifiers with deeper use of scikit-learn functionality, introducingautomated methods of feature selection, options for estimation including stochastic gradient descent, & advanced metrics for model evaluation.	
Session 1	**Decision Tree & Rule-Based Classification**
	• Algorithm for Decision Tree Induction • Methods for Expressing Attribute Test Conditions • Measures for Selecting an Attribute Test Condition • Characteristics of Decision Tree Classifiers • How a Rule-Based Classifier Works • Characteristics of Rule-Based Classifiers
Session 2	**Naïve Bayes & Logistic Regression**
	• Naïve Bayes motivation • Naïve Bayes Mathematical foundation • Basics of Probability Theory • Naïve Bayes Assumption • Logistic Regression motivation | Mathematical foundation • Logistic Regression as a Generalized Linear Model • Characteristics of Logistic Regression
Session 3	**Support Vector Machine SVM & KSVM**
	• SVM motivation • SVM Mathematical foundation • Margin of a Separating Hyperplane • Linear SVM | Nonlinear SVM • Soft-margin SVM • Characteristics of SVM
Session 4	**Neighbor Classifiers (KNN)**
	• KNN motivation • KNN Mathematical foundation • KNN Algorithm • Characteristics of Nearest Neighbor Classifiers
Session 5	**Pandas Fundamentals**
	• Pandas Data Frame Basic • Pandas' data types: Index, Series, DataFrame | Axis • Operations and Methods on Panda's data types • Create Data Frame | Addition and Deletion | Contact and Merge • Export to CSV | Excel | Txt | Group By | Sorting | Stack & Unstack | Iterate Over Data Frame |Reshaping using Pivot | Selection & Indexing

Week 10	Supervised Learning - Regression
Learning Objectives: This week provides Regression Analysis (Supervised Learning with a continuous target variable). Regression analysis is a form of predictive modeling technique that investigates the relationship between a dependent (target) and independent variable (s) (predictor). This technique is used for forecasting, timeseries modeling & finding the causal effect relationship between the variables. Machine learning topics taught involve linear regression algorithms (Simple Linear Regression, Multiple LinearRegression & Polynomial Linear Regression, Decision Tree, and Support Vector Machine) and Regularization Techniques (Lasso, Ridge, Elastic Net, Least Squared) with deeper use of scikit-learn functionality.	
Session 1	**Simple Linear Regression & Polynomial Regression**
	• Preparing Data for Linear Regression • Making Predictions with Linear Regression • Polynomial Algorithm • Why use polynomial regression? • Polynomial Features
Session 2	**Multiple Linear Regression**
	• Multiple Linear Regression • The F-Statistic • Interpreting results of Categorical variables • Heteroscedasticity • Backward Elimination | Backward Elimination | Automatic Backward Elimination
Session 3	**Least Square/Lasso/Ridge/Elastic Net Regression**
	• OLS Regression- Theory | Implementation • Confidence Interval and OLS Regressions • Ridge Regression | LASSO Regression • Implement ANOVA on OLS Regression • Identify Multicollinearity | Partial Least Square Regression
Session 4	**Decision Tree Regression & Support Vector Machine SVR**
	• A Basic Algorithm to Build a Decision Tree • Methods for Expressing Attribute Test Conditions | Measures for Selecting an Attribute Test Condition • Algorithm for Decision Tree Induction | Characteristics of Decision Tree Regression • Margin of a Separating Hyperplane | Linear SVM | Soft-margin SVM | Nonlinear SVM • Characteristics of SVM
Session 5	**Using Pandas for Data Analysis & Data Wrangling 1**
	• Data types: Categorical, Continuous • EDA, Data Cleaning • Filtering, Row/Record/Column
Week 11	**Ensemble Methods (Bagging & Boosting & Stacking)**
Learning Objectives: Ensemble methods help to improve the predictive performance of Machine Learning models.In this Week, you will learn about different Ensemble methods that combine several Machine Learning techniques into one predictive model to decrease variance, bias or improve predictions. You will learn about selecting one model over another and boosting & its importance in Machine Learning. You willlearn how to convert weaker algorithms into stronger ones.	

Session 1	**Random Forest Voting & Averaging**
	• Random Forests • Empirical Comparison among Ensemble Methods
Session 2	**Bagging | Boosting | Gradient Boosting | Ada-Boost | Stacking**
	• Methods for Constructing an Ensemble Classifier • What is the Model Selection? | The need for Model Selection • Cross-Validation • What is Boosting? | Adaptive Boosting • How Boosting Algorithms work? • Types of Boosting Algorithms
Session 3	**Class Imbalanced Problem**
	• Building Classifiers - Class Imbalance • Data-based Approaches & Algorithmic approach | Evaluating Performance - Class Imbalance • Finding an Optimal Score Threshold | Aggregate Evaluation of Performance
Session 4	**Using Pandas for Data Analysis & Data Wrangling 1**
	• Column/Feature Manipulations • Statistical Foundations • Problematic Data - Outliers, Missing Data (Impute)
Session 5	**ML with RapidMiner | Excel**

	• **RapidMiner:**
	  ○ Applying the Model | Testing a Model | Validating a Model | Finding the Right Model | Optimization of the Model Parameters
	• **Excel**
	  ○ Ordering the Source Worksheet
	  ○ Using Lookup Functions (Vlookup) to Fill the Database Sheet
	  ○ Use Sumif to Complete the Database Sheet
	  ○ Using Index & Match as a Substitute for Vlookup
	  ○ Learn How to Find Mistakes with Countif
	  ○ Calculating Growth Rates in Excel
	• **Tableau:**
	  ○ Joining, Blending and Relationships | Dual Axis Charts | Joins with Duplicate Values | Joining on Multiple Fields | The Showdown: Joining Data vs. Blending Data in Tableau
	  ○ Creating Calculated Fields in a Blend (Advanced Topic) | Working with Relationships in Tableau

Week 12	**Model Evaluation & Optimization**

Learning Objectives: Model building is an iterative process. Employing Feature Engineering techniques along witha careful model selection exercise helps to improve the model. Further, tuning the model is an important step to arrive at the best possible result.

This Week talks about the steps and processes around the same, you will learn how to analyze the performance of each algorithm, and dive deep in core Machine Learning concepts like Cost Function, Object Function, Model Optimization, Model Tuning, Regularization, Gradient Boosting, Grid & Random Search.

Session 1	Model generalization: Assessing Predictive Accuracy for New Data
	• CM, ROC, Rank-Ordered Approach • R2, MSE, MAE, Median Error, Median Absolute error, Correlation • Reasons for Model Overfitting • Model Selection | Using a Validation Set • Incorporating Model Complexity | Estimating Statistical Bounds • Model Selection for Decision Trees
Session 2	**Evaluation of Classification Models & Regression Models**
	• Holdout Method & Cross-Validation • Presence of Hyper-parameters & Hyper-parameter Selection • Nested Cross-Validation • Pitfalls of Model Selection & Evaluation • Overlap between Training & Test Sets • Use of Validation Error as Generalization Error | Cluster Evaluation
Session 3	**Model Optimization through Parameter Tuning | XGBoost**
	• Estimating the Confidence Interval for Accuracy • Comparing the Performance of Two Models • XG-Boost
Session 4	**Data Wrangling & Visualizing with Pandas**
	• Data transformations, Data type, Scaling, Normalizing, Binning, Group by • Slicing, changing index, Data conversion, Joining and Merging, Concatenation, Columns, Pivoting, Melting, Handling Duplicates Distributions, Histograms, Box Plot, Line Plots, Bar Plots, Grouped Bar Plot, Scatter Plot
Session 5	**Team Projects 3rd Presentation (Data Preparation)**
	• Team Projects (Analytical Approach) • Team Leads' Third Presentation

Week 13 — Unsupervised Learning - Clustering Analysis

Learning Objectives: In this week, we will shift to Unsupervised Learning techniques (target variable is unknown). Unsupervised Learning finds hidden patterns or intrinsic structures in data. The end goal is less clear-cut than predicting an output based on a corresponding input. In this Week, you will learnabout commonly used clustering techniques like K-Means Clustering and Hierarchical Clustering along with anomalies detection algorithms with deeper use scikit-learn functionality.

Session 1	K-Mean Clustering
	• The Basic K-means Algorithm • Bisecting K-means • K-means & Different Types of Clusters • K-means as an Optimization Problem
Session 2	**Agglomerative Hierarchical Clustering**
	• Basic Agglomerative Hierarchical Clustering Algorithm • AHC Specific Techniques • The Lance-Williams Formula for Cluster Proximity • Key Issues in Hierarchical Clustering • AHC Outliers

Session 3	**DBSCAN | Mean Shift | Gaussian Mixed Models | Fuzzy C Means**
	• BIRCH | DBSCAN | Mini-Batch K-Means • Mean Shift | OPTICS |Spectral Clustering |Gaussian Mixture Model
Session 4	**Python Regular Expression**
	• Extracting data from the web using JSON, Google API, and XML • Data Extraction - Getting Data from the Internet
Session 5	**ML with RapidMiner | Excel**

	• **RapidMiner:**
	  ○ Logistic Regression | Times Series | Clusters | Anomaly detection | AssociationDiscoveries
	• **Excel**
	  ○ Introduction to Excel Charts
	  ○ Modifying Excel Charts - The Easy Way
	  ○ Creating a Bridge Chart in Excel
	  ○ Stacked Column Chart with a Secondary Axis
	  ○ Doughnut Chart | Area Chart | Bridge Chart
	• **Tableau:**
	  ○ Table Calculations, Advanced Dashboards | Storytelling
	  ○ Mapping: How to Set Geographical Roles | Creating Table Calculations for Gender
	  ○ Creating Bins and Distributions for Age | Leveraging the Power of Parameters
	  ○ How to Create a Tree Map Chart | Creating a Customer Segmentation Dashboard
	  ○ Advanced Dashboard Interactivity |Analyzing the Customer Segmentation Dashboard

Week 14	**Anomalies Detection & A/B Testing**
Learning Objectives: This Week is a continuation of the Unsupervised Machine Learning algorithms from the previous week. We go deep into Association Rule mining which has numerous applications that are widely used to help discover sales correlations in transactional data or in medical data sets. We also go deep to A | B Testing, also known as split tests, to cover the ins & outs of how to use Python & R toanalyze customer behavior & business trends as well as how to create, run, and analyze A | B tests to make proactive, data-driven business decisions.	
Session 1	**Classification Based Algorithms - KNN**
	• Characteristics of Anomaly Detection Problems | Anomaly Detection Methods • Spectrum of Anomaly detection techniques • Nearest Neighbor based Methods | Rule-based Anomaly Detection • Classification based Models for Anomaly Detection • Proximity-based Approaches
Session 2	**Nearest-Neighbor Based algorithms**

	• Nearest-Neighbor based algorithms.
	  ○ k-NN Global Anomaly Score | Local Outlier Factor (LOF)
	  ○ Connectivity based Outlier Factor (COF)
	  ○ Local Outlier Probability (LoOP)
	  ○ Influenced Outlierness (INFLO) |Local Correlation Integral (LOCI)
	• Clustering based algorithms.
	  ○ Cluster based Local Outlier Factor (CBLOF)
	  ○ Local Density Cluster based Outlier Factor (LDCOF)

Session 3	**Clustering Based algorithms**

	• Clustering based algorithms.
	  ○ Cluster based Local Outlier Factor (CBLOF)
	  ○ Local Density Cluster based Outlier Factor (LDCOF)

Session 4	**A/B Testing**
	• Overview of A | B Testing? • How A | B Testing Works • Policy and Ethics • Characterizing Metrics • Designing an Experiment • A | B Testing Process | Analyzing Results | Alternatives to A | B Testing
Session 5	**Python: BeautifulSoup Library**
	• Installing Beautiful Soup • Data extraction with BeautifulSoup • BeautifulSoup Usage | Filtering | BeautifulSoup Object • Extracting all the URLs found within a page 'a' tag | Extracting text from a page.
Week 15	**Association Rules | Recommender Systems**
Learning Objectives: This Week is a continuation of the Unsupervised Machine Learning algorithms from the previous week. We go deep into Association Rule mining which has numerous applications that are widely used to help discover sales correlations in transactional data or in medical data sets. We will also dive deep into recommender systems, which are software that select products to recommend to individual customers. You will learn how to produce successful recommender systems that use past product purchase and satisfaction data to make high-quality personalized recommendations.	
Session 1	**Association Rules - Apriori**
	• Strength of an association rule • The Apriori Principle • Frequent Itemset Generation in the Apriori Algorithm • Candidate Generation & Pruning • Computational Complexity
Session 2	**Association Rules - Eclat**
	• Horizontal vs Vertical Data Format • The Intuition of ECLAT Algorithm • Advantages of Eclat & Eclat vs Apriori
Session 3	**Recommender Systems: First Generation**
	• Matrix Factorization Model (Object recommendation) • Content Filtering | Collaborating Filtering (CF)
Session 4	**Recommender Systems: Second Generation**
	• CF Neighborhood-Based Approach • CF Location-Based Approach • Popularity Based | Content Based
Session 5	**Recommender Systems: Third Generation**
	• Collaborative filtering • Matrix Factorization (MF)

Week 16	Time Series Analysis

Learning Objectives: In this Week, you will learn about Time Series Analysis to forecast dependent variables based on time. Time series analysis is a statistical technique that deals with time-series data, or trend analysis. Timeseries data means that data is in a series of periods or intervals. You will learn different models for time series modeling such that you analyze a real time-dependent data forforecasting.

Session 1	AR - MA & EST Models
	• What is Time Series Analysis? • Importance of TSA | Components of TSA • Single Exponential Smoothing • Forecasting with Single Exponential Smoothing • Double Exponential Smoothing | Forecasting with Double Exponential Smoothing • Triple Exponential Smoothing
Session 2	**ARMA & ARIMA Models**
	• Forecasting • Relation between time series: Causality & time lags • Distinction between short & long run • Study of agent's expectations • Stationarity | ACF & PACF • Trend removal | Seasonal adjustment
Session 3	**SARIMAX Model**
	• The limitations of ARIMA • The SARIMA extension of ARIMA • Implementing SARIMA method using the Statmodels library
Session 4	**Python: Matplotlib Library**
	• Annotation | Autocorrelation | Curves Scales | Shapes | Box Plot | Violet Plot | Histograms | Pie Chart | Twin Axis | Figures | Color Map Color Palettes/Figure Aesthetics
Session 5	**Team Projects 4th Presentation (Analytical Approach Implementation)**
	• Team Projects (Analytical Approach Implementation) • Team Leads' Fourth Presentation

Week 17	Text Analysis & Natural Language Processing (NLP)

Learning Objectives: In this Week, we dive into Natural Language Processing or NLP which is one such technology penetrating deeply and widely in the market, irrespective of the industry and domains. It is extensively applied in businesses today and it is the buzzword in every engineer's life.

Session 1	Introduction to Natural Language Processing
	• What is Natural Language Processing? • History of Natural Language Processing • NLP Applications | Levels | Components • NLU vs NLG • NLP Pipeline and Tasks | NLP Toolkits and Libraries • NLP Challenges

Session 2	Essentials of Natural Language Processing
	• Basic Text Analysis • Tokenization | POS Tagging • Stop Word Removal • Text Normalization • Spelling Correction • Stemming | Lemmatization • Named Entity Recognition (NER) • Word Sense Disambiguation • Sentence Boundary Detection
Session 3	**NLP Feature Extractions**
	• Data Structures • NLP Pre Processing • The Bag of Words • Frequency Vector (Count of Vectorization) • One-Hot Encoding • Term Frequency – Inverse Term Frequency • Distributed Representation • Word Embedding
Session 4	**NLP with TextBlob & spaCy**
	• Introduction to TextBlob • Language Detection • POS | Word Inflection • Sentiment Analysis • Introduction to spaCy Library | Objects of spaCy Library • The Statistical Modeling | Processing Pipelines
Session 5	**Python: SciPy Library**
	• Introduction to SciPy, building on top of NumPy. • What are the characteristics of SciPy? • Various sub packages for SciPy like Signal, Integrate, Fatback, Cluster, Optimize, Stats and more, Bayes Theorem with SciPy.

Week 18 — Advanced Natural Language Processing (NLP)

Learning Objectives: In this Week, we dive into advanced Natural Language Processing which is one such technology penetrating deeply and widely in the market, irrespective of the industry and domains. It is extensively applied in businesses today and it is the buzzword in every engineer's life.

Session 1	Text Classification
	• Data preparation | Data Vectorization • Text Classification
Session 2	**Text Summarization**
	• What is Text Summarization? • Text Summarization Categories • Stages of Text Summarization

Session 3	Topic Modeling
	• What is Topic Modeling? • Topic Modeling Use cases • Topic Modeling Libraries • Latent Semantic Analysis (LSA) • Latent Dirichlet Allocation (LDA) • Hierarchical Dirichlet Process (HDP)
Session 4	**Sentiment Analysis**
	• What is Sentiment Analysis? • Types of Sentiment Analysis • Benefits of Sentiment Analysis • Examples of Sentiment Analysis • Challenges of Sentiment Analysis
Session 5	**Large Language Models (LLMs) - ChatGPT, GPT4, BERT**
	• What are LLMs? Why LLMs (such as GPT and Bard) is so special? • The Power of Attention Mechanisms | Encoder-Decoder Architectures • Demystifying Transformers | Learn and Understand Transformers • Scaled Dot Product Attention | Multi-Headed Attention • Transfer Learning Potential • Exploring the GPT Architecture: Foundations | Innovations • Practical Applications of Masked Multi-Head Attention • GPT Pre-Training Strategies and Techniques • Optimizing GPT Pre-Training for Performance
Week 19	**Deep Learning: ANN | CNN | Object Detection**
Learning Objectives: This week, we dive into Deep Learning - an Artificial Intelligence function that imitatesthe workings of the human brain in processing data and creating patterns for use in decision making. Deep Learning carries out the Machine Learning process using an 'Artificial Neural Net', which is composed of a number of levels arranged in a hierarchy. In this session, you will learn about the basic building blocks of ArtificialNeural Networks. You'll learn how Deep Learning Networks can be successfully applied to data for knowledge discovery, knowledge application, and knowledge-based prediction.	
Session 1	**Introduction to Deep Learning**
	• Deep Learning Explained • Biological Neural Network • Perceptron – Gradient Decent – Scholastic Gradient Descent
Session 2	**Artificial Neural Network (ANN)**
	• ANN Explained • Multi-layer Neural Network • Characteristics of ANN
Session 3	**Convolutional Neural Network (CNN)**
	• CNN Explained • Synergistic Loss Functions | Responsive Activation Functions • Regularization | Initialization of Model Parameters • Characteristics of Deep Learning

Session 4	Real Time Object Detection using YOLO V4
	• Introduction to YOLO • YOLO as a real-time object detector. • YOLO as an object detector in TensorFlow & Keras • How to train your custom YOLO object detection model • YOLO compared to other detectors.
Session 5	**Python: Torch Library 1**

	• How to Install PyTorch
	• PyTorch Deep Learning Model Lifecycle
	  ○ Prepare the Data
	  ○ Define the Model
	  ○ Train the Model
	  ○ Evaluate theModel
	  ○ Predictions.

Week 20	**Self-Organizing Maps | Boltzmann Machines | Autoencoders**
Learning Objectives: This week is a continuation of week 19. You will learn 3 more Deep Learning Models (SOMs, BM and AutoEncoders). In SOMs, you will learn How to return the specific features (like frauds) detected by the SOM and How to make a Hybrid Deep Learning Model. You will also learn the Energy-Based Model point of view, and then for the Practical Lectures we will focus more on the Probabilistic Graphical Model point of view. Finally, you will learn how to build an AutoEncoder from scratch with PyTorch and how to manipulate classes and objects to improve and tune your AutoEncoder.	
Session 1	**Introduction to SOMs, BMN & AutoEncoders**
	• Introduction to Self-Organizing Maps • Introduction to Boltzmann Machine • Introduction to Auto Encoder
Session 2	**Self-Organizing Maps**
	• How Self-Organizing Maps work? • How Self-Organizing Maps Learn? • Live SOM example • Energy-based Models (EBMs) • Restricted Boltzmann Machines (RBM) • Deep Boltzmann Machines (DBM)
Session 3	**Boltzmann Machines**
	• Energy-Based Models (EBM) • Deep Belief Networks • Deep Boltzmann Machine
Session 4	**AutoEncoders**
	• Training an Auto Encoder • Overcomplete hidden layers • Sparse Autoencoders | Denoising Autoencoders | Contractive Autoencoders • Stacked Autoencoders | Deep Autoencoders
Session 5	**Team Projects Final Presentation (Insights & Action Plan)**
	• Team Projects (Analytical Approach Implementation) • Team Leads' Final Presentation

Week 21	Sequence Learning & GANs
Learning Objectives: This Week talks about yet another interesting implementation of Neural Networks that revolves around equipping computers to understand human language. You will learn to work with text data and sequential data, and explore the interesting world of RNNs and LSTMs	
Session 1	**Recurrent Neural Networks (RNN)**
	• RNN Explained | Characteristics of RNN • Backpropagation • Gradient problem (Vanishing vs. Exploding) • Long Short-Term Memory Networks
Session 2	**RNN vs LSTM with Google Stock Price**
	• Transform Data for Time Series • MLP & CNN for Time Series Forecasting • LSTM & CNN-LSTM for Time Series Forecasting • Encoder-Decoder LSTM Multi-step Forecasting
Session 3	**Sentiment Analysis with LSTM**
	• Motivations • From logistic regression to neural networks • Word representations • Unsupervised word vector learning • Backpropagation Training • Learning word-level classifiers: POS & NER
Session 4	**GANS: Generative Adversarial Networks**
	• Challenge of GAN Loss • Standard GAN Loss Functions • Alternate GAN Loss Functions • Effect of Different GAN Loss Functions
Session 5	**Python: Torch Library 2**

	• How to Develop PyTorch Deep Learning Models
	  ○ How to Develop an MLP for Binary Classification | Multiclass Classification
	  ○ How to Develop an MLP for Regression
	How to Develop a CNN for Image Classification

Week 22	**General Deployment Considerations & MLOps**
Learning Objectives: In this Week, you will learn every aspect of how to put your models in production. You will learn all the steps and infrastructure required to deploy machine learning models professionally, and you will have at your fingertips, the sequence of steps that you need to follow to deploy a machine learning model, plus a project template with full code, that you can adapt to deploy your own models.	
Session 1	**Avoiding False Discoveries | MLOps Introduction**
	• Preliminaries: Statistical Testing • Modeling Null & Alternative Distributions • What is MLOps? • Machine learning industrialization challenges • MLOps Motivation: High-level view • MLOps challenges • MLOps Components Machine Learning Life Cycle

	• How does it relate to DevOps, AIOps, ModelOps, and GitOps? • Major Phases - what it takes to master MLOps. • CI/CD in Production Case Study
Session 2	**Introduction to ML and MLOps stages**

	• MLOps Maturity Model Detailed
	• MLOps and stages
	  ○ Versioning
	  ○ Testing Automation (CI/CD)
	  ○ Reproducibility
	  ○ Deployment Monitoring

Session 3	**MLOps Architectures**

	• MLOps Architectures
	  ○ Architectures - Open Source tools - Kubeflow, Apache Airflow, MLFlow,Metaflow, Kedro, ZenML, MLRun, CML
	  ○ Architectures - Cloud Native tools - AWS, GCP and Azure
	• List of tools involved in each stage (MLOps tool ecosystem)
	• Different Roles involved in MLOps (ML Engineering + Operations)

Session 4	**Introduction to CI/CD**
	• Introduction to CI and CD • CI/CD challenges in Machine Learning • Steps involved in the CI/CD implementation in ML lifecycle and workflow. • A glimpse of popular Tools used in the DevOps ecosystem on the Cloud.
Session 5	**Python: TensorFlow & Keras Library 1**
	• Overview of TensorFlow and TensorFlow libraries • Use cases for a machine learning service. • Using and applying your model • Training your model | Testing your model • Using TensorBoard to visualize model performance
Week 23/26	**Individual Capstone Project (1 Month)**

Learning Objectives: During the final week, you will be transitioned into full-time focus on your final, passionindividual capstone assignment that will require you to apply the tools you have learned in the program.

Duration: The capstone will run for ONE month starting from the 6th of January 2024 till the6th of February 2024.

Grading: The capstone project is a program requirement. To successfully complete the capstone project and receive the certificate, students must achieve “Meets Expectations” OR“Exceeds Expectations”.

Week 27 |28	Data Science & AI Interview Preparation

Learning Objectives: In this module, you will be introduced to 120+ Data Science Interview questions and answers to help you master the Data Science language and express yourself fluently. Topics covered:

- The Big Picture
- Optimization
- Data Pre-processing
- Sampling & Splitting
- Supervised Learning
- Unsupervised Learning
- Model Evaluation
- Ensemble Learning
- Business Applications

Week 29|30	Data Science & AI 24 Hours Real Challenges

Learning Objectives: In this module, you will be introduced to some 24-hour real take-home challenges used as hiring. process with some big companies:

- Capgemini Challenge
- SDSC Challenge
- Foot Locker Challenge
- Fatality Crashes Challenge
- PayPal Challenge

Week 31	Data Science Soft Skill Preparation

Learning Objectives: In this module, you will assess your Data Science soft skills in core topics of statistics, fundamentals of data science and programming, Machine Learning, and Neural Networks. The test is designed to help you identify your strengths and weaknesses.

Week 32|40	Introduction to ML and MLOps stages

Learning Objectives: InnovatiCS has developed a new course "CAP Preparation Course" as a bonus & great opportunity for all students at all experience levels to prepare for the Certified Analytics Professional (CAP®) certification or to refresh their knowledge and skills. In this module, you will explore all seven domains of the analytics practice as:

CERTIFIED ANALYTICS PROFESSIONAL

- Domain I: Business Problem Framing
- Domain II: Analytics Problem Framing
- Domain III: Working with Data
- Domain IV: Methodology Selection
- Domain V: Model Building
- Domain VI: Model Deployment
- Domain VII: Life-Cycle Management
- Mock exam

Certificate of Completion

Upon successful completion of the program, Big Bang Data Science Institute grants a **verified/certified digital** certificate of completion to participants. This program is graded as pass or fail; participants must receive 80% to pass and obtain the certificate of completion.

This course has been certified by the CPD Certification Service as conforming to continuing professional development principles. After successful completion of the program, your verified digital certificate will be email to you in the name you used when registering for the program. All certificate images are for illustrative purpose only and may be subject to change at the direction of InnovatiCS.

The CPD Certification Service: The CPD Certification Service was established in 1996 and is the leading independent CPD accreditation institution operating across industry sectors to complement the CPD policies of professional and academic bodies.

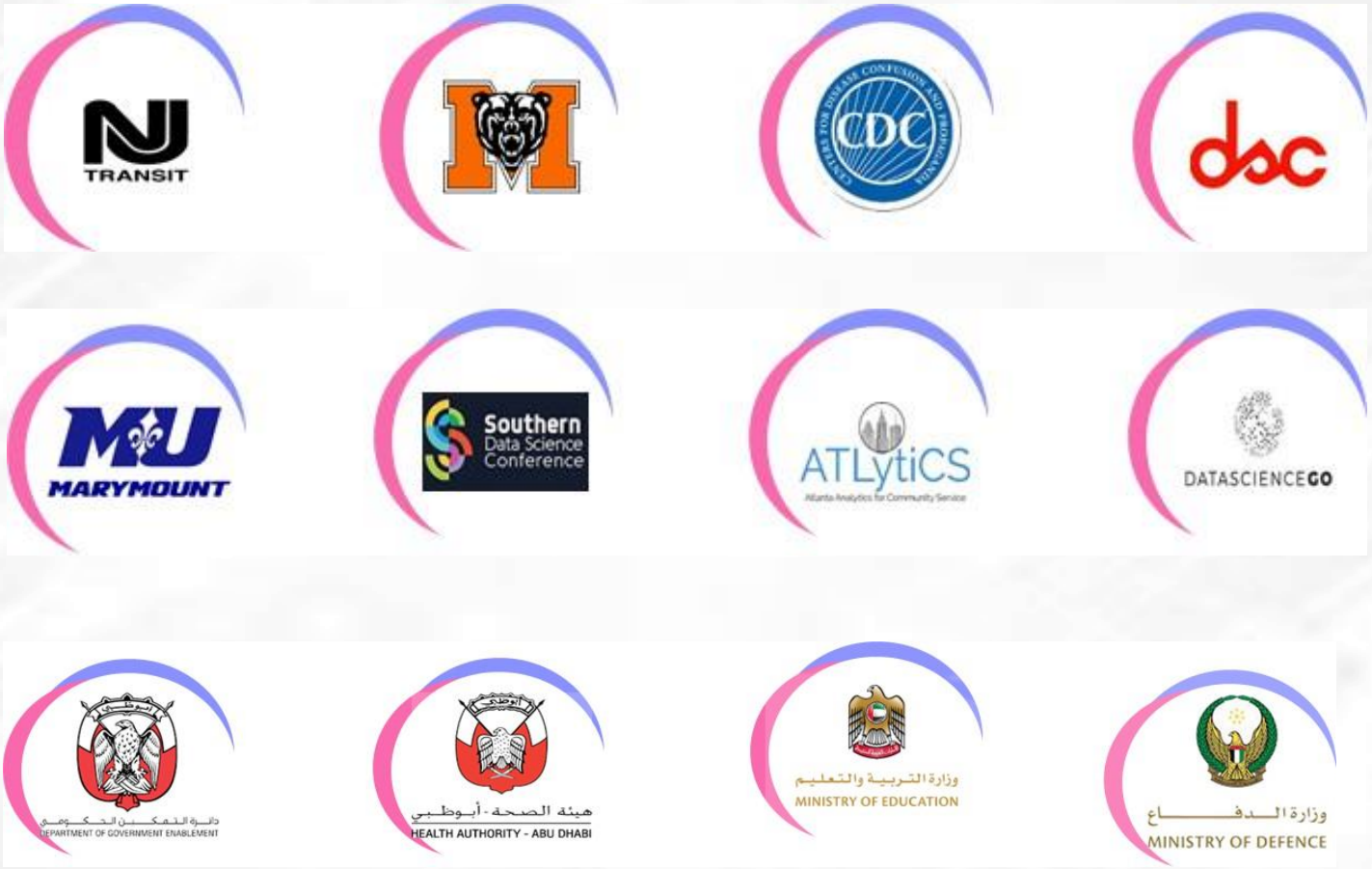
Testimonials

Here is another reason why you should trust **INNOVATICS** with your training. Watch this recording capturing **INNOVATICS** prestigiousAward from Internet 2.0 Conference - Dubai 2022, Internet 2.0 Conference **Dubai 2023** , Internet 2.0 Conference **Las Vegas** 2023.

Our Clients !

Our diverse clients and projects have continued to provide our team with the opportunity to grow a unique skillset. The following are few of our recently completed programs for corporates or universities.

OUR MOST RECENT CLIENTS INCLUDE:



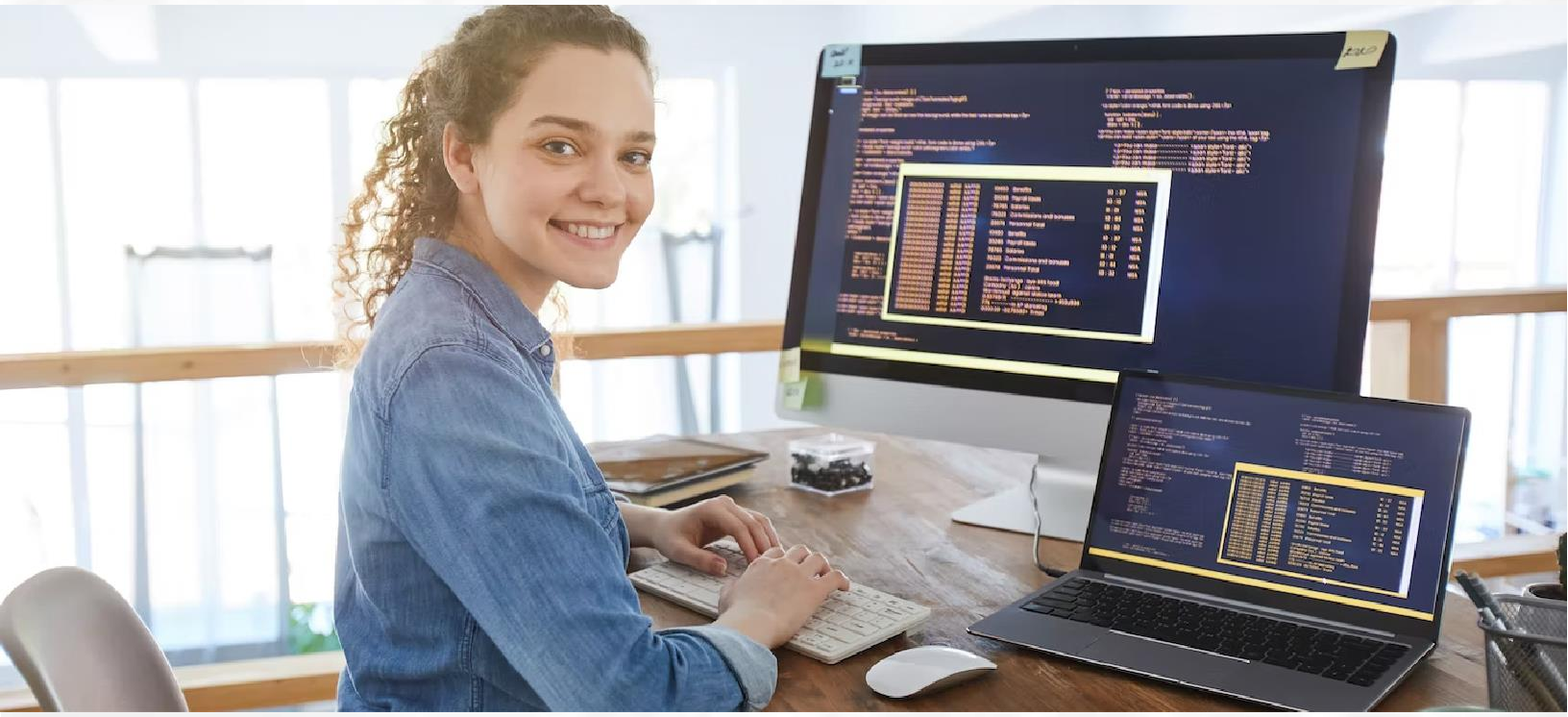
PROPOSAL FOR
DATA SCIENCE GO

About INNOVATICS

We are **INNOVATICS**, a holistic up-skilling platform driven by a unique, cohesive “Learn-Apply-Solve” framework. This innovative solution provides application-oriented immersive and interactive learning experience with extensive real-industry courses, cases, datasets and projects. It also ensures a blended pathway between industry and academia through simulation and context-ualisation.

INNOVATICS regularly presents at numerous conference workshops and until recently held regular monthly Meetups with industry experts as speakers.

We currently are a few multi-week, multi session courses that are live (then recorded) programs that participants have thoroughly enjoyed since we support our participants with almost endless one-on-one or group live support sessions.



Connect with a Program Advisor

Have questions about the program or how it fits in with your career goals?

SCHEDULE A CALL **+971 58 971 9975**

Visit our Website
www.innovatics.ai
S.smaoui@innovatics.ai