Data Science & AI - From Zero to Hero

A program you can repeat unlimited times at NO additional cost

Batch 17
January 21, 2023

Online LIVE Training Program

Big Bang Data Science Solutions
www.bbds.ma
register@bbds.ma
(315) 975-1661
2972 Webb Bridge Rd, Alpharetta, GA 30009

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.
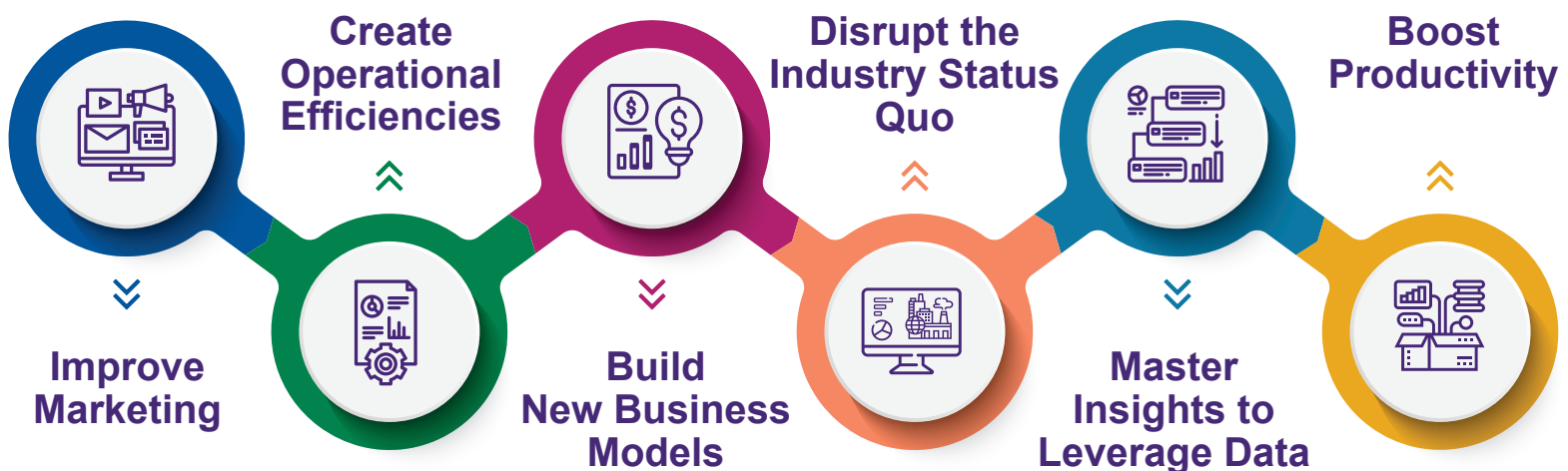
In the Data Science & Machine Learning program from zero to hero program, an executive education program led by the Big Bang Data, you will dive deep into learning the applied mathematical structure of core machine learning (ML) and artificial intelligence (AI) methods. The course focuses on the skills 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

Create Operational Efficiencies

Disrupt the Industry Status Quo

Boost Productivity

Improve Marketing

Build New Business Models

Master Insights to Leverage Data

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 become better 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 BBDS 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.

Participants Insights

- **Analytics** is a leadership problem

- **Analytics** can be used to address more business problems than they originally thought Analytics needs to be planned

- **Analytics** decisions must start with business problems, not with data

- **How** to quickly distinguish between good and bad analytics

Your team will help you to:

- **Identify** your abilities and preferences
- **Craft** a strong Linkedin profile
- **Increase** your visibility to prospective employers
- **Prepare** for job interviews in the field of data science
- **Develop** a portfolio to share with prospective employers that showcases your abilities and interests
- **Network** with peers, mentors, career coach, and the broader programming community

Who is this program for ?

BBDS 30-Week training program in Data Science & Artifitial Intelligence provides a hands-on introduction to the concepts, methods and processes of data analytics in business. It is designed for professionals looking to discover a new career path in data science or build on their existing technology career, and receive all of the career support that comes with it.

Upon completing this program, you will be ready to launch, build, or switch careers — to take advantage of new opportunities and be ready to face new challenges in the field of Data Science and Analytics. Whether it's smart grids, predictive marketing, automated factories, algorithmic trading, automation using machine learning, sensor data that is analyzed to create the internet of things, or healthcare data that is being analyzed to create new therapies … practically every industry, every company, and every professional is now using data to make decisions.

You will develop a data science portfolio that will help you join this growing community of data scientists, develop an online reputation and presence, and show prospective employers what you can do.

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.

Learn-by-Doing Approach gives Industry Readiness

- Horizontal applied learning through data case-based curriculum; hands-on practice via both point-click and coding; aided by learning videos, reading materials, access to professors, and structured mentorship.

- Virtual coding lab to build and implement/showcase one's ML based solutions; Practice on BBDS' cloud-based platform that enables faster learning, collaboration, and competitive hackathons.

- Vertical Domain Exposure across Banking, Retail, Manufacturing, etc.

- Industry standard virtual assessments and long-term

- Targeted career assistance via portfolio showcase and recruiter interface; Social Recognition and Personal Brand Creation.

Key Takeaways

- **Build** a working knowledge of data science

- **Identify** the value in utilizing analytics tools

- **Gather** the skills and confidence necessary to operate in a data-driven environment **Develop** the ability and

- intuition to judge "good" analytics"from "bad" analytics" **Understand** the importance of experimentation

- platforms to drive business growth **Learn** how to tell a persuasive story with data visualization tools

- Repeat unlimited time at no additional cost

PREREQUISITES:

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

WANTED: DATA SCIENTIST

In this role you will use data and your analytical ability to find and interpret rich data sources; manage large amounts of data; merge data sources; ensure consistency of datasets. You will create visualizations to aid in understanding data, build mathematical models using data, and produce and present data findings.

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.

Big Bang Data Science Solutions

Expertise : The Data Science from zero to hero course instructors are experts in data science, computer science and business transformation. They blend thought leadership with practical experience to craft a realistic, easy-to-understand course structure.

Engagement: The program follows a small group learning structure to foster direct interaction between peers and ins

Integration: Participants will develop a suite of interconnected learning modules and comprehensive skill-set.

Reputation: Previous participants have used this program to

land jobs in data science at:

1:1 Career Coaching

Small Group Mentoring Sessions

Live Webinars

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 BBDS 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

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

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

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

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

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

Launch - Build - Switch

Where do you fit in the expanding universe of Data science? What is your unique skill set?

One-to-one sessions with your **Career Coach** will help you develop your value proposition to employers.

Mentors will help you navigate the challenges specific to data science careers in small group sessions.

Regular **Webinars** will improve your networking and job search skills. You'll use these Webinars to develop and practice delivering your elevator pitch for different audiences, identify interview goals, and practice developing a rapport with prospective employers.

Program Faculty

Mo Medwani

Sr. Data Scientist, BBDS CEO

Founder of Big Data Science Solutions, is a PhD candidate 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.

Ed Bujak

Python Instructor

Data science and Python expert with over 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 electrical engineering and education, respectively. Ed is a public speak in everything related to programming, data science, and machine learning

Program Advisors

Shaukat Khan

VP of Public Relations

Dynamic professional with extensive experience in working with business partners in transforming and rebuilding teams. Areas of expertise include organizational assessments (including Outsourcing) and Production Control Support Functions. Experienced with different industries, including HR, insurance, retail healthcare (including compliance).

Noman Husain

Business Transformer

Business leader, entrepreneur, and change agent who galvanizes companies, communities, and individuals to thrive upon disruption and achieve excellence. He has over 20 years of experience in large-scale business transformation, is the founder of TRANSFORMability consulting agency and co-founder of IndustryX: Center for Operational Excellence, and has two master's degrees in mechanical engineering and industrial engineering, respectively.
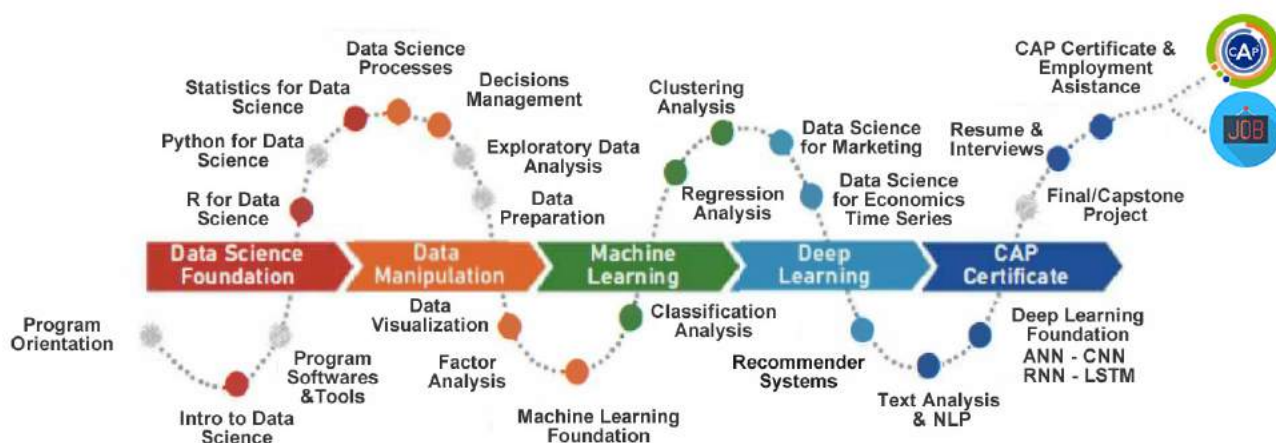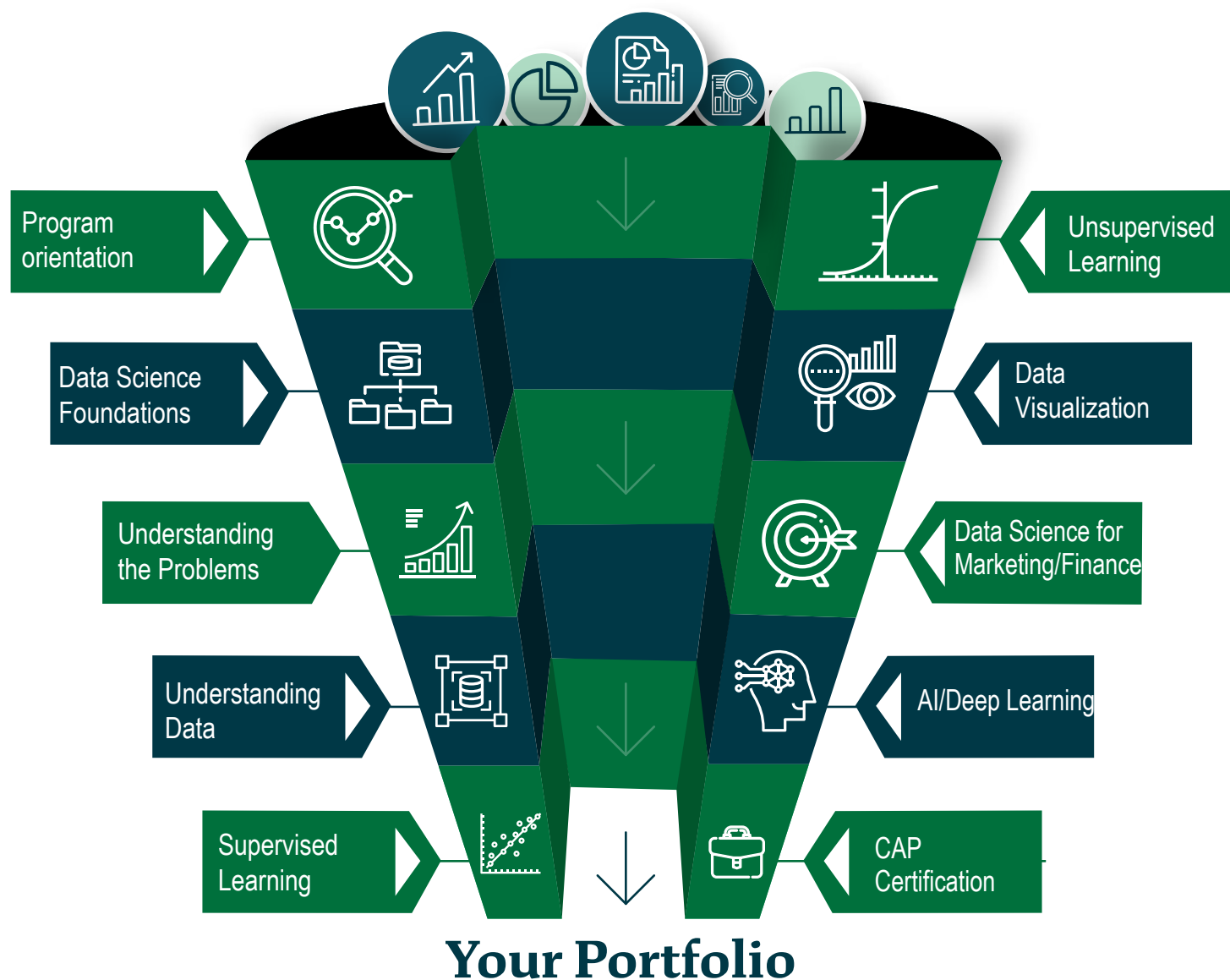
Fatima Ali

B.Dev. Manager

Business development initiative consultant that is responsible for client acquisition, marketing for the various programs. Extensive experience in HR, employee management & retention, employee training and development. Company projects streamlining and maintenance, record keeping and alignment of company goals with project delivery.

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.

Program orientation

Data Science Foundations

Understanding the Problems

Understanding Data

Supervised Learning

Unsupervised Learning

Data Visualization

Data Science for Marketing/Finance

AI/Deep Learning

CAP Certification

Your Portfolio

Program Orientation

Intro to Data Science

Python for Data Science

R for Data Science

Statistics for Data Science

Data Science Processes

Program Softwares &Tools

Decisions Management

Exploratory Data Analysis

Data Preparation

Data Visualization

Factor Analysis

Machine Learning Foundation

Clustering Analysis

Regression Analysis

Classification Analysis

Data Science for Marketing

Data Science for Economics Time Series

Recommender Systems

Text Analysis & NLP

Resume & Interviews

Final/Capstone Project

Deep Learning Foundation ANN - CNN RNN - LSTM

CAP Certificate & Employment Asistance

Data Science Foundation **Data Manipulation** **Machine Learning** **Deep Learning** **CAP Certificate**

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.

Program Learning Outcomes:

- Explore key concepts, tools, and techniques used in data science

- Explore examples of data science applications in the industry as a result of a dramatic growth in available data

- Learn or improve your coding in Python and R

- Apply the coding and data science knowledge acquired to real-world business challenges

- Experiment with data science tools and techniques

- Gain valuable career guidance in data science

Target Job Functions Include:

- Data Science Foundations
- R & Python Programming
- Machine Learning & AI
- Data Science Portfolio

Tools in the Program

python NumPy Keras TensorFlow matplotlib learn NLTK PyTorch

Seaborn Pandas statsmodels gensim R Studio ggplot2 R

tableau rapidminer MySQL SAS bigml Azure ML Flask H_2O

Why Now ?

The last few years have seen unprecedented growth in computational power, new users, new devices, new behaviors, new networks, platforms, and new data tools. The demand for data scientists — a 344 percent increase since 2013 alone — far outpaces the number of skilled applicants.

3X
Data Science jobs are expected to grow almost three times as fast as other job categories in the coming years.

(SOURCE: BUREAU OF LABOR STATISTICS)

Why Big Bang Data Science Solutions?

BBDS has a reputation for rigor, excellence, and cross-pollination across disciplines. What makes this program different?

- ✓ 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.

Your Fast Track to a Job in Data Science

Demand for data scientists crosses all industries as employers scramble to harness the power of data analytics to personalize their products, minimize risk, and expand into new markets.

The Professional Certificate in Applied Data Science & AI Online program from the **BBDS** positions professionals to take full advantage of this trend. You will gain valuable career guidance on interview preparation, networking, crafting a strong LinkedIn profile, and increasing your visibility to potential employers. You will also be able to learn from and network with your colleagues through peer discussions.

Duration
8 months Online
15-20 hours per week

Program Fee
$2,999

Schedule - 5 Time a week

Sat	9:00 am	to 12:00 pm est
Sun	9:00 am	to 12:00 pm est
Mon	8:00 pm	to 11:00 pm est
Wed	8:00 pm	to 11:00 pm est
Thur	8:00 pm	to 11:00 pm est

- Office Hours: Every day from 10:00 AM EST to 12:00 PM EST

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)
- BBDS 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	Projects, Teams & Team Leads

- Introduction to Data Science Project
- Projects Discussion (Milestone Projects Assignment)
- Projects List | Project Templates | Project Team

Session 4	Program Tools & Installation 1

- **Tableau**: Installation
- **Power BI** Installation
- **RapidMiner**: Installation

Session 5	Program Tools & Installation 2

- **R | RStudio**: Installation
- **MySQL**: Installation
- **Excel**

Week 2	Business Analytics & Business Understanding

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 this step can mean that a great deal of effort is put into producing the right answers to the wrong questions.

Session 1	Introduction to Data Science

- Data Science | Machine Learning | Deep Learning overview
- Certified Analytics Professional (CAP)
- Learning Path

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)
	    o Scientific Research Method
	    o Operations Research Method | Water Fall Method | CRISP-DM
	    o Exploration & Discovery
	    o 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 (R & Python Programming)

Learning Objectives: In this Week, students will be introduced to some of the programming languages commonly used, including Python and R. Students will work with the tools that professional Data Scientists work with, like Jupyter Notebooks, RStudio IDE, and others. Students will learn about what each tool is used for, what languages they can execute, and their features and limitations.

Session 1	Basic R Programming
	• R & RStudio • Fundamentals Working Directory | Assigning Objects | Comments • Data Structures (Vectors, Factors, Lists, Data frames, Matrix, Matrices)
Session 2	Advanced R Programming
	• Data Structures (Vectors, Factors, Lists, Data frames, Matrix, Matrices) • Functions & Packages

Session 3	Basic Python Programming

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

Session 4	Advanced Python Programming

	• Python files I/O Functions/Numbers/Strings and related operations
	   o Tuples and related operations
	   o Lists and related operations
	   o Dictionaries and related operations
	   o Sets and related operations

Session 5	Crash Course in SQL & Advanced SQL
	• Basic SQL Queries • Manipulating Values with Functions • SQL Theory • Working with Multiple Tables • Advanced Functions and Queries

Week 4	Data Science Foundations (Statistics - Probability - Algebra)

Learning Objectives: In this Week, students will build a solid foundation of statistics for Data Science, learn probability, distributions, hypothesis testing and algebra.

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.

Session 1	Introduction to Statistics
	• 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	Introduction - Git and GitHub
	- Introduction To Git - Installing Git | Configure Git - Basic commands - Exploring Git log | Git diff - Making corrections Git reset - Branching
Session 4	Introduction to Linear Algebra
	- Vectors | Matrices - Transpose of a matrix | Inverse of a matrix - Determinant of a matrix | Trace of a matrix - Dot product | Eigenvalues & Eigenvectors
Session 5	Team Projects 1st Presentation (Framing Business Problem)
	- Team Projects (Business Understanding) - Team Leads' First Presentation
Week 5	Data Science Foundations (Data Viz with Tableau & Power BI)

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
	- 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 the project 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 in (R & Python)

- Read data from different sources
- Get to Know the Data
    - 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
    - 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	EDA with RapidMiner	Excel

- **RapidMiner**
    - Installation | RapidMiner Basics | Studio - GUI Intro | Visualizing Data in RapidMiner | Marketplace – Extensions
- **Excel**
    - Data Entry Techniques in Excel
    - How to Make Your Spreadsheets Look Professional
    - Inserting, Deleting, and Modifying Rows & Columns
    - Excel Formulas for Beginners | Excel Functions
    - Work Efficiently by Using Cut, Copy, and Paste
    - Formatting | Pasting Values, Formulas and Formats with Paste Special

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 Data Understanding 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

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

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 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

Continuing from Week 8, you add several more Supervised Learning classification algorithms to your arsenals. Machine learning topics taught this week to involve parametric classification algorithms (Logistic Regression, Support Vector Machine, Nearest Neighbor Classifiers) with deeper use of R & 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
- Direct Methods for Rule Extraction
- Indirect Methods for Rule Extraction
- 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
- Logistic Regression 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
- Soft-margin SVM
- Nonlinear 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 Linear Regression & Polynomial Linear Regression) with deeper use of R & scikit-learn functionality.

This week is a continuation of Regression Analysis techniques learned in week 10. Machine learning topics taughtthis week to involve non-linear regression algorithms (Decision Tree and Support Vector Machine) and Regularization Techniques (Lasso, Ridge, Elastic Net, Least Squared) with deeper use of R and 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 (R & Python)

- 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

Week 12	Model Evaluation & Optimization

Learning Objectives: Model building is an iterative process. Employing Feature Engineering techniques along with a 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

- Estimating the Confidence Interval for Accuracy
- Comparing the Performance of Two Models
- XG-Boost in R & Python

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 learn about commonly used clustering techniques like K-Means Clustering and Hierarchical Clustering along with anomalies detection algorithms with deeper use of R & 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 | Association Discoveries
- **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

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 to analyze 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
- Characteristics of 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	Clustering 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	Statistics Based Techniques

- Z-score
- Modified Z-score
- Interquartile Range (IQR)
- Boxplot
- Histogram

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 all the 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: Collaborative Filtering
	• Matrix Factorization Model (Object recommendation) • Content Filtering | Collaborating Filtering (CF) • CF Neighborhood-Based Approach • CF Location-Based Approach

Session 4	Recommender Systems: Matrix Factorization & Hybrid R. Engine
	• Popularity Based • Content Based • Collaborative filtering • Matrix Factorization (MF)

Session 5	Python: Matplotlib Library
	• Annotation | Autocorrelation | Curves • Scales | Shapes | Box Plot | Violet Plot • Histograms | Pie Chart | Twin Axis | Figures | Color Map

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 for forecasting.

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
- Example of Triple Exponential Smoothing
- Exponential Smoothing Summary

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
- Detection of structural breaks
- Control of the process

Session 3	SARIMAX Model

- The limitations of ARIMA
- The SARIMA extension of ARIMA
- Implementing SARIMA method using the Statmodels library

Session 4	Python: Seaborn Library

- Visualizing Distribution to find Patterns
- Categorical Data and Multi-Panel data/
- Customizing the Facet Grid/Subset using the Paid Grid
- 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 — NLP & Text Analysis

- Basic feature extraction using text data
- Basic Text Pre-processing of text data
    - Lower Casing | Punctuation Removal | Stop Words Removal | Frequent Words Removal | Rare Words Removal | Spelling Correction | Tokenization | Stemming |Lemmatization
- Advance Text Processing
    - N-grams | Term Frequency | Inverse Document Frequency (TF-IDF) | Bag of Words |Sentiment Analysis | Word Embedding

Session 2 — Sentiments Analysis

- Motivations
- Why Sentiment Analysis?
- Types of Sentiment Analysis
    - Polarity Detection | Emotion Detection | Intent Detection

Session 3 — Topic Model using LDA

- Motivations
- Latent Dirichlet Allocation (LDA)
- Dirichlet Distribution | Dirichlet Output
- LDA & Matrix Factorization
- Nonnegative Matrix Factorization

Session 4 — Building an Intent Based Rasa Chatbot

- Introduction to Rasa
    - Intents, Entities, Slots, and Responses
- Domain File
- Train a Model

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	Deep Learning: ANN & CNN

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 Artificial Neural 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

- How to Install PyTorch
- PyTorch Deep Learning Model Lifecycle
  - Prepare the Data | Define the Model | Train the Model | Evaluate the Model |Predictions
- How to Develop PyTorch Deep Learning Models
  - How to Develop an MLP for Binary Classification
  - How to Develop an MLP for Multiclass Classification
  - How to Develop an MLP for Regression
- How to Develop a CNN for Image Classification

Week 19	Sequence Learning

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	Introduction to Transformer Model (BERT & GPT2)

- BERT
    - Pre-training Tasks
    - Input Embedding
    - Use BERT in Downstream Tasks
- OpenAI GPT-2
    - Zero-Shot Transfer
    - BPE on Byte Sequences
- Model Modifications

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 20	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 & General Deployment Considerations

- Preliminaries: Statistical Testing
- Modeling Null & Alternative Distributions
- Statistical Testing for Classification | Association Analysis | Cluster Analysis | Anomaly Detection

Session 2	ML Deployment with FLASK

- Docker Basics | Flask Basics
- Deployment Steps & Challenge
  - o Load your model with its weights - Preprocess your data - Perform the actual prediction - Handle the prediction response data
- Loading model directly into an application

Session 3	ML Deployment in Cloud - SAGEMAKER

- Calling an API
  - o Kubernetes - Custom REST-API with Flask | Django - AWS Lambda | Serverless
- Apache Beam
- Spark | Flink

Session 4	Python: TensorFlow & Keras Library 2

- 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

Session 5	Team Projects Final Presentation (Insights & Action Plan)

- Team Projects (Analytical Approach Implementation)
- Team Leads' Final Presentation

Week 25	Individual Capstone Project (1 Month)

Learning Objectives: During the final week, you will be transitioned into full-time focus on your final, passion individual 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 12th of January 2023 till the 19th of February 2023.

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".

Career Services

Resume Preparation | Interview Preparation | Employment Assistance Certificate Analytics Professional Training

Module 26	Data Science Interview Questions & Answers

Learning Objectives: In this module, you will be introduced to 120+ Data Science Interview questions and answers covering the topics like:

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

Module 27	Data Science Interview Real Challenges)

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

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

Module 28	Resume Preparation

Learning Objectives: In this module, you will learn how to build an affective Data Science resume:

- o Structure of your Data Science Resume
- o Adding Content and Information to your Data Science Resume
- o Get Feedback from Industry Experts
- o Build your Digital Presence

Module 29	Certified Analytics Professional (CAP Exam) Self-paced

Learning Objectives: BBDS 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 outlined by the CAP®Job Task Analysis. Apply concepts learned in each course module through a case study developed specifically for CAP® Prep users.

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

Module 30	Big Data & Hadoop Developer. – (Self-paced)

Learning Objectives: In this last module, you will be introduced to Big Data technologies. This is a self-paced course and topics covered, but not limited to:

- Introduction to Linux and Big Data Virtual Machine (VM)
- Understanding Big Data
- HDFS (The Hadoop Distributed File System)
- MapReduce Architecture
- Hadoop Streaming

- MR Algorithm and Data Flow
- MR algorithm s (Non- graph)
- Different types of NoSQL databases
- Pig and Hive
- Spark

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.

After successful completion of the program, your verified digital certificate will be emailed to you in the name you used when registering for the program. All certificate images are for illustrative purposes only and may be subject to change at the discretion of Big Bang Data Science Solutions

Testimonials

Here is another reason why you should trust Big Bang Data Science Solutions (BBDS) with your training. Watch this recording capturing BBDS prestigious Award from Internet 2.0 Conference - Dubai 2022
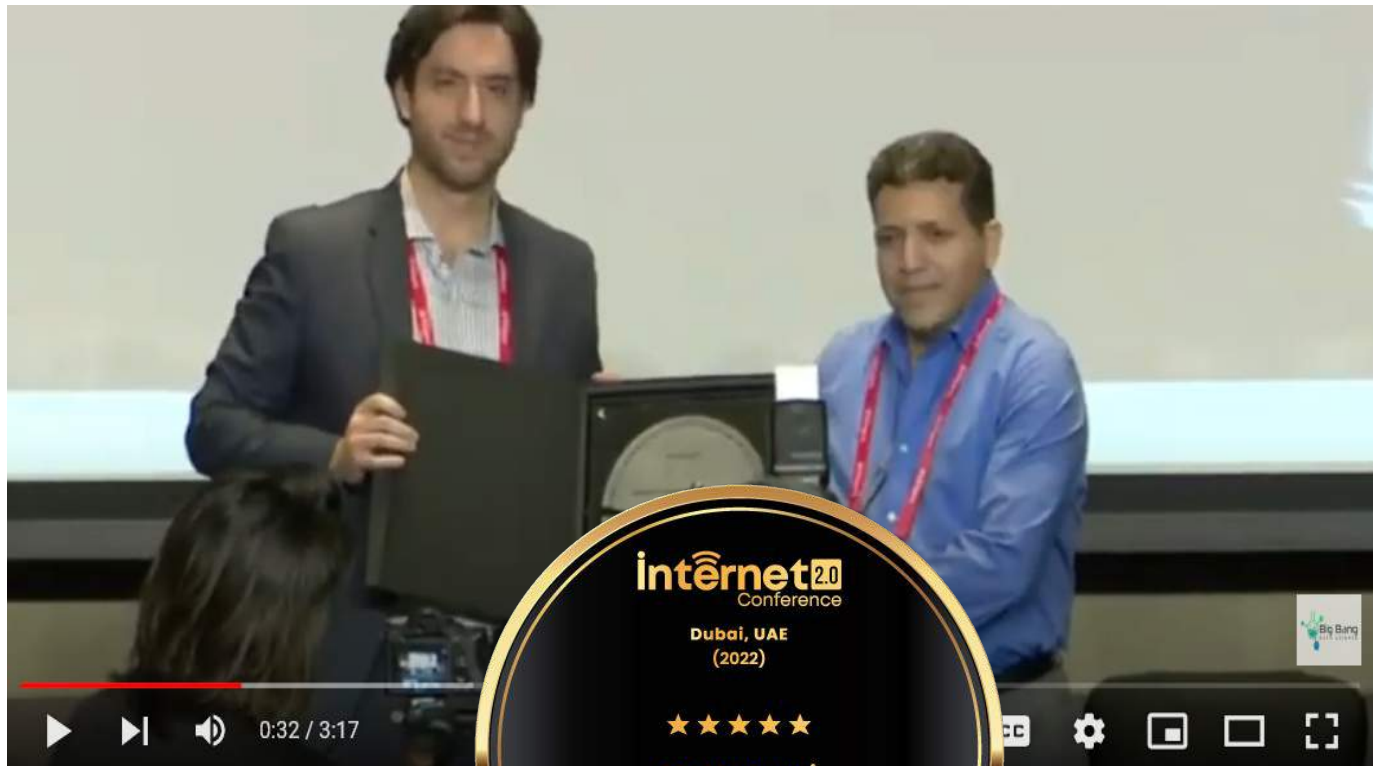
About BBDS...

We are Big Bang Data Science Solutions (BBDS), 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.

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

We currently offer 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?

+1 (315) 975-1661

✉ **register@bbds.ma**   🌐 **www.bbds.ma**

DURATION
30 Weeks

FORMAT
Online

PROGRAM FEE
US$2,999

ENROLL

SCHEDULE A CALL