



# Data Science & Machine Learning -

## Novice to Expert

Master Data Science and Machine Learning, from the critical foundations of statistics and probability to working hands-on with Machine Learning models using Python and R, two of the world's most popular programming languages.

January 9, 2021

30 Weeks

Batch - 13



**We're “Ridiculously Committed”!**  
**We will not quit until you achieve your learning goals.**



## ATC Innovation Center

2972 Webb Bridge Rd, Alpharetta, GA 30009  
[www.bbds.ma](http://www.bbds.ma) | [register@bbds.ma](mailto:register@bbds.ma) | 315-975-1661

Hundreds of online courses exist today. What many of them lack, however, is a commitment to helping you translate your knowledge into something tangible - the ability to excel and grow as an AI/ML professional.

To tackle this, the BBDS 30 Week Training Program in Data Science & Machine Learning has been designed to give you the academic rigor, learning support, and peer interaction of a full-time course with the flexibility of an online program.

The program uniquely combines a comprehensive curriculum, covering the most widely-used tools and techniques in the industry, with a hands-on learning approach. A structured learning journey keeps you on track throughout as you achieve your weekly learning milestones with your mentor and benefit from their rich professional experience.

Following a “**learn by doing**” pedagogy, the program offers you the opportunity to apply your skills and knowledge in real-time every week through interactive mentor-led practice sessions, quizzes, assignments, and hands-on projects. As you do so, you come to truly appreciate the nuances of data and build your portfolio in the process.

On a whole, the program empowers you with the skills, body of work, and job market insights you need to find the right career opportunities or lead AI and ML in your current organization. All this comes with the credibility, global advantage, and academic leadership of Big Bang Data Science Solutions.

## Faculty Profiles



**Mo Medwani**

**Data Scientist, Director – Training Programs**

- **Ph.D.** Candidate (Artificial Intelligence)
- 4 Master Degrees: MS in Machine Learning | MS in Data Science | MS in IT | MBA
- 20+ years of IT experience (Service Delivery Management)
- 8+ Years of experience in Data Science & related technologies



**Ed Bujak**

**Data Scientist, Python Instructor**

- 26 years of IT experience and 6 years of Data Science experience
- 20 years of education experience: Computer Science, Mathematics, Engineering
- 2 masters degrees (MS Electrical Engineering & Science, MS Education)
- Public speaker on everything programming, Data Science, Machine Learning

# About the Program

## What you will learn

- You will learn about Python & R and survey tools available in them for acquiring, cleaning, visualizing & analyzing data.
- You will learn how to extract data from files (CSV, HTML, JSON, XML) and databases (Mysql), the rudiments of data cleaning, and examine data analysis, machine learning and data visualization packages available in Python & R.
- You will learn models and methods used in machine learning and deep learning, and apply them to real-world situations ranging from identifying trending news topics to building recommendation engines, ranking sports teams, and plotting the path of movie zombies.
- You will learn Machine Learning types, Natural Language Processing (NLP) and network analysis available in Python | R | RapidMiner | SAS for Enterprise Miner...
- Certified Analytics Professional Training (CAP)

## Key Learning Outcomes

- Build your expertise in the most widely-used AI & ML tools and technologies.
- Acquire the ability to independently solve business problems using AI & ML.
- Master the skills needed to build machine learning and deep learning models.
- Develop know-how of the applications of AI in areas such as Computer Vision & NLP.
- Understand the possibilities and implications of AI in different industries.
- Build a substantial body of work and an industry-ready portfolio in AI & ML.

## What if you miss a class?

- If you miss a session, you can choose either or both of the two options:
- View the recorded session of the class available in your LMS account.
  - Schedule a 1:1 session to recap what you missed.

## Time & Schedule

- **Saturday : 9:00 AM EST to 12:00 PM EST**
- **Monday : 8:00 PM EST to 11:00 PM EST**
- **Wednesday : 8:00 PM EST to 11:00 PM EST**
- **Thursday : 8:00 PM EST to 11:00 PM EST**
- **Friday : 8:00 PM EST to 11:00 PM EST**
- **Office Hours: Every day from 10:00 AM EST to 2:00 PM EST**



**STARTS ON**

**January 9 , 2021**



**DURATION**

**8 months, LIVE  
15-18 hours per week**



**PROGRAM FEE**

**US\$2,999** or get US\$235 off with a referral

① [Flexible payment available](#)

① [Special group enrollment pricing](#)



# Key Features



## Interactive Learning

- ✓ Online – **Live** streaming
- ✓ Collaborative yet personalized learning in groups of 10-15



## Comprehensive Curriculum

- ✓ Designed to build industry-valued skills
- ✓ Machine Learning, Computer Vision, Natural Language Processing, Neural Networks



## Employment Assistance

- ✓ Resume preparation thru CVCompiler
- ✓ Interview Preparation
- ✓ Employment Assistance



## 24/7 Support

- ✓ Personalized 1:1 career coaching
- ✓ Mentoring Sessions and Office Hours with Instructors
- ✓ Repeat unlimited time at no additional cost



## Grades/Quizzes

- ✓ **Graded program:** 3 Quizzes per week (True | False & Multiple Choices)



## Weekly Labs

- ✓ **3 Lab** assignments per week: 1 Python, 1 R, & 1 Tableau assignment



## Team Projects

- ✓ **16+ Industry Relevant TEAM Projects:** Prediction-Clustering-Classification – TS



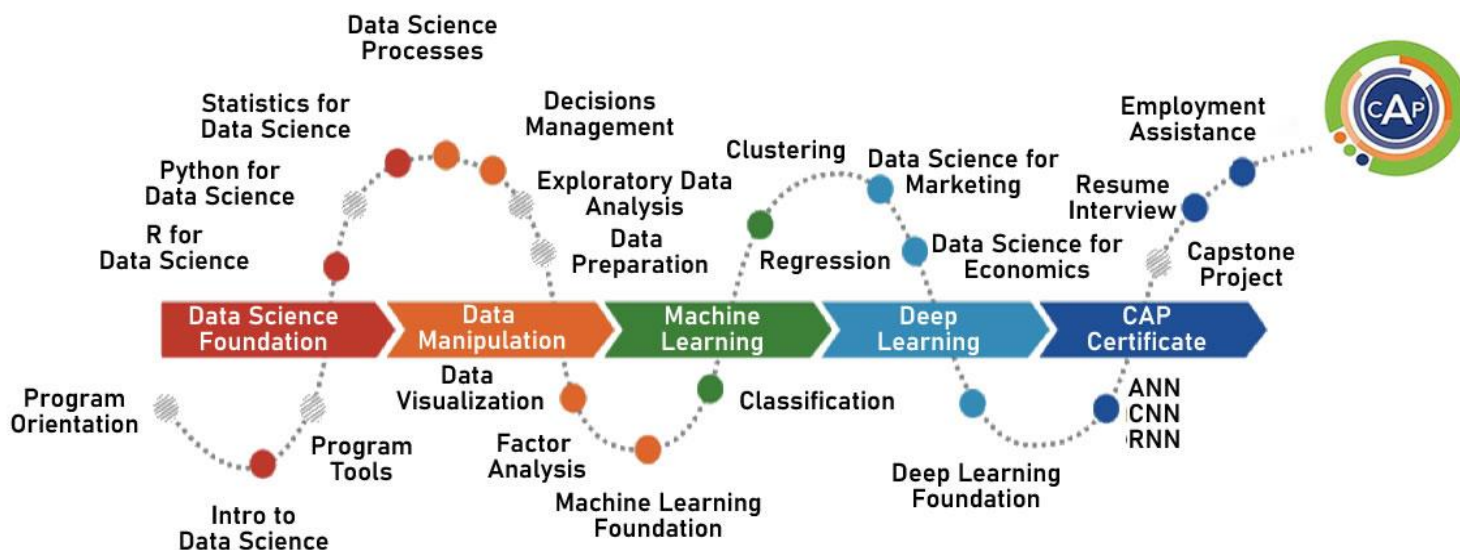
## Final Capstone Project

- ✓ **1 Final Personalized capstone project**

# Technologies & Tools covered



# Structured Learning Journey



# Program Syllabus

## Week 1

## Program Orientation – DataCamp Assignments

**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.

Also, you will have access to recorded video lectures on DataCamp to complete Data Science prerequisites courses (Python introductory, R introductory, SQL, Statistics, probability and Math) .

### Session 1

### Program Orientation – DataCamp Assignments

- Program Orientation (Agenda – Curriculum – DataCamp – Assignments - Projects)
- BBDS Website | Canvas | Slack walk through

### Session 2

### Why you should 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

### Program Data Science Projects – Collaborative teams

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

### Session 4

### Program Tools & Installation

- **Tableau:** Installation
- **Power BI** Installation
- **RapidMiner:** Installation
- **R | RStudio:** Installation

### Session 5

### Program Tools & Installation

- **MySQL:** Installation
- **BigML:** Account Creation
- **SAS For Enterprise Miner:** Account Creation & Installation
- **Azure** Installation

## Week 2

## Business Analytics & Business Understanding

**Learning Objectives:** The basic work-flow is now in place. You will dive into the first stage of the Cross-Industry Standard Process for Data Mining (CRISP-DM) process which is to understand what the customer wants to accomplish from a business perspective. Customers often have competing objectives and constraints that must be properly identified and balanced.

Your 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	What is Data Science, Deep Learning, Machine Learning?
	<ul style="list-style-type: none"> <li>• Data Science   Machine Learning   Deep Learning overview</li> <li>• Certified Analytics Professional (CAP)</li> <li>• Learning Path</li> </ul>
Session 2	Decisions Management & Problem Framing
	<ul style="list-style-type: none"> <li>• Quantifying Business Problem</li> <li>• Converting Business Problems into Analytics Solutions</li> <li>• Defining the objectives, assumptions, contains, analytical approaches</li> <li>• Designing the Analytics Base Table &amp; Implementing Features</li> <li>• Identifying Y variable &amp; Measures of Success</li> <li>• Stakeholders &amp; Analytical team</li> </ul>
Session 3	Introduction to Data Science Process - CRISP-DM
	<ul style="list-style-type: none"> <li>• Solutions Methodologies (Macro vs. Micro) <ul style="list-style-type: none"> <li>○ Scientific Research Method</li> <li>○ Operations Research Method   Water Fall Method   CRISP-DM</li> <li>○ Exploration &amp; Discovery</li> <li>○ Solutions are Dependent on Data   Solutions are Independent on Data</li> </ul> </li> <li>• CRISP-DM Process <ul style="list-style-type: none"> <li>○ Business Understanding   Data Understanding   Data Preparation   Modeling   Evaluation   Deployment</li> </ul> </li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li>• <b>Tableau:</b> <ul style="list-style-type: none"> <li>○ Introduction to Tableau   Your First Bar chart</li> <li>○ Connecting Tableau to a Data File - CSV File</li> <li>○ Navigating Tableau   Creating Calculated Fields</li> <li>○ Adding Colors</li> <li>○ Adding Labels and Formatting   Exporting Worksheets</li> </ul> </li> <li>• <b>Power BI</b> <ul style="list-style-type: none"> <li>○ Getting Started with Power BI</li> <li>○ Basic Concepts of Business Intelligence</li> <li>○ Connecting to Different Data Sources</li> </ul> </li> </ul>
Session 5	Crash Course in Python Programming
	<ul style="list-style-type: none"> <li>• Jupyter Notebook &amp; Notebook Azure</li> <li>• Introduction to Python <ul style="list-style-type: none"> <li>○ Introduction to a basic construct in Python</li> <li>○ Understanding indentation like tabs and spaces</li> <li>○ Overview of Python</li> <li>○ The Companies using Python</li> <li>○ Different Applications where Python is used</li> <li>○ Discuss Python Scripts on UNIX   Windows</li> <li>○ Values, Types, Variables</li> <li>○ Operands and Expressions</li> </ul> </li> </ul>

## Week 3

## Data Science Foundations

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

You will also learn basis statistics which is undoubtedly an essential element to become a Data Scientist. It is how the analysis is made towards better decisions and gaining deep insights into the data. Without statistics and mathematics, a Data Scientist is just guessing predictions and not formulating them.

Session 1	Basic R Programming
	<ul style="list-style-type: none"><li>• R &amp; RStudio</li><li>• Fundamentals Working Directory   Assigning Objects   Comments</li><li>• Data Structures (Vectors, Factors, Lists, Data frames, Matrix, Matrices)</li></ul>
Session 2	Advanced R Programming
	<ul style="list-style-type: none"><li>• Data Structures (Vectors, Factors, Lists, Data frames, Matrix, Matrices)</li><li>• Functions &amp; Packages</li></ul>
Session 3	Basic Python Programming
	<ul style="list-style-type: none"><li>• Python data types (basic and Boolean), conditional statements, functions, assignment operations<ul style="list-style-type: none"><li>○ Python built-in data types</li><li>○ Basic operators in Python</li><li>○ Loop and control statements like break, if, for, continue, else, range () and more</li></ul></li></ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"><li>• <b>RapidMiner:</b><ul style="list-style-type: none"><li>○ Installation   RapidMiner Basics   Studio - GUI Intro   Visualizing Data in RapidMiner   Marketplace – Extensions</li></ul></li><li>• <b>Case study # 1</b><ul style="list-style-type: none"><li>○ Academic Performance Analysis</li></ul></li></ul>
Session 5	Python Sequences and File Operations
	<ul style="list-style-type: none"><li>• Python files I/O Functions/Numbers/Strings and related operations<ul style="list-style-type: none"><li>○ Tuples and related operations</li><li>○ Lists and related operations</li><li>○ Dictionaries and related operations</li><li>○ Sets and related operations</li></ul></li></ul>

## Week 4

## Statistics for Data Science

**Learning Objectives:** In this module, we will build a Solid Foundation of Statistics for Data Science, Learn Probability, Distributions, Hypothesis Testing. 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 you to define statistics and essential terms related to it, explain measures of central tendency and dispersion, and comprehend skewness, correlation, regression, distribution. You will be able to make data-driven predictions through statistical inference.

Session 1	Introduction to Statistics & Probability
	<ul style="list-style-type: none"> <li>• Data</li> <li>• Levels of Measurement</li> <li>• Measures of Central Tendency</li> <li>• Population and Sample</li> <li>• Measures of Dispersion</li> <li>• Quartiles and IQR</li> </ul>
Session 2	Introduction to Distribution
	<ul style="list-style-type: none"> <li>• Introduction to Distribution</li> <li>• Uniform Distribution</li> <li>• Binomial Distribution</li> <li>• Poisson Distribution</li> <li>• Normal Distribution</li> <li>• Skewness</li> <li>• Standardization and Z Score</li> </ul>
Session 3	Advanced Statistics
	<ul style="list-style-type: none"> <li>• Central Limit Theorem</li> <li>• Hypothesis Testing</li> <li>• ANOVA - Analysis of Variance</li> <li>• Chi-Square Analysis</li> </ul>
Session 4	Team Projects 1st Presentation (Framing Business Problem)
	<ul style="list-style-type: none"> <li>• Team Projects (Business Understanding)</li> <li>• Team Leads' First Presentation</li> </ul>
Session 5	Python & R for Statistics
	<ul style="list-style-type: none"> <li>• Loading Data</li> <li>• Dataset Preparation   Dealing with Outliers   Distribution Histograms Box and Violin Scatter Matrix</li> <li>• Correlation   Mean   Median   Mode   Skewness   Kurtosis   Percentiles</li> <li>• Sampling and the Central Limit Theorem   Pearson's Chi2 Test</li> <li>• Exploratory Data Analysis Summary</li> </ul>

## Week 5 Data Understanding | Exploratory Data Analysis

**Learning Objectives:** Continuing with the topics from Week 4, you will be introduced the second phase of the Cross-Industry Standard Process for Data Mining (CRISP-DM) process model, you obtain data & verify that it is appropriate for the needs. You might identify issues that cause you to return to the 'Business Understanding' phase of the project & revise your plan. You may even discover flaws in the 'Business Understanding', another reason to rethink goals & plans.

The Data Understanding phase includes four tasks. These are Gathering data – Describing data – Exploring data – Verifying data quality.



Session 1	Data Types- Measure of Shape - Position - Dispersion
	<ul style="list-style-type: none"> <li>Measures of the Spread: Range – IQR – Variance – Standard Deviation</li> <li>Measures of Dispersion</li> <li>Measures of Position: Statistical Analysis (John T. 5 Numbers Summary)</li> <li>Measures of Relationships: Correlation   Variance &amp; Covariance</li> <li>Measures of Shapes: Skewness &amp; Kurtosis</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 2	Data Loading & Data Manipulation
	<ul style="list-style-type: none"> <li>Read data from different sources</li> <li>Get to Know the Data <ul style="list-style-type: none"> <li>Identifying Categorical Data: Nominal, Ordinal and Continuous</li> <li>Univariate   Bivariate   Multivariate Analysis</li> <li>Types of Data   Type of Central Tendency   Data distribution   Data Density</li> </ul> </li> </ul>
Session 3	Data Visualization (Numerical & Graphical Descriptive Statistics)
	<ul style="list-style-type: none"> <li>Data Visualization (Histograms, Bar Plot, Scatter Plot, Box Plot)</li> <li>Data Quality Report   Develop the Code Book <ul style="list-style-type: none"> <li>Summary of Data   Type of variable   Ranges of variables   Missing fields   Identify the primary list of variables to solve the business problem</li> </ul> </li> <li>Detecting missing values &amp; outliers   Duplicates &amp; redundant records</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li><b>Tableau:</b> <ul style="list-style-type: none"> <li>Time series, Aggregation, and Filters   Working with Data Extracts in Tableau</li> <li>Working with Time Series   Understanding Aggregation, Granularity, and Level of Detail   Creating an Area Chart &amp; Learning About Highlighting   Adding a Filter and Quick Filter</li> </ul> </li> <li><b>Power BI</b> <ul style="list-style-type: none"> <li>Data Transformation and Modeling   Basic Data Transformations   Splitting Columns and Changing Data Types   Conditional Columns and Merge Queries   Creating Calculated Columns   Manage Your Data Relationships</li> </ul> </li> </ul>
Session 5	OOPs in Python
	<ul style="list-style-type: none"> <li>Understanding the OOP paradigm like encapsulation, inheritance, polymorphism and abstraction</li> <li>What are access modifiers, instances, class members?</li> <li>Classes and objects   Function parameter and return type functions/Lambda expressions.</li> </ul>
Week 6	Data Preparation & Preprocessing (Transformations)
<p><b>Learning Objectives:</b> In the Data Preparation phase, you 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 reformatting data, making corrections to data and the combining of data sets to enrich data.</p> <p>You also learn how to fix data quality issues discovered through EDA &amp; Data visualization from the Data Understanding phase.</p>	

Session 1	Data Preparation & Processing (Categorical Attributes)
	<ul style="list-style-type: none"> <li>Encoding Categorical Data <ul style="list-style-type: none"> <li>Replacing values   Creating Dummy Variables   Encoding labels   One-Hot encoding   Binary encoding   Backward difference encoding   Miscellaneous features</li> </ul> </li> <li>Data Quality Report</li> <li>Develop the Code Book <ul style="list-style-type: none"> <li>Summary of Data   Type of variable   Ranges of variables   Missing fields   Identify the primary list of variables to solve the business problem</li> </ul> </li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 2	Data Preparation & Processing (Numeric Attributes)
	<ul style="list-style-type: none"> <li>Statistical Analysis (John T. 5 Numbers Summary)</li> <li>Variance &amp; Covariance   Correlation Analysis</li> <li>Dealing with Skewness &amp; Kurtosis   Missing Values &amp; Outliers   Duplicates &amp; redundant</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Data Preparation & Processing (Attributes Transformation)
	<ul style="list-style-type: none"> <li>Variable Conversion <ul style="list-style-type: none"> <li>Discretization</li> <li>Binning</li> </ul> </li> <li>Variable Transformation <ul style="list-style-type: none"> <li>Normalization   Standardization</li> <li>Factorizations   Binarization</li> </ul> </li> <li>Box-Cox Transformation</li> <li>Data Partitioning</li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li><b>RapidMiner:</b> <ul style="list-style-type: none"> <li>Turbo Prep – Introduction</li> <li>Data Cleansing   Merging Data</li> <li>Data Pivoting   Data Preparation   Connecting to Databases</li> </ul> </li> <li><b>Case study # 2</b> <ul style="list-style-type: none"> <li>Cities Population Analysis</li> </ul> </li> </ul>
Session 5	Intermediate Python
	<ul style="list-style-type: none"> <li>Data Structures for Analysis <ul style="list-style-type: none"> <li>Lists</li> <li>Dictionaries</li> <li>Mutability and iterations with examples on data structures</li> </ul> </li> </ul>
Week 7	Data Preparation & Preprocessing (Factor Analysis)
<p><b>Learning Objectives:</b> 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 &amp; many more.</p>	

Session 1	Features Engineering   Extraction   Selection
	<ul style="list-style-type: none"> <li>Goals of Feature Selection</li> <li>Classes of Feature Selection Methodologies <ul style="list-style-type: none"> <li>Correlation Method</li> <li>Machine Learning Method   Feature Importance</li> </ul> </li> <li>Regularization &amp; Embedded Feature Selection</li> <li>Feature Subset Selection   Feature Creation</li> </ul>
Session 2	Dimensionality Reduction (PCA & KPCA)
	<ul style="list-style-type: none"> <li>Factor Analysis: PCA   KPCA</li> <li>Aggregation   Sampling   Dimensionality Reduction</li> <li>Feature Subset Selection   Feature Creation</li> <li>Discretization &amp; Binarization   Variable Transformation   Data Partitioning</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Dimensionality Reduction (LDA & SVD)
	<ul style="list-style-type: none"> <li>Factor Analysis: LDA &amp; SVD</li> <li>Aggregation   Sampling   Dimensionality Reduction</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li><b>Tableau:</b> <ul style="list-style-type: none"> <li>Maps, Scatterplots, and Your First Dashboard   Relationships vs Joins (v00update)</li> <li>Joining Data in Tableau   Creating a Map, Working with Hierarchies</li> <li>Creating a Scatter Plot, Applying Filters to Multiple Worksheets   Creating a Dashboard   Adding an Interactive Action - Filter   Adding an Interactive Action - Highlighting</li> </ul> </li> <li><b>Power BI</b> <ul style="list-style-type: none"> <li>Power BI Desktop Visualizations   Introducing Various Visuals   Communicating with Charts</li> <li>R Visual Integration   Page Layout and Formatting</li> </ul> </li> </ul>
Session 5	Advanced Python
	<ul style="list-style-type: none"> <li>Generators, Comprehensions and Lambda Expression <ul style="list-style-type: none"> <li>Generators and Yield   Next and Ranges   The os. walk Generator   Searching the Filesystem   Reading Mp3 Tags   List Comprehensions</li> <li>List Comprehensions and Side-Effects   Conditional Comprehensions/Conditional Expressions</li> <li>Nested Comprehensions   The timeit Module/Map Intro   The Filter Function</li> <li>The Reduce Function   any and all/Named Tuples   any and all with Comprehensions</li> </ul> </li> </ul>
Week 8	Supervised Learning - Classification (Part 1)
<p><b>Learning Objectives:</b> This module broadens concepts learned in Data Understanding, Data Preparation &amp; Exploratory Data Analysis (EDA) by extending to Machine Learning where you will learn about the models &amp; methods used in machine learning &amp; apply them to real-world.</p> <p>The aim of supervised machine learning is to build a model that makes predictions based on evidence in the presence of uncertainty. In this session, you will learn about the different algorithms of supervised learning such as Decisions Tree, Rule-Based classifier &amp; Naive Bayes</p>	

<b>Session 1</b>	<b>Introduction to Machine Learning (Baseline Algorithms Analysis)</b>
	<ul style="list-style-type: none"> <li>• Introduction to Machine Learning from different perspectives</li> <li>• Understand the Need of Machine Learning</li> <li>• Understand Machine Learning major elements</li> <li>• Understand Machine Learning tasks &amp; types of Machine Learning</li> <li>• Understand Types of Algorithms &amp; types of Analysis</li> <li>• Understand Algorithms Performance</li> </ul>
<b>Session 2</b>	<b>Decision Tree &amp; Rule-Based Classification in R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• Algorithm for Decision Tree Induction</li> <li>• Methods for Expressing Attribute Test Conditions</li> <li>• Measures for Selecting an Attribute Test Condition</li> <li>• Characteristics of Decision Tree Classifiers</li> <li>• How a Rule-Based Classifier Works</li> <li>• Properties of a Rule Set</li> <li>• Direct Methods for Rule Extraction</li> <li>• Indirect Methods for Rule Extraction</li> <li>• Characteristics of Rule-Based Classifiers</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
<b>Session 3</b>	<b>Naïve Bayes - R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• Naïve Bayes motivation</li> <li>• Naïve Bayes Mathematical foundation</li> <li>• Basics of Probability Theory</li> <li>• Naïve Bayes Assumption</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
<b>Session 4</b>	<b>Team Projects 2nd Presentation (Data Understanding &amp; Data Viz)</b>
	<ul style="list-style-type: none"> <li>• Team Projects (Data Understanding &amp; Data Visualization)</li> <li>• Team Leads' Second Presentation</li> </ul>
<b>Session 5</b>	<b>Introduction Python Packages</b>
	<ul style="list-style-type: none"> <li>• Python: Intro to Five Most common Python Packages <ul style="list-style-type: none"> <li>○ Pandas</li> <li>○ NumPy</li> <li>○ Matplotlib</li> <li>○ Seaborn</li> <li>○ NLTK</li> <li>○ Sckit-learn</li> </ul> </li> </ul>

# BREAK WEEK - 1

3/6/2021

Week 9	Supervised Learning - Classification (Part 2)
<p><b>Learning Objectives:</b> 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 &amp; scikit-learn functionality, introducing automated methods of feature selection, options for estimation including stochastic gradient descent, &amp; advanced metrics for model evaluation.</p>	
Session 1	Logistic Regression in R & Python
	<ul style="list-style-type: none"> <li>• Logistic Regression motivation</li> <li>• Logistic Regression Mathematical foundation</li> <li>• Logistic Regression as a Generalized Linear Model</li> <li>• Characteristics of Logistic Regression</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 2	Support Vector Machine SVM & KSVM in R & Python
	<ul style="list-style-type: none"> <li>• SVM motivation</li> <li>• SVM Mathematical foundation</li> <li>• Margin of a Separating Hyperplane</li> <li>• Linear SVM</li> <li>• Soft-margin SVM</li> <li>• Nonlinear SVM</li> <li>• Characteristics of SVM</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Neighbor Classifiers (KNN) in R & Python - Nonparametric
	<ul style="list-style-type: none"> <li>• KNN motivation</li> <li>• KNN Mathematical foundation</li> <li>• KNN Algorithm</li> <li>• Characteristics of Nearest Neighbor Classifiers</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li>• <b>Azure ML</b> <ul style="list-style-type: none"> <li>○ Introduction Azure ML</li> <li>○ Get Familiar with Azure Machine Learning</li> <li>○ Introduction to Azure Machine Learning</li> <li>○ Introduction to supervised machine learning</li> <li>○ Introduction to Azure Machine Learning</li> <li>○ Azure Machine Learning Algorithms</li> </ul> </li> <li>• <b>Case study # 3</b> <ul style="list-style-type: none"> <li>○ World's GDP Analysis</li> </ul> </li> </ul>



Session 5	SQL for Python
	<ul style="list-style-type: none"> <li>Using SQL for Python, SQL workbench, working with multiple tables</li> <li>Data Extraction</li> </ul>
Week 10	Supervised Learning - Regression (SL & PL)
<p><b>Learning Objectives:</b> 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, time series modeling &amp; finding the causal effect relationship between the variables.</p> <p>Machine learning topics taught involve linear regression algorithms (Simple Linear Regression, Multiple Linear Regression &amp; Polynomial Linear Regression) with deeper use of R &amp; scikit-learn functionality.</p>	
Session 1	Simple Linear Regression R & Python
	<ul style="list-style-type: none"> <li>Preparing Data for Linear Regression</li> <li>Making Predictions with Linear Regression</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Multiple Linear Regression R & Python
	<ul style="list-style-type: none"> <li>Multiple Linear Regression</li> <li>The F-Statistic</li> <li>Interpreting results of Categorical variables</li> <li>Heteroscedasticity</li> <li>Backward Elimination   Backward Elimination   Automatic Backward Elimination</li> </ul>
Session 3	Polynomial Linear Regression R & Python
	<ul style="list-style-type: none"> <li>Polynomial Algorithm</li> <li>Why use polynomial regression?</li> <li>Polynomial Features</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li><b>Tableau:</b> <ul style="list-style-type: none"> <li>Joining, Blending and Relationships   Dual Axis Charts   Joins with Duplicate Values   Joining on Multiple Fields   The Showdown: Joining Data vs. Blending Data in Tableau</li> <li>Creating Calculated Fields in a Blend (Advanced Topic)   Working with Relationships in Tableau</li> </ul> </li> <li><b>Power BI</b> <ul style="list-style-type: none"> <li>Power BI Service   Introduction to Power BI Service</li> <li>Uploading Reports to Power BI Service and Quick Insights   Natural Language Querying</li> <li>Alerts and Notifications in Power BI Service   Personal Gateway</li> </ul> </li> </ul>
Session 5	Python Web scraping (JSON & XML)
	<ul style="list-style-type: none"> <li>Extracting data from the web using JSON, Google API, and XML</li> <li>Data Extraction - Getting Data from the Internet</li> </ul>

Week 11	Supervised Learning - Regression (ML & Regularization)
<p><b>Learning Objectives:</b> This week is a continuation of Regression Analysis techniques learned in week 10. Machine learning topics taught this 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.</p>	
Session 1	Least Square   Lasso   Ridge   Elastic Net Regression R & Python
	<ul style="list-style-type: none"> <li>• OLS Regression- Theory   Implementation</li> <li>• Confidence Interval and OLS Regressions</li> <li>• Ridge Regression   LASSO Regression</li> <li>• Implement ANOVA on OLS Regression</li> <li>• Identify Multicollinearity</li> <li>• Partial Least Square Regression</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 2	Decision Tree Regression R & Python
	<ul style="list-style-type: none"> <li>• A Basic Algorithm to Build a Decision Tree</li> <li>• Methods for Expressing Attribute Test Conditions</li> <li>• Measures for Selecting an Attribute Test Condition</li> <li>• Algorithm for Decision Tree Induction</li> <li>• Characteristics of Decision Tree Classifiers</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Support Vector Machine (SVR) R & Python
	<ul style="list-style-type: none"> <li>• Margin of a Separating Hyperplane</li> <li>• Linear SVM</li> <li>• Soft-margin SVM</li> <li>• Nonlinear SVM</li> <li>• Characteristics of SVM</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li>• <b>Azure ML</b> <ul style="list-style-type: none"> <li>○ Azure Classification Machine Learning</li> <li>○ Introduction to Azure ML Studio</li> <li>○ Explore dataset   Prepare the dataset</li> <li>○ Build the Azure ML experiment   Run the Azure ML experiment</li> <li>○ Visualize the results   Deploy the web service</li> </ul> </li> <li>• <b>RapidMiner:</b> <ul style="list-style-type: none"> <li>○ Applying the Model   Testing a Model   Validating a Model   Finding the Right Model   Optimization of the Model Parameters</li> </ul> </li> </ul>

Session 5	Python: BeautifulSoup Library
	<ul style="list-style-type: none"> <li>Installing BeautifulSoup</li> <li>Data extraction with BeautifulSoup</li> <li>BeautifulSoup Usage   Filtering   BeautifulSoup Object</li> <li>Extracting all the URLs found within a page 'a' tag   Extracting all the text from a page</li> </ul>
Week 12	Ensemble Methods & Imbalanced Data
<p><b>Learning Objectives:</b> Ensemble methods help to improve the predictive performance of Machine Learning models. In this module, you will learn about different Ensemble methods that combine several Machine Learning techniques into one predictive model in order to decrease variance, bias or improve predictions.</p> <p>You will learn about selecting one model over another, and Boosting &amp; its importance in Machine Learning. You will learn how to convert weaker algorithms into stronger ones.</p>	
Session 1	Random Forest Voting & Averaging R & Python
	<ul style="list-style-type: none"> <li>Random Forests</li> <li>Empirical Comparison among Ensemble Methods</li> <li>R &amp; Python hands-on training with real-world business problem</li> </ul>
Session 2	Bagging   Boosting   Gradient Boosting   Ada-Boost
	<ul style="list-style-type: none"> <li>Methods for Constructing an Ensemble Classifier</li> <li>What is the Model Selection?   The need for Model Selection</li> <li>Cross-Validation</li> <li>What is Boosting?</li> <li>Adaptive Boosting</li> <li>How Boosting Algorithms work?</li> <li>Types of Boosting Algorithms</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Class Imbalanced Problem
	<ul style="list-style-type: none"> <li>Building Classifiers - Class Imbalance</li> <li>Data-based Approaches &amp; Algorithmic approach</li> <li>Evaluating Performance - Class Imbalance</li> <li>Finding an Optimal Score Threshold</li> <li>Aggregate Evaluation of Performance</li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 4	Team Projects 3rd Presentation (Analytical Approach)
	<ul style="list-style-type: none"> <li>Team Projects (Analytical Approach)</li> <li>Team Leads' Third Presentation</li> </ul>
Session 5	Python: NumPy Library
	<ul style="list-style-type: none"> <li>Array Creation   Printing Array</li> <li>Changing Shape of an Array   Basic Operation   Universal function   View   Copy   Automatic Reshaping... and more</li> </ul>

**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 module 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 assessments
	<ul style="list-style-type: none"> <li>• CM, ROC, Rank-Ordered Approach</li> <li>• R2, MSE, MAE, Median Error, Median Absolute error, Correlation</li> <li>• Reasons for Model Overfitting</li> <li>• Model Selection   Using a Validation Set</li> <li>• Incorporating Model Complexity   Estimating Statistical Bounds</li> <li>• Model Selection for Decision Trees</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 2	Model Evaluation
	<ul style="list-style-type: none"> <li>• Holdout Method &amp; Cross-Validation</li> <li>• Presence of Hyper-parameters &amp; Hyper-parameter Selection</li> <li>• Nested Cross-Validation</li> <li>• Pitfalls of Model Selection &amp; Evaluation</li> <li>• Overlap between Training &amp; Test Sets</li> <li>• Use of Validation Error as Generalization Error</li> <li>• Cluster Evaluation</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Model Comparison   XG-Boost
	<ul style="list-style-type: none"> <li>• Estimating the Confidence Interval for Accuracy</li> <li>• Comparing the Performance of Two Models</li> <li>• XG-Boost in R &amp; Python</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li>• <b>Tableau:</b> <ul style="list-style-type: none"> <li>○ Table Calculations, Advanced Dashboards   Storytelling</li> <li>○ Mapping: How to Set Geographical Roles   Creating Table Calculations for Gender</li> <li>○ Creating Bins and Distributions for Age   Leveraging the Power of Parameters</li> <li>○ How to Create a Tree Map Chart   Creating a Customer Segmentation Dashboard</li> <li>○ Advanced Dashboard Interactivity  Analyzing the Customer Segmentation Dashboard</li> </ul> </li> <li>• <b>Power BI</b> <ul style="list-style-type: none"> <li>○ Working with Excel in Power BI   Using Excel Data in Power BI   Connecting to Excel Workbook on OneDrive for Business   Pinning Excel Tables or Visuals</li> </ul> </li> </ul>

Session 5	Python: Pandas Library
	<ul style="list-style-type: none"> <li>• Pandas Data Frame Basic</li> <li>• Create Data Frame   Addition and Deletion   Contact and Merge</li> <li>• Export to CSV   Excel   Txt</li> <li>• Group By   Sorting   Stack &amp; Unstack   Iterate Over Data Frame   Pandas Time Series   Reshaping using Pivot   Selection &amp; Indexing</li> </ul>

## Week 14    Unsupervised Learning - Clustering Analysis | Anomalies Detection

**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 module, 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 - R & Python
	<ul style="list-style-type: none"> <li>• The Basic K-means Algorithm</li> <li>• Bisecting K-means</li> <li>• K-means &amp; Different Types of Clusters</li> <li>• K-means as an Optimization Problem</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 2	Agglomerative Hierarchical Clustering - Density Methods – DBSCAN
	<ul style="list-style-type: none"> <li>• Basic Agglomerative Hierarchical Clustering Algorithm</li> <li>• AHC Specific Techniques</li> <li>• The Lance-Williams Formula for Cluster Proximity</li> <li>• Key Issues in Hierarchical Clustering</li> <li>• AHC Outliers</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Anomalies Detection - R & Python
	<ul style="list-style-type: none"> <li>• Characteristics of Anomaly Detection Problems</li> <li>• Characteristics of Anomaly Detection Methods</li> <li>• Spectrum of Anomaly detection techniques</li> <li>• Nearest Neighbor based Methods</li> <li>• Rule-based Anomaly Detection</li> <li>• Classification based Models for Anomaly Detection</li> <li>• Proximity-based Approaches</li> <li>• Strengths &amp; Weaknesses</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li>○ <b>SAS For Enterprise Miner:</b> <ul style="list-style-type: none"> <li>○ Introduction SAS for Enterprise Miner   Loading Data   Data Manipulation   Feature Selections   Data Split</li> </ul> </li> </ul>



	<ul style="list-style-type: none"> <li>• <b>BigML:</b> <ul style="list-style-type: none"> <li>○ Logistic Regression   Times Series   Clusters   Anomaly detection   Association Discoveries</li> </ul> </li> </ul>
<b>Session 5</b>	<b>Python: Matplotlib Library</b>
	<ul style="list-style-type: none"> <li>• Annotation   Autocorrelation   Curves</li> <li>• Scales   Shapes   Box Plot   Violet Plot</li> <li>• Histograms   Pie Chart   Twin Axis   Figures   Color Map</li> </ul>
<b>Week 15</b>	<b>Association Rules   A   B Testing   Recommender Systems</b>
<p><b><u>Learning Objectives:</u></b> This module 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.</p> <p>We also go deep to A   B Testing, also known as split tests, to cover the ins &amp; outs of how to use Python &amp; R to analyze customer behavior &amp; business trends as well as how to create, run, and analyze A   B tests to make proactive, data-driven business decisions.</p> <p>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.</p>	
<b>Session 1</b>	<b>Association Rules - R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• Strength of an association rule</li> <li>• The Apriori &amp; Eclat Principle</li> <li>• Frequent Itemset Generation in the Apriori Algorithm</li> <li>• Candidate Generation &amp; Pruning</li> <li>• Computational Complexity</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
<b>Session 2</b>	<b>A   B Testing R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• Overview of A   B Testing?</li> <li>• How A   B Testing Works</li> <li>• Policy and Ethics</li> <li>• Characterizing Metrics</li> <li>• Designing an Experiment</li> <li>• A   B Testing Process   Analyzing Results   Alternatives to A   B Testing</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
<b>Session 3</b>	<b>Recommender Systems R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• Matrix Factorization Model (Object recommendation)</li> <li>• Content Filtering   Collaborating Filtering (CF)</li> <li>• CF Neighborhood-Based Approach</li> <li>• CF Location-Based Approach</li> <li>• R &amp; Python hands-on training with real-world business problems</li> <li>• </li> </ul>

Session 4	Tableau – Power BI - RapidMiner - Analytics tools - Case Studies
	<ul style="list-style-type: none"> <li>• <b>Tableau:</b> <ul style="list-style-type: none"> <li>○ Advanced Data Preparation   What Format Your Data Should Be In</li> <li>○ Data Interpreter   Pivot   Splitting a Column into Multiple Columns</li> <li>○ Metadata Grid   Fixing Geographical Data Errors in Tableau</li> </ul> </li> <li>• <b>Power BI</b> <ul style="list-style-type: none"> <li>○ Publishing and Sharing   Introduction to Content Pack, Security, and Groups</li> <li>○ Creating a Group   Creating, Using, and Editing Content Pack</li> <li>○ Row-Level Security (RLS)   Integrate and Share OneDrive for Business with Power BI</li> </ul> </li> </ul>
Session 5	Python: Seaborn Library
	<ul style="list-style-type: none"> <li>• Visualizing Distribution to find Patterns</li> <li>• Categorical Data and Multi-Panel data/</li> <li>• Customizing the Facet Grid/Subset using the Paid Grid</li> <li>• Color Palettes/Figure Aesthetics</li> </ul>
Week 16	Time Series Analysis
<p><b><u>Learning Objectives:</u></b> In this module, 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. Time series data means that data is in a series of particular periods or intervals.</p> <p>You will learn different models for time series modeling such that you analyze a real time-dependent data for forecasting.</p>	
Session 1	AR - MA & EST Models - R & Python
	<ul style="list-style-type: none"> <li>• What is Time Series Analysis?</li> <li>• Importance of TSA   Components of TSA</li> <li>• Single Exponential Smoothing</li> <li>• Forecasting with Single Exponential Smoothing</li> <li>• Double Exponential Smoothing</li> <li>• Forecasting with Double Exponential Smoothing</li> <li>• Triple Exponential Smoothing</li> <li>• Example of Triple Exponential Smoothing</li> <li>• Exponential Smoothing Summary</li> </ul>
Session 3	ARMA & ARIMA Models - R & Python
	<ul style="list-style-type: none"> <li>• Forecasting</li> <li>• Relation between time series: Causality &amp; time lags</li> <li>• Distinction between short &amp; long run</li> <li>• Study of agent's expectations</li> <li>• Stationarity</li> <li>• ACF &amp; PACF</li> <li>• Trend removal</li> <li>• Seasonal adjustment</li> <li>• Detection of structural breaks</li> <li>• Control of the process</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>

Session 3	SARIMAX Model - R & Python
	<ul style="list-style-type: none"> <li>The limitations of ARIMA</li> <li>The SARIMA extension of ARIMA</li> <li>Implementing SARIMA method using the Statmodels library</li> </ul>
Session 4	Team Projects 4th Presentation (Analytical Approach Implementation)
	<ul style="list-style-type: none"> <li>Team Projects (Analytical Approach Implementation)</li> <li>Team Leads' Fourth Presentation</li> </ul>
Session 5	Python: SciPy for scientific computing
	<ul style="list-style-type: none"> <li>Introduction to Scipy, building on top of NumPy</li> <li>What are the characteristics of SciPy?</li> <li>Various sub packages for SciPy like Signal, Integrate, Fatback, Cluster, Optimize, Stats and more, Bayes Theorem with SciPy.</li> </ul>

# BREAK WEEK - 2

5/8/2021

Week 17	Text Analysis & Natural Language Processing (NLP)
<p><b>Learning Objectives:</b> In this module, 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.</p>	

Session 1	Text Analysis - R & Python
	<ul style="list-style-type: none"> <li>Basic feature extraction using text data</li> <li>Basic Text Pre-processing of text data <ul style="list-style-type: none"> <li>Lower Casing   Punctuation Removal   Stop Words Removal   Frequent Words Removal   Rare Words Removal   Spelling Correction   Tokenization   Stemming   Lemmatization</li> </ul> </li> <li>Advance Text Processing <ul style="list-style-type: none"> <li>N-grams   Term Frequency   Inverse Document Frequency (TF-IDF)   Bag of Words   Sentiment Analysis   Word Embedding</li> </ul> </li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 2	Sentiments Analysis - R & Python
	<ul style="list-style-type: none"> <li>Motivations</li> <li>Why Sentiment Analysis?</li> <li>Types of Sentiment Analysis <ul style="list-style-type: none"> <li>Polarity Detection   Emotion Detection   Intent Detection</li> </ul> </li> <li>R &amp; Python hands-on training with real-world business problems</li> </ul>
Session 3	Topic Model using LDA - R & Python
	<ul style="list-style-type: none"> <li>Motivations</li> <li>Latent Dirichlet Allocation (LDA)</li> </ul>

	<ul style="list-style-type: none"> <li>• Dirichlet Distribution   Dirichlet Output</li> <li>• LDA &amp; Matrix Factorization</li> <li>• Nonnegative Matrix Factorization</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
<b>Session 4</b>	<b>Tableau – Power BI - RapidMiner - Analytics tools - Case Studies</b>
	<ul style="list-style-type: none"> <li>• <b>Azure ML</b> <ul style="list-style-type: none"> <li>○ Azure Regression Machine Learning</li> <li>○ Predicting housing prices with Azure ML</li> <li>○ Explore dataset   Prepare the dataset   Build the Azure ML experiment</li> <li>○ Run the Azure ML experiment   Visualize the results   Deploy the web service</li> </ul> </li> <li>• <b>BigML:</b> <ul style="list-style-type: none"> <li>○ Supervised vs Unsupervised   Ensembles</li> <li>○ Logistic Regression   Deepnets   Times Series</li> </ul> </li> </ul>
<b>Session 5</b>	<b>Python for Deep Learning: PyTorch Library</b>
	<ul style="list-style-type: none"> <li>• How to Install PyTorch</li> <li>• PyTorch Deep Learning Model Life-Cycle <ul style="list-style-type: none"> <li>○ Prepare the Data   Define the Model   Train the Model   Evaluate the Model   Predictions</li> </ul> </li> <li>• How to Develop PyTorch Deep Learning Models <ul style="list-style-type: none"> <li>○ How to Develop an MLP for Binary Classification</li> <li>○ How to Develop an MLP for Multiclass Classification</li> <li>○ How to Develop an MLP for Regression</li> <li>○ How to Develop a CNN for Image Classification</li> </ul> </li> </ul>
<b>Week 18</b>	<b>Deep Learning: ANN &amp; CNN</b>
<p><b><u>Learning Objectives:</u></b> This week, we dive into Deep Learning - an Artificial Intelligence function that imitates the workings of the human brain in processing data and creating patterns for use in decision making.</p> <p>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.</p>	
<b>Session 1</b>	<b>Introduction to Deep Learning</b>
	<ul style="list-style-type: none"> <li>• Deep Learning Explained</li> <li>• Biological Neural Network</li> <li>• Perceptron – Gradient Decent – Scholastic Gradient Descent</li> </ul>
<b>Session 2</b>	<b>Artificial Neural Network (ANN) - R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• ANN Explained</li> <li>• Multi-layer Neural Network</li> <li>• Characteristics of ANN</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>

<b>Session 4</b>	<b>Convolutional Neural Network (CNN)</b>
	<ul style="list-style-type: none"> <li>• CNN Explained</li> <li>• Synergistic Loss Functions</li> <li>• Responsive Activation Functions</li> <li>• Regularization</li> <li>• Initialization of Model Parameters</li> <li>• Characteristics of Deep Learning</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
<b>Session 4</b>	<b>Tableau – Power BI - RapidMiner - Analytics tools - Case Studies</b>
	<ul style="list-style-type: none"> <li>• <b>Tableau:</b> <ul style="list-style-type: none"> <li>○ Clusters, Custom Territories, Design Features   Custom Territories Via Groups</li> <li>○ Custom Territories Via Geographic Roles   Adding a Highlighter   Clustering in Tableau</li> <li>○ Cross-Database Joins   Modeling with Clusters   Saving Your Clusters</li> </ul> </li> <li>• <b>Power BI</b> <ul style="list-style-type: none"> <li>○ Power BI Mobile</li> <li>○ Introduction to Power BI Mobile</li> <li>○ Mobile Dashboard Layout</li> <li>○ Sharing and Annotating in Power BI Mobile</li> <li>○ Using Custom Charts from Marketplace</li> </ul> </li> </ul>
<b>Session 5</b>	<b>Python for Deep Learning: Keras Library</b>
	<ul style="list-style-type: none"> <li>• Set up your environment.</li> <li>• Install Keras.</li> <li>• Import libraries and modules.</li> <li>• Load image data from MNIST.</li> <li>• Preprocess input data for Keras.</li> <li>• Preprocess class labels for Keras.</li> <li>• Define model architecture.</li> <li>• Compile model.</li> <li>• Fit model on training data.</li> <li>• Evaluate model on test data.</li> </ul>
<b>Week 19</b>	<b>Deep Learning (RNN) for NLP &amp; Time Series</b>
<p><b>Learning Objectives:</b> This module talks about yet another interesting implementation of Neural Networks that revolves around equipping computers to understand human language.</p> <p>You will learn to work with text data and sequential data, and explore the interesting world of RNNs and LSTMs</p>	
<b>Session 1</b>	<b>Recurrent Neural Networks (RNN) - R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• RNN Explained   Characteristics of RNN</li> <li>• Backpropagation</li> <li>• Gradient problem (Vanishing vs. Exploding)</li> <li>• Long Short-Term Memory Networks</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>



<b>Session 2</b>	<b>Recurrent Neural Networks for Time Series- R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• Transform Data for Time Series</li> <li>• MLP &amp; CNN for Time Series Forecasting</li> <li>• LSTM &amp; CNN-LSTM for Time Series Forecasting</li> <li>• Encoder-Decoder LSTM Multi-step Forecasting</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
<b>Session 2</b>	<b>Recurrent Neural Networks for NLP- R &amp; Python</b>
	<ul style="list-style-type: none"> <li>• Motivations</li> <li>• From logistic regression to neural networks</li> <li>• Word representations</li> <li>• Unsupervised word vector learning</li> <li>• Backpropagation Training</li> <li>• Learning word-level classifiers: POS &amp; NER</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>
<b>Session 4</b>	<b>Tableau – Power BI - RapidMiner - Analytics tools - Case Studies</b>
	<ul style="list-style-type: none"> <li>• <b>Azure ML:</b> <ul style="list-style-type: none"> <li>○ Deploy a Web Application with Azure Machine Learning</li> <li>○ Deploy a web application with Azure ML</li> <li>○ Introduction to Azure cloud</li> <li>○ Deploy &amp; Review Azure ML web application</li> <li>○ Test drive Azure ML web application</li> </ul> </li> <li>• <b>Case study # 4</b> <ul style="list-style-type: none"> <li>○ Crime Data Analysis</li> </ul> </li> </ul>
<b>Session 5</b>	<b>Python for Deep Learning: TensorFlow Library</b>
	<ul style="list-style-type: none"> <li>• Overview of TensorFlow and TensorFlow libraries</li> <li>• Use cases for a machine learning service</li> <li>• Using and applying your model</li> <li>• Training your model</li> <li>• Testing your model</li> <li>• Using TensorBoard to visualize model performance</li> </ul>
<b>Week 20</b>	<b>General Deployment Considerations &amp; intro to Big Data</b>
<b>Learning Objectives:</b> In this module, 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.	
<b>Session 1</b>	<b>Avoiding False Discoveries</b>
	<ul style="list-style-type: none"> <li>• Preliminaries: Statistical Testing</li> <li>• Modeling Null &amp; Alternative Distributions</li> <li>• Statistical Testing for Classification   Association Analysis   Cluster Analysis   Anomaly Detection</li> <li>• R &amp; Python hands-on training with real-world business problems</li> </ul>

Session 2	General Deployment Considerations
	<ul style="list-style-type: none"> <li>• Docker Basics   Flask Basics</li> <li>• Deployment Steps &amp; Challenge</li> <li>○ Load your model with its weights - Preprocess your data - Perform the actual prediction - Handle the prediction response data <ul style="list-style-type: none"> <li>• Loading model directly into an application</li> <li>• Calling an API</li> </ul> </li> <li>○ Kubernetes - Custom REST-API with Flask   Django - AWS Lambda   Serverless <ul style="list-style-type: none"> <li>• Apache Beam</li> </ul> </li> <li>○ Spark   Flink</li> </ul>
Session 3	Introduction to Big Data
	<ul style="list-style-type: none"> <li>• The job market around Big Data</li> <li>• What is Big Data?</li> <li>• Big Data use cases</li> <li>• Big Data Ecosystem</li> <li>• HDFS   What is MapReduce   Pig, Hive etc.</li> <li>• NoSQL Databases</li> </ul>
Session 4	Team Projects Final Presentation (Insights & Action Plan)
	<ul style="list-style-type: none"> <li>• Team Projects (Analytical Approach Implementation)</li> <li>• Team Leads' Final Presentation</li> </ul>
Session 5	Team Projects Codes share
	<ul style="list-style-type: none"> <li>• Share and Compare Team Project codes with the Business solutions</li> <li>• Lesson learned</li> </ul>
Week 21-25	Individual Capstone Project (1 Month)
<p><b>Learning Objectives:</b> 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.</p> <p><b>Duration:</b> The capstone will run for ONE month starting from the 12th of June 2021 till the 19th of July 2021.</p> <p><b>Grading:</b> 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”.</p>	



## ATC Innovation Center

2972 Webb Bridge Rd, Alpharetta, GA 30009

[www.bbds.ma](http://www.bbds.ma) || [register@bbds.ma](mailto:register@bbds.ma) || 315-975-1661

# Career Services

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

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

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

## Week 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:

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

## Week 28

## Resume Preparation

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

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

## Week 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

## Week 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 algorithms (Non-graph)
- Different types of NoSQL databases
- Pig and Hive
- Spark

# BBDS Advantage



We are Big Bang Data Science Solutions (BBDS), a professional Data Science Solutions company based in Atlanta GA. We are specializing in Data Science, Machine Learning and advanced analytics technologies, including software engineers/programming, for the past 5 years.

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.

## Key Facts about Artificial Intelligence and Machine Learning:

- **Between 2015 and 2018, the number of job postings with "AI" or "Machine Learning" increased by nearly 100%.**

*Indeed, 2018*

- **AI was among the top 5 in-demand hard skills in 2019.**

*LinkedIn, 2019*

- **The global machine learning market is expected to grow from USD 1.41 billion in 2017 to USD 8.81 billion by 2022.**

*Research and Markets Report, 2017*

- **86% of executives at fast-growing companies say AI is important to their company's success.**

*Cognizant Report, 2018*

**Showcase Your Competence  
with a BBDS Certificate**





# Ready to Advance Your Career?

**APPLY NOW**

## Speak to 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**