

Program with 3 International Accreditations



# Certificate in DATA SCIENCE & AI

From Zero to Hero 40-Week Training Program

LIVE ONLINE  
TRAINING PROGRAM

**BATCH 22**

**40 WEEKS**

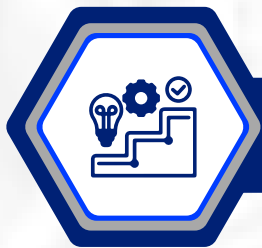
**22<sup>ND</sup> | FEBRUARY  
2025**

+1 (678) 209 9780

[www.innovatics.ai](http://www.innovatics.ai)

[register@innovatics.ai](mailto:register@innovatics.ai)





## PROGRA OVERVIEW

To maintain a competitive edge, organizations today must acquire practical expertise in Data Science, Machine Learning, Deep Learning, NLP, Gen-AI, and Artificial Intelligence. Careers in fields such as engineering, data science, and technical management increasingly rely on the advanced analytical and programming skills that these areas provide.

The Data Science & AI "Zero to Hero" program, an executive education course led by InnovatiCS, offers an immersive learning experience in the core methods of Data Science, Machine Learning, Deep Learning, NLP, Gen-AI, and Artificial Intelligence. This program emphasizes the hands-on skills needed to work effectively with machine learning models, including classification, regression, clustering, and optimization techniques.

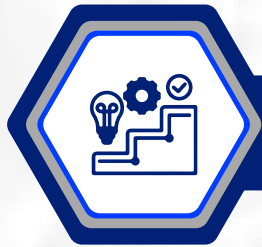
Throughout the course, you will apply these methods to real-world problems, gaining the practical expertise needed to thrive in machine learning-focused roles. What sets our program apart is its practical, applied approach, addressing both the "how" and the "why" of machine learning, and equipping you with the technical skills to succeed in any field that uses these technologies.

### You will leave this program with.....

- Key insights for leveraging analytics to drive growth, enhance efficiency, and boost productivity.
- The language and intuition needed to collaborate effectively with data scientists.
- The tools and techniques required to apply analytics practically and solve specific business challenges.

## Benefits of Developing a Data-Driven Organization





# LEARNING JOURNEY

## Data Science Foundations

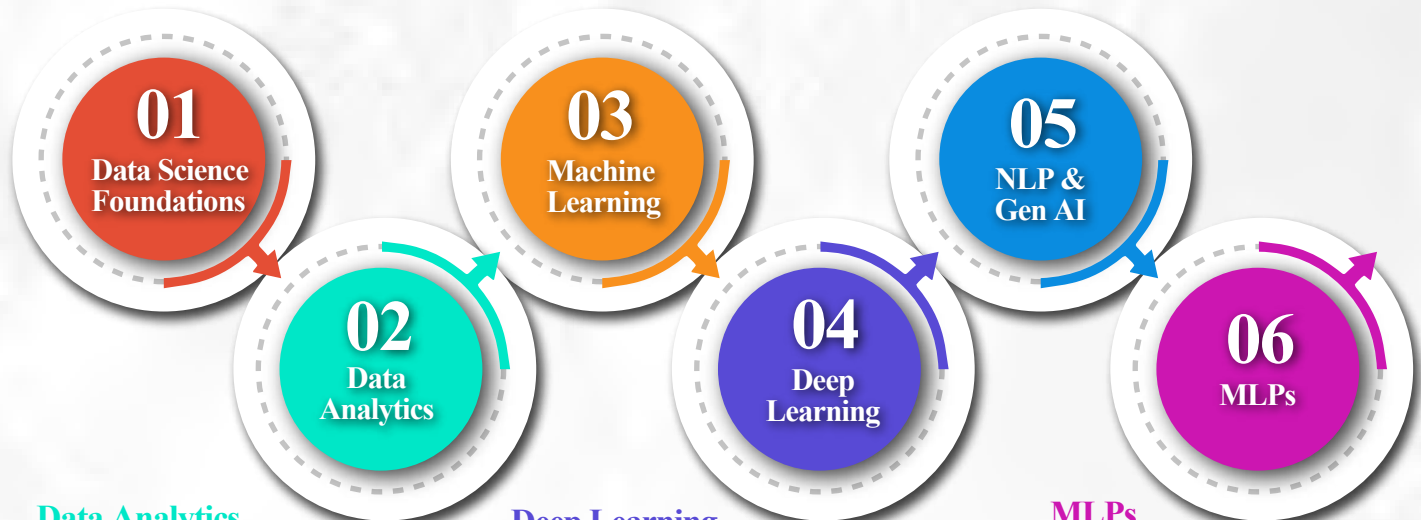
- Programming Foundation
- Statistics Foundation
- Business Understanding
- Decision Management

## Machine Learning

- Classification | Regression
- Clustering | A/B Testing | Association Rules
- ML Explainability & Interpretability
- Time Series Analysis | Recommended Sys.

## NLP & Gen AI

- NLP
- Gen AI –Chat GPT
- Prompt Engineering
- Cloud Computing



## Data Analytics

- Data Understanding
- Data Visualization
- Data Preparation & Processing
- Data Transformation
- Factor Analysis

## Deep Learning

- Advanced Python for Deep Learning
- ANN –CNN –RNN – GANs
- Self-Organizing Maps –Boltzmann Machine
- Autoencoder –Sequence Learning

## MLPs

- Deployment -Flask
- Capstone Project
- Soft-skills –Resume Preparation
- Interview Preparation
- Job Placement assistance



# TOOLS IN THE PROGRAM





# DATA SCIENCE AND AI PROGRAM



01

## InnovatiCS Certified/Verified Certificate

A certified and verified Certificate in Data Science and AI from InnovatiCS upon completing the program, participants will also obtain and will be trained to get the following 4 Certificates



02

## CPD-UK

350 Hours of Accreditation from Continuing Professional Development (CPD-UK)



03

## IABAC

Certificate in Data Science and AI from International Association of Business Analytics Certifications (IABAC)  
Candidate must take an exam and pay a fee



04

## HLACT

500 Hours of Accreditation from Higher Learning Accreditation Consultant and Training (HLACT) with a Certificate . Candidate must take an exam and pay a fee



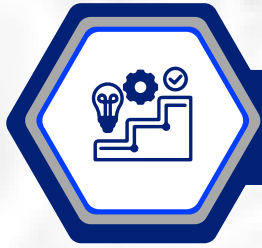
05

## CAP

Certified Analytics Professional (CAP) from INFORMS.  
Candidate must qualify, take an exam, and pay a fee







## PROGRAM IS FOR...

InnovatiCS' 40-Week Training Program in Data Science & AI is designed for individuals looking to launch a career in Data Science and AI or advance their existing technology careers, with extensive support provided throughout the program.

Graduates will be equipped with the skills to tackle real-world challenges in Data Science and AI across various industries, positioning them for new career opportunities.

Additionally, the program helps participants build a robust Data Science portfolio, enhance their online presence, and effectively showcase their skills to prospective employers



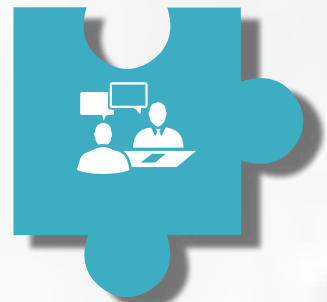
Engineers / Associates / IT Professionals: Software engineers in IT/ITES Startup teams building ML products/services



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



## PROGRAM SCHEDULE

**3 TIMES A WEEK**

**THR (Optional)**

**9:00  
AM EST**

**8:00  
PM EST**

**40  
WEEKS**

**10-15 HOURS  
PER WEEK**

**SAT & SUN**

**9:00  
AM EST**

**12:00  
PM EST**



## PROGRAM INSTRUCTORS



### Mo Medwani Ph. D.

#### Adjunct - Chief Data Scientist

The founder of InnovatiCS, Dr. Mo Medwani, holds a PhD in Artificial Intelligence and is a highly skilled data scientist with a passion for transforming data into actionable products. With over 22 years of experience in service delivery management, he also holds four master's degrees in Data Science, IT, Machine Learning, and Business Administration. Dr. Medwani brings over 11 years of expertise in AI specializing in Data Science, ML, DL, Gen-AI, NLP, and IT service delivery management



### Muhammed Mujeeb

#### Data Engineering

BS in Computer Science and a Master of Business Administration (MBA). Microsoft Microsoft Certified SQL Server Specialist. Azure Cloud Fundamentals Certified. Computer Programming and Database Development Over 24 years of experience in Software Development, Database Development, and Cyber Security, and a strong background in all varieties of Application Development and Application Security and 10 years of training experience Currently working as Manager, Security Business Intelligence at LexisNexis.



### Muhammed Rafique

#### MLOPs

Holding an M. Tech and an MS, Mohammed Rafique has amassed vast experience in cutting-edge technologies throughout his distinguished career. He has architected comprehensive end-to-end large enterprise platforms for some of the most prominent Fortune 50 clients, demonstrating his exceptional ability to handle complex technological challenges. His expertise spans a wide array of modern technologies, including Big Data, Cloud computing, DevOps practices, Artificial Intelligence.



### Azhar Usmani Ph.D.

#### Cloud Computing

Seasoned technology leader, holds a PhD, and is an AWS expert with two decades of experience. He excels in software development, architecture, and security, emphasizing simplicity and automation in system design. Azhar's achievements include migrating legacy platforms to cloud-native infrastructure and promoting public cloud adoption. His skill set covers programming languages, Kubernetes, and AWS, while his commitment to learning and mentoring remains unwavering. Azhar's career showcases his adaptability, innovation, and infectious passion.



### Talha Saleem Ahmed

#### Python | AI

Professor of Computer Science at Gloucestershire College - UK with MSc in Applied Science and MSc in Biotechnology. Professionally adept in Data Science, Artificial Intelligence, Machine Learning, Web Technology, Software Engineering, Biostatistics, Bioinformatics, Business Development, programming, and computing. Skilled in data analysis, software application development, predictive modeling, data processing, advanced data mining algorithms, strategic enterprise architecture, and proficient in various programming languages.



## PROGRAM ADVISORS



### Salwa Smaoui

#### Digital Transformation Leader

Salwa Smaoui is a digital transformation leader with over three decades of experience in the world of technology, marketing, and business development, with most of her career spent at some of the world's finest global companies. A successful 15-year career at Microsoft saw Salwa lead many key initiatives starting in her native country Tunisia leading the country to unprecedented business growth through building and executing an impactful capacity-building, community reach and citizenship plan. Further a field, Salwa was then commissioned to lead the Microsoft Western Europe Windows and Surface business, with responsibility to 14 Western European countries, after which Salwa became the Government and Smart City Industry leader at Microsoft Middle East and Africa. A role that propelled her to lead revenue growth, digital transformation, and smart city strategy for governments across 79 countries. This opened an opportunity to join Core42, a G42 Group Company in Abu-Dhabi, as Chief Commercial and Marketing Officer. She is world's most dynamic and inspirational digital transformation strategies. Salwa launched Baobab Campus an Education Platform that curates and validates content for AI skills and Mental health.



### Dr. Amnah Alazzah

#### Influential Technology and Leadership Expert in the Middle East

Dr. Amnah Alazzah is a distinguished trainer and coach specializing in technology and leadership, with over 15 years of experience in the Middle East. She is renowned for her ability to deliver impactful training programs to both international and local audiences. As the founder of a specialized firm providing cutting-edge technology solutions and training, Dr. Alazzah is a pivotal figure in the industry, frequently speaking at regional and international conferences and serving as a panelist. Dr. Alazzah's educational background is extensive and impressive. She holds a bachelor's degree in computer science, a master's in media studies, and an MBA. Additionally, she has completed numerous courses in business administration, project management, leadership, and human resources. This diverse education underpins her expertise and enables her to deliver comprehensive training programs. Throughout her career, Dr. Alazzah has been a staunch advocate for integrating technology into education. She has trained university faculties and professionals, helping them transition from traditional methods to modern, technology-enhanced learning approaches. Her efforts have significantly impacted the adoption of advanced tools in educational settings, promoting a more effective and engaging learning environment.



### Nadin Ouahid

#### Influential Marketing Strategist –North Africa

Nadin Ouahid holds a prestigious Master of Science degree in Marketing and Communication, which has provided her with a solid foundation in understanding market dynamics and effective communication strategies. With a robust background in real estate, Nadin has excelled as a sales and marketing representative, where she specialized in client relations and market analysis. Her keen ability to understand client needs and analyze market trends has significantly contributed to her success in the industry. Currently, Nadin serves as an operations associate for a federal contractor company, where she focuses on enhancing operational efficiency and streamlining processes. In this role, she applies her expertise in marketing and communication to improve internal workflows and ensure that the company meets its operational goals. Her innovative approach to problem-solving and process optimization has led to notable improvements in the company's operational performance. Nadin's diverse experience in both real estate and federal contracting highlights her versatility and adaptability in various professional environments. Her commitment to excellence and continuous improvement is evident in her work, making her a valuable asset to any organization.



### Antony Marshal Pereira

#### Pioneering Leader and Innovator in Digital Transformation

Antony Marshal Pereira, Founder & Managing Director of Digirabia and Chief Digital Officer at InnovatiCS US, is a distinguished leader with over two decades of experience in digital innovation. He has an extensive background in the telecom sector, having worked with leading brands such as BPL Mobile, Hutchinson Telecom, Vodafone Idea (VI), and Ooredoo in India & Kuwait. Since launching Digirabia in Kuwait in 2016, his visionary approach was recognized in 2022 when he received the "Internet 2.0 - Outstanding Leadership Award" in Dubai, honoring his exceptional contributions to the field of technology. Marshal has also taken on a pivotal role at InnovatiCS US as Chief Digital Officer, where he is tasked with revamping the company's entire digital strategy. He is spearheading a complete overhaul of the company's online presence, enhancing digital sales and marketing performance with a focus on data science, AI, cloud computing, and cybersecurity training. In addition to leading Digirabia, he is dedicated to helping businesses increase their customer base, sales, and revenue through strategic consulting and coaching, covering areas such as digital marketing, ad tech, mar tech, and data-driven AI marketing.

# PROGRAM MODULES

## DATA SCIENCE FOUNDATIONS



## 01 MODULE

### Week 1

### Program Orientation

**Learning Objectives:** The program kicks off with a lecture titled 'Why Learning Data Science & AI is an Absolute Must!' In this session, you'll be introduced to the data science process, explore the key components of a data science portfolio, and gain insights into the various types of analytics. You'll also learn about the daily responsibilities of a data scientist, essential quantitative and statistical techniques, and discover the career path to help you start your journey in data science.

#### Session 1

#### Program Orientation

- Program Orientation (Agenda – Curriculum).
- InnovatiCS Website | Canvas | Slack walkthrough.

#### Session 2

#### Why should you become a Data Scientist?

- Data Explosion | Why Data Science? | What is Data Science? | Type of Analytics.
- Data Science Portfolio | Data Science Process | Career in Data Science.

#### Session 3

#### Introduction to Data Science & AI

- Introduction to Data Science.
- Introduction to Data Analytics.
- Introduction to Business Intelligence.
- Certified Analytics Professional (CAP).

#### Session 4

#### Projects, Teams & Team Leads

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

#### Session 5

#### Program Tools & Installation

- **Jupyter Notebook** | **MySQL**: Installation | **Tableau**: Installation | **RapidMiner**: Installation

### Week 2

### Business Understanding (Strategy | Literacy | Problem Framing)

**Learning Objectives:** The basic workflow is now in place. You will dive into the first stage of the Cross-Industry Standard Process for Data Mining (CRISP-DM), which focuses on 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 analytical projects. Neglecting this step can result in a great deal of effort being put into producing the right answers to the wrong questions. This section will also provide an overview of the importance of data strategy as a tool for enabling organizations to make better-informed decisions, improve business processes, and gain new revenue streams. Data literacy involves articulating a problem that can potentially be solved using data.

#### Session 1

#### Introduction to Artificial Intelligence (ML)

- Introduction to Machine Learning from different perspectives.
- Understand the Need for Machine Learning & Learning major elements.
- Understand Machine Learning tasks & types of Machine Learning.
- Understand Types of Algorithms, Algorithms Performance & types of Analysis.



Session 2 Introduction to Artificial Intelligence (DL & GenAI)	
	<ul style="list-style-type: none"> <li>• Deep Learning Explained.</li> <li>• Biological Neural Network.</li> <li>• Understand Gen AI &amp; LLMs.</li> </ul>
Session 3 Introduction to 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.</li> <li>• Business Understanding   Data Understanding   Data Preparation   Modeling  Evaluation   Deployment.</li> </ul>
Session 4 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, and 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 5 Introduction to Big Data (self-paced)	
	<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 &amp; Big Data Ecosystem.</li> <li>• HDFS   What is MapReduce   Pig, Hive etc.</li> <li>• NoSQL Databases.</li> </ul>
Week 3 Python for Data Science & ML (Self-paced)	
<p><b>Learning Objectives:</b> This week, students will be introduced to the Python programming language, which is commonly used in data science. Students will work with the tools that professional data scientists use, such as Jupyter Notebooks, integrated development environments (IDEs), and others. They will learn about the purpose of each tool, the types of code they can execute, and their features and limitations.</p>	
Session 1 Basic Python Programming	
	<ul style="list-style-type: none"> <li>• Introduction to Python &amp; Jupyter Notebook</li> <li>• Data Structure: Sting   Lists   Tuples   Sets   Dictionaries.</li> <li>• Conditional Statements   Python built-in data types   Basic operators in Python.</li> <li>• Loop and control statements like a break, if, for, continue, else, range (), and more.</li> <li>• Functions, Assignment &amp; Operations.</li> <li>• Conditions &amp; Loops.</li> </ul>
Session 2 Python: NumPy Library	
	<ul style="list-style-type: none"> <li>• Introducing arrays.</li> <li>• Your first NumPy array   Creating arrays from scratch.</li> <li>• A range array   Array dimensionality.</li> <li>• 3D array creation   The fourth dimension   Flattening and reshaping.</li> <li>• NumPy data types   The dtype argument   Anticipating data types.</li> </ul>

Session 3		Python: Pandas Library
		<ul style="list-style-type: none"> <li>• Introducing DataFrames   Inspecting a DataFrame.</li> <li>• Parts of a DataFrame   Sorting and sub-setting.</li> <li>• Sorting rows   Sub-setting columns   Sub-setting rows.</li> <li>• Sub-setting rows by categorical variables   New columns   Adding new columns.</li> </ul>
Session 4		Python: Matplotlib
		<ul style="list-style-type: none"> <li>• Introduction to data visualization with Matplotlib.</li> <li>• Using the matplotlib.pyplot interface.</li> <li>• Adding data to an Axes object   Customizing your plots.</li> <li>• Customizing data appearance   Customizing axis labels and adding titles.</li> <li>• Small multiples   Creating a grid of subplots.</li> <li>• Creating small multiples with plt.subplots   Small multiples with shared y-axis.</li> </ul>
Session 5		Python: Regex
		<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> <li>• Introduction to string manipulation   Palindromes   String operations.</li> <li>• Normalizing reviews   Split lines or split the line?</li> <li>• Finding and replacing   Finding a substring   Replacing negations.</li> </ul>
Week 4		Statistics - Probability – SQL for Data Science
<p><b>Learning Objectives:</b> This week, students will build a solid foundation in statistics for data science, learning about probability, distributions, hypothesis testing, and SQL. Statistics is the science of assigning a probability to an event based on experiments and the application of quantitative principles to the collection, analysis, and presentation of numerical data. This course will help students master the fundamentals of data science, statistics, and machine learning. Students will learn to define statistics and essential related terms, explain measures of central tendency and dispersion, and comprehend skewness, correlation, regression, and distribution. They will also be able to make data-driven predictions through statistical inference. Additionally, students will learn about creating and working with tables, databases, Data Definition Language (DDL), and Data Manipulation Language (DML), including SELECT, INSERT, UPDATE, and DELETE statements in SQL.</p>		
Session 1		Introduction to Statistics
		<ul style="list-style-type: none"> <li>• Understanding Data.</li> <li>• Levels of Measurement   Measures of Dispersion.</li> <li>• Population and Sample.</li> <li>• Measures of Central Tendency   Quartiles and IQR.</li> </ul>
Session 2		Advanced Statistics / Probability & Information Theory
		<ul style="list-style-type: none"> <li>• Introduction to Distribution.</li> <li>• Uniform Distribution   Binomial Distribution   Poisson Distribution   Normal Distribution.</li> <li>• Skewness   Standardization and Z Score.</li> <li>• Central Limit Theorem   Hypothesis Testing.</li> <li>• ANOVA - Analysis of Variance   Chi-Square Analysis.</li> </ul>
Session 3		Crash Course in SQL (Tables and Constraints) (Self-paced)
		<ul style="list-style-type: none"> <li>• Creating, and Working with Tables.</li> <li>• Default Constraints.</li> <li>• Cascading referential integrity.</li> <li>• Check Constraint   Identity column   Unique key constraint.</li> <li>• Get the last generated id in the SQL server.</li> </ul>

Session 4	SQL Statements Group by Clause (Self-paced)
	<ul style="list-style-type: none"> <li>• SQL Statements (Select, Insert, Delete, Update)   Group by Clause.</li> <li>• Difference between where and having in SQL server.</li> <li>• Basic Joins   Advance Joins   Self Joins.</li> <li>• Different ways to replace NULL   Coalesce function.</li> <li>• Union and Union All.</li> </ul>
Session 5	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>

# DATA MANIPULATION & ANALYTICS



## 02 MODULE

Week 5	Data Understanding   EDA & Statistical Analysis
<p><b>Learning Objectives:</b> In this week, students will be introduced to the second phase of the Cross-Industry Standard Process for Data Mining (CRISP-DM) model. Students will obtain data and verify its appropriateness for their needs. They will learn to identify issues that may require returning to the 'Business Understanding' phase of the project to revise the plan. Students may even discover flaws in the 'Business Understanding,' prompting a reconsideration of goals and plans. The Data Understanding phase includes four tasks: gathering data, describing data, exploring data, and verifying data quality.</p>	
Session 1	Introduction to Data Literacy and Biases
	<ul style="list-style-type: none"> <li>• Understanding Bias in Data. <ul style="list-style-type: none"> <li>◦ Types of Data Bias (e.g., Sampling Bias, Selection Bias, Confirmation Bias).</li> <li>◦ Identifying and Mitigating Bias in Data Collection.</li> <li>◦ Recognizing Bias in Data Analysis and Interpretation.</li> <li>◦ Impact of Bias on Decision Making.</li> </ul> </li> <li>• What Exactly is Data Literacy? <ul style="list-style-type: none"> <li>◦ Why We Need Data Literacy.</li> <li>◦ Data-driven Decision Making.</li> </ul> </li> <li>• Benefits of Data Literacy and How to Get Started.</li> </ul>
Session 2	Data Loading & Manipulation
	<ul style="list-style-type: none"> <li>• Read data from different sources.</li> <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>
Session 3	Data Types- Measure of Shape - Position - Dispersion
	<ul style="list-style-type: none"> <li>• Measures of the Spread: Range – IQR – Variance – Standard Deviation <ul style="list-style-type: none"> <li>◦ Measures of Dispersion   Measures of Position   Measures of Relationships   Measures of Shapes</li> </ul> </li> </ul>
Session 4	Data Visualization (Numerical Data)
	<ul style="list-style-type: none"> <li>• Data Visualizatio <ul style="list-style-type: none"> <li>◦ Histograms</li> <li>◦ Bar Plot</li> <li>◦ Scatter Plot</li> <li>◦ Box Plot</li> </ul> </li> </ul>

Session 5 Data Visualization (Graphical Descriptive Statistics)	
	<ul style="list-style-type: none"> <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> <li>◦ Detecting missing values &amp; outliers   Duplicates &amp; redundant records</li> </ul> </li> </ul>
Week 6 Data Science Foundations (Visual Data Sensemaking)	
<p><b>Learning Objectives:</b> Before you can present information to others, you must understand its story. In this week, students will learn the concepts, principles, and practices of visual data sense-making. The skills taught rely primarily on vision, enabling students to interactively use graphs to find and examine meaningful patterns and relationships in quantitative data.</p>	
Session 1 Building a Framework for Visual Data Sensemaking	
	<ul style="list-style-type: none"> <li>• History of Data Visualization</li> <li>• Prerequisites for Data Sensemaking</li> <li>• Thinking with our eyes</li> <li>• Visual Data Sensemaking Interaction and Navigation</li> </ul>
Session 2 Data fitness, Variation, and Relationships (Categorical Variables)	
	<ul style="list-style-type: none"> <li>• Variation within Categorical Variables</li> <li>• Relationships among Categorical Variables</li> <li>• Variation across space   Variation through time</li> </ul>
Session 3 Data fitness, Variation, and Relationships (Quantitative Variables)	
	<ul style="list-style-type: none"> <li>• Variation within Quantitative Variables</li> <li>• Relationships among Quantitative Variables</li> </ul>
Session 4 Data fitness, Variation, and Relationships (Multiple Variables)	
	<ul style="list-style-type: none"> <li>• Relationships among multiple Variables and Perspectives</li> <li>• Multi-Perspective View</li> <li>• Multivariate Relationships</li> </ul>
Session 5 Interactive Dashboards (Tableau)	
	<ul style="list-style-type: none"> <li>• Tableau Interactive dashboards</li> </ul>
Week 7 Data Preparation & Preprocessing	
<p><b>Learning Objectives:</b> In the Data Preparation phase, students learn the process of cleaning and transforming raw data before processing and analysis. This important step often involves reformatting data, making corrections, and combining data sets to enrich the data. Students will also learn how to fix data quality issues discovered through exploratory data analysis (EDA) and 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: Replacing values   Creating Dummy Variables   Encoding labels   One -Hot encoding   Binary encoding   Backward difference encoding   Miscellaneous Features</li> <li>• 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.</li> </ul>
Session 2 Data Preparation & Processing (Numeric Attributes & Transformation)	
	<ul style="list-style-type: none"> <li>• Statistical Analysis (John T. 5 Numbers Summary)</li> <li>• Variance &amp; Covariance</li> <li>• Correlation Analysis</li> </ul>

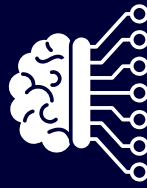


Session 3 Data Preparation & Processing (Missing Values)	
	<ul style="list-style-type: none"> <li>• Dealing with Skewness &amp; Kurtosis</li> <li>• Missing Values</li> <li>• Variable Conversion: Discretization   Binning</li> </ul>
Session 4 Data Preparation & Processing (Outliers   Duplicates)	
	<ul style="list-style-type: none"> <li>• Dealing with Outliers &amp; Duplicates &amp; redundant</li> <li>• Variable Conversion: Discretization   Binning</li> <li>• Variable Transformation: Normalization   Standardization   Factorizations   Binarization</li> <li>• Box-Cox Transformation</li> <li>• Data Partitioning</li> </ul>
Session 5 Data Prep with RapidMiner	
	<ul style="list-style-type: none"> <li>• <b>RapidMiner:</b> <ul style="list-style-type: none"> <li>◦ Turbo Prep – Introduction   Data Cleansing   Merging Data</li> <li>◦ Data Pivoting   Data Preparation   Connecting to Databases</li> </ul> </li> <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> </ul>
Week 8 Factor Analysis & Dimensionality Reduction	
<b>Learning Objectives:</b> Data preprocessing includes imputing missing values, handling outliers, removing duplicates, addressing redundant data, analyzing skewness and kurtosis, examining distributions, assessing correlation, performing feature selection and extraction, conducting generic wrangling and data manipulation, scaling data, factorization, binarization, applying transformations (such as Box-Cox), normalization, and many more techniques.	
Session 1 Principal Component Analysis	
	<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   Machine Learning Method   Feature Importance</li> </ul> </li> <li>• Regularization &amp; Embedded Feature Selection   Feature Subset Selection   Feature Creation</li> <li>• Factor Analysis: PCA   KPCA. • Aggregation   Sampling   Dimensionality Reduction</li> <li>• Feature Subset Selection   Feature Creation</li> <li>• Discretization &amp; Binarization   Variable Transformation   Data Partitioning</li> </ul>
Session 2 Singular Value Decomposition & Linear Discriminant Analysis	
	<ul style="list-style-type: none"> <li>• Factor Analysis: LDA &amp; SVD</li> <li>• Aggregation   Sampling   Dimensionality Reduction</li> </ul>
Session 3 ISOMAP   Locally Linear   Modified Locally Linear Embedding	
	<ul style="list-style-type: none"> <li>• Introduction to Manifold Learning (t-SNE, LLE, Isomap)</li> <li>• Local Linear Embedding (LLE)               <ul style="list-style-type: none"> <li>◦ Intuition   Least squares problem</li> </ul> </li> <li>• Eigenvalue problem</li> </ul>
Session 4 Data Prep with RapidMiner	
	<ul style="list-style-type: none"> <li>• <b>RapidMiner:</b> <ul style="list-style-type: none"> <li>◦ Data Cleansing   Merging Data</li> <li>◦ Data Pivoting   Data Preparation   Connecting to Databases</li> </ul> </li> <li>• <b>Table au:</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> </ul>

**Session 5****Team Projects 2nd Presentation (Data Understanding & Data Viz)**

- Team Projects (Data Understanding & Data Visualization)
- Team Leads' Second Presentation

# MACHINE LEARNING



## 03 MODULE

**Week 9****Supervised Learning - Classification**

**Learning Objectives:** This week broadens the concepts learned in Data Understanding, Data Preparation, and Exploratory Data Analysis (EDA) by extending into Machine Learning. You will learn about the models and methods used in machine learning and apply them to a real-world scenario. Supervised machine learning aims to build a model that makes predictions based on evidence in the presence of uncertainty. In this session, you will learn about different algorithms of supervised learning, such as Decision Trees, Rule-Based Classifiers, Naive Bayes, Logistic Regression, Support Vector Machines, and Nearest Neighbor Classifiers. You will also explore deeper use of sci-kit-learn functionality, automated methods of feature selection, options for estimation including stochastic gradient descent, and advanced metrics for model evaluation.

**Session 1****Decision Tree & Rule-Based Classifiers**

- Algorithm for Decision Tree Induction
- Measures for Selecting an Attribute Test Condition
- Characteristics of Decision Tree Classifiers
- How a Rule-Based Classifier Works
- Characteristics of Rule-Based Classifiers

**Session 2****Naïve Bayes Classifiers**

- Naïve Bayes motivation
- Naïve Bayes Mathematical foundation
- Basics of Probability Theory
- Naïve Bayes Assumption

**Session 3****Logistic Regression Classifiers**

- Logistic Regression motivation | Mathematical foundation
- Logistic Regression as a Generalized Linear Model
- Characteristics of Logistic Regression

**Session 4****Support Vector Machine SVM & KSVM Classifiers**

- SVM motivation | Characteristics of SVM
- SVM Mathematical foundation
- Margin of a Separating Hyperplane
- Linear SVM | Nonlinear SVM | Soft-margin SVM

**Session 5****Neighbor Classifiers (KNN)**

- KNN motivation
- KNN Mathematical Foundation
- KNN Algorithm
- Characteristics of Nearest Neighbor Classifiers

Week 10	Supervised Learning - Regression
<b>Learning Objectives:</b> This week focuses on Regression Analysis, a form of supervised learning with a continuous target variable. Regression analysis is a predictive modeling technique that investigates the relationship between a dependent (target) variable and independent (predictor) variables. This technique is used for forecasting, time series modeling, and identifying causal relationships between variables. The machine learning topics covered include linear regression algorithms (Simple Linear Regression, Multiple Linear Regression, Polynomial Linear Regression), Decision Trees, Support Vector Machines, and regularization techniques (Lasso, Ridge, Elastic Net, and Least Squares) with in-depth use of scikit-learn functionality.	
Session 1	Simple Linear Regression & Polynomial Regression
	<ul style="list-style-type: none"> <li>• Preparing Data for Linear Regression</li> <li>• Making Predictions with Linear Regression</li> <li>• Polynomial Algorithm</li> <li>• Why use polynomial regression?</li> <li>• Polynomial Features</li> </ul>
Session 2	Multiple Linear Regression
	<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	Least Square/Lasso/Ridge/Elastic Net Regression
	<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   Partial Least Square Regression</li> </ul>
Session 4	Decision Tree Regression
	<ul style="list-style-type: none"> <li>• Algorithm for Decision Tree Induction</li> <li>• Characteristics of Decision Tree Regression</li> <li>• Methods for Expressing Attribute Test Conditions</li> <li>• Measures for Selecting an Attribute Test Condition</li> </ul>
Session 5	Support Vector Machine SVR
	<ul style="list-style-type: none"> <li>• Margin of a Separating Hyperplane   Linear SVM</li> <li>• Soft-margin SVM   Nonlinear SVM</li> <li>• Characteristics of SVM</li> </ul>
Week 11	Ensemble Methods (Bagging & Boosting & Stacking)
<b>Learning Objectives:</b> Ensemble methods help to improve the predictive performance of machine learning models. This week, you will learn about different ensemble methods that combine several machine-learning techniques into one predictive model to decrease variance, reduce bias, or improve predictions. You will learn about selecting one model over another, the concept of boosting, and its importance in machine learning. Additionally, you will learn how to convert weaker algorithms into stronger ones.	
Session 1	Random Forest Voting & Averaging
	<ul style="list-style-type: none"> <li>• Random Forests (RF).</li> <li>• Out of Bag Evaluation   OOB Score vs Test Set Score.</li> <li>• Train an RF Classifier/Regressor.</li> <li>• Evaluate the RF Classifier/Regressor.</li> <li>• Visualizing features importance.</li> </ul>

Session 2		Bagging   Boosting   Gradient Boosting   Ada-Boost   Stacking
		<ul style="list-style-type: none"> <li>• Define the bagging classifier   How do Boosting Algorithms work?</li> <li>• What is the Model Selection?</li> <li>• Evaluate Bagging performance.</li> <li>• Methods for Constructing an Ensemble Classifier.</li> <li>• Types of Boosting Algorithms.</li> </ul>
Session 3		Gradient Boosting   Ada-Boost   Stacking
		<ul style="list-style-type: none"> <li>• What is Boosting?</li> <li>• Gradient Boosting Vs. AdaBoost.</li> <li>• Cross-Validation.</li> <li>• Prediction with Stacking.</li> </ul>
Session 4		Class Imbalanced Problem
		<ul style="list-style-type: none"> <li>• Building Classifiers - Class Imbalance.</li> <li>• Data-based Approaches &amp; Algorithmic approach   Evaluating Performance - Class Imbalance</li> <li>• Finding an Optimal Score Threshold   Aggregate Evaluation of Performance</li> </ul>
Session 5		ML with RapidMiner
		<ul style="list-style-type: none"> <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> <li>• <b>Tableau:</b> <ul style="list-style-type: none"> <li>◦ Joining, Blending and Relationships   Dual Axis Charts   Joins with Duplicate Values</li> <li>◦ 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</li> </ul> </li> </ul>
Week 12		Model Evaluation & Optimization
<p><b>Learning Objectives:</b> Model building is an iterative process. Employing feature engineering techniques along with careful model selection helps to improve the model. Additionally, tuning the model is an important step to achieve the best possible results. This week covers the steps and processes involved in this approach. You will learn how to analyze the performance of each algorithm and dive deep into core machine learning concepts such as cost function, objective function, model optimization, model tuning, regularization, gradient boosting, and grid and random search.</p>		
Session 1		Model generalization: Assessing Predictive Accuracy for New Data
		<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> </ul>
Session 2		Evaluation of Classification Models & Regression Models
		<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   Cluster Evaluation</li> </ul>
Session 3		Model Optimization & Extreme Gradient Boosting
		<ul style="list-style-type: none"> <li>• Estimating the Confidence Interval for Accuracy</li> <li>• Introducing XGBoost   XGBoost: Fit/Predict</li> <li>• When should I use XGBoost?</li> <li>• Comparing the Performance of Two Models</li> <li>• Measuring accuracy   Measuring AUC</li> </ul>



Session 4		Model Explainability/Interpretability
		<ul style="list-style-type: none"> <li>• Comparing Explainability and Interpretability</li> <li>• What is Explainability in Machine Learning?</li> <li>• Different Types of Explanations</li> <li>• Approaches to Explainability in Machine Learning</li> <li>• What is Interpretability in Machine Learning?</li> <li>• Mechanistic Interpretability: A New Frontier</li> <li>• Why are Interpretability and Explainability Important?</li> <li>• Final Thoughts on Interpretability vs. Explainability</li> </ul>
Session 5		Team Projects 3rd Presentation (Data Preparation)
		<ul style="list-style-type: none"> <li>• Team Projects (Analytical Approach)</li> <li>• Team Leads' Third Presentation</li> </ul>
Week 13		Unsupervised Learning - Clustering Analysis
<p><b>Learning Objectives:</b> This week, we will shift to unsupervised learning techniques, where the 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. You will learn about commonly used clustering techniques such as K-Means Clustering and Hierarchical Clustering, along with anomaly detection algorithms, with deeper use of scikit-learn functionality.</p>		
Session 1		K-Mean Clustering
		<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> </ul>
Session 2		Agglomerative Hierarchical Clustering
		<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> </ul>
Session 3		DBSCAN   Mean Shift Clustering
		<ul style="list-style-type: none"> <li>• BIRCH   DBSCAN   Mini-Batch K-Means</li> <li>• Mean Shift   OPTICS   Spectral Clustering   Gaussian Mixture Model</li> </ul>
Session 4		Gaussian Mixed Models   Fuzzy C Means Clustering
		<ul style="list-style-type: none"> <li>• BIRCH   DBSCAN   Mini-Batch K-Means</li> <li>• Mean Shift   OPTICS   Spectral Clustering   Gaussian Mixture Model</li> </ul>
Session 5		ML with RapidMiner
		<ul style="list-style-type: none"> <li>• <b>RapidMiner:</b> <ul style="list-style-type: none"> <li>• Logistic Regression   Times Series   Clusters   Anomaly detection</li> <li>• Association Discoveries</li> </ul> </li> <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> </ul> </li> <li>• How to Create a Tree Map Chart   Creating a Customer Segmentation Dashboard Advanced Dashboard Interactivity   Analyzing the Customer Segmentation Dashboard</li> </ul>

**Week 14****Anomalies Detection & A/B Testing/ Association Rules/R-Systems**

**Learning Objectives:** This week is a continuation of the unsupervised machine learning algorithms from the previous week. We will delve deeper into Association Rule Mining, which has numerous applications, such as discovering sales correlations in transactional data and insights into medical datasets. Additionally, we will cover A/B Testing (also known as split tests), focusing on using Python to analyze customer behavior and business trends. You will learn how to create, run, and analyze A/B tests to make proactive, data-driven business decisions. Furthermore, we will explore recommender systems, which are software tools that select products to recommend to individual customers. You will learn how to produce successful recommender systems using past product purchase and satisfaction data to make high-quality personalized recommendations.

**Session 1****Anomalies Detections**

- Characteristics of Anomaly Detection Problems | Anomaly Detection Methods
- Classification based Models for Anomaly Detection
- 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)
- Clustering-based algorithms: Cluster-based Local Outlier Factor (CBLOF)
- Local Density Cluster-based Outlier Factor (LDCOF)

**Session 2****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 3****Association Rules - Apriori | Eclat | FP-Growth**

- Strength of an association rule
- The Apriori Principle
- Frequent Itemset Generation in the Apriori Algorithm
- Candidate Generation & Pruning
- Computational Complexity
- Horizontal vs Vertical Data Format
- The Intuition of ECLAT Algorithm
- Advantages of Eclat & Eclat vs Apriori

**Session 4****Recommender Systems: Generations 1/2/3/4**

- Matrix Factorization Model (Object recommendation)
- Content Filtering | Collaborating Filtering (CF)
- CF Neighborhood-Based Approach
- CF Location-Based Approach

**Session 5****Recommender Systems: Generations 5/6/7/8**

- Popularity Based | Content-Based
- Collaborative filtering
- Matrix Factorization (MF)

## Week 15

## Time Series Analysis

**Learning Objectives:** 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 data collected at different periods or intervals, known as time-series data or trend analysis. You will explore various models for time series modeling to analyze real-time dependent data for forecasting. By the end of the week, you will have a comprehensive understanding of time series analysis and be equipped with the skills to build and evaluate time series models for various applications.

### Session 1

### Time Series Processing & Assumptions

- What is Time Series Analysis?
- Importance of TSA
- Components of TSA
- TSA Assumptions & Processing

### Session 2

### AR - MA & EST Models

- Single Exponential Smoothing
- Forecasting with Single Exponential Smoothing
- Double Exponential Smoothing | Forecasting with Double Exponential Smoothing
- Triple Exponential Smoothing

### Session 3

### ARMA & ARIMA Models

- Forecasting
- Relation between time series: Causality & time lags
- Distinction between short & long run
- Study of agent's expectations
- Stationarity | ACF & PACF
- Trend removal | Seasonal adjustment

### Session 4

### SARIMAX Model

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

### Session 5

### Time Series with Prophet

- Introduction to Prophet
- Advantages of Prophet
- Installation of Prophet
- Time Series Forecasting with Prophet
- Plotted the forecasted components
- Adding ChangePoints to Prophet
- Adjusting Trend

# DEEP LEARNING



## 04 MODULE

## Week 16

## Advanced Python for Deep Learning (Self-paced)

**Learning Objectives:** This week, we dive into Advanced Python for Deep Learning and GenAI. Because the code is concise and readable, it makes it a perfect match for deep learning applications. Its simple syntax also enables applications to be developed faster when compared to other programming languages. Another major reason for using Python for deep learning is that the language can be integrated with other systems coded in different programming languages. This makes it easier to blend it with AI projects written in other languages.

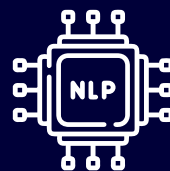
Session 1 Python: BeautifulSoup Library	
	<ul style="list-style-type: none"> <li>• Installing BeautifulSoup</li> <li>• Data extraction with BeautifulSoup</li> <li>• BeautifulSoup Usage   Filtering   BeautifulSoup Object</li> <li>• Extracting all the URLs found within a page 'a' tag   Extracting text from a page.</li> </ul>
Session 2 Python: SciPy	
	<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>
Session 3 Python: TensorFlow   Keras	
	<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   Testing your model</li> <li>• Using TensorBoard to visualize model performance</li> </ul>
Session 4 Python: Pytorch	
	<ul style="list-style-type: none"> <li>• Using User-Defined Triton Kernels with torch.compile</li> <li>• Large Scale Transformer model training with Tensor Parallel (TP)</li> <li>• Accelerating BERT with semi-structured (2:4) sparsity</li> <li>• torch.export Tutorial with torch.export.Dim</li> <li>• Extension points in nn.Module for load_state_dict and tensor subclasses</li> </ul>
Session 5 Lang Chain	
	<ul style="list-style-type: none"> <li>• Foundations of LangChain and Large Language Models (LLMs)</li> <li>• Advanced Techniques in LLMs and LangChain Integration</li> <li>• Data Management</li> <li>• Vector stores, Indices, and Prompts</li> <li>• Constructing and Deploying LangChain Applications</li> </ul>
Week 17 Deep Learning: ANN   CNN   GANs	
<p><b>Learning Objectives:</b> This week, we dive into Deep Learning, an artificial intelligence function that mimics the workings of the human brain in processing data and creating patterns for use in decision-making. Deep Learning uses an artificial neural network, composed of multiple levels arranged in a hierarchy, to carry out the machine learning process. In this session, you will learn about the basic building blocks of artificial neural networks. You will explore how deep learning networks can be successfully applied to data for knowledge discovery, knowledge application, and knowledge-based prediction.</p>	
Session 1 Introduction to Deep Learning	
	<ul style="list-style-type: none"> <li>• Deep Learning Explained</li> <li>• Biological Neural Network</li> <li>• Perceptron</li> <li>• Gradient Decent</li> <li>• Scholastic Gradient Descent</li> </ul>
Session 2 Artificial Neural Network (ANN)	
	<ul style="list-style-type: none"> <li>• ANN Explained</li> <li>• Multi-layer Neural Network</li> <li>• Characteristics of ANN</li> </ul>



Session 3 Convolutional Neural Network (CNN)	
	<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> </ul>
Session 4 Generative Adversarial Networks (GANs)	
	<ul style="list-style-type: none"> <li>• Challenge of GAN Loss</li> <li>• Standard GAN Loss Functions</li> <li>• Alternate GAN Loss Functions</li> <li>• Effect of Different GAN Loss Functions</li> </ul>
Session 5 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>
Week 18 Self-Organizing Maps   Boltzmann Machines   Autoencoders	
<p><b>Learning Objectives:</b> This week is a continuation of week 19. You will learn about three additional deep learning models: Self-Organizing Maps (SOMs), Boltzmann Machines (BM), and Autoencoders. In SOMs, you will learn how to detect specific features, such as fraud, and how to create a hybrid deep learning model. For Boltzmann Machines, you will explore the energy-based model perspective and, in practical lectures, focus on the probabilistic graphical model perspective. Finally, you will learn how to build an autoencoder from scratch using PyTorch, including how to manipulate classes and objects to improve and tune your autoencoder.</p>	
Session 1 Introduction to SOMs	
	<ul style="list-style-type: none"> <li>• Introduction to Self-Organizing Maps</li> <li>• How Do SOMs Work?</li> <li>• Users of SOMs</li> <li>• SOMs Architecture</li> <li>• Pros &amp; Cons of SOMs</li> </ul>
Session 2 Self-Organizing Maps	
	<ul style="list-style-type: none"> <li>• How do Self-Organizing Maps work?</li> <li>• How Self-Organizing Maps Learn?</li> <li>• Live SOM example</li> <li>• Energy-based Models (EBMs)</li> <li>• Restricted Boltzmann Machines (RBM)</li> <li>• Deep Boltzmann Machines (DBM)</li> </ul>
Session 3 Introduction to BMN & AutoEncoders	
	<ul style="list-style-type: none"> <li>• Introduction to Boltzmann Machine</li> <li>• Users &amp; BMNs Architecture</li> <li>• Pros &amp; Cons of BMNs</li> <li>• Introduction to Auto Encoder</li> <li>• Users of Auto Encoder s</li> <li>• Auto Encoder's Architecture</li> <li>• Pros &amp; Cons of Auto Encoder s</li> </ul>
Session 4 Boltzmann Machines	
	<ul style="list-style-type: none"> <li>• How Do BMNs Work?</li> <li>• Energy-Based Models (EBM)</li> <li>• Deep Belief Networks</li> <li>• Deep Boltzmann Machine</li> </ul>

Session 5		AutoEncoders
		<ul style="list-style-type: none"> <li>• How Do Auto Encoders Work?</li> <li>• Training an Auto Encoder</li> <li>• Overcomplete hidden layers</li> <li>• Sparse Autoencoders</li> <li>• Denoising Autoencoders</li> <li>• Contractive Autoencoders</li> <li>• Stacked Autoencoders</li> <li>• Deep Autoencoders</li> </ul>
Week 19		Sequence Learning & GANs
<p><b>Learning Objectives:</b> This week, we explore another fascinating application of neural networks: equipping computers to understand human language. You will learn to work with text data and sequential data, delving into the world of Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs).</p> <p>RNNs are designed to recognize patterns in sequences of data, making them ideal for tasks involving time series, language modeling, and more. LSTMs, a type of RNN, are particularly effective at learning long-term dependencies, which are crucial for understanding context in text and speech.</p>		
Session 1		Recurrent Neural Networks (RNN)
		<ul style="list-style-type: none"> <li>• RNN Explained</li> <li>• Characteristics of RNN</li> <li>• Backpropagation</li> <li>• Gradient problem</li> <li>• Vanishing vs. Exploding</li> </ul>
Session 2		Long Short-Term Memory Networks (LSTM)
		<ul style="list-style-type: none"> <li>• Long Short-Term Memory Networks</li> <li>• A Quick Look into LSTM Architecture</li> <li>• Why does LSTM outperform RNN?</li> <li>• Deep Learning about LSTM gates</li> </ul>
Session 3		Deep RNN - Bidirectional RNN - GRU
		<ul style="list-style-type: none"> <li>• Introduction to Gated Feedback Recurrent Neural Networks (GRU)</li> <li>• Bidirectional RNN • Deep Recurrent Networks</li> <li>• Vanishing and Exploding gradients.</li> <li>• Methods to prevent the vanishing(exploding) gradient</li> </ul>
Session 4		RNN vs LSTM with Google Stock Price
		<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> </ul>
Session 5		Sentiment Analysis with LSTM
		<ul style="list-style-type: none"> <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> </ul>

# NATURAL LANGUAGE PROCESSING



05  
MODULE

## Week 20

## Text Analysis & Natural Language Processing (NLP)

**Learning Objectives:** This week, we dive into Natural Language Processing (NLP), a technology that is deeply and widely penetrating the market across various industries and domains. NLP is extensively applied in businesses today and has become a buzzword in every engineer's life. Its applications range from customer service automation to sophisticated data analysis, making it an indispensable skill in the modern tech landscape. As NLP continues to evolve, it is transforming the way we interact with machines and manage information.

### Session 1

### Introduction to Natural Language Processing

- What is Natural Language Processing?
- History of Natural Language Processing
- NLP Applications | Levels | Components
- NLU vs NLG
- NLP Pipeline and Tasks | NLP Toolkits and Libraries
- NLP Challenges

### Session 2

### Essentials of Natural Language Processing 1

- Basic Text Analysis
- Tokenization | POS Tagging
- Stop Word Removal
- Text Normalization
- Spelling Correction

### Session 3

### Essentials of Natural Language Processing 2

- Stemming | Lemmatization
- Named Entity Recognition (NER)
- Word Sense Disambiguation
- Sentence Boundary Detection

### Session 4

### NLP Feature Extractions

- Data Structures
- NLP Pre Processing
- The Bag of Words
- Frequency Vector (Count of Vectorization)
- One-Hot Encoding • Term Frequency – Inverse Term Frequency
- Distributed Representation
- Word Embedding

### Session 5

### NLP with TextBlob & spaCy (NER)

- Introduction to TextBlob
- Language Detection
- POS
- Word Inflection
- Sentiment Analysis
- Introduction to spaCy Library
- Objects of spaCy Library
- The Statistical Modeling
- Processing Pipelines

Week 21	Advanced Natural Language Processing (NLP)
<b>Learning Objectives:</b> This week, we dive into advanced Natural Language Processing (NLP) modeling. As NLP continues to evolve, it is transforming the way we interact with machines and manage information. Advanced NLP techniques, such as deep learning-based models, are pushing the boundaries of what machines can understand and generate. Models like BERT, GPT-3, and Transformer networks are revolutionizing tasks like language translation, text generation, and question answering. Mastering these advanced techniques will enable you to tackle complex NLP challenges and create state-of-the-art solutions.	
Session 1	Text Classification
	<ul style="list-style-type: none"> <li>• Text Data Preparation</li> <li>• Data Vectorization</li> <li>• Text Classification</li> </ul>
Session 2	Text Summarization
	<ul style="list-style-type: none"> <li>• What is Text Summarization?</li> <li>• Text Summarization Categories</li> <li>• Stages of Text Summarization</li> </ul>
Session 3	Topic Modeling
	<ul style="list-style-type: none"> <li>• What is Topic Modeling?</li> <li>• Topic Modeling Use cases</li> <li>• Topic Modeling Libraries</li> <li>• Latent Semantic Analysis (LSA)   Latent Dirichlet Allocation (LDA)   Hierarchical Dirichlet Process (HDP)</li> </ul>
Session 4	Sentiment Analysis
	<ul style="list-style-type: none"> <li>• What is Sentiment Analysis?</li> <li>• Types of Sentiment Analysis</li> <li>• Benefits of Sentiment Analysis</li> <li>• Examples of Sentiment Analysis</li> <li>• Challenges of Sentiment Analysis</li> </ul>
Session 5	Team Projects 5th Presentation (Action Plan)
	<ul style="list-style-type: none"> <li>• Team Projects (Action Plan)</li> <li>• Lesson Learned</li> <li>• Team Leads' Fourth Presentation</li> </ul>

# GENERATIVE AI



**06**  
MODULE

Week 22	GenAI - LLMs - sLLMs
<b>Learning Objectives:</b> This week, we dive into Generative AI, a cutting-edge field that focuses on creating models capable of generating new content, such as text, images, and music. Generative AI leverages advanced techniques like generative adversarial networks (GANs) and transformer models to produce highly realistic and creative outputs. By mastering these technologies, you will be able to develop applications that can autonomously generate content, offering immense value across various industries.	
Session 1	Language Models and Transformer-based Generative Models
	<ul style="list-style-type: none"> <li>• What are LLMs? Why LLMs (such as GPT and Bard) is so special?</li> <li>• The Power of Attention Mechanisms   Encoder-Decoder Architectures</li> <li>• Demystifying Transformers   Learn and Understand Transformers</li> </ul>



	<ul style="list-style-type: none"> <li>• Scaled Dot Product Attention   Multi-Headed Attention</li> <li>• Exploring the GPT Architecture: Foundations   Innovations</li> <li>• Practical Applications of Masked Multi-Head Attention</li> <li>• GPT Pre-Training Strategies and Techniques</li> <li>• Optimizing GPT Pre-Training for Performance</li> </ul>
<b>Session 2</b>	<b>Generative AI with LLMs</b>
	<ul style="list-style-type: none"> <li>• LLMs and Generative AI Project Lifecycle</li> <li>• LLM Pre-Training and Scaling</li> <li>• Fine-tuning LLMs with Specific Instructions</li> <li>• Efficient Fine-Tuning of Parameters</li> <li>• Reinforcement Learning from Human Response</li> </ul>
<b>Session 3</b>	<b>LLMs for Search, Prediction, and Generation</b>
	<ul style="list-style-type: none"> <li>• Search Query Completion</li> <li>• Next Word Prediction</li> <li>• Word Embeddings</li> <li>• Transformers</li> <li>• Generating Text</li> <li>• Stacking Attention Layers</li> </ul>
<b>Session 4</b>	<b>Interacting with Data Using LangChain and RAG</b>
	<ul style="list-style-type: none"> <li>• LangChain Foundations</li> <li>• Benefits of using LangChain</li> <li>• Using LangChain to Develop LLM Applications</li> <li>• Value Propositions of LangChain   Components of LangChain</li> <li>• Off-the-Shelf Chains in LangChain</li> <li>• Build and Deploy LLM powered applications using LangChain</li> </ul>
<b>Session 5</b>	<b>LangChain for LLM Application Development</b>
	<ul style="list-style-type: none"> <li>• Understanding Retrieval-Augmented Generation (RAG)</li> <li>• Document Loading and Splitting</li> <li>• Vector Stores and Embeddings</li> <li>• Retrieval   Question Answering with Chatbots</li> <li>• Building RAG Models using LangChain</li> </ul>
<b>Week 23</b>	<b>Chatbots - ChatGPT &amp; Prompt Engineering</b>
<b>Learning Objectives:</b> This week, we dive into Chatbots, focusing on ChatGPT and Prompt Engineering, which are revolutionizing human-computer interactions. You will explore how ChatGPT leverages advanced natural language processing to generate human-like responses and how prompt engineering can enhance the performance and accuracy of these chatbots. By understanding and applying these techniques, you'll be able to design and implement sophisticated chatbots that can effectively communicate and assist users in various applications.	
<b>Session 1</b>	<b>Unveiling ChatGPT: Conversing with Superintelligence</b>
	<ul style="list-style-type: none"> <li>• Introduction to Generative AI</li> <li>• Introduction to ChatGPT and OpenAI</li> <li>• Unleashing the Power of ChatGPT</li> <li>• The Applications of ChatGPT</li> <li>• Human-AI Collaboration and the Future</li> <li>• Engaging with ChatGPT</li> <li>• Wrapping Up and Looking Ahead</li> </ul>

Session 2 ChatGPT: Pre-processing and Fine-tuning	
	<ul style="list-style-type: none"> <li>• Overview of language models</li> <li>• Understanding the architecture of the GPT model</li> <li>• GPT models: advantages and disadvantages</li> <li>• Overview of the pre-trained GPT models available for fine-tuning</li> <li>• Training of ChatGPT   Data preparation   Model architecture</li> <li>• Hyperparameter tuning   Training process</li> </ul>
Session 3 Prompt Engineering and ChatGPT Plugins	
	<ul style="list-style-type: none"> <li>• Introduction to Prompt Engineering</li> <li>• Why Prompt Engineering?  What is Prompt Engineering?</li> <li>• Applications of Prompt Engineering</li> <li>• Types of Prompting   Priming Prompts   Prompt Decomposition</li> <li>• How to Get Better Responses from ChatGPT</li> <li>• ChatGPT Plugins</li> </ul>
Session 4 Building an AI AI-powered chatbot	
	<ul style="list-style-type: none"> <li>• Creating a Chatbot using Chatterbot</li> <li>• Building an Interactive Chatbot using Rasa</li> <li>• Developing your own Chatbot using ChatGPT API</li> </ul>
Session 5 Popular Generative AI Tools	
	<ul style="list-style-type: none"> <li>• Dall-E 2   Midjourney</li> <li>• Bard   Hugging Face</li> <li>• NLG Cloud   Copy.ai</li> <li>• Tome   Codeium   WriteSonic</li> </ul>
Week 24	GenAI - Tools & Applications
<b>Learning Objectives:</b> This week, we learn Generative AI, focusing on the tools and applications that are transforming industries. You will explore various generative AI models and tools, such as GANs, VAEs, and diffusion models, and understand how they are used to create innovative applications in fields like art, entertainment, and data augmentation. By mastering these tools, you will be equipped to develop cutting-edge solutions that leverage the power of generative AI to produce creative and practical outcomes.	
Session 1 Leveraging Generative AI for Fraud Detection	
	<ul style="list-style-type: none"> <li>• Overview of Gen AI in Fraud Detection</li> <li>• Email Fraud Detection using the GAN model.</li> <li>• Best Practices</li> </ul>
Session 2 Using Midjourney for Generative AI Art	
	<ul style="list-style-type: none"> <li>• Artistic Exploration with Gen AI Midjourney</li> <li>• Getting Started with Midjourney</li> <li>• Midjourney User Interface</li> <li>• Crafting Visuals using Midjourney</li> </ul>
Session 3 GitHub Copilot for Developers	
	<ul style="list-style-type: none"> <li>• GitHub Copilot: Introduction, Installation, and Configuration</li> <li>• GitHub Copilot: Improving Developer Efficiency</li> <li>• Application Scenarios, Problem-Solving, and Summary</li> </ul>
Session 4 Generative AI: Privacy and Protection Perspectives	
	<ul style="list-style-type: none"> <li>• Introduction to Data Privacy</li> <li>• Privacy Challenges and Regulations in Generative AI</li> <li>• Safeguarding Data Privacy at Your Workplace</li> <li>• Legal and Ethical Considerations</li> </ul>

Session 5		Generative Privacy and Protection Perspectives & Cyber Security
		<ul style="list-style-type: none"> <li>• Introduction to Data Privacy</li> <li>• Privacy Challenges and Regulations in Generative AI</li> <li>• Safeguarding Data Privacy at Your Workplace</li> <li>• Legal and Ethical Considerations</li> <li>• Overview of Cybersecurity - Essentials of Cybersecurity</li> <li>• Gen AI Applications in Cybersecurity - Intrusion Detection Systems</li> <li>• Automating Security Operations using Gen AI</li> <li>• Enhancing Network and Endpoint Security with Gen AI</li> <li>• Anticipated Trends and Challenges in the Future</li> <li>• Components, Strongly Connected Components, Label Propagation</li> </ul>
Week 25		GenAI in Cloud
<p><b>Learning Objectives:</b> This week, we dive into Machine Learning, Deep Learning, and Generative AI in the Cloud, focusing on how cloud platforms are revolutionizing these technologies. You will explore cloud-based tools and services, such as AWS SageMaker, Google AI Platform, and Azure Machine Learning, that enable scalable and efficient deployment of machine learning and deep learning models. By leveraging these cloud solutions, you will be able to build, train, and deploy sophisticated AI models, including generative AI, with enhanced performance and accessibility.</p>		
Session 1		Generative AI on the Cloud
		<ul style="list-style-type: none"> <li>• Overview of generative AI - key concepts like prompting and few-shot learning.</li> <li>• Demo generative models like GPT-3, DALL-E, and others.</li> <li>• Discuss business use cases and implications of generative AI.</li> <li>• Introduction to Amazon AI Services</li> <li>• Overview of AWS AI services like SageMaker, Comprehend, Lex, Polly etc.</li> </ul>
Session 2		Amazon SageMaker   DeepComposer
		<ul style="list-style-type: none"> <li>• Amazon SageMaker</li> <li>• Managed service to build, train, and deploy ML models quickly.</li> <li>• Build models with built-in algorithms or bring your own.</li> <li>• AWS Deep Composer</li> <li>• ML service to create original music. Learn generative AI for music composition.</li> </ul>
Session 3		LLMs Vis APIs
		<ul style="list-style-type: none"> <li>• Intro to LLMs like GPT-3 and Codex</li> <li>• How they are trained and optimized for few-shot learning</li> <li>• Using LLMs in Python with HuggingFace Transformers</li> <li>• Fine-tuning and prompting LLMs for custom applications.</li> <li>• Hosting LLMs on SageMaker and accessing via APIs</li> </ul>
Session 4		Amazon Comprehend   Transcribe
		<ul style="list-style-type: none"> <li>• Amazon Comprehend</li> <li>• Natural language processing (NLP) service for text analysis</li> <li>• Amazon Transcribe</li> <li>• Automatic speech-to-text transcription service using ML.</li> <li>• Amazon recognition</li> <li>• Image and video analysis service for object and scene detection, facial analysis etc.</li> </ul>
Session 5		AWS DeepLens   AWS DeepRacer
		<ul style="list-style-type: none"> <li>• Monitoring and optimizing generative model performance.</li> <li>• Bias detection and mitigation in generative models</li> <li>• Explainability and interpretability of generative AI</li> <li>• MLOps - automating deployment, monitoring, and governance of generative models</li> <li>• Course conclusion and next steps for applying generative AI.</li> </ul>

- Amazon Lex - Build conversational chatbots with the same technology as Alexa.
- Amazon Polly - Turn text into lifelike speech with ML.
- Amazon Forecast - Time series forecasting using ML.
- AWS DeepLens - ML-enabled video camera for object detection and more.
- AWS DeepRacer - Autonomous 1/18th scale race car to test RL models.

# MLOps



## 07 MODULE

### Week 26

### General Deployment Considerations & MLOps

**Learning Objectives:** This week, you will learn every aspect of how to put your models into production. You will cover all the steps and infrastructure required to deploy machine learning models professionally, and you will have at your fingertips a sequence of steps to follow for deploying a machine learning model. Additionally, you will receive a project template with full code that you can adapt to deploy your models.

#### Session 1

#### Avoiding False Discoveries | MLOps Introduction

- Avoiding False Discoveries
- Preliminaries: Statistical Testing
- Modeling Null & Alternative Distributions
- Champion Models Vs. Challenge Models
- Machine learning industrialization challenges

#### Session 2

#### Model Deployment with Flask/Postman

- What is MLOps?
- MLOps Motivation: High-level view | MLOps challenges
- MLOps Components Machine Learning Life Cycle
- Machine Learning Deployment with Flask
- Machine Learning Deployment with Postman

#### Session 3

#### Introduction to ML and MLOps stages

- How does it relate to DevOps, AIOps, ModelOps, and GitOps?
- Major Phases - what it takes to master MLOps.
- CI/CD in Production Case Study
- MLOps Maturity Model Detailed
  - MLOps and stages | Versioning | Testing Automation (CI/CD) | Reproducibility
- Deployment Monitoring

#### Session 4

#### MLOps Architectures

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

#### Session 5

#### Introduction to CI/CD

- Introduction to CI and CD
- CI/CD challenges in Machine Learning
- Steps involved in the CI/CD implementation in ML lifecycle and workflow.
- A glimpse of popular Tools used in the DevOps ecosystem on the Cloud.



# MINI PROJECTS & CAPSTONE PROJECT



08  
MODULE

## Week 27-30

## Individual Capstone Project (1 Month)

**Learning Objectives:** During the final week, you will be transitioned into a 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 8th of August 2025 till the 8th of September 2025

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

## Data Science Interview Questions & Answers

The program includes real Data Science interview questions designed to prepare you for landing your dream job. These questions span various topics, including understanding the big picture, model optimization, data pre-processing, sampling, and splitting, supervised and unsupervised learning, model evaluation, ensemble learning, and practical business applications. This comprehensive approach ensures you are well-equipped to tackle the challenges in a data science interview:

- 1- The Big Picture
- 2 - Optimization
- 3 - Data Pre-processing
- 4 - Sampling & Splitting
- 5- Supervised Learning
- 6 - Unsupervised Learning
- 7 - Model Evaluation
- 8 - Ensemble Learning |
- 9 - Business Applications

## Data Science Interview Real Challenges

To prepare you for landing your dream job, the program includes 48 hours of take-home, real industry challenges. These challenges range from understanding the big picture and optimizing models to data pre-processing, sampling, and splitting. Key areas covered include mastering supervised and unsupervised learning, evaluating models, leveraging ensemble techniques, and applying insights to real-world business applications:

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

## Week 31-40

## Certified Analytics Professional Training Bonus (Self-paced)

End the program with a self-paced and supported training Certified Analytics Professional (CAP) certification. CAP is a globally recognized credential that validates your expertise in the entire analytics process, from problem framing to data analysis and model deployment. Earning the CAP certification demonstrates your proficiency in transforming data into valuable insights and your commitment to the analytics profession. The program will help you master.

- Domain 1: Business Problem Framing
- Domain 2: Analytics Problem Framing
- Domain 3: Data
- Domain 4: Methodology (Approach) Selection
- Domain 5: Model Building
- Domain 6: Deployment



## CAREER SERVICES



**Resume  
Preparation**



**Interview  
Preparation**



**Preparation  
Employment  
Assistance**



**Certificate Analytics  
Professional  
Training (Free)**

## ATC Innovation Center

925 North Point Pkwy Suite 130, Alpharetta, GA 30005

 [www.innovatics.ai](http://www.innovatics.ai)  +1 (678) 209 9780  [register@innovatics.ai](mailto:register@innovatics.ai)





## CERTIFICATION & ACCREDITATION

Upon successful completion of the program, InnovatiCS grants a verified/certified digital certificate in Data Science and AI to participants. This program is graded as pass or fail; participants must receive 80% to pass and obtain the certificates and the accreditations.



This program has been certified by 3 International body knowledge: The International Association of Business Analytics Certification (**IABAC**), Continuing Professional Development (**CPD**) and Higher Learning Accreditation Consultant and Training (**HLACT**) as conforming to continuing professional development principles. After successful completion of the program, you will get up to **4 international certificates in Data Science and AI**. Your verified digital certificate will be email to you in the name you used when registering for the program. All certificate images are for illustrative purpose only and may be subject to change at the direction of InnovatiCS

## ABOUT US

We are InnovatiCS, a comprehensive up-skilling platform driven by our unique, integrated "Learn-Apply-Solve" framework. This innovative approach offers an application-focused, immersive learning experience, featuring real-world industry courses, case studies, datasets, and projects. Our program fosters a seamless blend of industry and academic knowledge through simulations and conceptualization.

InnovatiCS regularly presents at leading conferences and workshops, and until recently, hosted monthly meetups featuring industry experts as speakers. Our live, multi-week, multi-session courses, which are recorded for flexibility, have received high praise from participants, who benefit from our nearly unlimited one-on-one and group support sessions.

## OUR OFFICES



**UAE**

**+971 58 971 9975**



**USA**

**+1 678-209-9780**



**MOROCCO**

**+212 666128384**



**QATAR**

**+974 3322 2966**



**KUWAIT**

**+965 6636 9857**

**VISIT OUR WEBSITE**

 [www.innovatics.ai](http://www.innovatics.ai)