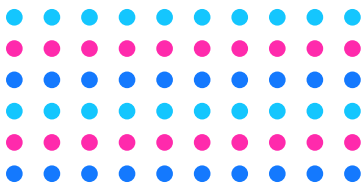# MastersCampus Academy

**9 months**

# Professional Data Science and AI Engineering Program

# Why Learn Data Science and AI?

## 01. The Power of Data Science

Data science is transforming industries by enabling businesses to leverage data for decision-making, innovation, and problem-solving. Here's why mastering data science is crucial:

- **Data-Driven Decision Making:** Leveraging data science allows organizations to make informed decisions based on actionable insights, improving strategic positioning and operational efficiency.

- **Problem Solving:** Data scientists use statistical and computational techniques to tackle complex issues and provide solutions across various industries.

- **Innovation:** Data science drives innovation by analyzing data to develop new products and services, impacting sectors such as technology, finance, healthcare, and retail.

# 02

# The Future of AI

- **Automation:** AI automates routine tasks, enhancing efficiency and reducing operational costs across industries.

- **Enhanced User Experiences:** AI technologies like recommendation systems, chatbots, and virtual assistants offer personalized and engaging user experiences.

- **Advanced Analytics:** AI algorithms provide predictive insights from large datasets, aiding strategic decision-making.

- **Innovation in Products and Services:** AI is at the forefront of technological advancements, including autonomous vehicles, smart devices, and advanced healthcare solutions.

## 03. The Dominance of AI in the Industry

- **Healthcare:** AI enhances diagnostics, personalized medicine, and patient care through advanced analytics.

- **Finance:** AI improves fraud detection, risk management, and automated trading systems.

- **Retail:** AI optimizes inventory management, customer service, and personalized marketing strategies.

- **Manufacturing:** AI drives predictive maintenance, quality control, and supply chain optimization.

- **Transportation:** AI transforms logistics, autonomous vehicles, and route optimization.

## 04.  Average Salaries for Data Roles in India

**For Freshers:**

- **Data Analyst:** ₹4.5L - ₹8L per annum
- **Data Scientist:** ₹8L - ₹15L per annum
- **Machine Learning Engineer:** ₹10L - ₹18L per annum
- **Business Intelligence (BI) Developer:** ₹6L - ₹12L per annum
- **Data Engineer:** ₹8L - ₹16L per annum
- **AI Research Scientist:** ₹12L - ₹22L per annum
- **ML Ops Engineer:** ₹9L - ₹17L per annum

**For Experienced Professionals:**

- **Data Analyst:**
  - **3-5 Years:** ₹6L - ₹12L per annum
  - **5-10 Years:** ₹12L - ₹20L per annum
- **Data Scientist:**
  - **3-5 Years:** ₹12L - ₹20L per annum
  - **5-10 Years:** ₹20L - ₹35L per annum
- **Machine Learning Engineer:**
  - **3-5 Years:** ₹15L - ₹25L per annum
  - **5-10 Years:** ₹25L - ₹40L per annum
- **Business Intelligence (BI) Developer:**
  - **3-5 Years:** ₹8L - ₹15L per annum
  - **5-10 Years:** ₹15L - ₹25L per annum
- **Data Engineer:**
  - **3-5 Years:** ₹12L - ₹20L per annum
  - **5-10 Years:** ₹20L - ₹35L per annum
- **AI Research Scientist:**
  - **3-5 Years:** ₹20L - ₹30L per annum
  - **5-10 Years:** ₹30L - ₹50L per annum

- **ML Ops Engineer:**
  - **3-5 Years:** ₹12L - ₹22L per annum
  - **5-10 Years:** ₹22L - ₹35L per annum

## 05. Future Job Creation

- **Job Growth:** Demand for data scientists, machine learning engineers, and AI specialists is projected to increase by 30-40% over the next decade.

- **Emerging Roles:** New roles such as AI ethicists, data privacy officers, and AI trainers are emerging, reflecting the expanding scope of these fields.

- **Cross-Industry Demand:** Data and AI professionals will be needed across various sectors to drive innovation and improve operations.
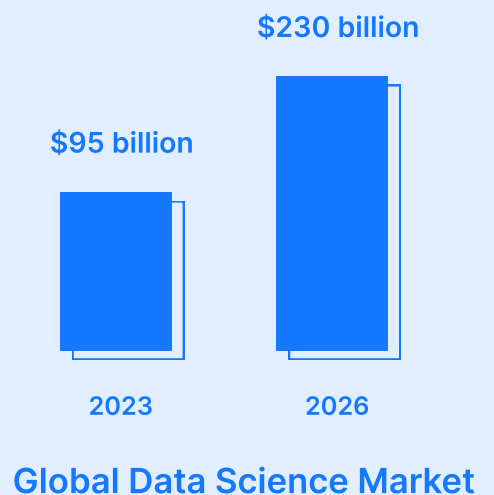
## 06. Market Size and Growth

The data science and AI markets are experiencing significant growth.

Valued at around **$136 billion** in 2023, projected to grow at a CAGR of **38%** to surpass **$500 billion** by 2027.

Global Data Science Market expected to grow at a CAGR of **26%**.

**$230 billion**

**$95 billion**

2023      2026

**Global Data Science Market**

## 07. Impact on Other Roles

The advancement of data science and AI will transform and impact various roles:

- **Job Transformation:** Traditional roles will evolve to incorporate data-driven decision-making and automation.

- **New Skill Requirements:** Professionals will need to acquire data literacy and AI skills to stay relevant.

- **Interdisciplinary Opportunities:** Data science and AI will intersect with fields like engineering, business management, and social sciences, creating new career paths

# MastersCampus
## Academy

## Program Overview

The **Professional Data Science and AI Engineering Program**, delivered entirely online by MasterCampus Academy, is crafted to equip students with the essential skills and knowledge required for advanced data roles. This program is designed for those aiming to excel in data science and AI before completing their engineering or bachelor's degrees

# Why Choose MasterCampus Academy for the Professional Data Science and AI Engineering Program?

## 01. Industry-Leading Expertise

**Collaborations with Top Tech Giants:** Partnering with leading technology companies such as Dell, Manhattan Associates, Amazon, NVIDIA, and Verizon, our curriculum is developed with insights from industry leaders. This ensures that the program incorporates the latest technologies and best practices in data science and AI.

## 02. Cost-Effective Education

**Affordable Excellence:** We offer a high-quality education comparable to master's programs at a fraction of the cost. Our aim is to provide advanced education that is both accessible and affordable, helping students gain significant skills without the financial strain of traditional programs.

## 03. Comprehensive and Practical Curriculum

**From Foundations to Advanced Techniques:** The program covers foundational statistics and data analysis, advanced machine learning, deep learning, MLOps, and generative AI. It features hands-on projects and capstone projects that address real-world problems, ensuring practical application of theoretical concepts.

## 04. Flexibility and Accessibility

**Online Learning for Convenience:** Delivered entirely online, the program offers flexibility for students to learn at their own pace, making it ideal for those balancing academic or professional commitments.

## 05. 100% Placement Assistance

**Career Support:** We provide comprehensive placement assistance, including resume building, interview preparation, and job search support. Our goal is to ensure that students are well-prepared and confident in securing roles in data science and AI.

## 06. Industry-Relevant Projects

**Real-World Experience:** Capstone projects across domains like healthcare, finance, and retail offer students practical experience that enhances employability and prepares them for real-world scenarios.

## 07. Proven Track Record

**Quality Education:** With a strong history of delivering quality training in data science and AI, our programs are designed to meet the evolving needs of the industry, equipping students for success in the competitive job market.

## 08. Expert Faculty and Mentors

**Guidance from Professionals:** The program features instructors and mentors with extensive experience in data science, machine learning, AI, and related fields, providing personalized guidance and real-world insights.

## 09. Global Learning Community

**Inclusive Network:** Join a diverse global community of learners and professionals. MasterCampus Academy fosters an inclusive and collaborative environment where students can network, share ideas, and learn from peers and experts worldwide.

## 10. Cutting-Edge Technology and Tools

**Hands-On Experience:** Gain practical experience with the latest tools and technologies used in the industry, including advanced machine learning frameworks, AI platforms, and data visualization tools, ensuring proficiency with modern data technologies.

## Uniqueness & Unmatched Value: MastersCampus Academy's Professional Data Science and AI Engineering Program

**Tailored Learning Paths for All Experience Levels**

1. **Customizable for Freshers to Experienced Professionals**
   - **For Freshers (0-2 Years):** This program offers specialized training for entry- level roles in data and business analytics. Freshers will gain foundational skills and practical knowledge needed for roles in data

- analysis, business intelligence, and related fields.

- **For Mid-Level Professionals (2-3 Years):** Designed to advance skills for professionals with some industry experience, the program focuses on product analytics and advanced analytics roles. It provides deeper insights into handling complex data challenges and making strategic decisions.

- **For Experienced Professionals (3+ Years):** Tailored modules and advanced training help experienced professionals transition into data science roles. The curriculum includes advanced data science techniques, machine learning, and AI applications to prepare them for higher-level roles and leadership positions.

## 2. Continuous Career Support and Guidance

- **Personalized Career Assistance:** We offer targeted support based on the learner's experience level. Freshers receive guidance on entering the industry, while mid-level and experienced professionals get assistance for career advancement and role transitions.

- **Ongoing Job Assistance:** We provide continuous career support, including resume building, interview preparation, and job search assistance, ensuring that learners are well-prepared and confident in securing and advancing in their data science careers.

## 3. Adaptive Curriculum and Practical Experience

- **Relevant Content for Every Stage:** The curriculum adapts to different levels of expertise, offering foundational concepts for beginners and advanced techniques for seasoned professionals. This approach ensures that learners acquire relevant skills applicable to their career stage.

- **Real-World Projects:** Industry-relevant and role-specific projects are included to address real-world challenges. These projects enhance learners' practical experience and job readiness, tailored to their level of expertise.

## 4. Flexible Learning and Career Development

- **Adaptable to Changing Needs:** The program is designed to be flexible,

- accommodating different learning paces and career goals. We adapt the curriculum and support based on individual needs and industry trends.

- **Dynamic Career Development:** We offer continuous career development resources, keeping learners updated with industry trends and additional training opportunities to ensure they remain competitive in the evolving data science field.

## 5. Integrated Learning and Career Services

- **Holistic Approach:** By integrating educational content with career development services, we ensure that learners not only gain knowledge but also effectively apply it to secure and advance in their desired roles.

- **Focus on Individual Goals:** Understanding that each learner's career path is unique, the program is designed to accommodate individual career aspirations and provide tailored support to help achieve them.

## 6. Unique Project Design for Experienced Professionals

- **Customized Projects Based on Prior Experience:** For professionals with significant industry experience, we design unique projects that align with their previous roles and domain expertise. This approach ensures that the projects are relevant and leverage their existing knowledge, enhancing the applicability of new skills in their current or desired roles.

- **Domain-Specific Experience Integration:** The program includes projects that are tailored to the specific domains and companies where the professionals have worked. This customization allows them to tackle complex problems and create solutions that are directly relevant to their industry, reinforcing their learning experience and demonstrating their expertise in a context that employers value.

- **Enhanced Professional Growth:** By focusing on domain-specific projects, we help experienced professionals deepen their expertise, make meaningful contributions to their field, and advance their careers with a portfolio of work that highlights their specialized skills and experience.
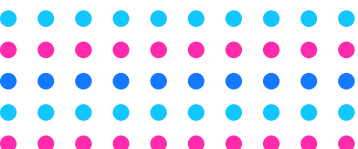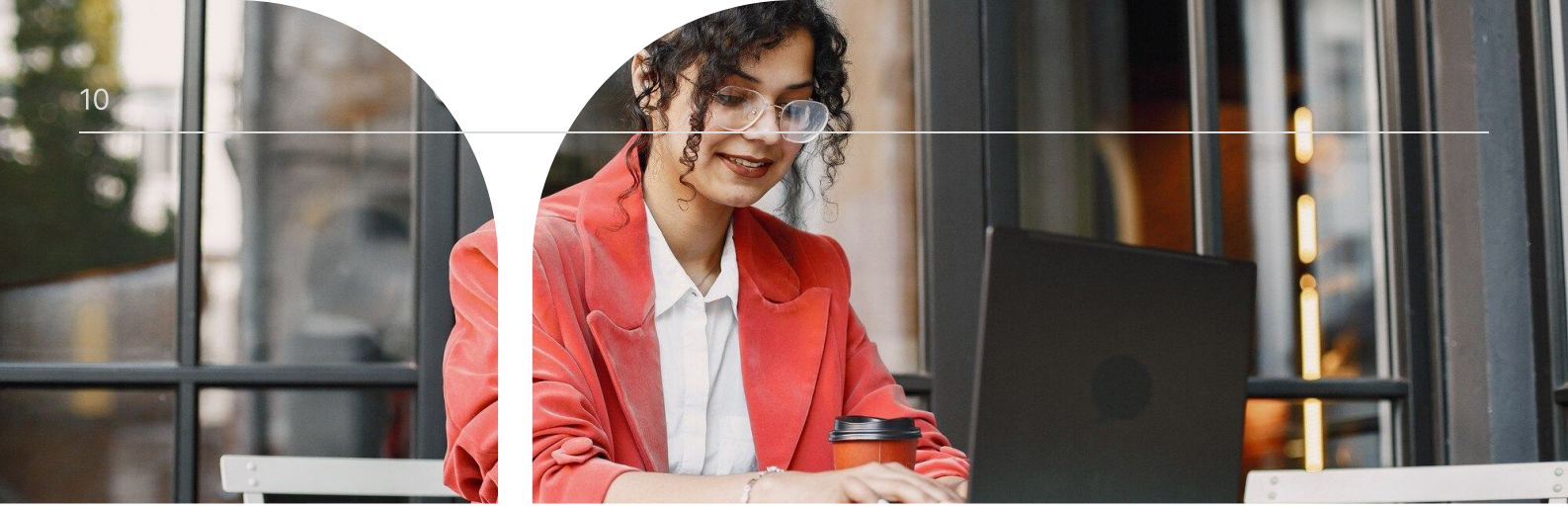
# Professional Data Science and AI Engineering Program

## Overview of the Comprehensive Data Science Engineering Program

The Data Science Engineering program is a rigorous and in-depth four-semester journey designed to transform you into an expert in data science, machine learning, AI, and emerging technologies like Generative AI and Prompt Engineering. This program combines theoretical knowledge, practical experience, and industry-relevant skills to prepare you for advanced roles in the fast-evolving data science landscape.

**Phase 1:** Foundations of Data Science Engineering

**Key Learning Areas:**

- **Python for Data Structures and Algorithms (DSA):** Gain a strong foundation in data structures, algorithms, and problem-solving techniques using Python. Develop proficiency in implementing and optimizing fundamental data structures and algorithms essential for efficient data processing.

- **SQL for Data Querying and Database Management:** Master the basics of SQL for querying and managing relational databases. Learn how to efficiently handle data management tasks and prepare for more complex data operations in subsequent semesters.

**Outcomes:**

- Solid understanding of Python and SQL, which are crucial for data

manipulation, analysis, and database management. Strong problem-solving skills and the ability to work with foundational data structures and algorithms.

## Semester 2: Advanced Data Analytics and Business Intelligence

**Key Learning Areas:**

- **Statistics:** Deepen your understanding of statistical analysis, from foundational concepts like probability and hypothesis testing to advanced regression analysis and inferential statistics.

- **Excel for Data Analysis:** Master advanced Excel functions, data manipulation techniques, and automation through VBA to efficiently manage and analyze large datasets.

- **SQL for Data Querying:** Enhance your SQL skills, focusing on complex queries, data aggregation, and relational database management.

- **Power BI for Data Visualization and Business Intelligence:** Develop expertise in Power BI to create dynamic, interactive dashboards that transform data into actionable insights, supporting strategic decision-making.

**Outcomes:**

- Advanced skills in statistical analysis, data visualization, and business intelligence.

- Ability to analyze complex datasets and present findings effectively to support data-driven business strategies.

## Semester 3: Comprehensive Machine Learning, Deep Learning, and MLOps

**Key Learning Areas:**

- **Machine Learning:** Build expertise in supervised and unsupervised learning algorithms. Learn to preprocess data, conduct exploratory data analysis, and apply machine learning techniques such as regression, classification, clustering, and dimensionality reduction.

- **Deep Learning:** Dive into deep learning by mastering neural networks, convolutional neural networks (CNNs), and recurrent neural networks (RNNs). Specialize in applications like computer vision or natural language processing (NLP).

- **MLOps:** Learn to operationalize machine learning models with a focus on continuous integration, deployment (CI/CD), scalability, and security. Develop the skills needed to manage the full lifecycle of machine learning models in production.

**Outcomes:**

- Expertise in machine learning and deep learning, with the ability to apply these techniques to solve real-world problems.

- Proficiency in MLOps, enabling you to deploy, monitor, and maintain machine learning models effectively in production environments.



**Semester 3:** Generative AI, Prompt Engineering, and Capstone Project

**Key Learning Areas:**

- **Generative AI:** Explore the cutting-edge field of Generative AI, including models like GANs (Generative Adversarial Networks) and transformers. Learn to create AI models that generate new content, such as text, images, and music, pushing the boundaries of traditional machine learning applications.

- **Prompt Engineering:** Develop specialized skills in Prompt Engineering, focusing on optimizing interactions with advanced language models like GPT. Learn techniques to design effective prompts, improve model outputs, and

- harness AI capabilities in natural language processing tasks.

- **Capstone Project:** Integrate all the knowledge and skills from the previous semesters by tackling a comprehensive, real-world project. The capstone project involves solving complex data science and AI problems, applying advanced techniques, and operationalizing models using MLOps principles.

**Outcomes:**

- Advanced proficiency in Generative AI and Prompt Engineering, positioning you at the forefront of AI innovation.

- Practical experience in solving industry-specific data science and AI problems through the capstone project, demonstrating your ability to handle large-scale data science initiatives.
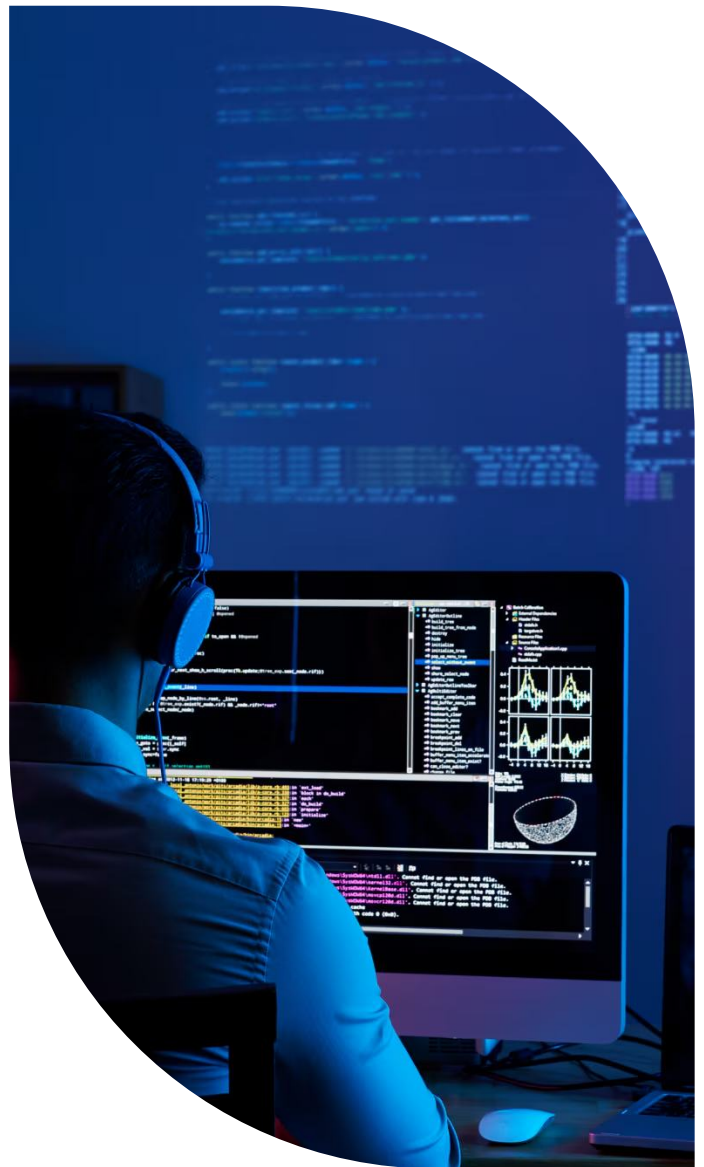
# Program Benefits, Effectiveness, and Career Opportunities

**Benefits:**

- **Comprehensive Skill Development:** TThe program covers a broad spectrum of topics, from foundational programming and statistical analysis to advanced machine learning, deep learning, Generative AI, and Prompt Engineering. This holistic approach ensures you gain a well-rounded and cutting-edge skill set.

- **Practical, Hands-On Experience:** The inclusion of capstone projects and real-world applications provides practical experience, preparing you for the challenges and demands of the industry.

- **Specialization and Flexibility:** The program allows you to specialize in areas such as computer vision, NLP, or Generative AI, tailoring your learning to your career goals and emerging industry trends.
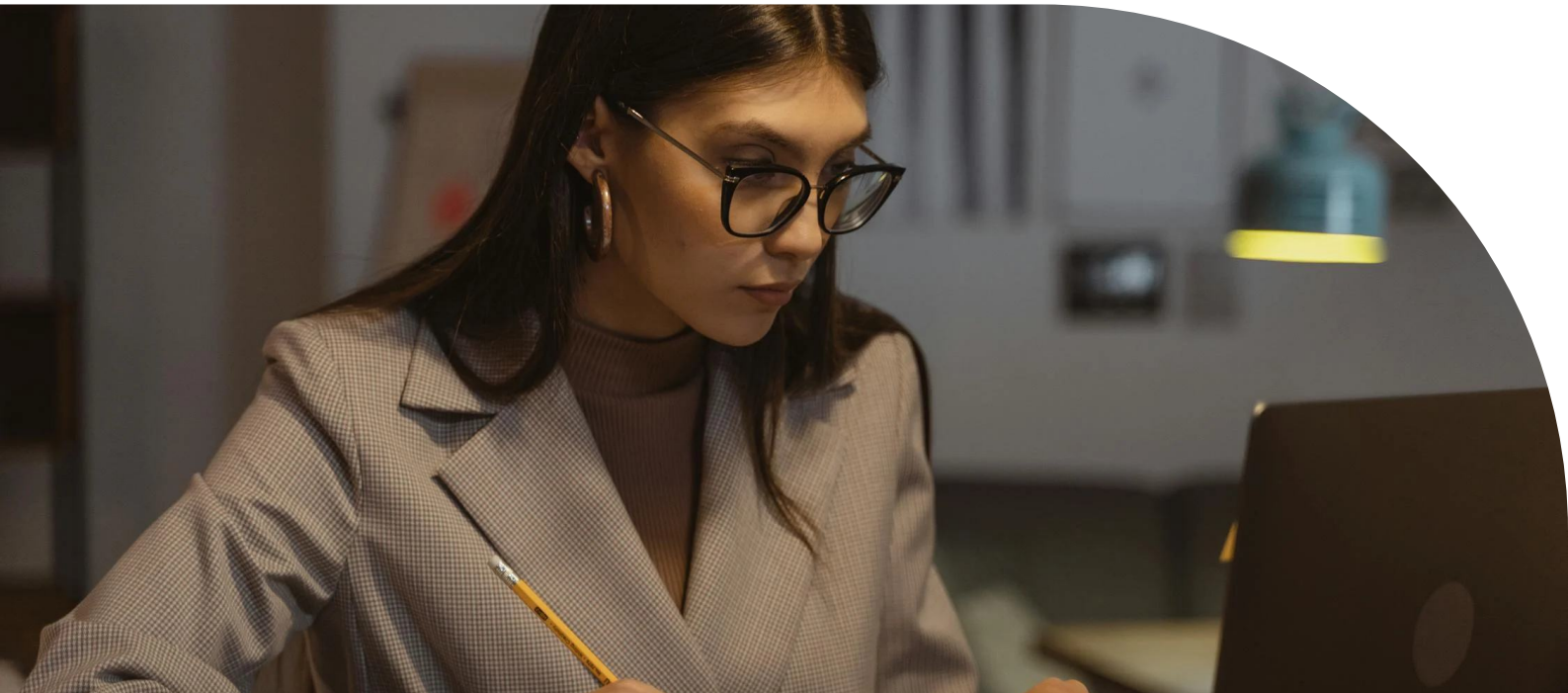


## Effectiveness:

- **Industry-Relevant Curriculum:** The curriculum is designed to meet the current and future demands of the data science and AI industries, ensuring you are learning the most relevant and in-demand skills.

- **Project-Based Learning:** The emphasis on hands-on projects throughout the program bridges the gap between theoretical knowledge and practical application, making you job-ready upon graduation.

- **Integration of Advanced AI Techniques:** The inclusion of Generative AI and Prompt Engineering prepares you for the next wave of AI advancements, giving you a competitive edge in the job market.

## Career Opportunities:

- **Data Scientist:** Utilize your skills in Python, SQL, machine learning, and statistics to analyze data, build models, and provide actionable insights for businesses.

- **Machine Learning Engineer:** Focus on developing, deploying, and maintaining machine learning models, leveraging your expertise in both ML algorithms

- and MLOps practices.

- **AI Specialist:** Specialize in emerging AI technologies like Generative AI and Prompt Engineering, working on cutting-edge projects that push the boundaries of what AI can achieve.

- **MLOps Engineer:** Manage the lifecycle of machine learning models in production, ensuring they are scalable, secure, and continuously integrated into business processes.

- **Business Intelligence Analyst:** Use your expertise in Power BI, Excel, and SQL to create data visualizations and reports that support strategic decision-making in businesses.

**Conclusion**

The Data Science Engineering program offers a comprehensive and in-depth learning experienc that takes you from foundational knowledge to advanced expertise in data science, machine learning and AI. By the end of the program, you will be equipped with the technical skills, practical experience, and cutting-edge knowledge necessary to excel in a variety of high-demand roles in the data science and AI fields. Whether you aspire to be a data scientist, machine learning engineer, AI specialist, o MLOps engineer, this program provides the comprehensive training needed to achieve your career goals and stay ahead in the ever-evolving landscape of technology and innovation.

The **Professional Data Science and AI Engineering Program** is a rigorous 10-month course designed to provide a comprehensive education in data science and artificial intelligence. This program is tailored to equip students with essential skills and knowledge, ranging from foundational concepts to advanced applications, preparing them for various roles in the industry.

**Curriculum Breakdown:**

**1. Programming Foundations and Core Skills**

- **Python for Data Science:** Master Python programming with a focus on libraries and tools essential for data science, including NumPy, Pandas, and Scikit-learn.

- **Data Structures and Algorithms (DSA):** Develop a deep understanding of core data structures (e.g., arrays, linked lists, trees) and algorithms (e.g.,

- sorting, searching) crucial for efficient problem-solving.

- **Dynamic Programming:** Learn and apply dynamic programming techniques to solve complex problems by breaking them into manageable subproblems.

- **Pseudo-coding:** Enhance your problem-solving skills through pseudo-coding, which helps in designing algorithms and planning solutions before implementation.

- **SQL for Data Querying and Database Management:** Gain expertise in writing complex SQL queries, managing databases, and performing data aggregation and transformation.

## 2. Foundations of Data Science

- **Statistics:** Dive into both basic and advanced statistical techniques, including probability distributions, hypothesis testing, and regression analysis.

- **Calculus and Linear Algebra:** Acquire fundamental concepts in calculus (e.g., differentiation, integration) and linear algebra (e.g., matrix operations, vector spaces) necessary for understanding machine learning algorithms.

- **Data Science Tools:** Get introduced to essential tools and environments used in data science, such as Jupyter Notebooks and version control systems.

## 3. Advanced Data Analytics and Business Intelligence

- **Advanced Statistics:** Explore sophisticated statistical techniques and their practical applications for in-depth data analysis.

- **Excel for Data Analysis:** Learn advanced Excel functions, data manipulation, and automation to enhance data analysis capabilities.

- **POWER BI for Data Visualization:** Develop skills in creating dynamic dashboards and visualizations to effectively communicate data insights.

- **Hands-on Business Analytics:** Apply your skills to real-world case studies and projects, gaining practical experience in business analytics.

## 4. Machine Learning & MLOps

- **Advanced Machine Learning Projects:** Work on implementing and optimizing various machine learning algorithms, including supervised, unsupervised, and deep learning techniques.

- **MLOps Integration and Deployment:** Learn about setting up CI/CD pipelines, automating model versioning, and deploying machine learning models into production environments.

- **Capstone Project:** Undertake a comprehensive project that integrates machine learning and MLOps skills to solve real-world problems, culminating in a final presentation.

## 5. Generative AI & Prompt Engineering

- **Generative AI:** Study advanced generative models such as GANs, VAEs, and Flow-based models, and explore their applications in generating images, audio, and text.

- **Prompt Engineering:** Learn to craft and optimize prompts for natural language processing (NLP) models, focusing on practical applications and ethical considerations.

- **Capstone Project:** Apply generative AI and prompt engineering techniques to tackle complex projects, showcasing your ability to handle cutting-edge AI technologies.

## Conclusion

The **Professional Data Science and AI Engineering Program** is meticulously designed to provide a thorough education in data science and AI. It combines programming foundations, mathematical principles, and advanced analytics with practical applications, ensuring students are well-prepared for various roles in the industry.

## Learning and Career Opportunities:

- **Programming and Core Skills:** Establish a strong base in Python, DSA, dynamic programming, pseudo-coding, and SQL, essential for all subsequent learning and applications.

- **Foundations of Data Science:** Prepare for roles such as Data Analyst,

- Business Intelligence Analyst, or Data Consultant with skills in statistics, calculus, and linear algebra.

- **Advanced Analytics and Business Intelligence:** Gain expertise in data analysis and visualization, opening opportunities in business analytics and data consulting roles.

- **Machine Learning and MLOps:** Focus on practical machine learning and operationalizing models, leading to roles like Machine Learning Engineer, Data Scientist, or MLOps Engineer.

- **Generative AI and Prompt Engineering:** Develop specialized skills for roles in Generative AI Research, NLP Engineering, and AI Product Development.

**Overall Benefits:**

- **Holistic Skill Development:** Comprehensive coverage from foundational principles to advanced applications, ensuring a deep understanding of data science.

- **Practical Experience:** Hands-on projects and real-world applications enhance problem-solving skills and prepare students for industry challenges.

- **Career Readiness:** Extensive training across various aspects of data science and AI positions graduates for a wide range of roles in the tech industry, research, and beyond.

## Phase1: Comprehensive Data Structures, Algorithms, and SQL

### Overview

Foundations of Data Science Engineering is designed to provide a strong base in essential programming and database management skills, crucial for advanced data science learning. This semester focuses on Python for Data Structures and Algorithms (DSA) and SQL for Data Querying and Database Management, preparing you for more complex topics in subsequent semesters.

### Key Learning Areas:

- **Python for Data Structures and Algorithms (DSA):**
  - **Foundational Knowledge:** Understand fundamental data structures and

- algorithms, their importance in computer science and data science, and basic problem-solving techniques using Python.

- **Python Programming:** Refresh your knowledge on Python basics, including data types, control flow, functions, and modules.

- **Data Structures:** Learn about lists, stacks, queues, and linked lists, including their operations and implementations. Explore trees, heaps, and hash tables, along with their applications and operations.

- **Algorithms:** Study sorting and searching algorithms, recursion, backtracking, and dynamic programming. Gain hands-on experience solving algorithmic problems and applying data structures in real-world scenarios.

- **Capstone Project:** Implement a comprehensive project incorporating multiple data structures and algorithms, culminating in a practical assessment through problem-solving exercises and project presentation.

- **SQL for Data Querying and Database Management:**

  - **SQL Fundamentals:** Learn SQL syntax for basic operations such as querying, filtering, and aggregating data. Understand data manipulation, constraints, and basic database design concepts.

  - **Advanced SQL Techniques:** Explore subqueries, common table expressions, window functions, and advanced data modeling. Develop skills in query optimization and database management, including indexing, transactions, and concurrency control.

  - **Specialized Topics:** Gain insights into data warehousing, business intelligence, and big data SQL extensions. Understand the integration of SQL with NoSQL databases and other big data tools.

  - **Capstone Project:** Design and implement a database solution, including data warehousing and business intelligence components. Present your findings and insights from the project.

## Objectives

- Develop proficiency in Python programming and data structures, laying a solid foundation for complex data science tasks.

- Gain comprehensive skills in SQL for effective data querying and management.

- Apply programming and database concepts to practical scenarios and real-world problems.

- Prepare for more advanced topics in data science and engineering, setting the stage for specialization and complex data analysis.
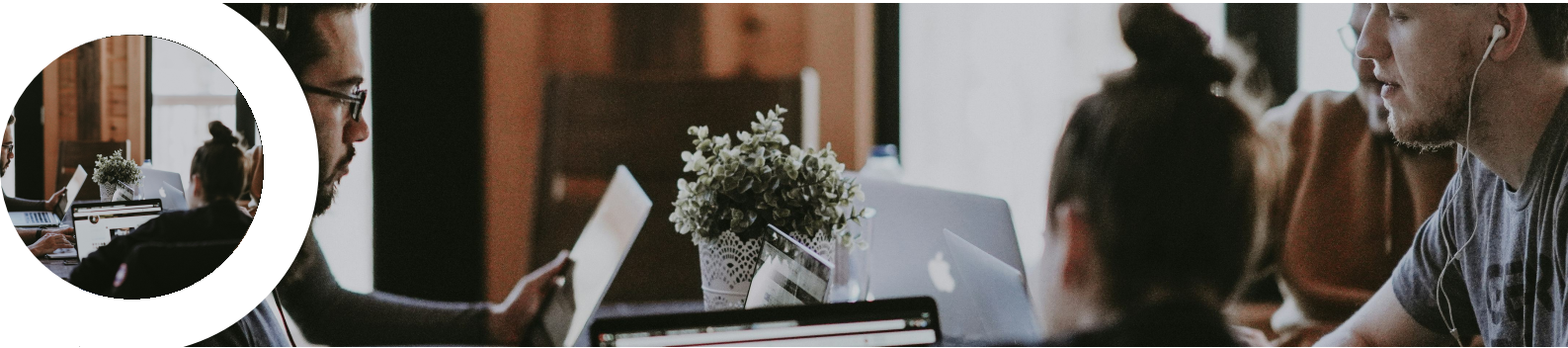


# Module 1: Python for Data Structures and Algorithms (DSA)

**Introduction to Data Structures and Algorithms**

- Overview of Data Structures and Algorithms
- Importance in Computer Science and Data Science
- Introduction to Problem-Solving Techniques
- Python's Role in Implementing Data Structures
- Python Basics Review
  - Data Types and Basic Operations
  - Control Flow (Loops, Conditionals)
  - Functions and Modules

**Basic Data Structures**

- **Lists**
  - Operations: Append, Insert, Remove, Slice
  - List Comprehensions
- **Stacks**
  - Concept and Implementation Using Lists
  - Stack Operations: Push, Pop, Peek

- **Queues**
  - Concept and Implementation Using Lists and collections.deque
  - Queue Operations: Enqueue, Dequeue, Peek
- **Linked Lists**
  - Singly Linked Lists: Nodes and Pointers
  - Operations: Insertion, Deletion, Traversal

**Intermediate Data Structures**

- **Trees**
  - Binary Trees: Structure and Traversal (In-Order, Pre-Order, Post-Order)
  - Binary Search Trees (BST): Insertion, Search, Deletion
  - AVL Trees: Self-Balancing and Rotations
- **Heaps**
  - Concept of Heaps and Heap Operations
  - Implementation Using Lists
  - Heap Sort Algorithm
- **Hash Tables**
  - Concept and Hash Functions
  - Handling Collisions: Chaining and Open Addressing
  - Implementation Using Python Dictionaries

**Algorithms Basics**

- **Sorting Algorithms**
  - Bubble Sort, Selection Sort, Insertion Sort
  - Merge Sort and Quick Sort
  - Comparison and Performance Analysis
- **Searching Algorithms**
  - Linear Search
  - Binary Search
  - Performance Comparison

- **Recursion and Backtracking**
  - Concept of Recursion and Base Cases
  - Common Problems: Factorial, Fibonacci Sequence
  - Backtracking Algorithms: N-Queens, Subset Sum

## Advanced Data Structures and Dynamic Programming

- **Graphs**
  - Representation: Adjacency Matrix and Adjacency List
  - Traversal Algorithms: Depth-First Search (DFS), Breadth-First Search (BFS)
  - Shortest Path Algorithms: Dijkstra's and Bellman-Ford
- **Tries**
  - Concept and Implementation
  - Operations: Insertion, Search, Deletion
  - Applications in Text Processing
- **Segment Trees and Fenwick Trees**
  - Concept and Applications
  - Implementation and Use Cases
- **Dynamic Programming**
  - Introduction to Dynamic Programming
  - Key Concepts: Overlapping Subproblems and Optimal Substructure
  - Common Problems: Knapsack Problem, Longest Common Subsequence, Longest Increasing
  - Subsequence, Matrix Chain Multiplication
  - Top-Down and Bottom-Up Approaches

## Problem Solving and Application

- **Algorithmic Problem Solving**
  - Approach to Solving Algorithmic Problems
  - Common Problem-Solving Techniques: Divide

- and Conquer, Dynamic Programming
- **Hands-On Projects**
  - Implementing Data Structures in Real-World Scenarios
  - Solving Complex Problems Using Learned Algorithms
  - Examples: Text Search Algorithms, Data Processing Tasks
- **Capstone Project**
  - Comprehensive Project Incorporating Multiple Data Structures and Algorithms
  - Real-World Problem Scenario and Solution Implementation
- **Final Assessment**
  - Review and Practical Test of Concepts Learned
  - Problem-Solving Exercises and Coding Challenges
  - Project Presentation
  - Peer Review and Feedback Session

## Module 2: SQL for Data Querying and Database Management

**Basic Level**

- **Introduction to SQL**
  - Overview and Importance of SQL
  - Basic SQL Syntax: SELECT, INSERT, UPDATE, DELETE
- **Retrieving Data with SQL**
  - Writing Simple Queries: Filtering and Sorting Results
  - Using Aggregate Functions: COUNT, SUM, AVG, MIN, MAX
  - Working with Multiple Tables: Joins (INNER, LEFT, RIGHT)
- **Data Manipulation with SQL**
  - Inserting, Updating, and Deleting Data
  - Data Types and Constraints
  - Query Optimization Techniques

## Intermediate Level

- **Advanced Query Techniques**
  - Subqueries and Nested Queries
  - Common Table Expressions (CTEs) and Recursive Queries
  - Window Functions: ROW_NUMBER, RANK, LAG, LEAD
- **Data Modeling and Database Design**
  - Relational Database Concepts: Tables, Relationships, Keys
  - Creating and Managing Tables: Constraints, Indexes
  - Normalization and Denormalization
- **Transactions and Concurrency Control**
  - ACID Properties
  - Managing Transactions: COMMIT, ROLLBACK, SAVEPOINT
  - Concurrency Issues and Locking Mechanisms

## Advanced Level

- **Performance Tuning and Optimization**
  - Indexing Strategies: Clustered, Non-clustered
  - Query Execution Plans
  - SQL Profilers and Monitoring Tools
- **Advanced Database Administration**
  - Backup and Recovery Strategies
  - Security and Access Control: Users, Roles, Permissions
  - Database Maintenance
- **Advanced Topics in SQL**
  - Stored Procedures and Functions
  - Triggers and Events
  - Dynamic SQL and Parameterized Queries

## Expert Level

- **Data Warehousing and Business Intelligence**

- Designing Data Warehouses: Star Schema, Snowflake Schema
- ETL Processes
- OLAP and Data Cubes

- **Big Data SQL**
  - Introduction to NoSQL Databases: MongoDB, Cassandra
  - SQL Extensions for Big Data: Hive, Presto
  - Polyglot Persistence

- **Capstone Project**
  - Designing and Implementing a Database Solution
  - Presentation of Findings and Insights

**Benefits to Learners:**

This combined curriculum of Python for Data Structures and Algorithms (DSA) and SQL for Data Querying and Database Management is designed to equip learners with a strong foundation in two critical areas of data science: computational efficiency and data management.

1. **Foundational Knowledge in Programming and Data Structures:**
   - By mastering Python and various data structures, learners gain the ability to write efficient code that can handle complex data manipulation tasks. This is essential for optimizing data processing workflows in real-world data science projects.

2. **Expertise in Algorithms:**
   - Understanding algorithms and their implementation in Python allows learners to solve problems efficiently, which is a vital skill in data science, especially in areas like machine learning, where algorithm performance can significantly impact model effectiveness.

3. **Proficiency in SQL and Database Management:**
   - SQL is the backbone of data manipulation in the industry. Mastering SQL allows learners to efficiently query, update, and manage databases, a skill that is crucial for handling large datasets in data science.

4. **Practical Application and Problem-Solving:**
   - The curriculum emphasizes hands-on projects and real-world scenarios,

- preparing learners to apply their skills directly in the industry. This practical experience is invaluable for a career in data science, where theoretical knowledge must be coupled with the ability to solve actual business problems.

5. **Career-Ready Skills**

- By the end of this semester, learners will have developed a comprehensive skill set that makes them ready for roles such as data scientists, data engineers, and database administrators. These roles require both the ability to implement complex algorithms and manage data effectively, both of which are covered in this curriculum.

# Phase2: Advanced Data Analysis & Business Intelligence

## Overview

Advanced Data Analytics and Business Intelligence focuses on elevating your data science skills by integrating advanced statistical analysis, data visualization, and business intelligence techniques. This is designed to build on the foundational knowledge you acquired in Comprehensive Data Structures, Algorithms, and SQL, enhancing your ability to perform complex data analysis and make data-driven decisions in real-world business environments.

## Key Learning Areas:

- **Statistics from Basic to Expert Level:**

  - Deepen your understanding of statistical concepts, from foundational principles to advanced techniques. You will explore probability distributions, hypothesis testing, regression analysis, and more, gaining the skills needed to conduct rigorous data analysis.

- **Excel for Data Analysis (Basic to Advanced):**

  - Master Excel as a powerful tool for data management and analysis. Learn advanced Excel functions, data manipulation techniques, and automation through macros and VBA to efficiently handle large datasets.

- **POWER BI for Data Visualization and Business Intelligence (Basic to Advanced):**

  - Develop expertise in Power BI to create dynamic, interactive dashboards and visualizations. Learn how to transform data into actionable insights,

- supporting strategic decision-making processes in business contexts.
- **Hands-on Business Analytics & Product Analytics:**
  - Apply your analytical skills to real-world business scenarios. Engage in case studies and projects that simulate industry challenges, enabling you to practice solving problems and making data-driven decisions that impact business outcomes.

**Objectives**

- Equip you with advanced analytical and technical skills necessary to thrive in data-driven roles.
- Enhance your ability to visualize and communicate data insights through powerful tools like Power BI.
- Provide practical, hands-on experience in applying data science concepts to solve real business problems.
- Prepare you for a range of career opportunities in data analytics, business intelligence, and related fields.

**Advanced Data Analysis & Business Intelligence** is designed to ensure that by the end of the program, you are not only proficient in advanced data analysis techniques but also capable of using these skills to influence and drive business strategies, making you a valuable asset in any organization.

## Module 1: Statistics from Basic to Expert Level

**Basic Level**

- **Introduction to Statistics**
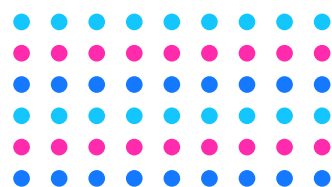
- Overview of Descriptive and Inferential Statistics
- Types of Data: Nominal, Ordinal, Interval, Ratio
- Data Collection Methods: Surveys, Experiments, Observations

- **Descriptive Statistics**
  - Measures of Central Tendency: Mean, Median, Mode
  - Measures of Dispersion: Range, Variance, Standard Deviation
  - Data Visualization: Histograms, Box Plots, Scatter Plots

- **Probability Basics**
  - Introduction to Probability Theory
  - Rules of Probability: Addition and Multiplication Rules
  - Conditional Probability and Bayes' Theorem
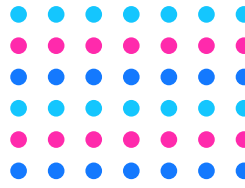
## Intermediate Level

- **Probability Distributions**
  - Discrete Distributions: Binomial, Poisson
  - Continuous Distributions: Normal, t-distribution, Chi-square
  - Understanding and Applying Z-scores

- **Inferential Statistics**
  - Sampling Techniques and Sampling Distributions
  - Central Limit Theorem
  - Confidence Intervals for Means and Proportions

- **Hypothesis Testing**
  - Formulating Null and Alternative Hypotheses
  - Type I and Type II Errors
  - Performing t-tests, Chi-square Tests, ANOVA

## Advanced Level

- **Correlation and Regression Analysis**
  - Pearson and Spearman Correlation Coefficients

- Simple Linear Regression: Model Building and Interpretation
- Multiple Regression Analysis: Assumptions and Diagnostics
- **Advanced Hypothesis Testing**
  - Non-parametric Tests: Mann-Whitney U, Wilcoxon Signed-Rank
  - Advanced ANOVA: Two-way ANOVA, Repeated Measures ANOVA
  - Statistical Power and Sample Size Determination
- **Multivariate Statistics**
  - Principal Component Analysis (PCA)
  - Factor Analysis
  - Cluster Analysis

**Expert Level**

- **Time Series Analysis**
  - Decomposition of Time Series: Trend, Seasonality, Residuals
  - Autoregressive Integrated Moving Average (ARIMA) Models
  - Exponential Smoothing Techniques
- **Bayesian Statistics**
  - Introduction to Bayesian Inference
  - Prior, Likelihood, Posterior Distribution
  - Applications of Bayesian Methods
- **Advanced Statistical Modeling**
  - Generalized Linear Models (GLM)
  - Survival Analysis
  - Logistic Regression and Odds Ratio

# Module 2: Excel for Data Analysis

**Basic Level**

- **Getting Started with Excel**
  - Introduction to Excel Interface
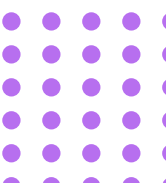
- Workbook and Worksheet Management
- Ribbon, Tabs, and Menus Navigation
- Cells, Rows, and Columns Basics
- **Basic Operations**
  - Creating and Saving Workbooks
  - Entering Data: Text, Numbers, and Dates
  - Simple Formatting: Fonts, Colors, and Cell Styles
- **Basic Formulas and Functions**
  - Using SUM, AVERAGE, COUNT Functions
  - Basic Arithmetic Operations (Addition, Subtraction, Multiplication, Division)
  - Utilizing the Formula Bar
- **Data Management**
  - Sorting and Filtering Data
  - Basic Charts Creation (Column, Line, Pie)

**Intermediate Level**

- **Enhancing Data Management**
  - Advanced Formatting Techniques
  - Conditional Formatting
  - Custom Cell Styles and Themes
- **Intermediate Formulas and Functions**
  - Implementing IF Statements
  - Utilizing VLOOKUP and HLOOKUP
  - Text Functions (CONCATENATE, LEFT, RIGHT)
- **Data Analysis Tools**
  - Creating and Using Pivot Tables and Pivot Charts
  - Data Validation Techniques
  - Grouping and Subtotals
- **Visualization**
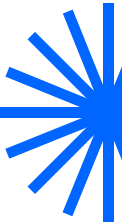
- Creating and Customizing Charts
- Using Sparklines
- Implementing Slicers with PivotTables

**Advanced Level**

- **Advanced Data Analysis**
  - Complex Formulas and Functions
  - Using INDEX and MATCH
  - Advanced Lookup Functions (XLOOKUP)
  - Array Formulas
- **Data Modeling and Power Query**
  - Introduction to Power Query
  - Data Transformation and Loading
  - Merging and Appending Queries
- **Power Pivot and DAX**
  - Creating Data Models
  - Introduction to DAX (Data Analysis Expressions)
  - Measures and KPIs
- **Automation and Macros**
  - Introduction to VBA (Visual Basic for Applications)
  - Recording and Running Macros
  - Editing Macro Code
- **Advanced VBA Techniques**
  - Creating User Forms
  - Error Handling Techniques
  - Interacting with Other Office Applications (Word, Outlook)

**Collaboration and Security**

- **Collaborating with Others**
  - Sharing Workbooks

- - Tracking Changes and Comments
  - Co-Authoring in Real Time
- **Protecting Your Work**
  - Password Protecting Workbooks and Worksheets
  - Setting Permissions
  - Data Encryption
- **Tips and Tricks**
  - Keyboard Shortcuts for Efficiency
  - Customizing the Quick Access Toolbar
  - Best Practices for Organizing Your Workbook and Maintaining Data Integrity

**Phase 1:** Hands-on Business Analytics & Product Analytics

**Industrial-Level Case Study: Business Analytics & Product Analytics**

- **Scenario Overview**
  - Analyze and optimize product offerings and business strategies for a retail company.
- **Tasks:**
  - **Data Exploration and Descriptive Analysis**
    - Analyze historical sales data for trends and seasonality.
    - Use descriptive statistics and visualizations to understand product performance and customer behavior.
  - **Hypothesis Testing and A/B Testing**
    - Formulate and validate hypotheses about customer preferences and product features.
  - **Predictive Modeling and Forecasting**
    - Build regression models to predict sales and forecast demand and inventory.
  - **Dashboard Creation and Reporting**

- Develop interactive dashboards to monitor KPIs using Excel.
- **Automation and Efficiency**
  - Automate tasks using Excel macros and VBA scripts.

# Module 3: POWER BI for Data Visualization and Business Intelligence

**Basic Level**

- **Introduction to POWER BI**
  - Overview and Importance in Business Intelligence
  - Understanding POWER BI Desktop and Service
- **Getting Started with POWER BI**
  - Connecting to Data Sources: Excel, SQL Server, CSV, Web Data
  - Importing and Transforming Data using Power Query Editor
  - Creating Basic Visualizations: Bar, Line, Pie Charts
- **Basic Data Modeling and Relationships**
  - Introduction to Data Modeling
  - Creating Relationships between Tables
  - Managing Data Models: Measures and Calculated Columns

**Intermediate Level**

- **Advanced Data Transformations**
  - Data Cleansing and Advanced Query Editor Techniques
  - Merging and Appending Queries
  - Working with Parameters and Functions
- **Advanced Visualizations**
  - Customizing Visualizations: Formatting, Themes, Templates
  - Using Custom Visuals from AppSource
  - Interactive Reports and Dashboards: Drill-down, Drill-through, Bookmarks
- **DAX (Data Analysis Expressions) Basics**
  - Introduction to DAX: Calculated Columns, Measures

- DAX Functions: Aggregates, Logical, Date and Time, Statistical

**Advanced Level**

- **Advanced DAX Techniques**
  - Using Variables in DAX
  - Time Intelligence Functions: TOTALYTD, SAMEPERIODLASTYEAR
  - Advanced DAX Patterns
- **Data Modeling Optimization**
  - Optimizing Data Models for Performance
  - Role-playing Dimensions and Hierarchies
  - Handling Many-to-Many Relationships
- **POWER BI Service and Collaboration**
  - Publishing Reports and Dashboards to POWER BI Service
  - Setting up Data Gateways
  - Sharing and Collaboration: Workspaces, Sharing Reports

**Expert Level**

- **Advanced Data Analysis with POWER BI**
  - Advanced Analytics: Forecasting, What-If Analysis
  - Integrating Machine Learning Models: Azure Machine Learning, R, Python
  - Real-time Data Streaming and Dashboarding
- **Governance and Security**
  - Implementing Row-level Security (RLS)
  - Data Encryption and Compliance
  - Managing and Auditing Data Access
- **Capstone Project**
  - Designing and Implementing a Business Intelligence Solution
  - Presentation of Findings and Insights

## Conclusion

Congratulations on completing the second semester of our Data Science Engineering program! You've mastered statistical concepts from basic to expert levels, developed advanced skills in Excel and Power BI, and applied business and product analytics techniques through hands-on projects. These competencies will empower you to excel in data-driven decision-making and analysis, significantly enhancing your data science career prospects.

## Benefits of Completing Advanced Data Analysis & Business Intelligence

**Comprehensive Data Structures, Algorithms, and SQL (Phase 1) & Advanced Data Analysis & Business Intelligence (Phase 2)** In **Phase 1,** you laid the foundation for your data science journey by mastering key concepts in probability, statistics, calculus, and linear algebra. These core mathematical and analytical skills are essential for understanding complex data structures, algorithms, and statistical models.

In **Phase 2**, you build on this foundation by diving deep into advanced statistical techniques, data analysis, and business intelligence tools. This semester's focus on practical applications, using Excel and Power BI, equips you with hands-on experience in analyzing and visualizing data, making data- driven decisions, and

By combining the theoretical knowledge from phase 1 with the applied skills from phase 2, you develop a well-rounded expertise in data science. This integrated approach ensures that you not only understand the underlying mathematical principles but also know how to apply them to extract actionable insights from data.

**Benefits of Completing phase 1 and 2**

Completing both semesters positions you as a versatile data science professional with a comprehensive skill set:

1. **Strong Analytical Foundation:**

   - Your deep understanding of statistical and mathematical principles (from phase 1) enables you to accurately interpret and analyze complex datasets.

   - Advanced statistical techniques (from phase 2) enhance your ability to perform rigorous data analysis and hypothesis testing.

2. **Practical Data Analysis Skills:**

   - Mastery of Excel (from phase 2) allows you to efficiently manage and analyze data, perform complex calculations, and automate repetitive tasks.

   - Proficiency in Power BI (from phase 2) enables you to create dynamic and interactive visualizations, dashboards, and reports, making it easier to communicate insights and support business decisions.

3. **Business Intelligence and Data-Driven Decision-Making:**

   - You gain the ability to apply business analytics techniques to real-world scenarios, bridging the gap between data science and business strategy.

   - The hands-on experience with case studies in business and product analytics (from phase 2) prepares you to tackle industry-specific challenges, enhancing your problem-solving capabilities.

# Job Opportunities

Upon completing Phase 1 and 2, you will be well-prepared for a variety of roles in the data science and analytics field, including:

1. **Data Analyst:**

- Utilize your advanced Excel skills and statistical knowledge to clean, analyze, and visualize data, providing actionable insights to stakeholders.

## 2. Business Intelligence (BI) Analyst:

- Leverage your Power BI expertise to create and manage dashboards, perform ad-hoc analyses, and deliver reports that drive strategic decisions.

## 3. Statistical Analyst:

- Apply your statistical expertise to perform hypothesis testing, regression analysis, and advanced modeling, supporting research and business initiatives.

## 4. Product Analyst:

- Use your combined skills in statistics, Excel, and Power BI to analyze product performance, customer behavior, and market trends, informing product development and marketing strategies.

## 5. Junior Data Scientist:

- With your solid foundation in mathematics, statistics, and data analysis, begin exploring machine learning algorithms, predictive modeling, and data-driven innovation.

## 6. Market Research Analyst:

- Analyze market conditions, trends, and customer preferences using statistical tools and business analytics to help companies understand market dynamics and identify opportunities.

## Program Benefits and Career Growth

- **Comprehensive Skill Set:** The combination of foundational knowledge and applied skills makes you highly versatile, capable of tackling a wide range of data-driven challenges across industries.

- **Industry-Relevant Tools:** Mastery of widely-used tools like Excel and Power BI ensures you are ready to contribute from day one in most data-centric roles.

- **Increased Employability:** The practical focus on business and product analytics, along with hands-on experience in real-world scenarios, significantly enhances your job prospects in a competitive market.

- **Career Flexibility:** The skills acquired through Phase 1 and 2 are applicable across multiple industries, including finance, healthcare, retail, manufacturing

By the end of Phase 2, you will have a strong and marketable skill set that positions you as a capable and effective data professional, ready to make an impact in the field of data science and analytics.

**Phase 3: Comprehensive Machine Learning, Deep Learning, and MLOps** focuses on advancing your technical expertise in machine learning and deep learning while integrating the principles of MLOps. This semester builds on the knowledge and skills developed in Phase 2, equipping you with advanced techniques and practical experience necessary for tackling complex data science challenges and deploying scalable machine learning solutions.

**Key Learning Areas:**

- **Introduction to Machine Learning:**

  - **Foundational Knowledge:** Expand your understanding of machine learning fundamentals, including data preprocessing, exploratory data analysis (EDA), and feature engineering. Learn to build and evaluate machine learning models using a variety of algorithms.

  - **Supervised Learning Algorithms:** Master techniques for regression and classification, including linear regression, logistic regression, support vector machines (SVM), and ensemble methods such as random forests and gradient boosting.

- **Unsupervised Learning Algorithms and Recommender Systems:**

  - **Clustering and Dimensionality Reduction:** Delve into unsupervised learning methods such as k-means, hierarchical clustering, and principal component analysis (PCA) to uncover hidden patterns in data.

  - **Recommender Systems:** Learn to build recommendation engines using collaborative filtering and content-based filtering techniques, and analyze time series data for forecasting.

- **Optimization Algorithms:**

  - **Gradient Descent Variants:** Understand optimization techniques crucial for training machine learning models, including stochastic gradient descent (SGD) and advanced methods like Adam and RMSprop.

  - **Advanced Optimization Techniques:** Explore additional methods such as evolutionary algorithms and simulated annealing to refine model performance.

- **Deep Learning Algorithms:**
  - **Neural Networks and Architectures:** Gain expertise in designing and training deep learning models, including feedforward neural networks (FNN), convolutional neural networks (CNN), and recurrent neural networks (RNN).
  - **Specialization:** Choose between specialization in computer vision or natural language processing (NLP), learning advanced techniques and applications relevant to each field.
- **MLOps:**
  - **Operationalizing ML Models:** Learn to manage the machine learning lifecycle, including continuous integration and deployment (CI/CD), model versioning, and monitoring.
  - **Scalability and Security:** Understand best practices for scaling machine learning infrastructure and ensuring security, as well as tools and techniques for automating and optimizing ML workflows.

**Objectives**

- **Advanced Technical Proficiency:** Develop advanced skills in machine learning and deep learning, with a focus on applying these techniques to complex data problems and real-world scenarios.
- **Specialization:** Gain in-depth knowledge in either computer vision or NLP, tailoring your expertise to meet industry demands and career interests.
- **Operational Excellence:** Learn the principles and practices of MLOps to manage, deploy, and maintain machine learning models effectively in production environments.
- **Practical Experience:** Engage in hands-on projects and case studies to apply your skills to industry-specific challenges, preparing you for impactful roles in data science and machine learning.

Phase 3 aims to enhance your capabilities in advanced data science techniques and operational practices, ensuring you are well-prepared to tackle sophisticated problems, lead machine learning initiatives, and contribute to data-driven decision-making processes in various industries.

## Module 1: Introduction to Machine Learning

- **Overview of Machine Learning**
  - Definition and Types of Machine Learning
  - Importance and Applications in Various Fields
- **Data Preprocessing**
  - Data Cleaning and Transformation
  - Handling Missing Data
  - Feature Scaling and Normalization
- **Additional Topics**
  - Exploratory Data Analysis (EDA)
  - Feature Engineering and Selection

## Module 2: Supervised Learning Algorithms

- **Machine Learning 1: Supervised**
  - Fundamentals and Metrics
  - Maximum Likelihood Estimation (MLE) and Maximum A Posteriori (MAP)
  - Confidence Interval
  - Classification Metrics (Accuracy, Precision, Recall, F1 Score, ROC-AUC)
  - Imbalanced Data Techniques (SMOTE, ADASYN)
- **Algorithms**
  - Linear Regression
  - Polynomial Regression
  - Ridge and Lasso Regression
  - Logistic Regression
  - Support Vector Machines (SVM)

- Decision Trees
- Random Forests
- K-Nearest Neighbors (KNN)
- Naive Bayes
- Bagging Techniques
- **Additional Topics**
  - Model Evaluation and Validation Techniques
  - Hyperparameter Tuning for Supervised Models

## Module 3: Unsupervised Learning Algorithms and Recommender Systems

- **Machine Learning 2: Unsupervised and Recommender Systems**
  - Clustering
    - Introduction to Clustering and k-Means
    - k-Means++
    - Hierarchical Clustering
    - Gaussian Mixture Models (GMM)
  - Dimensionality Reduction
    - Principal Component Analysis (PCA)
    - t-SNE
  - Anomaly, Outlier, and Novelty Detection
  - Recommender Systems
    - Collaborative Filtering
    - Content-Based Filtering
  - Time Series Analysis
    - ARIMA
    - Exponential Smoothing
    - Prophet
- **Additional Topics**
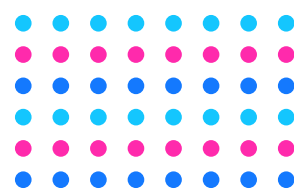
- Advanced Clustering Techniques
- Evaluation Metrics for Recommender Systems

## Module 4: Optimization Algorithms

- **Optimization Algorithms**
  - Gradient Descent Variants
    - Batch Gradient Descent
    - Stochastic Gradient Descent (SGD)
    - Mini-Batch Gradient Descent
    - Adam, RMSprop, Adagrad
  - Advanced Optimization Techniques
    - Conjugate Gradient Methods
    - Evolutionary Algorithms
    - Simulated Annealing
    - BFGS and L-BFGS
- **Additional Topics**
  - Optimization for Neural Networks
  - Practical Considerations in Optimization

## Module 5: Deep Learning Algorithms

- **Fundamentals of Deep Learning**
  - Basics of Neural Networks
    - Perceptrons
    - Hidden Layers
    - Activation Functions and Loss Functions
  - Introduction to TensorFlow
  - Introduction to Keras
  - Forward and Back Propagation

- Multilayer Perceptrons (MLP)
- Callbacks and TensorBoard
- Optimization Techniques in Deep Learning
- Hyperparameter Tuning
- **Additional Topics**
  - Advanced Neural Network Architectures
  - Transfer Learning in Deep Learning

## Module 6: Specialization in Computer Vision (CV) or Natural Language Processing (NLP)

- **Option 1: Computer Vision (CV)**
  - Convolutional Neural Networks (CNN)
    - Architecture and Layers
    - Applications in Image Recognition
  - Advanced CV Techniques
    - Transfer Learning
    - Object Detection (YOLO, SSD)
    - Image Segmentation (U-Net, Mask R-CNN)
- **Option 2: Natural Language Processing (NLP)**
  - Text Preprocessing
    - Tokenization
    - Stopword Removal
    - Lemmatization
  - Text Representation
    - Bag-of-Words
    - TF-IDF
    - Word Embeddings (Word2Vec, GloVe)
  - NLP Applications
    - Named Entity Recognition (NER)

- Sentiment Analysis
- Text Classification (e.g., Spam Detection, Topic Modeling)
- **Additional Topics**
  - Advanced Techniques in Computer Vision or NLP
  - Real-world Applications and Case Studies

## Module 7: MLOps

- **Introduction to MLOps**
  - Overview of MLOps
  - Definition and Importance of MLOps
  - Key Components and Principles
  - Benefits of MLOps in Modern Enterprises
- **Environment Management**
  - Setting Up ML Environments
  - Virtual Environments and Docker
  - Managing Dependencies
  - Introduction to Kubernetes for ML
- **Continuous Integration and Continuous Deployment (CI/CD) for ML**
  - CI/CD Pipelines for Machine Learning
    - Building CI/CD Pipelines
    - Tools and Technologies (Jenkins, GitHub Actions, GitLab CI)
    - Automating Testing and Validation
    - Model Versioning and Management
- **Model Training and Experimentation**
  - Efficient Model Training
  - Distributed Training Techniques
  - Leveraging Cloud Services for Training (AWS, GCP, Azure)
  - Experiment Management
    - Tracking Experiments and Results

- Tools for Experiment Management (Comet, Weights & Biases)
- **Model Deployment and Monitoring**
  - Deploying Machine Learning Models
    - Strategies for Model Deployment (Batch, Real-time, On-device)
    - Serving Models (TensorFlow Serving, FastAPI)
  - Monitoring and Maintenance
    - Monitoring Model Performance
    - Detecting and Handling Model Drift
    - Automating Retraining Pipelines
- **Scalability and Security**
  - Scalability in MLOps
    - Scaling Infrastructure for ML Workloads
    - Best Practices for Scalability
  - Security in MLOps
  - Ensuring Data and Model Security
  - Implementing Secure ML Practices
- **Advanced Topics in MLOps**
  - Continuous Monitoring and Automated Feedback Loops
  - Advanced Model Optimization
  - Cost Management in MLOps

**Benefits of the 3rd Semester for Learners**

The 3rd semester of your Data Science Engineering program is a critical phase where students transition from foundational knowledge to advanced, specialized skills in machine learning, deep learning, and MLOps. This semester equips learners with a comprehensive set of tools and techniques to handle complex data science challenges. Here's how it will benefit them:

1. **Advanced Machine Learning and Deep Learning Skills**

- **Deepened Knowledge:** Students will gain a deep understanding of both supervised and unsupervised learning algorithms. By learning about advanced techniques such as Gradient Descent Variants, optimization

- optimization methods, and deep learning architectures (e.g., CNNs, RNNs), they will be well-prepared to tackle complex problems in various domains.

- **Practical Experience:** Through hands-on projects and case studies, learners will apply machine learning models to real-world data, enhancing their practical skills and preparing them for industry challenges.

## 2. Specialization Opportunities

- **Focused Learning:** The option to specialize in Computer Vision (CV) or Natural Language Processing (NLP) allows students to delve deeply into a field of their choice, providing them with the expertise needed to excel in specific areas of data science.

- **Industry-Relevant Skills:** Specializing in CV or NLP aligns with current industry demands, making learners more competitive candidates for roles requiring expertise in these advanced areas.

## 3. Mastery of MLOps

- **Operational Excellence:** Understanding MLOps will enable students to manage the entire lifecycle of machine learning models, from development to deployment. They will learn about CI/CD pipelines, model versioning, and automated testing, which are essential for maintaining and scaling machine learning systems in production environments.

- **Efficiency and Scalability:** Knowledge in MLOps will help students build scalable and secure machine learning solutions, ensuring that models are efficiently managed and monitored over time.

## 4. Real-World Application

- **Capstone Projects:** The semester includes a capstone project that integrates machine learning, deep learning, and MLOps concepts. This project will allow students to apply their skills to real- world problems, providing them with valuable experience and a portfolio piece that showcases their abilities to potential employers.

- **Industry Readiness:** Exposure to industry-specific projects in sectors like retail, consumer banking, and supply chain ensures that students are prepared for the practical challenges they will face in their careers.

## 5. Comprehensive Skill Integration

- Holistic Approach: By combining modules on machine learning algorithms, optimization techniques, deep learning, and MLOps, students will develop a well-rounded skill set. This integration prepares them for a variety of roles in data science and machine learning, enhancing their versatility and adaptability in the field.

## 6. Career Advancement

- **Job Opportunities:** The advanced skills and specializations gained during this semester open doors to a wide range of job opportunities, including roles such as Machine Learning Engineer, Data Scientist, MLOps Engineer, and AI Specialist.

- **Competitive Edge:** Mastery of these advanced topics provides a significant competitive edge in the job market, positioning learners as highly qualified candidates for top positions in data science and artificial intelligence.

By the end of the Phase 3, students will be equipped with cutting-edge skills and practical experience, making them well-prepared for challenging and rewarding careers in data science and machine learning.

# Final Phase : Generative AI & Prompt Engineering

**Overview:**

The final semester focuses on advanced topics in Generative AI and Prompt Engineering. This comprehensive module will deepen your knowledge and skills in creating and fine-tuning generative models and optimizing prompt strategies for natural language processing. You will apply these techniques to practical projects, preparing you for advanced roles in AI research and development.

**Key Learning Areas:**

- **Module 1: Introduction to Generative AI**
  - **Overview of Generative AI:**
    - History and evolution of generative models.
    - Types of generative models: GANs, VAEs, and Flow-based Models.
    - Applications across various domains.

- **Foundations:**
  - Generative Adversarial Networks (GANs): Basic principles, architecture, and training.
  - Variational Autoencoders (VAEs): Basics and variations.
  - Flow-based Models: Introduction to RealNVP and Glow.
- **Module 2: Advanced Generative Models**
  - **Deep Dive into GANs:**
    - Architecture: Generator, Discriminator, and loss functions.
    - GAN Variants: DCGANs, CycleGANs, and StyleGANs.
    - Training challenges and stabilization techniques.
  - **Variational Autoencoders (VAEs):**
    - VAE Architecture: Encoder, Decoder, and Latent Space.
    - Variants: Conditional VAEs, Beta-VAEs, and their applications.
  - **Flow-based Models:**
    - Principles behind normalizing flows.
    - RealNVP and Glow: Implementation and applications.
- **Module 3: Natural Language Generation**
  - **Introduction to NLP Generative Models:**
    - Overview of language models and sequence-to-sequence models.
    - Transformers: Architecture, attention mechanisms.
  - **Pre-trained Models:**
    - BERT, GPT, T5: Overview, fine-tuning, and transfer learning.
- **Module 4: Computer Vision and Image Generation**
  - **Image Generation with GANs:**
    - Techniques and architectures for high-quality image generation.
    - Style transfer and super-resolution.
    - Advanced topics: Object Detection (YOLO, SSD) and Image Segmentation (U-Net, Mask R-CNN).
- **Module 5: Music and Audio Generation**
  - **Generative Models for Audio:**

- **Foundations:**
  - Generative Adversarial Networks (GANs): Basic principles, architecture, and training.
  - Variational Autoencoders (VAEs): Basics and variations.
  - Flow-based Models: Introduction to RealNVP and Glow.
- **Module 2: Advanced Generative Models**
  - **Deep Dive into GANs:**
    - Architecture: Generator, Discriminator, and loss functions.
    - GAN Variants: DCGANs, CycleGANs, and StyleGANs.
    - Training challenges and stabilization techniques.
  - **Variational Autoencoders (VAEs):**
    - VAE Architecture: Encoder, Decoder, and Latent Space.
    - Variants: Conditional VAEs, Beta-VAEs, and their applications.
  - **Flow-based Models:**
    - Principles behind normalizing flows.
    - RealNVP and Glow: Implementation and applications.
- **Module 3: Natural Language Generation**
  - **Introduction to NLP Generative Models:**
    - Overview of language models and sequence-to-sequence models.
    - Transformers: Architecture, attention mechanisms.
  - **Pre-trained Models:**
    - BERT, GPT, T5: Overview, fine-tuning, and transfer learning.
- **Module 4: Computer Vision and Image Generation**
  - **Image Generation with GANs:**
    - Techniques and architectures for high-quality image generation.
    - Style transfer and super-resolution.
    - Advanced topics: Object Detection (YOLO, SSD) and Image Segmentation (U-Net, Mask R-CNN).
- **Module 5: Music and Audio Generation**
  - **Generative Models for Audio:**

- Introduction and implementation of WaveNet and WaveRNN.
- Music generation techniques using RNNs and Transformers.
- Real-world applications in music and audio synthesis.

**Deliverables:**

- **Project Report:** Summarize findings, methodologies, and results from each module.
- **Presentation:** Prepare and deliver a presentation on a Generative AI project.

**Prompt Engineering:**

- **Module 6: Introduction to Prompt Engineering**
  - **Overview of Prompt Engineering:**
    - Definition, importance, and types of prompts.
  - **Prompt Design:**
    - Effective prompting techniques and optimization strategies.
- **Module 7: Fundamentals of NLP and Prompting**
  - **Basics of Natural Language Processing (NLP):**
    - Text preprocessing and representation techniques.
  - **Prompt Engineering with Pre-trained Models:**
    - Crafting prompts for GPT and BERT.
    - Prompt tuning techniques.
- **Module 8: Advanced Prompt Engineering Techniques**
  - **Complex Prompting Strategies:**
    - Conditional and few-shot prompting.
    - Interactive prompting and iterative refinement.
  - **Use Cases and Applications:**
    - Practical applications in Q&A, content generation, and more.
- **Module 9: Ethics and Considerations in Prompt Engineering**
  - **Ethical Implications:**
    - Addressing biases and ensuring responsible use.

- **Best Practices:**
  - Techniques for effective and ethical prompting.
- **Module 10: Capstone Project in Generative AI and Prompt Engineering**
  - **Project Initiation:**
    - Define project scope, objectives, and datasets.
    - Plan and outline methodologies.
  - **Project Execution:**
    - Implement generative models and prompt engineering techniques.
    - Evaluate, refine, and optimize project outcomes.
  - **Final Presentation:**
    - Develop a comprehensive report and deliver a formal presentation showcasing project findings and insights.

**Objectives:**

- Equip students with advanced skills in Generative AI and Prompt Engineering.
- Develop the ability to create, fine-tune, and deploy generative models and prompts.
- Provide practical experience through real-world projects and case studies.
- Prepare students for specialized roles in AI research, development, and application.
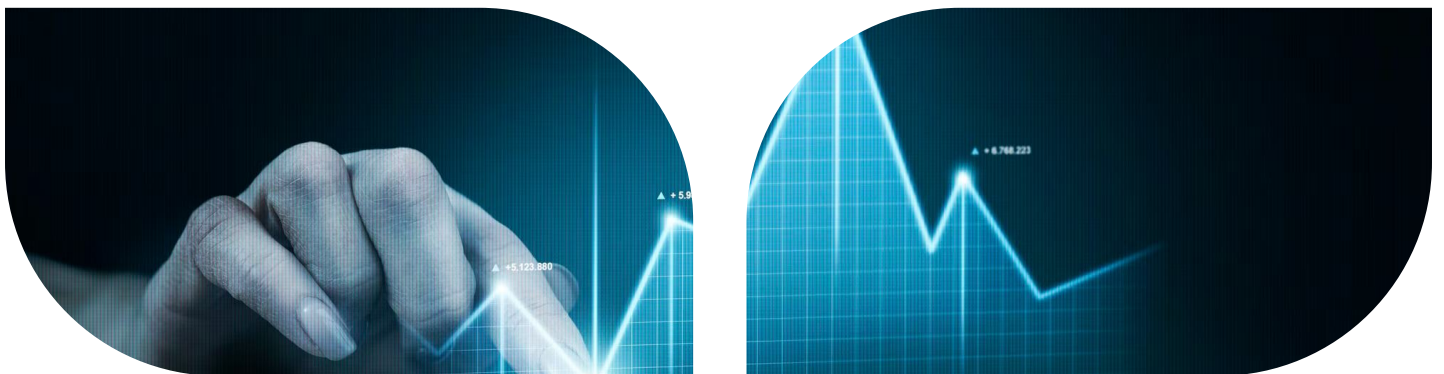


**Program Key Benefits::**

- **Advanced Expertise:** Gain deep knowledge and hands-on experience with cutting-edge generative models and prompt engineering techniques.

- **Practical Skills:** Apply theoretical knowledge to real-world projects, enhancing practical skills and problem-solving abilities.

- **Career Readiness:** Prepare for advanced roles in AI and data science, with skills applicable to various industries and research areas.

- **Comprehensive Learning:** Integrate concepts from multiple domains, including computer vision, natural language processing, and audio generation, for a holistic understanding of AI technologies.

# Capstone Project Examples by Module

1. **Statistics and Data Analysis**

   - **Project Topic:** Exploratory Data Analysis (EDA) and Statistical Modeling of a Real-world Dataset

   - **Deliverables:**

     - Comprehensive report on data insights

     - Visualization dashboards illustrating key findings

2. **Machine Learning and Deep Learning**

   - **Project Topic:** Building and Optimizing a Machine Learning Model for Predictive Analytics

   - **Deliverables:**

     - Detailed model performance metrics

     - Complete code implementation with documentation

3. **SQL and Database Management**

- **Project Topic:** Designing and Implementing a Database System for a Real-world Scenario

- **Deliverables:**

  - Entity-Relationship (ER) diagram

  - SQL queries and database implementation

4. **Business Intelligence and Data Visualization**

- **Project Topic:** Creating Interactive Dashboards Using Power BI or Tableau

- **Deliverables:**

  - Interactive dashboard

  - Presentation of data insights and business recommendations

5. **MLOps**

- **Project Topic:** Developing a CI/CD Pipeline for Machine Learning Model Deployment

- **Deliverables:**

  - CI/CD pipeline configuration

  - Deployment documentation and best practices

6. **General Artificial Intelligence (Gen AI)**

- **Project Topic:** Designing an AI System with Cognitive Abilities (e.g., Reasoning, Natural Language Understanding)

- **Deliverables:**

  - System architecture design

  - Demonstration of AI capabilities and performance

7. **Generative AI**

- **Project Topic:** Creating an Application Using Generative Models (e.g., Image Generation, Text-to-Speech)

- **Deliverables:**

  - Working prototype of the application

  - User guide and technical documentation

8. **Prompt Engineering**

- **Project Topic:** Implementing a Prompt-based NLP System for a Specific Application (e.g., Question Answering, Text Generation)
- **Deliverables:**
  - System documentation and setup
  - Performance evaluation and testing results

# Capstone Project Examples by Domain and Industry Problem

1. **Retail Analytics**
   - **Project Topic:** Customer Segmentation and Predictive Modeling for Sales Forecasting
   - **Industry Problem:** Enhancing customer targeting and personalization to improve sales and marketing strategies.

2. **Healthcare Analytics**
   - **Project Topic:** Predictive Modeling for Disease Diagnosis Using Patient Data
   - **Industry Problem:** Improving healthcare decision-making and patient care through accurate predictive analytics.

3. **Financial Analytics**
   - **Project Topic:** Fraud Detection System Using Machine Learning Algorithms
   - **Industry Problem:** Reducing financial risks and identifying fraudulent activities in banking and finance sectors.

4. **Supply Chain Analytics**
   - **Project Topic:** Optimization of Supply Chain Operations Using Predictive Analytics

- **Industry Problem:** Enhancing supply chain efficiency and reducing operational costs.

5. **Marketing Analytics**

   - **Project Topic:** Customer Lifetime Value Prediction and Churn Analysis

   - **Industry Problem:** Optimizing marketing strategies and budget allocation to maximize ROI and reduce churn.

6. **Social Media Analytics**

   - **Project Topic:** Sentiment Analysis and Topic Modeling on Social Media Data

   - **Industry Problem:** Understanding customer sentiment and brand perception to guide marketing and product development.

7. **Manufacturing Analytics**

   - **Project Topic:** Predictive Maintenance Using IoT Sensor Data

   - **Industry Problem:** Minimizing downtime and improving production processes through predictive maintenance.

8. **Environmental Analytics**

   - **Project Topic:** Predictive Modeling for Climate Change Impact Assessment

   - **Industry Problem:** Supporting environmental sustainability initiatives through accurate climate impact predictions.

# Final Advanced Capstone Project

1. **Advanced Capstone Project Overview**

   - Integration of skills and knowledge acquired throughout the program.

   - Addressing complex problems that require a combination of various data science and AI techniques.

2. **Project Scope Definition**

   - Defining a comprehensive project scope based on student interests and industry relevance.

   - Ensuring the project involves significant challenges and opportunities for showcasing skills.

## 3. Execution and Presentation

- Developing the project through iterative phases, including research, implementation, and testing.

- Presenting the final project to peers, mentors, and industry experts, demonstrating proficiency in data science and AI.