

upGrad



International  
Institute of Information  
Technology Bangalore

Post Graduate Programme in

# Data Scienc & AI

(Executive)



Now integrated with

**Generative AI**

/Prompt:

/Prompt:

/Prompt:

/Prompt:



# Table of Contents

## The Era Of Generative AI

About upGrad

Why upGrad?

Program Highlights

Faculty and Industry Experts

upGrad Learning Experience

Industry Projects

Learning Path

Executive PG Programme Curriculum

Meet the Class

Career Support

Our Alumni Work at

Career Transitions

Experience upGrad Offline

Hear from our Learners

Program Details and Admission Process

# The Era Of Generative AI



Usually, this first page is reserved for “About **upGrad**”. But the world is at the cusp of Generative AI rapidly changing the world as we know it. At upGrad, we’ve always believed in imparting learners the skills necessary to thrive in the fast-evolving world of technology. We are hence quite thrilled to pioneer Generative AI as an elective in **Executive PG Programme in Data Science**. With this key inclusion of Generative AI, learners will delve deeper into the fascinating realm of using **Data Science** to build practical applications like conversational AI chat bots, image creators, and content recommenders amongst others, to solve real-world challenges. So dive into this brave new world of Generative AI and Large Language Models with us, and watch yourself transform into a 10x Data Scientist.

“ IIT Bangalore prides itself in constantly updating cutting-edge topics to its curriculum. Our faculty has shaped this exciting Generative AI elective along with upGrad’s industry experts, thus ensuring both academic rigour as well as incorporating the latest advancements in tech.”

**Dr. V. Sridhar,**  
Head-Faculty, IITB

“ As an organisation that asks professionals to stay updated with the latest skills, we had to be one of the first to teach Generative AI. With this move, we are excited to witness the impact that Generative AI will have on the future, as well as the value our learners will bring to the field with this essential skill.”

**Mayank Kumar, Co-founder & MD**  
upGrad

# About upGrad

*upGrad has delivered over 20 million hours of learning, delivering programs by collaborating with universities across the world including Duke CE, IIT Bangalore and Deakin Business School among others.*

Online education is a fundamental disruption that will have a far-reaching impact. **upGrad** was founded taking this into consideration. upGrad is an online education platform to help individuals develop their professional potential in the most engaging learning environment.

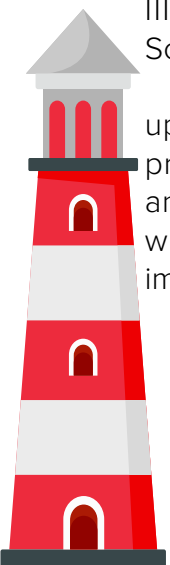
Since its inception, upGrad has delivered over 20 million hours of learning, delivering programs by collaborating with universities across the world, including Duke CE, IIT Bangalore and Deakin Business School among others.

upGrad is focused on helping working professionals in their bid to learn, grow and move up in their careers through a wide range of programs designed to improve their expertise.

**IITB** is a renowned university offering programs specialising in data science, machine learning and artificial intelligence. The IITB faculty includes an average of 15+ years of experience.

The faculty covers the conceptual depths of topics such as Data Science, Machine Learning and Artificial Intelligence, and Big Data Analytics. These will be complemented by industry-relevant case studies from major industry verticals by industry leaders with 8+ years of experience from upGrad's industry network.

Furthermore, our strong placement network, industry mentorship and the credibility of an Executive PG Programme will provide you with just the right push to accelerate your career in Data Science!



# Why upGrad?



**85+**

Countries Learners Base

**433%**

Highest Salary Hike



**50%**

Avg. Salary Hike



**300+**

Hiring Partners



**10 million+**

Learners



**700+**

Industry Experts

# Program Highlights

An illustration of an iceberg floating in a red ocean. The iceberg is white and grey, with a small black penguin on its base. Five callout lines with dots point to different parts of the iceberg, each leading to a text box. Two birds are flying in the sky above the iceberg.

## **Equivalent to NSQF (National Skill Qualification Framework) level 8**

Do an Executive PG Programme from IITB that satisfies NSQF level 8 criteria.

## **Executive PG Programme from IITB and Alumni Status**

Get certified by IITB and gain alumni status on successful completion of the program.

## **Tools & Languages**

Learn 14 + Programming Tools & Languages such as Python, Tableau, MySQL, Keras, Tensorflow and more.

## **5 Specialisations**

Choose from 5 specialisations such as Natural Language Processing, Deep Learning, Business Intelligence/ Data Analytics, Business Analytics, Data Engineering, based on your background and career aspirations and get the learning you want.

## **Blended Learning**

Learn with the ease and flexibility of recorded sessions as well as live sessions, designed to ensure a wholesome learning experience.



# Faculty and Industry Experts



**Dr. Debabrata Das**  
Director, IIITB

Dr. Debabrata Das is Director of IIITB. He has received his PhD from IIT-KGP. His main areas of research are IoT and Wireless Access Network.



**Chandrashekar Ramanathan**  
Dean Academics, IIITB

Prof. Chandrashekar has a PhD from Mississippi State University and experience of over 10 years in several multinational organisations.



**S. Anand**  
CEO, Gramener

An alumnus of IIT Madras, IIM Bangalore and LBS London, Anand is among the top 10 data scientists in India with 20 years of experience.



**Tricha Anjali**  
Ex-Associate Dean, IIITB

Prof. Anjali has a PhD from Georgia Institute of Technology as well as an integrated MTech (EE) from IIT Bombay.



**Behzad Ahmadi**  
Data Scientist Walmart Labs

An M. Tech graduate and PhD from Jersey Institute of Technology, Behzad possesses tremendous years of experience in Data Science and ML.



**Kautuk Pandey**  
Visa, Ex- Apple

Kautuk has 10+ years of experience working in Data Science. He is a seasoned professional in Big Data, AWS, Pyspark and other technologies.



**Prof. G. Srinivasaraghavan**  
Professor, IIITB

Prof. Srinivasaraghavan has a PhD in Computer Science from IIT-K and 18 years of experience with Infosys and several other MNCs.



**Mirza Rahim Baig**  
Analyst Lead, Zalando

Mirza is a veteran professional with 10+ years of experience in applications of data science, machine learning in e-commerce and healthcare.



**Sajjan Kedia**  
Ex- Data Science Lead, Myntra

Sajjan graduated from IIT, BHU and has tons of experience in Data Science, Big Data, Spark, Machine Learning and Natural Language Processing.



**Rajesh Sabapathy**  
Sr Director, Data Science, UHG Group

Rajesh has 10+ years of experience leading Data Science teams in various domains solving complex problems using Deep Learning & ML technique.



**Vishwa Mohan**  
LinkedIn, Ex- Walmart

An alumnus of IIT Varanasi, Vishwa has 10+ years of experience working in multiple MNCs for scaling solutions.



**Ankit Jain**  
ML Engineering Manager, Meta

An alumnus of IIT Bombay, UCB, and HBS with over 9 years of experience. Ankit has been recognised as 40 Under40 Data Scientist for 2022.



**Ujjayini Mitra**  
Head of Analytics, Zee5

An alumnus of McKinsey and Co, Flipkart and Bharati Airtel with over 11 years of experience.





# upGrad Learning Experience



## Student Support Team

- We have a dedicated/ Student Support Team for handling your queries via email or callback requests
- This support is available 7 days a week, 24x7

## Expert Feedback

- Personalised expert feedback on assignments and projects
- Regular live sessions by experts to clarify concept-related doubts

## Industry Networking

- Live sessions by experts on various industry topics
- One-on-one discussion and feedback sessions with industry mentors

## Industry Mentors

- Receive unparalleled guidance from industry mentors, teaching assistants and graders
- Receive one-on-one feedback on submissions and personalised feedback on improvement

## upGrad BaseCamp (PRE-COVID)

- Fun-packed, informative and career building workshop sessions by industry professionals and professors
- Group activities with your peers and alumni

## Q&A Forum

- Timely doubt resolution by industry experts and peers
- 100% expert-verified responses to ensure quality learning

# New Additions



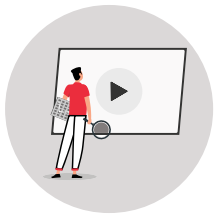
## Career Essential Soft-skills Program

1. Excel your personal & professional life with upGrad's Soft Skills Program
2. Study Three fundamental Skills - Interview & Job Search, Corporate & Business Communication and Problem Solving
3. Get access to 40+ learner hours of soft skills content delivered by the best faculty & Industry experts

## 30-Hour Programming Bootcamp for Non-tech Learners

1. Non-tech background? No need to fear Programming anymore
2. A 30-hour Python Programming bootcamp, focusing on developing Basic + Intermediate Python Programming Concepts to assist non-tech learners.
3. A blended learning experience delivered via Interactive live sessions and assessments

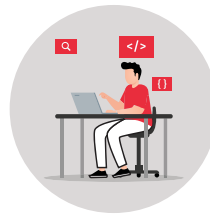
# Industry Projects



**IMDb Movie Analysis**



**Uber Supply-Demand Gap**



**Lead Scoring**



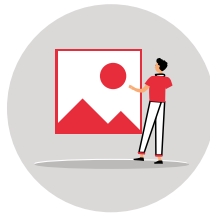
**Fraud Detection**



**Creditworthiness of  
Customers**



**Speech Recognition**



**Image Captioning**



**Gesture Recognition**



**Social Media Listening**



**Telecom Churn**



**Retail Giant Sales  
Forecasting**



**And many more!**



# Learning Path



## Preparatory Course

0 week

Tools: Python, Excel



## Data Toolkit

12 weeks

Tools: Python, Excel, MySQL



## Machine Learning

10 weeks

Tools: Python, Excel



## Choose any of the 5 Specialisations

22 weeks (with 4 weeks of Capstone)



## Natural Language Processing

Tools: ChatGPT, OpenAI API, Dall-E, Midjourney, Copilot, Flask



**Executive PG Programme in Data Science (Natural Language Processing)**



## Deep Learning

Tools: ChatGPT, OpenAI API, Dall-E, Midjourney, Copilot, Flask



**Executive PG Programme in Data Science (Deep Learning)**



## Business Analytics

Tools: ChatGPT, OpenAI API, Dall-E, Midjourney, Copilot, Flask



**Executive PG Programme in Data Science (Business Analytics)**



## Business Intelligence/ Data Analytics

Tools: Python, Power BI, Excel, MySQL, MongoDB, Shiny, Tableau



**Executive PG Programme in Data Science (Business Intelligence/ Data Analytics)**



## Data Engineering

Tools: Hadoop, HBase, Sqoop, Hive, Flume, PySpark, Spark, Airflow



**Executive PG Programme in Data Science (Data Engineering)**

# Executive PG Programme in Data Science

## COMMON CURRICULUM

### PRE-PROGRAM PREPARATORY CONTENT

#### 1. DATA ANALYSIS IN EXCEL

1. INTRODUCTION TO EXCEL
2. DATA ANALYSIS IN EXCEL - I:  
FUNCTIONS, FORMULAE, AND  
CHARTS
3. DATA ANALYSIS IN EXCEL - II:  
PIVOTS AND LOOKUPS

Taught by one of the most renowned data scientists in the country (S.Anand, CEO, Gramener), this module takes you from a beginner-level Excel user to an almost professional user.

---

#### 2. ANALYTICS PROBLEM SOLVING

1. THE CRISP-DM FRAMEWORK  
- BUSINESS AND DATA  
UNDERSTANDING
2. CRISP-DM FRAMEWORK  
- DATA PREPARATION,  
MODELLING, EVALUATION  
AND DEPLOYMENT

This module covers concepts of the CRISP-DM framework for business problem-solving.

---

### COURSE 1: DATA TOOLKIT

#### 1. INTRODUCTION TO PYTHON

1. UNDERSTANDING THE  
UPGRAD CODING CONSOLE
2. BASICS OF PYTHON
3. DATA STRUCTURES IN  
PYTHON
4. CONTROL STRUCTURE AND  
FUNCTIONS IN PYTHON
5. OOP IN PYTHON

Build a foundation for the most in-demand programming language of the 21st century.

**2 WEEKS**

## 2. PROGRAMMING IN PYTHON

1. LOGIC AND SYNTAX BUILDING	Learn how to approach and solve logical problems using programming.	1 WEEK
2. DATA STRUCTURES: LISTS, STRINGS, DICTIONARIES, AND STACKS		
3. TIME COMPLEXITY		
4. SEARCHING AND SORTING		
5. TWO POINTERS		
6. RECURSION		

---

## 3. PYTHON FOR DATA SCIENCE

1. INTRODUCTION TO NUMPY	Learn how to manipulate datasets in Python using Pandas which is the most powerful library for data preparation and analysis.	1 WEEK
2. INTRODUCTION TO MATPLOTLIB		
3. INTRODUCTION TO PANDAS		
4. GETTING AND CLEANING DATA		

---

## 4. DATA VISUALISATION IN PYTHON

1. INTRODUCTION TO DATA VISUALISATION	Humans are visual learners, and hence no task related to data is complete without visualisation. Learn to plot and interpret various graphs in Python and observe how they make data analysis and drawing insights easier.	1 WEEK
2. DATA VISUALISATION USING SEABORN		

---



## 5. EXPLORATORY DATA ANALYSIS

1. DATA SOURCING	Learn how to find and analyse the patterns in the data to draw actionable insights.	<b>1 WEEK</b>
2. DATA CLEANING		
3. UNIVARIATE ANALYSIS		
4. BIVARIATE ANALYSIS AND MULTIVARIATE ANALYSIS		

---

## 6. CREDIT EDA CASE STUDY

1. PROBLEM STATEMENT	Solve a real industry problem through the concepts learnt in exploratory data analysis.	<b>1 WEEK</b>
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

## 7. INFERENCE STATISTICS

1. BASICS OF PROBABILITY	Build a strong statistical foundation and learn how to 'infer' insights from a huge population using a small sample.	<b>1 WEEK</b>
2. DISCRETE PROBABILITY DISTRIBUTIONS		
3. CONTINUOUS PROBABILITY DISTRIBUTIONS		
4. CENTRAL LIMIT THEOREM		

---

## 8. HYPOTHESIS TESTING

1. CONCEPTS OF HYPOTHESIS TESTING - I: NULL AND ALTERNATE HYPOTHESIS, MAKING A DECISION, AND CRITICAL VALUE METHOD	Understand how to formulate and validate hypotheses for a population to solve real-life business problems.	<b>1 WEEK</b>
2. CONCEPTS OF HYPOTHESIS TESTING - II: P-VALUE METHOD AND TYPES OF ERRORS		
3. INDUSTRY DEMONSTRATION OF HYPOTHESIS TESTING: TWO-SAMPLE MEAN AND PROPORTION TEST, A/B TESTING		

## 9. DATA ANALYSIS USING SQL

1. DATABASE DESIGN	Data in companies is definitely not stored in excel sheets! Learn the fundamentals of databases and extract information from RDBMS using the structured query language.	1 WEEK
2. DATABASE CREATION IN MYSQL WORKBENCH		
3. QUERYING IN MYSQL		
4. JOINS AND SET OPERATIONS		

---

## 10. ADVANCED SQL & BEST PRACTICES

1. WINDOW FUNCTIONS	Apply advanced SQL concepts like windowing and procedures to derive insights from data and answer pertinent business questions.	1 WEEK
2. CASE STATEMENTS, STORED ROUTINES AND CURSORS		
3. QUERY OPTIMISATION AND BEST PRACTICES		
4. PROBLEM-SOLVING USING SQL		

---

## 11. SQL ASSIGNMENT: RSVP MOVIES

1. PROBLEM STATEMENT	In this assignment, you will work on a movies dataset using SQL to extract exciting insights.	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

## COURSE 2 - MACHINE LEARNING I

### 1. LINEAR REGRESSION

1. SIMPLE LINEAR REGRESSION	Venture into the machine learning community by learning how one variable can be predicted using several other variables through a housing dataset where you will predict the prices of houses based on various factors.	2 WEEKS
2. SIMPLE LINEAR REGRESSION IN PYTHON		
3. MULTIPLE LINEAR REGRESSION		
4. MULTIPLE LINEAR REGRESSION IN PYTHON		
5. INDUSTRY RELEVANCE OF LINEAR REGRESSION		

## 2. LINEAR REGRESSION ASSIGNMENT

1. PROBLEM STATEMENT	Build a model to understand the factors on which the demand for bike-sharing systems vary on and help a company optimise its revenue.	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

## 3. LOGISTIC REGRESSION

1. UNIVARIATE LOGISTIC REGRESSION	Learn your first binary classification technique by determining which telecom operator customers are likely to churn versus those who are not to help the business retain customers.	2 WEEKS
2. MULTIVARIATE LOGISTIC REGRESSION: MODEL BUILDING AND EVALUATION		
3. LOGISTIC REGRESSION: INDUSTRY APPLICATIONS		

---

## 4. CLASSIFICATION USING DECISION TREES

1. INTRODUCTION TO DECISION TREES	Learn how the human decision-making process can be replicated using a decision tree and tune it to suit your needs.	1 WEEK
2. ALGORITHMS FOR DECISION TREES CONSTRUCTION		
3. HYPERPARAMETER TUNING IN DECISION TREES		

---

## 5. UNSUPERVISED LEARNING: CLUSTERING

1. INTRODUCTION TO CLUSTERING	Learn how to group elements into different clusters when you don't have any pre-defined labels to segregate them through K-means clustering, hierarchical clustering, and more.	1 WEEK
2. K-MEANS CLUSTERING		
3. HIERARCHICAL CLUSTERING		
4. OTHER FORMS OF CLUSTERING: K-MODE, K-PROTOTYPE, DB SCAN		

---

## 6. BASICS OF NLP AND TEXT MINING

1. REGEX AND INTRODUCTION TO NLP	Do you get annoyed by the constant spam in your mailbox? Wouldn't it be nice if we had a program to check your spelling? In this module learn how to build a spell checker & spam detector using techniques like phonetic hashing, bag-of-words, TF-IDF, etc.	1 WEEK
2. BASIC LEXICAL PROCESSING		
3. ADVANCED LEXICAL PROCESSING		

---

## 5. BUSINESS PROBLEM SOLVING

1. INTRODUCTION TO BUSINESS PROBLEM SOLVING	Learn how to approach open-ended real-world problems using data as a lever to draw actionable insights.	1 WEEK
2. BUSINESS PROBLEM SOLVING: CASE STUDY DEMONSTRATIONS		

---

## 7. CASE STUDY: LEAD SCORING

1. PROBLEM STATEMENT	Help the Sales team of your company identify which leads are worth pursuing through this classification case study.	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

# SPECIALISATION: DEEP LEARNING

## COURSE 3 - MACHINE LEARNING II

### 1. BAGGING & RANDOM FOREST

1. POPULAR ENSEMBLES	Learn how powerful ensemble algorithms can improve your classification models by building random forests from decision trees.	1 WEEK
2. INTRODUCTION TO RANDOM FORESTS		
3. FEATURE IMPORTANCE IN RANDOM FORESTS		
4. RANDOM FORESTS IN PYTHON		

---

### 2. BOOSTING

1. INTRODUCTION TO BOOSTING AND ADABOOST	Learn about ensemble modelling through bagging and boosting and, understand how weak algorithms can be transformed into stronger ones.	1 WEEK
2. GRADIENT BOOSTING		

---

### 3. MODEL SELECTION & GENERAL ML TECHNIQUES

1. PRINCIPLES OF MODEL SELECTION	Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, along with general machine learning techniques like feature engineering, model evaluation, and many more.	1 WEEK
2. MODEL EVALUATION		
3. MODEL SELECTION: BEST PRACTICES		

---

### 4. PRINCIPAL COMPONENT ANALYSIS

1. PRINCIPAL COMPONENT ANALYSIS AND SINGULAR VALUE DECOMPOSITION	Understand important concepts related to dimensionality reduction, the basic idea and the learning algorithm of PCA, and its practical applications on supervised and unsupervised problems.	1 WEEK
2. PRINCIPAL COMPONENT ANALYSIS IN PYTHON		

---

## 5. ADVANCED REGRESSION

- |                                  |  |               |
|----------------------------------|--|---------------|
| 1. GENERALISED LINEAR REGRESSION | In this module, take a more advanced look at regression models and learn the concepts related to regularisation. | <b>1 WEEK</b> |
| 2. REGULARISED REGRESSION        |  |               |
- 

## 6. TIME SERIES FORECASTING (OPTIONAL)

- |   |  |               |
|---|--|---------------|
| 1. INTRODUCTION TO TIME SERIES AND ITS COMPONENTS | In this module, you will learn how to analyse and forecast a series that varies with time. | <b>0 WEEK</b> |
| 2. WORKING WITH STATIONARY TIME SERIES            |  |               |
| 3. END-TO-END ANALYSIS OF TIME SERIES             |  |               |
- 

## 7. ADVANCED ML CASE STUDY

**1 WEEK**

- |                      |   |
|----------------------|---|
| 1. PROBLEM STATEMENT | Build a regularized regression model to understand the most important variables to predict house prices in Australia. |
| 2. EVALUATION RUBRIC |   |
| 3. FINAL SUBMISSION  |   |
| 4. SOLUTION          |   |
- 

# COURSE 4 - ADVANCED MACHINE LEARNING AND DEEP LEARNING

## 1. INTRODUCTION TO NEURAL NETWORKS AND ANN

- |   |   |                |
|---|---|----------------|
| 1. STRUCTURE OF NEURAL NETWORKS             | Learn the most sophisticated and cutting-edge technique in machine learning - Artificial Neural Networks or ANNs. | <b>2 WEEKS</b> |
| 2. FEED FORWARD IN NEURAL NETWORKS          |   |                |
| 3. BACKPROPAGATION IN NEURAL NETWORKS       |   |                |
| 4. MODIFICATIONS TO NEURAL NETWORKS         |   |                |
| 5. HYPERPARAMETER TUNING IN NEURAL NETWORKS |   |                |



## 2. CONVOLUTIONAL NEURAL NETWORKS

1. INTRODUCTION TO CONVOLUTIONAL NEURAL NETWORKS	Learn the basics of CNN and OpenCV and how to classify image data using various architectures which you will then implement using Python and Keras.	1 WEEK
2. BUILDING CNNs WITH PYTHON AND KERAS		
3. CNN ARCHITECTURES AND TRANSFER LEARNING		
4. STYLE TRANSFER AND OBJECT DETECTION		

---

## 3. CONVOLUTIONAL NEURAL NETWORKS -INDUSTRY APPLICATIONS

1. INDUSTRY DEMONSTRATION: USING CNNs WITH FLOWERS IMAGES	Apply CNNs to Computer Vision tasks like detecting anomalies in chest X-Ray scans.	1 WEEK
2. INDUSTRY DEMONSTRATION: USING CNNs WITH X-RAY IMAGES		

---

## 4. OBJECT DETECTION & IMAGE SEGMENTATION

1. FUNDAMENTALS OF OBJECT DETECTION	Learn the applications of DL in computer vision through industry-relevant detection algorithms such as RCNNs, YOLO and SSD.	1 WEEK
2. REGION-BASED DETECTORS		
3. ONE-SHOT DETECTORS		
4. CUSTOM OBJECT DETECTION		
5. SEMANTIC SEGMENTATION		

## 5. RECURRENT NEURAL NETWORKS (OPTIONAL)

1. WHAT MAKES A NEURAL NETWORK RECURRENT	Ever wondered what goes behind machine translation, sentiment analysis, and speech recognition? Learn how RNN helps in areas having sequential data like text, speech, videos, and a lot more.	1 WEEK
2. VARIANTS OF RNNs: BIDIRECTIONAL RNNs AND LSTMS		
3. BUILDING RNNs IN PYTHON		

---

## 6. GESTURE RECOGNITION

1. TWO ARCHITECTURES: 3D CONVS AND CNN-RNN STACK	Make a Smart TV system which can control the TV with the user's hand gestures as the remote control	1 WEEK
2. UNDERSTANDING GENERATORS		
3. STARTER CODE WALKTHROUGH		
4. PROBLEM STATEMENT AND FINAL SUBMISSION		

---

## COURSE 5 - GENERATIVE AI

<b>1. “FUNDAMENTALS OF TRANSFORMERS ARCHITECTURE, GENERATIVE AI, CHATGPT &amp; PROMPT ENGINEERING USING NON REASONING, CHAIN OF THOUGHT &amp; ADVANCED TECHNIQUES”</b>	<b>1 WEEK</b>
--	---------------

<b>2. PRODUCT DEVELOPMENT USING OPENAI APIS, FINE TUNING USING STAR TECHNIQUE IN PYTHON</b>	<b>1 WEEK</b>
---	---------------

<b>3. INTEGRATING SPEECH USING WHISPER API AND APPLICATION DEPLOYMENT USING FLASK</b>	<b>1 WEEK</b>
---	---------------

<b>4. “FUNDAMENTALS OF DESIGN, PHOTOGRAPHY, PRODUCT DEVELOPMENT USING STABLE DIFFUSION IN PYTHON &amp; CREATE PIXXELCRAFT AI TO ENABLE FAST-TRACK DIGITISATION FOR OFFLINE E-COMMERCE BUSINESSES BY GENERATING HIGH-QUALITY IMAGES AI FOR A LARGE PRODUCT PORTFOLIO”</b>	<b>1 WEEK</b>
--	---------------

<b>5. “APPLICATIONS OF LLMS IN DATA SCIENCE PROJECTS &amp; AUTOMATING NEWS RECOMMENDATION USING GPT3 AND COPILOT POWERED MACHINE LEARNING APPLICATIONS OF LLMS”</b>	<b>1 WEEK</b>
---	---------------

<b>6. INTERVIEW GYNIE AI: CHATBOT DEVELOPMENT PROJECT</b>	<b>1 WEEK</b>
---	---------------

## COURSE 6 - CAPSTONE PROJECT

### CAPSTONE PROJECT

1. AN OVERVIEW OF THE DOMAIN AND ASSOCIATED CONCEPTS
2. PROBLEM STATEMENT
3. EVALUATION RUBRIC
4. MID SUBMISSION
5. FINAL SUBMISSION
6. SOLUTION

Choose from a range of real-world industry-woven projects on advanced topics like Recommendation Systems, Fraud Detection, Emotion Detection from faces, Social Media Listening, and Speech Recognition among many others.

**4 WEEKS**

---

## SPECIALISATION: NATURAL LANGUAGE PROCESSING

### COURSE 3 - MACHINE LEARNING II

#### 1. BAGGING & RANDOM FOREST

1. POPULAR ENSEMBLES
2. INTRODUCTION TO RANDOM FORESTS
3. FEATURE IMPORTANCE IN RANDOM FORESTS
4. RANDOM FORESTS IN PYTHON

Learn how powerful ensemble algorithms can improve your classification models by building random forests from decision trees.

**1 WEEK**

---

#### 2. BOOSTING

1. INTRODUCTION TO BOOSTING AND ADABOOST
2. GRADIENT BOOSTING

Learn about ensemble modelling through bagging and boosting, and understand how weak algorithms can be transformed into stronger ones.

**1 WEEK**

### 3. MODEL SELECTION & GENERAL ML TECHNIQUES

1. PRINCIPLES OF MODEL SELECTION	Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, along with general machine learning techniques like feature engineering, model evaluation, and many more.	1 WEEK
2. MODEL EVALUATION		
3. MODEL SELECTION: BEST PRACTICES		

---

### 4. PRINCIPAL COMPONENT ANALYSIS

1. PRINCIPAL COMPONENT ANALYSIS AND SINGULAR VALUE DECOMPOSITION	Understand important concepts related to dimensionality reduction, the basic idea and the learning algorithm of PCA, and its practical applications on supervised and unsupervised problems.	1 WEEK
2. PRINCIPAL COMPONENT ANALYSIS IN PYTHON		

---

### 5. ADVANCED REGRESSION

1. GENERALISED LINEAR REGRESSION	In this module, take a more advanced look at regression models and learn the concepts related to regularisation.	1 WEEK
2. REGULARISED REGRESSION		

---

### 6. TIME SERIES ANALYSIS (OPTIONAL)

1. INTRODUCTION TO TIME SERIES AND ITS COMPONENTS	In this module, you will learn how to analyse and forecast a series that varies with time.	2 WEEKS
2. WORKING WITH STATIONARY TIME SERIES		
3. END-TO-END ANALYSIS OF TIME SERIES		

---

## 7. ADVANCED ML CASE STUDY

1. PROBLEM STATEMENT	Build a regularised regression model to understand the most important variables to predict house prices in Australia.	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

## COURSE 4 - ADVANCED MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING

### 1. NEURAL NETS FOR NLP

1. UNDERSTANDING NEURAL NETWORKS	Learn the most sophisticated and cutting-edge technique in machine learning - Artificial Neural Networks or ANNs.	1 WEEK
2. LOSS FUNCTIONS AND BACK PROPAGATION		
3. UNDERSTANDING TENSORFLOW		
4. CASE STUDY: IMDB MOVIE REVIEW CLASSIFICATION		

---

### 2. SYNTACTIC PROCESSING

1. INTRODUCTION TO SYNTACTIC PROCESSING	Learn how to analyse the syntax or the grammatical structure of sentences using POS tagging and Dependency parsing.	1 WEEK
2. PARSING		
3. INFORMATION EXTRACTION		
4. CONDITIONAL RANDOM FIELDS		

---



### 3. SYNTACTIC PROCESSING

1. PROBLEM STATEMENT	Use the techniques such as POS tagging and Dependency parsing to extract information from unstructured text data.	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

### 4. SEMANTIC PROCESSING

1. INTRODUCTION TO SEMANTIC PROCESSING	Learn the most interesting area in the field of NLP and understand different techniques like word-embeddings and topic modelling to build an application that extracts opinions about socially relevant issues.	2 WEEKS
2. DISTRIBUTIONAL SEMANTICS		
3. INDUSTRY APPLICATIONS OF DISTRIBUTIONAL SEMANTICS		
4. TOPIC MODELLING		

---

### 5. APPLIED DL IN NLP

1. INTRODUCTION TO MACHINE TRANSLATION	Apply the concepts of DL in natural language processing problems through encoder-decoder architecture and NMTs, and implement them in TensorFlow.	1 WEEK
2. ATTENTION-BASED NMT MODEL		
3. CUSTOM MODEL BUILDING IN TENSORFLOW		

---

### 6. CASE STUDY: AUTOMATIC TICKET CLASSIFICATION

1. PROBLEM STATEMENT	Categorise support tickets with the help of Unsupervised learning and Topic modelling.	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

## COURSE 5 - GENERATIVE AI

**1. “FUNDAMENTALS OF TRANSFORMERS ARCHITECTURE, GENERATIVE AI, CHATGPT & PROMPT ENGINEERING USING NON REASONING, CHAIN OF THOUGHT & ADVANCED TECHNIQUES”**

**1 WEEK**

---

**2. PRODUCT DEVELOPMENT USING OPENAI APIS, FINE TUNING USING STAR TECHNIQUE IN PYTHON**

**1 WEEK**

---

**3. INTEGRATING SPEECH USING WHISPER API AND APPLICATION DEPLOYMENT USING FLASK**

**1 WEEK**

---

**4. “FUNDAMENTALS OF DESIGN, PHOTOGRAPHY, PRODUCT DEVELOPMENT USING STABLE DIFFUSION IN PYTHON & CREATE PIXXELCRAFT AI TO ENABLE FAST-TRACK DIGITISATION FOR OFFLINE E-COMMERCE BUSINESSES BY GENERATING HIGH-QUALITY IMAGES AI FOR A LARGE PRODUCT PORTFOLIO”**

**1 WEEK**

---

**5. “APPLICATIONS OF LLMS IN DATA SCIENCE PROJECTS & AUTOMATING NEWS RECOMMENDATION USING GPT3 AND COPILOT POWERED MACHINE LEARNING APPLICATIONS OF LLMS”**

**1 WEEK**

---

**6. INTERVIEW GYNIE AI: CHATBOT DEVELOPMENT PROJECT**

**1 WEEK**

## COURSE 6 - CAPSTONE PROJECT

### 1. CAPSTONE PROJECT

1. AN OVERVIEW OF THE DOMAIN AND ASSOCIATED CONCEPTS	Choose from a range of real-world industry-woven projects on advanced topics like Recommendation Systems, Fraud Detection, Emotion Detection from faces, Social Media Listening, and Speech Recognition among many others.	4 WEEKS
2. PROBLEM STATEMENT		
3. EVALUATION RUBRIC		
4. MID SUBMISSION		

---

## SPECIALISATION: BUSINESS ANALYTICS

### COURSE 3 - ADVANCED MACHINE LEARNING

#### 1. BAGGING & RANDOM FOREST

1. POPULAR ENSEMBLES	Learn how powerful ensemble algorithms can improve your classification models by building random forests from decision trees.	1 WEEK
2. INTRODUCTION TO RANDOM FORESTS		
3. FEATURE IMPORTANCE IN RANDOM FORESTS		
4. RANDOM FORESTS IN PYTHON		

---

#### 2. MODEL SELECTION & GENERAL ML TECHNIQUES

1. PRINCIPLES OF MODEL SELECTION	Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, along with general machine learning techniques like feature engineering, model evaluation, and many more.	2 WEEKS
2. MODEL BUILDING AND EVALUATION		
3. FEATURE ENGINEERING		
4. CLASS IMBALANCE		

---

### 3. TIME SERIES FORECASTING

1. INTRODUCTION TO TIME SERIES AND ITS COMPONENTS	In this module, you will learn how to analyse and forecast a series that varies with time.	2 WEEKS
2. SMOOTHING TECHNIQUES		
3. INTRODUCTION TO AR MODELS		
4. BUILDING AR MODELS		

---

### 4. MODEL SELECTION CASE STUDY

1. PROBLEM STATEMENT	Apply your business acumen to the newly learnt machine learning techniques, and select the right model most appropriate for a provided business scenario.	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

## COURSE 4 - DATA VISUALISATION AND STORYTELLING

### 1. VISUALISATION USING TABLEAU

1. DATA EXPLORATION IN TABLEAU	Learn basic visualisation techniques using the most in-demand visualisation tool in the industry.	1 WEEK
2. VISUALISING AND ANALYSING DATA IN TABLEAU WITH BASIC PLOTS		

---

## 2. ADVANCED EXCEL

1. EXCEL FUNCTIONS	Learn the advanced concepts in Excel and start to perform data analysis like a pro!	1 WEEK
2. DATA ANALYSIS IN EXCEL		
3. ADVANCED TOOLS AND VISUALISATIONS		

---

## 3. VISUALISATION USING POWERBI

1. POWERBI: INTRODUCTION AND SETUP	Take your visualisation game a step forward by understanding how to operate PowerBI.	1 WEEK
2. VISUALISING AND ANALYSING DATA IN POWERBI		
3. DATA TRANSFORMATIONS USING POWERBI		

---

## 4. STRUCTURED PROBLEM SOLVING USING FRAMEWORKS

1. INTRODUCTION TO STRUCTURED PROBLEM SOLVING	Learn how to attack a business problem using various structured frameworks like 5W, 5WHYS, and SPIN.	1 WEEK
2. INTERVIEWING AND FRAMEWORKS - I: 5W AND 5WHYS		
3. INTERVIEWING AND FRAMEWORKS - II: SPIN		
4. INDUSTRY DEMONSTRATIONS ON FRAMEWORKS		
5. UNDERSTANDING BUSINESS MODEL CANVAS AND ISSUE TREE FRAMEWORK		
6. INDUSTRY DEMONSTRATIONS ON ISSUE TREE FRAMEWORK		
7. SPECIALISED FRAMEWORKS FOR BUSINESS PROBLEMS: 7PS, 5CS, ETC.		

---

## 5. DATA STORYTELLING

1. INTRODUCTION TO DATA STORYTELLING
2. COMPONENTS OF A GOOD STORY WITH DATA - UNDERSTANDING YOUR STAKEHOLDER AND STAKEHOLDER EMPATHY, LEVELS OF DETAILS FOR DIFFERENT STAKEHOLDERS - CXO/LEADERSHIP VS TEAM PRESENTATIONS, VISUALS, ETC.
3. GOLDEN RULES FOR DATA STORYTELLING

Learn how to effectively strategise, communicate, and fine-grain your data analysis projects and understand how to optimally present your findings to technical and non-technical stakeholders and upgrade your storytelling skills.

**1 WEEK**

---

## 6. AIRBNB CASE STUDY

1. PROBLEM STATEMENT
2. EVALUATION RUBRIC
3. FINAL SUBMISSION
4. SOLUTION

Use your newly learnt UI tools skills to analyse an AirBnB dataset to make important business decisions. But the analysis is just a small part; can you also effectively present it using Data Storytelling to the right stakeholders?

**1 WEEK**

---



## COURSE 5 - GENERATIVE AI

<b>1. “FUNDAMENTALS OF TRANSFORMERS ARCHITECTURE, GENERATIVE AI, CHATGPT &amp; PROMPT ENGINEERING USING NON REASONING, CHAIN OF THOUGHT &amp; ADVANCED TECHNIQUES”</b>	<b>1 WEEK</b>
<b>2. PRODUCT DEVELOPMENT USING OPENAI APIS, FINE TUNING USING STAR TECHNIQUE IN PYTHON</b>	<b>1 WEEK</b>
<b>3. INTEGRATING SPEECH USING WHISPER API AND APPLICATION DEPLOYMENT USING FLASK</b>	<b>1 WEEK</b>
<b>4. “FUNDAMENTALS OF DESIGN, PHOTOGRAPHY, PRODUCT DEVELOPMENT USING STABLE DIFFUSION IN PYTHON &amp; CREATE PIXXELCRAFT AI TO ENABLE FAST-TRACK DIGITISATION FOR OFFLINE E-COMMERCE BUSINESSES BY GENERATING HIGH-QUALITY IMAGES AI FOR A LARGE PRODUCT PORTFOLIO”</b>	<b>1 WEEK</b>
<b>5. “APPLICATIONS OF LLMS IN DATA SCIENCE PROJECTS &amp; AUTOMATING NEWS RECOMMENDATION USING GPT3 AND COPILOT POWERED MACHINE LEARNING APPLICATIONS OF LLMS”</b>	<b>1 WEEK</b>
<b>6. INTERVIEW GYNIE AI: CHATBOT DEVELOPMENT PROJECT</b>	<b>1 WEEK</b>

## COURSE 6 - CAPSTONE PROJECT

### 1. CAPSTONE PROJECT

1. POWER BI - OPTIONAL
2. AN OVERVIEW OF THE DOMAIN AND ASSOCIATED CONCEPTS
3. PROBLEM STATEMENT
4. EVALUATION RUBRIC
5. MID SUBMISSION
6. FINAL SUBMISSION
7. SOLUTION

Solve an end-to-end real-life industry problem from a wide variety of domains.

**4 WEEKS**

---

## SPECIALISATION: BUSINESS INTELLIGENCE / DATA ANALYTICS

## COURSE 3: ADVANCED DBS AND BIG DATA ANALYTICS

### 1. DATA MODELLING

1. DATABASE DESIGN RECAP
2. BUILDING BLOCKS OF DATA MODELLING
3. PROBLEM SOLVING USING DATA MODELLING
4. DATA MODELLING: OPTIONAL ASSIGNMENT

In this module, you will learn and use data modelling on a dataset to solve a business problem.

**1 WEEK**

---

### 2. ADVANCED SQL AND BEST PRACTICES

1. WINDOW FUNCTIONS
2. CASE STATEMENTS, STORED ROUTINES, AND CURSORS
3. QUERY OPTIMISATION AND BEST PRACTICES
4. PROBLEM SOLVING USING SQL

Apply advanced SQL concepts like windowing and procedures to derive insights from data and answer pertinent business questions.

**1 WEEK**

### 3. INTRODUCTION TO BIG DATA AND CLOUD

1. BIG DATA AND CLOUD COMPUTING	Understand the basics of big data and cloud and learn to work with an EMR cluster on a cloud-based service.	1 WEEK
2. AMAZON WEB SERVICES		
3. BIG DATA STORAGE AND PROCESSING - HADOOP		
4. EMR CLUSTER IN AWS		

---

### 4. ANALYTICS USING SPARK

1. EXPLORATORY DATA ANALYSIS WITH PYSPARK	Use PySpark to do EDA and Predictive Analysis using Spark's ML library.	2 WEEKS
2. PREDICTIVE ANALYSIS WITH SPARK MLLIB		

---

### 5. BIG DATA CASE STUDY

1. PROBLEM STATEMENT	Use your analytics skills to work on a large dataset in the cloud to solve an industry problem.	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

## COURSE 4 - DATA VISUALISATION AND STORYTELLING

### 1. VISUALISATION USING TABLEAU

1. DATA EXPLORATION IN TABLEAU	Learn basic visualisation techniques using the most in-demand visualisation tool in the industry.	1 WEEK
2. VISUALISING AND ANALYSING DATA IN TABLEAU WITH BASIC PLOTS		

---

## 2. ADVANCED EXCEL

1. EXCEL FUNCTIONS	Learn the advanced concepts in Excel and start to perform data analysis like a pro!	1 WEEK
2. DATA ANALYSIS IN EXCEL		
3. ADVANCED TOOLS AND VISUALISATIONS		

---

## 3. VISUALISATION USING POWERBI

1. POWERBI: INTRODUCTION AND SETUP	Take your visualisation game a step forward by understanding how to operate PowerBI.	1 WEEK
2. VISUALISING AND ANALYSING DATA IN POWERBI		
3. DATA TRANSFORMATIONS USING POWERBI		

---

## 4. STRUCTURED PROBLEM SOLVING USING FRAMEWORKS

1. INTRODUCTION TO STRUCTURED PROBLEM SOLVING	Learn how to attack a business problem using various structured frameworks like 5W, 5WHYS, and SPIN.	1 WEEK
2. INTERVIEWING AND FRAMEWORKS - I: 5W AND 5WHYS		
3. INTERVIEWING AND FRAMEWORKS - II: SPIN		
4. INDUSTRY DEMONSTRATIONS ON FRAMEWORKS		
5. UNDERSTANDING BUSINESS MODEL CANVAS AND ISSUE TREE FRAMEWORK		
6. INDUSTRY DEMONSTRATIONS ON ISSUE TREE FRAMEWORK		
7. SPECIALIZED FRAMEWORKS FOR BUSINESS PROBLEMS: 7PS, 5CS, ETC.		

---

## 5. DATA STORYTELLING

1. INTRODUCTION TO DATA STORYTELLING	Learn how to effectively strategise, communicate, and fine-grain your data analysis projects and understand how to optimally present your findings to technical and non-technical stakeholders and upgrade your storytelling skills.	1 WEEK
2. COMPONENTS OF A GOOD STORY WITH DATA - UNDERSTANDING YOUR STAKEHOLDER AND STAKEHOLDER EMPATHY, LEVELS OF DETAILS FOR DIFFERENT STAKEHOLDERS - CXO/LEADERSHIP VS TEAM PRESENTATIONS, VISUALS, ETC.		
3. GOLDEN RULES FOR DATA STORYTELLING		

---

## 6. AIRBNB CASE STUDY

1. PROBLEM STATEMENT	Use your newly learnt UI tools skills to analyse an AirBnB dataset to make important business decisions. But the analysis is just a small part; can you also effectively present it using Data Storytelling to the right stakeholders?	1 WEEK
2. EVALUATION RUBRIC		
3. FINAL SUBMISSION		
4. SOLUTION		

---

## COURSE 5: ADVANCED PROBLEM SOLVING AND PROGRAMMING

### 1. DATA STRUCTURES - SETS, DICTIONARIES, STACKS, QUEUES

1. IN-BUILT DATA STRUCTURES	Learn user-defined data structures -Stack, Queue, and Trees in Python that help in advanced data manipulation.	1 WEEK
2. STACK		
3. QUEUE		
4. TREES		

---

## 2. SEARCHING AND SORTING

1. <b>SEARCHING</b>	Learn most fundamental searching and sorting algorithms and design techniques	<b>1 WEEK</b>
2. <b>SORTING</b>		
3. <b>TWO POINTERS</b>		

---

## 3. ALGORITHM ANALYSIS + RECURSION

1. <b>ALGORITHM ANALYSIS</b>	Learn how to assess the efficiency of your code using algorithm analysis techniques and learn to write recursive algorithms	<b>1 WEEK</b>
2. <b>TIME AND SPACE COMPLEXITY</b>		
3. <b>RECURSION</b>		

---

## 4. ADVANCED DATABASE PROGRAMMING USING PANDAS

1. <b>ADVANCED DATA WRANGLING WITH PANDAS - I</b>	Learn and implement advanced wrangling functions and techniques in Pandas related to date-time, multi-columns aggregation, hierarchical indexing, and more.	<b>1 WEEK</b>
2. <b>ADVANCED DATA WRANGLING WITH PANDAS - II</b>		

---

## 5. PYTHON & SQL LAB

1. <b>SQL: TIMED TEST + ASSIGNMENT</b>	In this competitive assignment, you will solve a variety of programming questions in both SQL and Python in a timed environment. You will also demonstrate one of the questions through a video submission to help improve your interviewing skills.	<b>2 WEEKS</b>
2. <b>PYTHON: TIMED TESTS I &amp; II</b>		
3. <b>VIDEO SUBMISSION</b>		

---

## COURSE 6 - CAPSTONE PROJECT

### 1. CAPSTONE PROJECT

1. AN OVERVIEW OF THE DOMAIN AND ASSOCIATED CONCEPTS	Solve an end-to-end real-life industry problem from a wide variety of domains.	4 WEEKS
2. PROBLEM STATEMENT		
3. EVALUATION RUBRIC		
4. MID SUBMISSION		
5. FINAL SUBMISSION		
6. SOLUTION		

---

## SPECIALISATION: DATA ENGINEERING

### COURSE 3: DATA ENGINEERING - I

#### 1. DATA MANAGEMENT AND RELATIONAL DATABASE MODELLING

1. ENTERPRISE DATA MANAGEMENT	Understand the concepts of Data Management and learn to model data from a Relational Database.	1 WEEK
2. RELATIONAL DATABASE MODELLING		
3. NORMAL FORMS AND ER DIAGRAMS		

---

#### 2. INTRODUCTION TO BIG DATA(OPTIONAL)

1. 4VS OF BIG DATA	This module you will learn what big data is, its various characteristics, and its determining factors. You will also get an idea of the various sources of big data and the wide range of big data applications in different industries such as retail, healthcare, and finance.	0 WEEK
2. BIG DATA: INDUSTRY CASE STUDIES		

---

### 3. INTRODUCTION TO CLOUD AND AWS SETUP

1. INTRODUCTION TO CLOUD	Understand what is cloud and setup your AWS account which will be required during the program.	1 WEEK
2. AWS SETUP		

---

### 4. INTRODUCTION TO HADOOP AND MAPREDUCE PROGRAMMING

1. CONCEPTS RETAILED TO DISTRIBUTED COMPUTING	Understand the world of distributed data processing and storage with Hadoop. Learn to write MapReduce jobs in Python.	1 WEEK
2. HADOOP DISTRIBUTED FILE SYSTEM		
3. MAPREDUCE PROGRAMMING IN PYTHON		

---

### 5. ASSIGNMENT (OPTIONAL)

1. INTRODUCTION, PROBLEM STATEMENT AND GRADING RUBRICS	Solve an assignment to brush up on the skills learnt so far.	0 WEEK
--	--	--------

---

### 6. NOSQL DATABASES AND APACHE HBASE NOSQL DATABASES AND MONGODB (OPTIONAL)

1. CONCEPTS OF NOSQL DATABASES	Learn the concepts of NoSQL databases. Understand the working of Apache HBase.	1 WEEK
2. INTRODUCTION TO APACHE HBASE		
3. HBASE PYTHON API		
4. COMPARISON OF NOSQL DATABASES		

---



## 7. DATA WAREHOUSING (OPTIONAL)

- |  |   |        |
|--|---|--------|
| 1. INTRODUCTION TO DATA WAREHOUSE AND DATA LAKES       | Understand the intricacies behind designing a data warehouse and a data lake for use case(s). | 0 WEEK |
| 2. DESIGNING DATA WAREHOUSING FOR AN ETL DATA PIPELINE |   |        |
| 3. DESIGNING DATA LAKE FOR AN ETL DATA PIPELINE        |   |        |
- 

## 8. DATA INGESTION WITH APACHE SQOOP AND APACHE FLUME

- |   |  |        |
|---|--|--------|
| 1. INTRODUCTION TO DATA INGESTION         | Get familiar with the challenges involved in data ingestion. Use Sqoop and Flume to ingest structured and unstructured data into Hadoop. | 1 WEEK |
| 2. STRUCTURED DATA INGESTION WITH SQOOP   |  |        |
| 3. UNSTRUCTURED DATA INGESTION WITH FLUME |  |        |
- 

## 9. MAPREDUCE PROGRAMMING ASSIGNMENT

- |   |  |        |
|---|--|--------|
| 1. PROBLEM STATEMENT AND SAMPLE DATASET | Practise MapReduce Programming on a Big Dataset. | 1 WEEK |
| 2. SOLUTION                             |  |        |
- 

## COURSE 4 - DATA ENGINEERING - II

### 1. HIVE & QUERYING

- |   |   |         |
|---|---|---------|
| 1. FUNDAMENTALS OF APACHE HIVE          | Manage and query a data warehouse with Apache Hive. Learn to write optimised HQL for large-scale data analysis. | 2 WEEKS |
| 2. WRITING HQL FOR DATA ANALYSIS        |   |         |
| 3. PARTITIONING AND BUCKETING WITH HIVE |   |         |
-

## 2. ASSIGNMENT (OPTIONAL)

1. INTRODUCTION, PROBLEM STATEMENT AND GRADING RUBRICS	Solve an assignment to brush up the skills learnt so far.	0 WEEK
--	---	--------

---

## 3. AMAZON REDSHIFT

1. DATA WAREHOUSING WITH REDSHIFT	Learn to deploy a Redshift cluster and use it for querying data.	1 WEEK
2. ANALYSE DATA WITH REDSHIFT		

---

## 4. INTRODUCTION TO APACHE SPARK

1. SPARK ARCHITECTURE	Get introduced to Apache Spark, a lightning fast big data processing engine.	1 WEEK
2. RDD, DATAFRAME API, SPARK SQL		

---

## 5. PROJECT: ETL DATA PIPELINE

1. INTRODUCTION AND PROBLEM STATEMENT	Make use of Sqoop, Redshift & Spark to design an ETL data pipeline.	2 WEEKS
2. GRADING RUBRICS AND SUBMISSION		

---

## 6. AWS CLOUD INFRASTRUCTURE (OPTIONAL)

1. THE AWS CLOUD PLATFORM	Do a deep dive into AWS Cloud.	0 WEEK
2. BUILDING AND DEPLOYING VIRTUAL MACHINES		
3. AWS CLOUD STORAGE SOLUTIONS		
4. APPLICATION DEPLOYMENT		
5. CLOUD ADMINISTRATION AND SECURITY		
6. LOAD BALANCING AND BACKUP STRATEGIES		
7. CLOUD AUTOMATION		

*\*The Curriculum is subject to change as per the inputs from university or industry experts*

## COURSE 5 - DATA ENGINEERING - III

### 1. OPTIMISING SPARK FOR LARGE-SCALE DATA PROCESSING

- |   |   |        |
|---|---|--------|
| 1. RUNNING SPARK ON MULTINODE CLUSTER   | Use PySpark to create large-scale data processing applications. | 1 WEEK |
| 2. SPARK MEMORY & DISK OPTIMISATION     |   |        |
| 3. OPTIMISING SPARK CLUSTER ENVIRONMENT |   |        |
- 

### 2. APACHE FLINK(OPTIONAL)

- |  |   |        |
|--|---|--------|
| 1. INTRODUCTION TO APACHE FLINK        | Get Introduced to Apache Flink and learn query batch data.    | 0 WEEK |
| 2. BATCH DATA PROCESSING WITH FLINK    |   |        |
| 3. STREAM PROCESSING WITH APACHE FLINK | Use DataStream API to create a stream processing application. |        |
| 4. SQL API                             |   |        |
- 

### 3. REAL-TIME DATA STREAMING WITH APACHE KAFKA

- |   |   |        |
|---|---|--------|
| 1. INTRO TO REAL-TIME DATA PROCESSING ARCHITECTURES | Understand the producer-consumer architecture of Apache Kafka. Learn to set up a Kafka cluster for managing real-time data. | 1 WEEK |
| 2. FUNDAMENTALS OF APACHE KAFKA                     |   |        |
| 3. SETTING UP KAFKA PRODUCER AND CONSUMER           |   |        |
| 4. KAFKA CONNECT API & KAFKA STREAMS                |   |        |
-

## 4. REAL-TIME DATA PROCESSING USING SPARK STREAMING

1. SPARK STREAMING ARCHITECTURE	Learn about the real-time data processing architecture of Apache Spark. Build Spark Streaming applications to process data in real-time.	1 WEEK
2. SPARK STREAMING APIS		
3. BUILDING STREAM PROCESSING APPLICATION WITH SPARK		
4. COMPARISION BETWEEN SPARK STREAMING AND FLINK		

---

## 5. ASSIGNMENT (OPTIONAL)

1. INTRODUCTION, PROBLEM STATEMENT AND GRADING RUBRICS	Solve an assignment to brush up on the skills learnt so far.	0 WEEK
--	--	--------

---

## 6. BUILDING AUTOMATED DATA PIPELINES WITH AIRFLOW

1. FUNDAMENTS OF AIRFLOW	Automate Data Pipelines with Airflow.	1 WEEK
2. WORKFLOW MANAGEMENT WITH AIRFLOW		
3. AUTOMATING AN ENTIRE DATA PIPELINE WITH AIRFLOW		

---

## 7. ANALYTICS USING PYSPARK

1. EXPLORATORY DATA ANALYSIS WITH PYSPARK	Use PySpark to do EDA and Predictive Analysis using Spark's ML library.	1 WEEK
2. PREDICTIVE ANALYSIS WITH SPARK MLLIB		

---

## 8. PROJECT: REAL-TIME DATA PROCESSING

1. INTRODUCTION AND PROBLEM STATEMENT	Build an end-to-end real-time data processing application using Spark Streaming and Kafka.	1 WEEK
2. GRADING RUBRICS AND SUBMISSION		

## COURSE 6 - CAPSTONE PROJECT

### CAPSTONE PROJECT

1. AN OVERVIEW OF THE  
DOMAIN AND ASSOCIATED  
CONCEPTS

The capstone project will stitch all the  
components of data engineering together.

**4 WEEKS**

2. PROBLEM STATEMENT

3. EVALUATION RUBRIC

4. MID SUBMISSION

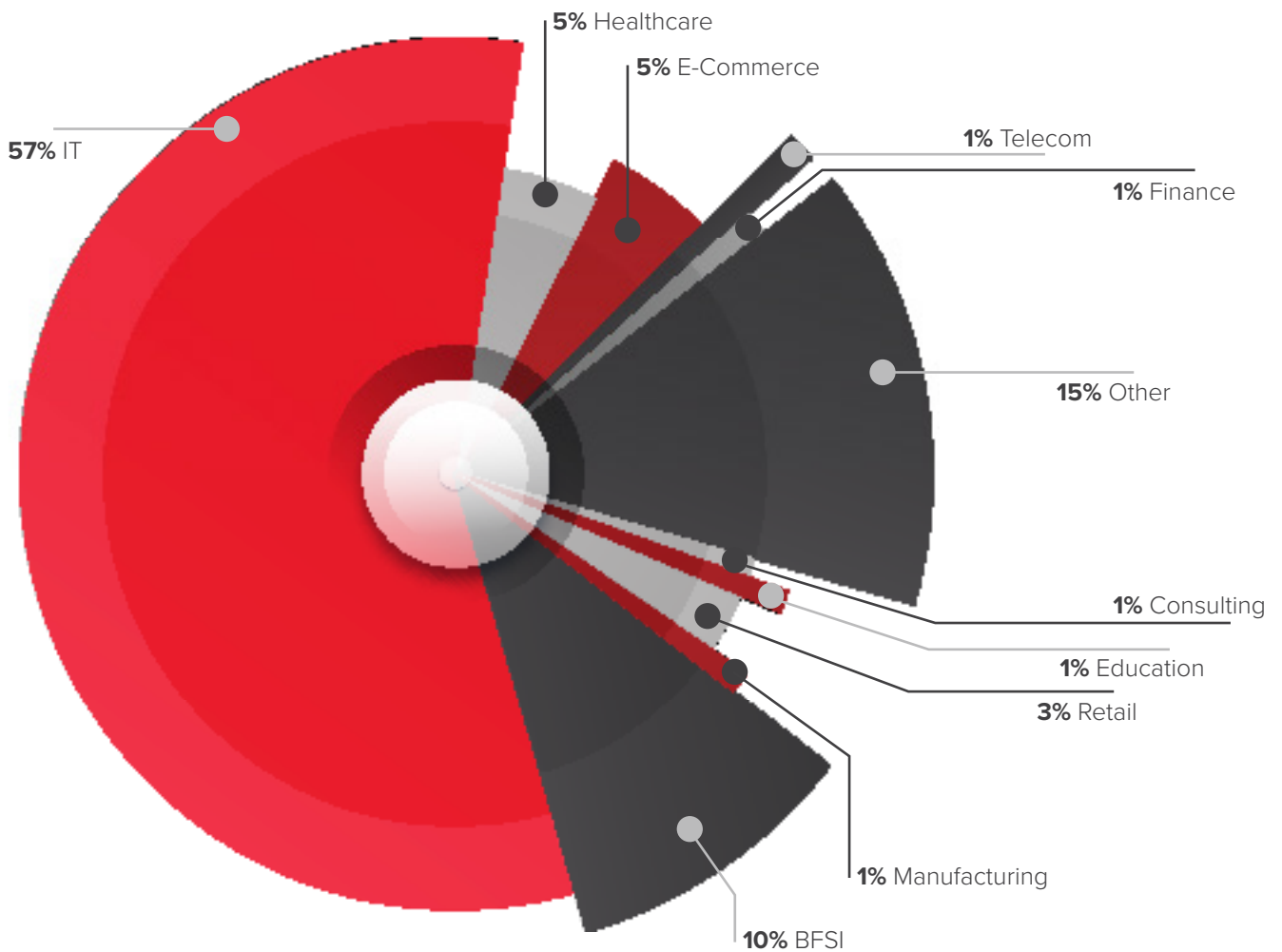
5. FINAL SUBMISSION

6. SOLUTION

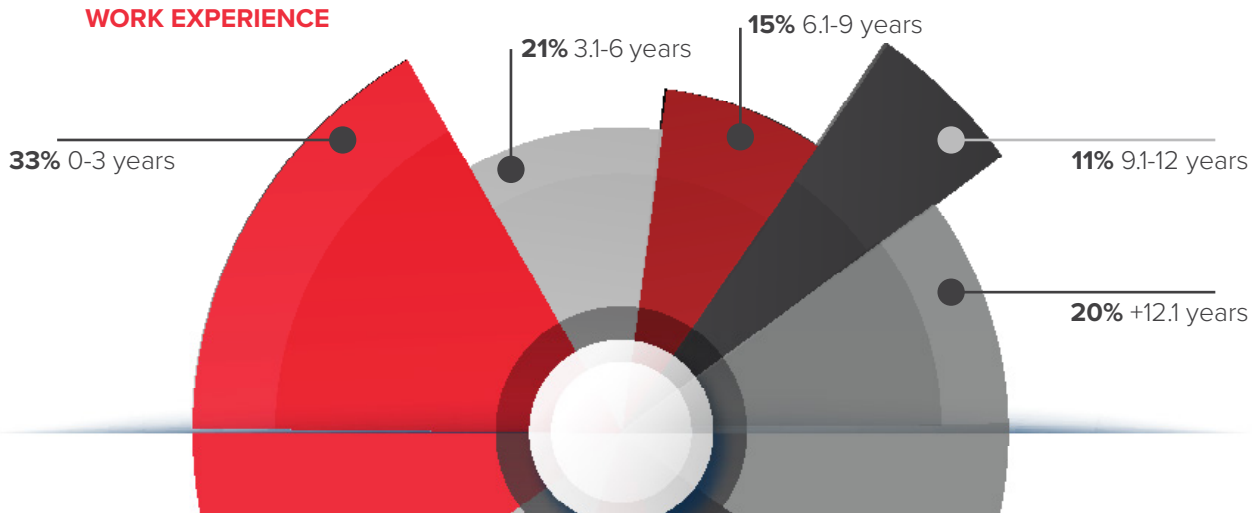
---

# Meet the Class

## INDUSTRIES OUR STUDENTS COME FROM



## WORK EXPERIENCE



# Career Support

## Jobs on Career Centre

Career Centre offers upGrad jobs across experience levels and CTC ranges.

- Easy apply feature for upGrad hiring partner vacancies.
- Create a resume at profile builder with one click to apply for various jobs.

## upGrad Elevate

- Recruitment Drive to connect you with the best talent admirers in the industry
- Get access to a wide range of opportunities and find the perfect job
- Apply your learnings to real industry problems

## Interview Preparation

Pre-recorded content on topics such as:

- Profile building, communications, etc.
- Problem-solving approach
- Approaching guesstimates
- Domain-specific interview question bank and much more.

## Profile Builder (AI-Powered)

An easy-to-use Resume, LinkedIn and Cover Letter preparation tool.

- Resume Score: AI-Driven Resume Score
- Real-time recommendations to improve
- Match your resume to the JD and check fitment
- LinkedIn Profile Review
- Cover Letter creation

## Just-In-Time Interview Prep (JIT)

For upcoming job interviews, JITs are conducted within 48 hours for eligible programs.

- Tailored to the job role and target domain
- Real-time feedback and tips for improvement

## High-Performance Coaching

Dedicated coaches working with you to identify best-suited career opportunities.

- Help you define your value proposition
- Lay out a Career Path and help you adhere to your timelines and goals
- Help you with interview preparations, finding jobs in the market, salary negotiations and other preparation as required

## Personalised Industry Session

90-minute sessions over the weekend by leading industry experts.

- Session categories: Career, Technical and Communications
- Doubt resolution
- Develop proof of concepts and apply theoretical concepts in the real world
- Assess skill levels
- Peer Networking
- Classroom element
- Business communication sessions and much more

## Career Mentorship Sessions

Get personalised career advice through 1-1 sessions with industry experts.

- Goal setting for better employment results

## PROGRAM SUMMARY

**10 Million+**

Learners

**INR 1.23 CR**

Highest Salary Package

**433%**

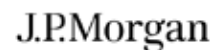
Highest Salary Hike

**50%**

Average Salary Hike

# Our Alumni Work at

upGrad has a network of over 100 companies that look to recruit graduates from our programs. Some of these well-known companies include:





# Career Transitions

**Sandeep Varma**

Assistant Manager



Assistant Manager  
Business Analyst(HSBC)

**Joseph Jeffrey**

Development Engineer



Quality Leader  
(IKEA of Sweden AB)

**Govind**

Consultant



Senior  
Associate Consultant  
(INFOSYS)

**Vandana Maurya**

Database Tech Lead



SQL Server Database  
Administration  
(Citius Tech Pvt Ltd)

**Souvik Mitra**

Data Analyst



Business Analyst  
(Xiaomi)

**Deepak Baliya**

Software Engineer



Sr Software Engineer  
(Oak North)

**Rohit Ambasta**

Data Analyst, Vendor BI  
(contractual position)



Senior Data Analyst  
(GE Renewable Energy))  
(Oct 2019)

**Prateek Aneja**

Software Engineer



Machine Learning  
Consultant  
(Tardid Technologies)

# Experience upGrad Offline



## UPGRAD BASECAMPS (PRE-COVID)

Held across all major cities in India, upGrad basecamps bring together learners, faculty and industry experts for a power-packed day of activities, career-building sessions and live group projects. Get to know your peers and faculty and hone your networking skills in an exciting environment.

## CAREER FAIRS

Attend regular hiring drives in major cities across India, giving you the opportunity to interview with upGrad's 300+ hiring partners, ensuring you get every opportunity you deserve.





## HACKATHONS

Team up and put your learning to use with our offline Hackathons: designed to help you apply concepts and meet, network, and grow!



# Hear from Our Learners

**Kunwar Alok, Experience: 15+ Years**

*"You may not believe it, but I had never done coding in my life. I did it during this course and was thrilled to see the outcomes of those codes. Just the way I used to get happy after solving good (tough) maths problem during my school years. Thanks to upGrad for providing a great service to people like us who at the age of 43 can dream of study with budding talents around."*

**Sachin Aggarwal, Experience: 18+ Years**

*"Learning with IITB and upGrad has been an experience like no other. Being an online program, you have your worries about how the program and teaching methods will be. My favourite part about the learning experience has been programming through well-designed and thoughtful content shared by IITB professors and industry experts on upGrad platforms. Kudos to upGrad."*

**Sidharth Mahapatra, Experience: 3 Years**

*"The concepts of R programming and Machine Learning will be taught by Prof. Chandrasekhar Ramanathan and Prof. G Srinivasaraghavan respectively. Both of them have been listed in the top twenty most prominent Data Science academics published by Analytics India Magazine. So you need not worry about the quality of teaching in this program."*

**Harkirat Dhillon, Experience: 8 Years**

*"A dedicated studying regime is the key to be successful and pass the program. This program will help build a strong foundation for a successful transition into Data Science. Additionally, participating in Hackathons and Kaggle competitions to solve real-world problems will definitely give you an edge and land a job if one is willing to work hard."*

**Shravani Shahapure, Experience: 16 Years**

*"For someone who really wants to pursue a career in the field of Data Science, it is worth opting for the complete course by IIITB and upGrad. IIITB and upGrad's online program on Data Science gives many opportunities and develops students for their future as they provide the best professors, thought-provoking assignments and case studies."*

**Sagar Tekwani, Experience: 2 Years**

*"A very well-structured and well-balanced program content which you won't get in other programs/nano-degrees. Being a beginner in DS, I found the structure of the Executive PG Programme from IIITB and upGrad most helpful. They even teach you most of the prerequisites with prep sessions before you even start the course. Being a working professional, it was neither too difficult nor too easy to keep up with the pace of the course."*



# Program Details and Admission Process

## PROGRAM DURATION AND FORMAT

12 Months | Blended

## PROGRAM START DATES

Please refer to the website for program start dates.

[www.upgrad.com/data-science-pgd-iiitb/](http://www.upgrad.com/data-science-pgd-iiitb/)

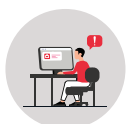
## PROGRAM FEE

Please refer to the program page for update fee details

## ELIGIBILITY

Bachelor's Degree with 50% or equivalent passing marks. No coding experience is required.

## WEEKLY COMMITMENT (12-15 hours/week)



### 6-7 HOURS

Asynchronous learning time.



### 6-7 HOURS

Assignments and projects.



### 1 LIVE SESSION

Every two weeks.

## SELECTION PROCESS



### STEP 1: Selection Test

Fill out an application and take a short 17-minute online test with 11 questions.



### STEP 2: Review and Shortlisting of Suitable Candidates

Our faculty will review all applications, considering the educational and professional background of an applicant and review the test scores where applicable. Following this, Offer Letters will be rolled out so you are assured of a great peer group to learn and network with.



### STEP 3: Enrollment for Access to Prep Content

Make a quick block payment with assistance from our loan partners where required, receive immediate access to the prepped content and begin your upGrad journey.

## FOR FURTHER INFORMATION, CONTACT

admissions@upgrad.com  
1800 210 2020  
We are available 24\*7

Disclaimer: Program fee and payment options are subject to change. Please refer to the website for updated details or speak to our admission counsellor.

## COMPANY INFORMATION

upGrad Education Private Limited  
Nishuvi, 75, Annie Besant Road,  
Worli, Mumbai - 400018.