upGrad

iiit-b International Institute of Information Technology Bangalore

Executive PG Program in

DATA SCIENCE

Get the Whole Picture

Table of Contents

About upGrad

upGrad has delivered over 20 million hours of learning, delivering programs by collaborating with universities across the world including Duke CE, IIT Madras, IIIT 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.

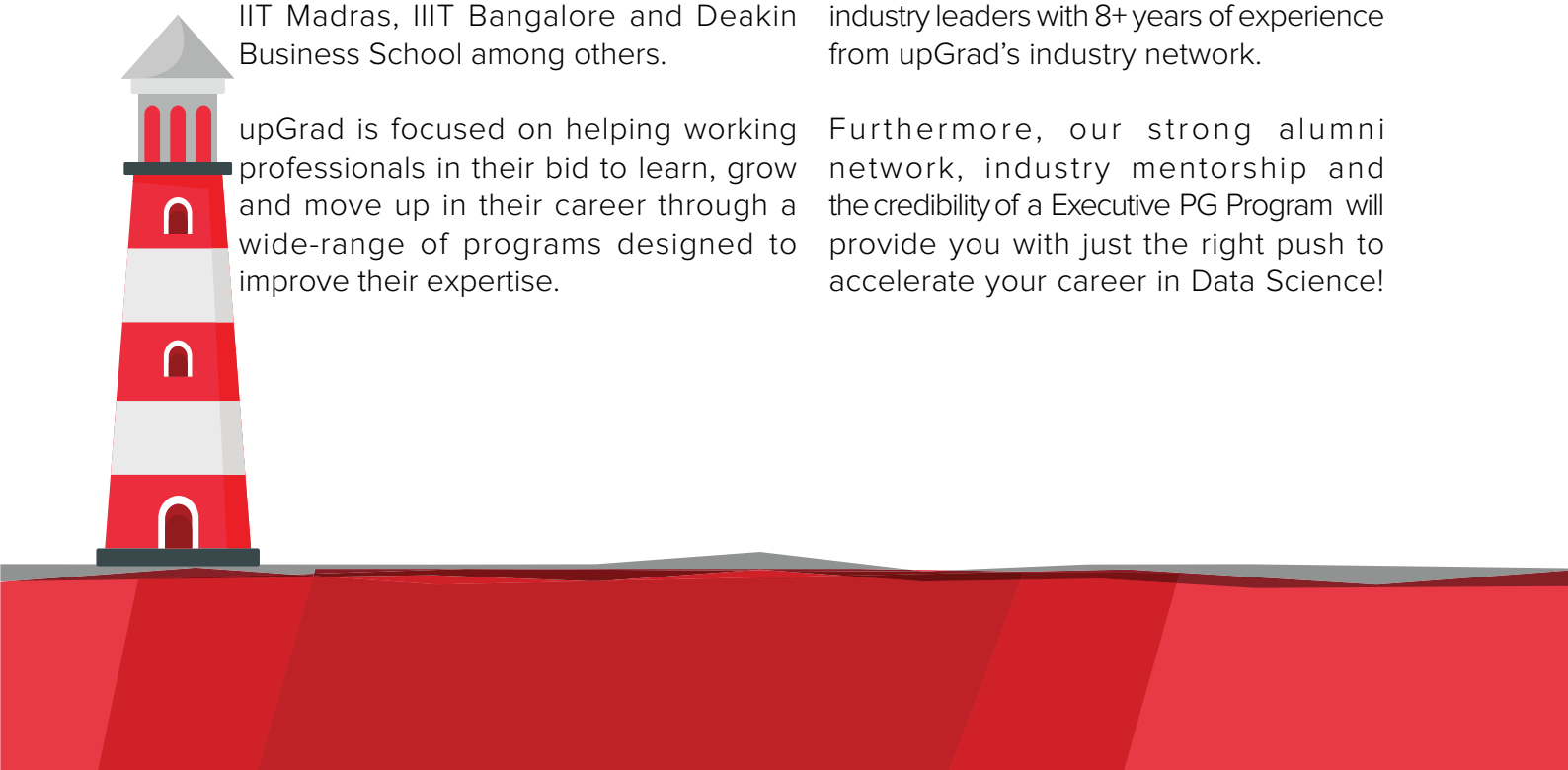
Since inception, upGrad has delivered over 20 million hours of learning, delivering programs by collaborating with universities across the world including Duke CE, IIT Madras, IIIT Bangalore and Deakin Business School among others.

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

IIITB is a renowned university offering programs specialising in data science, machine learning and artificial intelligence. The IIITB 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 alumni network, industry mentorship and the credibility of a Executive PG Program will provide you with just the right push to accelerate your career in Data Science!



Why upGrad?

400%
Highest Raise

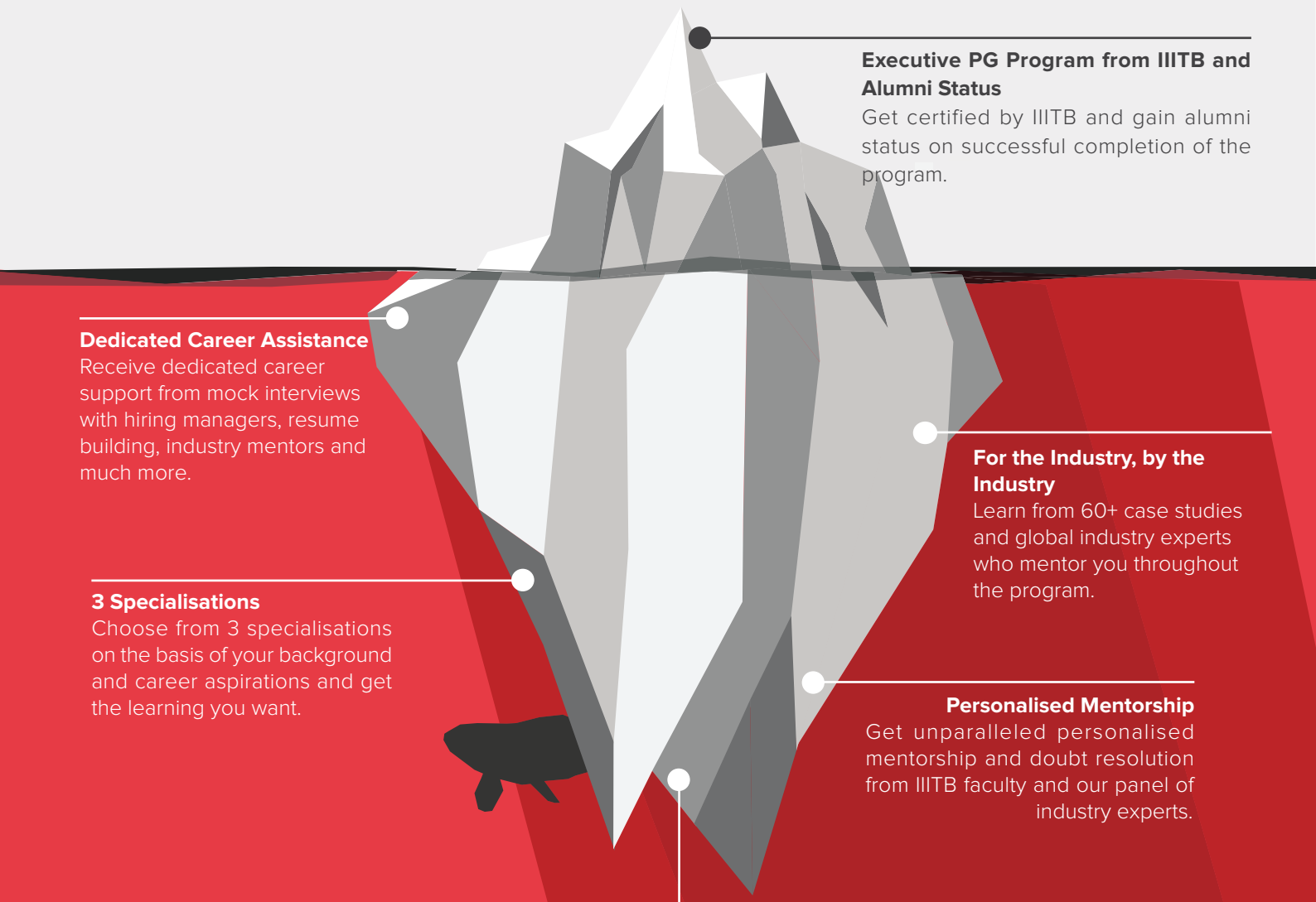
58%
Avg. Raise

87%
Program Completion Rates

40,000+
Learners

700+
Industry Experts

Program Highlights



Executive PG Program from IIITB and Alumni Status
Get certified by IIITB and gain alumni status on successful completion of the program.

Dedicated Career Assistance
Receive dedicated career support from mock interviews with hiring managers, resume building, industry mentors and much more.

For the Industry, by the Industry
Learn from 60+ case studies and global industry experts who mentor you throughout the program.

3 Specialisations
Choose from 3 specialisations on the basis of your background and career aspirations and get the learning you want.

Personalised Mentorship
Get unparalleled personalised mentorship and doubt resolution from IIITB faculty and our panel of industry experts.

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

Hindol Basu
CEO, Actify Data Labs

An alumnus of IIT and IIM with over 13 years of experience in analytics with industry leaders such as Citigroup and Tata Industries.

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

A gold medallist from IIM Bangalore, an alumnus of IIT Madras and London Business School, 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.

Sameer Dhanrajani
Co-Founder and CEO, AIQRATE

Sameer Dhanrajani is an AI and Analytics evangelist for Fortune 500 companies who has won several industry awards in the field of analytics.

Prof. Debabrata Das
Director, IIITB

Dr. Debabrata Das is serving as Director of IIIT Bangalore (IIITB). He has completed his Ph.D. degree from the Indian Institute of Technology Kharagpur. His main areas of research interest are IoT and Wireless Access Network's MAC, QoS, Power saving.

Prof. G. Srinivasaraghavan
Professor, IIIB

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

Ujjyaini Mitra
Head of Analytics, Zee5

An alumna of McKinsey and Co., Flipkart, and Bharti Airtel with over 11 years of experience.

Dinesh Babu Jayagopi
Associate Professor, IIITB

Prof. Dinesh has a PhD from EPFL Switzerland, MSc from IISc Bangalore in System Science and Signal Processing and BTech.

Kalpana Subbaramappa
Ex-AVP, Genpat

Kalpana is the ex-AVP of Decision Sciences at Genpact with over 20 years of experience.

Ankit Jain
Sr. Research Scientist, Uber

An alumnus of IIT Bombay, UCB and Harvard Business School with over 9 years of experience.

Sajan Kedia
Lead Data Scientist (Pricing), Myntra

An alumnus of IIT with over 7 years of experience at Watson at IBM Research, startups and Myntra.

Mirza Rahim Baig
Lead Business Analytics, Flipkart

Advanced Analytics professional with 8+ years of experience as a consultant in the e-commerce and healthcare domains.

Colin Hagemeyer
Data Scientist, Mentor Analytics,
San Jose, CA, USA

Colin has done his PhD from University of California. He brings with him 10+ years of rich international experience in the field of Data.

Behzad Ahmadi
Data Scientist, Walmart Labs
San Francisco Bay Area, USA

Behzad has done his PhD from New Jersey Institute of Technology. He has worked with Qualcomm, US and Walmart, US on different data driven roles.

Prof. Dr. V. Sridhar
Faculty In Charge, CPE, IIITB

Dr. Sridhar has a Ph.D. from the University of Iowa, U.S.A. He has been a member of Government of India committees on Telecom and IT and has published many peer reviewed articles in telecom and information systems journals.

upGrad Learning Experience

Coaching

- Student Support Team & upGrad Buddy
- Weekly real-time doubt clearing sessions
- Live Discussion forum for peer-to-peer doubt resolution monitored by technical experts
- Peer-to-peer international networking opportunities with an alumni pool of 10,000+
- Lab walk-throughs of 15+ industry-driven case studies
- 6 Employability Tests for industry readiness
- Access to the program for up to 3 years

Mentorship

60+ live interactive sessions with industry experts, fortnightly personalised group (1:8) mentorship sessions with global industry experts for pro-active mentoring.

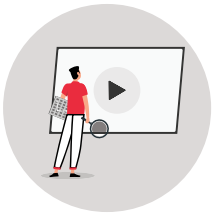
Format

Online format with weekly live sessions from global industry experts to help with topic walk-throughs, doubt resolution and personalised project feedback.

Hands-On Projects and Case Studies

60+ Industry Projects, Case Studies and Capstone Project to learn from. This will help the learners get job ready.

Industry Projects



IMDb Movie Analysis

Uber Supply-Demand Gap

Lead Scoring

Fraud Detection

Creditworthiness of Customers

Speech Recognition

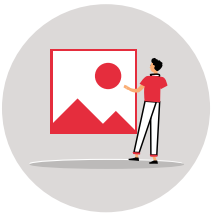
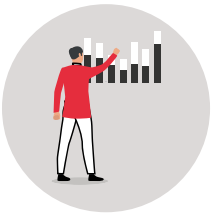


Image Captioning

Gesture Recognition

Social Media Listening

Telecom Churn



Interactive Market Campaign Analysis

Retail Giant Sales Forecasting

And many more!

Learning Path

Preparatory Course
2 weeks
Tools: Python, Excel

Data Toolkit
11 weeks
Tools: Python, Excel, mySQL

Machine Learning
9 weeks
Tools: Python, Excel

Choose any of the 3 Specialisations
25 weeks (with 6 weeks of Capstone)

Deep Learning
Tools: Python, Excel, TensorFlow

Business Intelligence/ Data Analytics
Tools: Python, Power BI, Excel, mySQL, MongoDB, Shiny, Tableau

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

Executive PG Program in Data Science (Deep Learning)

Executive PG Program in Data Science (Business Intelligence/ Data Analytics)

Executive PG Program 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

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

2. PROGRAMMING IN PYTHON

1. **LOGIC AND SYNTAX BUILDING**
2. **DATA STRUCTURES: LISTS, STRINGS, DICTIONARIES, AND STACKS**
3. **TIME COMPLEXITY**
4. **SEARCHING AND SORTING**
5. **TWO POINTERS**
6. **RECURSION**

Learn how to approach and solve logical problems using programming.

1 WEEK

3. PYTHON FOR DATA SCIENCE

1. **INTRODUCTION TO NUMPY**
2. **INTRODUCTION TO MATPLOTLIB**
3. **INTRODUCTION TO PANDAS**
4. **GETTING AND CLEANING DATA**

Learn how to manipulate datasets in Python using Pandas which is the most powerful library for data preparation and analysis.

1 WEEK

4. DATA VISUALIZATION IN PYTHON

1. **INTRODUCTION TO DATA VISUALIZATION**
2. **DATA VISUALISATION USING SEABORN**

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

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

5. EXPLORATORY DATA ANALYSIS

1. **DATA SOURCING**
2. **DATA CLEANING**
3. **UNIVARIATE ANALYSIS**
4. **BIVARIATE ANALYSIS AND MULTIVARIATE ANALYSIS**

Learn how to find and analyse the patterns in the data to draw actionable insights.

2 WEEKS

6. CREDIT EDA CASE STUDY

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

Solve a real industry problem through the concepts learnt in exploratory data analysis.

1 WEEK

7. INFERENTIAL STATISTICS

1. **BASICS OF PROBABILITY**
2. **DISCRETE PROBABILITY DISTRIBUTIONS**
3. **CONTINUOUS PROBABILITY DISTRIBUTIONS**
4. **CENTRAL LIMIT THEOREM**

Build a strong statistical foundation and learn how to 'infer' insights from a huge population using a small sample.

1 WEEK

8. HYPOTHESIS TESTING

1. **CONCEPTS OF HYPOTHESIS TESTING - I: NULL AND ALTERNATE HYPOTHESIS, MAKING A DECISION, AND CRITICAL VALUE METHOD**
2. **CONCEPTS OF HYPOTHESIS TESTING - II: P-VALUE METHOD AND TYPES OF ERRORS**
3. **INDUSTRY DEMONSTRATION OF HYPOTHESIS TESTING: TWO-SAMPLE MEAN AND PROPROTION TEST, A/B TESTING**

Understand how to formulate and validate hypotheses for a population to solve real-life business problems.

1 WEEK

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

9. DATA ANALYSIS USING SQL

1. DATABASE DESIGN
2. DATABASE CREATION IN MYSQL WORKBENCH
3. QUERYING IN MYSQL
4. JOINS AND SET OPERATIONS

Data in companies is definitely not stored in excel sheets! Learn the fundamentals of database and extract information from RDBMS using the structured query language.

1 WEEK

10. ADVACED SQL & 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

11. SQL ASSIGNMENT: RSVP MOVIES

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

In this assignment, you will work on a movies dataset using SQL to extract exciting insights.

1 WEEK

COURSE 2 - MACHINE LEARNING - I

1. LINEAR REGRESSION

1. SIMPLE LINEAR REGRESSION
2. SIMPLE LINEAR REGRESSION IN PYTHON
3. MULTIPLE LINEAR REGRESSION
4. MUTLIPLE LINEAR REGRESSION IN PYTHON
5. INDUSTRY RELEVANCE OF 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

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

2. LINEAR REGRESSION ASSIGNMENT

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

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

---

3. LOGISTIC REGRESSION

1. **UNIVARIATE LOGISTIC REGRESSION**
2. **MULTIVARIATE LOGISTIC REGRESSION: MODEL BUILDING AND EVALUATION**
3. **LOGISTIC REGRESSION: INDUSTRY APPLICATIONS**

Learn your first binary classification technique by determining which customers of a telecom operator are likely to churn versus who are not to help the business retain customers.

2 WEEKS

---

4. CLASSIFICATION USING DECISION TREES

1. **INTRODUCTION TO DECISION TREES**
2. **ALGORITHMS FOR DECISION TREES CONSTRUCTION**
3. **HYPERPARAMETER TUNING IN DECISION TREES**

Learn how the human decision making process can be replicated using a decision treeand tune it to suit your needs.

1 WEEK

---

5. UNSUPERVISED LEARNING: CLUSTERING

1. **INTRODUCTION TO CLUSTERING**
2. **K-MEANS CLUSTERING**
3. **HIERARCHICAL CLUSTERING**
4. **OTHER FORMS OF CLUSTERING: K-MODE, K-PROTOTYPE, DB SCAN**

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

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

6. BASICS OF NLP AND TEXT MINING

1. **REGEX AND INTRODUCTION TO NLP**
2. **BASIC LEXICAL PROCESSING**
3. **ADVANCED LEXICAL PROCESSING**

Do you get annoyed by the constant spams in your mailbox? Wouldn't it be nice if we had a program to check your spellings? In this module learn how to build a spell checker & spam detector using techniques like phonetic hashing, bag-of-words, TF-IDF, etc.

0 WEEK

---

7. BUSINESS PROBLEM SOLVING

1. **INTRODUCTION TO BUSINESS PROBLEM SOLVING**
2. **BUSINESS PROBLEM SOLVING: CASE STUDY DEMONSTRATIONS**

Learn how to approach open ended real world problems using data as a lever to draw actionable insights.

1 WEEK

---

8. CASE STUDY: LEAD SCORING

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

Help the Sales team of your company identify which leads are worth pursuing through this classification case study.

1 WEEK

SPECIALISATION: DEEP LEARNING

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**
2. **MODEL EVALUATION**
3. **MODEL SELECTION: BEST PRACTICES**

Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, alongwith general machine learning techniques like feature engineering, model evaluation, and many more.

1 WEEK

---

4. PRINCIPAL COMPONENT ANALYSIS

1. **PRINICIPAL COMPONENT ANALYSIS AND SINGULAR VALUE DECOMPOSITION**
2. **PRINCIPAL COMPONENT ANALYSIS IN PYTHON**

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

5. ADVANCED REGRESSION

1. GENERALIZED LINEAR REGRESSION
2. REGULARIZED REGRESSION

In this module, take a more advanced look at regression models and learn the concepts related to regularization.

1 WEEK

---

6. ADVANCED ML CASE STUY

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

Build a regularized regression model to understand the most important variables to predict the house prices in Australia.

1 WEEK

COURSE 4 - ADVANCED MACHINE LEARNING AND DEEP LEARNING

1. TIME SERIES ANALYSIS

1. INTRODUCTION TO TIME SERIES AND ITS COMPONENTS
2. WORKING WITH STATIONARY TIME SERIES
3. END-TO-END ANALYSIS OF TIME SERIES

In this module, you will learn how to analyse and forecast a series that varies with time.

2 WEEKS

---

2. INTRODUCTION TO NEURAL NETWORKS AND ANN

1. STRUCTURE OF NEURAL NETWORKS
2. FEED FORWARD IN NEURAL NETWORKS
3. BACKPROPAGATION IN NEURAL NETWORKS
4. MODIFICATIONS TO NEURAL NETWORKS
5. HYPERPARAMETER TUNING IN NEURAL NETWORKS

Learn the most sophisticated and cut-ting-edge technique in machine learning - Artificial Neural Networks or ANNs

3 WEEKS

3. NEURAL NETWORK ASSIGNMENT

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

Build a neural network from scratch in Tensorflow to identify the type of skin cancer from image.

1 WEEK

COURSE 5 -ADVANCED DEEP LEARNING AND COMPUTER VISION

1. CONVOLUTIONAL NEURAL NETWORKS

1. INTRODUCTION TO CONVOLUTIONAL NEURAL NETWORKS
2. BUILDING CNNS WITH PYTHON AND KERAS
3. CNN ARCHITECTURES AND TRANSFER LEARNING
4. STYLE TRANSFER AND OBJECT DETECTION

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.

2 WEEKS

---

2. CONVOLUTIONAL NEURAL NETWORKS -INDUSTRY APPLICATIONS

1. INDUSTRY DEMONSTRATION: USING CNNS WITH FLOWERS IMAGES
2. INDUSTRY DEMONSTRATION: USING CNNS WITH X-RAY IMAGES

Apply CNNs to Computer Vision tasks like detecting anomalies in chest X-Ray scans.

1 WEEK

---

3. OBJECT DETECTION & IMAGE SEGMENTATION (OPTIONAL)

1. FUNDAMENTALS OF OBJECT DETECTION
2. REGION-BASED DETECTORS
3. ONE-SHOT DETECTORS
4. CUSTOM OBJECT DETECTION
5. SEMANTIC SEGMENTATION

Learn the applications of DL in computer vision through industry-relevant detection algorithms such as RCNNs, YOLO and SSD.

0 WEEKS

4. RECURRENT NEURAL NETWORKS

1. **WHAT MAKES A NEURAL NETWORK RECURRENT**
2. **VARIANTS OF RNNS: BIDIRECTIONAL RNNS AND LSTMS**
3. **BUILDING RNNS IN PYTHON**

Ever wondered what goes behind machine translation, sentiment analysis, speech recognition? Learn how RNN helps in these areas having sequential data like text, speech, videos, and a lot more.

1 WEEK

---

5. GESTURE RECOGNITION

1. **TWO ARCHITECTURES: 3D CONVS AND CNN-RNN STACK**
2. **UNDERSTANDING GENERATORS**
3. **STARTER CODE WALKTHROUGH**
4. **PROBLEM STATEMENT AND FINAL SUBMISSION**

Make a Smart TV system which can control the TV with user's hand gestures as the remote control

2 WEEKS

COURSE 5 - CAPSTONE PROJECT

1. 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, Speech Recognition among many others.

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**
2. **AMAZON WEB SERVICES**
3. **BIG DATA STORAGE AND PROCESSING - HADOOP**
4. **EMR CLUSTER IN AWS**

Understand the basics of big data and cloud and learn to work with an EMR cluster on a cloud-based service.

1 WEEK

4. ANALYTICS USING SPARK

1. **EXPLORATORY DATA ANALYSIS WITH PYSPARK**
2. **PREDICTIVE ANALYSIS WITH SPARK MLLIB**

Use PySpark to do EDA and Predictive Analysis using Spark's ML library.

2 WEEKS

---

5. BIG DATA CASE STUDY

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

Use your analytics skills to work on a large dataset in cloud to solve an industry problem.

1 WEEK

COURSE 4: DATA VISUALISATION AND STORYTELLING

1. VISUALISATION USING TABLEAU

1. **DATA EXPLORATION IN TABLEAU**
2. **VISUALISING AND ANALYSING DATA IN TABLEAU WITH BASIC PLOTS**

Learn basic visualisation techniques using the most in-demand visualization tool in the industry.

1 WEEK

---

2. ADVANCED EXCEL

1. **EXCEL FUNCTIONS**
2. **DATA ANALYSIS IN EXCEL**
3. **ADVANCED TOOLS AND VISUALISATIONS**

Learn the advanced concepts in Excel and start to perform data analysis like a pro!

1 WEEK

---

3. VISUALISATION USING POWERBI

1. **POWERBI: INTRODUCTION AND SETUP**
2. **VISUALISING AND ANALYSING DATA IN POWERBI**
3. **DATA TRANSFORMATIONS USING POWERBI**

Take your visualization game a step forward by understanding how to operate PowerBI.

1 WEEK

---

4. STRUCTURED PROBLEM SOLVING USING FRAMEWORKS

1. **INTRODUCTION TO STRUCTURED PROBLEM SOLVING**
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.**

Learn how to attack a business problem using various structured frameworks like 5W, 5WHYs, and SPIN.

1 WEEK

---

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

---

5. 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:ADVANCED PROBLEM SOLVING AND PROGRAMMING

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

1. IN-BUILT DATA STRUCTURES
2. STACK
3. QUEUE
4. TREES

Learn user defined data structures -Stack, Queue, Trees in Python that help in advanced data manipulation

1 WEEK

---

2. SEARCHING AND SORTING

1. SEARCHING
2. SORTING
3. TWO POINTERS

Learn most fundamental searching and sorting algorithms and design techniques

1 WEEK

3. ALGORITHM ANALYSIS + RECURSION

1. **ALGORITHM ANALYSIS**
2. **TIME AND SPACE COMPLEXITY**
3. **RECURSION**

Learn how to assess the efficiency your code using algorithm analysis techniques and learn to write recursive algorithms

1 WEEK

---

4. ADVANCED DATABASE PROGRAMMING USING PANDAS

1. **ADVANCED DATA WRANGLING WITH PANDAS - I**
2. **ADVANCED DATA WRANGLING WITH PANDAS - II**

Learn and implement advanced wrangling functions and techniques in Pandas related to date-time, multi-columns aggregation, hierarchical indexing, and more.

1 WEEK

---

5. PYTHON & SQL LAB

1. **SQL: TIMED TEST + ASSIGNMENT**
2. **PYTHON: TIMED TESTS I & II**
3. **VIDEO SUBMISSION**

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.

2 WEEKS

COURSE 6 - CAPSTONE PROJECT

1. 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**

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

4 WEEKS

SPECIALISATION: DATA ENGINEERING

COURSE 3: DATA ENGINEERING - I

1. DATA MANAGEMENT AND RELATIONAL DATABASE MODELLING

1. **ENTERPRISE DATA MANAGEMENT**
2. **RELATIONAL DATABASE MODELLING**
3. **NORMAL FORMS AND ER DIAGRAMS**

Understand the concepts of Data Management and learn to model data from a Relational Database.

1 WEEK

---

2. INTRODUCTION TO BIG DATA(OPTIONAL)

1. **4VS OF BIG DATA**
2. **BIG DATA: INDUSTRY CASE STUDIES**

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

---

3. INTRODUCTION TO CLOUD AND AWS SETUP

1. **INTRODUCTION TO CLOUD**
2. **AWS SETUP**

Understand what is cloud and setup your AWS account which will be required duing the program.

1 WEEK

---

4. INTRODUCTION TO HADOOP AND MAPREDUCE PROGRAMMING

1. **CONCEPTS RETAILED TO DISTRIBUTED COMPUTING**
2. **HADOOP DISTRIBUTED FILE SYSTEM**
3. **MAPREDUCE PROGRAMMING IN PYTHON**

Understand the world of distributed data processing and storage with Hadoop. Learn to write MapReduce jobs in Python.

1 WEEK

5. ASSIGNMENT (OPTIONAL)

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

---

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

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

---

7. DATA WAREHOUSING (OPTIONAL)

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

---

8. DATA INGESTION WITH APACHE SQOOP AND APACHE FLUME

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

9. MAP REDUCE PROGRAMMING ASSIGNMENT

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

COURSE 4 - DATA ENGINEERING - II

1. HIVE & QUERYING

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

---

2. ASSIGNMENT (OPTIONAL)

1. NTRODUCTION, 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 2. ANALYZE DATA WITH REDSHIFT	Learn to deploy a Redshift cluster and use it for querying data.	1 WEEK

---

4. INTRODUCTION TO APACHE SPARK

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

5. PROJECT: ETL DATA PIPLINE

1. **INTRODUCTION AND PROBLEM STATEMENT**
2. **GRADING RUBRICS AND SUBMISSION**

Make use of Sqoop, Redshift & Spark to design an ETL data pipeline.

2 WEEKS

---

6. AWS CLOUD INFRASTRUCTURE (OPTIONAL)

1. **THE AWS CLOUD PLATFORM**
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**

Do a deep dive into AWS Cloud

0 WEEKS

COURSE 5 - DATA ENGINEERING - III

1. OPTIMISING SPARK FOR LARGE SCALE DATA PROCESSING

1. **RUNNING SPARK ON MULTINODE CLUSTER**
2. **SPARK MEMORY & DISK OPTIMISATION**
3. **OPTIMISING SPARK CLUSTER ENVIRONMENT**

Use PySpark to create large scale data processing applications.

1 WEEK

2. APACHE FLINK(OPTIONAL)

1. **INTRODUCTION TO APACHE FLINK**
2. **BATCH DATA PROCESSING WITH FLINK**
3. **STREAM PROCESSING WITH APACHE FLINK**
4. **SQL API**

Get Introduced to Apahce Flink and learn query batch data

Use DataStream API to create a stream processing application

0 WEEK

3. REAL-TIME DATA STREAMING WITH APACHE KAFKA

1. **INTRO TO REAL-TIME DATA PROCESSING ARCHITECTURES**
2. **FUNDAMENTALS OF APACHE KAFKA**
3. **SETTING UP KAFKA PRODUCER AND CONSUMER**
4. **KAFKA CONNECT API & KAFKA STREAMS**

Understand the producer-consumer architecture of Apache Kafka. Learn to set up a Kafka cluster for managing real-time data.

1 WEEK

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

4. REAL-TIME DATA PROCESSING USING SPARK STREAMING

1. **SPARK STREAMING ARCHITECTURE**
2. **SPARK STREAMING APIS**
3. **BUILDING STREAM PROCESSING APPLICATION WITH SPARK**
4. **COMPARISION BETWEEN SPARK STREAMING AND FLINK**

Learn about the real-time data processing architecture of Apache Spark. Build Spark Streaming applications to process data in real-time.

1 WEEK

5. ASSIGNMENT (OPTIONAL)

1. **INTRODUCTION, PROBLEM STATEMENT AND GRADING RUBRICS**

Solve an assignment to brush up the skills learnt so far.

0 WEEK

6. BUILDING AUTOMATED DATA PIPELINES WITH AIRFLOW

1. **FUNDAMENTS OF AIRFLOW**
2. **WORKFLOW MANAGEMENT WITH AIRFLOW**
3. **AUTOMATING AN ENTIRE DATA PIPELINE WITH AIRFLOW**

Automate Data Pipelines with Airflow.

1 WEEK

7. ANALYTICS USING PYSPARK

1. **EXPLORATORY DATA ANALYSIS WITH PYSPARK**
2. **PREDICTIVE ANALYSIS WITH SPARK MLLIB**

Use PySpark to do EDA and Predictive Analysis using Spark's ML library.

1 WEEK

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

8. PROJECT: REAL TIME DATA PROCESSING

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

COURSE 5 - CAPSTONE PROJECT

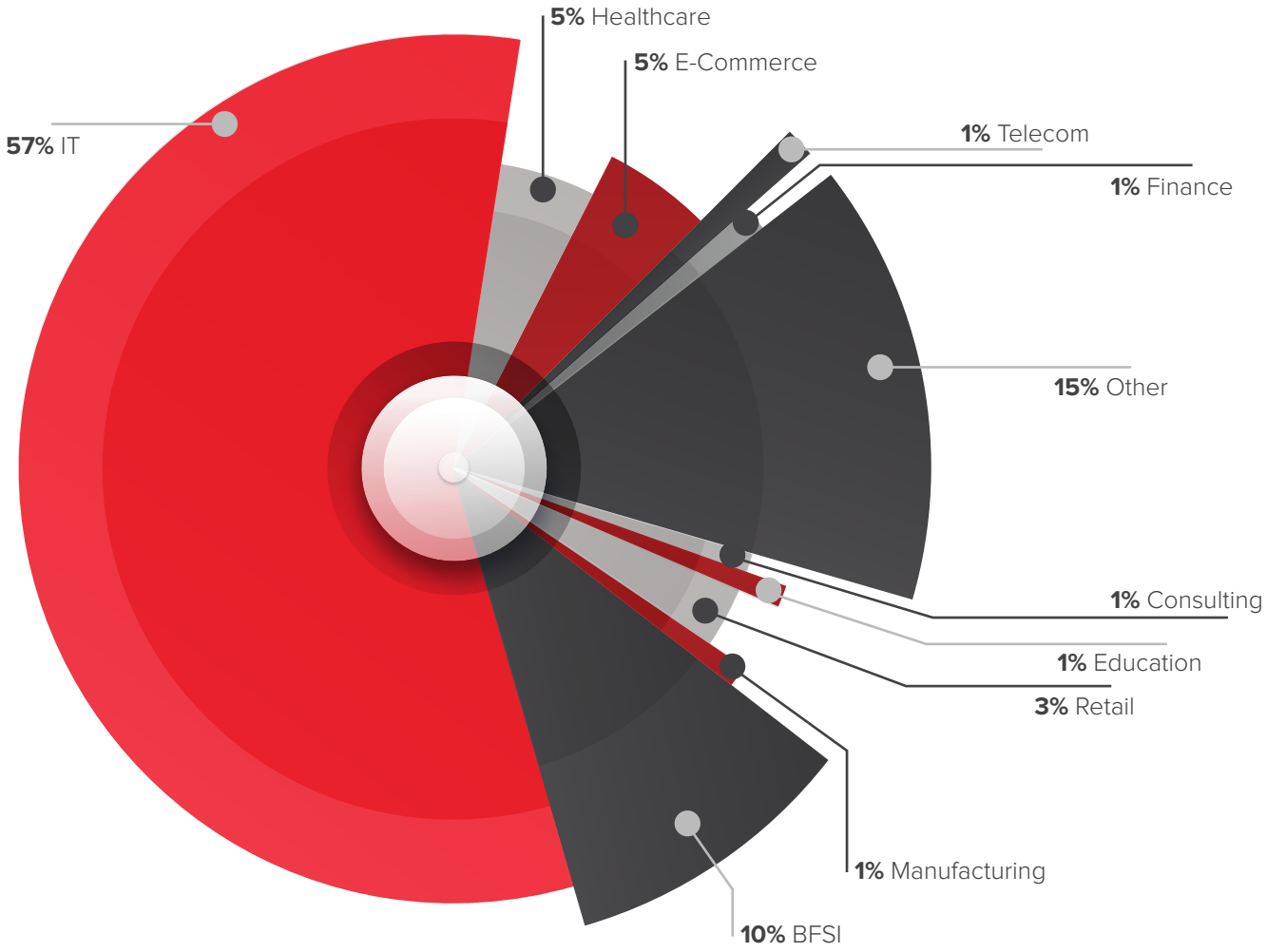
1. CAPSTONE PROJECT

1. AN OVERVIEW OF THE DOMAIN AND ASSOCIATED CONCEPTS 2. PROBLEM STATEMENT 3. EVALUATION RUBRIC 4. MID SUBMISSION 5. FINAL SUBMISSION 6. SOLUTION	The capstone project will stich all the components of data engineering together.	4 WEEKS

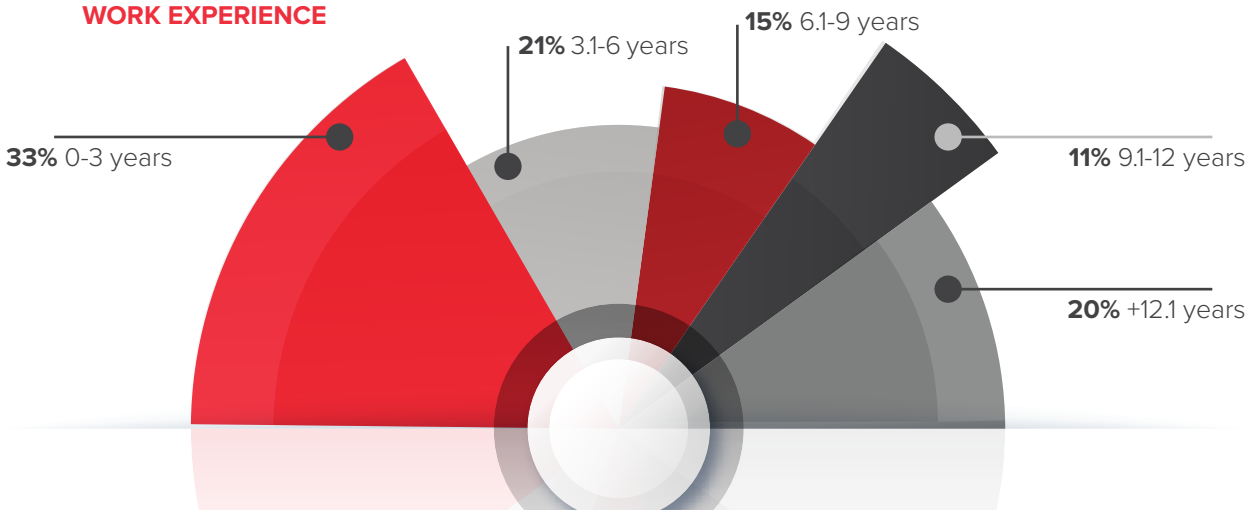
**The Curriculum is subject to change as per the inputs from university or industry experts*

Meet the Class

INDUSTRIES OUR STUDENTS COME FROM



WORK EXPERIENCE



Career Support

Mock Interviews

Get company and role-specific preparation with our mock interview sessions.

Resume Review

Obtain specific, personalised inputs on your resume structure and content.

Personalised Mentorship

Get mentored by an experienced data science global industry experts and receive personalised feedback with 4 one-to-one calls.

Career Mentor

Get a dedicated career mentor to help track your weekly company application targets, coach you on your profile, and support you during your career transition journey.

Post Graduation Career Support

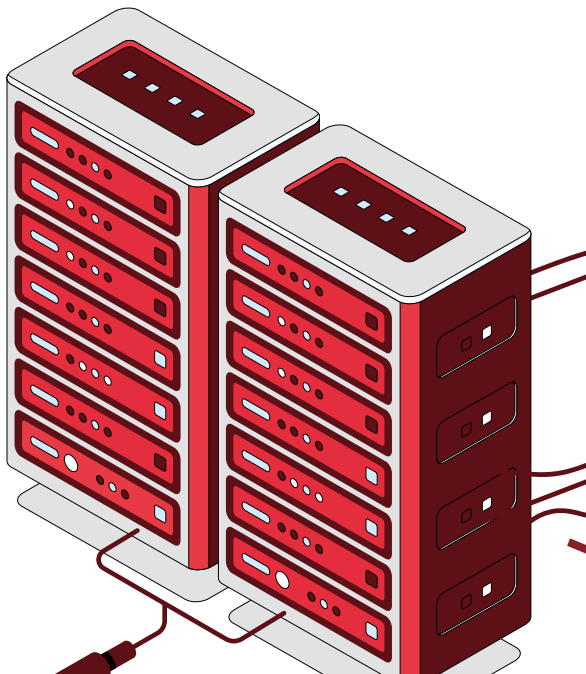
Career sessions are continued post graduation because we want you to have all the help you need. These include: live sessions with industry mentors to guide you, dedicated mentorship, and access to upGrad's career assistance resources.

Company-Specific Preparation

Get company-specific guidance with access to a carefully curated pool of interview resources per company to ensure you are interview-ready for the company of your choice.

Live Profile Building Workshops

Have live sessions of how to build your profile - be it your resume or GitHub - with hands-on sessions on Git and Github you can boost your profile. And also include your work from Kaggle & OpenCV.



Our Alumni Work at

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

Career Transitions

Abhishek Singh
MIS Executive (F&B)
Batch: March 2018

↓

Senior Data Analyst, Myntra
(Oct 2019)

Ayush Modi
Associate Consultant (IT)
Batch: March 2018

↓

Marketing Analyst,
Globalization Partners
(July 2019)

Ashish Y
Asst. Manager
(Manufacturing/Production)
Batch: Sept 2018

↓

Business Analyst, Arvind
(Sept 2019)

Jai Krishna
Fresher
Batch: Sept 2018

↓

Business Analyst, Quantzig
(Nov 2019)

Shadab Hussain
Data Scientist (IT)
Batch: Sept 2018

↓

Analyst, TheMathCompany
(July 2019)

Damodar Bandi
Global Supply Chain Analyst
(Software)
Batch: Dec 2018

↓

Sr. Business Analyst, Data
Semantics
(Aug 2019)

Shubhadip B
Technical Project Manager
(Tech Solutions)
Batch: June 2018

↓

Sr. Technical Project Manager,
Aurionpro
(Oct 2019)

Deepshikha
Senior Associate (Tech)
Batch: March 2019

↓

Data Analyst, Amazon
(Nov 2019)

Rohit W
Quality Assurance Tester
(BFSI)
Batch: Jan 2019

↓

Python Automation Engineer,
Credit Suisse
(July 2019)

Hariharan S
AVP (BFSI)
Batch: March 2018

↓

Project Manager, HSBC
(Aug 2019)

Mohit Mamgain
Internship (IT)
Batch: March 2018

↓

Data Analyst, Shine.com (Aug
2019)

Aakash Dusane
Software Engineer (Software)
Batch: Dec 2018

↓

Data Scientist, Quantzig
(Aug 2019)

Sri Harsha Ravi
Principal Data Structure Engineer (IT)
Batch: Sept 2018

↓

Senior Data Scientist, IHS Markit
(Nov 2019)

Sujit Nalawade
Software Engineer (IT)
Batch: Sept 2018

↓

Data Analyst, Xoriant

Sudha Choudhary
Fresher
Batch: June 2018

↓

Internship - AI Engineer, Athancare
(July 2019)

Nishant Chalasany
Project Manager (Agro)
Batch: March 2019

↓

Analytics Operations Lead, Syngenta
(July 2019)

Ganesh Varanasi
Analyst (BFSI)
Batch: March 2018

↓

Data Scientist, Innominds
(July 2019)

Anshul Srivastava
Business Analyst (BFSI)
Batch: Sept 2018

↓

Associate Consultant, Fractal
(Aug 2019)

Anshul Kumar
Analyst (BFSI)
Batch: March 2019

↓

Data Science Intern, Merkle Sokrati
(Aug 2019)

Sylvester Pinto
Senior Software Engineer (IT)
Batch: Dec 2018

↓

Data Associate, J.P. Morgan
(June 2019)

Ansuman Das
Risk Analyst (IT)
Batch: March 2019

↓

Specialist Data Analyst, Novartis
(Aug 2019)

And many more...

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 coming out of those codes. Just the way I used to get happy after solving a good (tough) math problem during my school age. Thanks to upGrad for providing a great service to people like us who at the age of 43 can dream to study with budding talents around."

Sachin Aggarwal, Experience: 18+ Years
"Learning with IIITB 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 IIITB professors and industry experts on upGrad platforms. Kudos to upGrad."

Savita Upadhaya, Experience: 4 Years
"It has been an amazing journey with upGrad till now. Starting with their course material to live sessions to mentor support, each helps you to always be on track and progress efficiently with the Data Science course. My sincere thanks to the entire team of upGrad and Professors of IIITB for showing me the path and direction for my dream to become a Data Analyst."

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 list of the top twenty most prominent Data Science academics published by Analytics India Magazine. So you need not worry about quality of teaching in this program."

Tuhin Pal, Experience: 5 Years

"I appreciate the platform upGrad has provided and how they have arranged modules and assignments. Modules are locked until you complete the previous one, so it feels like clearing a semester and going to the next one."

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."

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 Executive PG Program 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 | Online

PROGRAM FEE
Please refer to the website for the program fee.

PROGRAM START DATES
Please refer to the website for program start dates.

ELIGIBILITY
Bachelor's degree, no coding experience required
(50% or equivalent passing marks)

WEEKLY COMMITMENT (12-15 hours/week)

6-7 HOURS
Asynchronous learning time.

6-7 HOURS
Assignments and projects.

1 LIVE SESSION/COACHING SESSION
Every week.

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 a great peer group to learn and network with.

STEP 3: Enrollment for Access to Prep Content
Make a quick block payment, receive immediate access to the prep content and begin your upGrad journey.

For any queries, reach us on the following numbers:

+44 7380 344510, info.emea@upgrad.com - Europe, Middle East and Africa
+1 (240) 719-6120, admissions@upgrad.com - North and South America
+65 6232 6730, query@upgrad.com - Asia Pacific except India
+62 21 50864249, query@upgrad.com - Indonesia
+84-28-7300-8884, connect@upgrad.com - Vietnam