

Advanced Certificate Programme in

# DATA SCIENCE

Get the Right Picture



# Table of Contents

- 2** About upGrad and IIITB
- 3** Why upGrad?
- 4** Program Highlights
- 5** New Additions
- 6** Faculty and Industry Experts
- 8** upGrad Learning Experience
- 9** Learning Path
- 10** Program Curriculum
- 24** Industry Projects
- 25** Meet the Class
- 26** Career Support
- 27** Our Alumni Work at
- 28** Career Transitions
- 30** Offline Meet-Ups & Career Assistance
- 31** Hear from Our Learners
- 33** Program Details and Admission Process

# About upGrad and IIITB

*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.

**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 placement network, industry mentorship and the credibility of an Advanced Certificate Programme will provide you with just the right push to accelerate your career in Data Science!

# Why upGrad?

**INR 1.23 Cr Highest Salary**

**300+ Hiring Partners**

**433% Highest Hike**

**700+ Industry Experts**

**50% Avg Salary Hike**

**2+ Million Learners**

Gate 16

Gate 17

Airline SU 123456

Airline SU 123456

# Program Highlights



## Certification from IIITB

Get certified by IIITB on successful completion of the Advanced Certificate Programme in Data Science.



## For The Industry, By The Industry

Learn from 7+ case studies and industry experts to mentor you throughout the program.



## Dedicated Career Assistance

Receive 360 degree career support from mock interviews with hiring managers, resume building, career fairs, industry mentors and much more.



## Personalised Mentorship

Get unparalleled personalised mentorship and doubt resolution from our panel of industry experts.



## Blended Learning

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



## Articulate to a an Executive PG Programme in Data Science

Upon successful completion of the Advanced Certificate Programme in Data Science, get an opportunity to upgrade to the NASSCOM Certified Executive PG Programme in Data Science from IIIT Bangalore wherein, on completing the program, you would receive alumni status from IIIT Bangalore and get a certification from NASSCOM FutureSkills. You can choose to specialise in any of the 5 specialisations of the Executive PG Programme

**Learner is required to achieve 3 out of 4 CGPA to be eligible for the articulation**



# New Additions

## **30-Hour Programming Bootcamp for Non-Tech Learners**

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

## **Career Essential Soft-skills Program**

- Excel at your personal & professional life with upGrad's soft skills program.
- Study three fundamental skills - Interview & Job Search, Corporate & Business Communication and Problem Solving.
- Get access to 40+ learning hours of soft skills content delivered by the best faculty & Industry experts.



# 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

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



**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

**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**

Lead Analyst,  
Zalando

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

**Sajan Kedia**

Ex- Data Science Lead,  
Myntra

Sajan 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.





# upGrad Learning Experience

## Job Opportunities

### upGrad Opportunities

- upGrad Elevate: Virtual hiring drive giving you the opportunity to interview with upGrad's 300+ hiring partners  
Job Opportunities Portal: Gain exclusive access to upGrad's Job Opportunities portal which has 100+ openings from upGrad's hiring partners at any given time
- Be the first to know vacancies to gain an edge in the application process
- Connect with companies that are the best match for you

## Learning Support

### Industry Expert Guidance

- Interactive Live Sessions with leading industry experts covering curriculum + advanced topics
- Personalised Industry Session in small groups (of 10-12) with industry experts to augment program curriculum with customized industry based learning

### Student Support

- Student Support is available 7 days a week, 24\*7
- You can write to us via [studentsupport@upgrad.com](mailto:studentsupport@upgrad.com) or for urgent queries, use the "Talk to Us" option on the Learn platform.

## Practical Learning and Networking

### Networking & Learning Experience

- Live Discussion forum for peer to peer doubt resolution monitored by technical experts
- Peer to peer networking opportunities with a alumni pool of 10000+
- Lab walkthroughs of industry-driven projects
- Weekly real-time doubt clearing sessions

## Career Assistance

### Career Mentorship Sessions (1:1)

- Get mentored by an experienced industry expert and receive personalised feedback to achieve your desired outcome

### High Performance Coaching (1:1)

- Get a dedicated career coach after the program to help track your career goals, coach you on your profile, and support you during your career transition journey

### AI Powered Profile Builder

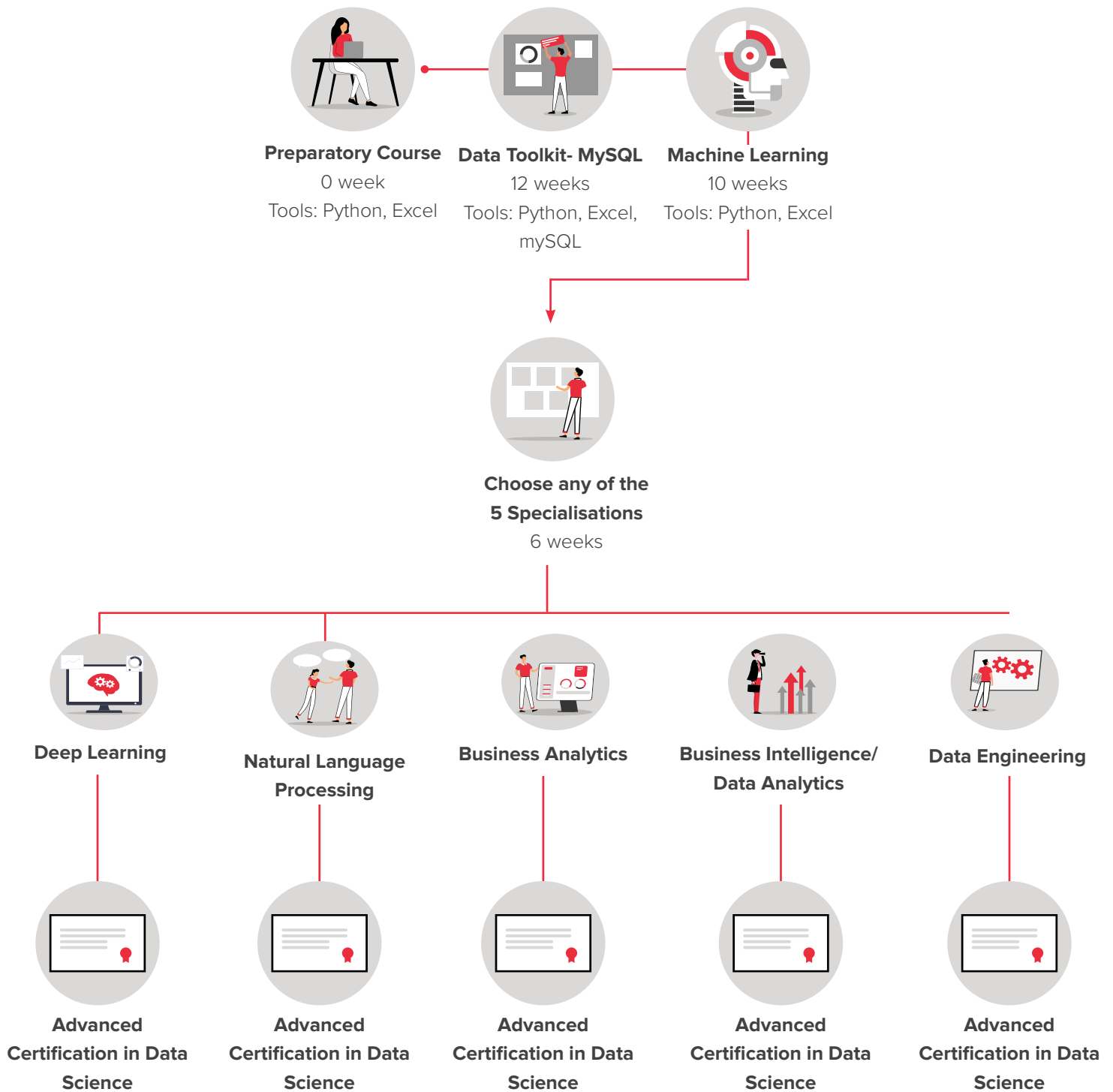
- Obtain specific, AI powered inputs on your resume and LinkedIn structure along with content on real time basis

### Interview Preparation

- Get access to Industry Experts and discuss any queries before your interview  
Career bootcamps to refresh your technical concepts and improve your soft skills



# Learning Path



# Program Curriculum

## PRE-PROGRAMME PREPARATORY CONTENT

### 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.

### 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

### INTRODUCTION TO PYTHON

1. UNDERSTANDING THE UPGRADE 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**

## PROGRAMMING IN PYTHON

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

## 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  |   |        |

## DATA VISUALIZATION IN PYTHON

|                                       |  |        |
|---------------------------------------|--|--------|
| 1. INTRODUCTION TO DATA VISUALIZATION | 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   |  |        |

## EXPLORATORY DATA ANALYSIS

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

### 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**

### INFERENCEAL 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**

### 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 PROPORTION TEST, A/B TESTING**

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

**1 WEEK**

### 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**

## 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. | <b>1 WEEK</b> |
| 2. CASE STATEMENTS, STORED ROUTINES AND CURSORS |   |               |
| 3. QUERY OPTIMISATION AND BEST PRACTICES        |   |               |
| 4. PROBLEM-SOLVING USING SQL                    |   |               |

## SQL ASSIGNMENT: RSVP MOVIES

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

## COURSE 2 - MACHINE LEARNING I

### 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. | <b>2 WEEKS</b> |
| 2. SIMPLE LINEAR REGRESSION IN PYTHON      |   |                |
| 3. MULTIPLE LINEAR REGRESSION              |   |                |
| 4. MULTIPLE LINEAR REGRESSION IN PYTHON    |   |                |
| 5. INDUSTRY RELEVANCE OF LINEAR REGRESSION |   |                |

### 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. | <b>1 WEEK</b> |
| 2. EVALUATION RUBRIC |   |               |
| 3. FINAL SUBMISSION  |   |               |
| 4. SOLUTION          |   |               |

## 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. | <b>2 WEEKS</b> |
| 2. MULTIVARIATE LOGISTIC REGRESSION: MODEL BUILDING AND EVALUATION |  |                |
| 3. LOGISTIC REGRESSION: INDUSTRY APPLICATIONS                      |  |                |

## 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. | <b>1 WEEK</b> |
| 2. ALGORITHMS FOR DECISION TREES CONSTRUCTION |   |               |
| 3. HYPERPARAMETER TUNING IN DECISION TREES    |   |               |

## 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. | <b>1 WEEK</b> |
| 2. K-MEANS CLUSTERING                                      |   |               |
| 3. HIERARCHICAL CLUSTERING                                 |   |               |
| 4. OTHER FORMS OF CLUSTERING: K-MODE, K-PROTOTYPE, DB SCAN |   |               |

## 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. | <b>1 WEEK</b> |
| 2. BASIC LEXICAL PROCESSING      |   |               |
| 3. ADVANCED LEXICAL PROCESSING   |   |               |

## 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. | <b>1 WEEK</b> |
| 2. BUSINESS PROBLEM SOLVING: CASE STUDY DEMONSTRATIONS |  |               |

---

## 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. | <b>1 WEEK</b> |
| 2. EVALUATION RUBRIC |   |               |
| 3. FINAL SUBMISSION  |   |               |
| 4. SOLUTION          |   |               |

---

# SPECIALISATION - DEEP LEARNING

## COURSE 3 - MACHINE LEARNING II

### BAGGING & RANDOM FOREST

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

---

### 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. | <b>1 WEEK</b> |
| 2. GRADIENT BOOSTING                     |  |               |

---



## MODEL SELECTION & GENERAL ML TECHNIQUES

|   |  |               |
|---|--|---------------|
| 1. <b>PRINCIPLES OF MODEL SELECTION</b>   | 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. | <b>1 WEEK</b> |
| 2. <b>MODEL EVALUATION</b>                |  |               |
| 3. <b>MODEL SELECTION: BEST PRACTICES</b> |  |               |

## PRINCIPAL COMPONENT ANALYSIS

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

## ADVANCED REGRESSION

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

## ADVANCED ML CASE STUDY

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

## OPTIONAL

1. **INTRODUCTION TO NEURAL NETWORKS AND ANN**
2. **BACKPROPAGATION IN NEURAL NETWORKS**
3. **HYPERPARAMETER TUNING IN NEURAL NETWORKS**

# SPECIALISATION - NATURAL LANGUAGE PROCESSING

## COURSE 3 - MACHINE LEARNING II

### BAGGING & RANDOM FOREST

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

### 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. | <b>1 WEEK</b> |
| 2. GRADIENT BOOSTING                     |  |               |

### 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. | <b>1 WEEK</b> |
| 2. MODEL EVALUATION                |  |               |
| 3. MODEL SELECTION: BEST PRACTICES |  |               |

### 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 in supervised and unsupervised problems. | <b>1 WEEK</b> |
| 2. PRINCIPAL COMPONENT ANALYSIS IN PYTHON                        |  |               |

## ADVANCED REGRESSION

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

## ADVANCED ML CASE STUDY

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

## OPTIONAL

1. **SYNTACTIC PROCESSING - I**
2. **SYNTACTIC PROCESSING - II**

# SPECIALISATION - BUSINESS ANALYTICS

## COURSE 3 - ADVANCED MACHINE LEARNING

### BAGGING & RANDOM FOREST

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

### MODEL SELECTION & GENERAL ML TECHNIQUES

- |   |  |                |
|---|--|----------------|
| 1. <b>PRINCIPLES OF MODEL SELECTION</b> | 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. | <b>2 WEEKS</b> |
| 2. <b>MODEL BUILDING AND EVALUATION</b> |  |                |
| 3. <b>FEATURE ENGINEERING</b>           |  |                |
| 4. <b>CLASS IMBALANCE</b>               |  |                |

## 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

## MODEL SELCTION CASE STUDY

### 1. PROBLEM STATEMENT

Apply your business acumen to the newly learnt machine learning techniques, and select the most appropriate model for a provided business scenario.

**1 WEEK**

### 2. EVALUATION RUBRIC

### 3. FINAL SUBMISSION

### 4. SOLUTION

# SPECIALISATION - BUSINESS INTELLIGENCE/ DATA ANALYTICS

## COURSE 3: ADVANCED DBS AND BIG DATA ANALYTICS

### DATA MODELLING

### 1. DATABASE DESIGN RECAP

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

**1 WEEK**

### 2. BUILDING BLOCKS OF DATA MODELLING

### 3. PROBLEM SOLVING USING DATA MODELLING

### 4. DATA MODELLING: OPTIONAL ASSIGNMENT

## ADVANCED SQL AND BEST PRACTICES

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

## 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. | <b>1 WEEK</b> |
| 2. AMAZON WEB SERVICES                      |   |               |
| 3. BIG DATA STORAGE AND PROCESSING - HADOOP |   |               |
| 4. EMR CLUSTER IN AWS                       |   |               |

## ANALYTICS USING SPARK

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

## BIG DATA CASE STUDY

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

# SPECIALISATION - DATA ENGINEERING

## COURSE 3 - DATA ENGINEERING - I

### 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  |  |        |

### 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 |  |        |

### 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             |  |        |

### 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            |   |        |

**ASSIGNMENT (OPTIONAL)**

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

**NOSQL DATABASES AND APACHE HBASE****NOSQL DATABASES AND MONGODB (OPTIONAL)**

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

**DATA WAREHOUSING (OPTIONAL)**

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

**DATA INGESTION WITH APACHE SQOOP AND APACHE FLUME**

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

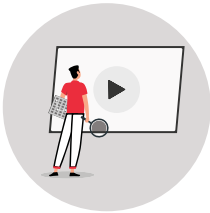
**MAPREDUCE PROGRAMMING ASSIGNMENT****1. PROBLEM STATEMENT AND  
SAMPLE DATASET**

Practise MapReduce Programming on a Big Dataset.

**1 WEEK****2. SOLUTION**



# Industry Projects



**IMDb Movie Analysis**



**Uber Supply-Demand**



**Lead Scoring**



**Fraud Detection**



**Creditworthiness of  
Customers**



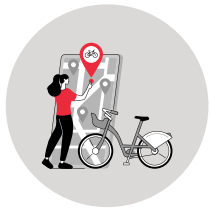
**Speech Recognition**



**Credit EDA Assignment**



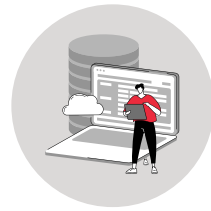
**RSVP CaseStudy**



**Bike Sharing Assignment**



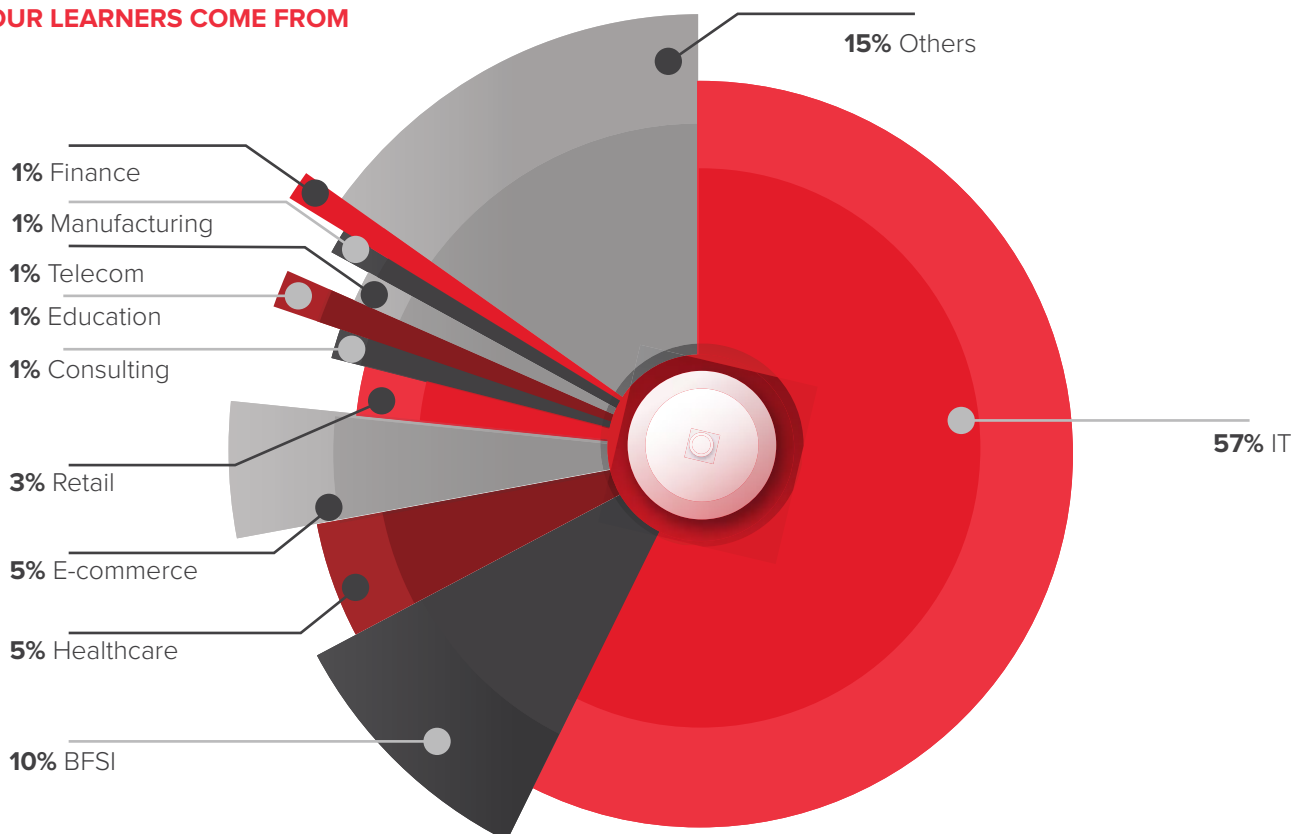
**Lead Scoring Case Study**



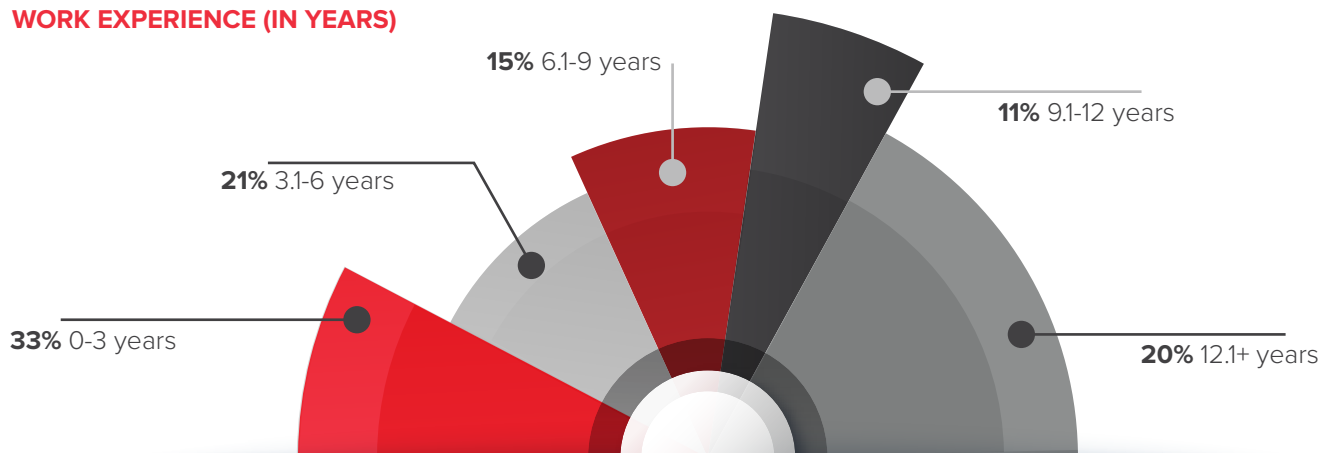
**Big Data Case Study**

# Meet the Class

## INDUSTRIES OUR LEARNERS COME FROM



## WORK EXPERIENCE (IN YEARS)



**A few of the companies our students are from:** Accenture, Amazon, Cognizant, Deloitte, Infosys, Microsoft, Wipro, EY, CitiBank, Cisco, Thomson Reuters

# 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 guestimates
- 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

# Our Alumni Work at

## PROGRAM SUMMARY

# 2+ Million

Enrolled Learners

# INR 1.23<sub>CR</sub>

Highest Salary Hike

# 433%

Highest Salary Hike

# 50%

Average Salary Hike

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



## 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...**

# Offline **Meet-ups & Career Assistance**



## **UPGRAD BASECAMPS**

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 have never done coding in my life. I did it during this program and was thrilled to see the outcomes of those codes. Just the way I used to get happy after solving good (tough) math problems during my school age. Thanks to upGrad for doing a great service to people like us who at the age of 43 can dream to studying 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 program material to live sessions to mentor support helps one to always be on track and progress efficiently with Data Science program. My sincere thanks to the entire team of upGrad and Profs of IIITB to show me the path and direction towards 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 top twenty prominent data science academicians list published by Analytics India Magazine. So you need not worry about the quality of teaching in this program."*



**Tuhin Pal, Experience: 5 Years**

*"I appreciate the platform upGrad how they have arranged the modules and the assignments are quality ones. You will relish your college days again as the exams felt like semester ones, where you can't talk to anybody. Modules are locked till you complete the previous one so until you clear a semester, you can't move to the next one."*

**Harkirat Dhillon, Experience: 8 Years**

*"A dedicated regimen for studying the program and keep learning is the key to being successful and passing the program. This program will help build a strong foundation for a successful transition to data science. Additionally, participating in hackathons and Kaggle competitions to solve real-world problems will give you an edge and land a job if one is willing to work hard."*

**Shravani Shahapure, Experience: 16 Years**

*"For someone who passionately wants to pursue a career in the field of data science, it is worth to opt for the complete program by IIIT B and upGrad. Data science is an experimental science. We need to develop the right kind of thinking ability to approach problems. And to develop this ability, we need expert's direction. Knowing tools won't solve the problems always, we need to use them wisely and correctly. IIITB and upGrad's online program on data science gives this opportunity and develop 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 powered by upGrad most helpful. They even teach you most of the prerequisites with prep sessions before you even start the program. Being a working professional, it was neither too difficult nor too easy to keep up with the pace of the program."*



# Program Details and Admission Process

## PROGRAM DURATION AND FORMAT

8-8.5 Months | Online

## PROGRAM FEE

INR 99,000 (Inclusive of taxes)

## PROGRAM START DATES

Please refer to the [website](#) for program start dates.

## ELIGIBILITY

Bachelor's Degree with a minimum 50% or equivalent passing marks.

## SELECTION PROCESS



### Step: 1

#### Fill the Application form

Fill out an application giving your basic profile details



### Step: 2

#### Shortlisting & Block Payment to reserve seat

University will shortlist basis profile. Receive offer letter and proceed with block amount INR 15,000/-



### Step: 3

#### Payment

Balance to be closed within 7 days or batch start date whichever is sooner

## 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.