MASTERS DEGREE IN
ARTIFICIAL INTELLIGENCE

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, 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 its inception, upGrad has delivered over 20 million hours of learning, delivering programs by collaborating with universities across the world including LJMU, IIIT Bangalore and Deakin Business School among others. And it doesn't end there.

upGrad, in collaboration with IIIT Bangalore, a renowned university and Katz School of Science and Health at Yeshiva University, USA is excited to offer a one-of-its-kind, academically rigorous and industry relevant Masters Degree in Artificial Intelligence.

The faculty includes an average of 15+ years of experience. The faculty covers the conceptual depths of topics such as Data Science, Machine Learning and AI, 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.

Our aim is simple: We strive to create high-impact, on-campus hands-on experiences that prepare students for meaningful and productive careers.

Ronnie Screwvala
Co-founder & Executive Chairman

upGrad

ABOUT
YESHIVA UNIVERSITY

Founded 1886

More than 70,000 Alumni

Accredited by Middle States Commission on Higher Education (MSCHE)

QS World University Rankings 2023, **246**

R2 Doctoral University with High Research Activity

US News & World Report #68

Situated in one of the most vibrant cities, New York, Yeshiva University's graduate and affiliate schools offer tremendous opportunities for students to grow. Founded in 1886, the university provides degrees in various disciplines, including science, legal studies, economics, social work, psychology, medicine, etc. With 11 graduate and undergraduate schools, the university has a strong network of over 70,000 alumni. Yeshiva University caters to students worldwide and prepares them to be future leaders under highly experienced and qualified faculty.

About us
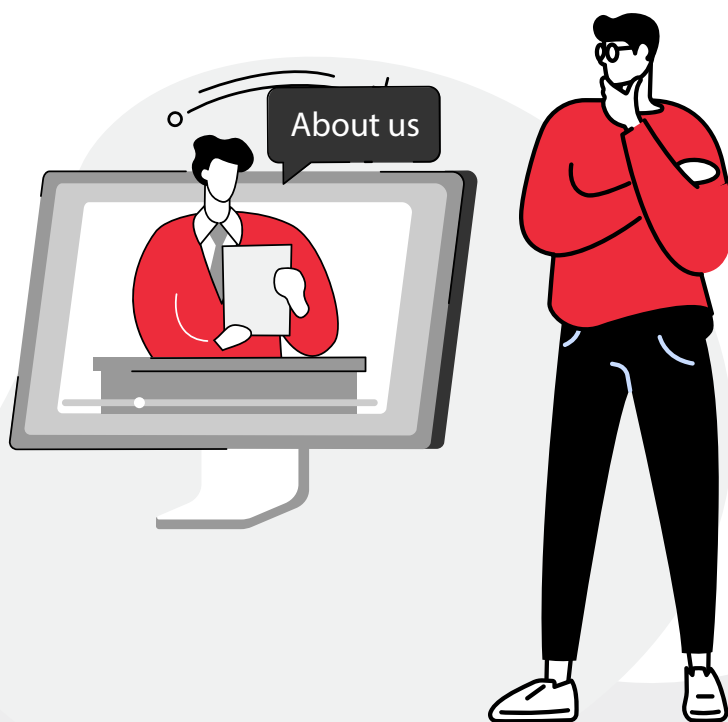
ABOUT
KATZ SCHOOL OF SCIENCE AND HEALTH

Katz School of Science and Health is one of the 11 schools of Yeshiva University. The Katz School offers programs that prepare students for advanced education in various disciplines. The industry-leading faculty at the Katz School helps learners polish their skills and build a professional network. The skills and strategies you will harness at the Katz School of Science and Health will allow you to push boundaries in many in-demand and high-wage industries and shape your path to success. The Katz School is where bold, purpose-driven people come to connect, create and explore. Katz faculty, students, and alumni are advancing scholarly knowledge and transforming lives through pioneering research, citywide initiatives, new technologies, and innovative clinical diagnostics.

ACHIEVEMENTS AT A GLANCE:

- The university is ranked 246 as per the QS World University Rankings 2023.

- Ranked 68th in the National Universities by US News & World Report.

- The university is accredited by the Middle States Commission on Higher Education (MSCHE).

- Over 70,000 strong alumni network.

- Offers 65+ majors and minors in various disciplines.

- Has 11 graduate and undergraduate schools and 3 affiliate schools to impart quality education.

- More than 2,000 staff across the four university campuses

- A variety of scholarships are available for local and international students.

ELIGIBILITY CRITERIA
FOR TRANSFER TO YESHIVA UNIVERSITY

- An aggregate score of 3.0 CGPA and above in the Advanced Certificate Program in Data Science with a minimum of 3.0 GPA in each course.

- IELTS of >= 6.5 or TOEFL of >=90 or Duolingo acceptable score is 110.

PROGRAM
HIGHLIGHTS

Dual Accreditation and Alumni Status
Get certified by IIITB and Katz School of Science and Health at Yeshiva University, USA.

High Employment Potential
Starting salaries of Data Scientists in USA is $100,000 as per indeed.com

Customisable Curriculum
Choose from 8 electives at Katz School of Science and Health at Yeshiva University, USA on the basis of your background and career aspirations and get the learning you want.

Complete your course in 1 year
Only 9 subjects to be done which makes it easier to complete course in 1 year.

Optional Practical Training (OPT)
STEM-approved degree, international students may be eligible for up to 36 months of OPT (post-study work visa).

FACULTY AND INDUSTRY
EXPERTS AT IIITB

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 the CitiGroup, Tata Industries etc.

Chandrashekar Ramanathan

Dean | Academics
International Institute of Information Technology Bangalore

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

S. Anand

CEO
Gramener

A Gold medalist from IIM Bangalore, an alumnus of IIT Madras and London School of Business, 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 Tech as well as an integrated M.Tech. (EE) from IIT Bombay. Her research interests are computers and wireless technology.

Dr. Debabrata Das

Director, IIITB

Dr. Debabrata Das is currently Director of IIITB. He has received his PhD degree from IIT-Kharagpur. His main areas of research interest are IoT and Wireless Access Network's MAC, QoS, Power saving.

Prof. G. Srinivasaraghavan

Professor
International Institute of Information
Technology Bangalore

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

Ankit Jain

Sr. Research Scientist
Uber Ai Labs

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

Dinesh Babu Jayagopi

Assistant Professor
International Institute of Information
Technology Bangalore

Prof. Dinesh has a PhD from Ecole
Polytechnic Federate Switzerland,
M.Sc. from IISc Bangalore in System
Science and Signal Processing, and
B.Tech.

Kalpana Subbaramappa

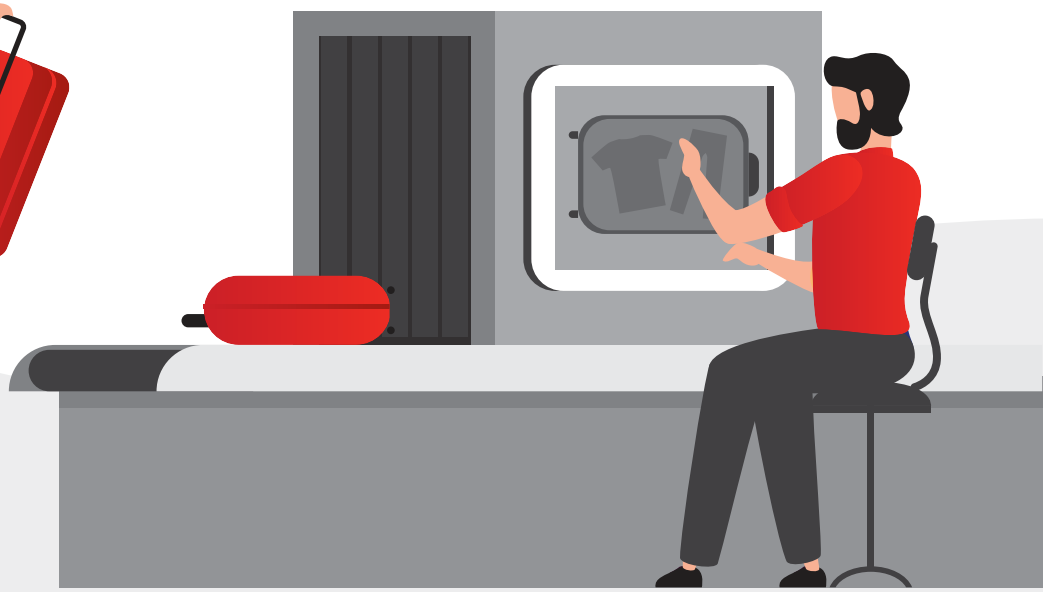ex-AVP | Decision Science
Genpact

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

Mirza Rahim Baig

Lead | Analyst
Zalando

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

FACULTIES AT
YESHIVA UNIVERSITY

Dr. Pablo Roldan

Dr. Pablo Roldan, is director of the artificial intelligence master's degree program at the Katz School. He is an accomplished researcher in mathematics, statistics and computer science. He has received research grants from the National Science Foundation; the Spanish Ministry of Science; and the European Regional Development Fund. He has consulted on machine learning for companies that work in medical devices, molecular design and treasury management. He holds a Ph.D. in applied mathematics and an M.S. in physics engineering from the KTH Royal Institute of Technology and a B.S. in computer science from the Universitat Politecnica de Catalunya.

Michael Schulte

Michael Schulte is a health data scientist. His research focuses on incorporating social determinants of health into healthcare quality and improving how quality is measured and evaluated. As a consultant at CareFirst Blue Cross/Blue Shield he is developing innovative ways for measuring and improving health care quality using predictive models and data-driven intervention and anticipation programs. Previously he was senior health data analyst for Medicare STARS at Johns Hopkins HealthCare, where he developed Tableau dashboards and executive-level reports. He holds an M.A. in applied economics, and a B.A. in mathematics and philosophy from Western Michigan University.

Dr. Sai Kadiyala

Dr. Kadiyala is an Artificial Intelligence postdoctoral research fellow at the Katz School. Prior to joining Yeshiva University, he was research scientist at System Security Group, Institute for Infocomm Research, A-STAR Singapore, where his work involved developing anomaly detection algorithms for program behavior analysis and characterization of malware. He has also held postdoctoral fellowship positions in the fields of electrical engineering and cybersecurity at Nanyang Technological University in Singapore. He holds a Ph.D. in computer science engineering from the Indian Institute of Technology, Kharagpur.

James Topor

James Topor is managing director of Mirus Global Advisory Services and is a management consultant and data scientist with expertise in data analytics, predictive modeling, machine learning, machine vision, and systems integration and development. His major clients have included businesses in the financial services, pharmaceutical, telecom and manufacturing industries throughout North America, Europe, Asia and Australia. He holds an M.S. in data analytics from the City University of New York, an MBA in international business from George Washington University, an M.S. in computer science from Virginia Tech and a B.S. in computer science from Eastern Connecticut State University.

UPGRAD
LEARNING EXPERIENCE

EXPERT FEEDBACK AND LEARNING

- TA sessions
- Regular live sessions by experts to clarify concept related doubt
- Receive unparalleled guidance from industry mentors, teaching assistants and graders

STUDENT SUPPORT TEAM

- Student support available 7 days a week, 24*7
- You can write to us via studentsupport@upgrad.com or for urgent queries, use the "Talk to us" option on the learn platform

Q/A FORUM

Peer-to-peer discussion forum where you can post your queries and your peers/faculty/teaching assistants answer your queries within a day. Regular Q&A sessions with faculty to get clarification on conceptual doubts.

NEW
ADDITIONS

Career Essential Soft-skills Program

- Excel in 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+ learner hours of soft skills content delivered by the best faculty & Industry experts

INDUSTRY
PROJECTS
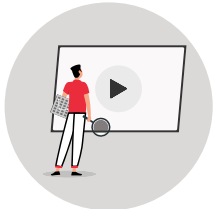
IMDb Movie Analysis

Uber Supply-Demand Gap

Lead Scoring

Fraud Detection

Creditworthiness of Customers

Speech Recognition
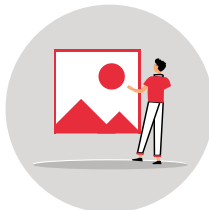
Image Captioning

Gesture Recognition
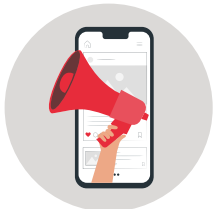
Social Media Listening

Telecom Churn
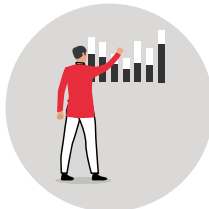
Interactive Market Campaign Analysis

Retail Giant Sales Forecasting

And many more!

LEARNING
PATH

Preparatory Course
Tools: Python,
Excel

**IELTS Preparation
& Exam**

Data Toolkit
13 weeks
Tools: Python,
Excel, mySQL

Machine Learning
10 weeks
Tools: Python, Excel
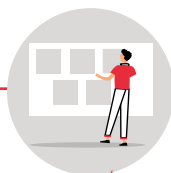
Course 3
6 weeks
Tools: Python, Excel
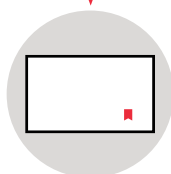
Advanced Certification in Data Science

Journey in Katz **(Year 2)**
Katz School
of Science and Health
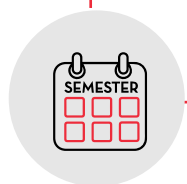
Sem 1
Numerical Methods, Data
Management, Elective
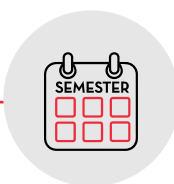
Sem 2
Neural Networks, Artificial
Intelligence, Capstone, Elective

Masters Degree in Artificial Intelligence
at Katz School of Science
and Health at Yeshiva University

MASTERS DEGREE IN ARTIFICIAL INTELLIGENCE

1. Advanced Certificate in Data Science from IIIT, Bangalore (8 months)

PROGRAM CURRICULUM

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   2 WEEKS

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.

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

Programming in Python

<div style="float:right">**2 WEEKS**</div>

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.

Python for Data Science

1 WEEK

1. Introduction to NumPy
2. Introduction to Matplotlib
3. Introduction to Pandas
4. Getting and Cleaning Data

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.

Data Visualization in Python

1 WEEK

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.

Exploratory Data Analysis

1 WEEK

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.

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

Credit Eda Case Study   1 WEEK

1. Problem Statement
2. Evaluation Rubric
3. Final Submission
4. Solution

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

Inferential Statistics   1 WEEK

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.

Hypothesis Testing   1 WEEK

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.

Data Analysis Using Sql   1 WEEK

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.

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

Advaced Sql & Best Practices   1 WEEK

1. Window Functions
2. Case Statements, Stored Routines and Cursors
3. Query Optimisation And Best Practices
4. Problem-Solving Using

Apply advanced SQL concepts like window-ing and procedures to derive insights from data and answer pertinent business ques-tions.

Sql Assignment: Rsvp Movies   1 WEEK

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.

Course 2 - Machine Learning I

Linear Regression   2 WEEKS

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 vari-able can be predict-ed using several other variables through a housing dataset where you will predict the prices of houses based on various factors.

Linear Regression Assignment   1 WEEK

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.

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

Logistic Regression

<div align="right">

2 WEEKS

</div>

1. Univariate Logistic Regression
2. Multivariate Logistic Regression: Model Building and Evaluation
3. Logistic Regression: Industry Applications

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

Classification Using Decision Trees

<div align="right">

1 WEEK

</div>

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.

Unsupervised Learning: Clustering

<div align="right">

1 WEEK

</div>

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.

Basics Of NLP and Text Mining

<div align="right">

1 WEEK

</div>

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 pho-net-ic hashing, bag-of-words, TF-IDF, etc.

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

Business Problem Solving

<div style="text-align: right">**1 WEEK**</div>

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.

Case Study: Lead Scoring

<div style="text-align: right">**1 WEEK**</div>

1. Problem Statement
2. Evaluation Rubric
3. Final Submission
4. Solution

Help the Sales team of your company iden-tify which leads are worth pursu-ing through this classification case study.

Specialisation - Deep Learning

Course 3 - Machine Learning II

Bagging & Random Forest

<div style="text-align: right">**1 WEEK**</div>

1. Popular Ensembles
2. Introduction to Random Forests
3. Feature Importance in Random Forests
4. Random Forests in Python

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

Boosting

<div style="text-align: right">**1 WEEK**</div>

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.

Model Selection & General ML Techniques

<div style="text-align: right">**1 WEEK**</div>

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 com-plexity, alongwith general machine learning techniques like feature engi-neering, model evaluation, and many more.

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

Principal Component Analysis   **1 WEEK**

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.

Advanced Regression   **1 WEEK**

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.

Advanced ML Case Study   **1 WEEK**

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.

Specialisation - Natural Language Processing

Course 3 - Machine Learning II

Bagging & Random Forest   **1 WEEK**

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.

Boosting   **1 WEEK**

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.

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

Model Selection & General ML Techniques

1 WEEK

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.

Principal Component Analysis

1 WEEK

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.

Advanced Regression

1 WEEK

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.

Advanced ML Case Study

1 WEEK

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.

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

Specialisation - Business Analytics

Course 3 - Advanced Machine Learning

Bagging & Random Forest   1 WEEK

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.

Model Selection & General ML Techniques   2 WEEKS

1. Principles of Model Selection
2. Model Building and Evaluation
3. Feature Engineering
4. Class Imbalance

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.

Time Series Forecasting   2 WEEKS

1. Introduction to Time Series and its Components
2. Smoothing Techniques
3. Introduction to AR Models
4. Building AR Models

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

Model Selction Case Study   1 WEEK

1. Problem Statement
2. Evaluation Rubric
3. Final Submission
4. Solution

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

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

Specialisation - Business Intelligence/Data Analytics

Course 3 - Advanced DBS and Big Data Analytics

Data Modelling   1 WEEK

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.

Advanced SQL and Best Practices   1 WEEK

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

---

Introduction to Big Data and Cloud   1 WEEK

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.

Analytics using Spark   2 WEEKS

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.

Big Data Case Study   1 WEEK

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.

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

Specialisation - Data Engineering

Course 3 - Data Engineering - I

Data Management and Relational Database Modelling

1 WEEK

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.

Introduction to Big Data(Optional)

0 WEEK

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.

Introduction to Cloud and AWS Setup

1 WEEK

1. Introduction to Cloud
2. AWS Setup

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

Introduction to Hadoop and MapReduce Programming

1 WEEK

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.

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

Assignment (Optional)

<div align="right">**0 WEEK**</div>

1. Introduction, Problem Statement and Grading Rubrics

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

NoSQL Databases and Apache HBase
NoSQL Databases and MongoDB (Optional)

<div align="right">**1 WEEK**</div>

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.

Data Warehousing (Optional)

<div align="right">**0 WEEK**</div>

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.

Data Ingestion with Apache Sqoop and Apache Flume

<div align="right">**1 WEEK**</div>

1. Introduction to Data Ingestion
2. Structured Data Ingestion with Sqoop
3. Unstructured Data Ingestion with Flume

Get familiar with the challenges involed in data ingestion. Use Sqoop and Flume to ingest structured and unstructured data into Hadoop.

Map reduce Programming Assignment

<div align="right">**1 WEEK**</div>

1. Problem Statement and Sample Dataset
2. Solution

Practise MapReduce Programming on a Big Dataset.

STUDY ABROAD CURRICULUM

IELTS PREPARATION

- Preparation of IELTS Exam

ACADEMIC ENGLISH

- Language Development
- Academic Writing Skills
- Spoken Academic Communication
- Reading and Listening in Academic Context

RESEARCH METHODOLOGY

- Introduction to Research
- Research Sampling

ACADEMIC INTEGRITY & RESEARCH

- Introduction to Academic Integrity
- Introduction to Plagiarism
- Plagiarism : Advanced Understanding
- Ethical Considerations

ACADEMIC REFERENCING

- Academic Referencing
- Referencing Style Guides

FORMS OF ACADEMIC WRITING

- Different forms of Long Format Writing
- Dissertation or Thesis Writing
- Writing SOPs

CRITICAL THINKING

- Being a Critical Thinker
- Reasoning Skills
- Persuasion
- Critical Thinking as a Life Skill

NETWORKING AND PROFILE BUILDING IN FOREIGN DESTINATION

- Networking
- Profile Building

FINDING JOBS AND INTERNSHIPS

- How to Find Jobs And Internships in Foreign Destinations

IMMIGRATION AND VISA RULES

FINANCING EDUCATION

2. MASTERS DEGREE IN ARTIFICIAL INTELLIGENCE FROM YESHIVA UNIVERSITY, USA COURSE DESCRIPTION (YEAR 2)

CORE COURSES

1. DATA ACQUISITION AND MANAGEMENT — 3 CREDITS

Data Acquisition and Management focuses on the data structures, data design patterns, algorithms, methods, and best practices for the pre-modeling phases of data science workflows, including problem formulation, gather, analyze, explore, model, and communicate, analytics programming focuses on the gather, analyze, and explore workflow steps. This comprises the "data wrangling" work which is where most data scientists spend the majority of their time. Because data science is iterative, this preparatory work informs the modeling phase. Often, the creation and validation of new models requires going back for additional data, different data transformations, and exploration of data distributions. In short, every effective data scientist needs to master analytics programming. Course topics include reading from or writing to databases, text files, and the web; shaping data into "tidy" data frames, exploratory data analysis, data imputations, feature engineering, and feature scaling.

2. NUMERICAL METHODS — 3 CREDITS

Algorithms in machine learning and neural networks are built upon a strong foundation of linear algebra. For example, modern recommendation systems may have sparse matrices with millions of users and millions of items; matrix factorization methods make the underlying calculations tractable say this course builds a foundation of linear algebra concepts such as matrices, determinants, vectors and eigen values. Then it deepens it into data science applications around network analysis and logistic algorithms. In addition, some multi-variate calculus and graph theory topics are covered.

3. ARTIFICIAL INTELLIGENCE — 3 CREDITS

Prerequisites: Data Acquisition and Management; Computational Statistics and Probability Artificial Intelligence (AI) is an interdisciplinary field, integrating knowledge and methods from computer science, mathematics, philosophy, psychology, economics, neuroscience, linguistics, and biology. Intelligent agents mimic cognitive functions to implement intelligent behaviors such as perception, reasoning, communication, and acting in symbolic and computational models. AI is used in a wide range of narrow applications, from medical diagnosis to speech recognition to bot control.

The autonomous single, multiple, and adversarial agents that students build in this course will support fully observable and partially observable decisions in both deterministic and stochastic environments.

Topics covered include search, constraint satisfaction, Markov decision processes, planning, knowledge representation, reasoning under uncertainty, graphical models, and reinforcement learning. The techniques and technologies mastered here will provide the foundational knowledge for the ongoing study and application of AI in other applications across practice areas.

4. NEURAL NETWORKS AND DEEP LEARNING   3 CREDITS

Prerequisites: Machine Learning
Data scientists have been able to leverage better algorithms on faster hardware optimized with graphical processing units to deliver improved performance and accuracy in whole classes of applications that had been previously commercially unviable. The biggest beneficiaries are applications that require unstructured data, such as audio and or video processing. Deep neural networks have also provided gains for other complex applications, from recommendation systems to natural language processing. This course builds on the concepts in machine learning to train multi-layered neural networks.

5. AI CAPSTONE: R&D EXPERIENCE   3 CREDITS

The Capstone integrates prior coursework, research, colloquia, and professional experience, and provides the opportunity to synthesize theory with practice in an applied project, thesis, or equivalent activity. Examples include developing an AI application or methodology, publishing a research paper at a peer-reviewed conference, or creating a startup company through YU's Innovation lab—though students may propose other related projects based on their interests.

The Capstone will include four components: a brief proposal and project schedule; the main deliverable (e.g. thesis, conference paper, working system with analysis/code/data); and a final presentation to the student and faculty body. Faculty will provide students with mentorship and feedback at each stage of the work.

6. COMPUTATIONAL STATISTICS AND PROBABILITY*   3 CREDITS

Arguably, most of data science is statistical learning, which requires strong foundational knowledge in probability and statistics. And applying computational methods such as direct simulation, shuffling, bootstrapping, and cross-validation to statistical problems is often more intuitive, and intuitive and can provide solutions where analytical methods would prove computationally intractable. This course introduces students to the statistical analysis of data using modern computational methods and software. Probability, descriptive statistics, inferential statistics and computation methods such as simulations, sample distributions, shuffling, bootstrapping, and cross-validation will be covered.

7. PREDICTIVE MODELS*   3 CREDITS

Computational Statistics and Probability Predictive modeling answers the question, "What will happen next?" Linear regression and logistic regression are foundational predictive modeling methods, used to predict continuous and categorical output respectively. The main topics covered in this course include simple and multiple linear regression, variable selection and shrinkage methods, binary logistic regression, count regression, weighted least squares, robust regression, generalized least squares, multinomial logistic regression, generalized linear models, panel regression, and nonparametric regression.

8. MACHINE LEARNING**   3 CREDITS

Computational Statistics and Probability In classical programming, answers are obtained from rules and data. In machine learning, rules are obtained from data and answers. The widespread availability and sharing of data, and improvements in computing capacity, processing methods, and algorithms have given machine learning the power to deliver game-changing systems and technologies to organizations that compete on predictive, prescriptive, and/or autonomous analytics. In this course, we'll look at methods for using, tuning, and comparing machine learning algorithms, based on measures of performance, accuracy, and explainability. We'll also look at recent advances and trends in automated machine learning.

* Computational Statistics and Probability is waived on the basis of IIIT B Data Toolkit courses
* Predictive Models are waived on the basis of IIIT B Machine Learning course.
** Machine learning is waived If a learner has done Deep Learning Track/NLP/Business Analytics track at IIITB ACP program.

ELECTIVES (CHOOSE ELECTIVES AGGREGATING TO TOTAL 12 CREDITS)

1. BAYESIAN METHODS   3 CREDITS

Prerequisites: Data Acquisition and Management; Computational Statistics and Probability
Bayesian inference provides powerful tools to model random variables. While Bayesian methods often yield the most accurate theoretical results, historically analytical complexity made it challenging to apply Bayesian methods against less trivial problems. Now, the confluence of more powerful computing resources and improved computational algorithms make Bayesian methods the best choice for tackling some of the most complex data science problems.

Bayesian analysis is increasingly important in academic research, and research and is the preferred standard statistical analysis tool in data science practice. In this course, we'll build from Bayes' probability foundations to first applying Bayesian methods to infer binomial probabilities, then hierarchical models, and finally generalized linear models. We'll provide comparisons between frequentist approaches and Bayesian approaches. We'll build basic algorithms from scratch, as well as using high-performance Markov Chain Monte Carlo (MCMC) methods.

2. AI PRODUCT STUDIO   3 CREDITS

What is needed to convert a promising idea or research into a viable product or service? Bringing successful products to market is an experiential discipline that requires hands-on practice working through iterative workflow of a customer-driven product development lifecycle. In this course, students work with mentors to design products, develop customers, and create product development roadmaps. Students create and communicate hypotheses around customers, cost and revenue streams, activities, and value propositions. Agile project management, data-driven product design and customer feedback, and technical constraint identification are all covered.

3. NATURAL LANGUAGE PROCESSING   3 CREDITS

Prerequisites: Machine Learning; Neural Networks and Deep Learning
Natural Language Processing lives at the intersection of machine learning, artificial intelligence, and linguistics. It is the key to unlocking vast amounts of human-generated, unstructured data. The increased availability of corpuses of text data, the wide availability of cheap distributed systems, improvements in neural network algorithms, and increased access to graphical processing units (GPUs) have improved the performance and accuracy of entire families of once computationally intractable problems, making these commercially feasible.

This course explores a series of text and voice processing use cases, including sentiment analysis and topic modeling. It is the key to unlocking vast amounts of human-generated, unstructured data. Along the way, students gain experience working with supervised and unsupervised methods using both machine learning algorithms and deep neural networks.

4. DATA VISUALIZATION   3 CREDITS

Data scientists depend on data visualizations for their own exploratory analysis to support their modeling decisions--the mind can process visual information must faster than numbers. Data visualization is also important to inform—and often to persuade—other people about what can be inferred from the data.
These explanatory visualizations often require higher production values, interactivity, and guiding text. In this course, students apply the concepts, methods, and best practices of data visualization to create reproducible, code-based exploratory and explanatory data visualizations.

5. ADVANCED DATA ENGINEERING   3 CREDITS

As both the volume and the velocity of data increase exponentially, problems in both commerce and research become increasingly reliant on environments with distributed data storage and data processing capabilities. This requires rethinking how our entire approach to distributed envi-ron-ments. This course provides students with the concepts, data structures, and algorithms needed to implement data science applications in distributed computing environments. In this course, we will implement and apply distributed algorithms, data frames, and streaming. You will also learn how to choose appropriate distributed algorithms based on the characteristics of the problem and the system.

6. COMPLEX SYSTEMS: FINANCIAL TIME SERIES ANALYSIS   3 CREDITS

Provides a rigorous introduction to modeling and prediction of financial time series. The goals are to learn basic characteristics of financial data, understand the application of financial econometric models, and gain experience in analyzing financial time series. We begin with the basic concepts of linear time series analysis such as stationarity and autocorrelation function, introduce regres-sion models with time series errors, seasonality, unit-root non-stationarity, and long-memo-ry processes. We provide methods of analysis in the presence of conditional heteroscedasticity and serial correlations of asset returns. The course introduces heavy-tailed distributions, and their application to financial risk management.

In particular, we discuss modern valuations of credit risk. We introduce multivariate time series analysis and apply the concept of co-integration to investi-gate arbitrage opportunity in pairs trading. The course places great emphasis on empirical data analysis. We use real examples and exercises in R and Python are included. The course aims to broaden the horizons of students in applied mathematics and to provide conceptual background to students who are interested in a career in financial industry.

7. SPECIAL TOPICS   1-3 CREDITS

This course provides the opportunity to take on emerging theory, phenomena, and technologies in the field of artificial intelligence, machine learning, data science, and big data generally. This will be an advanced class, whether seminar style or project based.

8. INDEPENDENT STUDY   1-3 CREDITS

The course provides flexibility to learn more about a topic of interest outside of the formal course setting. The subject should be chosen in consultation with a faculty advisor who acts as the student's supervisor, and with the permission of the program director. The student is required to submit a course contract describing the course of study and its specific learning objectives.

9. ADVANCED DATA ENGINEERING***   3 CREDITS

As both the volume and the velocity of data increase exponentially, problems in both commerce and research become increasingly reliant on environments with distributed data storage and data processing capabilities. This requires rethinking how our entire approach to distributed environ-ments. This course provides students with the concepts, data structures, and algorithms needed to implement data science applications in distributed computing environments. In this course, we will implement and apply distributed algorithms, data frames, and streaming. You will also learn how to choose appropriate distributed algorithms based on the characteristics of the problem and the system.
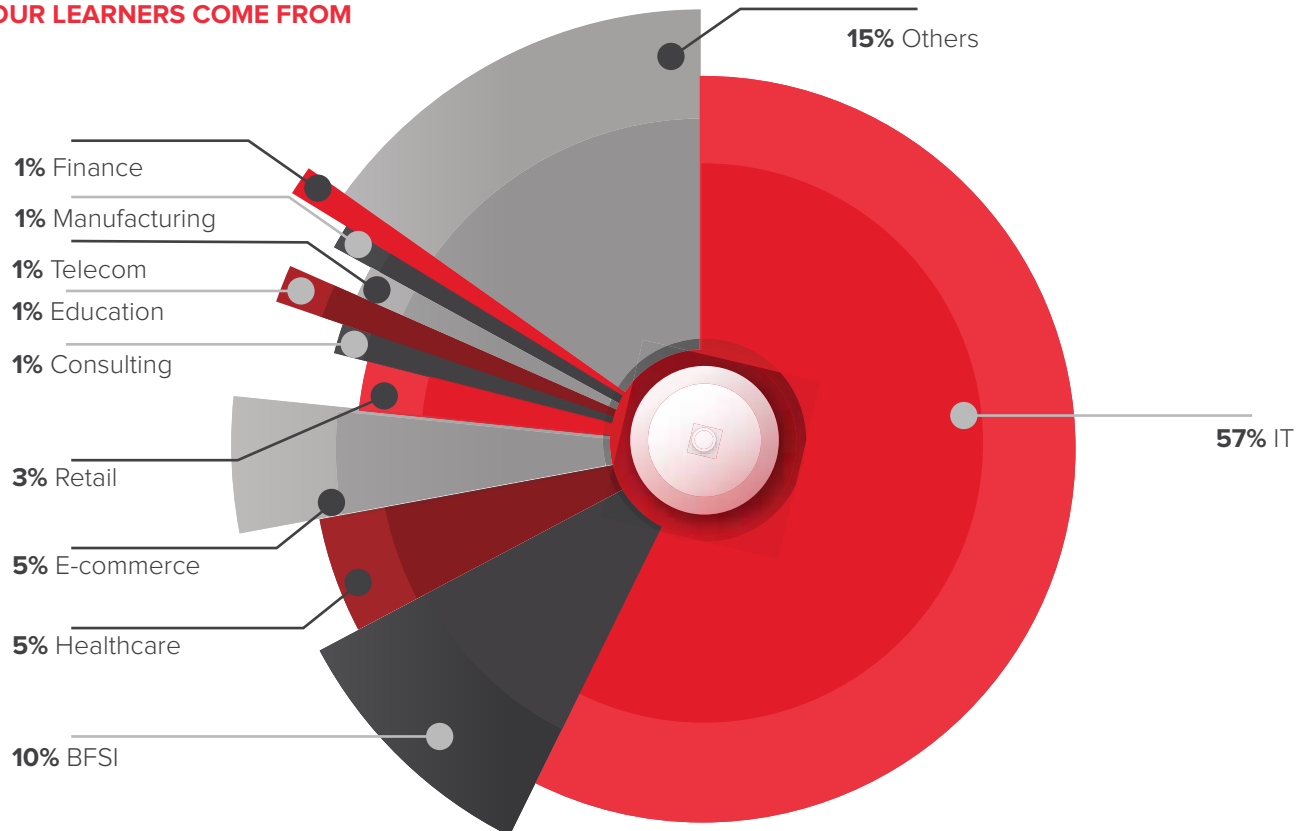
***Advanced Data Engineering is waived If a learner has done Data Engineering I Data Engineering track at IIITB ACP program.
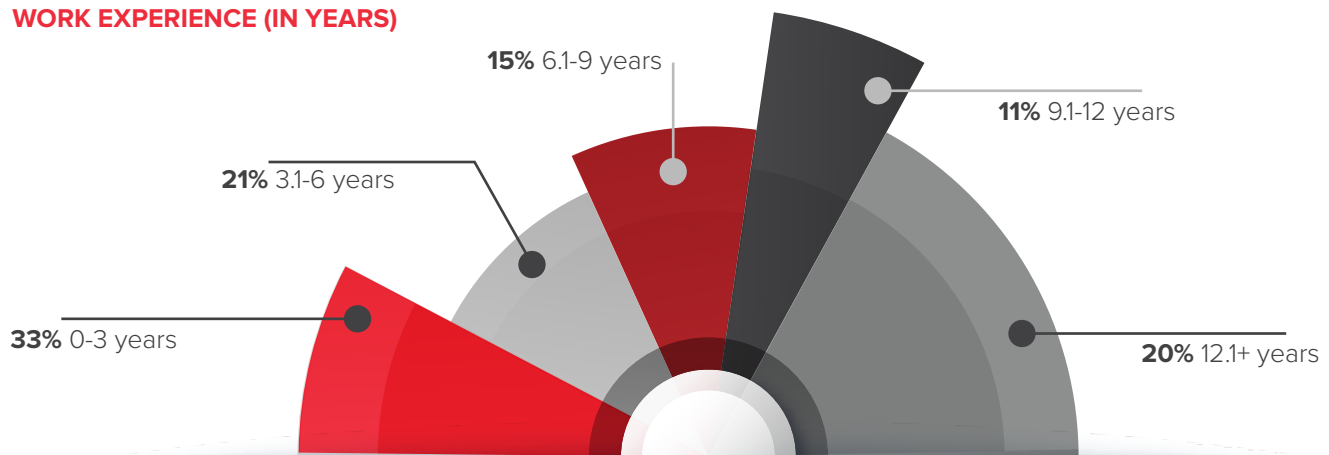
Note:
Electives offerings will vary each semester. Therefore, some choices will not be available for a particular cohort.

MEET
THE CLASS

INDUSTRIES
OUR LEARNERS COME FROM

15% Others

1% Finance

1% Manufacturing

1% Telecom

1% Education

1% Consulting

3% Retail

5% E-commerce

5% Healthcare

10% BFSI

57% IT

WORK EXPERIENCE (IN YEARS)

15% 6.1-9 years

11% 9.1-12 years

21% 3.1-6 years

33% 0-3 years

20% 12.1+ years

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

HEAR FROM
OUR LEARNERS (IIITB)

Kumar Alok, Experience: 15+ Years

"You may not believe but I have never done coding in my life. I did it during this program 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 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 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 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 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 top twenty prominent data science academicians list 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 how they have arranged the modules and the assignments are quality ones. You will relish your college days again as the exams are felt like semester ones, you can't talk to anybody . Modules are locked till 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 regimen for studying the program and keep learning is the key to be successful and pass the program. This program will help build a strong foundation for successful transition to 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 career in the field of data science, it is worth to opt for the complete program by IIIT B and upGrad. Data science is experimental science. We need to develop right kind of thinking ability to approach to the problems. And to develop this ability, we need experts 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 student for their future as they provide 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 Programme from IIITB and 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."

UPGRAD ABROAD
ONSHORE ALUMNI BENEFITS

Our support does not end up once you board a flight to your study destination but also continues in the University you join. Once you reach the University, you will be provided following additional support by upGrad abroad:

- **upGradabroad Onshore Buddy**

  Entering a foreign land can be a blissful experience when one knows that someone is always there to assist in attuning to the culture and other important aspects related to the country. With upGradabroad courses, the students will be entitled to one such service which is onshore upGradabroad buddy.

  As the name suggests, the onshore upGradabroad buddy will extend the support in many ways like helping them with the orientation session, arranging the city tours, guiding you with the local visa process, finding the accommodations, city registration, opening a bank account, and other mandatory services.

  Not only this, the onshore buddy will organize insightful sessions on destination country cultures so that the students can easily get along with the local students at the university. The onshore buddy will extend his support in finding part-time work at the university or outside University (if allowed) and providing guidance on how to apply for a post-study work visa / OPT (optional practical training).

- **Extended upGrad Career Services**

  There are also extended career services offered by upGradabraod and that include building a strong resume, organizing mock interviews, and providing guidance on the best practices to find onshore jobs.

- **Alumni Portal**

  As an upGradabroad student, the students are entitled to access the Alumni portal where they can build a professional network with other upGradabroad Alumni for assimilating information related to finding jobs or social integration.

- **upGrad Alumni Discount**

  Under the upGrad Alumni discount, the students can avail of great discounts on the courses that they would like to take up in the future which can help them in upskilling their careers and suit their resume for jobs abroad.

ADMISSION PROCESS

PROGRAM DURATION AND FORMAT

8 months Online | 12 months On-Campus in United States

PROGRAM FEE

Online Course Fee **INR 4,00,000** (incl. of all taxes)

PROGRAM FEE

On-Campus at Yeshiva University

- Indicative Tuition Fees - **$23,000**
  (It does not include semester fees or cost of textbooks)

- Indicative Cost of Living - **$22,490**

ELIGIBILITY CRITERIA

- Applicants with 4 year Bachelor's Degree in B.E/B.Tech (Only Computer Science/ IT) with a minimum of 60% marks. Maximum backlogs (ATKTs) accepted will be 5
- Applicants with 3 year Bachelor's Degree in BSc (IT/CS/Mathematics) or a BCA Degree with a minimum of 60% marks. Maximum backlogs (ATKTs) accepted will be 4.
- A 4-year Undergraduate degree is preferred. However, if a learner has completed 3 years of Undergraduate, then the awarding University must be NAAC A or A+ (in the year 2019/2020/2021) and preferably from a Technical background i.e. BCA/B.Sc/BIT.

PROGRAM START DATES

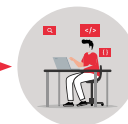Please refer to the website for program start dates.

SELECTION PROCESS

Step 1: Complete Application Form
Fill out an application form Online

Step 2: Get Shortlisted & Received your offer Letter
Our admissions committee will review your profile. Upon qualifying, an offer letter will be sent to you.

Step 3: Block your Seat & Begin the Prep Course
Block your seat with a payment of INR 25,000 to enroll on the programme. Begin with your Prep course and start your Artificial Intelligence journey!

upGrad

For further details, contact:
admissions@upgrad.com | 1800-210-2020

upGrad Education Private Limited
Nishuvi, 75, Dr. Annie Besant Road Worli, Mumbai - 400018
info@upgrad.com | 1800-210-2020

upgrad.com