upgrad.com

upGrad | iit-b ज्ञानमुत्तमम् & LIVERPOOL JOHN MOORES UNIVERSITY

Master of Science in Data Science

18 Months | Online

upGrad

upGrad is an online education platform to help individuals develop their professional potential in the most engaging learning environment. Online education is a fundamental disruption that will have a far-reaching impact. At upGrad, we are working towards transforming this online education wave into a tsunami! We are taking a full-stack approach of leveraging content, technology, marketing and services to offer quality education at scale in partnership with corporates & academics to offer a rigorous & industry relevant program.

Based on our market research and conversation with the industry, we have identified Data Science as one of the sectors with critical supply demand imbalance. Our vision is to design and deliver a quality online Master of Science in Data Science to drive the growth of the sector and make India a global hub for data science.

If you are reading this, you may wish to accelerate your career in Data Science. With upGrad, we promise to equip you with the perfect mix of business acumen and technical capabilities to help you achieve exactly the same.

Ronnie Screwvala

Co-founder & Chairman
upGrad

WHY DATA SCIENCE WITH UPGRAD AND LJMU

INTEGRATED MASTER'S FROM IIIT-B AND LJMU

Do a PG Diploma from IIIT Bangalore & Master of Science from LJMU

ONE-ON-ONE MENTORSHIP

Get mentored by a thesis supervisor for the dissertation project.

GLOBAL ACCESS TO JOBS

Get 360 degree career support and access to opportunities globally

GET ALUMNI STATUS + LIBRARY ACCESS

Earn Alumni status of IIIT Bangalore and LJMU, with digital library access from LJMU

INSIGHTS FROM LJMU FACULTY

PROF PAULO LISBOA
HOD - Applied Mathematics
LJMU

DR ATIF WARAICH
Faculty - Computer Science
LJMU

PROF DHIYA AL-JUMEILY
Associate Dean
LJMU

DR GABRIELA CZANNER
Faculty - Engineering and Tech
LJMU

CONCEPTS FROM INDUSTRY EXPERTS

TEJAS SANGHVI
Vice President
Fractal Analytics

ANSHUMAN GUPTA
Director - Data Science
Pitney Bowes

UJJYAINI MITRA
Head of Analytics
Viacom 18

S. ANAND
CEO
Gramener

SAMEER DHANRAJANI
CSO
Fractal Analytics

KALPANA SUBBARAMAPPA
Ex-Assis. VP, Decision Sciences
GENPACT

SAI ALLURI
PRO Analytics &
Strategy Manager
Uber

HINDOL BASU
Partner
Tata IQ

INSIGHTS FROM IIIT-B FACULTY

PROF. S. SADAGOPAN
Director
IIIT Bangalore

TRICHA ANJALI
Associate Professor
IIIT Bangalore

CHANDRASHEKAR RAMANATHAN
Dean (Academics)
IIIT Bangalore

DINESH BABU JAYAGOPI
Assistant Professor
IIIT Bangalore

G SRINIVASARAGHAVAN
Professor
IIIT Bangalore

MASTER OF SCIENCE IN DATA SCIENCE. HOW DOES IT WORK?

MASTER OF SCIENCE IN DATA SCIENCE

PG DIPLOMA IN DATA SCIENCE

12 months | 110 credits

International Institute of Information Technology Bangalore

Next

After successful completion of the PG Diploma in Data Science | upGrad and IIIT-B

RESEARCH METHODOLOGIES

2 months | 10 credits

LIVERPOOL JOHN MOORES UNIVERSITY

MASTER'S DISSERTATION

4 months | 60 credits

LIVERPOOL JOHN MOORES UNIVERSITY

PROGRAM CURRICULUM

This curriculum is subject to change based on inputs from IIIT-B and industry modules marked as (*) are optional.

CURRICULUM
PREPARATORY COURSE

DATA ANALYTICS in EXCEL
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
This module covers concepts of the CRISP-DM framework for business problem-solving.

COURSE 1 - DATA TOOLKIT

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

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

PROGRAMMING IN PYTHON
Learn how to approach and solve logical problems using programming.

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

VISUALIZATION IN PYTHON
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.

IMDB MOVIE ASSIGNMENT
Reinforce the concepts learnt in data science through this rigorous assignment involving the past hundred years of movie data.

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

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

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

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

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

COURSE 2 - MACHINE LEARNING

INTRODUCTION TO MACHINE LEARNING AND 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.

LINEAR REGRESSION ASSIGNMENT
Build a model to understand the factors car prices vary on and help a Chinese company enter the US car market.

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

UNSUPERVISED LEARNING: 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.

BUSINESS PROBLEM SOLVING
Learn how to approach open ended real world problems using data as a lever to draw action-able insights.

ASSIGNMENT: UNSUPERVISED + SUPERVISED
Apply the machine learning concepts learnt to solve a real-life predictive analytics problem.

CASE STUDY: LEAD SCORING
Help the Sales team of your company identify which leads are worth pursuing through this classification case study.

SPECIALISATION 1: NLP
COURSE 3 - MACHINE LEARNING II

TREE MODELS
Learn how the human decision making process can be replicated using a decision tree and other powerful ensemble algorithms.

MODEL SELECTION & GENERAL ML TECHNIQUES
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.

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

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

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

PRINCIPAL COMPONENT ANALYSIS
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.

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

TELECOM CHURN CASE STUDY
Solve the most crucial business problem for a leading telecom operator in India and southeast Asia - predicting customer churn.

COURSE 4 - NATURAL LANGUAGE PROCESSING

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

SYNTACTIC PROCESSING
'Learn how to analyse the syntax or the grammatical structure of sentences with the help of algorithms & techniques like HMMs, Viterbi Algorithm, Named Entity Recognition (NER), etc.'

SYNTACTIC PROCESSING ASSIGNMENT
Build a POS tagger for tagging unknown words using HMM's & modified Viterbi algorithm.

SEMANTIC PROCESSING
Learn the most interesting area in the field of NLP and understand different techniques like word-embeddings, LSA, topic modelling to build an application that extracts opinions about socially relevant issues (such as demonetisation) on social media platforms

CHATBOT CASE STUDY
Imagine if you could make restaurant booking without opening Zomato. Build your own restaurant-search chatbot with the help of RASA - an open source framework and deploy it on Slack.

COURSE 5 - CAPSTONE PROJECT

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

SPECIALISATION 2: DEEP LEARNING
COURSE 3 - MACHINE LEARNING II

TREE MODELS
Learn how the human decision making process can be replicated using a decision tree and other powerful ensemble algorithms.

MODEL SELECTION & GENERAL ML TECHNIQUES
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.

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

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

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

PRINCIPAL COMPONENT ANALYSIS
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.

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

TELECOM CHURN CASE STUDY
Solve the most crucial business problem for a leading telecom operator in India and southeast Asia - predicting customer churn.

COURSE 4 - DEEP LEARNING

INTRODUCTION TO NEURAL NETWORKS
Learn the most sophisticated and cutting-edge technique in machine learning - Artificial Neural Networks or ANNs

NEURAL NETWORKS ASSIGNMENT
Build a neural network from scratch in Numpy to identify handwritten digits.

SYNTACTIC PROCESSING ASSIGNMENT
Build a POS tagger for tagging unknown words using HMM's & modified Viterbi algorithm.

CONVOLUTIONAL NEURAL NETWORKS - INTRODUCTION AND INDUSTRY APPLICATIONS
Learn the basics of CNN and OpenCV and apply it to Computer Vision tasks like detecting anomalies in chest X-Ray scans, vehicle detection to count & categorise them to help the government ascertain the width and strength of the road.

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

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

COURSE 5 - CAPSTONE PROJECT

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

SPECIALISATION 3: BUSINESS ANALYTICS
COURSE 3 - MACHINE LEARNING II

TREE MODELS
Learn how the human decision making process can be replicated using a decision tree and other powerful ensemble algorithms.

MODEL SELECTION & GENERAL ML TECHNIQUES
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.

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

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

ADVANCED REGRESSION ASSIGNMENT
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
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.

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

TELECOM CHURN CASE STUDY
Solve the most crucial business problem for a leading telecom operator in India and southeast Asia - predicting customer churn.

COURSE 4 - BUSINESS REQUIREMENTS

ADVANCED SQL AND BEST PRACTICES
Apply advanced SQL concepts like windowing and procedures to derive insights from data and answer pertinent business questions

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

SYNTACTIC PROCESSING ASSIGNMENT
Build a POS tagger for tagging unknown words using HMM's & modified Viterbi algorithm.

STRUCTURED PROBLEM SOLVING USING FRAMEWORKS
Learn how to attack a business problem using various structured frameworks like 5W, 5WHYs, and 5PIN.

HYPOTHESIS FORMULATION
The module will equip you with a stepwise process for understanding a business problem and building hypotheses around it.

ASSIGNMENT
Apply your learnings from the course to solve a real-life business problem.

REVENUE AND OPERATIONAL COST MODELLING
Understand financial concepts such as revenue, cost of goods sold, profit, balance sheets, cash flow statements, and EBITDA. Also, get a brief introduction to economics concepts including supply and demand curves, cost curves, and a lot more!

EFFECTIVE COMMUNICATION STRATEGIES, FORMATS, AND TEMPLATES
Learn how to effectively strategise, communicate, and fine grain your data analysis projects.

PRESENTATIONS TO TECHNICAL AND NON-TECHNICAL STAKEHOLDERS
Understand how to optimally present your findings to technical and non-technical stakeholders and upgrade your storytelling skills.

BUSINESS CASE STUDY
Understand how a project in the industry is taken up and solved through a comprehensive business case study.

COURSE 5 - CAPSTONE PROJECT

CAPSTONE PROJECT
Solve an end-to-end real-life industry problem from a wide variety of domains like Marketing, Retail, E-Commerce, Supply Chain, Healthcare, BFSI, and many more.

SPECIALISATION 4:
BUSINESS INTELLIGENCE/ DATA ANALYTICS

COURSE 3 - DATA ANALYST TOOLKIT

INTRODUCTION TO DATABASES
Learn how data is stored and which database is optimal to use in a particular scenario

ADVANCED SQL AND BEST PRACTICES
Apply advanced SQL concepts like windowing and procedures to derive insights from data and answer pertinent business questions

DATA ACQUISITION ASSIGNMENT
Reinforce the concepts learnt in SQL and solve this comprehensive assignment involving hundreds of queries.

NOSQL DATABASES AND BEST PRACTICES
Take your knowledge of query languages a step further by learning about MongoDB - a NoSQL database which is becoming more and more popular in the industry.

INTRODUCTION TO CLOUD AND HIVE
Understand the basics of cloud and learn about the Hive Query Language.

SQL CASE STUDY
Understand how a project in the industry is taken up and solved through a comprehensive business case study.

COURSE 4 - STORYTELLING WITH ADVANCED VISUALIZATION

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

VISUALISATION USING TABLEAU
Learn advanced visualisation techniques using the most in-demand visualization tool in the industry.

INTERACTIVE MARKETING CAMPAIGN ANALYSIS
Apply the new found Excel and Tableau skills to solve an exciting business assignment.

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

INTRODUCTION TO R AND RSHINY
Get a brief introduction to another popular data science language and learn how to manipulate dataframes in R and learn to create attractive dashboards and web applications using ShinyR.

EFFECTIVE COMMUNICATION STRATEGIES, FORMATS, AND TEMPLATES
Learn how to effectively strategise, communicate, and fine grain your data analysis projects.

PRESENTATIONS TO TECHNICAL AND NON-TECHNICAL STAKEHOLDERS
Understand how to optimally present your findings to technical and non-technical stakeholders and upgrade your storytelling skills.

BUSINESS CASE STUDY
Understand how a project in the industry is taken up and solved through a comprehensive business case study.

COURSE 5 - CAPSTONE PROJECT

CAPSTONE PROJECT
Solve an end-to-end real-life industry problem from a wide variety of domains like Marketing, Retail, E-Commerce, Supply Chain, Healtcare, BFSI, and many more.

SPECIALISATION 5: DATA ENGINEERING

COURSE 3 - DATA ENGINEERING I

INTRODUCTION TO HADOOP AND MAPREDUCE PROGRAMMING
Understand the world of distributed data processing and storage with Hadoop. Learn to write MapReduce jobs in Python.

DATA MANAGEMENT AND RELATIONAL DATABASE MODELLING
"Understand the concepts of Data Management and learn to model data from a Relational Database"

NOSQL DATABASES AND APACHE HBASE
Learn the concepts of NoSQL databases. Understand the workings of Apache HBase.

DATA WAREHOUSING (OPTIONAL)
Understand the intricacies behind designing a data warehouse and a data lake for your use case.

DATA INGESTION WITH APACHE SQOOP AND APACHE FLUME
Get familiar with the challenges involed in data ingestion. Use Sqoop and Flume to ingest structured and unstructured data into Hadoop.

BUILDING AND QUERYING DATA WAREHOUSE WITH APACHE HIVE
Manage and query a data warehouse with Apache Hive. Learn to write optimized HQL for large scale data analysis.

CASE STUDY: INGESTION & WAREHOUSING
Make use of Sqoop, Flume, Hive and HBase to design an ETL data pipeline.

COURSE 4 - DATA ENGINEERING II

DATA PROCESSING WITH PYSPARK
Get introduced tp Apache Spark, a lighting fast big data processing engine. Use PySpark to create large scale data processing applications.

REAL-TIME DATA STREAMING WITH APACHE KAFKA
Understand the producer-consumer architecture of Apache Kafka. Learn to set up a Kafka cluster for managing real-time data.

"REAL-TIME DATA PROCESSING USING SPARK STREAMING"
Learn about the real-time data processing architecture of Apache Spark. Build Spark Streaming applications to process data in real-time.

ASSIGNMENT (OPTIONAL)
Use Kafka and Spark to develop a real-time data processing applicaiton.

BUILDING AUTOMATED DATA PIPELINES WITH OOZIE/AIRFLOW
Automate your Data Pipeline with Apache Airflow

ANALYTICS USING PYSPARK
Use PySpark to do EDA and Predictive Analysis of large datasets.

CASE STUDY: KAFKA, SPARK STREAMING AND PYSPARK
Build an end-to-end real-time data processing application using Spark Streaming and Kafka

COURSE 5 - CAPSTONE PROJECT

CAPSTONE PROJECT
The capstone project will stich all the components of data engineering together.

RESEARCH METHODOLOGY

(DURATION: 2 MONTHS) 10 CREDITS

WHAT IS RESEARCH?

Familiarise with different aspects of research

- Intro to research
- Importance of research
- Criticism in research and its importance
- Peer reviews in research and its importance

TYPES OF RESEARCH

Develop an understanding of various research design and techniques

- Descriptive vs Analytical
- Applied vs Fundamental
- Quantitative vs Qualitative
- Bayesian vs Frequentist Approach

RESEARCH PROCESS

Learn about the different steps in the research process and how to evaluate a literature

- Research question
- Hypothesis and aims
- Formulating a Problem
- Literature review

RESEARCH PROJECT MANAGEMENT

Learn how to plan the project timelines and arrange for data & software

- Understand the different steps involved in a project cycle
- Project Requirements on Data
- Identifying the milestones in a research project
- Learn how to track the progress using Gantt Chart

REPORT WRITING AND PRESENTATION

Master good scientific writing and proper presentation skills

- Art of writing papers
- Parts of a paper
- Tools to write papers
- Publishing papers: Journals + Seminars

SCIENTIFIC ETHICS

Develop an understanding of the ethical dimension in research

- Citation Methods and Rules
- Honor Code
- Research Claims
- IPh

MASTER'S DISSERTATION

(DURATION: 4 MONTHS) 60 CREDITS

MASTER'S THESIS

Articulate, research and present your project.

- Monthly Checkpoints
- Submission

An example of project outlines is here:

- Investigate the risk factors for eye disease from complex longitudinal datasets
- Investigate a diagnosis of eye diseases using imaging ophthalmic data
- Multi-task learning for drug design and discovery
- Using stacking for brain tumour discrimination
- Investigate dietary patterns and metabolite fingerprints of takeaway (fast) food consumers using PCA and Clustering methods
- Longitudinal studies to investigate the complex link between corporate environment engagement, green disclosure, business model transformation and supply chain performance
- Preventing credit card fraud through pattern recognition
- Developing a recommender system for a Media giant
- Using social media feed to place tweets regarding natural disasters on a map

PROGRAM DETAILS

PROGRAM STARTS
Please refer to the website for program start dates

DURATION
18 months

PROGRAM FLOW
12 months - PG Diploma in Data Science
2 months - Research Methodology
4 months - Master's Dissertation

WEEKLY COMMITMENT
12 hours per week

PROGRAM FEE
₹4,85,000 (Incl. of all taxes)
Flexible Payment Options Available

For further details, call us at +91 9372718764 or contact:

Kunal Naresh Babu
Chief Admissions Counsellor
kunal.babu@upgrad.com

ROHIT SHARMA
Program Director
mastersds@upgrad.com

upGrad

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