

A dark, moody photograph of a person from the waist up, looking slightly down and to the side. Their face and shirt are partially obscured by a projection of colorful digital code and data. The background is dark, suggesting a low-light environment.

MASTER'S COURSE IN BIG DATA AND DATA SCIENCE

Master Big Data and Data Science skills and take your career to the next level!



1.2 Million
Learners



1:1 Personalized
Mentorship



55% Average
Salary Hike

Master's Course in Big Data and Data Science

This master's course by Intellipaat, in collaboration with IBM and Microsoft, is designed by industry experts and will help you gain proficiency in Big Data and Data Science. Gain real-world knowledge of market-relevant concepts and technologies by working on industry-based projects.



Hottest Job of 21st Century



2.4 Million Job Postings

There is a global estimate of 2.4 million job postings for Big Data and Data Science roles by 2022



Skill Development

Big Data and Data Science professionals are equipped with various relevant skills fetching lucrative job offers



Growing Big Data and Data Science Industry

48.2% CAGR in the global Big Data and Data Science industry



Future-oriented Career

Big Data and Data Science is a budding field; a head start will prove to be beneficial



Popular Degree

48% of Big Data and Data Science professionals have a Master's degree



High Demand

By 2022, India and US will face a demand supply gap of 610,000 Big Data and Data Science professionals

Our Credentials



1.2 Million+

Aspiring Active Students



1,000+

Industry-expert Instructors



400+

Hiring Partners



200+

Corporates Upskilled



55%

Average Salary Hike



155+

Countries' Learners

About Program

Intellipaat's Big Data and Data Science Master's Course will provide you with in-depth knowledge of designing, developing, and deploying Data Science and Big Data applications in the real world. You will also get to work with real-time analytics, statistical computing, parsing machine-generated data, and creating NoSQL applications.



Learning Format

Online



322 Hours

Duration



Career Services

by Intellipaat



IBM and Microsoft Certifications

Key Highlights

- 322 Hrs Instructor-led Training
- 528 Hrs Projects and Exercises
- Certification from IBM and Microsoft
- Job Assistance
- 3 Guaranteed Interviews by Intellipaat
- 24*7 Support
- 381 Hrs Self-paced Videos
- 20+ Courses
- Flexible Schedule
- Lifetime Free Upgradation
- 1:1 with Industry Mentors
- No-cost EMI Option

Program Pedagogy



Instructor-led Training

Get trained by top industry experts



Hackathons

Get a sense of how real projects are built



24*7 Technical Support

Speak to Subject Matter Experts anytime and clarify your queries instantly.



Peer Networking and Group Learning

Improve your professional network and learn from peers through our innovative Peer Chat tool.



Self-paced videos

Learn at your own pace with world-class content



Gamified Learning

Get involved in group activities to solve real-world problems



Projects and Exercises

Get real-world experience through projects

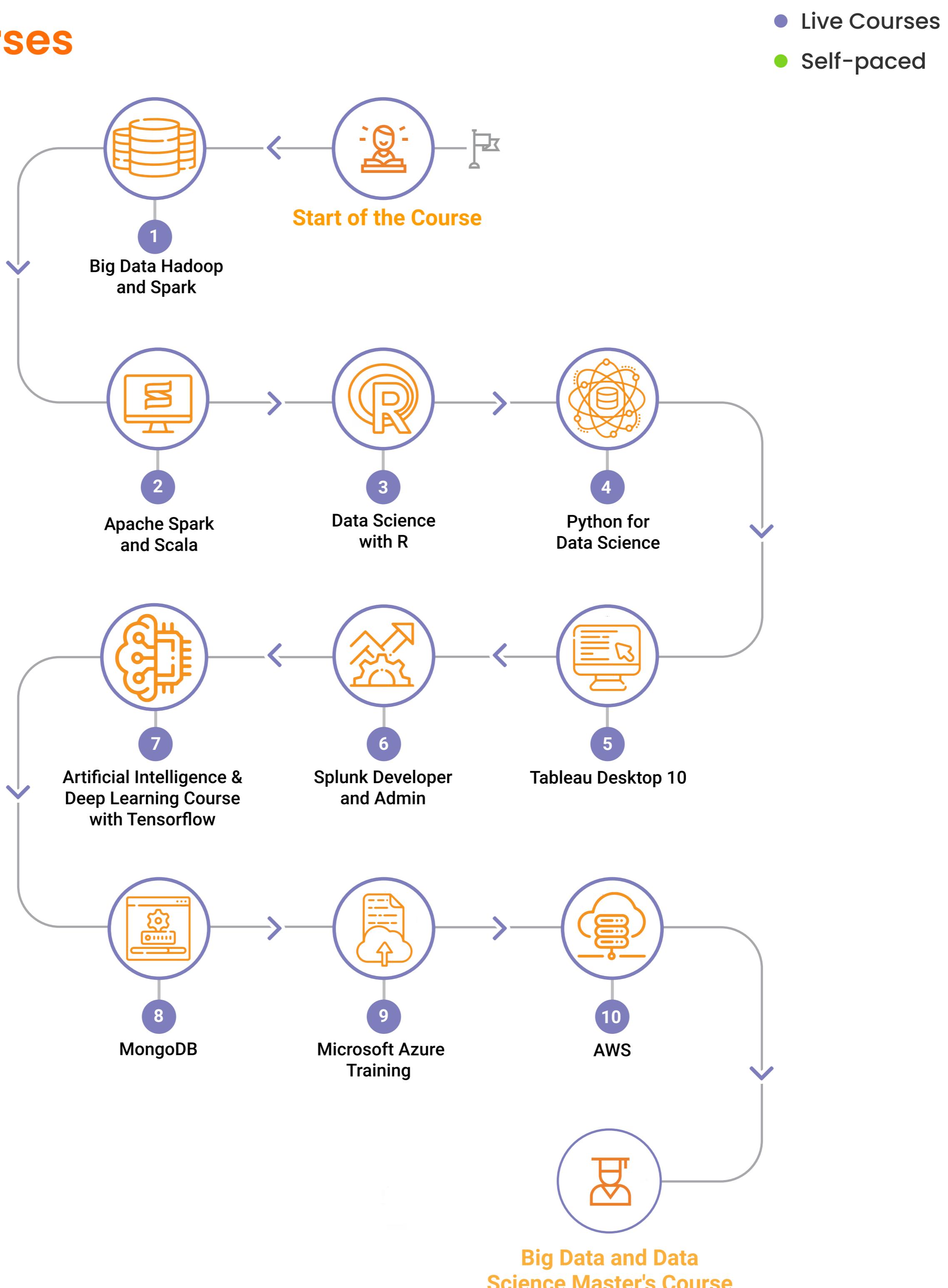


1:1 Personalized Learning

Hands-on exercises, project work, quizzes, and capstone projects

Learning Path

Live Courses



Self-paced Courses

- Machine Learning
- Solr
- Linux
- Java
- Apache Kafka
- SQL
- Apache HBase
- Apache Cassandra
- Couchbase

Program Curriculum

Module 1

Big Data Hadoop

- Hadoop Installation and Setup
- Introduction to Big Data Hadoop and Understanding HDFS and MapReduce
- Deep Dive in MapReduce
- Introduction to Hive
- Advanced Hive and Impala
- Introduction to Pig
- Flume, Sqoop, and HBase
- Writing Spark Applications Using Scala
- Use Case Bobsrockets Package

Following topics will be available only in self-paced mode:

- Hadoop Administration – Multi-node Cluster Setup using Amazon EC2
- Hadoop Administration – Cluster Configuration
- Hadoop Administration – Maintenance, Monitoring, and Troubleshooting
- ETL Connectivity with Hadoop Ecosystem (Self-paced)
- Hadoop Application Testing
- Roles and Responsibilities of Hadoop Testing Professional
- Framework called MRUnit for Testing of MapReduce Programs
- Unit Testing
- Test Execution
- Test Plan Strategy and Writing Test Cases for Testing Hadoop Application

Module 2

Apache Spark and Scala

Scala Course Content

- Introduction to Scala
- Pattern Matching
- Executing the Scala Code
- Classes Concept in Scala
- Case Classes and Pattern Matching
- Concepts of Traits with Examples
- Scala–Java Interoperability
- Scala Collections
- Mutable Collections vs Immutable Collections
- Use Case Bobsrockets Package

Spark Course Content

- Introduction to Spark
- Spark Basics
- Working with RDDs in Spark

- Aggregating Data with Pair RDDs
- Writing and Deploying Spark Applications
- Parallel Processing
- Spark RDD Persistence
- Spark MLlib
- Integrating Apache Flume and Apache Kafka
- Spark Streaming
- Improving Spark Performance
- Spark SQL and Data Frames
- Scheduling/Partitioning

Module 3

Data Science with R

- Introduction to Data Science with R
- Data Exploration
- Data Manipulation
- Data Visualization
- Introduction to Statistics
- Machine Learning
- Logistic Regression
- Decision Trees and Random Forest
- Unsupervised Learning
- Association Rule Mining and Recommendation Engines

Self-paced Course Content

- Introduction to Artificial Intelligence
- Time Series Analysis
- Support Vector Machine (SVM)
- Naïve Bayes
- Text Mining

Module 4

Python for Data Science

- Introduction to Data Science using Python
- Python basic constructs
- Maths for DS–Statistics & Probability
- OOPs in Python (Self paced)
- NumPy for mathematical computing
- SciPy for scientific computing
- Data manipulation
- Data visualization with Matplotlib
- Machine Learning using Python
- Supervised learning
- Unsupervised Learning
- Python integration with Spark (Self paced)

Program Curriculum

- Dimensionality Reduction
- Time Series Forecasting

Module 5

Tableau Desktop 10

- Introduction to Data Visualization and The Power of Tableau
- Architecture of Tableau
- Charts and Graphs
- Working with Metadata and Data Blending
- Advanced Data Manipulations
- Working with Filters
- Organizing Data and Visual Analytics
- Working with Mapping
- Working with Calculations and Expressions
- Working with Parameters
- Dashboards and Stories
- Tableau Prep
- Integration of Tableau with R

Module 6

Splunk Developer and Admin

- Splunk Development Concepts
- Basic Searching
- Using Fields in Searches
- Saving and Scheduling Searches
- Creating Alerts
- Scheduled Reports
- Tags and Event Types
- Creating and Using Macros
- Workflow
- Splunk Search Commands
- Transforming Commands
- Reporting Commands
- Mapping and Single Value Commands
- Splunk Reports and Visualizations
- Analyzing, Calculating and Formatting Results
- Correlating Events
- Enriching Data with Lookups
- Creating Reports and Dashboards
- Getting Started with Parsing
- Using Pivot
- Common Information Model (CIM) Add-On

Splunk Administration Topics

- Overview of Splunk
- Splunk Installation
- Splunk Installation in Linux
- Distributed Management Console
- Introduction to Splunk App
- Splunk Indexes and Users
- Splunk Configuration Files
- Splunk Deployment Management
- Splunk Indexes
- User Roles and Authentication
- Splunk Administration Environment
- Basic Production Environment
- Splunk Search Engine
- Various Splunk Input Methods
- Splunk User and Index Management
- Machine Data Parsing
- Search Scaling and Monitoring
- Splunk Cluster Implementation

Module 7

Artificial Intelligence & Deep Learning Course with Tensorflow

- Introduction to Deep Learning and Neural Networks
- Multi-layered Neural Networks
- Artificial Neural Networks and Various Methods
- Deep Learning Libraries
- Keras API
- TFLearn API for TensorFlow
- Deep Neural Networks (DNNs)
- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks(RNNs)
- GPU in Deep Learning
- Autoencoders and Restricted Boltzmann Machine (RBM)
- Deep Learning Applications
- Chatbots

Module 8

MongoDB

- Introduction to NoSQL and MongoDB
- MongoDB Installation
- Importance of NoSQL
- CRUD Operations
- Data Modeling and Schema Design
- Data Management and Administration

Program Curriculum

- Data Indexing and Aggregation
- MongoDB Security
- Working with Unstructured Data

Module 9

Microsoft Azure Training

- Introduction to Microsoft Azure
- Introduction to ARM & Azure Storage
- Introduction to Azure storage
- Azure Virtual Machines
- Azure App and Container services
- Azure Networking – I
- Azure Networking – II
- Authentication and Authorization in Azure using RBAC
- Microsoft Azure Active Directory
- Azure Monitoring

Module 10

AWS

- Introduction to Cloud Computing & AWS
- Elastic Compute and Storage Volumes
- Load Balancing, Autoscaling, and DNS
- Virtual Private Cloud
- Storage – Simple Storage Service (S3)
- Databases and In-memory Datastores
- Management and Application Services
- Access Management and Monitoring Services
- Automation and Configuration management
- AWS Migration

Self-paced

- Architecting AWS – Whitepaper
- DevOps on AWS
- Amazon FSx and Global Accelerator
- AWS Architect Interview Questions

Module 12

Apache HBase

- HBase Overview
- Architecture of NoSQL
- HBase Data Modeling
- HBase Cluster Components
- HBase API and Advanced Operations
- Integration of Hive with HBase
- File Loading with Both Load Utilities

Module 13

Apache Cassandra

- Advantages and Usage of Cassandra
- CAP Theorem and No SQL DataBase
- Cassandra Fundamentals, Data model, Installation, and Setup
- Cassandra Configuration
- Summarization, Node Tool Commands, Cluster, Indexes, Cassandra & MapReduce, Installing Ops-center
- Multi-cluster Setup
- Thrift/Avro/Json/Hector Client
- Datastax Installation Part, Secondary index
- Advance Modeling
- Deploying IDE for Cassandra applications
- Cassandra Administration
- Cassandra API and Summarization and Thrift

Module 14

Couchbase

- Introduction to Couchbase
- Single-node Implementation
- Couchbase Web Console
- Couchbase Multi-node Cluster
- Couchbase Command-line Interface

Module 15

Machine Learning

- Introduction to Machine Learning
- Supervised Learning and Linear Regression
- Classification and Logistic Regression
- Decision Tree and Random Forest
- Naïve Bayes and Support Vector Machine (Self-paced)
- Unsupervised Learning
- Natural Language Processing and Text Mining
- Introduction to Deep Learning
- Time-series Analysis

Module 16

Solr

- Fundamentals of Search Engine and Apache Lucene
- Analyzers in Lucene
- Exploring Apache Lucene
- Apache Lucene Demonstration
- Apache Lucene advanced

Program Curriculum

- Advanced Topics of Apache Lucene (Practical)
- Apache Solr
- Apache Solr Indexing
- Solr Indexing continued
- Apache Solr Searching
- Deep Dive into Apache Solr
- Apache Solr continued
- Extended Features
- Multicore
- Administration & SolrCloud

Module 17

Linux

- Introduction to Linux
- File Management
- Files and Processes
- Introduction to Shell Scripting
- Conditional, Looping Statements, and Functions
- Text Processing
- Scheduling Tasks
- Advanced Shell Scripting
- Database Connectivity
- Linux Networking

Module 18

Java

- Core Java Concepts
- Writing Java Programs using Java Principles
- Language Conceptuals
- Operating with Java Statements
- Concept of Objects and Classes
- Introduction to Core Classes
- Inheritance in Java
- Exception Handling in Detail
- Getting started with Interfaces and Abstract Classes
- Overview of Nested Classes
- Getting started with Java Threads
- Overview of Java Collections
- Understanding JDBC
- Java Generics
- Input/Output in Java
- Getting started with Java Annotations
- Reflection and its Usage

Module 19

Apache Kafka

- What is Kafka – An Introduction
- Multi-broker Kafka Implementation
- Multi-node Cluster Setup
- Integrate Flume with Kafka
- Kafka API
- Producers & Consumers

Module 20

SQL

- Introduction to SQL
- Database Normalization and Entity Relationship Model
- SQL Operators
- Working with SQL – Join, Tables, and Variables
- Deep Dive into SQL Functions
- Working with Subqueries
- SQL Views, Functions, and Stored Procedures
- Deep Dive into User-defined Functions
- SQL Optimization and Performance
- Advanced Topics
- Managing Database Concurrency
- Programming Databases using Transact – SQL
- Microsoft Courses – Study Material

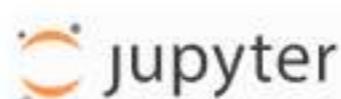
Program Curriculum

```
modifier ob.select=1  
bpy.context.scene.objects.active = modifier_ob  
print("Selected" + str(modifier_ob)) # modifier ob is  
selected ob.select = 0  
bpy.ops.object.select_all(action='DESELECT')
```

Skills to Master

- Big Data Hadoop
- Apache Spark & Scala
- Data Science with R
- Python for Data Science
- Tableau Desktop 10
- Splunk Developer & Admin
- Azure
- Artificial Intelligence and Deep
- MongoDB
- AWS
- Learning with TensorFlow

Tools to Master



Course Projects

Projects cover the following industries:



Retail



Social Media



Supply Chain



Entrepreneurship



E-commerce



Banking



Healthcare



Insurance

Working with NumPy

Beginner

Get hands-on learning to successfully work with NumPy to solve various Python problems. The project also requires learners to efficiently create 2D arrays and perform simple arithmetic operations on the arrays.

Visualizing and Analyzing the Customer Churn Dataset using Python

Beginner

Analyze data by building aesthetic graphs to make better sense of it. Work with bar plots and their applications, which also includes histogram graphs for data analysis and box plots and the outliers in them.

Deal with Financial Data

Beginner

Deploy Apache Spark and work with Spark MLlib. Perform regression, clustering, dimensionality reduction, and collaborative filtering to build a movie recommendation system.

Analyzing the Naming Trends using Python

Intermediate

Analyze the naming trends using Python. Use Python to extract data and understand the applications of data manipulation and the concepts of data visualization.

Performing Analysis on Customer Churn Dataset

Intermediate

Analyze employment reliability in the telecom industry. In this project, learners will also get to work on real-time analysis of data with multiple labels and data visualization for reliability.

Work with Customer Data of a UK Bank

Intermediate

Analyze and create a dashboard to gain insight from a UK bank's customer data in this real-life project. Build an asymmetric drop-down consisting of the region and the corresponding customer details.

Market Basket Analysis

Advance

Identify trends in a company's inventory dataset to increase sales numbers. The learners will also get to implement data extraction, data manipulation, etc., for market basket analysis.

Credit Card Fraud Detection

Advance

Perform data analysis on a banking dataset for several parameters. Use V4 predictor analysis, V7 predictor analysis, and data visualization techniques to calculate the probability of fraud activities.

Prediction of Loan Approval

Advance

Use banking datasets to analyze, clean, process, and visualize data. Once data analysis is completed, you will learn to implement Naive Bayes and Principal Component Analysis.

Analyzing the Trends of COVID-19 with Python

Advance

Use pandas to collect data from various files. Use Plotly to build interactive visualizations. Use Prophet from Facebook to efficiently develop time-series models.

Course Advisors / Faculty



Raghu Raman A V

Big Data Consultant

With over 10+ years of IT experience and corporate training, he is a subject matter expert in Big Data and Hadoop along with Cloud Computing. He has an exhaustive track record of training for major organizations including HCL, IBM, Capgemini, Wipro.



Suresh Paritala

Solutions Architect at Microsoft

Suresh is responsible for overseeing and directing Big Data and Analytics projects at Microsoft. He is experienced in advanced analytics and big data, prescriptive and predictive analysis having expertise in Hadoop and Spark.



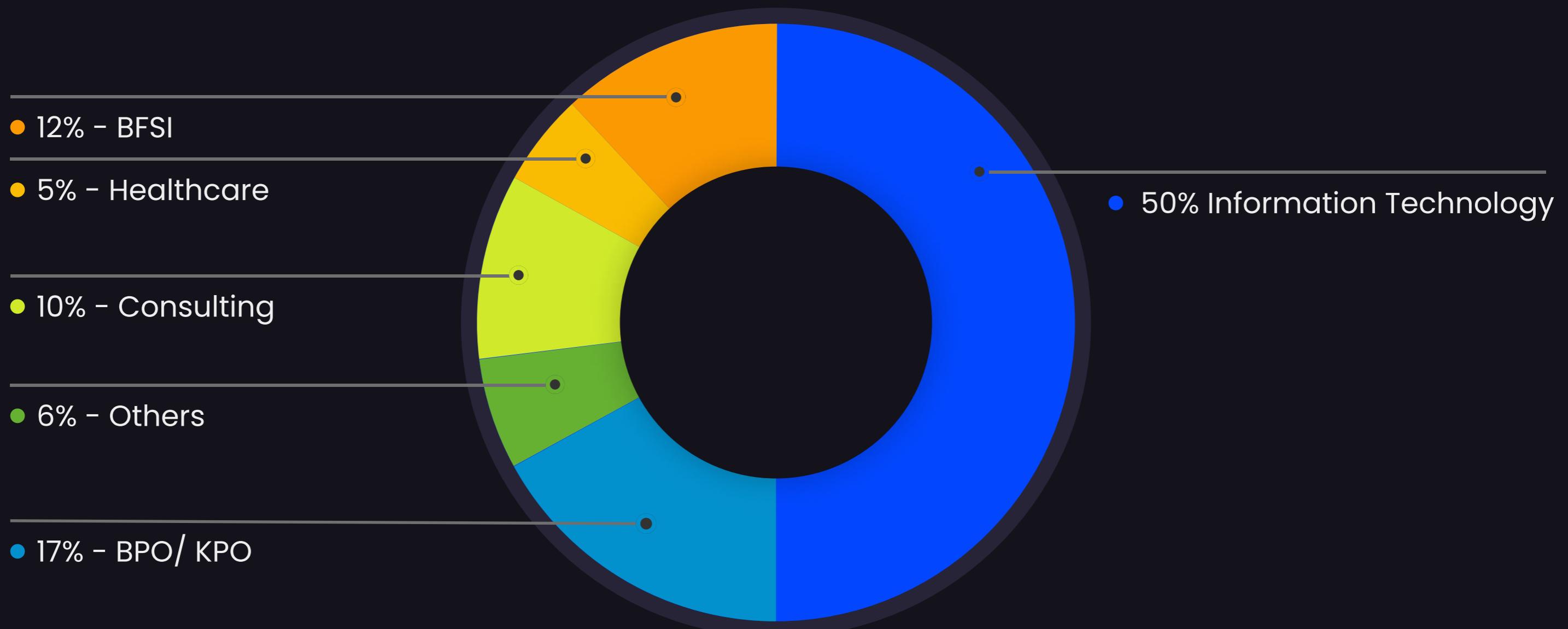
David Callaghan

Big Data Solutions Architect, USA

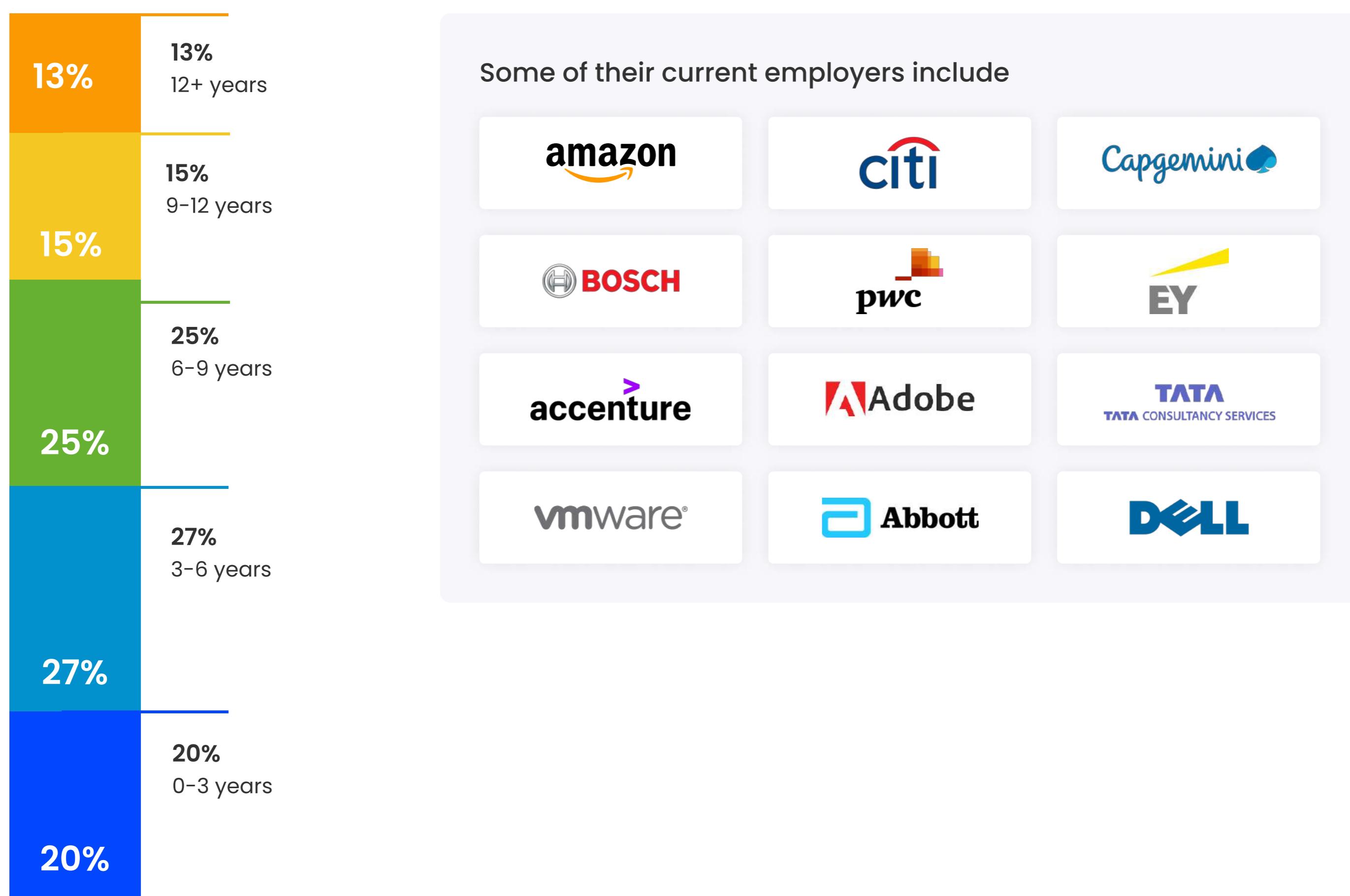
A Blockchain expert, David has been bringing integrated blockchain, particularly Hyperledger, Ethereum and Big Data solutions to the cloud. He has previously worked on Hadoop, AWS Cloud, Big Data and Pentaho Projects.

Meet the **Batch**

Industries Our Learners Come From



Work **Experience**



Glimpse of our Successful Transitions



4.6 ★★★★★



4.5 ★★★★★



4.38 ★★★★★

Career Transitions



Rajesh Venaganti
Senior Software Engineer

From
Associate
Consultant

To
Senior Software
Engineer



Pratik Kumar
Product Manager

From
Marketing Manager

To
Product Manager



Yogesh Kumar
Senior Software Engineer

From
Associate
Consultant

To
Senior Software
Engineer



Gaurav Saboo
Senior Technical Associate

From
Technical Analyst

To
Senior Technical
Associate



Ankit Kumar
Data Scientist

From
B.com Graduate

To
Data Scientist



Melwin Rodrigues
Data Scientist

From
Customer Service
Agent

To
Data Scientist



Shehzin Mulla
Research Analyst

From
Data Analyst

To
Research Analyst



D Rajesh Sahoo
Machine Learning Engineer

From
B. Tech – Civil
Engineering

To
Machine Learning
Engineer



Intellipaat Career Services



500+
Webinars



600+
Job Shares



400+
Hiring Partners



55%
Avg. Salary Hike

What Makes Us **Tick**



Career-oriented Sessions

Attend 25+ career-oriented sessions by industry mentors and plan your career trajectory



Profile Building

Craft a Big Data and Data Science resume and LinkedIn profile and make an impression on top employers



Mock Interview Preparation

Prepare with mock interviews including most asked questions by top employers



1:1 Mentoring Sessions

Get 1:1 guidance at every step in your career transition to Big Data and Data Science



Assured Interviews

Get job interviews with 400+ hiring partners including promising startups and top MNCs



Dedicated Job Portal Access

Get exclusive access to 200 job postings per month on Intellipaat's job portal



Job Fairs

Job fairs are conducted regularly to introduce learners to major organizations



Hackathons

Work in teams and get exclusive access to hackathons

Learner **Reviews**



DEVESH SINHA

Communication cell at JGEC

The course is comprehensive and has a variety of material like videos, PPTs and PDFs, that are neatly organized. Also, the support I received from the trainer during my learning was great.

“



Srikanth Penumadula

Senior Manager at Fidelity Investments

Best platform to learn. Course structure is well designed and on request I got access to two more projects that made my foundational knowledge stronger. Overall, I rate the course 4 on a scale of 5.

“



Vikrant Singh

Senior Big Data Consultant at Amazon

It was a wonderful learning experience to learn from the trainers at Intellipaat. They were hands-on and provided real-time scenarios. Intellipaat is the right place for learning latest technology.

“



About Intellipaat

Intellipaat is one of the leading online training providers with more than 1.2 million learners in over 155 countries. We are on a mission to democratize education as we believe that everyone has the right to quality education.

We create courses in collaboration with top universities and MNCs for employability like IIT Madras, University of Essex, University of Liverpool, IIT Roorkee, IIT Guwahati, SPJIMR, IBM, Microsoft, etc.

Our courses are delivered by SMEs & our pedagogy enables quick learning of difficult topics. 24/7 technical support & career services help learners to jump-start their careers.



Collaboration with IBM and Microsoft

IBM and Microsoft are leading innovators and the biggest players in creating innovative Big Data and Data Science tools. In this Master's course, top subject matter experts will share knowledge in the domain of Big Data and Data Science.

Benefits of this collaboration for learners:

- Industry-recognized certification from IBM
- Industry recognized certification from Microsoft
- Industry-specific case studies and project work



1.2 Million Learners & 200+ corporates across 155+ countries
upskilling on Intellipaat platform



Contact Us

INDIA

AMR Tech Park 3, Ground Floor, Tower B, Hongasandra Village,
Bommanahalli, Hosur Road, Bangalore, Karnataka 560068, India
Phone No: +91-7022374614

UK

Flat 16 Bluepoint Court, 203 Station Road, Harrow,
Middlesex HA1 2TS, UK

USA

1219 E. Hillsdale Blvd. Suite 205, Foster City, CA 94404
Phone No: 1-800-216-8930