Data Science Architect

Master's Course

In Collaboration with **IBM**

IntelliPaat

Table of Contents

> **Global Big Data market to reach $122 billion in revenue by 2025**
> — Frost & Sullivan

About the Program

Intellipaat's online master's in Data Science program lets you gain proficiency in Data Science. You will work on real-world projects in Data Science with R, Hadoop Dev, Admin, Test and Analysis, Apache Spark, Scala, Deep Learning, Tableau, Data Science with SAS, SQL, MongoDB and more. In this program, you will cover 10 courses and 53 industry-based projects with 1 CAPSTONE project. As a part of online classroom training, you will receive five additional self-paced courses co-created with IBM namely Deep Learning with TensorFlow, Build Chatbots with Watson Assistant, R for Data Science, Spark MLlIb, and Python for Data Science. Moreover, you will also get an exclusive access to IBM Watson Cloud Lab for Chatbots course. Enroll now and pursue your MS in Data Science online.

IBM

Collaborating with IBM

IBM is one of the leading innovators and the biggest player in creating innovative tools for Big Data Analytics. Top subject matter experts from IBM will share knowledge in the domains of Analytics and Big Data through this training program, which will help you gain the breadth of knowledge and industry experience.

Benefits for students from IBM

- Industry-recognized IBM certificate

- Access to IBM Watson for hands-on training and practice

- Industry in-line case studies and project work

> At Intellipaat, we are in a constant endeavor of changing the education culture to be more of an open-source collaborative environment.

IntelliPaat

About Intellipaat

Intellipaat is one of the leading online e-learning training providers with more than 600,000 learners across 55+ countries. We are on a mission to democratize education as we believe that everyone has the right to quality education.

Our courses are delivered by subject matter experts from top MNCs, and our world-class pedagogy enables to quickly learn difficult topics in no time. Our 24/7 technical support and career services will help learners jump-start their careers in their dream companies.

Key Features

232 HRS OF INSTRUCTOR-LED TRAINING

104 HRS OF SELF-PACED TRAINING

253 HRS OF REAL-TIME PROJECT WORK

LIFETIME ACCESS

24/7 TECHNICAL SUPPORT

INDUSTRY-RECOGNIZED CERTIFICATION

JOB ASSISTANCE THROUGH 80+ CORPORATE TIE-UPS

FLEXIBLE SCHEDULING

Career Support

SESSIONS WITH INDUSTRY MENTORS
Attend sessions from top industry experts and get guidance on how to boost your career growth

MOCK INTERVIEWS
Mock interviews to make you prepare for cracking interviews by top employers

GUARANTEED INTERVIEWS & JOB SUPPORT
Get interviewed by our 400+ hiring partners

RESUME PREPARATION
Get assistance in creating a world-class resume from our career services team

Why take up this course?

- Data Scientist is the best job of the 21st century – Harvard Business Review

- Global Big Data market to reach $122 billion in revenue by 2025 – Frost & Sullivan

This best Data Science master's program has been created keeping in mind the needs of the industry when it comes to the domain of Data Science. Today's Data Scientists need to have a diverse set of skills which include working with huge volumes of data, parsing that data and converting them into a format that is easily understandable, using which business insights can be derived. This training program lets you play multiple roles in the Big Data and Data Science domains and get hired for top-notch salaries.

Who should take up this training?

- Data Scientists, Machine Learning Professionals, and Software Developers

- Business Intelligence Professionals, Information Architects, and Project Managers

- Those looking to be a Data Science Architect

Courses offered

INSTRUCTOR LED TRAINING COURSES

1. Data Science with R

2. Python for Data Science

3. Machine Learning

4. Artificial Intelligence and Deep Learning with TensorFlow

5. Big Data Hadoop & Spark

6. Tableau Desktop 10

7. Data Science with SAS

SELF PACED LEARNING

1. Advanced Excel

2. MongoDB

3. MS-SQL

Course Advisors

Suresh Paritala

Solutions Architect at Microsoft, Texas

A Senior Software Architect at Microsoft, Texas, who has previously worked with IBM Corporation, Suresh Paritala has worked on Big Data, Data Science, Advanced Analytics, Internet of Things, and Azure, along with AI domains like Machine Learning and Deep Learning. He has successfully implemented high-impact projects in major corporations around the world.

David Callaghan

Big Data Strategist and Solutions Architect, Perficient, USA

An experienced Blockchain professional, who has been bringing integrated Blockchain, particularly Hyperledger and Ethereum, and Big Data solutions to the cloud, David Callaghan has previously worked on Hadoop, AWS Cloud, Big Data, and Pentaho projects that have had major impact on the revenues of marquee brands around the world.

Samanth Reddy

Data Scientist at ASCAP

A renowned Data Scientist who has worked with Google and is currently working at ASCAP, Samanth Reddy has a proven ability to develop Data Science strategies that have a high impact on the revenues of various organizations. He comes with strong Data Science expertise and has created decisive Data Science strategies for Fortune 500 corporations.

Program Curriculum

DATA SCIENCE WITH R

Learning Objective

This Data Scientist course online provides detailed learning through self-paced videos and live instructor-led sessions that help you gain skills in the shortest possible time. Data Scientists are among the highest-paid and most in demand professionals. This in-depth Data Scientist course covers 'What is Data Science?,' statistical methods, data acquisition and analysis, Machine Learning algorithms, predictive analytics, data modeling, etc. At the end of the course, you will work on building a recommendation engine for an ecommerce site and will work on a real-time capstone project.

Topics:

- 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

- Introduction to Artificial Intelligence (self-paced)

- Time Series Analysis (self-paced)

- Support Vector Machine (SVM) (self-paced)

- Naïve Bayes (self-paced)

- Text Mining (self-paced)

PYTHON FOR DATA SCIENCE

Learning Objective

Data Science with Python course helps you learn the python programming required for Data Science. This python for Data Science training you will master the technique of how Python is deployed for Data Science, work with Pandas library for Data Science, data cleaning, data visualization, Machine Learning, advanced numeric analysis, etc. along with real-world projects and case studies.

Topics:

- Introduction to Data Science using Python

- Python Basic Constructs

- Maths for DS-Statistics & Probability

- OOPs in Python

- 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)

- Dimensionality Reduction

- Time Series Forecasting

MACHINE LEARNING

Learning Objective

This Machine Learning online course is curated and developed by SMEs from top product-based companies to meet the needs of the current data-driven industry. It covers a detailed overview of various algorithms and techniques, such as regression, classification, time series modeling,

supervised and unsupervised learning, Natural Language Processing, etc. You will also use Python programming language to write code for implementing numerous algorithms in this certification training.

Topics:

- Introduction to Machine Learning

- Supervised Learning and Linear Regression

- Classification and Logistic Regression

- Decision Trees and Random Forest

- Naïve Bayes and Support Vector Machine (self-paced)

- Unsupervised Learning

- Natural Language Processing and Text Mining (self-paced)

- Introduction to Deep Learning

- Time Series Analysis (self-paced)

AI & DEEP LEARNING WITH TENSORFLOW

Learning Objective

Intellipaat offers a comprehensive Artificial Intelligence program that will help you work on today cutting-edge technology Artificial Intelligence (AI). As part of this best AI training, you will master various aspects of artificial neural networks, supervised and unsupervised learning, logistic regression with a neural network mindset, binary classification, vectorization, Python for scripting Machine Learning applications, and much more.

Topics:

- 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

- DNNs (Deep Neural Networks)

- CNNs (Convolutional Neural Networks)

- RNNs (Recurrent Neural Networks)

- GPU in Deep Learning

- Autoencoders and Restricted Boltzmann Machine (RBM)

- Deep Learning Applications

- Chatbots

BIG DATA HADOOP & SPARK

Learning Objective

It is a comprehensive Hadoop Big Data training course designed by industry experts considering current industry job requirements to help you learn Big Data Hadoop and Spark modules. This is an industry-recognized Big Data Hadoop certification training course that is a combination of the training courses in Hadoop developer, Hadoop administrator, Hadoop testing and analytics with Apache Spark. This Cloudera Hadoop and Spark training will prepare you to clear Cloudera CCA175 Big Data certification.

Topics:

- Hadoop Installation and Setup

- Introduction to Big Data Hadoop and Understanding HDFS and MapReduce

- Deep Dive into MapReduce

- Introduction to Hive

- Advanced Hive and Impala

- Introduction to Pig

- Flume, Sqoop, and HBase

- Writing Spark Applications Using Scala

- Spark Framework

- RDDs in Spark

- DataFrames and Spark SQL

- Machine Learning Using Spark (MLlib)

- Integrating Apache Flume and Apache Kafka

- Spark Streaming

- Hadoop Administration: Multi-node Cluster Setup Using Amazon EC2

- Hadoop Administration: Cluster Configuration

- Hadoop Administration: Maintenance, Monitoring, and Troubleshooting

- ETL Connectivity with the Hadoop Ecosystem (Self-Paced)

- Project Solution Discussion and Cloudera Certification Tips and Tricks

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

- 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

TABLEAU DESKTOP 10

Learning Objective

Intellipaat's Tableau certification training program helps you learn Tableau and gets you up to work on the concepts of data visualization with a firm understanding of the Tableau

architecture. You will become proficient in the concepts of filters, parameters, maps, graphs, dashboards, and table calculation. Furthermore, you will learn about data blending, data aggregation, and R connectivity with Tableau in this online Tableau course.

Topics:

- Introduction to Data Visualization and Power of Tableau

- Architecture of Tableau

- Working with Metadata and Data Blending

- Creation of Sets

- Working with Filters

- Organizing Data and Visual Analytics

- Working with Mapping

- Working with Calculations and Expressions

- Working with Parameters

- Charts and Graphs

- Dashboards and Stories

- Tableau Prep

- Integration of Tableau with R and Hadoop

DATA SCIENCE WITH SAS

Learning Objective

This SAS training course will help you in learning the domains of Business Analytics and Business Intelligence. Upon the completion of this SAS online training, you will have enough proficiency in reading spreadsheets, databases, using SAS functions for manipulating this data and debugging it. This Base SAS certification training also includes data mining, data analytics, modeling techniques, visualization of data, predictive analysis and extracting insights through real-world case studies.

Topics:

- Introduction to SAS

- SAS Enterprise Guide

- SAS Operators and Functions

- Compilation and Execution

- Using Variables

- Creation and Compilation of SAS Datasets

- SAS Procedures

- Input Statement and Formatted Input

- SAS Format

- SAS Graphs

- Interactive Data Processing

- Data Transformation Function

- Output Delivery System (ODS)

- SAS Macros

- PROC SQL

- Advanced Base SAS

- Summarization Reports

Self-paced Courses

As a part of online classroom training, you will receive five additional self-paced courses co-created with IBM, namely, Deep Learning with TensorFlow, Building Chatbots with Watson Assistant, R for Data Science, Spark MLlib, and Python for Data Science. Moreover, you will also get an exclusive access to IBM Watson Cloud Lab for the Chatbots course.

Project Work

Data Science with R

- Market Basket Analysis

- Credit Card Fraud Detection

- Loan Approval Prediction

- Netflix Recommendation System

- Case Study 1: Introduction to R Programming

- Case Study 2: Solving Customer Churn Using Data Exploration

- Case Study 3: Creating Data Structures in R

- Case Study 4: Implementing SVD in R

- Case Study 5: Time Series Analysis

Python for Data Science

- Analyzing the Trends of COVID-19 With Python

- Analyzing the Naming Trends Using Python

- Performing Analysis on Customer Churn Dataset

- Netflix-Recommendation System

- Python Web Scraping for Data Science

- OOPS in Python

- Working With NumPy

- Visualizing and Analyzing the Customer Churn dataset using Python

- Building Models With the Help of Machine Learning Algorithms

Machine Learning

- Analyzing the Trends of COVID-19 with Python

- Customer Churn Classification

- Creating a Recommendation System for Movies

- Case Study 1 - Decision Tree

- Case Study 2 - Insurance Cost Prediction (Linear Regression)

- Case Study 3 - Diabetes Classification (Logistic Regression)

- Case Study 4 - Random Forest

- Case Study 5 - Principal Component Analysis (PCA)

- Case Study 6 - K-means Clustering

AI & Deep Learning

- Auto-Encoder Assignment

- CNN Assignment

- Binary Classification on 'Customer_Churn' Using Keras

- Face Detection Project

- Keras Assignment

- MLP Assignment

- AI and Deep Learning Intro Assignment

- RNN Assignment

- TensorFlow Assignment

- TFLearn Assignment

Big Data Hadoop & Spark

- Working with MapReduce, Hive, and Sqoop

- Work on MovieLens Data For Finding the Top Movies

- Hadoop YARN Project: End-to-End PoC

- Table Partitioning in Hive

- Connecting Pentaho with Hadoop Ecosystem

- Multi-node Cluster Setup

- Hadoop Testing Using MRUnit

- Hadoop Web Log Analytics

- Hadoop Maintenance

- Twitter Sentiment Analysis

- Analyzing IPL T20 Cricket

- Movie Recommendation

- Twitter API Integration for Tweet Analysis

- Data Exploration Using Spark SQL – Wikipedia Data Set

Tableau Desktop 10

- Understanding the global covid-19 mortality rates

- Understand the UK bank customer data

- Understand Financial Data

- Understand Agriculture Data

Data Science with SAS

- Categorization of Patients Based on the Count of Drugs for Their Therapy

- Build Revenue Projections Reports

- Impact of Pre-paid Plans on the Preferences of Investors

- k-means cluster Analysis on Iris Dataset

Certification

After the completion of the course, you will get a certificate from IBM.

You will also get a certificate from Intellipaat.

Intellipaat Success Stories

Vishal Pentakota

The best part of this online course is the series of hands-on demonstrations the trainer performed. Not only did he explain each concept theoretically but he also implemented all those concepts practically. Great job! A must go for beginners.

Shreyashkumar Limbhetwala

I want to talk about the rich LMS that Intellipaat Data Science training offered. The extensive set of PPTs, PDFs, and other related course material were of the highest quality, and due to this my learning with Intellipaat was excellent. I could clear the Cloudera Data Scientist certification in the first attempt.

Giri Karnal

I had taken the Data Science master's program which is a combo of SAS, R, and Apache Mahout. Since there are so many technologies involved in the Data Science course, getting your query resolved at the right time becomes the most important aspect. But with Intellipaat, there was no such problem as all my queries were resolved in less than 24 hours.

Sharath Reddy Yellapati

The course material was very well organized. The trainer explained each module to me from the basics. All my queries were addressed very clearly. The trainer also made me realize how important this course is for beginners in the IT stream.

CONTACT US

INTELLIPAAT SOFTWARE SOLUTIONS PVT. LTD.

Bangalore

AMR Tech Park 3, Ground Floor, Tower B,
Hongasandra Village, Bommanahalli,
Hosur Road, Bangalore – 560068

USA

1219 E. Hillsdale Blvd. Suite 205,
Foster City, CA 94404

If you have any further queries or just want to have a conversation with us, then do call us.

IND: +91-7022374614 | US: 1-800-216-8930