






## About IntelliPaat


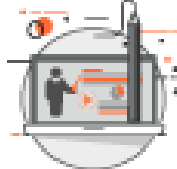
IntelliPaat is a fast growing professional training provider that is offering training in over 150 most sought-after tools and technologies. We have a learner base of 700,000 in over 32 countries and growing. For job assistance and placement we have direct tie-ups with 80+ MNCs.

### Key Features of IntelliPaat Training :

 <b>24X7</b>				
<b>Life Time Support and Assistance</b>	<b>Real Time Projects</b>	<b>Life Time Access and Free Upgrade</b>	<b>Job Assistance</b>	<b>Industry Recognised Certification</b>

## About the Course

This IntelliPaat Data Science specialization training course provides you detailed learning in data science, data analytics, project life cycle, data acquisition, analysis, statistical methods and machine learning. You will gain expertise to deploy Recommenders using Apache Mahout, data analysis, data transformation, experimentation and evaluation.

	<b>Instructor Led</b> <b>Duration – 36 Hrs</b> <b>Weekend Batch – 3 Hrs/Session</b>		<b>Self paced</b> <b>Duration – 28 Hrs</b>
---	---	---	---

## Why Take This Course ?

There is a serious shortage of Data Scientists and this is a major concern for Top MNCs around the world. All this means the major corporations are ready to pay top dollar salaries for professionals with the right Data Science skills.

- ❖ Data Scientist is the best job of the 21st century - Harvard Business Review
- ❖ Global Big Data market to reach \$122B in revenue by 2025 – Frost & Sullivan
- ❖ The US alone could face a shortage of 1.4 -1.9 million Big Data Analysts by 2018 – Mckinsey



## Course Contents

<p><b>Introduction to R</b></p> <ul style="list-style-type: none"> <li>❖ Installing R and R-Studio</li> <li>❖ R packages</li> <li>❖ R Operators</li> <li>❖ if statements and loops (for, while, repeat, break, next),</li> <li>❖ switch case</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ R studio &amp; its environment Introduction</li> <li>❖ Basic Operations</li> <li>❖ Conditional Statments</li> <li>❖ Control flow statments</li> </ul>
<p><b>Data Exploration, Data Wrangling and R Data Structure</b></p> <ul style="list-style-type: none"> <li>❖ Importing and Exporting data from external source</li> <li>❖ Data exploratory analysis</li> <li>❖ R Data Structure (Vector, Scalar, Matrices, Array, Data frame, List)</li> <li>❖ Functions</li> <li>❖ Apply Functions</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Data Types in R</li> <li>❖ Data Structures in R (vector, list, complex, data frame etc)</li> <li>❖ Importing files (text, csv, db, json)</li> <li>❖ Functions in R (build-in &amp; custom functions)</li> <li>❖ Data Wrangling (subset, NA impuation, sampling data into test and train, Normalization)</li> </ul>
<p><b>Data Visualization</b></p> <ul style="list-style-type: none"> <li>❖ Bar Graph (Simple, Grouped, Stacked)</li> <li>❖ Histogram, Pi Chart</li> <li>❖ Line Chart</li> <li>❖ Box (Whisker) Plot, Scatter Plot, Correlogram</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Visualization importance</li> <li>❖ Graphs with base pacakge: Bar Graph, Line Graph, box plot, scatter plot</li> <li>❖ ggplot2 package: Bar graph, linegraph, box plot, scatter plot, color &amp; shape importance</li> </ul>
<p><b>Introduction to Statistics</b></p> <ul style="list-style-type: none"> <li>❖ Terminologies of Statistics</li> <li>❖ Measures of Centers</li> <li>❖ Measures of Spread</li> <li>❖ Probability</li> <li>❖ Normal Distribution</li> <li>❖ Binary Distribution</li> <li>❖ Hypothesis Testing</li> <li>❖ Chi Square Test</li> <li>❖ ANOVA</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Hypothesis Testing(z-score, t-test, chi-square test) using R</li> <li>❖ Understand Mean, Median, Mode, variance &amp; standard devication in R</li> <li>❖ Learn probability, conditional probablility, Confiusion Matrix, bayes Theorem.</li> </ul>
<p><b>Predictive Modeling - 1</b></p> <ul style="list-style-type: none"> <li>❖ Supervised Learning - Linear Regression</li> <li>❖ Bivariate Regression</li> <li>❖ Multiple Regression Analysis</li> <li>❖ Correlation( Positive, negative and neutral)</li> <li>❖ Industrial Case Study</li> <li>❖ Machine Learning Use-Cases</li> <li>❖ Machine Learning Process Flow</li> <li>❖ Machine Learning Categories</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Data preprocessing</li> <li>❖ Build a Linear Regression model using R</li> <li>❖ Understand the results and evaluation Metrics (RMSE, MAPE)</li> </ul>

<p><b>Predictive Modeling - 2</b></p> <ul style="list-style-type: none"> <li>❖ Logistic Regression</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ What is Linear &amp; Non-Linear Data</li> <li>❖ Binary class</li> <li>❖ Multiclass classification</li> <li>❖ Create logistic model, Prediction, importance of Threshold</li> <li>❖ Evaluation metrics for logistic, Accuracy, Precision</li> <li>❖ ROC curve analysis</li> <li>❖ Kappa metric</li> </ul>
<p><b>Decision Trees</b></p> <ul style="list-style-type: none"> <li>❖ What is Classification and its use cases?</li> <li>❖ What is Decision Tree?</li> <li>❖ Understand Root and leaf node of Decision Trees</li> <li>❖ Algorithm for Decision Tree Induction</li> <li>❖ Understand Information Gain &amp; Entropy in Decision Trees &amp; their importance</li> <li>❖ rpart&amp; C5.0 Decision Tree Algorithms</li> <li>❖ Accuracy, Precision, Recall, F1 Score of classifier</li> <li>❖ Creating a Perfect Decision Tree</li> <li>❖ Confusion Matrix</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Entropy, Information Gain</li> <li>❖ rpart&amp; C5.0 Decision Tree Algorithms</li> <li>❖ Accuracy, Precision, Recall, F1 Score of classifier</li> </ul>
<p><b>Random Forest</b></p> <ul style="list-style-type: none"> <li>❖ Benefits of building Random forest vs a decision tree</li> <li>❖ Random Forest for feature importance and other usage.</li> <li>❖ Understand the Naive Bayes Classifier</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Understanding Random forest with an example</li> </ul>
<p><b>Unsupervised learning</b></p> <ul style="list-style-type: none"> <li>❖ What is Clustering &amp; its Use Cases</li> <li>❖ What is K-means Clustering</li> <li>❖ What is Canopy Clustering</li> <li>❖ What is Hierarchical Clustering</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Clusteing(Aglomerative or divisive)</li> <li>❖ K-means clustering and graphs</li> <li>❖ Hierarchical clustering</li> </ul>
<p><b>Association Analysis and Recommendation engine</b></p> <ul style="list-style-type: none"> <li>❖ Learn Market basket Analysis in order to Analyse the purchase patterns of various users of an e-commerce.</li> <li>❖ Learn how to Build Association Rules to increase the Sales</li> <li>❖ Recommendation system: With the purchase pattern of a customer recommend the product to another customer who blongs to a similar group.</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Building a recommendation engine using R</li> </ul>

<p><b>Sentiment Analysis</b></p> <ul style="list-style-type: none"> <li>❖ Introduction to Token, Lexicon, Parts of Speech Tagging (POS) &amp; Text corpus</li> <li>❖ Processing the Text corpus to drive insights (eg: Pre-release or post release sentiment of iPhone X)</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ R code to get twites using Twitter API</li> <li>❖ Sentiment Analysis of twites for either Positive sentiment or Negative sentiment.</li> </ul>
<p><b>Time Series</b></p> <ul style="list-style-type: none"> <li>❖ What is Time Series data</li> <li>❖ Time Series variables</li> <li>❖ Different components of Time Series data</li> <li>❖ Visualize the data to identify Time Series Components</li> <li>❖ Implement ARIMA model for forecasting</li> <li>❖ Exponential smoothing models</li> <li>❖ Identifying different time series scenario based on which different Exponential Smoothing model can be applied</li> <li>❖ Implement respective ETS model for forecasting</li> </ul>	<p><b>Hands on Exercises</b></p> <ul style="list-style-type: none"> <li>❖ Real Time hands on with stock price data of any organization.</li> <li>❖ Understand the Components of Time Series (Trend, Randomness, Sesionality)</li> <li>❖ Build an ARIMA Model to predict the future stock price.</li> </ul>

## Data Science Projects

---

### Project 1 – Understanding Cold Start Problem in Data Science

**Topics:** This project involves understanding of the cold start problem associated with the recommender systems. You will gain hands-on experience in information filtering, working on systems with zero historical data to refer to, as in the case of launching a new product. You will gain proficiency in working with personalized applications like movies, books, songs, news and such other recommendations. This project includes the following:

- ❖ Algorithms for Recommender
- ❖ Ways of Recommendation
- ❖ Types of Recommendation -Collaborative Filtering Based Recommendation, Content-Based
- ❖ Complete mastery in working with the Cold Start Problem.

### Project 2 – Recommendation for Movie, Summary

**Topics:** This is real world project that gives you hands-on experience in working with a movie recommender system. Depending on what movies are liked by a particular user, you will be in a position to provide data-driven recommendations. This project involves understanding recommender systems, information filtering, predicting 'rating', learning about user 'preference' and so on. You will exclusively work on data related to user details, movie details and others. The main components of the project include the following:

- ❖ Recommendation for movie
- ❖ Two Types of Predictions – Rating Prediction, Item Prediction
- ❖ Important Approaches: Memory Based and Model-Based

- ❖ Knowing User Based Methods in K-Nearest Neighbor
- ❖ Understanding Item Based Method
- ❖ Matrix Factorization
- ❖ Decomposition of Singular Value
- ❖ Data Science Project discussion
- ❖ Collaboration Filtering
- ❖ Business Variables Overview

## What makes us who we are

---



RiteshBhagwat

“With over a decade of experience as an Oracle Business Intelligence Architect, I was able to pursue my dream career by mastering Data Science and now there’s no stopping me” ....[Read More!](#)

[VIEW ALL SUCCESS STORIES](#)

[READ ALL REVIEWS](#)