IMARTICUS LEARNING

DATA SCIENCE PRODEGREE

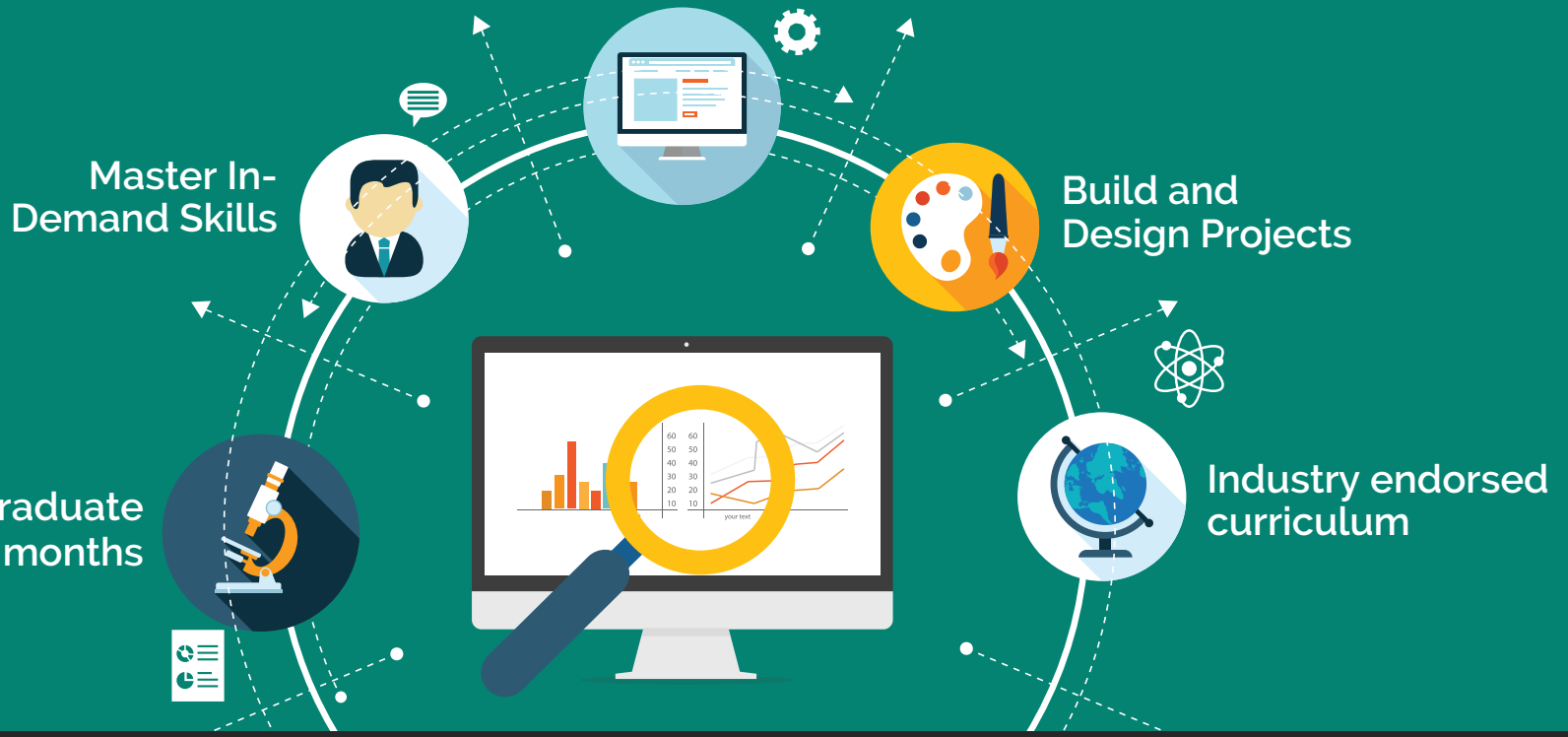
Knowledge Partner:

GENPACT

GENERATING IMPACT℠

Global Leader in Digitally-Powered Business Process Management & Services

INDUSTRY CREDENTIALS TO ACCELERATE CAREERS



DATA SCIENCE PRODEGREE

Knowledge Partner

GENPACT

GENERATING IMPACT℠

COMPREHENSIVE COVERAGE

The Data Scientist Prodegree is a 200-hour program that delivers a deep understanding of Data Analysis and Statistics, along with business perspectives and cutting-edge practices using SAS, R, Python, Hive and Spark.

ENDORSED BY GENPACT

The program is co-created with Genpact as the Knowledge Partner and comes with a cutting edge industry-aligned curriculum that is co-created with Genpact.

PROJECT BASED LEARNING

You will spend approximately 100 hours of this program getting hands-on with industry projects and build a portfolio of demonstrable work.

PROGRAM DELIVERY

The Prodegree is delivered in two modes: Classroom and Online (Live Virtual Classes) to cater to your learning preferences while ensuring maximum learning efficacy.

CAREER READINESS

We prepare you to be job-ready with interview preparation, resume building workshops and 1-1 mock interviews with industry experts.

INNOVATIVE LEARNING METHODOLOGY

Self Paced Instructor Videos

Active, self-paced, data-driven learning through HD videos

68 Hours

TWO DELIVERY MODES TO CHOOSE FROM

Classroom Delivery

Classroom training by expert faculty with industry credentials at our Imarticus centers

132 Hours

Online Delivery

Live Instructor-led Virtual Classes with expert faculty for real-time learning for learning as per your convenience

132 Hours

WHY A PRODEGREE?

Genpact Endorsed
Cutting edge industry-aligned curriculum and learning methodology vetted by Genpact

Experiential Learning
Delivery using combination of self-paced videos, live lectures, case studies, project work and simulations

Guest Lectures
Periodic guest lectures by Genpact Leadership on industry best practices

Customized Learning
Flexible learning paths to cater to both novices and experienced professionals

Verified Certification
Earn meaningful certifications validated by rigorous coach evaluations

- VERIFIED PRODEGREE CREDENTIAL
- JOB READY PROJECT PORTFOLIO
- EXPERIENTIAL LEARNING AT COMPETITIVE PRICE
- FLEXIBLE LEARNING PATHS

PLACEMENT AT GLOBALLY REPUTED COMPANIES

CURRICULUM

Semester 1: Base SAS — 48 Hours

Base SAS 24 Hours	Data Structures | Data Types | Type Casting | Data Import/Export | Loops & Functions | Procedures
Understanding Data 8 Hours	Definition of Data | Meaning of Variable | Understanding Data Types | Measures of Central Tendency in Data | Understanding Skewness in Data | Measures of Dispersion | Understanding Data Distribution
SAS - Store Data Analytics 8 Hours	Extracting Retail Data into SAS | Cleaning the Data | Transforming the Data | Loading the Data for Reporting and Analysis | Reporting and Presenting the Data
Project 1 Submission 8 Hours	**Project 1 - Store Data Analytics** Key Learning Outcomes - ETL, Analysis and Reporting using SAS

Semester 2: R Programming — 78 Hours

R Basics 32 Hours	Data Structures | Data Types | Data Import/Export | Conditional Statements | Loops & Functions
Business Case: Managing Credit Risk 1 Hours	Meaning of Credit Risk | Impact of Credit Default | Sources of Data for Managing Risk | Understanding Loss given Default | Understanding Default
Linear Regression 4 Hours	Covariance and Correlation in Data | Multivariate Analysis | Assumptions of Linearity | Hypothesis Testing | Limitations of Regression
Case Study on Loss Given Default using Linear Regression using R 4 Hours	Extract the Data in R | Univariate Analysis of Data | Apply Data Transformations | Bivariate Analysis | Multicollinearity in Data | Identify Hetroscedasticity | Modelling of Data | Model Significance Test | Build Model on Training Data Set
Logistic Regression 4 Hours	Reason for using Logistic Regression | The Logistic Transform | Logistic Regression Modelling | Model Optimization | Understanding the ROC Curve
Case Study on Default Modelling using Logistic Regression using R 8 Hours	Extract the Data in R | Univariate Analysis of Data | Apply Data Transformations | Bivariate Analysis | Identify Multicollinearity in Data | Identify Hetroscedasticity | Variable Significance Identification | Model Significance Test | Validate the Model Performance | Drawing the ROC Curve | Estimating the Classification Model Hit Ratio | Isolating the Classifier for Optimum Results
Support Vector Machines 4 Hours	Introduction to SVM | Classification as a Hyper Plane Location Problem | Motivation for Linear Support Vectors | SVM as a Quadriatic Optimization Problem | Non Linear SVMs | Introduction to Kernel Functions
Case Study on Default Modelling using Suppport Vector Machines using R 6 Hours	Build the Model | Training the Model | Testing and Validation | Tuning the Model
Business Case: Intrusion in IT Network 1 Hours	Meaning of Intrusion in IT Network | Cost of Intrusion | Meaning of Intrusion Detection System
Decision Tree & Ensemble Learning 8 Hours	Theory of Entropy & Information Gain | Stopping Rules | Overfitting Problem | Cross Validations for Overfittinig Problem | Prunning as a Solution for Overfitting | Ensemble Learning Notion | Concept of Bootstrap Aggregation | Concept of Random Forest

CURRICULUM

Case Study on Network Intrusion Detection using Decision Trees & Ensemble Learning using R 6 Hours	Validate the Model Performance Perform Cross Validations | Select the Best Split | Prune the Tree | Predict and Validate Performance of Model
Project 2 and 3 Submission 2 Hours	**Project 2 Credit Risk Analytics** Key Learning Outcomes - Linear and Logistics Regression, Support Vector Machines using R **Project 3 - Network Intrusion Detection** Key Learning Outcomes - Decision Trees & Ensemble using R

Semester - 3: Python — 28 Hours

Data Structures in Python Used for Data Analysis 8 Hours	Data Structures | Data Types | Data Import/Export | Numpy/Pandas | Loops & Functions
Case Study on Default Modelling using Logistic Regression using Python 8 Hours	Univariate Analysis of Data | Apply Data Transformations | Bivariate Analysis | Build Model on Training Data Set | Predict using Testing Data Set | Validate the Model Performance
Case Study on Default Modelling using Suppport Vector Machines using Python 4 Hours	Build the Model | Training the Model | Testing and Validation | Tuning the Model
Case Study on Network Intrusion Detection using Decision Trees & Ensemble Learning using Python 6 Hours	Extract the Data in Python | Convert Data into Array | Standardize the Independent | Variables | Predict using Testing Data Set | Validate the Model Performance Perform Cross Validations | Select the Best Split | Prune the Tree | Predict and Validate | Performance of Model | Improve Model Performance using Bagging and Random Forest
Project 2 and 3 Submission 2 Hours	**Project 2 - Credit Risk Analytics** Key Learning Outcomes - Logistics Regression, Support Vector & Machines using Python **Project 3 - Network Intrusion Detection** Key Learning Outcomes - Decision Trees & Ensemble using Python

Semester 4 : Hive, Spark , Tableau — 46 Hours

HIVE Basics 5 Hours	Set Up | Data Loading in HDFS | Creating HIVE Tables | Creating ORC File | Querying HIVE data
Business Case: Pricing Analytics 1 Hours	Understanding Pricing as a Business Function | Importance of Pricing Analytics | Sources of Pricing Data
Case Study on Pricing Analytics using HIVE 8 Hours	Load and Extract Data into HDFS | Transform Data in Hive | Load Data in Hive Tables | Write Queries Prepare Analysis Output | Show the Reports in Zepplin

CURRICULUM

SPARK Basics 5 Hours	Set Up | Creating a RDD | Querying a RDD
Business Case: Understanding Customer Life Cycle 1 Hour	Understanding Customer Life Cycle - Acquisition/Consumption/Saturation/Churn | Impact of Customer Churn | Churn Identification
Case Study on Telecom Customer Churn using Spark 8 Hours	Case Study on Telecom Customer Churn using Spark | Extract Data in Spark | Perform a Logistic | Stochastic Gradient Descent Model | Predict using Testing Data Set | Validate the Model Performance
Tableau Basics 8 Hours	Introduction to Visualization | Working with Tableau | Data Organization | Advanced Visualization | Mapping | Enterprise Dashboards | Data Presentation
Interview and Resume Preparation 4 Hours	Analytics Experts Guide You on How to Prepare for Technical Interview Round with Tips, Tricks and a Quick Refresher on Concepts Learnt
Mock Interviews - HR and Domain 4 Hours	1:1 or Panel Mock Interviews with Industry Veterans to Clear the HR and Technical Round of Interviews to Give You Confidence to Face Real World Scenarios
Project 4, 5 and 6 Submission 2 Hours	**Project 4 - Pricing Analytics** Key Learning Outcomes - Big Data Analytics and Reporting using HIVE **Project 5 - Telecom Churn Analytics** Key Learning Outcomes - In Memory Big Data and Logistics Regression using Spark **Project 6 - Reporting** Key Learning Outcomes -Visualization for Structured and Unstructured Data using Tableau

HANDS-ON PROJECTS

Students get hands-on with industry projects and build a portfolio of demonstrable work

Store Data Analytics & Reporting

Credit Risk Analytics

Network Intrusion Detection Analytics

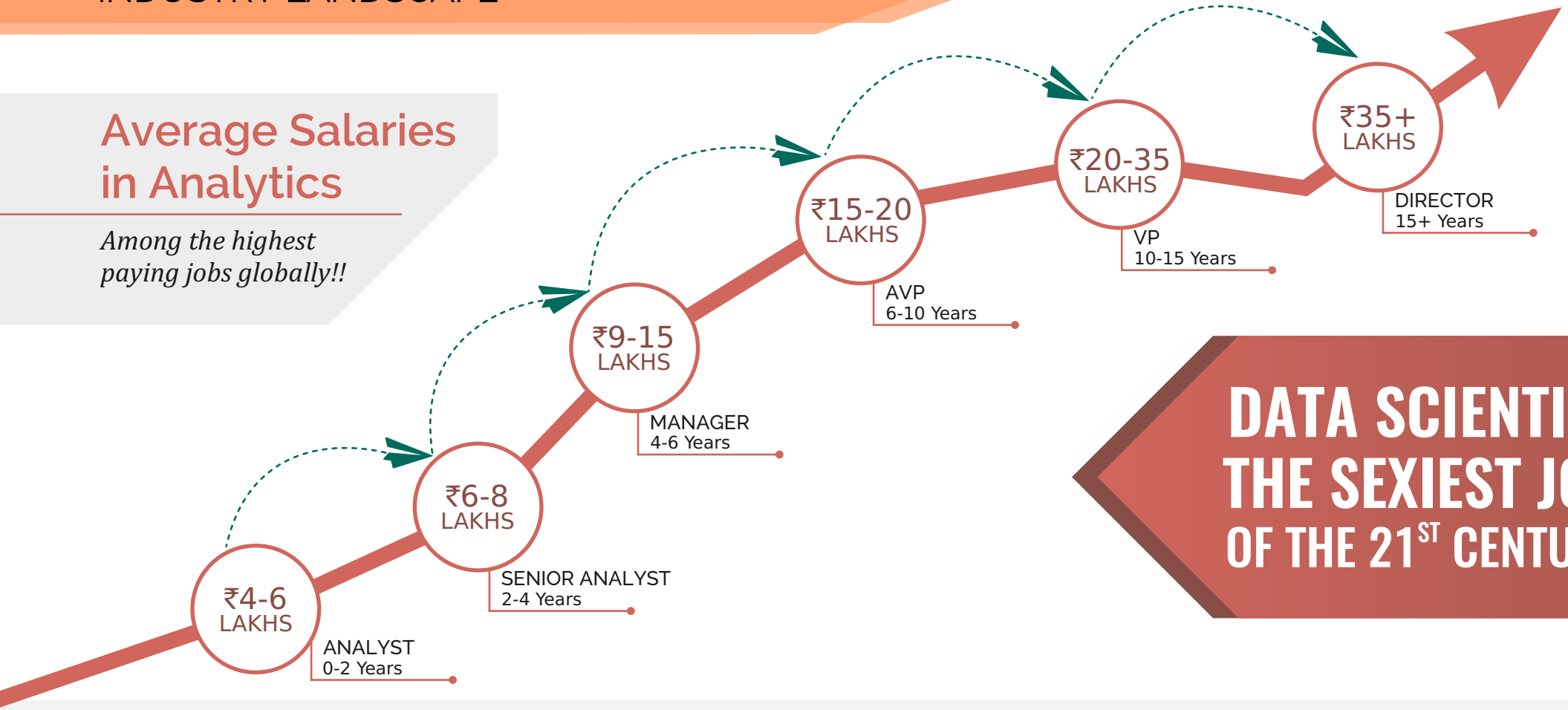
Pricing Analytics in Bullion/Commodity Market

Telecom Churn Analytics

Visualizing Vanilla, Analytical, Un/Structured Data

Average Salaries in Analytics

Among the highest paying jobs globally!!



DATA SCIENTIST THE SEXIEST JOB OF THE 21ST CENTURY

TOP COMPANIES HIRING FOR BUSINESS ANALYTICS

amazon · GENPACT · Flipkart · Google · VISA · Mu Sigma · Microsoft · airtel · Bank of America Merrill Lynch · mozilla Firefox

COLLABORATION WITH GENPACT

The Data Science prodegree is co-created with Genpact as the Knowledge Partner and comes with a cutting edge industry-aligned curriculum and learning methodology. You will benefit in terms of:

PROJECT EVALUATION

At the end of each semester, Genpact will evaluate project submissions and provide constructive feedback.

GUEST LECTURES & MENTORSHIP

Periodic guest lectures by Genpact management on key trends and real-world challenges plaguing the industry.

INDUSTRY-APPROVED CURRICULUM

You learn in-demand skills, sought after tools and techniques which are required by the Industry.

INDUSTRY SPEAK

Sidhartha Shishoo
Business HR Leader, Genpact

"There is a huge talent gap in Analytics in India. Our approach has been to work closely with Imarticus as our training partner. They bring expertise in training delivery, while Genpact brings in industry exposure. We are happy to support this initiative through guest lectures, projects, data sharing etc. The Prodegree also helps the creation of a lateral talent pool for analytics companies."

GENPACT
GENERATING IMPACT℠

ABOUT GENPACT

Genpact is a global leader in digitally-powered business process management and services across technology, analytics, and organizational design. The company boasts net revenues of US$2.46 billion with more than 70,000 employees spread across 25 countries and 1/5th of the Fortune Global 500 companies as its clients.

INDICATIVE FACULTY

Mohan Rai

Mohan has 10 years of experience in Core Analytics (Sales & IT). His key domain expertise lies in domains ranging from Banking, Construction, Real Estate, Automobile, Component Manufacturing and Retail. He has also worked with verticals ranging from Training, Research, Sales, Market Research, Business Consulting, Sales Planning, Market Strategy and IT. Mohan is a Director for S & R Analytics involved in Delivery of Analytics Consulting/Training and SIP Partners of TCS. He is also a visiting Faculty for Analytics at various Colleges and Institutes. Mohan holds degrees in Business Analytics and Intelligence from IIM-Bangalore, MBA in Marketing and BSC in Statistics.

Arun Upadhyay

Arun has over 14 years experience in Information Technology and has conducted SAS training for Infosys, Wipro, IBM, Genpact, ICICI Bank, Reliance Mutual Fund. Arun is a certified, accredited IT professional who has successfully trained more than 10,000 students in different technologies like SAS and R. He has previously worked as a trainer for companies such as Aptech, NIIT, Ultramax Infonet Education Pvt. Ltd., and Vistaar Systems Pvt. Ltd. He has cleared many Microsoft international certifications such as MCAD, MCPD, MCTS etc. and is also a Microsoft-certified trainer.

Roney Joseph

Roney is a Big Data and Hadoop trainer, mentor & consultant having 20+ years of experience in the field of Information Technology. He has been conducting training sessions on Big Data and Hadoop for corporates such as the Virtue Group since 2014 when Hadoop was in its infancy. He has set up offshore development teams for large IT organizations, systems and procedures and managed multiple accounts across different geographical locations and successfully led software development and testing projects on various platforms. He is the Founder of Xillon Infotech, which offers training and consulting on Big Data technologies. He is certified by IBM in Big Data Fundamentals.

Satya Srinivas

Satya has 25 years of experience aligning multi-million dollar Information Technology deployments with business strategy and operational processes for Fortune 1000 companies. In the past he has been a Management Consultant & a negotiator and has consulted in the areas of performance management in enterprise architecture, data mining & analytics, machine learning, pattern recognition, social media analytics and big data management and analytics for several start-ups as well as major corporate houses like Infosys and IBM. Satya is a BE, Electronics and Communication from University of Mysore and a MS, Computer Engineering from Florida Atlantic University.

and many more...

1 Resume Building

Refining and polishing the candidate's resume with insider tips to help them land their dream job

2 Interview Prep

Preparing candidates to ace HR and Technical interview rounds with model interview Q&A

3 Career Guidance

Doubt clearing about the industry and career opportunities

THIS PROGRAM IS IDEAL FOR

Recent Post Graduates

Bachelors or Masters in science, math, statistics or computer applications/IT

Experienced Professionals in Programming or Related Fields

With Less than Four Years of Experience

Individuals looking for Global Certifications

To enhance their Resume

Awards & Alliances

Leading Institute in SAS & Python Analytics Vidhya

Excellence Awards

National Awards in IT Excellence

TOP 4 Analytics Training Institute in India

N·S·D·C National Skill Development Corporation
Transforming the skill landscape

#1 Big Data Training Institute 2016

IMARTICUS LEARNING

Toll Free: 1800-267-7679

www.imarticus.org | elearning@imarticus.org