



IMARTICUS
LEARNING

BIG DATA & MACHINE LEARNING PRODEGREE

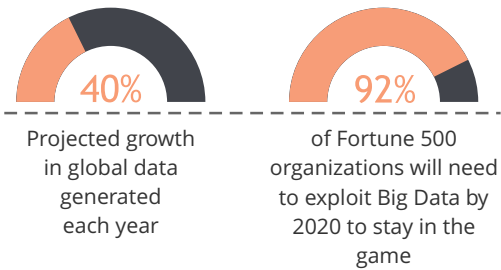
Imarticus Learning is an
EdTech Partner of:



INDUSTRY LANDSCAPE

Machine learning marks a major technological breakthrough in the field of computer science, big data and artificial intelligence. Machine learning has seen a phenomenal rise in terms of industry application over the last year because it is ideal for exploiting the opportunities hidden in Big Data.

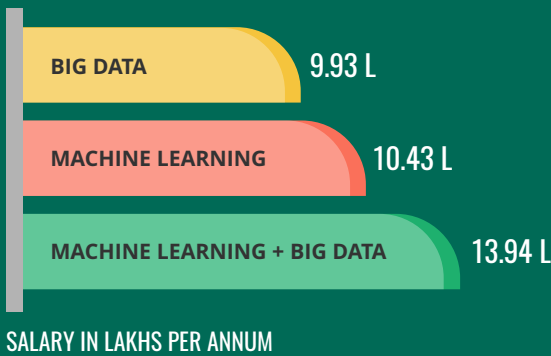
THE SKILLS GAP



IN DEMAND SKILL SETS

- Distributed Computing
- Predictive Modeling
- Math, Stats
- Machine Learning
- Storytelling
- Deep Learning

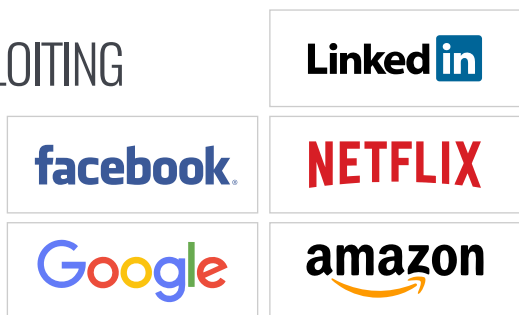
GROWING DEMAND



DATA SCIENTIST:
THE
SEXIEST
JOB IN THE
21ST CENTURY

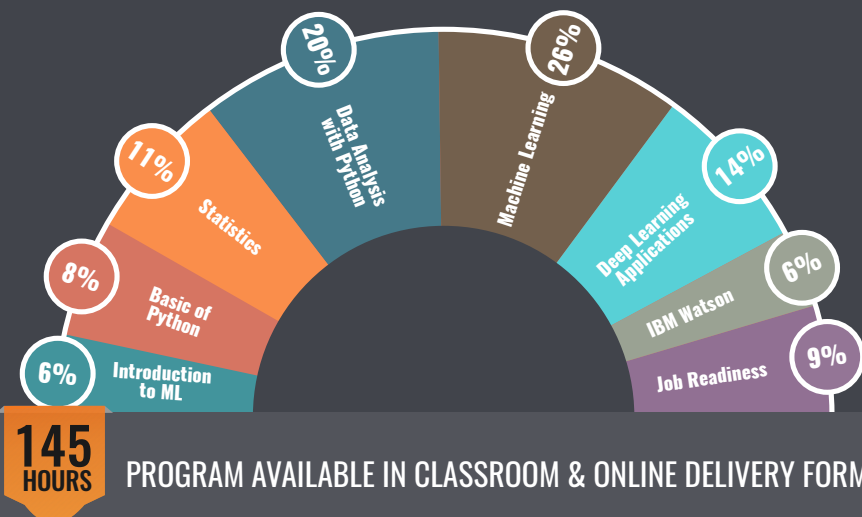
HARVARD BUSINESS REVIEW, OCT 2012

TOP COMPANIES EXPLOITING BIG DATA AND MACHINE LEARNING



OVERVIEW OF PROGRAM

The Big Data and Machine Learning Prodegree, in association with IBM as the EdTech Partner, is a first-of-its-kind 145-hour program providing in-depth exposure to Data Science, Machine Learning and Deep Learning through a rigorous industry-aligned curriculum featuring Python and IBM Watson.



CURRICULUM

LEVEL 1. INTRODUCTION TO STATISTICS

INTRODUCTION TO MACHINE LEARNING AND BIG DATA

ML Spectrum & Journey | Intro to Modeling Lifecycle | Intro to Supervised Learning | Intro to Unsupervised Learning | Big Data and Hadoop | Big Data and its Sources | Popular Tools Used for Big Data | RDBMS vs Hadoop | Hadoop Architecture and Ecosystem | HDFS Design and Architecture Overview | When to Use & Not Use Hadoop?

BASIC PROBABILITY AND TERMS

Rules of Probability | Permutations and Combinations | Bayes Theorem | Descriptive Statistics | Compound Probability | Conditional Probability

PROBABILITY DISTRIBUTIONS

Types of Distributions | Functions of Random Variables | Probability Distribution Graphs | Confidence Intervals

DATA TRANSFORMATIONS

Merge, Rollup, Transpose and Append | Missing Analysis and Treatment | Outlier Analysis and Treatment

EXPLORATORY DATA ANALYSIS

Summarizing and Visualizing the Important Characteristics of Data Hypothesis Testing | Visualizations | Univariates, Bivariates | Crosstabs, Correlation

LEVEL 2. BASICS OF PYTHON FOR MACHINE LEARNING

INTRODUCTION TO PYTHON

Python Basics | Spyder IDE | Jupyter Notebook | Floats and Strings | Simple Input & Output | Variables | Single and Multiline Comments

CONTROL STRUCTURES

Booleans and Comparisons | Conditional Statements (IF ELSE) Operator Precedence | Lists - Operations and Functions

FUNCTIONS AND MODULES

Function Arguments | Comments and doc strings | Functions as Objects | Modules | Standard Lib and Pip

EXCEPTIONS AND FILES

Exception Handling | Raising Exceptions | Assertions | Working With Files

LEVEL 3. DATA ANALYSIS WITH PYTHON FOR MACHINE LEARNING

PANDAS

Introduction to Pandas | IO Tools | Basics of NumPy | NumPy Functions | Pandas - Series and Dataframes

DATA VISUALIZATION

Basics of Data Visualization | Line Plots | Bar Charts | Pie Charts | Histograms | Scatter Plots | Parallel Coordinates

LOGISTIC REGRESSION

Implementing Logistic Regression with Python | Making Sense of Result Parameters: Wald Test, Likelihood Ratio Test Statistic, Chi-Square Test | Goodness of Fit Measures | Model Validation: Cross Validation, ROC Curve, Confusion Matrix | Use Cases

DECISION TREES

Implementing Decision Trees using Python | Homogeneity | Entropy | Information Gain | Gini Index | Standard Deviation Reduction | Visualizing & Pruning a Tree

TIME SERIES

Handling Time Series Data | Holt-Winters Model | ARIMA Model | ACF/PACF Functions

PROJECT 1

LINEAR REGRESSION

► Property Price Prediction Using Linear Regression

PROJECT 2

LOGISTIC REGRESSION

► Bankruptcy Prediction Using Logistic Regression

PROJECT 3

DECISION TREES

► Facebook Post Count Using Decision Tree

PROJECT 4

TIME SERIES

► Forecasting and Predicting the Sales of Furniture of the Superstore

LEVEL 4. INTRODUCTION TO MACHINE LEARNING

INTRODUCTION TO MACHINE LEARNING

Introduction to Machine Learning | Machine Learning Modelling Flow | How to Treat Data in ML | Parametric & Non-Parametric ML Algorithm | Types of Machine Learning | Performance Measures | Bias-Variance Trade-Off | Overfitting & Underfitting | Bootstrap Sampling | Bagging Aggregation | Boosting

SCIKIT LEARN

Introduction to SciKit Learn | Load Data into Scikit Learn | Run Machine Learning Algorithms Both for Unsupervised and Supervised Data | Supervised Methods: Classification & Regression | Unsupervised Methods: Clustering, Gaussian Mixture Models | Decide What's the Best Model for Every Scenario

OPTIMISATION TECHNIQUES

Constant Learning Rate Procedures | Adaptive Learning Procedures | Batch Gradient Descent | Mini-Batch Gradient Descent | Stochastic Gradient Descent | Nesterov Accelerated Gradient | Root Mean Squared Propagation | Adaptive Moment Estimation Procedure

ML ALGORITHM - SUPERVISED LEARNING AND UNSUPERVISED LEARNING

Linear Regression with Stochastic Gradient Descent | Logistic Regression with Stochastic Gradient Descent | K-Nearest Neighbour | Eager Methods vs. Lazy Methods | Nearest Neighbor Classification | Building kD-Trees | Support Vector Machine | Perceptron Algorithm. What is Clustering? | K-means Algorithm | Types of Clustering | Evaluating K-means Clusters

ENSEMBLE ALGORITHMS

Ensemble Techniques | Bootstrap Aggregation | Random Forest | Boosting

NEURAL NETWORKS

Neural Networks | The Biological Inspiration Perceptron Learning & Binary Classification | Backpropagation Learning | Learning Feature Vectors for Words | Object Recognition

LEVEL 5. INTRODUCTION TO DEEP LEARNING

KERAS

Keras for Classification and Regression in Typical Data Science Problems | Setting up KERAS | Different Layers in KERAS | Creating a Neural Network | Training Models and Monitoring Artificial Neural Networks

PROJECT 5

ANN ON KERAS

► Credit Default Using ANN on Keras

TENSORFLOW

Introducing Tensorflow | Neural Networks using Tensorflow Debugging and Monitoring | Convolutional Neural Networks | Unsupervised Learning

PROJECT 6

CNN ON TENSORFLOW

► Handwriting/Facial Recognition Using CNN on TensorFlow

RNN

Introducing Recurrent Neural Network | Application Areas | Case Study

LEVEL 6. JOB READINESS

RESUME WRITING

The Why, the What and the How of Resumes | Personal Branding Tips and Resources | Interview Skills | Using Social Media | CV Discussion

MOCK INTERVIEWS - DOMAIN

1:1 or Panel Mock Interviews with Faculty to Clear the Technical Round of Interviews to Give You Confidence to Face Real World Scenarios

CAPSTONE PROJECT

Students Present their Capstone Project in Front of their Peers and Industry Experts, who will Provide Constructive Feedback and Critique

Capstone Project in Deep Learning/Machine Learning

- Predicting Purchase Behaviour on E-Commerce Dataset
- Goal: Use the data of GroceryKart customer orders over time to predict which previously purchased products will be in a user's next order.
- Using multiple data sets, users are to use ML algorithms to determine:
 - 1) When do customers order the most?
 - 2) What are the top 5 products that are reordered?
 - 3) What is the reorder ratio for each department?
 - 4) Build a model to predict which previously purchased products will be in a user's next order.

KEY HIGHLIGHTS



COMPREHENSIVE COVERAGE

The Prodegree features 145 hours of impactful, supervised and unsupervised learning which forms the core of Machine Learning. Aspirants also learn about Data Science, Machine Learning and Deep Learning through hands-on practice on tools such as Python and IBM Watson.



COLLABORATION WITH IBM

The Big Data and Machine Learning Prodegree is co-created with IBM as the EdTech Partner, who are at the forefront of technology innovation.



PROJECT BASED LEARNING

The program provides an edge through our unique project-based methodology, focusing on real life projects. The Prodegree features seven projects covering tools such as Python and IBM Watson and teaches you how to apply predictive models to massive dataset typically found in healthcare, financial services, social media and many more!



VIRTUAL LABS AND CODING PLATFORM

Learn on a state-of-the-art virtual lab, with 24/7 access to all required software and datasets pre-installed. Agnostic of machine configuration, with no installation and compatibility issues, learn anytime, anywhere!



JOB READINESS

The Imarticus Career Assistance Services (CAS) team prepares you to be job-ready through extensive interview prep, resume building & mock interviews.



MENTORSHIP

Industry experts from leading companies act as your personal mentor to advise and guide you in your journey towards job-readiness.

TWO DELIVERY MODES TO CHOOSE FROM: CLASSROOM DELIVERY OR ONLINE DELIVERY

CASE STUDIES

Build valuable hands-on development experience which can be showcased to future recruiters.

MODELS	DATASET	TECHNIQUE
Linear Regression	Boston Dataset	Sklearn Linear Model & Gradient Descent Model
Logistic Regression	Iris Dataset	Sklearn Logistic Model & Stochastic Average Gradient Descent
Decision Tree & Random Forest	Bank Marketing Dataset	Decision Tree Classifier, Random Forest Classifier, Adaboost Classifier & Bagging Classifier
KNN	Breast Cancer Dataset	KNN Classifier & K Value
SVM	Default Of Credit Card Clients Dataset	SVM Classifier Using Different Kernels
K-Means	Cars dataset	K-Means Clustering
Neural Networks	Dow Jones Industrial Average (DIJA)	Neural Networks

FACULTY

SANDEEP AGARWAL

Sandeep has over 18 years of experience in IT and extensive hands-on expertise in application development involving analysis, design, development and maintenance with 10+ years of experience in data mining and BI and Big Data Hadoop. He has worked across multiple business domains such as Manufacturing, Retail, Banking and Insurance and has experience with large-scale, distributed systems design and development.

YOGESH PARTE

Yogesh is a research engineer with over 14 years of experience in algorithmic development and PoC demonstration using C/C++, Python and R. He is the Founder of Y P Consulting Services, which specializes in innovation engineering and technology applications. He holds a PhD. in Applied Mathematics from University of Paul Sabatier, France and has won over 30 awards for academic excellence.

SATYA SRINIVAS

Satya has 25 years of experience aligning multi-million dollar IT deployments with business strategy and operational processes for Fortune 1000 companies. He has expertise in performance management in enterprise architecture, data mining & analytics, machine learning, pattern recognition, social media analytics and Big Data management & analytics. Satya is a MS - Computer Engineering from Florida Atlantic University.

★ Indicative Faculty ★

PLACEMENT ASSISTANCE

Imarticus provides 100% assistance throughout the program to guide and help navigate ample career options and help you get job-ready from Day 1.



RESUME BUILDING

Refine and polish your resume with experts tips to help you land your coveted job



INTERVIEW PREP

Prepare you to ace HR and technical interview rounds with interview Q&A and mock interviews



MOCK INTERVIEWS

Preparing candidates to face interview scenarios through 1:1 mock interviews with industry veterans

COLLABORATION WITH IBM



COGNITIVE CLASS

Get access to IBM's state-of-the-art content made and delivered by the experts.



IBM PLATFORMS

Aspirants are provided access to IBM Cloud Platforms featuring IBM Watson and other software for 24/7 practice.



IBM CERTIFICATION

All candidates earn IBM Badges on completion of the Prodegree with an option of additional IBM certifications.

IBM is a cognitive solutions and cloud platform company that leverages the power of innovation, data and expertise to improve business and society. For more information, visit www.ibm.com

OUR CAMPUSES



MUMBAI



+91 22 61419595



BANGALORE



+91 80 45129999



CHENNAI



+91 44 43558466



THANE



+91 7400048942



HYDERABAD



+91 40 47123304/06



GURUGRAM



+91 1244874030



PUNE



+91 8433905205



DELHI



+91 8448584775



AHMEDABAD



+91 8980682998

FOR ONLINE DELIVERY INQUIRIES: 1800-267-7679



WWW.IMARTICUS.ORG



INFO@IMARTICUS.COM