



DATA SCIENCE &  
MACHINE LEARNING

---

---

**BOOTCAMP**

---

---

DUBAI

LEARN.NETWORK.GROW

# PROGRAM OVERVIEW

This bootcamp is a focused attempt to provide a solid foundation to everyone seeking to get started as a data scientist. It doesn't matter if you are a fresher or IT professional or non-IT professional, all you need is a passion for problem solving. After completing this course you will possess all the skills required to build analytics solutions using R.



## 200 HOURS PROGRAM

**Pre-Bootcamp** -> **30** Hours of self paced video based learning

**Bootcamp** -> **30** Hours of classroom learning over 3 days

**Post-Bootcamp** -> **5** hours of online doubt clearing + **60** hours of Capstone Project



## KEY FEATURES

6 real life industry projects

6 months access

Access to virtual cloud for practice

International award winning faculty

Separate doubt clearing sessions



## ELIGIBILITY

**Education** -> Graduate in Maths, Science, Commerce, Statistics, Economics or Management

**Work Exp.** -> Fresher / IT / Non-IT



## PREREQUISITES

**No prerequisites**

**Learning starts from scratch**

**Excel** -> Basic level

**Programming Exp.** -> Not required

# PROGRAM STRUCTURE



## GETTING READY

Before you come in, get ready for the Bootcamp. Introduction to data science.

A series of online tutorials will teach you the fundamentals of data science and introduce you to the basic statistics & excel

## 3 DAY BOOTCAMP

Learn and apply data science. 30 hours in 3 days. Build analytics solutions.

Our carefully crafted curriculum provides the right mix of theory, hands-on labs & projects. Our experienced instructors bring case from frontlines to solidify learning.

## BEYOND BOOTCAMP

More learning, capstone project, case studies & doubt clearing session.

Get started on 3 capstone projects. Solidify your learning by accessing several assignments & new case studies. Doubt clearing sessions will be conducted as well.

# PROGRAM OVERVIEW

## PRE BOOTCAMP

- **INTRODUCTION TO STATISTICS**
  - Self paced online classes
  - Online instructor support
  - Access 82 lectures & 9 hours of content 24/7
  - Cover basic statistical principles like mean, median, range, etc.
  - Learn theoretical aspects of statistical concepts
  - Real life industry case studies
- **DATA EXPLORATION USING R**
  - Self paced online classes
  - Online instructor support
  - Discover data types & data structures in R, vectors, arrays, matrices & more
  - 4 real life industry case studies
  - Practice exercise
- **DATA VISUALIZATION**
  - Visualization data in R using a variety of charts & graphs
  - Delve into descriptive & inferential statistics
  - Understand Linear Regression

## BOOTCAMP

- **R PROGRAMMING (DAY 1)**
  - Classroom session
  - 10 labs based on open data
- **LINEAR REGRESSION (DAY 2)**
  - Classroom session
  - Live project on predicting house price
- **LOGISTIC REGRESSION (DAY 2)**
  - Classroom session
  - Live project on predicting attrition
- **CLUSTERING (DAY 2)**
  - Classroom session
  - Live project on customer segmentation
- **MACHINE LEARNING (DAY 3)**
  - Classroom session
  - Live project on Facebook/ Twitter data

## POST BOOTCAMP

- **CAPSTONE PROJECTS & ASSIGNMENTS**
  - 2 predictive modeling capstone projects
  - 1 machine learning capstone project
  - Hands on exercises on R
  - Recordings of Bootcamp sessions
  - Access to learning management system
  - Mentorship for the Projects
  - Doubt Clearing sessions

# POST COMPLETION YOU WILL BE ABLE TO.....



EXPLORE DATA AND TELL THE STORY. REALIZE TRENDS & OUTLIERS.

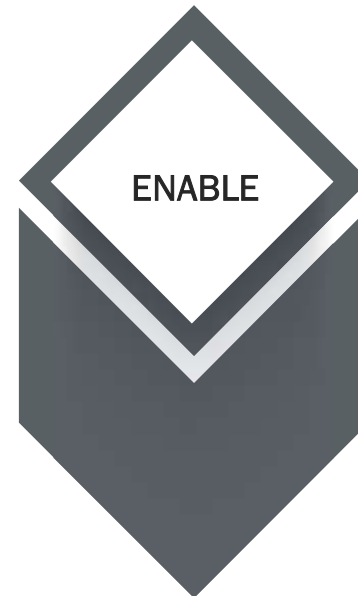
ANALYZE THE RELATIONSHIP BETWEEN DIFFERENT ASPECTS



WORK ON R AND BUILD ANALYTICS SOLUTIONS USING R



ENABLING DECISION MAKING THROUGH PREDICTIVE MODELING



DEMONSTRATE ANALYTICS SKILLS & GROW AS A DATA SCIENTIST



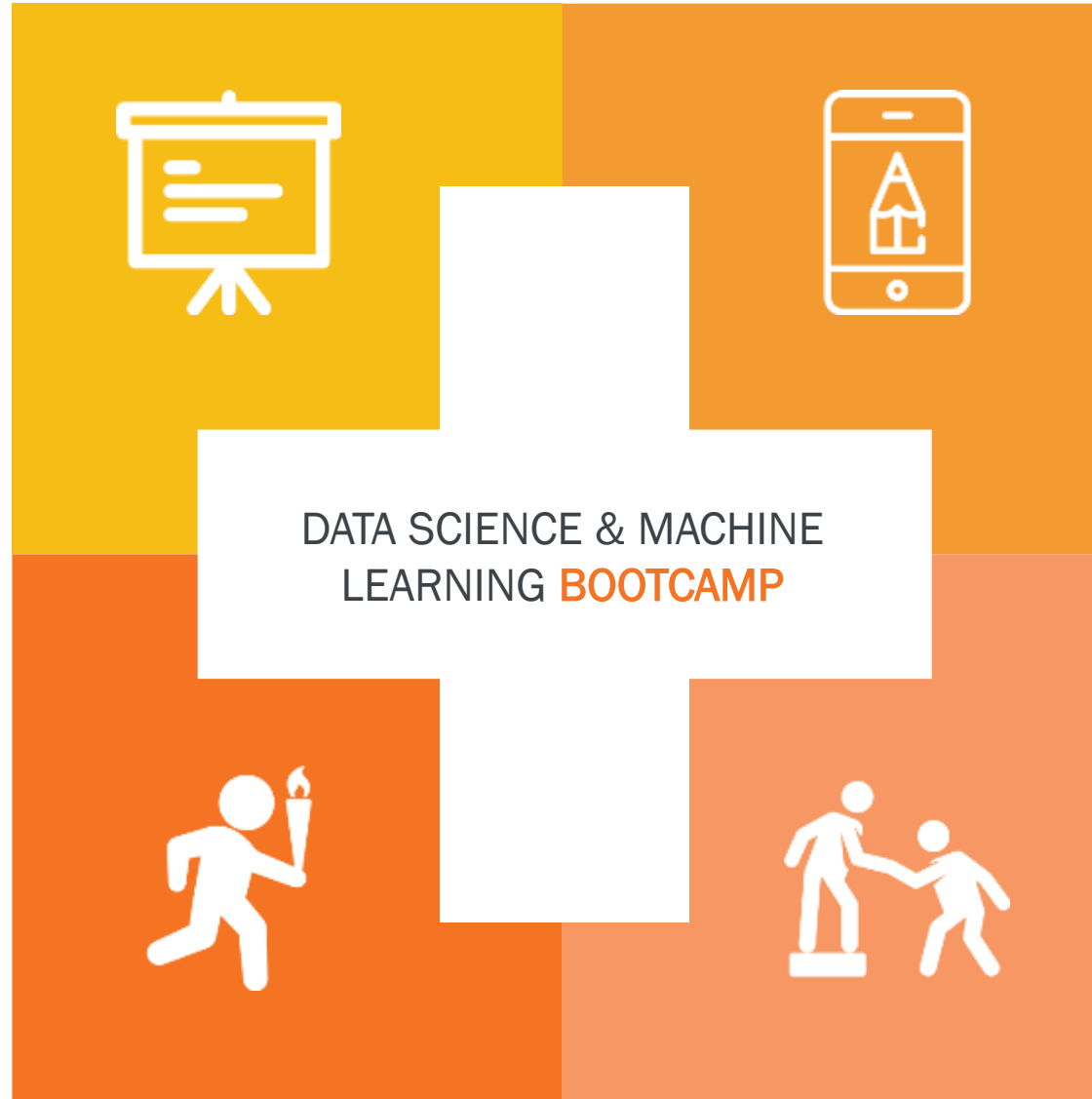
# WHY SHOULD YOU DO THIS BOOTCAMP?

## CUTTING EDGE CURRICULUM

COURSE CONTENT HAND CRAFTED (& *NOT COPY PASTED*) BY CONSULTING EXPERTS. EVERY THING IS LEARNT THROUGH PRACTICAL CASE STUDIES & INDUSTRY PROJECTS

## BUILDS SOLID FOUNDATION

3 FOCUSSED DAYS WITH US IS ALL YOU NEED TO GET STARTED ON ANALYTICS SOLUTION BUILDING USING DATA FROM YOUR ORGANIZATION OR FROM OPEN DATA SOURCES



## ON THE GO LEARNING

ANYTIME ACCESS ONLINE – IMMERSIVE LEARNING VIDEOS, RECORDED SESSIONS, ASSIGNMENTS, PROJECTS, SOLUTIONS & DOUBT CLEARING SUPPORT THROUGH OUR LEARNING SYSTEM

## INDUSTRY MENTORSHIP

RECEIVE 1-TO-1 GUIDANCE FROM INDUSTRY EXPERTS ON HOW TO GET STARTED IN ANALYTICS CAREER OR HOW TO GROW WITHIN YOUR CURRENT ORGANIZATION

# COURSE DETAIL – BOOTCAMP (3 Days)

## R PROGRAMMING

- R overview. Installation. Packages & walkthrough.
- Data structures (Vector, array, factors, data frames, lists).
- Arithmetic & logical operators.
- Subsetting. Missing, indefinite & infinite values.
- For loops. While loops. Nested loops.
- Disadvantage of using loops.
- Alternates to loops.
- Understand the structure of function. Build your own function.
- Package installation & updates.
- Access package functions.
- Hack a function.
- Build your own package.
- Save, load & delete objects.
- Data import & export in R
- Sort & rank.
- Data Aggregation.
- Merging.
- Apply, Lapply, Tapply,
- By, Replicate functions.
- Data visualization in R (plot, ggplot)

## CLUSTERING

- Introduction to clustering.
- Types of clustering & their uses.
- Data standardization
- K-Means clustering.
- Cluster profiling & validation

## LINEAR REGRESSION

- Introduction to linear regression technique & its uses.
- Details of ordinary least squares estimation technique.
- Modeling steps & variable handling.
- Model statistics interpretation.
- Outlier treatment
- Missing value treatment
- Covariance & Correlation
- Multicollinearity & Variance inflation factor
- Model development & validation
- Validation of linear regression assumptions.
- Metrics to measure model performance.

## LOGISTIC REGRESSION

- Introduction to logistic regression technique & its uses.
- Maximum likelihood estimation technique
- Modeling steps.
- Dependent variable definition.
- Variable handling.
- Model statistics interpretation.
- Model development & validation
- Concordance
- Sensitivity, specificity & Accuracy
- Misclassification matrix
- Gini coefficient & ROC curve
- KS statistic & Lift chart
- Weight of Evidence & Information Value.
- Variable reduction.
- Model statistics interpretation.
- Metrics to measure model performance

## MACHINE LEARNING

- Introduction to machine learning
- Decision Trees & Random Forest
  - Understanding decision trees.
  - Understanding classification rules.
  - Understanding Random forest.
  - Modeling using Random Forest.
- Classification using Nearest Neighbors
  - Understanding classification using nearest neighbors.
  - The kNN algorithm - Calculating distance,
  - Choosing an appropriate k,
  - Preparing data for use with kNN.
- Understanding neural networks.
  - Activation functions.
  - Network topology.
  - Training neural networks with backpropagation.
- Text Mining & Sentiment Analysis
  - Concepts and components of text mining
  - Text mining tasks and approaches
  - Access twitter data.
  - Build frequent term network
  - Topic modelling
  - Analysis of followers & retweets
  - Understand sentiment analysis and its key concepts.
  - Sentiment polarity.
  - Opinion summarization.
  - Feature extraction.



# 6 INDUSTRY PROJECTS- COVERED IN BOOTCAMP

## PREDICT HOUSE PRICES

- OBJECTIVE - Use 20+ attributes like number of bedrooms, bathrooms, area, views, age, etc to predict the price of house.
- SOLUTION - We will be building a linear regression model using R.

## PREDICT WHO IS GOING TO LEAVE A TELECOM SERVICE PROVIDER

- OBJECTIVE - Use 50+ attributes across customer behaviour, usage behaviour, package details, etc to predict who will attrite and go to other service provider
- SOLUTION - We will be building a logistic regression model using R.



## CLUSTER RETAIL CUSTOMERS

- OBJECTIVE - Use a large number of aspects like demographics, buying behaviour, value, etc. to realize different homogenous clusters of customers.
- SOLUTION - We will be building a clustering model using R.

## ANALYZE A TWITTER ACCOUNT

- OBJECTIVE - Access a twitter account and create several visualizations to draw interesting insights & perform sentiment analysis on tweets
- SOLUTION - We will be analysing real time data from a twitter account



# POST BOOTCAMP – 3 CAPSTONE PROJECTS

## PREDICT CLAIM AMOUNT FOR AN AUTO INSURANCE COMPANY

- OBJECTIVE - Use 15+ attributes like age, gender, marital status, car details, driving details, etc to predict the claim amount expected from a policy holder.
- SOLUTION - We will be building a linear regression model using R.

## PREDICT IF A TELECOM CUSTOMER IS GOING TO BECOME A HIGH VALUE CUSTOMER

- OBJECTIVE - Basis first 15 days of telecom usage, the company intends to predict who is going to become a high value customer, so that they can be provided differentiated service
- SOLUTION - We will be building a logistic regression model using R.



## DIAGNOSE BREAST CANCER

- OBJECTIVE - Use multiple patient attributes to diagnose breast cancer.
- SOLUTION - We will be building a clustering model using R.

# OUR VALUE PROPOSITION

## YOU WILL LEARN TO BUILD ANALYTICS SOLUTIONS. GUARANTEED!



**LEARN FROM  
'EXPERTS ONLY'**

IVY LEAGUE EDUCATION + 10+  
YEARS OF EXPERIENCE AT  
WORLD'S BEST CONSULTING  
COMPANIES



**REAL INDUSTRY  
PROJECTS**

WE LEVERAGE OUR CONSULTING  
PRACTICE TO BRING NOTHING BUT  
ONLY REAL INDUSTRY PROJECTS



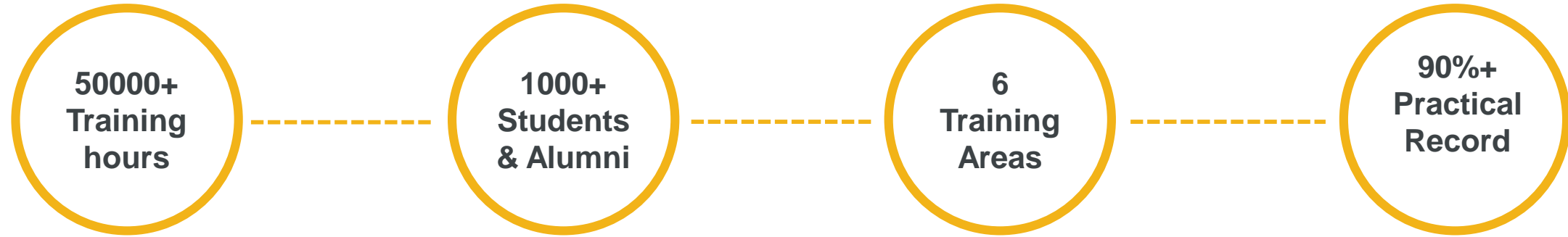
**PERSONALIZED  
LEARNING**

HANDHOLDING FOR ANALYTICS  
NOVICES & NEW EXPOSURE FOR  
ANALYTICS AMATUERS



**GUARANTEED 90%  
Practical**

# OUR DATA SCIENCE TRAINING OVERVIEW



- Rich experience in both online and classroom training. **100+ batches in last 2 years**
- Corporate training at **Fortune 500 companies**
- Analytics programs and guest lectures at **leading B Schools**

- Our students are trained to be successful in keeping with their skillsets and aspirations
- We groom students as **Data Engineers, Data Scientists, Analytics Consultants and Data Creatives**

- We have courses in **Business Analytics, Machine Learning, Big Data, Visualizations, Analytics Tools & Industry specific areas**
- Each area has multiple courses giving the opportunity to select micro specialization
- We use **industry cases only**

- Quality of our students has enabled them to **work across leading consulting & technology companies**
- We forge partnerships with our students well beyond their training helping them with guidance and business outreach



**Average Google & Facebook ratings**

**This is a result of our relentless focus to add value to our students**