



DATA SCIENCE CERTIFICATION COURSE

A Word from the CEO



Mr. Ram Tavva

– CEO and MD of ExcelR Solutions

An alumnus of IIM-C and a PMP-certified trainer with 25+ years of experience.

Whether you are looking for a pay raise, a career transition, or a new job, upskilling is an unceasing necessity. As a training provider, we understand that this need constantly changes with time and technological evolution, and we tailor courses accordingly. We believe that complete learning is a blend of theory and practice. Hence, we offer the kind of post-training support for your assignments and projects that makes us stand apart from our counterparts.

“Our mission is to accompany people in their journey to excellence and empower them to realize their dream careers.”

Our Greatest ACHIEVEMENTS



Best Training Institute by
SiliconIndia on Analytics 2015



Top 20 Most Promising
PMP Training Institutes in India 2018



Top 20 Most Promising
Business Intelligence(BI) Solution
Provider Companies.



4.8 Stars on Google Reviews
for our Training and Support.



We're a Deloitte Technology Fast 50 India Winner



We have been recognized as one of the fastest growing technology companies in the Deloitte Technology Fast 50 India Award Ceremony 2022.

This achievement for ExcelR, which has been accompanied by a staggering average growth rate of 257% in three years, marks an incredible milestone and has propelled our transformation into a global company.

This reflects our dedication towards “Raising Excellence” and demonstrating the company’s professionalism. Our top priority is to provide unparalleled training and consulting to students and professionals in India and abroad so they can shape their future in the right direction.

On the heels of this good news, we are more committed than ever to creating a state-of-the-art platform for all of you to succeed in your career and come out with flying colors.

We're Also a Partner of FutureSkills Prime, A MeitY-**NASSCOM**® Digital Skilling Initiative

We have successfully collaborated with FutureSkills Prime, a first-of-its-kind partnership between MeitY and NASSCOM, to build a digital skilling ecosystem and make India a global digital talent in emerging technologies.

This collaboration will facilitate continuous learning and knowledge enhancement of graduates and working professionals with regard to their aspirations and aptitude, improving their employability.

We’re proud to share in and contribute to the national interest, and we will continue to play a significant role in the digital upskilling revolution in the years to come.





Course Overview

Embark on a transformative journey with our Data Science Course. With over 220+ hours of intensive training, we delve deep into the core principles and advanced techniques that drive data-driven decision-making across industries. Along the path, you'll explore **Statistical Analysis, Hypothesis Testing, Linear and Logistic Regression, EDA, Unsupervised ML Algorithms, Machine Learning Models, Neural Networks, Bagging and Boosting, Text Mining, Forecasting**, and more. The course highlights, curriculum, and value-added courses presented here will tell you why you should become a data scientist and how our training will make a difference.

The Science, Scientists, and Course

Ever wondered how companies predict what you might buy next, personalize your social media feed, or even optimize traffic flow in a city? The answer lies in data science. With the mighty ocean of data and more drops added every nanosecond, it's no wonder the data scientists, professionals who transform data into insights, are among the highest-paid professions worldwide.

How do you unlock the secrets of this in-demand skillset? A data science course equips you with the tools and techniques to navigate this ocean effectively. You'll learn how to wrangle data, uncover patterns, tell data stories, and more, evolving into a valuable asset in the data-driven world.



Why Become a Data Scientist?

#3

High-Paying Careers
in India 2024

#11

Job Vacancies
in India

2.7 Lakh

Jobs Worldwide

11.5 M

Jobs by 2026

12 LPA

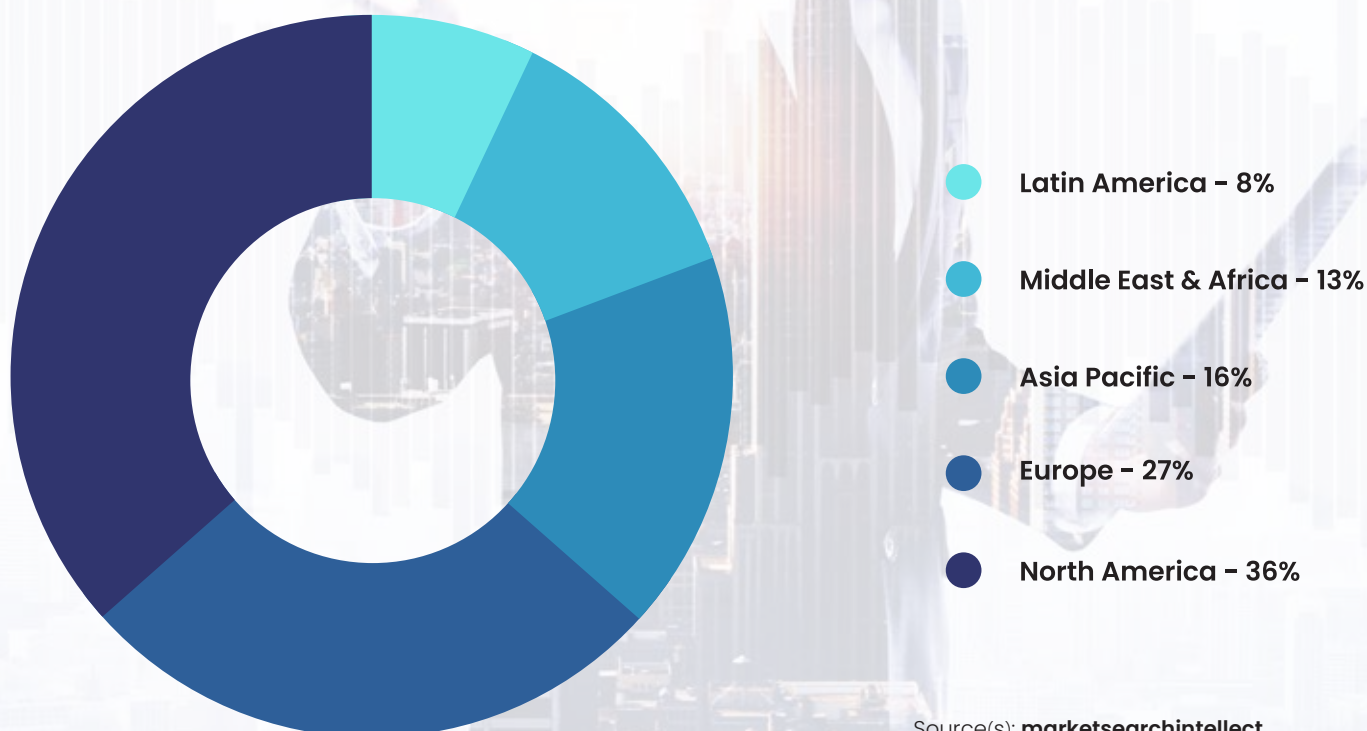
Average
Salary Package

36%

Employment Growth
from 2022 to 2032

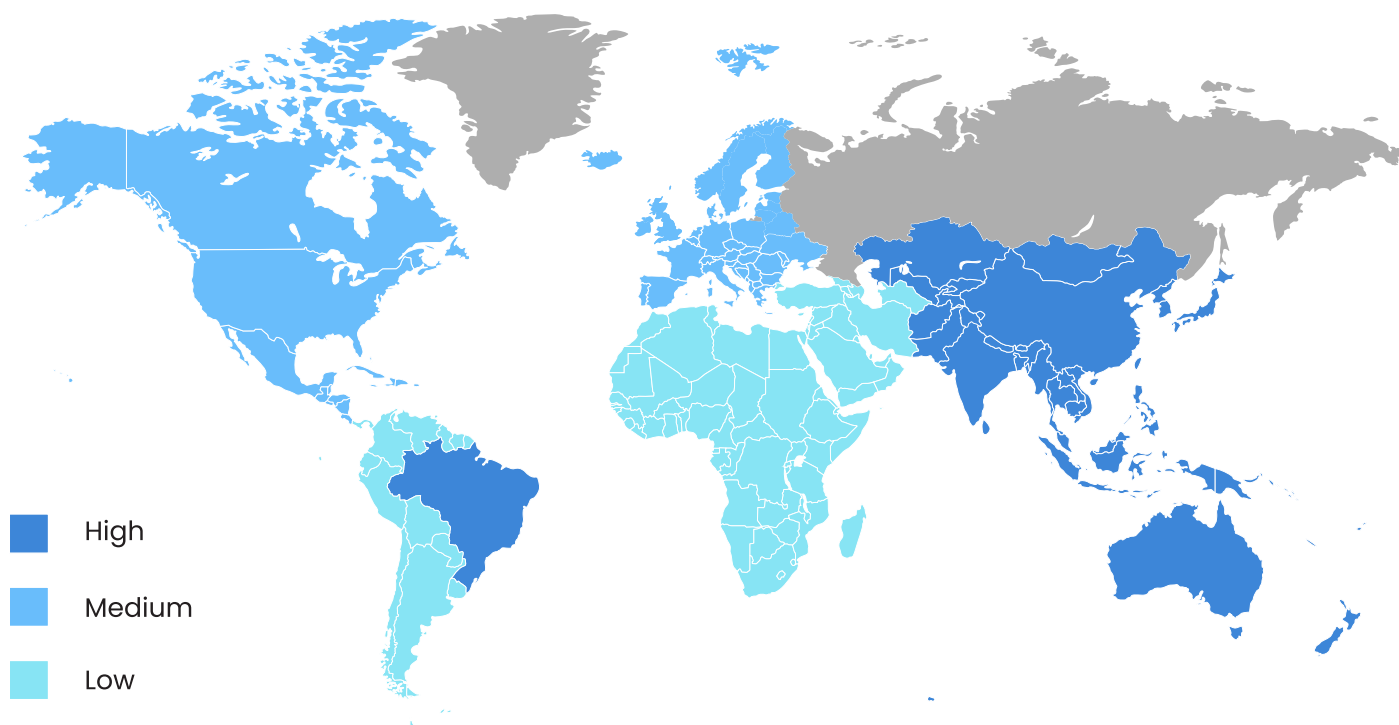
Source(s): **Financial Express, Statista, LinkedIn, The Economic Times, Glassdoor, BLS**

Data Science Industry Segmentation by Geography



Source(s): **marketsearchintellect**

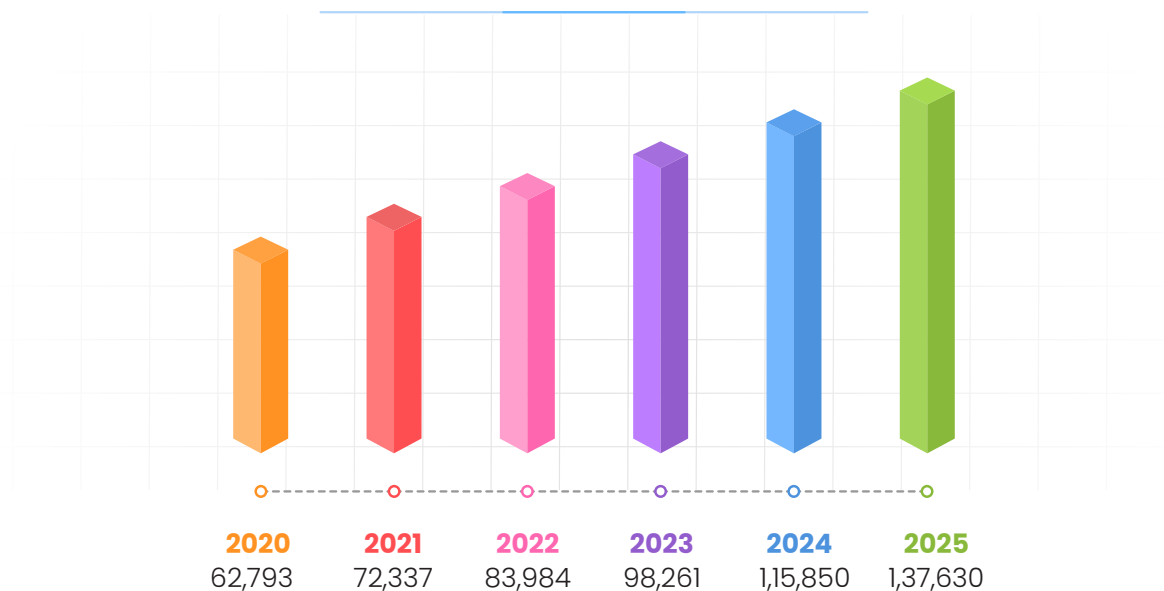
Data Science Market Growth (2024-2029)



Asia Pacific is estimated to have the highest growth rate over the forecast period. **(2024-2029)**

Source(s): Mordor Intelligence; Market Research Intellect

New Jobs in India by 2025

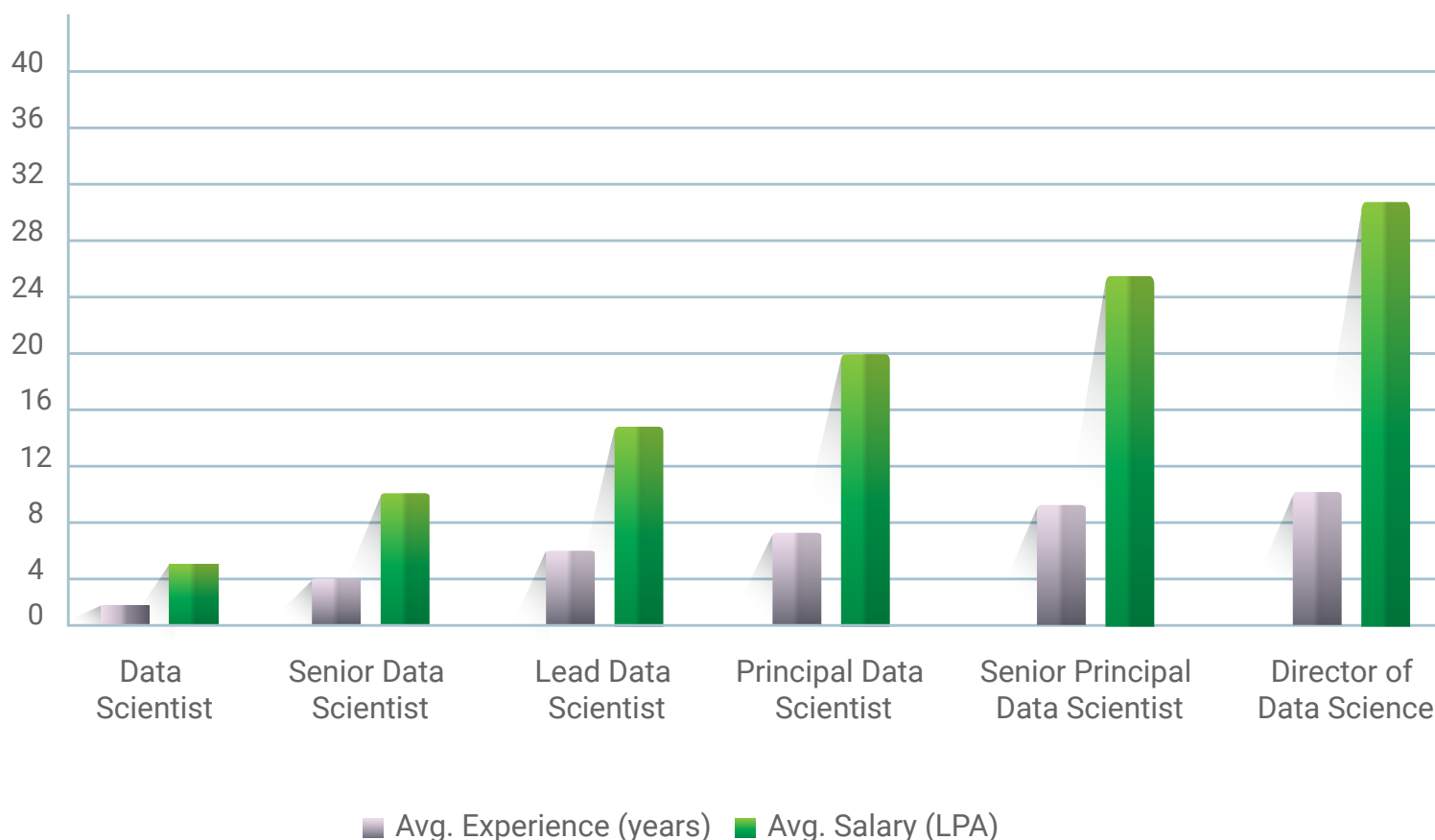


Source(s): **Analytics Insight**

Top Hiring Companies

Deloitte.**tcs** TATA
CONSULTANCY
SERVICES**ORACLE****amazon****DXC**
TECHNOLOGY**accenture****EY****MICRO**
FOCUS**magicbricks**.com**IBM****KONE****Mindtree**
A Larsen & Toubro Group Company**Coca-Cola****Tech**
Mahindra**Honeywell**
THE POWER OF CONNECTED**cognizant**

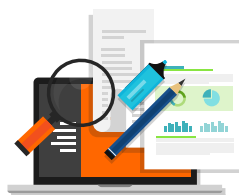
Career Progression and Salary Trends



Course Highlights



Course Duration
220+ Hours/ 6 Months



20+ Capstone Projects



16+ Assignments
& Case Studies



9+Tools Covered



30+ Lab Sessions



8 Value Added Courses

Each module is integrated with assignments for you to apply the best practices of data science. Upon completing the assignments, you'll get the opportunity to work on capstone projects and mould yourself into an industry-ready data scientist.

Tools Covered



Python



Pandas



Numpy



Scikit learn



Tableau



Apache
Spark



MySQL



Azure



ChatGPT

Skills Covered



Statistics



Data Analytics



Data Visualization



Machine Learning Algorithms



Ensemble Techniques



Text Mining/NLP



Forecasting Analytics



LLM & ChatGPT

Value Added Courses

Our course curriculum doesn't end with merely the core concepts of the {Course Name} Course. We also provide value-added courses like **Core Python, Tableau, Azure, ChatGPT**, and more, which go hand-in-hand with the core concepts of the data science course and prove essential in the data science world. This will help you gain a comprehensive skill set, which will give you a competitive edge over your peers during interviews.



Core Python

Mastering Core Python provides programming skills crucial for data scientists. It enables writing clear, efficient code for data manipulation and analysis tasks, enhancing productivity and versatility in handling diverse datasets. Thus, it's chosen as a value-add for this course.



Tableau

Exploring Tableau enhances data scientists' ability to create compelling visualizations, aiding in the effective communication of insights to stakeholders, thus bolstering the impact and persuasiveness of data-driven recommendations and analyses.



MySQL

Learning MySQL empowers data scientists with robust database management skills, facilitating seamless data storage, retrieval, and querying, thereby streamlining the data preprocessing phase and enhancing overall project efficiency.



Artificial Intelligence

Delving into Artificial Intelligence equips data scientists with advanced techniques like machine learning and natural language processing, expanding their toolkit for predictive modeling and text analytics. Hence, we've got it covered as a value-add for you!



Big Data Tools

Big Data Tools, in itself a complete course, adds extraordinary value to this data science course. Familiarizing with Big Data tools like Hadoop and Apache Spark equips data scientists with scalable solutions for handling large volumes of data efficiently.



Azure

Azure offers data scientists access to cloud-based services for scalable data storage, computation, and AI capabilities, providing a flexible and powerful platform for experimentation, deployment, and collaboration on data science projects.



Basics of R

This value-add complements a data scientist's Python skills, offering alternative tools for statistical computing and visualization, enabling them to leverage the strengths of both languages for a thorough data analysis.



ChatGPT

Engaging with ChatGPT exposes data scientists to natural language processing and conversational AI, thus allowing them to develop sophisticated AI-driven solutions for text-based data analysis and interaction.





Industry-Relevant Project Samples

Project 1: "Daily" Twitter Data Analysis for a Product

Domain: Social Media

As more and more people are expressing their views and opinions on various microblogging websites about various products and services. There has been a surge of data generated by the users, these websites have people sharing their thoughts daily.

Sentiment Analysis with the help of Natural Language Processing technique for identifying the sentiments of a product or service

Project 2: Natural Language Processing

Domain: E-commerce

Customers are looking for more information before buying a product on E-commerce websites. Amazon introduced a new feature 'question and answer' search field for products.

The project is to build an information retrieval system from Amazon products' data based on NLP techniques. Top 5 relevant answers to be retrieved based on input question

Project 3: Predicting Loan defaulters

Domain: Banking

Reducing the risk of fraudulent loans by carefully evaluating the risk & at the same time, increasing profits by rejecting only those loans which have the potential of defaulting.

Hinduja Finance

Project 4: Warranty Cost prediction

Domain: FMCG

The objective of the analysis is to predict an item when sold, what is the probability that a customer would file for a warranty, and to understand important factors associated with them.

Project 5: Predict flight delays

Domain: Aviation

Predict which flights will be delayed and by how long.

Flight delays cost the industry an estimated \$25 billion every year. More than 60% of frequent flyers cite delays among the things about air travel that they find most dismaying. And the costs are spread around — an extra \$25 in parking here, a missed business meeting there. Carriers, meanwhile, pay an estimated \$62 per minute in crew, fuel, maintenance and other costs. It adds up.

Your 3-in-1 Certifications

Certificate from **Nasscom**



Certificate from **ExcelR**



Internship Certificate from **AiVariant**





Curriculum:

Data Science and Value-Added Courses

Demo – Datascience, DA vs DS vs AI

Introduction

- What is Data Science?
- Use cases with Business Problem (Mobile/Banking) and How ML gives a solution
- Types of Roles
- what learnings are important
- Data entry (fill series)
- VAC courses offers
- Jumbo Pass
- Q & A

Project Life Cycle

- ML Project Life Cycle (Problem, Collecting the data, EDA,Cleaning,Transformation, Partition, Model fitting, Cross validation, Metrics, Deployment),

Fundamentals of Stats – Basic Stat

- Sample
- population
- Data types(continous, discrete)
- Central tendency
- spread
- shape of the data such histogram
- skewness
- kurtosis

Statistical Charts – Basic Stat

- Bargraph
- Box plot(IQR, Whisker lengths, outliers)
- Scatter plot(Positive , Negative, Neutral)
- correlation

Fundamentals of Python

- Intro to Python language
- Anaconda Installation(Jupyter, Spyder)
- Datatypes(Int, Float,dic,Set)
- Operators(Arthematic,comparision,Logical, Assignment)
- Data structures(List (types of list methods such as append ,extend ,insert ,remove , pop ,clear ,index ,count ,sort ,reverse), tuples,dictionary,set)
- What are Control structures (if, ifelse, if elif, Nested if)



Python – Loops, Functions, Numpy

- For loop
- functions
- numpy (scalar, array, vector, 1 dim, 2 dim, random int)
- converting numpy to panda
- giving column names
- Importing pandas (read_csv, head, tail, describe)

Exploring Pandas

- Pandas (info, selecting columns, dropping columns, groupby, concat(row and columns), merge, removing duplicates, filling blanks with mean)

Data Visualization: EDA

- EDA (showing graphs such as histogram)
- boxplot
- bargraph
- scatter plot
- heat map using matplotlib, seaborn) using Google collab with generative AI usage
- Giving an example dataset ask them to work in class

Advance Stats – Probability – Normal distribution

- Probability, Normal distribution theory, standardization, zscore, z tables, applications, python code, confidence Interval

Data Summarization: Pivot Report and Charts

- Probability
- Normal distribution theory
- standardization
- zscore
- z tables
- applications
- python code
- confidence Interval

Advance Stats – Hypothesis Testing

- Level of significance Hypothesis Testing (One sample Z test, Two sample Z test)
- t-test

Intro to Machine learning (Supervised Learning) and SLR

- Simple Linear Regression
- metrics such RMSE and R square – Working on Age vs Weight example



Multiple Linear Regression

- Intro to Regression models
- MLR – Assumptions of Linear Regression
- Variable selection
- Multicollinearity VIF

Logistic Regression

- what is meant by classification models ?
- When do we choose Logistic regression
- model fitting
- confusion matrix
- accuracy score – Working on Breast cancer case study
- Other metrics Sensitivity
- Specificity
- precision
- F1 score
- ROC curve
- AUC score

Data Transformation

- Data Transformation (Standard scaler, minmax scaler, label encoding, one hot encoding) and Data partition (Training and Test)

Model Validation Techniques

- Cross validation (Stratified K-Fold, K-Fold cross validation, Shuffle Split Cross-Validation)

Under fitting to Over fitting

- Variance Biased Trade-off (under fitting – causes Lack of training, best fit, over fitting – causes Noise in training data, Too many training epochs or iterations, too many variables)
- Visualizations (Underfitting, bestfit, Overfitting) and Feature Engineering – Working on Bangalore housing prices case study

Regularization

- Techniques such Lasso
- Ridge
- ElasticNet – Working on "Bangalore housing prices" case study



Classifiers - SVM

- Support vector machine (Hyperplane, Maximum margin classifier, Support Vectors, SVM for Linear Classification)
- SVM for Non-Linear Classification (polynomial, RBF, Sigmoid)

Decision Tree

- Decision Tree Structure (Root node, Internal nodes, terminal nodes)
- Gini Impurity
- Entropy and Information Gain (for classification)
- Overfitting and Underfitting in Decision Trees
- Pruning
- Hyperparameters - Working on Sales data set using python

Ensembled Techniques

- Ensemble Methods: Bagging and Random forests
- working on hyper parameters to control overfitting

Boosting methods

- Sequential methods: Gradient Boosting
- AdaBoost
- using Grid search CV
- XG Boost
- LightGBM

Deployment - Project Discussion

- Final project with Deployment

Unsupervised Learning

- "What are Dimensional Reduction Techniques ?"
 1. Purpose of PCA
 2. Eigenvectors/Eigen values
 3. Applications
 4. Advantages
 5. Working on case study"

Clustering

- Introduction to Clustering
- Distance Metrics
- Clustering Algorithms (K mean, dbscan)
- Choosing the Right Number of Clusters (Elbow Method, Silhouette Analysis)

Recommendation system

- what is Recommendation and why it is important?
- What is Collaborative Filtering (CF) And Content-Based Filtering ?

Time series

- Time series Concepts
- components
- Visualization
- Data partition
- Lagplot
- ARIMA models
- Python code on ARIMA models

Deep Learning – Artificial Neural Network

- Perceptron
- Single Layer Network
- activation functions
- Back propagation method
- Simple ANN code
- Multilayer Neural network
- Gradient Descent method
- optimizers
- learning rate – complete code with tensorflow

Deployment – Project Discussion

- Final project with Deployment

Deep Learning – Recurrent Neural Networks

- RNN – use cases
- vanishing and exploding problem
- Simple RNN code
- LSTM Architecture
- Working model
- LSTM vs GRU
- python code

Natural language processing

- What is Text Data
- Various forms
- Applications
- Text pre-processing (Tokenization, Normalization, Stopwords, Lemmatization, stemming)
- Visualization on preprocessed text data
- Text Representation: Bag of Words (BoW)
- Term Frequency-Inverse Document Frequency (TF-IDF)
- Sentiment Analysis
- Classification model using ML
- Named Entity Recognition (NER)
- What is Word Embedding?
- What are pre-trained word Embeddings
- Word2Vec (Skip gram, CBOW)
- real time applications
- example codes



Natural language processing

- Language Modeling: N-gram Models
- Neural Language Models
- application of RNNs
- LSTMs on Text data

Generative AI – Intro to LLM

- Large Language Models?
- Transfer Learnings in NLP
- what are pre-trained models?

Generative AI – Applications of LLM's

- what are transformers?
- Hugging Face transformers library and its use cases

Core Python

Introduction to Python, Variables

- Python Introduction – Programing Cycle of Python
- Python Installation
- Python IDE
- Variables
- Data types

Operators,Data Structures

- Operator –Arthmatic
- Comparison
- Assignment
- Logical
- Bitwise opeartor
- List
- Tuple
- Set
- Dictironary

Conditional Statements and Loops

- Conditional Statements (if, if-else, if elif, Nested if)
- Loops in Python (for, while)
- Loop Control Statements(break, continue, pass)



Function

- "Function – Define function
- Calling function
- pass by reference as value
- Function arguments
- Anonymous functions
- return statements
- Scope of variables – local & global
- Lambda
- map
- filter
- reduce

Modules and Packages

- Importing modules
- Creating user-defined modules
- Python Standard Library
- Installing packages using pip

Data Cleaning and Preparation

- Importing the data
- Handling Missing Data
- Filtering Out Missing Data
- Filling In Missing Data
- Data Transformation
- Removing Duplicates
- Data Type Conversion
- Detecting outliers using Boxplot
- Z score
- Handling Outliers (Capping, Transformation, Removal)
- Transforming Data Using a Function or Mapping, Replacing Values
- Feature Engineering such as Creating new variables
- Aggregations and groupings

Data Wrangling: Join, Combine, and Reshape

- Hierarchical Indexing
- Combining and Merging multiple datasets (merge())
- join()
- concat()
- Reshaping and Pivoting

Handling Time data using pandas

- Convert to datetime
- Extract attributes
- Create datetime range
- Resample data
- Time delta calculations
- Add time offset
- Time zone conversion
- Set datetime index
- Filter by date
- Handle missing time data



Exception Handling & Regular Expressions

- 1. Exception Handling :
Try, except, else, finally ,Built-in exceptions ,Raising exceptions,
Custom exceptions ,Hands-on error handling tasks
- 2. Regular expressions:
match function, search function , matching vs searching
Regular exp modifiers and patterns"

Object-Oriented Programming (OOP) – Part 1

- Class and Object, __init__ method
- Attributes and methods
- Hands-on: Create simple classes

Object-Oriented Programming (OOP) – Part 2

- Inheritance
- Polymorphism
- Hands-on: Real-world OOP examples

Object-Oriented Programming (OOP) – Part 3

- Encapsulation and Abstraction
- Hands-on: Real-world OOP examples

Advanced Topics

- Iterators and Generators
- Decorators

Tableau

Introduction to Tableau

- What is Tableau?
- What is Data Visualization?
- Tableau Products
- Tableau Desktop Variations
- Tableau File Extensions
- Data Types, Dimensions, Measures, Aggregation concept
- Tableau Desktop Installation
- Data Source Overview
- Live Vs Extract



Basic Charts & Formatting

- Overview of worksheet sections
- Shelves
- Bar Chart, Stacked Bar Chart
- Discrete & Continuous Line Charts
- Symbol Map & Filled Map
- Text Table, Highlight Table
- Formatting: Remove grid lines, hiding the axes, conversion of numbers to thousands, millions, Shading, Row divider, Column divider
- Marks Card

Filters

- What are Filters?
- Types of Filters
- Extract, Data Source, Context, Dimension, Measure, Quick Filters
- Order of operation of filters
- Cascading
- Apply to Worksheets

Calculations

- Need for calculations
- Types: Basic, LOD's, Table
- Examples of Basic Calculations: Aggregate functions, Logical functions, String functions, Tableau calculation functions, numerical functions, Date functions
- LOD's: Examples
- Table Calculations: Examples

Data Combining Techniques

- What are Data Combining Techniques?
- Types
- Joins, Relationships, Blending, and Union

Custom Charts

- Dual Axis
- Combined Axis
- Donut Chart
- Lollipop Chart
- KPI Cards (Simple)
- KPI Cards (With Shape)

Groups, Bins, Hierarchies, Sets, Parameters

- What are Groups? Purpose
- What are Bins? Purpose
- What are Hierarchies? Purpose
- What are Sets? Purpose
- What are Parameters? Purpose and examples



Analytics & Dashboard

- Reference Lines
- Trend Line
- Overview of Dashboard: Tiled Vs Floating
- All Objects overview, Layout overview
- Dashboard creation with formatting

Dashboard Actions & Tableau Public

- Actions: Filter, Highlight, URL, Sheet, Parameter, Set
- How to save the workbook to Tableau Public website?

MySQL

Introduction to MySQL

- Introduction to Databases
- Introduction to RDBMS
- Explain RDBMS through normalization
- Different types of RDBMS
- Software Installation (MySQL Workbench)

SQL Commands and Data Types

- Types of SQL Commands (DDL, DML, DQL, DCL, TCL) and their applications
- Data Types in SQL (Numeric, Char, Datetime)

DQL & Operators

- SELECT: LIMIT, DISTINCT, WHERE, AND, OR, IN, NOT IN, BETWEEN, EXIST, ISNULL, IS NOT NULL, Wild Cards, ORDER BY

Case When Then and Handling NULL Values

- Usage of Case When Then to solve logical problems and handling NULL Values (IFNULL, COALESCE)

Group Operations & Aggregate Functions

- Group By, Having clause
- COUNT, SUM, AVG, MIN, MAX, COUNT
- String functions
- Date & Time function



Constraints

- NOT NULL
- UNIQUE
- CHECK
- DEFAULT
- Primary key
- Foreign Key (Both at column level and table level)

Joins

- Inner
- Left
- Right
- Cross
- Self joins
- Full outer join

DDL Commands

- Create
- Drop
- Alter
- Rename
- Truncate
- Modify
- Comment

DML & TCL Commands

- DML: Insert, Update, and Delete
- TCL: Commit, Rollback, Savepoint, and Data Partitioning

Indexes and Views

- Indexes (Different Type of Indexes) and Views in SQL

Stored Procedure

- Procedure with IN Parameter
- Procedure with OUT parameter
- Procedure with INOUT parameter

Function, Constructs

- User Defined Function
- Window Functions – Rank, Dense Rank, Lead, Lag, Row_number



Union, Intersect, Sub-query

- Union
- Union all
- Intersect
- Sub Queries
- Multiple Query

Exception Handling, Loops, Cursor

- Handling exceptions in a query
 - CONTINUE Handler
 - EXIT handler
- Loops: Simple, Repeat, While
- Cursor

Triggers

- Triggers – Before | After DML Statement

MLOps

Intro to MLOps

- What is MLOps,
- Different stages in MLOps
- ML project lifecycle
- Job Roles in MLOps

Design and Development

- What is Development stage of an ML workflow
- Pipelines and steps
- Artifacts
- Materializers
- Parameters & Settings

Execution

- Stacks & components
- Orchestrators
- Artifact stores
- Flavors etc

Management

- ML Server infrastructure
- Server deployment
- Metadata tracking
- Collaborations
- Dashboards



ChatGPT

Introduction to ChatGPT

- "History and Development of ChatGPT"
- Examples of ChatGPT use in various industries
- Basics of Transformers
- Key concepts and principles of Generative AI
- Examples of Generative AI models including ChatGPT, open source LLM's, Prompting basics,
- Overview of Different ChatGPT models

ChatGPT Prompting, Generic Applications

- "Prompt Techniques"
- Few-shot Prompting
- Zero Shot prompting
- One-Shot Prompting
- Chain of Thought Prompting ChatGPT applications in everyday life such as writing translation and creativity
- Explore ChatGPT potential for Education, Work and Business Use Cases

ChatGPT in Everyday Tools and ML AI applications

- Utilizing ChatGPT for Excel, word, powerpoint, web development, data analysis, programming, Dashboards, ChatGPT projects etc.

ChatGPT for Interview Preparation and Job readiness

- Seeking jobs, career changes, working on resume, and updation, networking, job search strategies using ChatGPT
- LinkedIn Profile Optimization"

API Integration & Custom Workflow Automation

- Introduction to OpenAI API & usage limits Authentication
- Endpoint usage Integrating GPT with Python, Google Sheets, Excel, Power BI Zapier, Make, LangChain basics

Artificial Intelligence

Fundamentals of Mathematics and Revision of ANN

- Linear Algebra: Vectors, matrices, dot product, matrix multiplication
- Calculus: Derivatives, partial derivatives, chain rule (for backpropagation)
- Review on ANN"



Detail explanation on Artificial Neural Networks

- What is Gradient Descent in detail? Connectivity of Calculus in Back propagation. Weight & Bias Updates
- Types of Loss functions(MSE, Binary Cross entropy (Binary & Muliti)
- Overfitting Solutions(Dropout, Early stopping)
- Types of Optimizers and acitvations functions its applications
- Example Case study

CNN

- CNN,DeepConvolutionModel,DetectionAlgorithm, CNN FaceRecognition
- Working on MNIST data set"

Web Scrapping

- Introduction to Web Scraping & Web Basics,
- Python Libraries for Web Scraping (requests, BeautifulSoup)
- HTML & Web Page Structure Basics

Web Scrapping Application integrated with ML

- "Selecting a website and extract the data
 1. To Extract Image, Reviews, Ratings, and Price Tags
 2. Store in Structured Format
 3. a) Image classification from images
b) Sentiment analysis from Reviews
c) Regression model from Prices."

Big Data Tools

Intro to Big data technologies

- What is big data,
- characteristics of big data
- technologies in big data etc

Intro to Spark environment

- what is spark environment
- spark documentation
- installation of spark
- spark concepts



Integration of Spark Platform

- Integration with different languages like python , r, scala, etc. Introducing pyspark environment , pyspark basics and functions

Pyspark Concepts

- Pyspark RDD structures, dataframe modules, sql modules , examples , exercise problems, working on datasets

Pyspark ML Concepts

- Pyspark ML libraries, Regression models, linear and logistic regression and clustering basics, tree based models, ensemble concepts

Pyspark ML Applications

- Pyspark ML applications, with excercises, visualizations

Databricks environment

- What is databricks, account creation, cluster creation, working on pyspark applications in databricks with r, python and scala

Intro to AWS cloud

- What is aws cloud, account creation , understanding basic aws environment and knowledge

Hadoop environment

- What is hadoop
- hadoop architecture
- creating hadoop environment on AWS cloud
- install java
- install hadoop and related concepts

Hadoop applications

- Running applications like map reduce on data
- getting insights
- doing analysis
- word count problems etc



Azure

Intro to Cloud Computing

- What is cloud computing
- why it is important
- cloud services
- applications
- benefits
- architectures

Intro to Azure Cloud Platform

- What is Azure
- Why Azure
- Azure services
- Azure core architecture
- core azure services domains
- creation of azure account

Azure Cloud Applications

- Intro to AI/ML services
- What is azure ml designer studio
- developing ml models
- python and r applications in studio

Azure Cloud Services

- Resource groups
- virtual machine concepts
- storage service
- web apps
- databricks environment
- azure sql databases
- billing etc.

Azure open AI Studio

- What is azure open ai
- open ai documentation
- how to use azure open ai studio
- creating applications
- different models in azure open ai



Basics of R

Introduction to R, Installation of R Studio, Data types in R

- Data types(Numeric, Char, logical, Complex, Vector, List, Matrix, Factor, Array, Data frame)
- Relational Operators
- Logical Operators

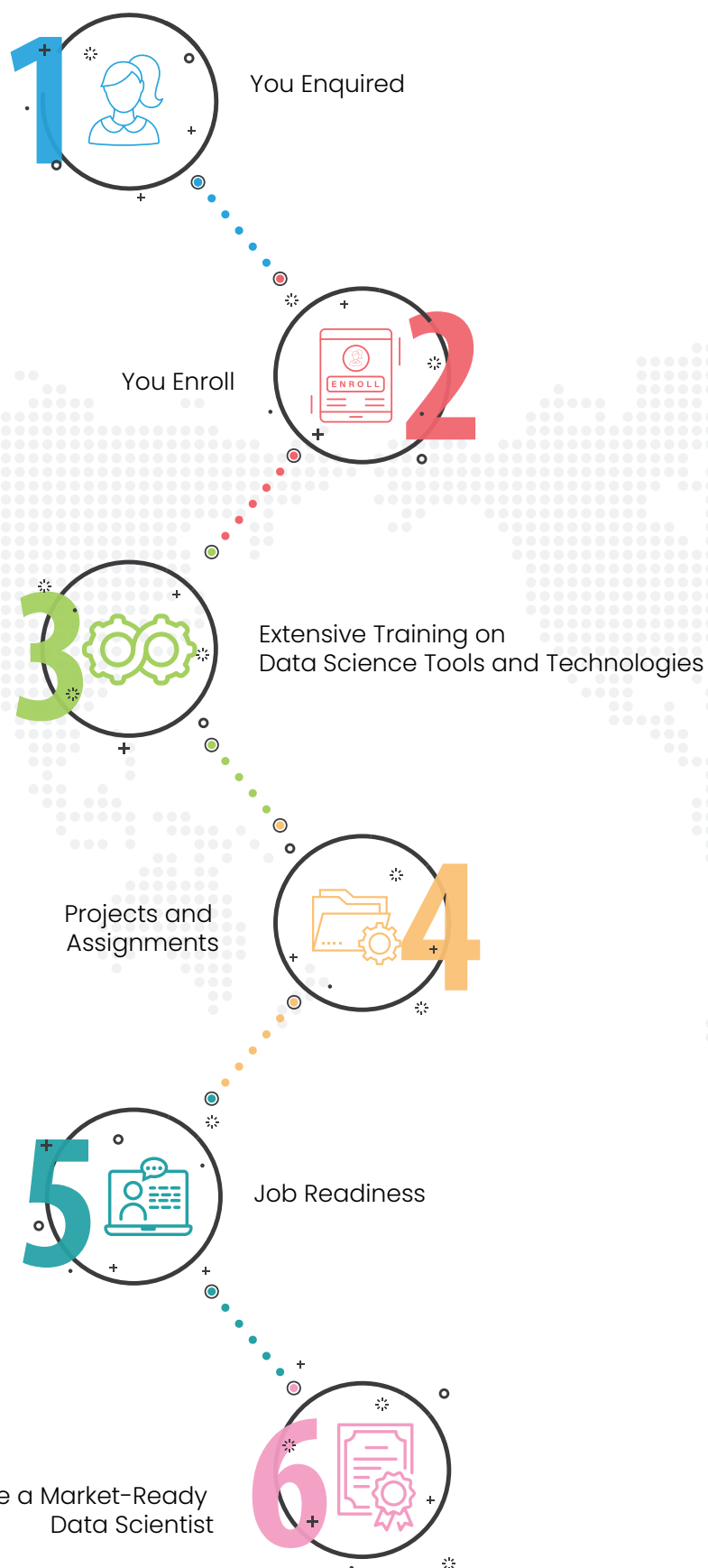
Decision Making statements, Loops, Functions

- If, Ifelse, For loop, While loop, Repeat, Functions

Built in Functions in R, Joins, dplyr and ggplot2

- Merging data frames
- Analyzing Iris Dataset using apply functions dplyr package(Filter, Set, Arrange)
- Data Visualization using ggplot2
- Scatterplot
- Histogram
- Boxplot

Learning Path



Why ExcelR?



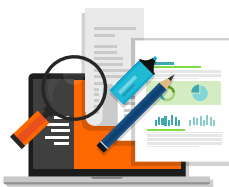
Top-Notch Faculty

Our faculty is our strength. We have certified data science trainers, with over 12 years of experience, who have a strong passion for training. The trainers are always at your disposal to assist you throughout the learning process.



Jumbo Pass

With our complimentary Jumbo Pass, you get access to attend one classroom batch and unlimited live online batches for one year from different trainers of your choice, for better absorption and assimilation of concepts.



Industry-Based Course Curriculum

Considering the unmet skill gaps in industry, our course modules are aligned to the latest technological trends, and cover all the essential topics and tools.



Dedicated Post-Training Support

Learners gain hands-on experience when they start on their assignments and case studies. We provide dedicated mentors to help with any queries during this phase.



Lifetime Access to the Learning Management System

We provide lifetime access to the Learning Management System, so you can revise concepts by going through the recorded sessions as and when you want.



Job Assistance Program

Our data science training doesn't stop with theory and practice. We also take responsibility for making you job-ready. Our placement support is included with the course.



4.8/5 Star Reviews on Google

We have had many happy learners over the years, and our ratings reflect our ethos of striving to always provide a high-quality learning experience.

Job Readiness Program

Our dedicated placement teams have a decade-plus of experience transitioning students from classroom to workplace. Here's how it goes:



Resume Preparation Session

Our mentors will guide you through preparing your personalized resumes connecting your talents, skills, previous experience, etc.



LinkedIn Profile Optimization

LinkedIn is the platform where you get the tickets to your dream job, and we help you optimize your profile there in the finest way so you bag the best salary in the market.



Interview Preparatory Sessions From Day 1

We provide intensive interview preparation right from Day 1 to prepare candidates for interviews with our network of 2000+ hiring partners.



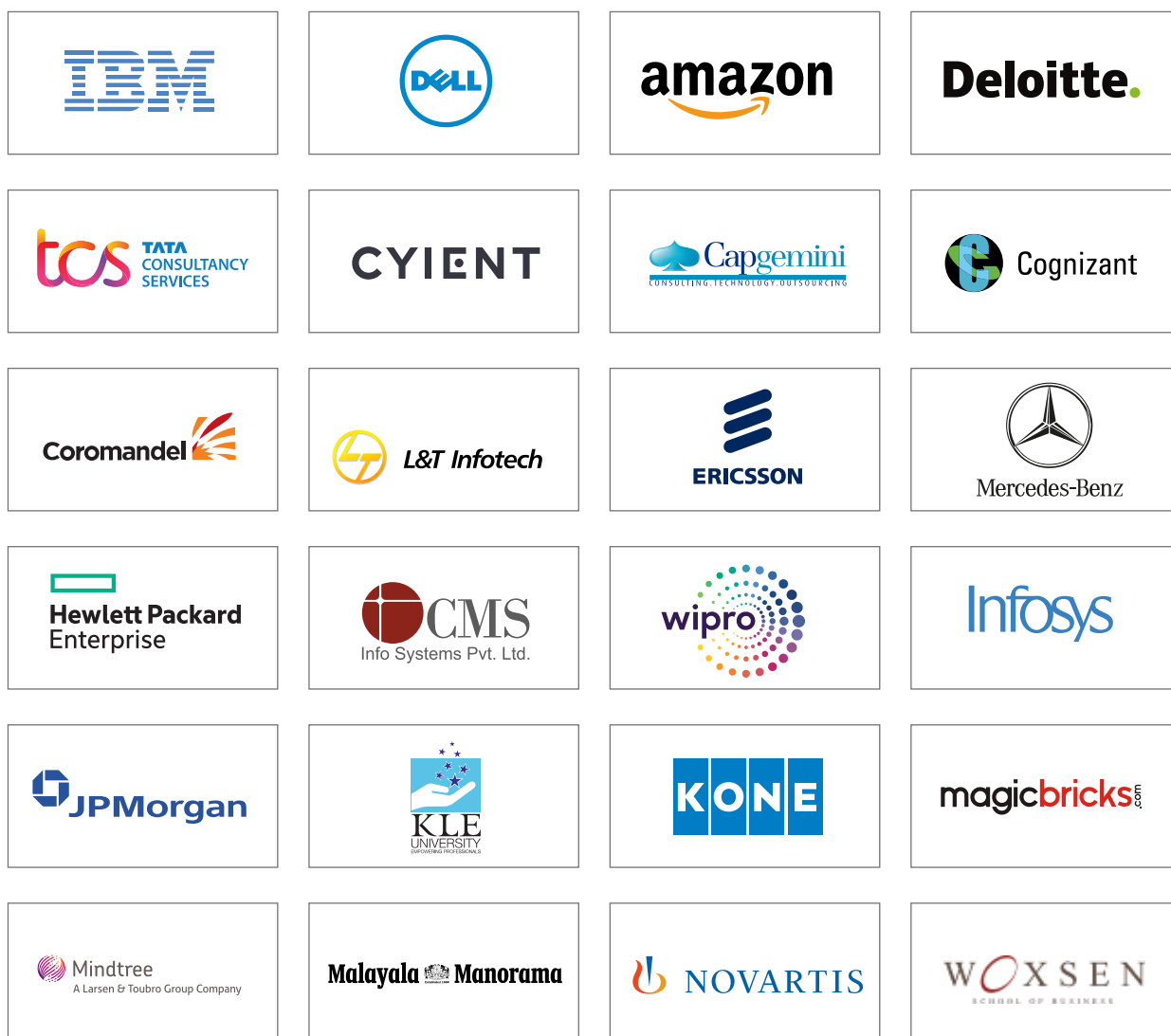
Mock Interviews

A panel of industry experts will conduct a SWOT analysis and present you with the feedback and ratings of how you performed in the interview and the areas you need to improve.

Academic Partners



Our Clientele





and many more...

CLASSROOM	CLASSROOM	LIVE VIRTUAL	LIVE VIRTUAL
with IIT Madras Pravartak Certificate	without IIT Madras Pravartak Certificate	with IIT Madras Pravartak Certificate	without IIT Madras Pravartak Certificate
₹99999/-	₹78999/-	₹95999/-	₹68999/-
₹79999 Incl. of GST	₹59999 Incl. of GST	₹75999 Incl. of GST	₹54999 Incl. of GST

Classroom training is at ExcelR campus and Optional 2-day Campus Immersion in the beautiful, state-of-the-art IIT Madras.

Payment Options

- ☒ UPI
- ☒ Netbanking
- ☒ Debit / Credit Cards
- ☒ All Major Credit Cards Are Accepted (One-Time Payment)

EMI Partners



Loans

All Major Bank Credit / Debit Cards Accepted for EMIs too.



For further information regarding EMI, please feel free to contact our career counsellors.

We don't just TRAIN

We build CAREERS



INDIA



Bangalore



Chennai



Hyderabad



Pune



Mumbai

Overseas



Malaysia
Kuala Lumpur



USA
Houston



UK
Reading RG1 5DS



Singapore