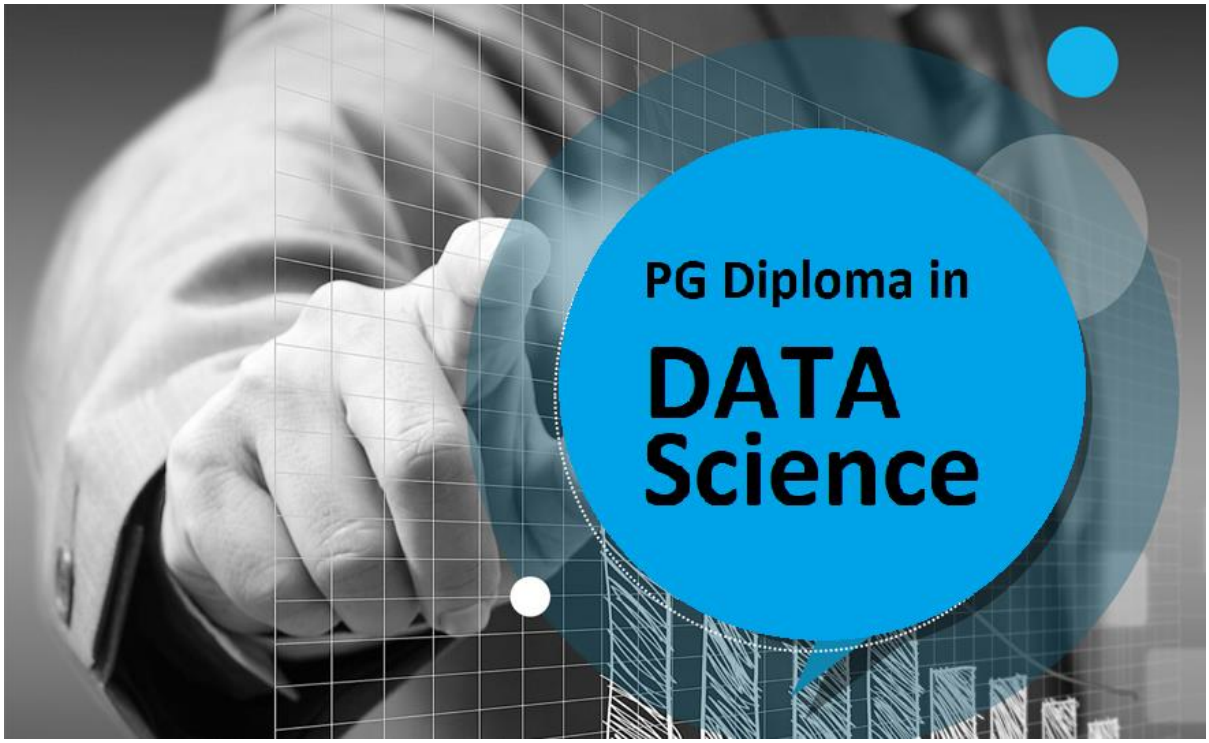
iAnalyst

Presents



Call: (+91) 8669694488
Email: enquiry2ianalyst@gmail.com
Visit: www.ianalyst.solutions

Address

1. **Wakad –** Flat No. 106, Vrundavan Society, Opp. Polarise Healthcare, Datta Mandir Road, Wakad
2. **Pimple Saudagar –** **Opening on 1st September** at Vision9, Kunal Icon Road

Program Features

PGD-DS is a 6-month comprehensive program that combines Fundamentals such as Statistics, Linear Algebra and Data science, Machine learning, Deep Learning and Computer Vision to prepare candidates for various roles related to Data Science such as Data Scientists, Machine Learning engineers, Machine Learning architects, Technology architects, Solution Engineers, Technology Consultants, Chief Technology Officers etc.

Wide Audience

The program is designed for wide audience from fresher to experienced professionals. delivered on either Weekdays or Weekends.

Industry Exposure

The program gives the participants an opportunity to learn from some of the Industry's leading minds in Machine Learning. Guest Lectures, Case studies and Problem Walkthroughs ensure that participants get an understanding of what's happening in the industry. Every candidate gets a mentor who provides guidance and mentorship for capstone project which further enhances the learning experience.

Extensive Hands On

The program includes extensive hands-on approach where participants have to work on several challenging problems after each sessions, case studies, mini projects and capstone project. By working on problems hands on, candidates gets good understanding of the concepts taught in the class sessions and gain expertise by applying to live problems.

Classroom Learning

The program spans 150+ hours of content in Data Science delivered through in-person classroom sessions. The classroom schedule will be shared at the start of the batch to provide understanding of different sessions and which days it will be covered.

Capstone and Mini Projects

The projects are covered during the training after the topics. Our participants have often hailed the capstone project as one of the hallmarks of the learning experience because of the quality of work they get to do and also the mentorship that they receive through Great Lakes faculty and industry experts.

InHouse Projects

iAnalyst works with corporate clients and companies on their Data Science projects. This program gives our an opportunity to work on good PoC and Live industry Projects done at iAnalysts Internship Program. This entirely depends on the projects and programs undertaken at the time of the Batch Program.

Program Highlights

R and Python

R and Python are the two most used tools by Data Scientists. Considering the job market we have included R and Python for this program. This session includes details of R and Python from Installation, Language Fundamentals till details needed for Data Science Hands On.

Linear Algebra

Linear algebra forms the fundamentals required to learn Data Science. We will learn how to represent the data in different dimensions, vectors and matrices. Different operations related to these such as matrix decomposition, matrix arithmetic etc.

Statistics

This topic covers the statistical concepts required for Data Science to understand the data and distribution of Data.

Exploratory Data Analysis

This topic covers doing data preparation and data visualization, transformation of Data etc.

Machine Learning – Supervised and Unsupervised

In this topic we kick start the core Machine learning. All of the Supervised and Unsupervised machine learning algorithms will be covered.

Neural Networks

Artificial Neural Networks, Convolutional and Recurrent Neural Networks using Tensor flow and Keras.

Computer Vision

The topics cover pattern recognition using neural networks and computer vision.

Time Series

Time Series is a special type of data and this separate section covers all about Time Series and Prediction.

Natural Language Processing

This topic covers Natural Language Processing to process string and Text data for different applications such as sentiment analysis, identifying positive and negative reviews.

For the Detailed Course Curriculum, Please refer at the end

Faculty Profile

Learn Expert Data Science Trainers

iAnalysts faculty have been ranked amongst the best data science faculty. All of our faculty have significant industry experience and Expertise in Data Science.

Sachin Adnaik

Sachin is an PhD in Statistics with several years of experience in Data Science. He has conducted several batches of Data Science, Machine Learning, Six Sigma Black Belt, Minitab, SPSS, SAS batches in Pune. He also works in reputed colleges in Pune as Guest Lecturer.

Corporate Training conducted for clients including Indian Statistical Institute, John Deer, Red Hat, Vodafone etc.

Admission Process

Prerequisite

Bachelor's Degree with a **minimum of 50%** aggregate marks at graduation or equivalent

Nice to Have, But not mandatory

Comfort with at **programming language**
Familiarity with college-level **mathematics and statistics**

Student Testimonials

I completed Data Science course from Sachin Sir, the highlight of the course is the curriculum and projects. It covered everything from stats, linear algebra till neural networks and related projects. I am happy that I chose iAnalyst… Technology Architect @ Cognizant

Each of the lectures were effective mix of information, demo, real-life situations, and PPTs.... The concepts discussed in this class will stick in my mind because of the unique way in which they were discussed and practiced the number of examples..... After taking this class, I feel that the field of Data Science is AWESOME and having bright future. I want to thank you for this. - Product Development Specialist @FIS Pune

Clear & deep concept about Statistics & Machine Learning along with R & Python. Cool & co-operative behaviour of Sachin. It helped me to understand the deep concept of machine learning & statistics. I enjoyed It and build my career in Data science … Employee from Vodafone

Detailed Program CURRICULUM

1. R

Installation on Windows, Mac, Linux
Rstudio, Rmarkdown, RShiny
Packages/Libraries, Installing libraries
Data Types, Vectors, Matrices, Lists, Data Frames, Factors
Default functions: Numeric, Character, String, Date Time, Arithmetic, R Pipe Operator
Loops/statements: for loop, if, if-else and nested if,
while, Select, Break, Continue, Repeat, Return
File Handling, Data files (Create, Import, Export)
Vector Indexing, Slicing
Apply Functions
Creating User Defined functions
Default Datasets in R
Additional topics on R

2. Python

Course Overview
Intro to Python
Python Data Types
Control Statements
Loops, Functions, Classes
Tuples, lists, dictionaries, and sets.
Read/Write different types of Files
Modules and Libraries.
Range()
List comprehension
Lambda Expression
Map and filter
Methods
Important libraries for DS (NumPy,
SciPy, Pandas, IPython, Matplotlib,
Seaborn etc.)
Accessing SQL, NOSQL databases

3. Linear Algebra

Vector and Matrices
Understanding vector and matrices, addition, subtraction and product of vectors and matrices, Metrics Operations – Dot Product, Determinant and Inverse of a matrix, Metrics Decomposition. Eigen Values and Eigen Vectors, real life applications
Functions - types of functions, Error in predicting numeric value as a function
Distance Measurements – Euclidean, Minkowaski, Manhattan,
Derivatives/Differentiation - Meaning of differentiation, differentiating functions and visualization
Gradient Descent, Maxima, Minima, Local and Global Maxima and Minima, Optimization Functions
Integration - Meaning of Integration, Integrating functions, Use of integration in probability theory.

4. Statistics & Probability

Introduction to Statistics, What is Statistics? Branches of Statistics, Population and Sample, Types of measurements & Data, Classification and Tabulation, Frequency Distribution, Diagrammatic and graphical representation.

Descriptive Statistics: Introduction, Measures of Central Tendency (AM, Weighted Mean, GM, HM, Median, Mode), Measures of Partition (Quartiles, Deciles, Percentiles),Measures of Variation (Range, Mean Deviation about Mean, Median ,Mode; Quartile Deviation – Inter Quartile Ranges, Standard Deviation, Coefficient of Variation), Skewness and Kurtosis

Probability: Permutation and combination, Addition and Multiplication Theorem, Random Variable, Independent Variable, Joint & Conditional probability, Bayes Theorem and application of Bayes theorem

Random Variables & Probability: Probability Mass Function (PMF),Probability Density Function (PDF),Cumulative Distribution Function (CDF), Mathematical Expectation, Applications
Probability Distributions: Discrete and Continuous distributions, Bernoulli Trials and Distribution, Binomial, Poisson, Uniform, Normal, Applications,

Sampling Theory: Concept, Types of Sampling, Sampling Methods,(SRS, Stratified, Systematic, 2-Stage, multi-stage and Cluster) Sampling Distribution (Chi-Square, t, F), Sample size, Central Limit Theorem
Estimation: Point Estimation, Interval Estimation / Confidence Intervals

Testing of Hypothesis: Statistical Hypothesis, Null / alternate Level of significance, Testing procedure, Hypothesis Testing for Proportions in Population parameters, p-value, Z tests, t tests, Chi Square Test ,F tests

Non-parametric tests – Sign Test, Wilcoxon signed-rank test, Mann-Whitney U Test, Kruskal–Wallis test / Non Parametric ANOVA
Hypothesis Test Errors
A/B Testing, ANOVA, Two Way ANOVA
Correlation and Regression: Karl Pearson Correlation Coefficient, Spearman's Rank Correlation.
Regression Analysis: Introduction, Simple Linear Regression, Interpretation of Regression analysis results (Hypothesis related to coefficients, p value, R square, Adj R square)

5. Data preparation

Data type conversion, Binarization and Categorization of Numeric variables
Missing value identification and imputing
Outlier detection and removing.

6. Data Visualization

Summarize and visualize different types of data by appropriate chart/graph, Bar chart, Pie chart, histogram, density plot, boxplot, scatterplot.
Advanced data visualization and Interactive plots

7. Feature Engineering, Dimensionality Reduction Basics

Min Max, Normal, Binary, Range, Box-Cox, Standardise
PCA, TSNE, Singular Value Decomposition, AIC – BIC, Multi Collinearity

8. Machine Learning

Machine Learning Model Development Basics
Training, Testing, Cross Validation, Bootstrap Resampling
Curse of Dimensionality
Classification vs Regression
Supervised Learning Introduction
Unsupervised Learning Introduction
Regularization L1, L2, Bias Variance
Loss Functions – Hinge, 0-1, Logistic Loss
Grid Search, Random Search, Parameter Tuning, Saddle Point
Model Evaluation: R square, Adj R square, S2, AIC, BIC, Misclassification, ROC, AUC, MSE, RMSE, MAE etc.
Multiclass, Cat/Numeric, Outliers, Bias-Variance Trade-off, Feature Importance, Imbalanced Data, Handling New Features, Large Dimensions, Best / Worst Cases, Space and Time Complexity

9. Supervised Learning – Classification and Regression Techniques

Logistic Regression
Discriminant Analysis
Decision Trees
Support Vector Machine
Naïve Bayes Classifier
K-NN
Linear Regression, Polynomial Regression

10. Unsupervised Learning

Association Rules – APRIORI, Elcat, FP Tree,
Market Basket Analysis
Lift, Support and Confidence
Extracting, Visualization and Interpreting Rules
Clustering
Hierarchical
Non Hierarchical - K Means, K Medoids,
Density Based Clustering
Ensembles: Stacking, Bagging, Boosting, Random Forest, xgboost, Adaboost

11. Neural Networks

Artificial Neural Network, Convolutional and Recurrent Neural Network
Tensorflow, Keras Installation
Perceptron, Multi-Layer Perception, Activation Function, Vanishing – Exploding Gradient Problem
Dropout Regularization, Forward and Backward Propagation, ANN for Regression / Classification
Weights Initialization, Batch Normalization – Adagrad, Ada Delta, Adam, RMS Prop
Auto Encoders - Variation, Denoising, Sparse, Restricted Boltzmann

12. Time Series Analysis

Introduction to Time Series
Components of Time Series
Fitting and Forecasting, Moving Average (MA) model, Exponential Smoothing model, Auto Regressive (AR) model, ARMA and ARIMA models

13. Text Mining

Text Indexing, Inverted Index, Bookean Query Processing, Phrase Query, Proximity Query, Relevance Ranking, Page Analysis – Page Ranks

14. Natural Language Processing

NTLK, scikit-learn, keras for NLP – Basics, Installation
TF-IDF, BoW, Uni-Bi-n Grams
Text Processing – Stemming, Stop Words, Tokenization, Lemmetization, Word2Vec

15. Additional Topics

Markov Chain Model, Hidden Markov Models
Gaussian Mixure Model