GAMAKA
Artificial Intelligence Solutions

Office No 309, Paranjape – The Business Hub, Karve Road, Kothrud, Pune - 411038
Email: itplacements@gamakaai.com Cell: 91-7378483656. WhatsApp: 91-7378493293
www.gamakaai.com

List of Projects/Case Studies

1. Forecasting Stock and Commodity Prices
2. Analysing Wine Types and Quality
3. Build your own image recognition model with TensorFlow
4. Customer Segmentation and Effective Cross Selling
5. Predict fraud with data visualization & predictive modelling
6. Analysing Movie Reviews Sentiment
7. Analysing Music Trends and Recommendations
8. Spam Detection
9. Build your own Recommendation System
10. Build your own Python predictive modelling, regression analysis & machine learning Model
11. Football Players (Estimating Population Mean from a Sample)
12. Election Polling (Estimating Population Proportion from a Sample)
13. A Medical Study (Hypothesis Test for the Population Mean)
    Employee Behavior (Hypothesis Test for the Population Proportion)
14. A/B Testing (Comparing the means of two populations
15. Customer Analysis (Comparing the proportions of 2 populations)
16. Predictive medicine: prognosis and diagnostic accuracy
17. Virtual assistance for patients and customer support
18. Creation of drugs - allows choosing, which experiments should be done and incorporates all the new information in a continuous learning loop
19. Clustering algorithms for customer segmentation
20. Discovering similarities across my Spotify music using data, clustering and visualization
21. An End-to-End Project on Time Series Analysis and Forecasting with Python
22. Using LSTMs to forecast time-series
23. Evolution of a salesman: A complete genetic algorithm tutorial for Python
24. A Machine Learning Approach — Building a Hotel Recommendation Engine
25. How To Create Data Products That Are Magical Using Sequence-to-Sequence Models

GAMAKA
Artificial Intelligence Solutions

Office No 309, Paranjape – The Business Hub, Karve Road, Kothrud, Pune - 411038
Email: itplacements@gamakaai.com Cell: 91-7378483656. WhatsApp: 91-7378493293
www.gamakaai.com

Module 1: Data Science – Python

Duration: 90 Hours with hands on tutorials

12 Case Studies with Internship

Mode of training: Online/Offline(classroom)/Combo (Offline & Online)

Python Environment Setup and Essentials Hadoop Fundamentals

- Anaconda Python Distribution – Windows, Mac OS, Linux
- Jupyter Notebook Installation
- Variable Assignment
- Understanding Data Types: Integer, Float, String, None, Boolean, Typecasting
- Tuples: Create, Access, and Slice
- Dicts: Create, View, Access, and Modify
- Studying Basic Operations: 'in', '+', '*

Computing with Python – NumPy and SciPy

- Mathematical Computing with Python - NumPy
- Understanding NumPy
- ndarray: Purpose, Properties, Types
- ndarray: Class and Attributes
- How to Access Array Elements?
- Indexing, Slicing, Iteration, Indexing with Boolean Arrays
- Studying Universal Functions
- What is Shape Manipulation?
- Linear Algebra
- Scientific Computing with Python – SciPy
- Understanding SciPy
- Studying SciPy Sub-packages
- Sub-Packages: Integration and Optimize
- Sub-Packages: Statistics, Weave, I O
- Linear Algebra

Data Manipulation with Python

- Data Manipulation and Machine Learning with Python
- Data Manipulation with Python – Pandas
- Understanding Pandas
- Defining Data Structures
- Data Operations and Data Standardization
- Pandas: File Read and Write Support
- SQL Operation
- Machine Learning with Python – Scikit
- Natural Language Processing with Scikit
- NLP Environment Setup & Applications
- NLP Sentence Analysis & Libraries
- Scikit – Built-in Modules & Feature Extraction
- Scikit – Grid Search & Parameters

Fundamentals of Machine Learning

- Overview & Terminologies
- What is Machine Learning?
- Why Learn?
- When is Learning required?
- Data Mining
- Application Areas and Roles
- Types of Machine Learning
- Supervised Learning
- Unsupervised Learning
- Reinforcement learning

GAMAKA
Artificial Intelligence Solutions

Office No 309, Paranjape – The Business Hub, Karve Road, Kothrud, Pune - 411038
Email: itplacements@gamakaai.com Cell: 91-7378483656. WhatsApp: 91-7378493293
www.gamakaai.com

Machine Learning Concepts & Terminologies

- Steps in developing a Machine Learning application
- Key tasks of Machine Learning
- Modelling Terminologies
- Learning a Class from Examples
- Probability and Inference
- PAC (Probably Approximately Correct) Learning
- Noise
- Noise and Model Complexity
- Triple Trade-Off
- Association Rules
- Association Measures
- Sample Algorithms

Simple Linear Regression

- Correlation
- Regression
- Model Assumptions
- Estimation Process
- Least Squares Method
- The Coefficient of Determination
- Correlation and Regression
- Simple Linear Regression Assignments

Multiple Regression Analysis

- Introduction
- Design Requirements
- Assumptions
- Independence
- Normality, Homoscedasticity, Linearity
- Multiple Regression
- Formal Statement of the Model
- Estimating parameters of the model
- F-test for the overall fit of the model
- Multiple regression model Building
- Selecting the best Regression equation
- Examples/Use Cases
- Interpreting the Final Model
- Multicollinearity and its Diagnostics
- Examples/Use Cases
- Qualitative Independent Variables
- Indicator variables
- Interpretation of Regression Coefficients
- Examples/Use Cases
- Regression Diagnostics and Residual Analysis
- Multiple Linear Regression Using R & Python
- Multiple Regression Assignment

Logistic Regression Analysis

- Theory Behind Logistic Regression
- Assessing the Model and Predictors
- When and Why do we Use Logistic Regression?
- Binary
- Multinomial
- Interpreting Logistic Regression
- Sample size requirements
- he logistic function & Interpretation
- Methods for including variables
- Computational method

Maximum Likelihood Estimation

- Bernoulli distribution
- Multinomial distribution
- Gaussian distribution
- Assessing the Model
- Assessing Changes in Models
- Assessing Predictors
- Methods of Regression
- Complete Separation
- Overdispersion
- MLE using Python

Office No 309, Paranjape – The Business Hub, Karve Road, Kothrud, Pune - 411038
Email: itplacements@gamakaai.com Cell: 91-7378483656. WhatsApp: 91-7378493293
www.gamakaai.com

Decision Trees

- Understanding the Concept
- Internal decision nodes
- Terminal leaves.
- Tree induction: Construction of the tree
- Classification Trees
- Entropy
- Selecting Attribute
- Information Gain
- Partially learned tree
- Overfitting
- Causes for over fitting
- Overfitting Prevention (Pruning) Methods
- Reduced Error Pruning
- Decision trees - Advantages & Drawbacks
- Ensemble Models

Random Forests

- Introduction & Motivation
- Ensemble Methods - Bagging, Boosting & Random Forests
- Ensemble Classifiers
- Ensemble Models
- How random forests work?
- Gini Index
- Operation of Random Forest
- Random forest algorithm
- Common variables for random forests
- Random Forest – practical consideration
- Random Forest – Features, Advantages and Disadvantages
- Limitations of random forest
- Random Forest using Python

Support Vector Machine

- Problem Definition
- Separating Hyperplanes
- Linear separable case
- Formula for the Margin
- Finding the optimal hyperplane
- The optimization problem
- The Legrangin Dual Problem
- Importance of the Support Vectors
- VC dimension
- Non-linear SVM
- Mapping the data to higher dimension
- The Kernel Trick
- Important Kernel Issues
- Soft Margin
- The primal optimization problem
- The Dual Formulation
- The “C” Problem: Overfitting and Underfitting
- Model selection procedure
- SVM For Multi-class classification
- Applications of SVM
- Advantages & Drawbacks of SVM

Bayesian Theory

- Axioms of Probability Theory
- Conditional Probability
- Independence
- Joint Distribution
- Baye’s Rule
- Bayesian Categorization
- Generative Probabilistic Models
- Naïve Bayes Generative Model
- Naïve Bayesian Categorization
- Example & Exercises
- Naïve Bayes Classifier using Python

K-Nearest Neighbor (K-NN)

- Non-parametric methods
- How to Choose k or h

Office No 309, Paranjape – The Business Hub, Karve Road, Kothrud, Pune - 411038
Email: itplacements@gamakaai.com Cell: 91-7378483656. WhatsApp: 91-7378493293
www.gamakaai.com

- k-Nearest Neighbor Estimator
- Strengths and Weaknesses

Boosting

- Gradient Boosting
- ADA Boost

Dimensionality Reduction

- Principal Components Analysis (PCA)
- Singular Value Decomposition (SVD)
- Latent Dirichlet Analysis (LDA)
- Latent Dirichlet Analysis (LDA

K Means Clustering

- Parametric Methods Recap
- Clustering
- Direct Clustering Method
- Mixture densities
- Classes v/s Clusters
- Non-Hierarchical Clustering
- K-Means
- Distance Metrics
- K-Means Algorithm
- K-Means Objective
- Color Quantization
- Vector Quantization
- Encoding/Decoding
- Soft Clustering
- Expectation Maximization (EM)
- EM Algorithm
- Feature Selection vs Extraction
- Seed Choice
- Uses of Clustering
- Clustering as Pre-processing

Time Series

- The Art of Forecasting
- Forecasting Approaches
- Qualitative Forecasting Methods
- Quantitative Forecasting Methods
- Time Series & its Components
- Trend
- Cyclical
- Seasonal
- Irregular
- Smoothing Methods
- Moving Average Method
- Exponential Smoothing Method
- Forecast Effect of Smoothing Coefficient
- Linear Time-Series Forecasting Model
- Forecast using Trend Models
- The Linear Trend Model
- Time Series Plot
- Seasonality Plot
- Trend Analysis
- Quadratic Time-Series Forecasting Model

Quadratic Time-Series Model Relationships

- Quadratic Trend Model
- Exponential Time-Series Forecasting Model
- Exponential Weight
- Exponential Trend Model
- Autoregressive Modeling
- Time Series Data Plot
- Auto-correlation Plot
- Evaluating Forecasts
- Quantitative Forecasting Steps
- Forecasting Guidelines
- Pattern of Forecast Error
- Residual Analysis

GAMAKA
Artificial Intelligence Solutions

Office No 309, Paranjape – The Business Hub, Karve Road, Kothrud, Pune - 411038
Email: itplacements@gamakaai.com Cell: 91-7378483656. WhatsApp: 91-7378493293
www.gamakaai.com

TensorFlow for Deep Learning with Python

- Understand how Neural Networks Work
- Build your own Neural Network from Scratch with Python
- Use TensorFlow for Classification and Regression Tasks
- Use TensorFlow for Image Classification with Convolutional Neural Networks
- Use TensorFlow for Time Series Analysis with Recurrent Neural Networks
- Use TensorFlow for solving Unsupervised Learning Problems with AutoEncoders
- Learn how to conduct Reinforcement Learning with OpenAI Gym
- Create Generative Adversarial Networks with TensorFlow

Data Visualization and Web Scraping

- Data Visualization and Matplotlib
- Python Libraries
- Features of Matplotlib
- Line Properties Plot with (x, y)
- Set Axis, Labels, and Legend Properties
- Alpha and Annotation
- Multiple Plots and SubPlots
- Python Web Scraping and Data Science
- The Parser
- Searching & Modifying the Tree
- Printing, Formatting, Encoding