



## Data Science & AI

**Regd. Off.** Plot A-40, Unit No.  
207-301, 2nd & 3rd Floor,  
I-Thum Tower A, Sector-62, Noida 201309  
Nearest Metro Station Noida Electronic City

**Website:** [www.trainingbasket.in](http://www.trainingbasket.in)  
**E-mail:** [support@trainingbasket.in](mailto:support@trainingbasket.in)  
**Helpline:** +91-9015887887

## PYTHON REFRESHER

### Introduction

- Python, Anaconda and relevant packages installations
- Structure of Python Program (Comments, Indentation)
- Variables, Keywords and Data types in Python
- Standard Input and Output
- Operators
- Numbers/ Strings

### Basic Data Structures

- Mutable /Immutable
- Lists/ Tuples
- Sets/ Dictionaries
- List Comprehensions

### Control Statements

- If, Else, Elif statements
  - While and For Loop
  - Break and Continue
  - Nested Loops
  - Pattern Making

### Functions

- Basic Functions
- Built-In
- User Defined Functions
- \*Args/ \*\*Kwargs

### Advance functions

- Maps/ Filters/ Reduce
- Closures / Decorators
- Generators

### File Handling

### Exception Handling

### Modules and Packages

- Web Framework(Django)
- Database Handling

**Object Oriented Programming**  
**Multithreading / Multiprocessing**  
**Advance DataStructures**

- Linked Lists
- Doubly Linked Lists
- Stacks
- Queues/ Dequeues

## **NUMPY, PANDAS, STATISTICS & MATHEMATICA**

**Numerical Python (Numpy)**

- Indexing/ Slicing
- Broadcasting
- Appending/ Inserting on Axis
- Mathematical and Statistical operations
- Sort/ Conditions
- Transpose operations
- Joining/ Splitting
- Linear Algebra

**Data Manipulation with Pandas**

- Data Extraction
- Series/ DataFrame Creations
- Indexing and Slicing
- Conditions/ Grouping/ Imputation
- Append/ Concat/ Merge/ Join
- DateTime Functionalities and Resampling
- Window Functions
- Excel functions

**Data Visualization**

- Customization of Matplotlib/ Seaborn
- Scatterplots/ Barplots/ Histograms/ Density Plots
- 3D plots
- Boxplotting and Outlier Detection
- Visualizing Linear Relationships
- Plotting with Pandas

**Introduction to Statistics**

- Probability
- Basics( Mutually Exclusive/ Joint Probability)
- Conditional Probability
- Dependent/ Independent

- Logs/ Odds
- Bayes Theorem
- Descriptive/ Inferential
- Mean, Median, Mode
- Variance/ Standard Deviation
- Co-variance/ Correlation (Pearson/ Spearman)
- Central Limit Theorem

### **Types of Distributions**

- PDF/ PMF/ CDF
- Uniform/ Normal/ Skewed Distributions
- Binomial/ Bernoulli Distribution
- Poisson/ Exponential Distributions

### **Hypothesis Testing**

- Null/ Alternative Hypothesis
- Z-test/ T-test/ Chi2-test
- p-value
- F-test/ Anova
- Scipy.Stats/ Statsmodels

### **Bayesian Statistics**

- Difference between Bayesian and Frequentist
- MCMC
- PyMC3

### **Mathematical**

- Linear Algebra
- Calculus

## **MACHINE LEARNING**

### **Introduction to Machine Learning**

- Difference Between AI, ML and DL
- Applications of Machine Learning
- Categorization of Machine Learning
  - Supervised / Unsupervised / Semi Supervised
  - Parametric vs Non Parametric
  - Geometric/ Rule Based/ Gaussian
  - Reinforcement Learning
- Flow Operation (Pipelining)

## Supervised (Regression/ Classification)

### Regression

- Linear Regression
- Polynomial Regression
- Lasso Regression
- Ridge Regression
- Stepwise Regression
- Bayesian Regression

### Classification

- Logistic Regression
- KNN
- SVM (Support Vector Machines)
- Decision Tree
- Naive Bayes
- LDA
- Classification for Imbalanced Dataset

## Unsupervised Learning

- K-means
- PCA
- Hierarchical Clustering
  - Agglomerative (Bottom-Top)
  - Divisive (Top-Bottom)
- DBSCAN/ HDBSCAN

## Scikit Learn

- Introduction to SciKit Learn (sklearn)
- Sample Dataset in SciKit Learn
- Artificial Generation of Dataset

## EDA and DataWrangling

- Null Values Imputation
- Outlier Detection
- Univariate/ Bivariate/ Multivariate Analysis
- Encoding
  - Label Encoder
  - Ordinal Encoding
  - One Hot Encoding
- Scaling
  - Binarizer
  - MinMaxScaling
  - Normalizer
  - StandardScaler



### Feature Selection and Dimensional Reduction

- Feature Selection
  - Filter Methods
  - Wrapper Methods
  - Embedding(Ridge /Lasso)
  - Ensembling
- Dimensional Reduction
  - Factor Analysis/ t-SNE/ UMAP
  - PCA/ ICA/ ISOMAP
  - AutoEncoders

### Modelling

- Train/Test Split
- Assumptions
- Basic Modeling
- Under fitting, Over fitting, Bias and Variance
- Loss Functions
- cross validation

### Cross-Validation and HyperParameter Tuning

- Holdout Validation
- K-fold cross Validation
- Stratified Kfold
- Cross\_val\_score
- GridSearchCV
- RandomizedSearchCV

### Evaluation Metrics and Improvement Techniques

- MSE/ MAE/ R2/ Adjusted R2
- Accuracy measurement
- Confusion Matrix
- Precision/ Sensitivity/ Specificity/ F1Score
- AUC/ ROC
- AIC and BIC

### Criteria to Select Models

- Bias vs Variance Trade offs
- Evaluation Metrics
- Cross-Validation

### Linear Regression

- Introduction to Linear Regression
- Understanding the real meaning of Linear Regression

- Multiple Linear Regression and Non-linear Regression
- Cost Function (Sum of Square Error)
- Gradient Descent based approach

### **Polynomial Regression**

- Introduction to Polynomial Regression
- When to use Polynomial regression
- Evaluation based on RMSE/ R<sup>2</sup>
- Hands-on on Iris dataset

### **Regularization with Lasso/Ridge Regression**

- Problems with Large Features
- Why penalty is inducted
- Difference between L1 and L2
- Cost Function

### **Step Regression**

- Statistical Significance
- Use of t-Test
- explanatory variables
- Forward Selection
- Backward Selection Rule

### **Logistic Regression:**

- Logistic regression vs Linear Regression
- Threshold Value
- Hands on experience titanic dataset

### **K Nearest Neighbours:**

- Introduction to KNN algorithm
- Decision boundary KNN Vs Decision tree
- What is the best K
- KNN Problems
- Feature selection using KNNs
- KNN using sklearn
- Digits classification using KNN in Python
- IRIS dataset classification using KNN

### **Support Vector Machine (SVM):**

- Linear Classifiers
- Margin of SVM's
- SVM optimization
- SVM for Data which is not linear separable



- Learning non-linear patterns
- Kernel Trick
- SVM Parameter Tuning
- Linear SVM using Python
- SVM with RBF kernel with Python
- Hinge Loss

### **Decision Tree:**

- Introduction to Decision tree
- Details of tree induction
- GINI index computation
- Entropy and information gain
- Pruning
- Metrics for performance Evaluation
- Iris Decision Tree Example

### **Naive Bayes**

- Bayes Theorem
- Multinomial
- Burnoulli
- Gaussian

### **Ensembling:**

- Voting/ Averaging
- Bagging/ Boosting
- Random Forest
- AdaBoost
- Gradient Boosting
- XGB

### **Random Forest:**

- Comparison between Random Forest and Decision Tree
- Feature Importance
- Use Cases

### **Time Series:**

- Stationarity
- AutoRegression/ AutoCorrelation
- ACF vs PACF Plots
- Smoothing Time Series
- Dicky Fuller Test
- Time Series Decomposition
- Modelling and Forecasting
- AR/ MA/ ARIMA/ SARIMA



## Unsupervised Learning:

### K Means:

- Applications of Clustering
- Understanding Euclidean Distance
- Basics of Clustering
- Elbow Method
- Hierarchical clustering
- K-means Algorithm example

### Dimensionality Reduction using PCA:

- What is PCA?
- Understanding Matrix Transformations
- Eigen Values and Eigen Vectors

### Hierarchical Clustering

- Agglomerative
- Devisive

### DBSCAN

- Reachability
- Connectivity
- epsilon and r

### Gaussian Mixed Models

## DEEP LEARNING

### Artificial Neural Networks In Python

- Perceptron and relate it with Logistic Regression
- Multiple layer Neural network
- Similarities and Differences with Basic ML
- Forward Propagation
- Back Propagation Algorithm
- Vanishing Gradient and Exploding Gradient

### Activation Functions:

- Non Linearity
- Sigmoid / Tanh Function
- Relu / Leaky Relu / Gelu
- Softmax Functiona

## Optimizers

- Gradient Descent
- Stochastic Gradient Descent
- Momentum
- AdaGrad
- RMSProp
- NAG
- Adam/ Nadam

## New Strategies for Optimizing

- Shuffling and Curriculum Learning
- Batch Normalization
- Early stopping

## Tensorflow:

- Tensors
- Session
- Placeholders and Variables
- Hands on with Tensorflow
- Graphs

## Keras:

- Purpose of Keras
- Sequential vs Functional
- Model Creating
- Using while Regression and Classification

## Pytorch

- Tensors
- Autograd
- Graphs
- Pytorch.nn
- Control Flow and Weight Sharing

## OpenCV

- Read/ Write Images
- Gray to BGR
- Filter2D
- Scaling/ Rotation
- Laplace Transformation



### Types of Networks

- Feed Forward Networks
- Fully Connected Networks
- Recurrent Neural Networks
- Convolutional Networks
- RBM
- Deep Belief Networks

### CNN(Convolutional Neural Network)

- Convolution/ Filters/ Pooling
- Back Propagation in CNN
- Masking and Rol
- Types Of CNN
- RCNN, FastRCNN, YOLO

### Architectures

- LeNet/ Alexnet
- VGG 16/ 19
- ResNet
- MobileNet

### Recurrent Neural Networks

- Classical RNN
- LSTM/GRU
- Vanishing Gradient
- Exploding Gradient
- Bidirectional RNN

### Transfer Learning

- ImageNet
- Need of Transfer Learning
- Freezing of Layers
- Reusing of Structure

### AutoEncoders

- Encoder vs Decoder
- Difference with PCA
- KL Divergence
- Variable Auto Encoders

### GANs

- Generators
- Discriminator
- GAN Structure
- GAN Training

### NLP Natural Learning Process

- WordEmbedding
- Frequency vs Production based embedding
- Count Vector/ TFIDF/ Co Occurrence
- Bag of Words / Skip Gram
- Word2Vec / GloVe

### NLTK

- Morphological/ Lexical Analysis
- Syntactic and Semantic Analysis
- Tagging

### Attention Mechanism

- Soft and Hard Attention
- Local and Global attention
- Monotonic Alignment and Predictive Alignment
- Multi headed Attention

### Miscellaneous

- Semi Supervised Learning
- Transformers/Bert/GPT3
- Graphs and GNN

## REINFORCEMENT LEARNING AND ARTIFICIAL SUPER INTELLIGENCE

### Element of Reinforcement Learning

- Limitation and Scope
- Environment
- Agent, State, Action
- Reward and Punishment
- Model based vs Model free
- On Policy vs Off Policy
- Policy based vs Value based
- Greedy policy
- Bellman Equation
- Tabular/ Approximate Solutions

### OpenAI/Gym

- Environments
- Observations
- Spaces
- Cart Pole Environments

### Dynamic Programming

- Policy Evaluation/ Improvement/ Iteration
- Value Iteration
- Asynchronous Dynamic Programming

### Markov Decision Process

- Goals and rewards
- Returns and Episodes
- Optimal Policy/ Value

### Monte Carlo Methods

- Prediction of Actions
- Importance Sampling
- Discounting
- Pre-rewards

### Temporal Difference

- Advantages of TD
- Q Learning
- Sarsa
- DQN (Deep Q Learning)
- Maximization Bias and Double Learning

### Miscellaneous

- Raspberry Pi
- Carla
- Genetic Algorithms
- PSO
- Ant Algorithms



**Q: Why should I choose Training Basket over other training providers?**

Ans. Training Basket provides a unique amalgamation of quality, convenience, flexibility and cost. Training Basket has some of the best trainers in the industry. Our trainers excel not just in depth and width of knowledge, but also in their patience and ability to explain difficult concepts in simple terms.

Training Basket has made serious investments with long-term vision for ensuring good environmental factor for studies by first buying their own suites in the prestigious iThum Towers in NOIDA, Sec-62. This lets us configure our labs and classroom suited best to our student's comfort and focused studies.

**Q: What is the criterion for availing the Training Basket job assistance program?**

Ans: All Training Basket students who have successfully completed their training in any of our courses are directly eligible for placement assistance.

**Q: Which are the companies that Training Basket has placed students in the past?**

Ans: We have exclusive tie ups with MNC's like Ericsson, Cisco, Cognizant, Tech Mahindra, MEON, Bingo, Genpact etc.

**Q: Do I need a prior industry experience in getting an interview opportunity?**

Ans: There is no need to have prior opportunity for getting an interview call. The successful completion of any industry level technology training at Training Basket is like an industry experience. This training makes you confident to clear interviews and we also conduct in-house mock interviews on our online assessment platform where we assess our student's skills by testing their code online or industry specific assessment before sending them for interviews.

**Q: How does Training Basket assist in placement?**

Ans: You will be guided on creating an attractive template based resume. You will get opportunity to attend free personality development program and mock interviews conducted by our SME's to boost your confidence for real interviews. Plus you will be given our level assessment platform where we assess our student's skills by testing their code online or industry specific assessment.

**Q: If I don't clear in first attempt, will I get another chance?**

Ans: Yes, for sure. Your resume will be active on our job portal and will be visible to all our associates and clients. Training Basket will continue to send your resume to future job requirements matching your profile till you land a job.

**Q: Does Training Basket Guarantee job through it's job assistance program ?**

Ans: Training Basket does not guarantee job placement but it will continue to assist you on best efforts basis to place you in it's affiliated companies' network.

# COMPANIES WHERE OUR STUDENTS ARE PLACED

**Tech  
Mahindra**



**Source Soft Solutions**  
We only believe in the best

**Chi Networks™**

**ADS-eVER**

**MEON**  
DELIVERING GREAT SURFACES

**BioMax™**



**Bingo**  
Change • Innovate • Lead

**Dimensions**  
Five Dimensions Infotech

**WTS Webtech Solutions**

**HCL**

**IBM**

**Infosys**

**accenture**

**HUAWEI**

**Reliance  
Jio Infocomm**

**pinga™ solutions**

**DATA BRIDGE**  
MARKET RESEARCH

**CoreIP**  
Technology Pvt Ltd

**iSOURCE**

**network  
solutions**



**Shailers**  
Solutions...for your need



**SAMSUNG**

**indiatech**  
Solutions