



2nd Floor, Sri Sai Arcade, Beside Aditya Trade Centre, Ameerpet, Hyderabad. [www.sathyatech.com](http://www.sathyatech.com)  
Ph : 9100920092 / 9100940094 | E-mail : [info@sathyatech.com](mailto:info@sathyatech.com) | [www.sathyatech.com](http://www.sathyatech.com)

## DATA SCIENCE (MACHINE LEARNING, DEEP LEARNING & AI)

### INTRODUCTION TO DATA SCIENCE

- Need of Data Science
- History of Data Science
- What is Data Science
- Data Science vs Data Analytics
- What is Data Analytics
- What is Data Analysis
- Data Mining
- Decision Making
- Loops
- Numbers
- Strings
- Functions
- Date & Time
- Exception Handling

### INTRODUCTION TO MACHINE LEARNING

- What is machine learning
- Types of learning
- Supervised Machine Learning
- Unsupervised Machine Learning
- Machine learning algorithms
- Flow of Supervised and Unsupervised Machine Learning
- Simple Linear Regression
- Multiple Linear Regression
- Logistic Regression
- K-Nearest Neighbour
- Support Vector Machine
- Decision Tree
- Random Forest
- Ensemble Machine Learning
- Naïve Bayes
- Clustering
- K-Means
- Hierarchical Clustering

### PYTHON

- Basics
- Overview
- Environment Setup
- Syntax
- Variables
- Data Types
- Type Conversion
- Operators

### Intermediate

- Lists
- List Manipulation
- List Methods
- Packages
- Tuples
- Dictionary Manipulation

### Data Science Essentials

- Numpy
- Introduction
- Numpy Package
- Numpy Object
- Data Types
- Array Attributes
- Array from Numerical Ranges
- Indexing & Slicing
- Advanced Indexing
- Iterating over array
- Array manipulation
- String Functions
- Arithmetic Operations
- Statistical Functions

### Pandas

- Introduction
- Pandas Package
- Series
- DataFrame
- Panel
- Descriptive Statistics
- Indexing and Selecting Data
- Iteration

- Sorting
- Aggregations
- Missing Data
- GroupBy
- Merging/Joining
- Concatenation
- Date Functionality
- Pandas – Visualization
- Pandas – IO Tools
- CSV to DataFrame
- Loc and iloc
- DataFrame Filtering

### Manipulating DataFrames with Pandas

- Extracting and Transforming Data
- Reshaping Data
- Grouping Data

### Data Visualization using Python

- Matplotlib
- Bar Graph
- Histogram
- Scatter Plot
- Pie Chart

### Statistics and Mathematical Essentials for Data Science

- Measure of Central Tendency
- Mean
- Mode
- Median
- Range
- Inter Quartile Range
- Variance
- Standard Deviation
- Correlation
- Regression Models in Machine Learning
- Residuals
- Correlation Coefficients (Pearson)
- Accuracy Measurement
- Least Square Regression
- Root Mean Square Error
- Coefficient of Determination (R2 Score)
- Cost Function
- Gradient Descent
- Hypothesis Testing and p-values

- T-values
- Z-score
- Create Dummy Variables
- Cross Validation
- Confusion Matrix
- Compute Precision, Recall, F-Measure and Support
- TPR, FPR, FNR, TNR
- Accuracy
- Learning rate
- Sensitivity and Specificity
- ROC Curve (Receiver Operating Characteristic)
- Receiver Operating Characteristic (ROC) curves
- Area under the Curve (AUC)
- Calculating similarity based on Euclidean/Manhattan Distance
- Calculation of Entropy and Information Gain
- Calculation of Gini Index
- Basic Probability
- Randomness
- Conditional Probability
- Naïve Bayes Theorem
- Multiplication rule for dependent and independent events
- Differential Equations and Partial Derivations
- Linear Algebra :
- Correlation, Covariance
- Matrices and Vectors
- Addition and Scalar Multiplication
- Matrix Vector Multiplication
- Matrices Multiplication
- Matrix Transformations
- Inverse and Transpose of Matrices
- Eigen Values and Eigen Vectors

### Machine Learning using Python

- Regression
- Linear Regression
- What is Regression
- Types of Regression
- Model Description
- Ordinary Least Square method
- Import and Read the Data

- Sorting
- Aggregations
- Missing Data
- GroupBy
- Merging/Joining
- Concatenation
- Date Functionality
- Pandas – Visualization
- Pandas – IO Tools
- CSV to DataFrame
- Loc and iloc
- DataFrame Filtering

### Manipulating DataFrames with Pandas

- Extracting and Transforming Data
- Reshaping Data
- Grouping Data

### Data Visualization using Python

- Matplotlib
- Bar Graph
- Histogram
- Scatter Plot
- Pie Chart

### Statistics and Mathematical Essentials for Data Science

- Measure of Central Tendency
- Mean
- Mode
- Median
- Range
- Inter Quartile Range
- Variance
- Standard Deviation
- Correlation
- Regression Models in Machine Learning
- Residuals
- Correlation Coefficients (Pearson)
- Accuracy Measurement
- Least Square Regression
- Root Mean Square Error
- Coefficient of Determination (R2 Score)
- Cost Function
- Gradient Descent
- Hypothesis Testing and p-values

- T-values
- Z-score
- Create Dummy Variables
- Cross Validation
- Confusion Matrix
- Compute Precision, Recall, F-Measure and Support
- TPR, FPR, FNR, TNR
- Accuracy
- Learning rate
- Sensitivity and Specificity
- ROC Curve (Receiver Operating Characteristic)
- Receiver Operating Characteristic (ROC) curves
- Area under the Curve (AUC)
- Calculating similarity based on Euclidean/Manhattan Distance
- Calculation of Entropy and Information Gain
- Calculation of Gini Index
- Basic Probability
- Randomness
- Conditional Probability
- Naïve Bayes Theorem
- Multiplication rule for dependent and independent events
- Differential Equations and Partial Derivations
- Linear Algebra:
- Correlation, Covariance
- Matrices and Vectors
- Addition and Scalar Multiplication
- Matrix Vector Multiplication
- Matrices Multiplication
- Matrix Transformations
- Inverse and Transpose of Matrices
- Eigen Values and Eigen Vectors

### Machine Learning using Python

- Regression
- Linear Regression
- What is Regression
- Types of Regression
- Model Description
- Ordinary Least Square method
- Import and Read the Data

- Perform Exploratory Data Analysis
- Interpreting Model Coefficients
- Feature Selection
- Training and Testing the data
- Model Evaluation Using Train/Test Split
- Training the model
- Predicting Test data
- Model Evaluation Metrics for Regression
- Use Case – Linear Regression using Advertising Dataset and Housing Dataset

### Logistic Regression

- Introduction
- Data Exploration
- Data Visualization
- Feature Selection (Recursive Feature Elimination)
- Implementing the Model
- Logistic Regression Model Fitting
- Predicting Test Set Results and Calculate Accuracy
- Cross Validation
- Confusion Matrix
- Compute Precision, Recall, F-Measure and Support
- ROC Curve (Receiver Operating Characteristic)
- Classification Report
- Logistic Regression Hypothesis
- Use Case – Logistic Regression using Banking dataset

### K-Nearest Neighbor

- Understanding classification using Nearest Neighbor
- Find K-Nearest Neighbors
- Rescale using min-max normalization
- Diagnosing cancer with K-NN algorithm
- Import/Load Data
- Exploring and Preparing the data
- Transformation – Normalizing numeric data
- Data preparation – creating training and test datasets
- Training a model on the data

- Evaluating model performance
- Improve model performance

### Support Vector Machine (SVM)

- Goal of Support Vector Machine (SVM)
- Support Vector Machine – Basics
- Advantages and Disadvantages of SVMs
- Hyperplane and Margin
- Classification with Hyperplanes
- Linear Separable Case
- Non-Separable Case
- Linear SVM
- Kernel and Radial Functions
- Constructing the Maximal Margin Classifier
- Use Case – SVM using cancer dataset

### Decision Tree and Random Forest

- Understanding decision trees
- Calculation of Entropy and Information Gain
- Choosing the best split
- Pruning the decision tree
- Collect data
- Exploring and preparing the data
- Training a model on the data
- Evaluating model performance
- Improving model performance
- Boosting the accuracy of decision trees
- What is Random Forest algorithm?
- Advantages of Random Forest algorithm
- Use Case – Decision Tree and Random Forest in Medicine

### PROBABILISTIC LEARNING – CLASSIFICATION USING NAÏVE BAYES

- Understanding naïve Bayes
- Basic concepts of Bayesian methods
- Probability
- Joint probability
- Conditional probability with Bayes' theorem
- The naïve Bayes algorithm
- The naïve Bayes classification
- Using numeric features with naïve

## Bayes

- Naïve Bayes algorithm Example
- Collecting data
- Exploring and preparing the data
- Training a model on the data
- Evaluating model performance
- Improving model performance

## FINDING GROUPS OF DATA- CLUSTERING WITH K-MEANS

- Understanding clustering
- Clustering as a machine learning task
- The K-means algorithm for clustering
- Using distance to assign and update cluster
- Choosing the appropriate number of cluster
- Finding segments using K-means clustering
- Collecting data
- Exploring and preparing the data
- Data preparation –dummy coding missing values
- Data preparing –imputing missing values
- Training a model on the data
- Evaluating model performance
- Improving model performance
- Principal Component Analysis (PCA)
- Dimensionality Reduction
- Use Case – KMeans Clustering using Wholesale Customers dataset

## DIMENSIONALITY REDUCTION AND VISUALIZATION

- What is Dimensionality reduction?
- Row Vector and Column Vector
- How to represent a data set?
- How to represent a dataset as a Matrix.
- Data Pre-processing: Feature Normalisation
- Mean of a data matrix
- Data Pre-processing: Column Standardization
- Co-variance of a Data Matrix

## PCA(PRINCIPAL COMPONENT ANALYSIS)

- Why learn PCA?
- Geometric intuition of PCA
- Mathematical objective function of PCA
- Eigen values and Eigen vectors (PCA): Dimensionality reduction
- PCA for Dimensionality Reduction and Visualization

## Deep Learning

- Introduction to Deep Learning
- Building
- Neural Networks Architecture
- Convolutional Neural Networks (CNN)

## Artificial Neural Networks (ANN)

- Deep Learning with Keras & Tensorflow
- Image Classification with Keras

## Artificial Intelligence

- Natural Language Processing
- Introduction to NLP and NLTK
- Preprocessing data using tokenization
- Stemming text data
- Converting text to its base form using lemmatization
- Building a bag-of-words model
- Building a text classifier
- Text to Features
- TF-IDF Extraction
- Word Vectors
- Analysing the sentiment of sentence

## Building Recommendation Engines

- What is Recommendation Engine
- Types of Recommendation Engines
- Collaborative Filtering
- Item Based Collaborative Filtering
- User Based Collaborative Filtering
- Content Based Filtering

## Optical Character Recognition

- Extraction of text from PDF
- Extraction of text from image

**For Online Classes visit us [www.sathyatech.com](http://www.sathyatech.com)**