

Data Science

Get Ready for Data Science Jobs

Python for machine learning

- Introduction to Python
- Different Applications where Python is Used
- Fundamentals of Python Programming
- Values, Types, Variables
- Conditional Statements
- The Companies using Python
- Operations and Expressions
- Loops
- Demonstrating Conditional Statements
- Demonstrating Loops
- Packages Installation
- Different Applications where Python is Used
- Conditional Statements
- Values, Types, Variables
- Conditional Statements
- Python Functions Practice
- NumPy Library - Creating NumPy Array
- Bar Graph, a Pie Chart to Show Information
- Pandas Library - Creating Series and Data
- Basic Functionalities of a Data Object
- Concatenation of Data Objects
- Exploring a Dataset
- Merging of Data objects
- Aggregation
- Analyzing a dataset
- Pandas Function

Statistics & Probability

- Descriptive Statistics and Inferential Statistics
- Sample and Population
- Variables and Data Types
- Percentiles
- Measure of Central Tendency
- Measures of Spread
- Skewness, Kurtosis
- Sampling Distribution
- Introduction to Probability, Discrete and Continuous Distributions
- Standard Probability Distribution Functions
- Bernoulli, Binomial-Distributions
- Descriptive Statistics and Inferential Statistics in Python
- Normal Distributions
- Test of Hypothesis - Z Test, T-Test and Chi-square Test
- Confidence Interval
- Variance, Covariance, Correlation
- Degrees of Freedom
- Annova Test
- Python Revision (Numpy, Pandas, Scikit Learn, Matplotlib)

Introduction to R Programming

- Introduction to R
- Data Types in R
- How To Install R & R Studio
- Data Structures in R
- Programming Statistical
- R-Packages
- Predictive Analysis in R

Machine Learning

- What is Machine Learning?
- Machine Learning Use-Cases
- Machine Learning Categories
- Machine Learning Process Flow
- Machine Learning types
- **Supervised Learning**
- Implementing Different Types of Supervised Learning Algorithms
- What are Classification and its Use Cases?

- Confusion Matrix
- Evaluating model output
- Implementation of Logistic regression
- What is Decision Tree?
- Algorithm for Decision Tree Induction
- Creating a Perfect Decision Tree
- What is a Random Forest?
- Advanced Methods in Machine Learning
- Dimensionality Reduction
- **Unsupervised Learning**
- What is Clustering & its Use Cases?
- What is K-means Clustering?
- How does the K-means algorithm work?
- Implementing K-means Clustering
- What is C-means Clustering?
- Implementation of Clustering - various types
- What is Hierarchical Clustering?
- Implementing Hierarchical Clustering
- How to do optimal clustering



Dimensionality Reduction

- Introduction to Dimensionality
- Why Dimensionality Reduction
- Implementing Dimensionality Reduction Technique
- PCA
- Scaling Dimensional Model
- Feature Engineering
- Regularization
- Under fitting Vs Overfitting
- Boosting Bagging and Random Forest
- Cross Validation

Model Selection and Boosting

- What is the Model Selection?
- The need for Model Selection
- Cross-Validation
- What is Boosting?
- How Boosting Algorithms work?

Association Rules Mining and Recommendation Systems

- What are Association Rules?
- Calculating Association Rule Parameters
- Association Rule Parameters
- How does Recommendation Engines work?
- Content-Based Filtering
- Collaborative Filtering
- Market Basket Analysis
- Apriori Algorithm

Time Series Analysis

- What is Time Series Analysis?
- Components of TSA
- Importance of TSA
- TSA in Python
- TSA Forecasting
- AR model
- MA model
- ARMA model
- ARIMA model
- Generating the ARIMA plot
- Stationarity
- Converting a non- stationary data to stationary
- ACF & PACF
- Plot ACF and PACF
- Implementing the Dickey-Fuller Test

Natural Language Processing

- Tokenization and Text Normalization
- Exercise: Tokenization and Text Normalization
- Exploring Text Data
- Part of Speech Tagging and Grammar Parsing
- Exercise: Part of Speech Tagging and Grammar Parsing
- Implementing Text Pre-processing Using NLTK
- Exercise: Implementing Text Pre-processing Using NLTK
- Natural Language Processing Techniques Using spaCy
- Introduction to Text Feature Engineering
- Count Vector, TFIDF Representations of Text
- Exercise: Introduction to Text Feature Engineering
- Understanding Vector Representation of Text
- Exercise: Understanding Vector Representation of Text

- Understanding Word Embeddings
- Word Embeddings in Action - Word2Vec
- Word Embeddings in Action - Glove
- Chat GPT Use case

Reinforcement Learning

- What is Reinforcement Learning
- Elements of Reinforcement Learning
- Why Reinforcement Learning
- Implement Reinforcement Learning using Python

Introduction to Deep Learning

- The deep in deep learning
- How deep learning works
- What deep learning achieved so far?
- Early neural network kernel methody
- What makes deep learning different
- Modern machine learning landscape,

Deep learning for computer vision

- Image formation
- Image representation
- Linear filtering
- Image sampling
- Edge detection
- Edges to Blobs and corner detection
- Image segmentation
- Human visual systems
- Feature matching
- Image description matching
- From traditional vision to computer vision
- Feedforward Neural network and back propagation
- Gradient Descent and variant
- Regularization in Neural Network
- Convolutional Neural Network
- Backpropagation in CNN
- Evolution of CNN architecture
- Recent CNN architecture
- CNN for object Detection
- CNN for segmentation

- Recurrent Neural Network
- Backpropagation RNNS
- LSTMs for time series

Data Visualization using (Qlik sense)

Optional Tools:

Snowflake (Basics)

Alteryx (Basics)

Data Iku (Basics)