



Industry 4.0 : Artificial Intelligence, Data Science & Machine Learning with Python

Recommended duration: 5 full days

Prerequisites:

Participants in this Machine Learning online course should have:

- Familiarity with the fundamentals of Python programming

Course Outline

1. Introduction to Artificial Intelligence and Machine Learning

- ❖ Artificial Intelligence
- ❖ What is Machine Learning?
- ❖ Machine Learning algorithms
- ❖ Supervised Versus Unsupervised Learning
- ❖ Machine Learning Algorithms
 - Regression
 - Classification
 - Clustering
- ❖ Applications of Machine Learning
- ❖ Machine learning examples
- ❖ Setting up Anaconda & Python Notebooks.
- ❖ Working on notebooks for Data Science

2. Techniques of Machine Learning

- ❖ Supervised learning
- ❖ Unsupervised learning

3. Mathematics & Statistics Refresher

- ❖ Concepts of linear algebra
- ❖ Euclidean and Non-Euclidean geometry
- ❖ Introduction to Calculus
- ❖ Probability and statistics

- ❖ Distributions, CDF, PDF
- ❖ Mean, Median, Mode
- ❖ Standard Deviation, quartiles, percentiles
- ❖ Variable Relationships & Estimation
- ❖ Hypothesis Testing

4. Accessing/Importing and Exporting Data

- ❖ Importing Data from various sources (Csv, txt, excel...etc) Viewing Data objects
- ❖ Exporting Data to various formats
- ❖ Important python modules: numpy, pandas, etc.

5. Data Manipulation - Cleaning-Munging

- ❖ Cleansing Data with Python
- ❖ Data Manipulation steps(sorting, filtering, duplicates, merging, appending, derived variables, sampling, Data type, conversions, renaming, formatting etc)
- ❖ Data manipulation tools(Operators, Functions, Packages, control structures, Loops, arrays etc)
- ❖ Python Built-in Functions (Text, numeric, date, utility functions)
- ❖ Python User Defined Functions
- ❖ Stripping out extraneous information
- ❖ Normalizing data
- ❖ Formatting data
- ❖ Important Python modules for data manipulation (Pandas, Numpy, re, math, string, datetime. etc)

6. Data Analysts-Visualization

- ❖ Introduction exploratory data analysis
- ❖ Descriptive statistics, Frequency Tables and summarization
- ❖ Univariate Analysis (Distribution of data & Graphical Analysis)
- ❖ Bivariate Analysis(Cross Tabs, Distributions & Relationships, Graphical Analysis}
- ❖ Creating Graphs-
 - Bar plot
 - Pie plot

- Count plot
- Line chart
- Histogram
- Boxplot
- Scatter
- Density
- Violine Plot
- Swarmplot
- Distplot
- Pair plot
- Heatmap

❖ Important Packages for Exploratory Analysis

- NumPyArrays
- Matplotlib
- Seaborn
- Pandas
- scipy.stats

Supervised Learning :

7. Linear Regression

- ❖ Introduction- Applications
- ❖ Assumptions of Linear Regression
- ❖ Building Linear Regression Model
- ❖ Understanding standard metrics (Variable significance, Rsquare/Adjusted R-square, Global hypothesis ,etc)
- ❖ Assess the overall effectiveness of the model
- ❖ Validation of Models {Re running Vs. Scoring}
- ❖ Standard Business Outputs (Decile Analysis, Error distribution (histogram), Model equation, drivers etc.)
- ❖ Interpretation of Results- Business Validation
- ❖ Implementation on new data

8. Logistic Regression

- ❖ Introduction- Applications
- ❖ Linear Regression Vs. Logistic Regression

- ❖ Building Logistic Regression Model
- ❖ Understanding standard model metrics
- ❖ Validation of Logistic Regression Models
- ❖ Multiple Logistic Regression
- ❖ Standard Business Outputs
- ❖ Interpretation of Results- Business Validation
- ❖ Implementation on new data

9. Time Series Forecasting

- ❖ Introduction -Applications
- ❖ Time Series Components(Trend, Seasonality, Cyclicity and Level) and Decomposition
- ❖ Classification of Techniques
- ❖ Basic Techniques- Averages, Smoothing, etc
- ❖ Advanced Techniques- AR Models, ARIMA, etc

10. Decision Trees

- ❖ Decision Trees - Introduction-Applications
- ❖ Construction of Decision Trees through Simplified Examples
- ❖ Generalizing Decision Trees; Information Content and Gain Ratio; Dealing with Numerical variables; other
- ❖ Measures of Randomness
- ❖ Pruning a Decision Tree; Cost as a consideration ;
- ❖ Unwrapping Trees as Rules
- ❖ Decision Trees – Validation
- ❖ Over fitting- Best Practices to avoid

11. Ensemble Learning

- ❖ Concept of Ensembling
- ❖ Manual Ensembling Vs. Automated Ensembling
- ❖ Methods of Ensembling (Stacking, Mixture of Experts)
- ❖ Bagging (Logic, Practical Applications)
- ❖ Random forest (Logic, Practical Applications)
- ❖ Boosting (Logic, Practical Applications)
- ❖ Ada Boost
- ❖ Gradient Boosting Machines (GBM)

- ❖ XGBoost

12. Naive Bayes

- ❖ Concept of Conditional Probability
- ❖ Bayes Theorem and Its Applications
- ❖ Naive Bayes for classification
- ❖ Applications of Naive Bayes in Classifications

13. Model Evaluation, Improvements & Performance Metrics

- ❖ Data Split Practices
- ❖ Cross Validation
- ❖ K-Fold Validation
- ❖ Confusion Matrix
- ❖ ROC Curves
- ❖ Mean Absolute/Square Errors & R-Square
- ❖ Ensemble Learning & Model Stacking

14. Kernel Learning

- ❖ Support Vector Machines
- ❖ Principal Component Analysis
- ❖ Ridge Regression
- ❖ Spectral Clustering

15. Support Vector Machines

- ❖ Motivation for Support Vector Machine & Applications
- ❖ Support Vector Regression
- ❖ Support vector classifier (Linear & Non-Linear)
- ❖ Interpretation of Outputs and Fine tune the models with hyper parameters
- ❖ Validating SVM models

16. Unsupervised Learning: Segmentation

- ❖ What is segmentation & Role of ML in Segmentation?
- ❖ Clustering algorithms
- ❖ Concept of Distance and related math background



- ❖ K-Means Clustering
- ❖ Hierarchical Clustering

17. Natural Language Processing

- ❖ What is NLP & How to solve NLP problems
- ❖ NLP Feature Engineering & Modelling
- ❖ How to process any raw data file.
- ❖ Build models for solving practical real world problems.

18. Deep Learning - Artificial Neural Networks (ANN)

- ❖ Motivation for Neural Networks and Its Applications
- ❖ Perceptron and Single Layer Neural Network, and Hand Calculations
- ❖ Learning In a Multi Layered Neural Net: BackPropagation and Conjugant Gradient Techniques
- ❖ Introducing & Using Tensorflow
- ❖ Neural Networks for Regression
- ❖ Neural Networks for Classification
- ❖ Interpretation of Outputs and Fine tune the models with hyper parameters
- ❖ Validating ANN models

19. End to End ML Implementation and Use Case specific discussions.