Dice Analytics Presents

Data Science & Machine Learning

About the Course

In this hybrid training, you will learn about machine learning algorithms and its applications. Further, you will also be guided how to use the machine learning algorithms in Python. This course will cover data sets from multiple domains and how to apply Machine Learning algorithms to the available data, how to get value out of Machine Learning algorithms, and how to present the output of those algorithms.

By the end of this training, you will have enough knowledge and hands-on expertise in Python to use and apply them in the real world around you. Also, you will be able to get prepared for certifications of Data Camp and Cognitive AI.

Who Should Attend?

- Graduate or Masters Students who want to start their career in the Data Science domain
- People who are working in the BI domain and want to advance their career in the field of Data Science.
- Executive who want to build a Data Science department in their start-ups/organizations

About the Instructors

- **Mr. Ahmed Rasheed:** Data Scientist - Team Lead - Machine Learning Engineer.

  Experienced Machine Learning Engineer and Data Scientist with a demonstrated history of working in the research and client-based industry.

COURSE OUTLINE

WEEK 1

- Basics of Data Science Flow
- Anaconda Installation
- Intro to Jupyter Notebook
- Intro to Python
- Python Objects & Data Structure
- Subsetting (Strings, Lists, Dictionaries)
- Python Comparison Operators
- Python Statements
- Methods & Functions
- Importing Data in Python
- NumPy & Pandas Basics in Python
- Subsetting Data frames in Pandas
- Data Aggregation using Group By, Pivot and Melt
- Hands-On Assignments of Python

WEEK 2

- Interactive Discussions on Last Week's Assignments
- Types of Variables
- Data Visualizations (Scatter plot, Histogram, Bar plots, Line plots, Heat maps)
- Data Centricity (Mean, Modes, Median, STD, Variance, Interquantile Range)
- Box plot
- Data Transformation (Log, Natural Log, Min Max)
- Data Cleaning in Python
- Visualization on Matplotlib
- Visualization on Seaborn
- Exploratory Data Analysis of Titanic dataset
- Feature Engineering
- Techniques of Filling Missing values in EDA
- Correlation Matrix
- Hands-on Assignment of EDA

WEEK 3

- Interactive Discussions on Last Week's Assignments
- What is Probability
- Conditional Probability (Disjoint Events + General Addition Rule).
- Dependence vs Independence
- Probability Trees & Bayesian Inference with their examples
- Machine Learning Basics
- Machine Learning Playground
- Supervised Learning
- Train Test Splittng
- Overfitting vs Underfitting
- Cross Validation using K-folds
- Linear Regression
- Gradient Descent, Ordinary Least Squares
- Project-1 Assigned to Students

WEEK 4

- Interactive Discussions on Project
- Multivariate Regression
- Residual Plots, R square, Adjusted R Square
- Polynomial Regression
- Model Complexity, Model Selection
- Lasso Regularization
- Ridge Regularization
- Classification
- Logistic Regression
- Confusion Matrix
- True Positive, True Negative, False Positive , False Negative
- Precision, Accuracy, Recall, F Measure
- ROC Curve, AUC, TPR, FPR
- Project-2 Assigned to Students

WEEK 5

- Interactive Discussions on Project
- Decision Trees
- Information Gain, Gini Index, Chi Square
- Random Forest
- Grid Search CV of Random Forest Hyper-parameters
- What is Boosting
- What is Bagging
- AdaBoost on Python
- Multi-Classification and Analyzing its Confusion Matrix
- Unsupervised Learning
- Clustering
- K-Means algorithm
- Elbow Analysis, Internal Indexes, Silhouette Score
- Project-3 Assigned to Students

WEEK 6

- Interactive Discussions on Project
- External Indexes, Adjusted Rand Index
- Cluster Profiling using Radar Chart
- Feature Scaling
- DBSCAN Algorithm
- Cluster Validation using DBCV
- Hierarchical clustering
- Average vs Complete vs Ward linkage
- Dendrogram Creation and Reading clusters
- External Indexes, Adjusted Rand Index
- Hierarchical clustering Use Cases
- Association Rules
- Apriori Algorithm
- Support, Confidence, Lift, Leverage, Conviction

WEEK 7

- Interactive Discussions on Project
- Visualizing Association Rules
- Network Graph Theory
- Social Network Analysis by Network Graph
- Dimensionality Reduction Concept
- Principal Component Analysis (PCA)
- Principal Vectors/Components
- Composite Features
- Maximal Variance
- Info Loss and Principal Component Analysis
- Image Classification using PCA
- Model Deployment Basics
- Flask App Introduction
- Model Deployment on Flask App

WEEK 8

- Data Science Test
- Project & Presentation
- Self-learning Path Guidance

TOOLS

• Python • Anaconda • Jupyter

Duration: 8 Weeks
Contact Us: 051-8356065-6
Email: info@diceanalytics.pk