COURSE CONTENT –
ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

1.Introduction to Data Science and AI & ML

- **Data Science, AI & ML**
- **Use Cases in Business and Scope**
- **Scientific Method**
- **Modeling Concepts**
- **CRISP-DM Method**

2. R Essentials (Tutorial) Programming

- **Commands and Syntax**
- **Packages and Libraries**
- **Introduction to Data Types**
- **Data Structures in R - Vectors, Matrices, Arrays, Lists, Factors, Data Frames**
- **Importing and Exporting Data.**
- **Control structures and Functions**

3. Descriptive Statistics

- **Data exploration (histograms, bar chart, box plot, line graph, scatter plot)**
- **Qualitative and Quantitative Data**
- **Measure of Central Tendency (Mean, Median and Mode),**
- **Measure of Positions (Quartiles, Deciles, Percentiles and Quantiles),**
- **Measure of Dispersion (Range, Median, Absolute deviation about median, Variance and Standard deviation), Anscombe's quartet**
- **Other Measures: Quartile and Percentile, Interquartile Range**

Tripple One Solutions

COURSE CONTENT –
ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

4. Statistical Analysis Initial Data Analysis

- Relationship between attributes: Covariance, Correlation Coefficient, Chi Square

- Measure of Distribution (Skewness and Kurtosis), Box and Whisker Plot (Box Plot and its parts, Using Box Plots to compare distribution) and other statistical graphs

5. Probability

- Probability (Joint, marginal and conditional probabilities)

- Probability distributions (Continuous and Discrete)

- Density Functions and Cumulative functions

6.Data Acquisition

- Gather information from different sources.

- Internal systems and External systems.

- Web APIs, Open Data Sources, Data APIs, Web Scrapping

- Relational Database access (queries) to process/access data.

7. Data Pre-processing and Preparation

- Data Munging, Wrangling

- Plyr packages

- Cast/Melt

8. Data Quality and Transformation

- Data imputation

- Data Transformation (minmax, log transform, z-score transform etc.,).

- Binning, Classing and Standardization.

- Outlier/Noise& Anomalies

Tripple One Solutions

9. Handling Text Data

- **Bag-of-words**

- **Regular Expressions**

- **Sentence Splitting and Tokenization**

- **Punctuations and Stop words, Incorrect spellings**

- **Properties of words and Word cloud**

- **Lemmatization and Term-Document TxD computation**

- **Sentiment Analysis (Case Study)**

10. Principles of Big Data

- **Introduction to Big Data**

- **Challenges of processing Big Data (Volume, Velocity and Variety perspective)**

- **Use Cases**

11. Big Data Frameworks – Hadoop, Spark and NoSQL

- **Processing, Storage andProgramming Framework**

- **Hadoop eco-system Components and their functions**

- **Essential Algorithms (Word count, Page Rank, IT-IDF)**

- **Spark: RDDs, Streaming and Spark ML**

- **NoSQL concepts (CAP, ACID, NoSQL types)**

12. Data Visualization

- **Principles of data visualization - different methods of presenting data in business analytics.**

- **Concepts of Size, Shape, Color**

13. Sampling and Estimation

- **Sample versus population**

- **Sample techniques (simple, stratified, clustered, random)**

- **Sampling Distributions**

- **Parameter Estimation**

- **Unbalanced data treatment**

14. Inferential Statistics

- **Procedure for statistical testing, etc.**

- **Test of Hypothesis (Concept of Hypothesis testing, Null Hypothesis and Alternative Hypothesis)**

- **Cross Tabulations (Contingency table and their use, Chi-Square test, Fisher's exact test)**

- **One Sample t test (Concept, Assumptions, Hypothesis, Verification of assumptions, Performing the test and interpretation of results)**

- **Independent Samples t test**

- **Paired Samples t test**

15. Linear Regression

- **Regression basics: Relationship between attributes using Covariance and Correlation**

- **Relationship between multiple variables**

- **Residual Analysis**

- **Identifying significant features, feature reduction using AIC, multi-collinearity**

- **Non-normality and Heteroscedasticity**

- **Hypothesis testing of Regression Model**

- **Confidence intervals of Slope Ÿ R-square and goodness of fit**

Tripple One Solutions

COURSE CONTENT –
ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

16. Multiple Linear Regression

- **Polynomial Regression**
- **Regularization methods**
- **Lasso, Ridge and Elastic nets**
- **Categorical Variables in Regression**

17. Non-Linear Regression

- **Logit function and interpretation**
- **Types of error measures (ROCR)**
- **Logistic Regression in classification**

18. Forecasting models

- **Trend analysis**
- **Cyclical and Seasonal analysis**
- **Smoothing; Moving averages; Box-Jenkins, Holt-winters, Auto-correlation; ARIMA**

19. Foundations for ML

- **ML Techniques overview**
- **Validation Techniques (Cross-Validations)**
- **Feature Reduction/Dimensionality reduction**
- **Principal components analysis (Eigen values, Eigen vectors, Orthogonality)**

20. Clustering

- **Distance measures**
- **Different clustering methods (Distance, Density, Hierarchical)**
- **Iterative distance-based clustering;**
- **Dealing with continuous, categorical values in K-Means**
- **K-Medoids, k-Mode and density-based clustering**

Tripple One Solutions

COURSE CONTENT –
ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

21. Classification Naïve Bayes Classifier

- K-Nearest Neighbors

- Support Vector Machines

- Decision Trees

- Ensembles methods

22. Association Rule mining

- The applications of Association Rule Mining: Market Basket, Recommendation Engines, etc.

- A mathematical model for association analysis

- Apriori: Constructs large item sets with mini sup by iterations; Interestingness of discovered association rules

- FP-trees

23. Foundations for AI

- AI: Application areas

- AI Basics (Divide and Conquer, Greedy, Branch and Bound, Gradient Descent)

- NN basics (Perceptron and MLP, FFN, Backpropagation)

24. Convolution Neural Networks

- Image classification

- Text classification

- Image classification and hyper-parameter tuning

- Emerging NN architectures

Tripple One Solutions

COURSE CONTENT –
ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

Tripple One Solutions

25. Recurrent Neural Networks

- **Building recurrent NN**
- **Long Short-Term Memory**
- **Time Series Forecasting**

26. Deep Learning

- **Auto-encoders and unsupervised learning**
- **Stacked auto-encoders and semi-supervised learning**
- **Regularization - Dropout and Batch normalization**

27. Case Studies

CONTACT US-

Phone no.- 8851517505
Email-trippleonesolutions@gmail.com
Website- www.trippleonesolutions.com

Tripple One Solutions