Data Science

Introduction to Data Science – Basic Fundamentals and Concepts – Statistics and Probability - Basics of Machine Language – Types – Algorithms – Supervised Models – Linear Regression – Logistic Regression – Classification – Decision Trees – Random Forests – K Nearest Neighbour – Naïve Bayes Algorithm for Spam detection – Support Vector Machine – Unsupervised Learning – k-Means Clustering - Association Rules Models – Reinforcement Learning - Deep Learning – Neural Network Types – Artificial Neural Network – Convolution Neural Network – Recurrent Neural Network

- **Artificial Intelligence**

  - Knowledge Base is implemented

  - enable to think, decide and react (Self Driven Cars / Robots - Mechanical, Electrical, Electronics and Knowledge Base Combination.)

- **Machine Learning is the Subset of Artificial Intelligence**
  - Statistical Models to draw intelligent / relevant Inferences from the Data Sets.
  - Statistical Estimation and Inference, Deterministic Probability and Stochastic Probability, Time Series.
  - Supervised Learning (Labelled Variable – Output is Known)
    - Linear Regression
    - Multiple Linear Regression
    - Logistic Regression (Categorical Variables)
  - Classification Models
    - Decision Trees (CART Classification Regression Analysis Test) – Random Forests – K Nearest Neighbour (KNN) – Naïve Bayes Algorithm – Support Vector Machine (SVM)
    - Unsupervised Models
      - k-Means Clustering - Association Rules Models
    - Reinforcement (Semi Supervised Model)

- **Deep Learning is the subset of Machine Learning**
  - Artificial Neural Network (Numerals)
  - Convolution Neural Network (Images)
  - Recurrent Neural Network (Time Series)
- **Data Science is the fusion of Deep, Machine, Artificial Intelligent Concepts** and developed on Mathematics and Statistics and Programming Languages
  - R, Python, SAS.
  - Differential Equations – Linear Algebra – Relational Calculus - Matrices - Distributions Theory - Statistical Estimation and Inference, Deterministic Probability and Stochastic Probability

The Data Scientists get credentials to connect to any RDBMS Database (Oracle, SQL Server, DB2 and any open source DB and Hadoop Big Data Architecture). The role of the Data Scientists is to make the request for the data sets based on the business requirements to the DBA. The role of DBA is to provide them the data file in the form of flat files (* .CSV file). **DBA need not involve in Data analytics and building Predictive models**. Hence **the role of Data Scientists is towards building predictive models and data analytics using Machine Learning and Deep Learning techniques.**