Code Mania 2019

Machine Learning

Module - 1: Introduction to Machine learning

1. What is Data Science?
2. Life of an Data Scientist?
3. Use case
4. Business case

Module - 2 : Statistics Basic Statistical Concepts and Measure

1. About Data, Statistical terminology,
2. Central Tendency, Variance and Spread
3. Basic Probability Concepts of Probability Distribution
    - Probability basics, Bayes Theorem,Confusion Metrics
    - Probability Distribution basics
4. Discrete probability Distribution

    - Geometric, Bernoulli, Binomial, Poisson
5. Continuous Probability Distribution
    - Exponential, Normal, t-distribution
6. Inferential Statistics
    - Central limit Theorem, Sampling Distributions
    - Confidence interval and Hypothesis testing.
7. Statistical Hypothesis testing
    - t-test, z-test, f-test, chi-square and ANOVA
8. Hands-On Practical
    - Python for Statistics on Distribution.

Module - 3: Python for Data Science

1. Python setup
2. Python Basics
3. Python Data Structure
4. Programming fundamentals
5. Data Analysis using Python
    - Importing Data set into Jupyter
    - Introduction to Numpy
    - Introduction to pandas
    - Introduction of Matplotlib
6. Data Cleaning
    - Identifying missing values and imputing
    - Data Formatting
    - Binning
    - Label Encoding
    - Dummies Creation
    - Data Splitting

Module - 4: Machine Learning

1. Introduction to supervised and unsupervised learning
2. Introduction to Regression and Classification
3. Supervised Learning: Regression
    - Simple Linear Regression
    - Multiple linear Regression
    - Polynomial Regression
    - Support Vector Regression
    - Time Series
4. Hands-on Practical:

- You will be implementing each Regression algorithms in python.
- We will take a dataset implement necessary preprocessing steps, building model
- Evolution Metrics and Model Overfitting,Underfitting Concepts

Module - 5: Supervised Learning(Classification)

1. Logistic Regression
2. Decision Tree
3. Support Vector Machine
4. K-Nearest Neighboring
5. Naive-Bayes
6. Ensemble
    - Bagging and Boosting
    - Random Forest
7. Hands-on Practical:
    - You will be implementing each Classification algorithms in python..
    - We will take a dataset implement necessary preprocessing steps, building model
    - Evolution Metrics and Model Overfitting,Underfitting Concepts

Module - 6: Unsupervised Learning

1. Clustering
    - K-means Clustering
    - Hierarchical Clustering
2. Hands-On Practical:
    - You will be implementing each Classification algorithms in python.
    - We will take a dataset implement necessary preprocessing steps, building model
    - Evolution Metrics and Model Overfitting Underfitting Concepts

Module - 7: Rules based learning

1. Apriori
2. Collaborative filtering
3. Market Basket Analysis
4. Dimension Reduction:
    - Principal component Analysis
5. Hands-On Practical:

- You will be implementing each Classification algorithms in python.
- We will take a dataset implement necessary preprocessing steps, building model
- Evolution Metrics and Model Overfitting,Underfitting Concepts

Module - 8: Introduction to Data Science Experience

1. Setting up IBM data platform
2. Getting started with DSX
3. Getting started with Data Catalog
4. Getting Started with IBM Data Refinery
5. Organize resources in a project
    - Set up a project
    - Watson Data Platform projects
    - Project Collaborators
    - Add associated services
6. Prepare data
    - Add data to a project
    - Refine data
    - Ingest streaming data
7. Working with Jupyter Notebooks
    - Create notebooks
    - Code and run notebooks
    - Share and publish notebooks
8. Machine Learning Flows
    - Creating Machine Learning flows with IBM SPSS
    - Creating Machine Learning flows with Spark MLlib
9. Watson Machine Learning
    - Setting up your machine learning environment
    - Building models
    - Deploying machine learning
10. Visualizations
    - Pixie Dust
    - Brunel visualization
    - SPSS model visualization
11. Predictive Analytics Algorithms
    - Classification and regression
    - Clustering
    - Forecasting
    - Survival analysis