ALTALUNE TECHNOLOGY

MACHINE LEARNING AND DATA SCIENCE

- Course Introduction
- Environment Setup
- Jupyter Overview
- Python Crash Course
- Python for Data Analysis – Numpy
    - Introduction
    - Numpy Array
    - Array indexing
    - Numpy operations
- Python for Data Analysis – Pandas
    - Introduction
    - Series
    - DataFrame
    - Missing Data
    - Groupby
    - Merging, Joining, and Concatenating
    - Operations
    - Data Input and Output
    - Pandas Exercise
- Python for Data Visualization – Matplotlib
    - Introduction
    - Plotting graphs using matplotlib
    - Matplotlib Exercise
- Exploratory Data Analysis:
    - Univariate and Bivariate Analysis
    - Missing Values Treatment
    - Outlier Treatment
    - Feature Scaling
    - Feature Engineering
- Introduction to Machine Learning
    - Machine Learning Overview
    - Supervised and Unsupervised Learning
    - Training and Testing models
    - Evaluation – Regression
    - Evaluation – Classification
    - Model Selection
    - Bias Variance Trade-off
    - Cross Validation
    - Hyperparametric Optimization
- Supervised Machine Learning Algorithms
    - Linear Regression
    - Logistic Regression
    - Decision Trees and Random Forest

  - K Nearest Neighbors
  - Naïve Bayes
  - Support Vector Machines
  - Ensemble Methods
- Supervised Learning Project
- Unsupervised Machine Learning Algorithms
  - K-means Clustering
  - Hierarchical Clustering
  - Density based clustering
    - Gaussian Mixture Models
  - Principal Component Analysis
- Recommender System
- Natural Language Processing
  - Text Processing
  - Tokenization
  - Parts of Speech Tagging
  - Stemming
  - Named Entity Recognition
  - Feature Extraction