Data Science

Introduction to Data Science

- Overview of Data Science
- The Data Science Pipeline: Collection, Cleaning, Analysis, Visualization,
- Key Tools and Technologies (Python, Jupyter Notebooks)

Programming for Data Science

- Basics of Python (Variables, Data Types, Control Structures)
- Libraries for Data Science (NumPy, Pandas)
- Data Manipulation with Pandas
- Visualization with Matplotlib and Seaborn

Data Collection and Cleaning

- Data Import and Export (CSV, Excel, JSON, SQL)
- Data Cleaning and Preprocessing
- Handling Missing Values and Outliers
- Data Transformation and Feature Engineering

Exploratory Data Analysis (EDA)

- Understanding Data Distributions
- Data Visualization Techniques (Histograms, Box Plots, Heatmaps)
- Correlation Coefficient

Data Science

Machine Learning Fundamentals

- Introduction to Machine Learning
- Supervised vs. Unsupervised Learning
- Model Evaluation Metrics (Accuracy, Precision, Recall, F1 Score, ROC-AUC)
- Cross-Validation Techniques

Supervised Learning Techniques

- Linear Regression
- Logistic Regression
- Decision Trees
- Random Forests
- Support Vector Machines (SVM)
- K-Nearest Neighbors (KNN)

Unsupervised Learning Techniques

- Clustering (K-Means, Hierarchical Clustering)
- Dimensionality Reduction (PCA, t-SNE)
- Association Rule Learning (Apriori Algorithm)

Data Science

Advanced Topics

- Ensemble Methods (Boosting, Bagging)
- Neural Networks and Deep Learning (Introduction to TensorFlow/Keras)
- Natural Language Processing (NLP) Basics
- Time Series Analysis and Forecasting

Real-World Applications and Projects

- Case Studies from Various Industries (Healthcare, Finance, Marketing, etc.)
- Project Work: End-to-End Data Science Projects
- Best Practices for Data Science Project Management

Tools and Technologies

- Introduction to SQL for Data Querying
- Version Control with Git and GitHub

Thank You

Dev's Planet Software Solutions