DataQuest Solutions

Machine Learning With Python

Lesson 1. Python Foundations for ML

- Numpy - Numerical Operations
- Pandas - Data Manipulation & Preprocessing
- Matplotlib / Seaborn - Basic Data Visualization for EDA

Lesson 2. Data Preprocessing

- Handling Missing Values
- Encoding Categorical Variables
- Feature Scaling (Normalization & Standardization)
- Train-test Split from sklearn
- Feature Selection Techniques (filter, wrapper, embedded methods)

Lesson 3. Exploratory Data Analysis

- Understanding Distributions, outliers & correlations
- Creating Summary Statistics
- Visualization (scatter plots, histograms, heatmaps)

Lesson 4. Supervised Learning Algorithms

- Linear Regression
- Logistic Regression
- Decision Trees
- Random Forest
- Gradient Boosting Machines
- Support Vector Machines
- K-Nearest Neighbors
- Neural Networks

Lesson 5. Unsupervised Learning Techniques

- K-Means Clustering
- Hierarchical Clustering
- Dimensionality Reduction (PCA)

Lesson 6. Model Evaluation Techniques

- Classification Metrics (accuracy, precision, recall, F1-score, ROC-AUC)
- Regression Metrics (RMSE, MAE, R^2)
- Confusion Matrix
- Cross Validation

Lesson 7. Model Tuning & Optimization

- Grid Search
- Randomized Search
- Bayesian Optimization

Lesson 8. Hands-on Project

- Working with real datasets (Kaggle, UCI, ML Repository)
- Building Complete ML pipelines (pipeline from sklearn)
- Saving and Loading Models (pickle, joblib)