

Assignment 6

1. Data Ingestion Pipeline:

- a. Design a data ingestion pipeline that collects and stores data from various sources such as databases, APIs, and streaming platforms.
- b. Implement a real-time data ingestion pipeline for processing sensor data from IoT devices.
- c. Develop a data ingestion pipeline that handles data from different file formats (CSV, JSON, etc.) and performs data validation and cleansing.

Answer: a.

2. Model Training:

- a. Build a machine learning model to predict customer churn based on a given dataset. Train the model using appropriate algorithms and evaluate its performance.
- b. Develop a model training pipeline that incorporates feature engineering techniques such as one-hot encoding, feature scaling, and dimensionality reduction.
- c. Train a deep learning model for image classification using transfer learning and fine-tuning techniques.

Answer: a.

3. Model Validation:

- a. Implement cross-validation to evaluate the performance of a regression model for predicting housing prices.
- b. Perform model validation using different evaluation metrics such as accuracy, precision, recall, and F1 score for a binary classification problem.
- c. Design a model validation strategy that incorporates stratified sampling to handle imbalanced datasets.

Answer: b.

4. Deployment Strategy:

- a. Create a deployment strategy for a machine learning model that provides real-time recommendations based on user interactions.
- b. Develop a deployment pipeline that automates the process of deploying machine learning models to cloud platforms such as AWS or Azure.
- c. Design a monitoring and maintenance strategy for deployed models to ensure their performance and reliability over time.

Answer: b.