DATA-236 Sec 12 - Distributed Systems for Data Engineering

HOMEWORK 11

Total Point: 10

Instructions:

- Include screenshots of the code and output. Ensure that the code and its output are displayed together, one below the other.
- Submission should be in PDF Format.
- Please name your submission file as {last_name}_HW11.pdf

Q1. Shipping Microservice **(5 Marks)**
Build a shipping microservice that listens to the "order-confirmed" topic and creates a shipping record. (Refer the demo code for orders topic).
Requirements:

- Connect to Kafka as a consumer.
- Subscribe to the topic: "order-confirmed"
- Save record to db with tracking id and status
    - Generate a tracking Id: (e.g SHIP-XYZ)
    - Set the status to 'Pending'
    - Save the record

NOTE: The shipping schema should consist of
itemId: String
itemName: String
Quantity: Number
trackingId: String
status: { type: String, default: "pending" }
createdAt: { type: Date, default: Date.now }

Q2. ML-OPS (5 Marks)
Create a mini-MLOps pipeline using your own model and deploy a FastAPI-based prediction endpoint.

- Train a Model - Use a simple dataset like Iris, Wine, or your own mini regression/classification problem, Save the model as a .pkl file using joblib

- Serve Using FastAPI - Build an API with /predict endpoint using FastAPI. Accept input as JSON and return the predicted output

- Set Up GitHub Repo- Upload your project files: app.py, model.pkl, requirements.txt, etc.

- Add CI Workflow - Create .github/workflows/ci.yml. It should install dependencies, run the FastAPI script or sanity check model load.

Submit the screenshot of your CI workflow.yml and the actions page of your GitHub repo with the pipeline successful.