END-TO-END MLOPS PROJECT – INSTRUCTION SHEET (LOCAL & OPEN-SOURCE TOOLS ONLY)

1. PROJECT OBJECTIVE

Build a complete ML + MLOps pipeline using only open-source tools running locally. Students may choose any dataset with at least 300 rows for tabular data or 100 samples for text/images.

2. ALLOWED TOOLS (OPEN-SOURCE & LOCAL)

Python, Git/GitHub, DVC, MLflow (local), Prefect (local), Docker, FastAPI/Flask, Streamlit, Evidently AI.

3. REQUIRED COMPONENTS

A. Problem Definition & Dataset Selection

B. Exploratory Data Analysis (EDA)

C. Data Preprocessing & DVC Tracking

D. Model Development with MLflow Tracking

E. Prefect Pipeline Orchestration

F. Repository Structure & Version Control

G. CI/CD using GitHub Actions (local simulation)

H. Local Model Deployment via FastAPI/Flask/Streamlit

I. Containerization using Docker

J. Local Monitoring using Evidently

4. DELIVERABLES

- GitHub repository

- PDF final report

- 5–10 minute demo video

- Dockerized local deployment

5. EVALUATION CRITERIA (100 MARKS)

Problem Definition (5), EDA (10), Preprocessing+DVC (10),

Model + MLflow (15), Prefect Workflow (15), CI/CD (10),

Dockerization (10), Deployment (10), Monitoring (10), Documentation (5)