MLOps End-to-End Pipeline - Student Handout

This handout provides a complete step-by-step guide to building an end-to-end MLOps pipeline using open-source tools:
Feature Store, DVC, MLflow, FastAPI, Docker, and GitHub Actions.

1. Project Overview

- Feature Store (SQLite)
- DVC for data and model versioning
- MLflow and Optuna for experiment tracking
- FastAPI for model deployment
- Docker for containerization
- GitHub Actions for CI/CD automation

2. Setup Instructions

1. Clone the repository:
   git clone https://github.com/renuka1983/mlops_docker_demo.git
   cd mlops_docker_demo

2. Create and activate a virtual environment:
   python3 -m venv .venv
   source .venv/bin/activate
   pip install -r requirements.txt

3. Install DVC and pull project data:
   pip install "dvc[gdrive]"
   dvc pull

3. Run the Pipeline

Run all stages of the pipeline:
   dvc repro

Show current metrics:

```
dvc metrics show
```

Launch MLflow UI:

```
mlflow ui --backend-store-uri ./models/mlruns
```

4. API and Docker

Start FastAPI service:

```
uvicorn api.model_api:app --reload
```

Docker build and run:

```
docker build -t mlops-demo .
docker run -d -p 8000:8000 mlops-demo
```

5. CI/CD and Verification

GitHub Actions automates the DVC pipeline, builds Docker images, and tests the FastAPI service after each push to main.

Verification Commands:

```
dvc metrics show
dvc repro
docker ps
curl http://127.0.0.1:8000/docs
```