MLOps Pipeline for California Housing Price Prediction

#	Name	BITS Email ID
1	Priyanka Asati	2023ac05700@wilp.bits-pilani.ac.in
2	Achyut Kumar Dwivedi	2023ac05784@wilp.bits-pilani.ac.in
3	Ram Lal Singh	2023ac05814@wilp.bits-pilani.ac.in
4	Shubham Yadav	2023ac05241@wilp.bits-pilani.ac.in

Goal: Build, track, package, deploy & monitor regression models using **DVC, MLflow, FastAPI, Docker, GitHub Actions, SQLite, Prometheus & Grafana**, ensuring reproducibility, automated deployment & basic monitoring.

Dataset: sklearn.datasets.fetch_california_housing (saved as data/raw/california_housing.csv)

Component	Tool / Tech	Purpose
Version Control	Git + GitHub	Track code, pipeline, configs
Data Versioning	DVC	Track & reproduce dataset changes
Experiment Tracking	MLflow	Track parameters, metrics, artifacts
Model Registry	MLflow	Store and version best models
API Service	FastAPI	Serve predictions as REST API
Containerization	Docker	Portable deployment
CI/CD	GitHub Actions	Auto build, test, push image
Logging	SQLite	Store predictions & errors
Monitoring	Prometheus Client	Export metrics for dashboard
UI	Streamlit	Simple web UI to submit features and view predictions

Key Features -

- **Data versioning:** DVC tracks raw CSV; **auto-retrain** on data/code changes via ***dvc repro***.
- **Experiment tracking:** MLflow logs params/metrics/artifacts; **best model** registered with alias best.
- **Serving:** FastAPI loads models:/housing_best_model@best; JSON in → prediction out; Streamlit
- **Containers:** docker-compose.yml brings up **mlflow**, **api**, **prometheus**, **Grafana, Streamlit**.
- **CI/CD:** GitHub Actions builds & pushes API image to Docker Hub.
- **Code Quality** (Lint)/ **Format** (Black)/ **Unit Test** (Pytest) on Git Push
- **Monitoring:** /metrics/prom for Prometheus; **Grafana** dashboard ready.
- **Logging:** All requests stored in **SQLite** at logs/api.db.

Artefacts –

- **GitHub Repo:** https://github.com/asati-priyanka/mlops-housing-project
- **Docker Hub Image: asatipriyanka/housing-api:latest (docker_hub_link)**
- **MLFLOW URL:** http://localhost:5500/#/experiments
- **SQLite + Prometheus integration: http://127.0.0.1:9090**
- **Grafana-ready dashboard: Granfana_Dashboard_Link**
- **Streamlit UI:** Interactive Streamlit UI (http://localhost:8501) calling Docker API (http://api:8000/predict)