Capstone Project – End-to-End ML Pipeline for Automotive AI

This project builds a complete machine learning workflow using MLflow, covering everything from training to deployment and monitoring—all using a public automotive dataset.

Dataset: Auto MPG Dataset
Dataset URL: https://www.kaggle.com/datasets/uciml/autompg-dataset

1. Train & Pick the Best Model
Data Prep: Clean the Auto MPG data (handle missing values, normalize features).
Training: Experiment with 5 models (e.g., Linear Regression, Random Forest, etc.).
MLflow Tracking: Log metrics (RMSE, MAE) and compare models in the MLflow UI.
Save the Best: Register the top-performing model in the MLflow Model Registry.

2. Deployment
Local Test: Serve the model as a REST API (e.g., predict MPG for a new car's specs).

3. Monitor in Production
Track Metrics: Log real-world prediction accuracy over time.
Drift Detection: Alert if new car data differs too much from training data

4. Final Deliverable
A fully documented MLflow pipeline showing:
Model comparisons
Version history and rollback demonstration.
Monitoring dashboards for fuel efficiency predictions