UMBC

DATA 606 - Capstone Project

Data Fusion for Predicting Highway Maintenance and Deterioration Trends

Team Members:
Chandra Sekhar Katipalli
Sindura Reddy Challa
Sanjana Reddy Soma

Project Overview

Project Goal:

To develop a predictive model for forecasting road deterioration using integrated datasets of traffic, pavement conditions, and freight movements from HPMS and FAF (2013–2022).

Problem Statement:

Road maintenance decisions are often reactive due to limited predictive insights. Our project aims to provide data-driven forecasts of road deterioration, enabling proactive maintenance and optimized budget allocation.

Approach:

We performed data fusion between HPMS and FAF datasets, engineered relevant features, and applied machine learning models (Decision Tree, XGBoost, KNN) to predict pavement deterioration using International Roughness Index (IRI) as a key indicator.

Deployment Scope

- The project was deployed as a **public GitHub repository**, making all components accessible for reuse, collaboration, and extension.
- Repository includes:
  - Cleaned and merged datasets (HPMS + FAF sample)
  - Python scripts and Jupyter notebooks for data processing, feature engineering, model training and evaluation
  - Trained model artifacts and evaluation reports
  - README documentation explaining workflow, usage, and results
- Deployment focuses on **reproducibility, transparency and ease of access** for researchers and stakeholders.
- No live API or dashboard deployed yet; groundwork set for potential integration into a visualization platform or decision-support tool in the future.

Deployment Goals

- **Ensure reproducibility**: Allow users to reproduce results from raw data to final model predictions via scripts and instructions.
- **Public accessibility**: Share the project openly to support further research and real-world application by transportation agencies or policymakers.
- **Documentation clarity**: Provide a well-documented repository with step-by-step instructions for running analysis, training models, and generating outputs.
- **Enable collaboration and extension**: Facilitate future work by making code modular and repository structure clear for other developers or researchers.
- **Model preservation**: Store trained model files alongside evaluation metrics for easy reuse or retraining.

UMBC

Github Repository

Technology Stack

Frontend	Flask (lightweight web framework), HTML with Ajax, CSS for UI
Backend	Python (data processing, routing), Pandas, NumPy
Model	TCN with hyperparameter tuning and regularization
Visualization	Matplotlib, Seaborn, Folium (interactive maps)
Deployment	hosted on GCR with build triggers connected with GitHub (on commit). .Yaml file for build configuration. Dockerfile for build instructions
Data	HPMS & FAF datasets (2013–2021), preprocessed and scaled

Dashboard Features

- Route selection (via dropdown menu) to choose highway route.
- Section filtering to specify road segment for analysis.
- IRI prediction output for the selected route/section.
- Trend visualization of IRI over time.
- Color-coded map display with maintenance suggestion indicators.

Workflow Behind the Dashboard

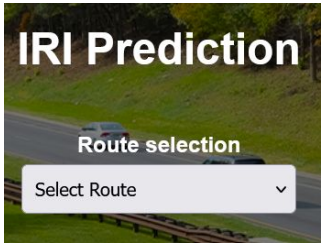
Input: User selects route and section (input stage).

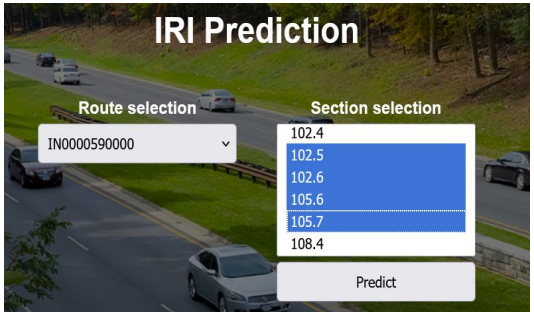
Data Processing: Filter and preprocess the selected data.

Inference: Run the trained model to predict IRI (inference stage).

Output: Update visual elements (map, trend graph) and prediction table (output stage).

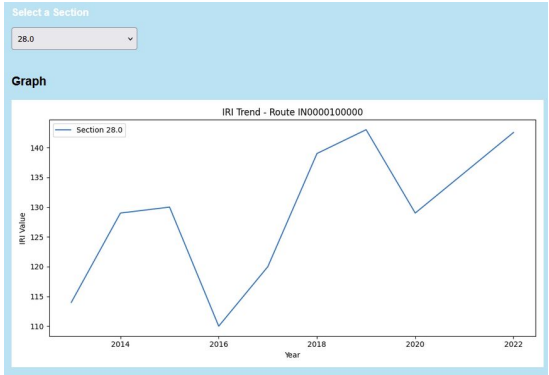
Dashboard Overview





Prediction Table

ROUTE_ID	BEGIN_POIN	END_POINT	Predicted IRI
IN0000100000	28.0	28.1	142.53
IN0000100000	28.1	28.2	144.00
IN0000100000	28.2	28.3	147.04
IN0000100000	29.0	29.1	147.27
IN0000100000	29.1	29.2	128.89



Future Work: Enhancements & Next Steps

1. **Expand Domain Knowledge & Features**
    - Integrate more factors like material properties, maintenance history, construction records and expert insights to improve model accuracy.
2. **Improve Prediction Visualization**
    - Develop interactive dashboards or GIS maps to display predictions more intuitively and highlight critical deterioration zones.
3. **Include Geospatial Analysis**
    - Apply spatial methods to analyze geographic patterns and regional deterioration hotspots, leveraging GIS-based insights.
4. **Broaden Road Coverage**
    - Extend models to other road types: Principal Arterial – Other Freeways/Expressways, Minor Arterial, Major/Minor Collector, Local roads.
5. **Account for Climate Change**
    - Incorporate climate data (temperature, precipitation, extreme events) to model long-term deterioration impacts and adaptive maintenance strategies.

UMBC

THANK YOU