

Title: Road Closure Severity Prediction

Member 1	Rajat Belgundi	rajatrb@vt.edu
Member2	Yash Kulkarni	yashpandharish@vt.edu

Introduction:

Maintenance of roads and construction of new roads or pavements is very important to improve the day-to-day life of the people. These events many times require roads to be closed temporarily or traffic to be diverted onto alternate routes. Due to a variety of reasons, these road closures get prolonged and become an inconvenience for the general public. Every department of transportation in the country always aims to reduce the impact of such road closures to ensure a smooth daily travel routine of the travelers.

The reasons for prolonged road closures could be many. To name a few, environmental reasons - wind, precipitation cold, time of the day, presence of traffic signals or crosswalks, airport or train station in the vicinity and many more. Regardless of the reason, finding ways to speed up the road work, reduce traffic bottlenecks to alleviate the inconvenience caused by the event is the ultimate goal.

We aim to solve this problem by making use of machine learning based classification techniques on the US Road Construction and Closures dataset. Having a machine learning model to classify a particular closure event will help the transport department address the highest and least impacted routes effectively by taking prompt action. During the data analysis we also aim to derive insights from the data to further study the impact of various stimuli on the road closure event. The dataset has been provided by [1] in which the authors also run experiments on spatial data with deep learning.

Project Problem Statement:

Implementing machine learning classification techniques to predict the severity of the road closure event or construction work to assist the department of transportation in either speeding up the current work or offering alternate routes for diverting traffic on busy routes.

Data Set:

US Road Construction and Closures (2016 - 2021)

Available on Kaggle:

<https://www.kaggle.com/datasets/sobhanmoosavi/us-road-construction-and-closures>

About the dataset:

Description

The following is a countrywide dataset of road construction and closure events, which covers 49 states of the US. Construction events in this dataset could be any roadwork, ranging from fixing pavements to substantial projects that could take months to finish. The data is collected from Jan 2016 to Dec 2021, using multiple APIs that provide streaming traffic incident (or event) data. These APIs broadcast traffic data captured by a variety of entities, such as the US and state departments of transportation, law enforcement agencies, traffic cameras, and traffic sensors within the road-networks. Currently, there are about 6.2 million construction and closure records in this dataset.

Columns: 47 (original) can be increased in feature engineering step as we have a variety of features - categorical, datetime, numerical, boolean etc

Rows: 6.2 M rows

Preprocessing steps:

1. Exploratory Data Analysis
2. Data Cleaning and Preprocessing
3. Feature Engineering
4. Feature Scaling

Methods and Models:

1. Logistic Regression
2. Random Forest
3. Decision Tree
4. KNN Classification
5. Support Vector Machine

References:

[1] Karimi Monsefi, Amin, Sobhan Moosavi, and Rajiv Ramnath. "Will there be a construction? Predicting road constructions based on heterogeneous spatiotemporal data.", In Proceedings of the 30th ACM SIGSPATIAL 2022.