A Comparison of text classification models applied to highway closure severity

Neha Alluri

Purpose

- Highway closure

  = roadway congestion

- How can we understand

  Highway closure data?

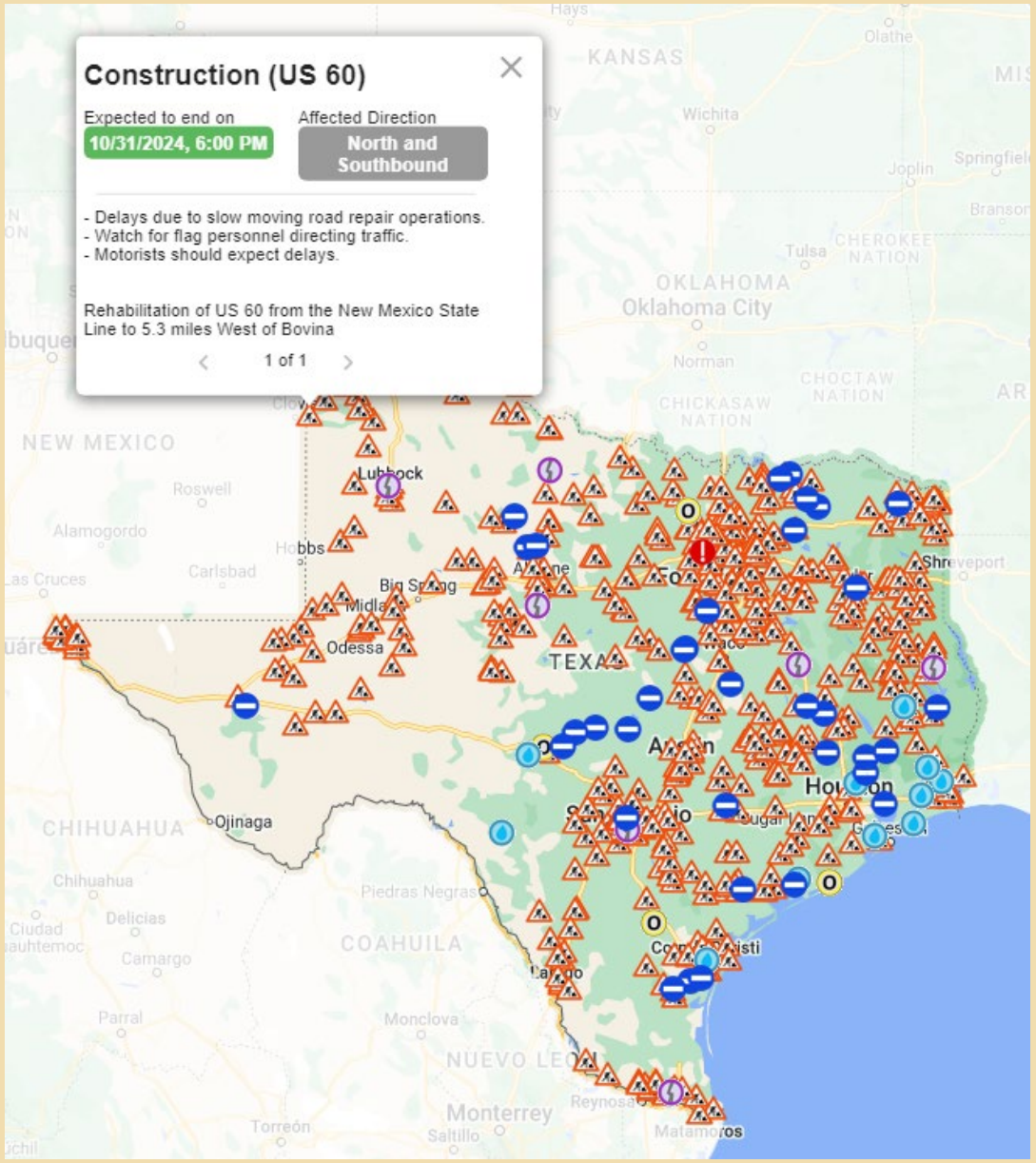
- How is this information useful to

  us?

Key Characteristics

- Multi-label text classification
- 4 Severity Categories:
  - Not Severe
  - Slightly Severe
  - Mostly Severe
  - Severe
- HCRS Database

text classification



Take data from TxDOT Highway Conditions

n Data → Split into Testing and trainig sets

Train and Test model → Store Model

Predict new labels for unseen descriptions

Framework for multi-label text classification

Information Retreaval

Data Set Creaton

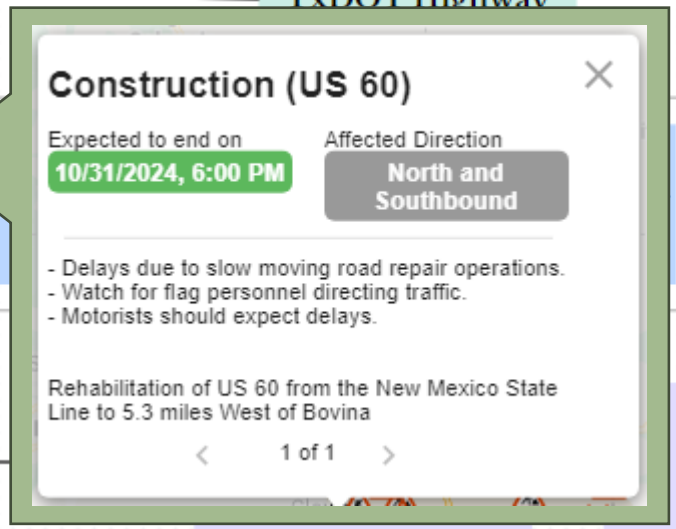
Model Building

Model Analysis

Write information into HCRS Database

Take data from TxDOT Highway

Pull Descriptions From All entries



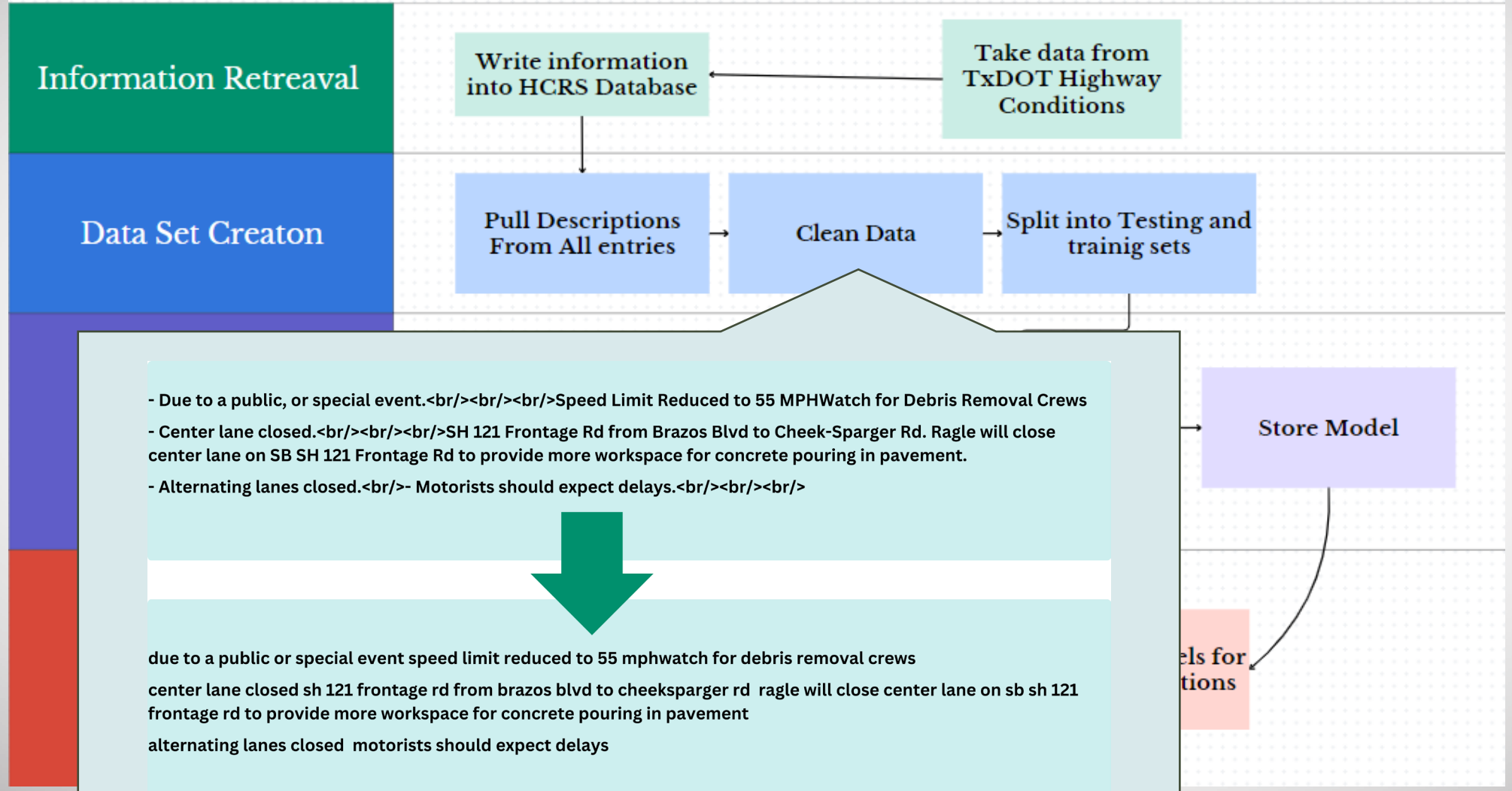
Set up model based on seperate libraries

Store Model

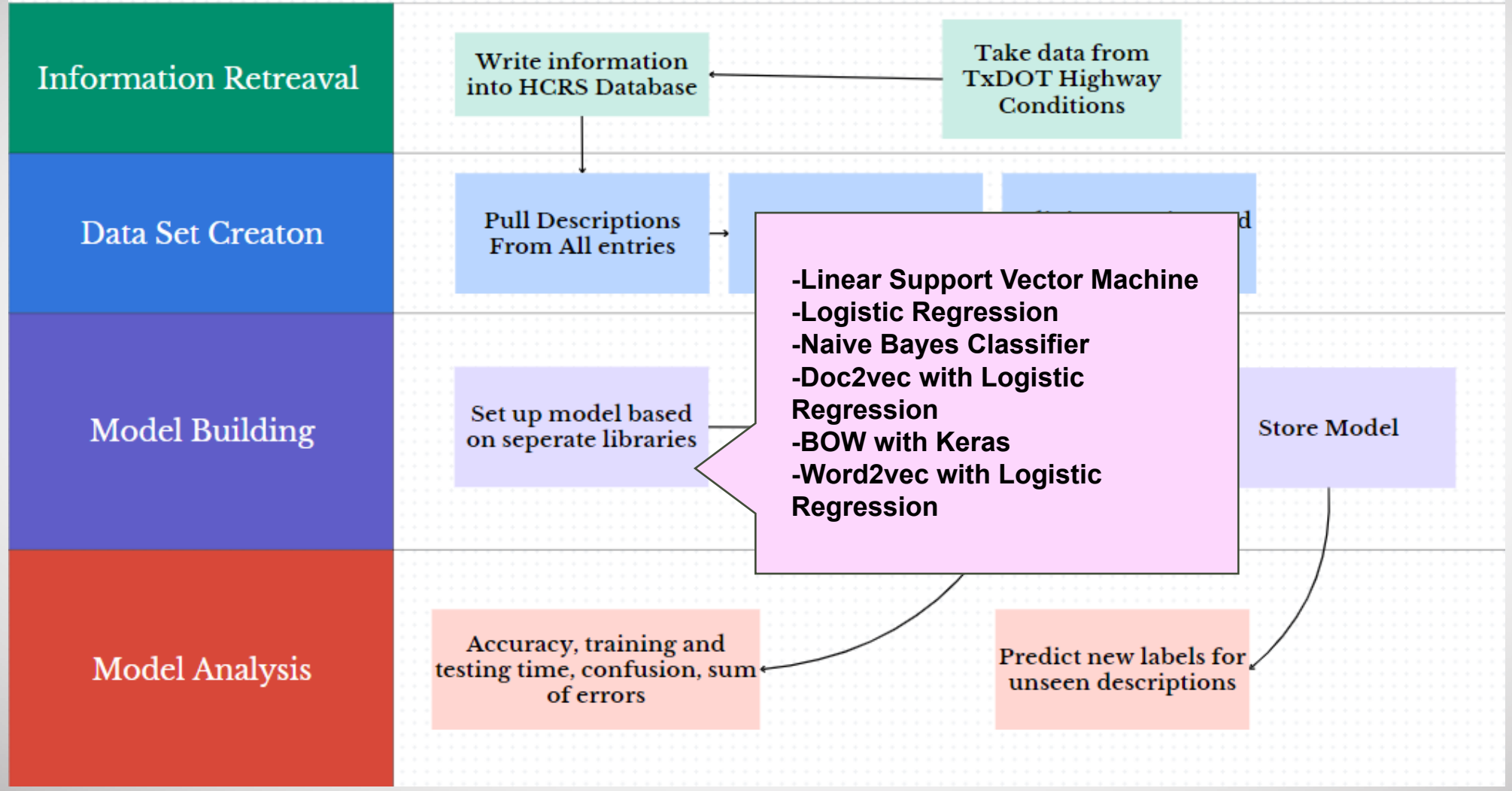
Accuracy, training and testing time, confusion, sum of errors

Predict new labels for unseen descriptions

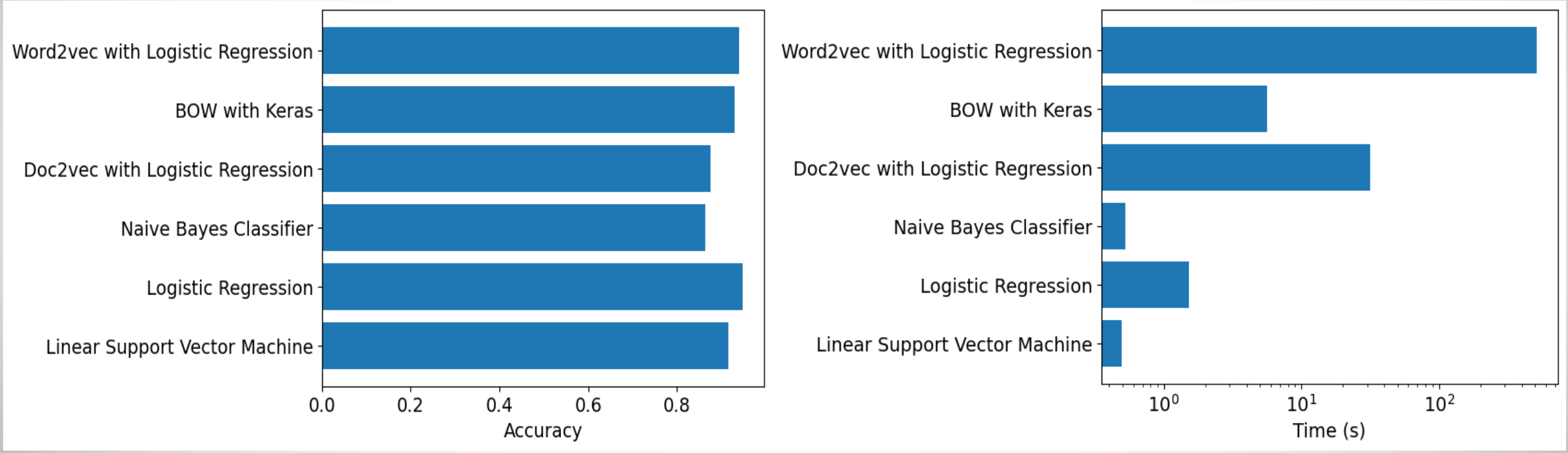
Framework for multi-label text classification



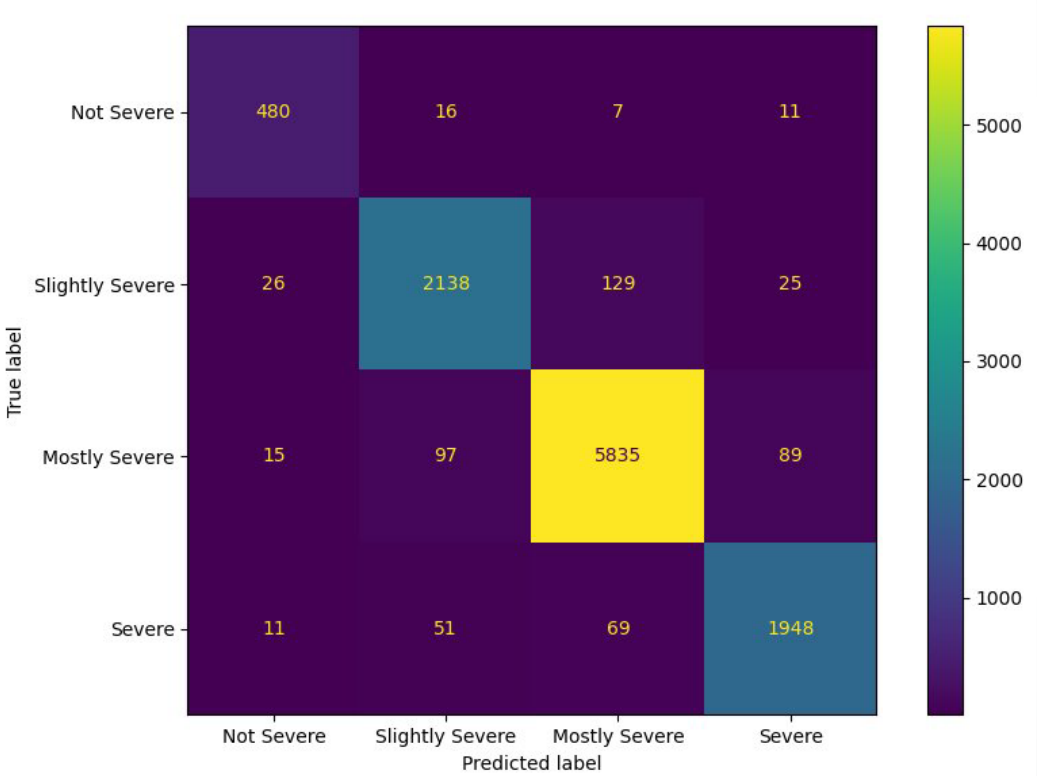
Framework for multi-label text classification



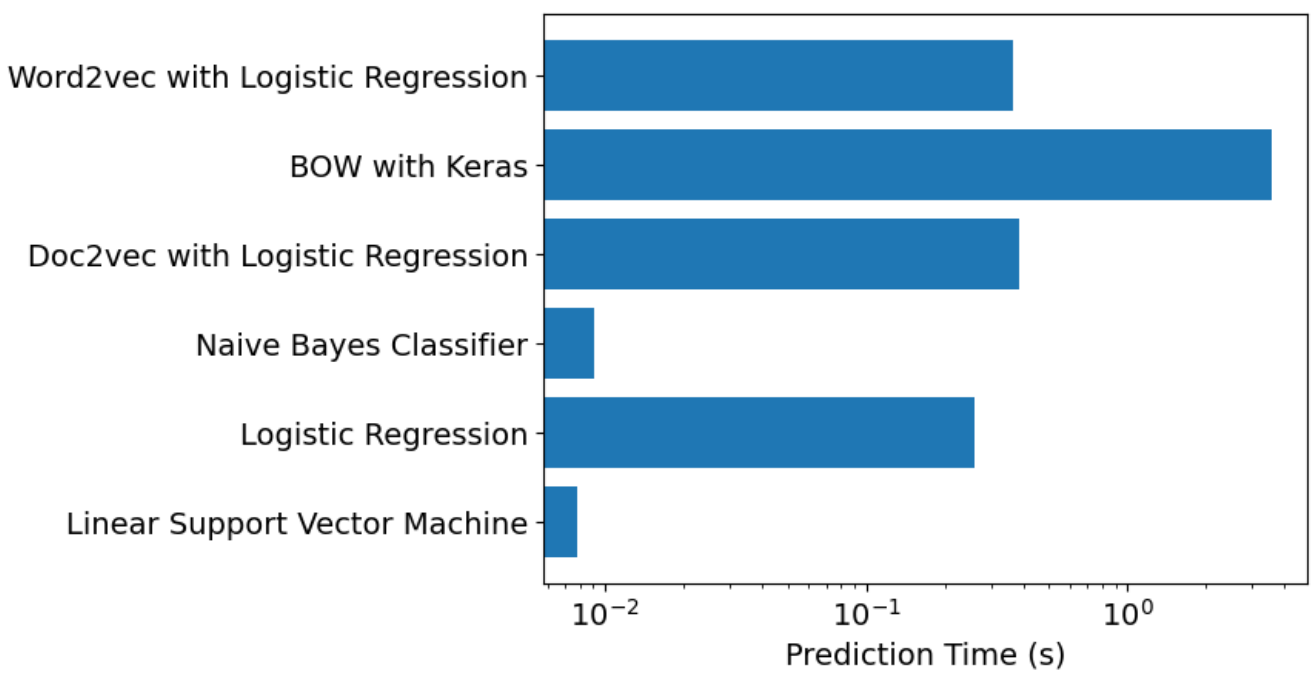
Results – Time and Accuracy



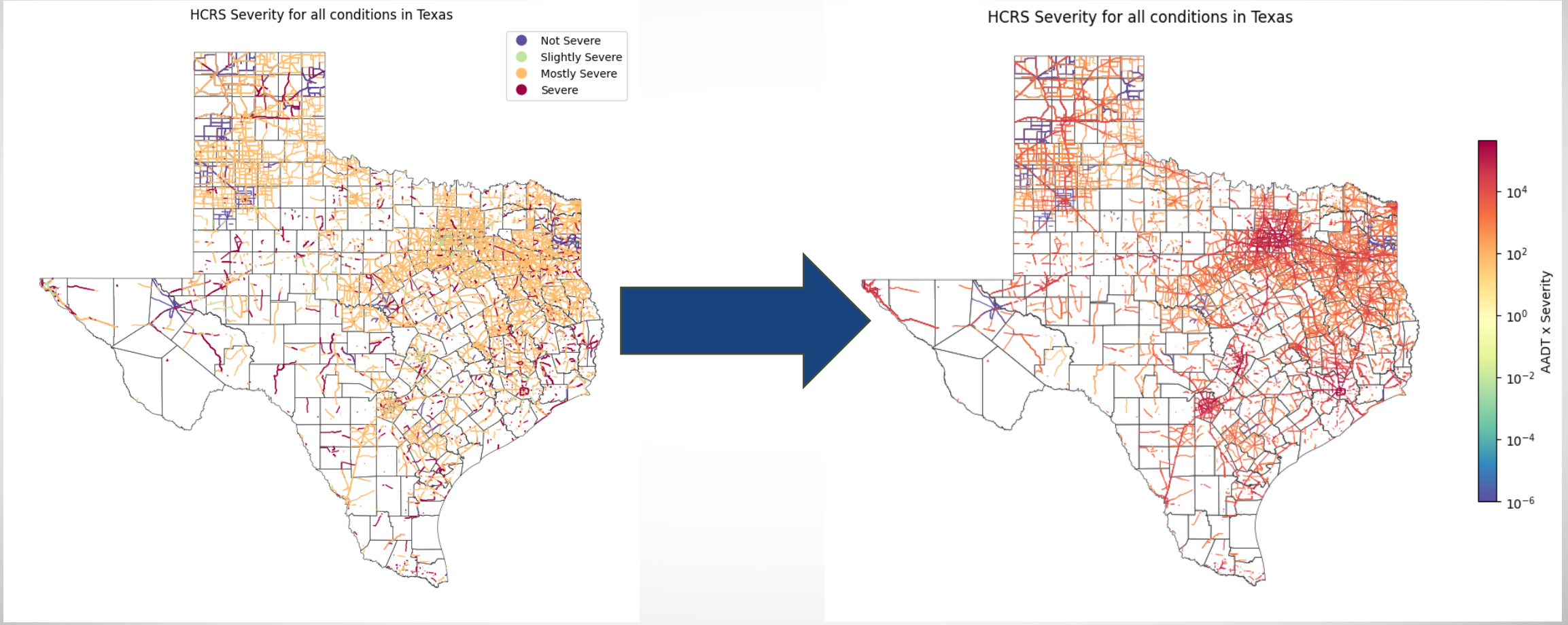
Results – Confusion, Prediction Time and Sum of Errors



Logistic Regression – Visualized Confusion Matrix



Applications



Further Research

- The process of data transformation could lead to new insights.
- The results gained from this research could help prepare for road closures in advance.
- Analyze patterns based on conditions to identify high risk highways.
- Study real-time effects of highway closures based on daily traffic.

Thank You

Neha Alluri

Texas A&M Transportation Institute

Questions?