Applied Analytics Project

Analyzing US Accident Data to Predict High-Risk Areas and Times in Massachusetts

Week 11 – Explain the model, analyze risk, bias and ethical considerations

Major: Applied Analytics

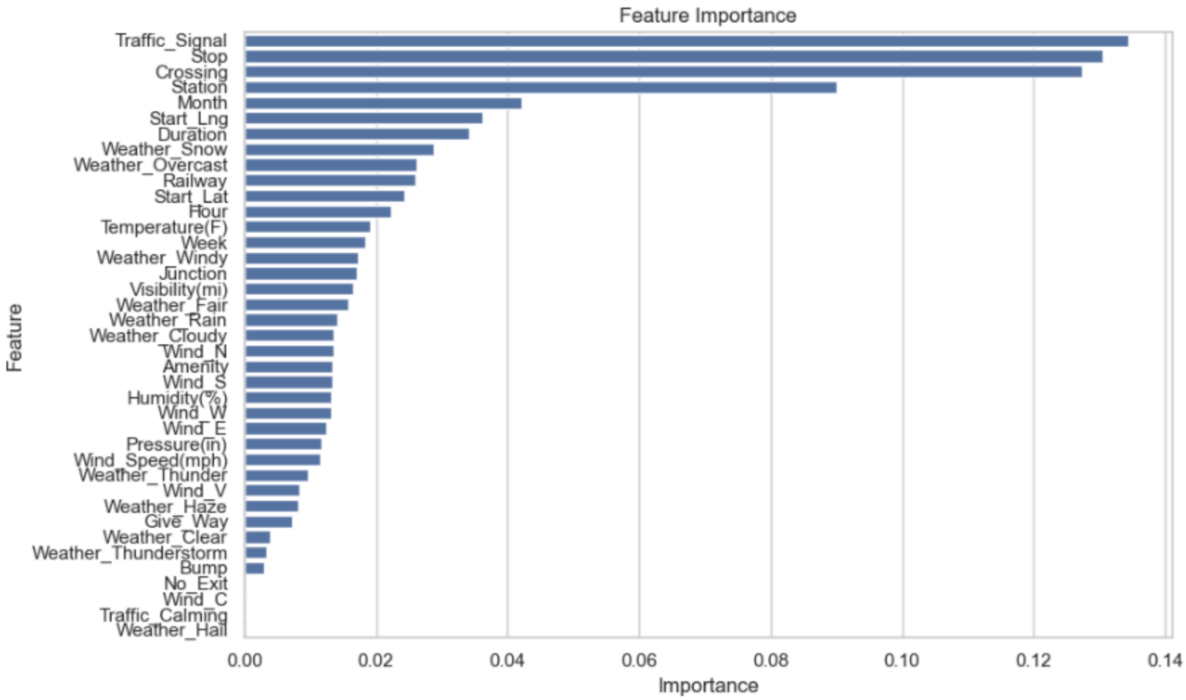
Name: Gefan Wang, Chenhe Shi, Tianchen Liu

Date: 04.13.2025

1. Identification of Important Features

Based on the XGBoost model, the most important predictors contributing significantly to the model's accuracy and predictive power are:

- **Traffic_Signal:** Presence of traffic signals appears to be the most influential feature, possibly due to its role in controlling intersections and traffic behavior.
- **Stop:** Locations near stop signs are associated with accident severity, likely due to abrupt stops and intersection complexity.
- **Crossing:** Pedestrian or road crossings increase the risk profile, influencing the severity outcome.
- **Station:** Proximity to transit stations may indicate high-traffic areas, affecting accident dynamics.
- **Month:** Seasonality and time of year appear to correlate strongly with accident severity, likely due to weather and travel trends.
- **Start_Lng & Start_Lat:** Geolocation is a significant predictor, pointing to geographical zones with higher risk.
- **Duration:** The length of an accident, derived from time between start and end, is also a meaningful indicator of severity.
- **Weather-related features (e.g., Snow, Overcast, Windy):** Various weather conditions moderately influence outcomes, especially adverse ones like snow and wind.
- **Temperature(F), Visibility(mi), and Humidity(%):** These environmental factors still contribute but play a more supporting role compared to road and temporal features.



2. Explanation of Random Predictions

Let's examine five randomly selected predictions and discuss the specific influential factors:

- **Prediction 1:** Predicted severity is high primarily due to very low visibility, prolonged accident duration, and adverse weather conditions.
    - To reduce the severity, visibility could be improved by approximately 2 miles, and accident duration reduced by half.
- **Prediction 2:** Moderate severity driven largely by high humidity levels and the presence of intersections and traffic signals.
    - Lowering humidity by around 15% or altering the traffic infrastructure by removing intersections would significantly decrease severity.
- **Prediction 3:** Low severity was predicted for a location without a traffic signal or crossing and with minimal weather interference. However, adding a station or crossing to the same context would likely increase predicted severity.
- **Prediction 4:** The model predicted high severity due to a combination of adverse weather (snow), a stop sign, and high accident duration. Adjusting any of these, such as reducing the duration by 50%, could lower the severity prediction.

- **Prediction 5:** A moderate prediction was made for an accident in February (a high-risk month per model importance) near a station with overcast weather. Shifting the month or removing proximity to the station could reduce the predicted severity.

3. Protected Categories in Dataset

Since our dataset focuses on environmental conditions, geographical locations, and accident-related details, it does not explicitly contain typical protected categories like race, gender, ethnicity, or age. Therefore, there are no protected categories that have been intentionally used in model training or predictions.

4. Model Bias Analysis

Although explicit protected categories are absent, implicit biases could still exist due to correlations between accident locations and socioeconomic or demographic variables. Geographical areas might disproportionately represent specific demographic groups, potentially introducing spatial bias. Additionally, conditions related to the cars (e.g., vehicle type, age, and maintenance status), road conditions (e.g., quality, signage, and road design), and driver-related factors (e.g., experience level, age, and behavior patterns) could significantly influence the model's predictions. This implicit bias could inadvertently disadvantage certain communities or populations if resources are unevenly allocated due to these biased predictions.

5. Bias Removal Strategies

Several strategies can be implemented to mitigate potential bias：:

- **Feature Neutralization:** Reducing the influence of highly localized data by employing generalized regional features rather than specific coordinates can minimize spatial bias.
- **Data Balancing:** Utilizing techniques like Synthetic Minority Over-sampling Technique (SMOTE) to create balanced data across different geographic and severity categories can reduce bias.
- **Regular Audits and Monitoring:** Implement periodic evaluations and fairness audits across different regions and communities to continuously identify and address emerging biases.
- **Explainable AI:** Providing clear explanations for predictions can help stakeholders understand and challenge potential biases.

- **Incorporating Additional Contextual Features:** Explicitly including or adjusting for factors related to vehicles, road conditions, and driver behaviors can help neutralize biases arising from these implicit correlations.

Implementing these strategies would result in fairer, more equitable predictions, enhancing trust and reliability.

6. Other Risks and Stakeholder Impact

For our risks and impacts on stakeholders include:

- **Customers and Citizens:** Erroneous severity predictions may lead to delayed emergency responses or misallocation of resources, potentially endangering lives and property.
- **Environmental Impact:** Inaccurate assessments of environmental influences could cause inefficient deployment of resources, impacting environmental protection initiatives.
- **Government and Infrastructure Planning:** Over-reliance on the model without frequent updates and validations could result in suboptimal planning and resource allocation, exacerbating traffic congestion or mismanagement of infrastructure projects.
- **Data Privacy and Security:** The extensive use of geographical data raises concerns regarding privacy, especially if location data inadvertently reveals personal or demographic information.

To mitigate these risks, regular validation, updating predictive models, and transparent communication of the model's limitations and uncertainties are essential. Stakeholder engagement, including citizens and policymakers, can help enhance the model's robustness and applicability in practical scenarios.