

Understanding the Leading Causes of Road Traffic Accident

Maureen and Co.



Overview

Focus

This presentation focuses on understanding the factors that contribute to severe road traffic accidents using data-driven analysis.

Objective

The objective is to identify patterns and conditions under which accidents are more likely to result in severe outcomes, translating insights into actionable recommendations.

Objectives



Analyze

Explore crash data to understand patterns and trends in accidents.



Develop

Build a predictive machine learning model to classify cause of car crash.



Evaluate

Assess model performance using accuracy, precision, and recall metrics.



Recommend

Provide data-driven recommendations for improving road safety.

Business & Data Understanding

- ✓ Reducing the severity of road traffic accidents is a critical priority for public safety and healthcare systems.
- ✓ Severe accidents place immense strain on emergency services and the economy at large.
- ✓ **Key Question:** What are the primary contributory causes of car crashes?
- ✓ We utilized historical crash data containing road conditions, weather, and speed limits.



Data Preparation

- ✓ **Focused Analysis:** Concentrated on driver-related crash records to pinpoint behavioral factors.
- ✓ **Dataset Merging:** Combined crashes, people, and vehicles datasets into a comprehensive single source.
- ✓ **Feature Engineering:** Grouped related crash causes into broader categories to reduce complexity.

Exploratory Data Analysis

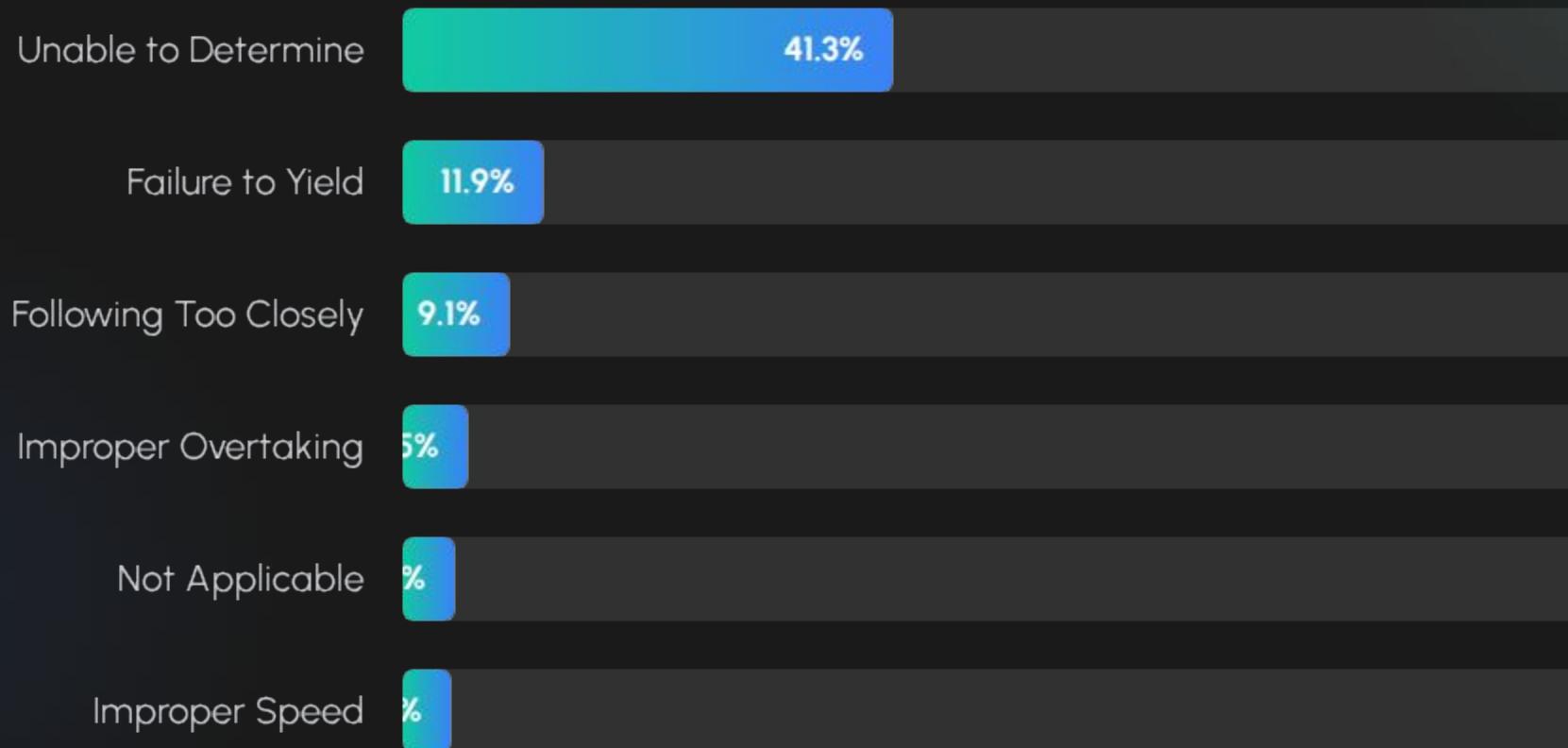
We explored the data to understand accident severity and spot clear patterns. Key factors analyzed included road surface, weather conditions, and speed limits.

Key Insight: Severe accidents were found to be more common on poor road surfaces, in adverse weather conditions, and at higher speeds.

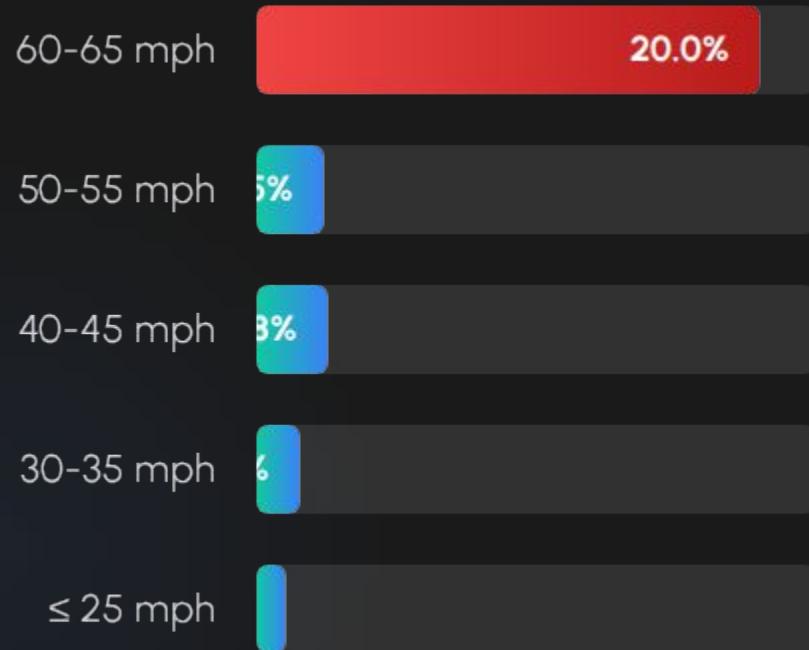
These insights guided the selection of the most critical variables for our predictive modeling.



Top Contributory Causes of Crashes



Severe Injury Rate by Speed Limit



Speed Kills

- ✓ The likelihood of severe injury increases drastically with speed.
- ✓ The highest risk (~20%) is observed in **60-65 mph** zones.
- ✓ This data strongly correlates higher speed limits with more catastrophic accident outcomes.

Modelling Approach



Feature Selection

Focused on location, time, and weather factors to pinpoint exactly when and where accidents are most likely to occur.



Class Grouping

Grouped related causes into 6 core categories to reduce noise and improve the model's focus on key patterns.



Imbalance Handling

Used label encoding and sample weighting to penalize errors in minority classes, preventing overfitting.

Model Evaluation

Performance Metrics

The XGBoost model achieved a weighted average accuracy of **51%**.

- ✓ **71% Recall** for Traffic Violations.
- ✓ **87% Recall** for Environmental Causes.
- ✓ The model is strong at identifying risk conditions on these two classes.

High Recall Areas

87% **71%**

Environment

Violations

Top Feature Importance



Road surface conditions and specific crash types are the most significant predictors of severity.

Recommendations



Targeted Enforcement

Deploy increased traffic enforcement at high-risk locations identified by the model.



Driver Education

Develop mandatory refresher courses targeting high-risk behaviors like improper backing.



Infrastructure

Prioritize upgrades at recurring accident spots and improve traffic control devices.



Environment

Deploy variable speed limits that adjust based on real-time weather conditions.

Next Steps

- ✓ **Real-World Testing:** Implement the recommendations in a pilot area to measure the actual reduction in accident rates.
- ✓ **Regular Updates:** Routinely update the prediction models as new data becomes available to maintain accuracy.
- ✓ **Stakeholder Engagement:** Present findings to city planners and traffic authorities for policy integration.



Conclusion

Accidents are driven by a combination of driver behavior, environmental conditions, and infrastructure. With targeted interventions and improved data, these insights can support effective strategies to reduce accident severity.

Questions?

Thank you for your attention.

Maureen and Co.