ANALYSIS OF TRAFFIC ACCIDENTS CAUSES, INSIGHTS AND RECOMMENDATIONS

INTRODUCTION

- **Problem Statement:** The City of Chicago faces serious road safety issues due to frequent traffic accidents, prompting a need to analyze crash data to identify their main causes and develop targeted solutions to improve safety for residents.

- **Objective:** Determine the main factors behind traffic accidents and provide practical recommendations for their prevention.
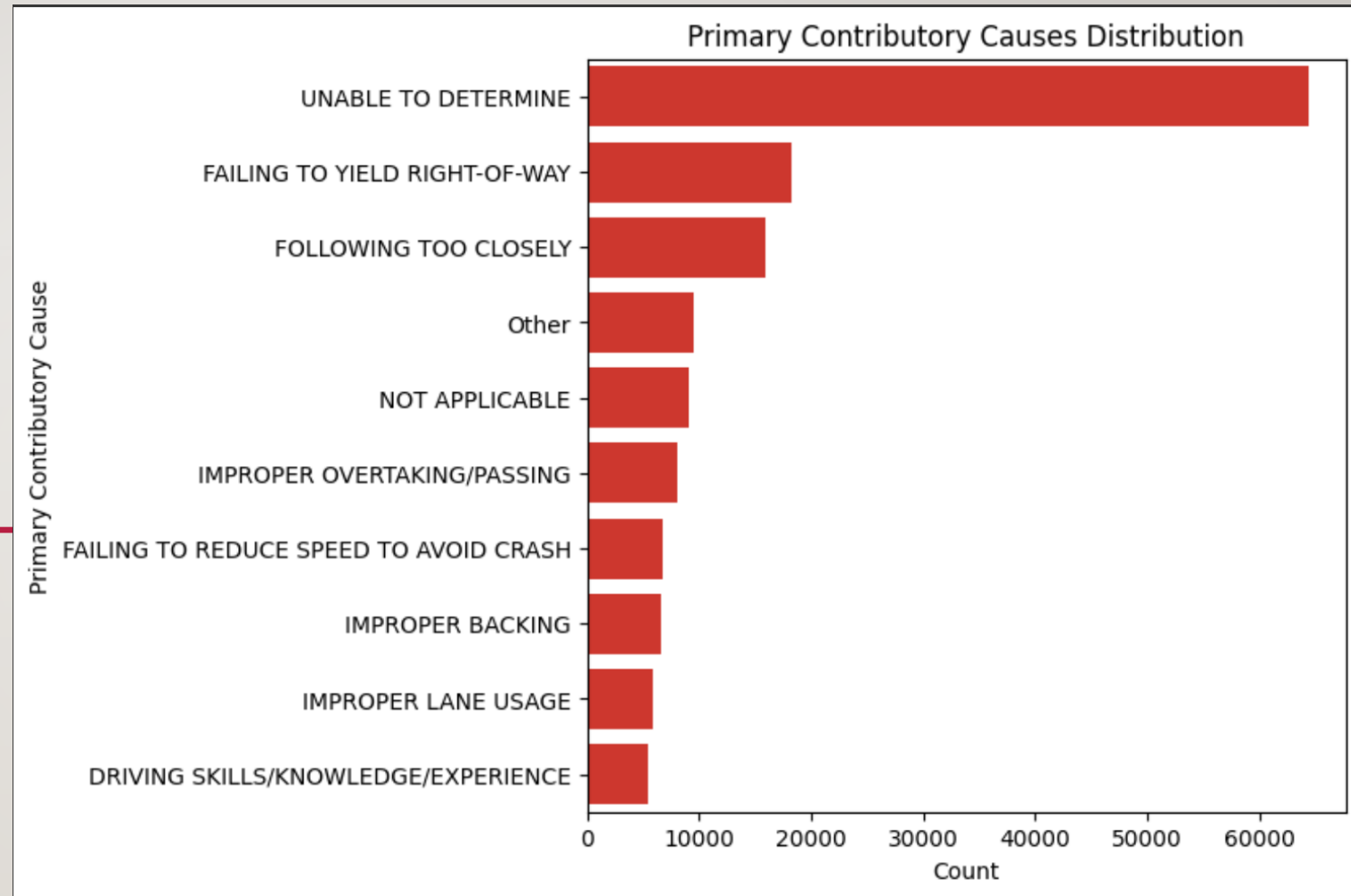
DATA PREPARATION

- Addressed missing data through imputation or column removal,

- Engineered cyclical and interaction features for time-related complexity, and

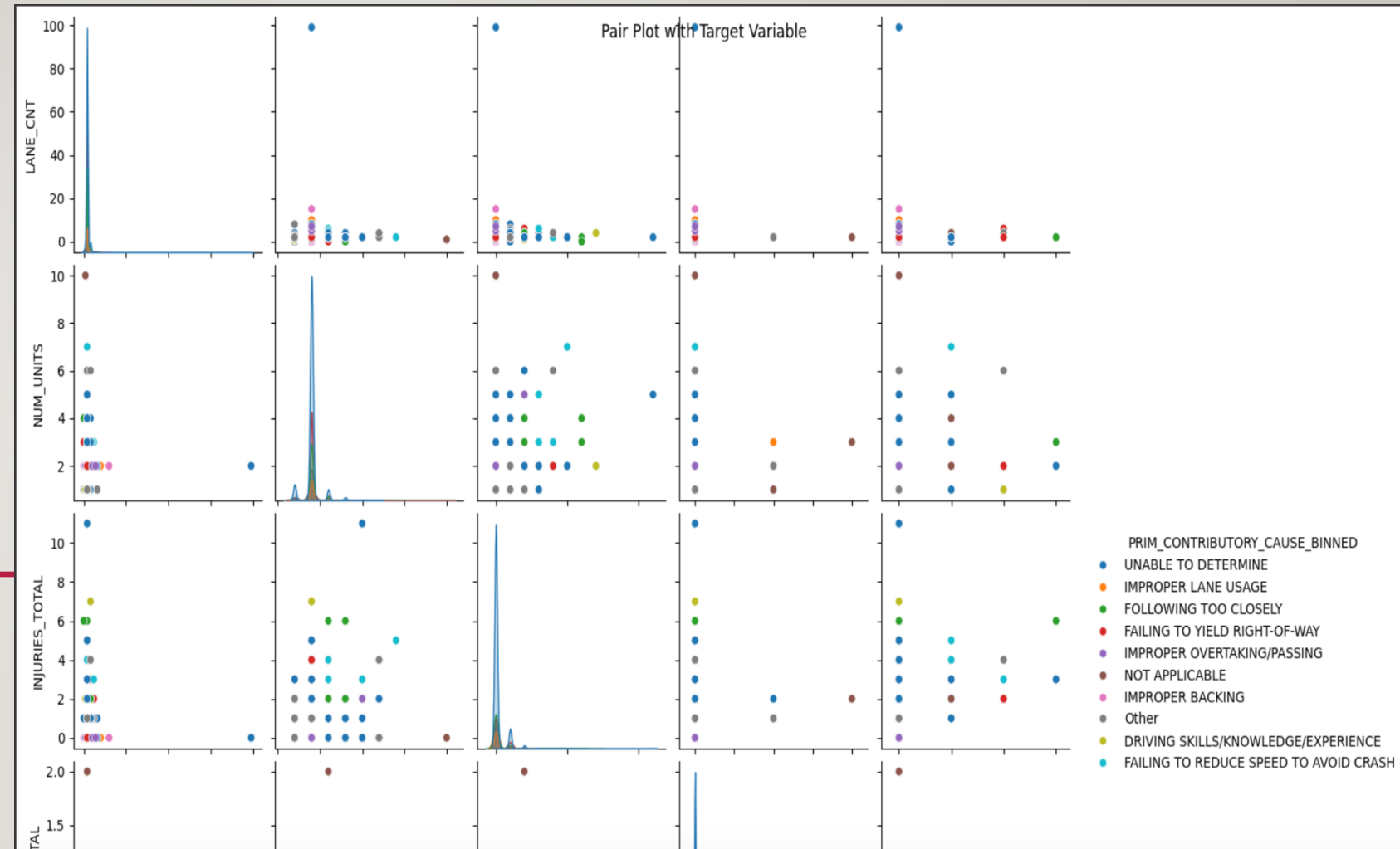- Grouped rare causes into an "Other" category to optimize model performance.

EXPLORATORY DATA ANALYSIS (EDA)

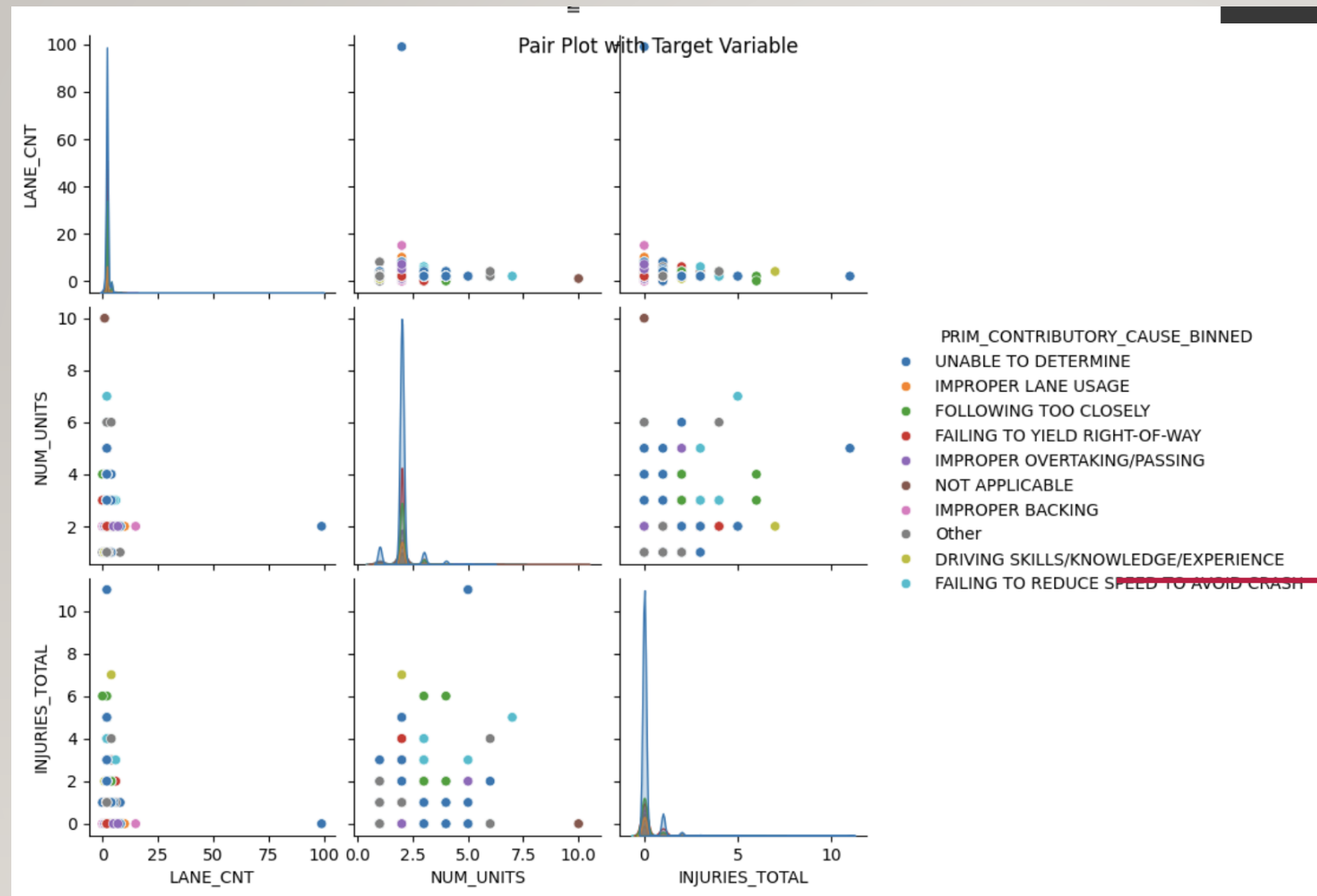# EXPLORATORY DATA ANALYSIS (EDA)

Correlation Heatmap of Numerical Features
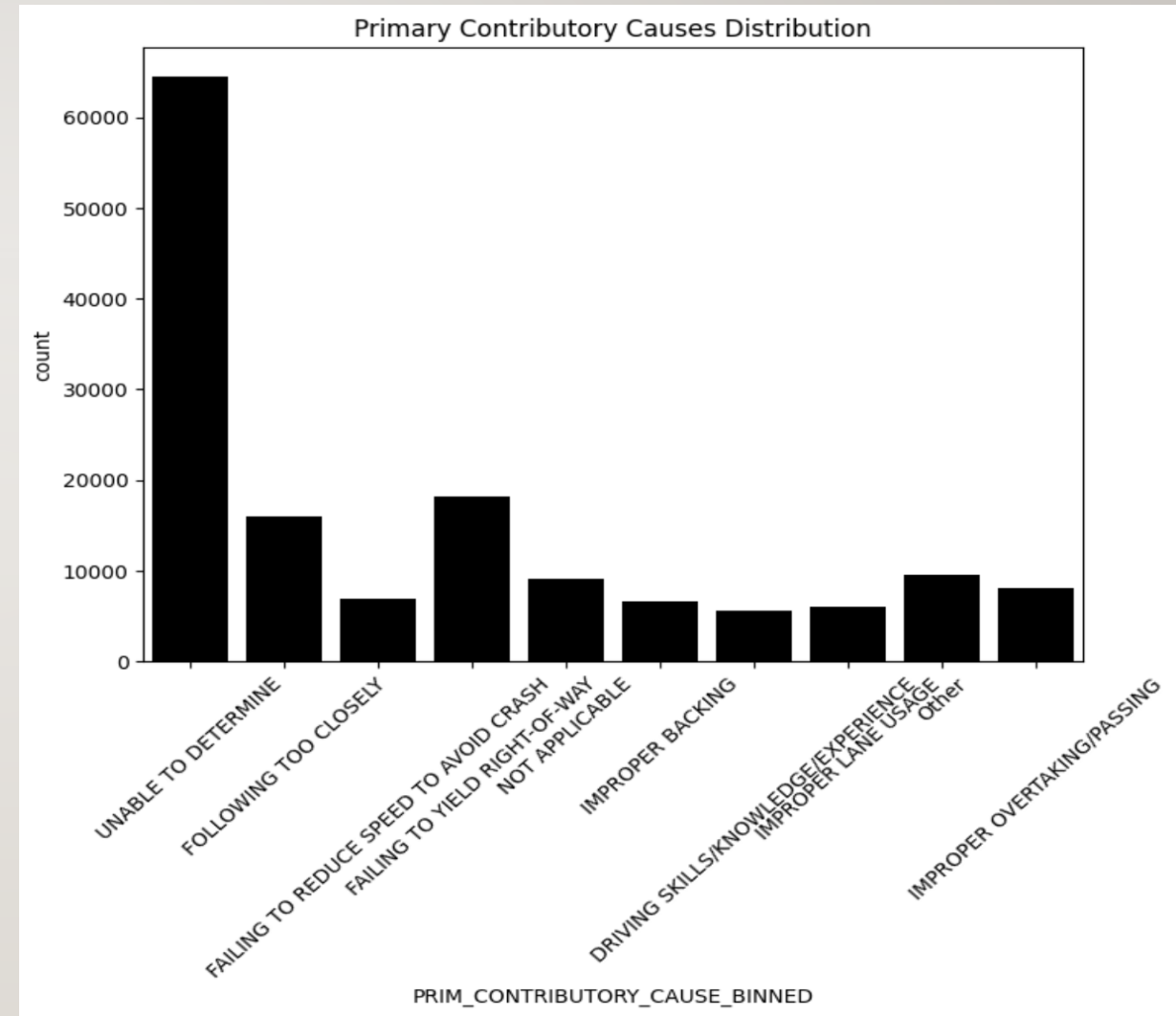
EXPLORATORY DATA ANALYSIS (EDA) UNIVARIATIVE ANALYSIS

EXPLORATORY DATA ANALYSIS (EDA) MULTIVARIATIVE ANALYSIS

Pair Plot with Target Variable

EXPLORATORY DATA ANALYSIS (EDA) MULTIVARIATIVE ANALYSIS

MODELING APROACH & EVAPUATION METRICS

Models tested include: Gradient Boosting, Random Forest, Lasso Logic Regression, Decision Tree, Ridge Classifier,

Evaluation Metrics

- Accuracy: The proportion of correct predictions made by the model.

- Precision: The ability of the model to correctly identify positive instances.

- Recall: The ability of the model to capture all positive instances.

- F1-Score: The harmonic mean of precision and recall, balancing the two.

- Confusion Matrix: A detailed breakdown of the model's performance across different classes.

MODEL COMPARISON

From all the models mentioned above that were tested, Gradient boosting emerged the best given the size of the data in play for this project.

- The Model combines the strengths the flexibility of Decision Trees, allowing it to model complex non-linear relationships, handle interactions between features and generalize well on unseen data. Its performance were further enhanced through Hyperparameter Tuning, which likely led to it outperforming the other models tested.

MODEL COMPARISON CHART

Model Performance Comparison
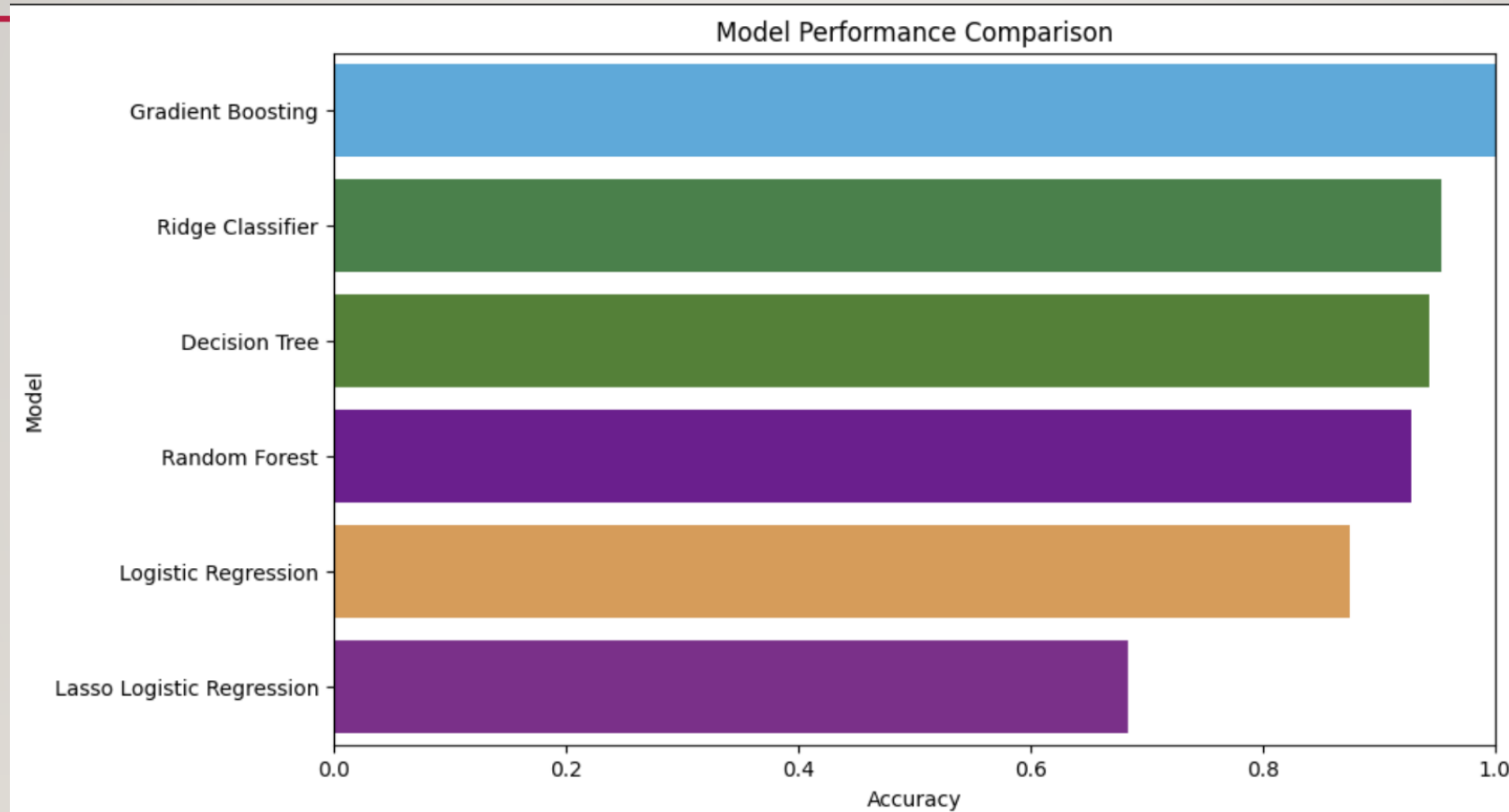
DATA MODELING

Several machine learning models will be trained and evaluated, including:

1. Logistic Regression

2. Regularized Logistic Regression

3. Decision Trees

4. Random Forest

5. Gradient Boosting

BEST MODEL

- Best model in our case as indicated on the chart is Gradient Boosting

- Gradient Boosting emerged as the top-performing model for this problem primarily due to the large size of the dataset. It leverages the strengths of ensemble learning and the flexibility of Decision Trees, enabling it to capture complex non-linear relationships, handle feature interactions effectively, and generalize well to unseen data. Additionally, its performance was further optimized through hyperparameter tuning, which likely contributed to its superior results compared to the other models.

RECOMMENDATIONS

- Policy Adjustments: Modify traffic regulations and safety measures based on observed patterns, such as enhancing signage or enforcing speed limits in high-risk zones.

- Enhanced Traffic Monitoring: Focus on targeted interventions in areas with high accident rates, addressing the identified contributory factors to mitigate risks effectively.

- Model Deployment: Deploy the top-performing model in a real-time traffic monitoring system to deliver proactive alerts and help reduce accident rates.

- Further Research: Expand this analysis by incorporating more recent data and additional features, such as vehicle types and driver demographics, to improve model accuracy and insights.

CONCLUSION

- The analysis of the traffic crash data has provided valuable insights into the primary contributory causes of accidents and the critical factors influencing road safety. By identifying the top features and causes associated with accidents, we can make informed recommendations to enhance road safety in the City of Chicago

- The insights gained from this analysis provide a clear pathway to making the city's roads safer for everyone by reducing traffic accidents, enhancing road safety, and ultimately save lives.

NEXT STEPS

- Putting into effect the recommendations made so as to enhance the road safety and reduce incidences and accidents on our roads

THANK YOU

Q & A