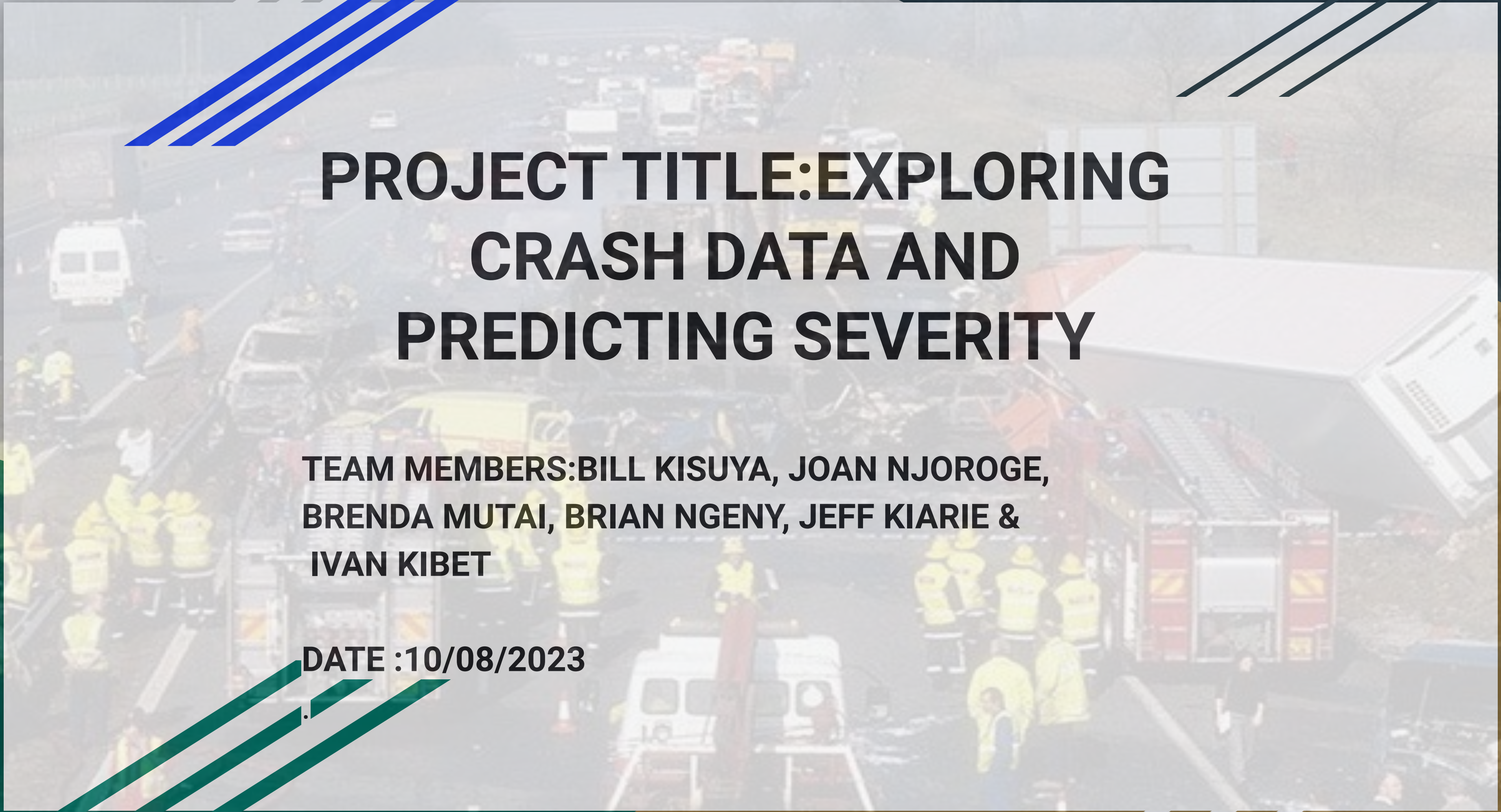

PROJECT TITLE:EXPLORING CRASH DATA AND PREDICTING SEVERITY

TEAM MEMBERS:BILL KISUYA, JOAN NJOROGE, BRENDA MUTAI, BRIAN NGENY, JEFF KIARIE & IVAN KIBET

DATE :10/08/2023

.

Table of Contents

INTRODUCTION

Welcome to the presentation on

"Exploring Traffic Crash Data and Predicting Severity."

In this project, we delve into a comprehensive analysis of traffic crash data to uncover insights and develop predictive models for crash severity.

Our aim is to contribute valuable insights to improve road safety and reduce the frequency and impact of traffic accidents.

DATASET OVERVIEW

We begin by presenting an overview of the dataset used for our analysis.

The dataset consists of information on various attributes related to traffic crashes, including weather conditions, road surface, vehicle types, and injury classifications.

This dataset serves as the foundation for our exploratory analysis and predictive modeling.

RESEARCH QUESTIONS

Our project addresses the following key research questions:

1. What factors significantly influence the severity of traffic crashes, and can they be used to predict severity levels?
2. How does driver behavior (e.g., distraction, speeding, impaired driving) contribute to different types of crashes, and can we identify patterns of behavior associated with higher crash risks?
3. How does the usage of safety equipment impact crash occurrences and outcomes?
4. Are there specific time periods and environmental conditions associated with higher crash rates?
5. Can we identify geographic hotspots with higher crash frequencies and recommend targeted prevention strategies?
6. How do road surface conditions and defects influence the likelihood and severity of crashes, and can we recommend road improvements based on this analysis?

DATA EXPLORATION

Our initial exploration of the dataset revealed important patterns and insights.

We observed correlations between variables such as weather conditions and crash severity, as well as temporal patterns related to crash occurrences during certain hours and days of the week.

Additionally, we investigated the influence of road surface conditions and defects on crash likelihood and severity.

MODEL APPROACH

Our modeling approach involved a combination of exploratory data analysis, preprocessing, and machine learning techniques.

We employed various models, including Random Forest Classifier, Decision Tree Classifier, Logistic Regression, and K-Nearest Neighbors, to address different research questions and predict outcomes.

In Model 1, we focused on predicting crash severity using features such as speed limit, weather conditions, and age of individuals involved.

The model achieved an accuracy of approximately 90%, with detailed precision, recall, and F1-score metrics provided for each severity class.

Model 2 explored the impact of safety equipment usage on crash outcomes, such as seatbelt and airbag deployment.

The model demonstrated promising results in predicting injuries based on safety equipment usage, with recall and precision scores reflecting its performance.

Model 3 analyzed temporal and environmental patterns, including crash occurrences by time of day and day of the week.

The KNeighborsClassifier model provided insights into time-based patterns and successfully predicted crash occurrences.

Model 4 identified geographic hotspots with higher crash frequencies and recommended prevention strategies.

Logistic Regression was employed to predict crash occurrence in specific locations, enabling targeted safety interventions.

Model 5 investigated the influence of road surface conditions and defects on crash likelihood and severity.

K-Nearest Neighbors model was utilized to analyze and predict the impact of road conditions on crash outcomes, achieving an accuracy of approximately 51%.

Model Performance Comparison

A comparative analysis of model performance highlighted the strengths and limitations of each approach.

We evaluated accuracy, precision, recall, F1-score, and other relevant metrics to assess the reliability of our predictive models.

Recommendations

Our insights lead to several key recommendations:Emphasize seatbelt and safety equipment usage to reduce injury severity.

Implement targeted traffic control measures in identified geographic hotspots.

Consider road surface improvements based on analysis of road conditions and defects

Next Steps

Moving forward, we suggest the following next steps:Incorporate real-time data for more accurate predictions and intervention strategies.

Expand the analysis to include additional variables, such as road conditions and driver behavior.

Collaborate with local authorities to implement and assess the effectiveness of recommended strategies.

Conclusion

In conclusion, our project demonstrated the power of data analysis and predictive modeling in understanding and addressing traffic crash severity.

By identifying influencing factors, patterns, and potential road improvements, we can contribute to safer roads and improved road safety measures.

Thank You

Thank you for your attention and interest in our presentation.

For further discussions, questions, or collaborations, please feel free to reach out to us for more information