Forecasting Motor Vehicle Crash Incidents in New York City Using Prophet

Avinash Bunga

Master of Science in Information Systems and Business Analytics

Park University

CIS607HOF1P2024 Applied Business Forecasting

Professor: Dr. Abdelmonaem Jornaz

September 8, 2024

Forecasting Motor Vehicle Crash Incidents in New York City Using Prophet

Introduction

Motor vehicle crashes are a significant issue in New York City, leading to injuries, fatalities, and economic losses. These incidents disrupt the daily lives of pedestrians, cyclists, and drivers, posing ongoing challenges for city planners and public safety officials who aim to make the roads safer. With accurate predictions of when and where crashes are likely to occur, authorities can take proactive measures to prevent them. This project focuses on using historical crash data to forecast future incidents, providing insights that can help reduce accidents and improve overall traffic safety (Data.Gov, 2024).

The Aim to Predict: The primary aim of this study is to work on a predictive model that forecasts motor vehicle crashes in New York City using the Prophet model. The model identifies patterns related to time, location, and contributing factors, allowing for the prediction of high risk periods. By understanding these patterns, the project supports data driven decisions for deploying preventive measures, allocating resources, and enhancing road safety.

Link to the Dataset: The data for this project is sourced from the Data.gov website, which provides publicly available datasets. This dataset records detailed information on motor vehicle collisions reported by the NYPD, including crash dates, times, locations, contributing factors, and vehicle types Motor Vehicle Collisions - Crashes Dataset (Data.Gov, 2024).

Forecasting Model to Be Used: The Prophet model is chosen for this project because it is easy to use, interpretable, and effective for time series forecasting.

Why Prophet?

- Prophet is specifically designed to model time series data with trends and seasonal patterns, which are common in crash data.

- It automatically handles missing data, outliers, and shifts in trends, making it suitable for real world datasets like this one.
- The model provides clear visual outputs that separate the trend, seasonality, and other factors, making it easy to understand what drives the forecasts (Wakamiya, 2020).

Implementation Plan:

1. **Data Preparation**: Clean and format the crash data, ensuring the dates are correct and addressing any anomalies that could affect forecasting.
2. **Model Training**: Train the Prophet model using the cleaned data, focusing on identifying seasonal and trend patterns relevant to crash occurrences.
3. **Forecasting**: Generate predictions of future crash occurrences to identify high risk times and suggest preventive measures.
4. **Interpretation**: Analyze the outputs to understand how different factors, such as time of day or specific months, impact crash rates.
5. **Evaluation**: Validate the model's predictions against actual past data to assess accuracy and make any necessary adjustments (Wakamiya, 2020).

Conclusion

This study aims to provide actionable insights that can help reduce motor vehicle crashes in New York City. By forecasting crash occurrences, the project supports the city's efforts in improving traffic safety and protecting its residents. The resulting forecasting tool will highlight high risk periods, guiding interventions that can save lives and enhance the overall safety of New York City's streets.

References

Data.Gov (2024, September 6). *Motor Vehicle Collisions - Crashes*. https://catalog.data.gov/dataset/motor-vehicle-collisions-crashes

Wakamiya, C. (2020). *Introduction to Time Series Forecasting*. Berkeley. https://datax.berkeley.edu/wp-content/uploads/2020/09/slides-m215-time-series.pdf