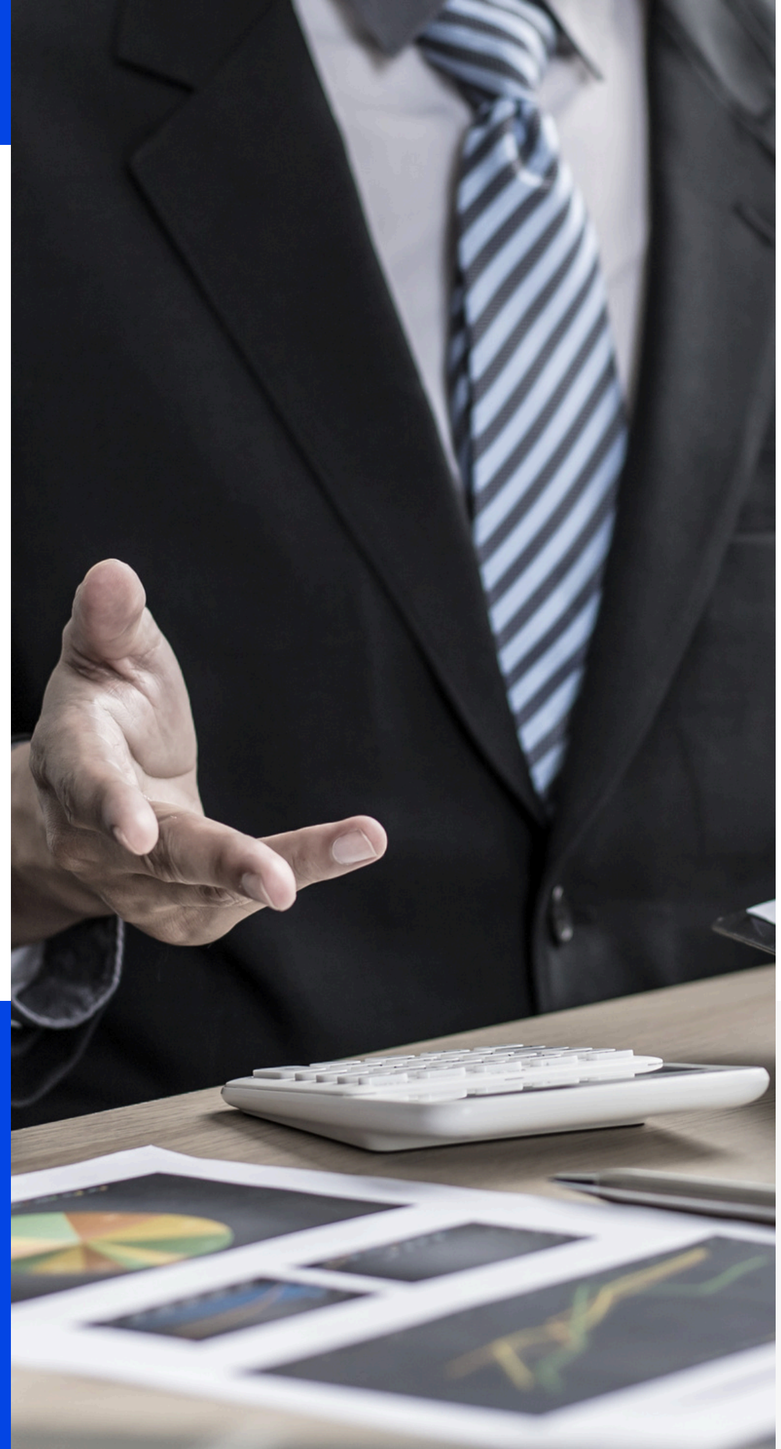
Luwate Inc.

MAKE ROADS SAFE

December 2023

medium.com/Luwate

github.com/Luwate



CONTENT

OVERVIEW

Road accidents are the leading cause of death for individuals under the age of 30. Therefore it is in the interest of all parties to attempt to reduce this figure. This project aims to design a model that will predict the primary contributory cause of car accidents based on available data such as vehicle information, occupant details, road conditions, and environmental factors.

Luwate Inc.

01 **Problem Statement**

The goal is to build a predictive model that accurately identifies the primary contributory cause of accidents using available data on vehicles, drivers, road conditions, and environmental factors.

02 **Expected Benefits**

Through the success of the project the Transportation Safety Board will be able to understand what factors lead to accidents and will inform policy creation in order to make the roads safer.

03 **Business Success Criteria**

Success of the project will be evaluated through creation of a model which can predict the primary cause of an accident where one has occurred.

THE DATA

Luwate Inc.

Source

The data has been collected by the Chicago Police department and dates between 2015 and December 2024. It has 899,000 records with 48 columns

Scope

The most prominent risk of this project is the existence of an unbalanced dataset. The causes of accidents may not be evenly spread hence, some factors may be more likely than others, leading to issues in the modelling process.

Mitigation

In order to mitigate this SMOTE will be used in the event of an unbalanced target variable

GOALS

The project will be considered a success following the completion and evaluation of a decision tree classifier to predict the main reason for accidents.

Luwate Inc.

01 **Data Preparation**

Clean the data

02 **Exploratory Data Analysis**

Perform Exploratory Data Analysis

03 **Model Development**

Develop a decision tree and iterate upon it to improve performance metrics. In addition, the model will be evaluated.

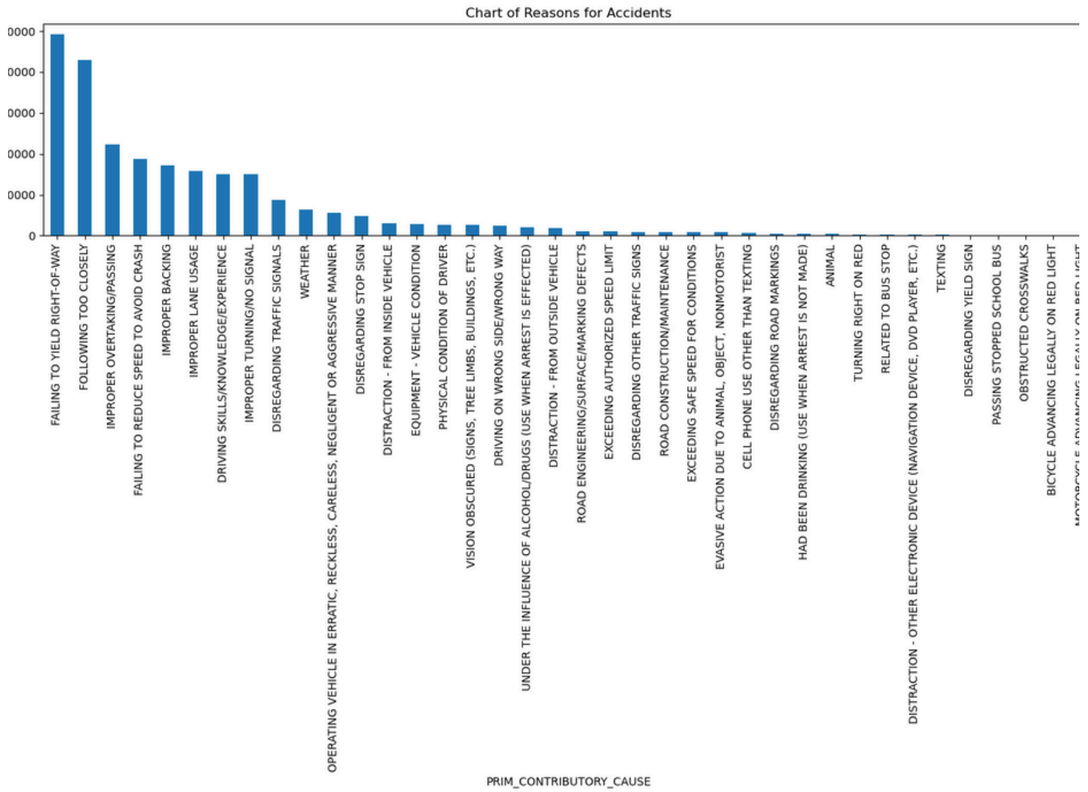
Luwate Inc.

ANALYSIS

github.com/Luwate

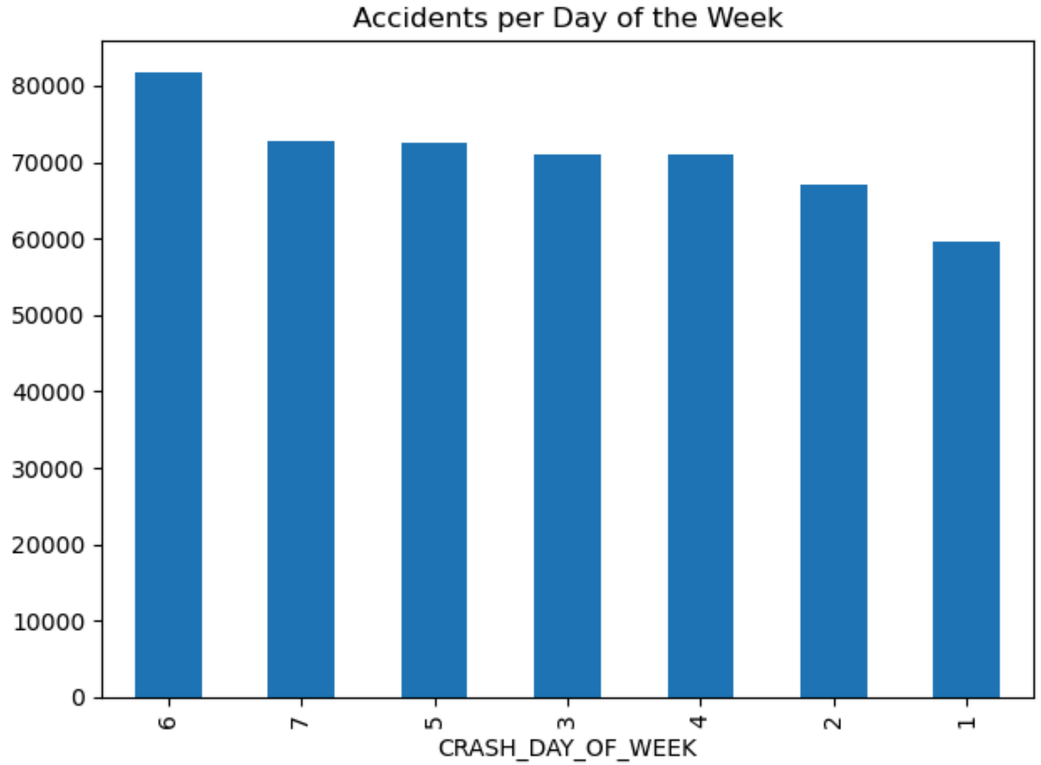
ANALYSIS

Luwate Inc.



Causes of accidents

The majority of accidents are as a result of human error.

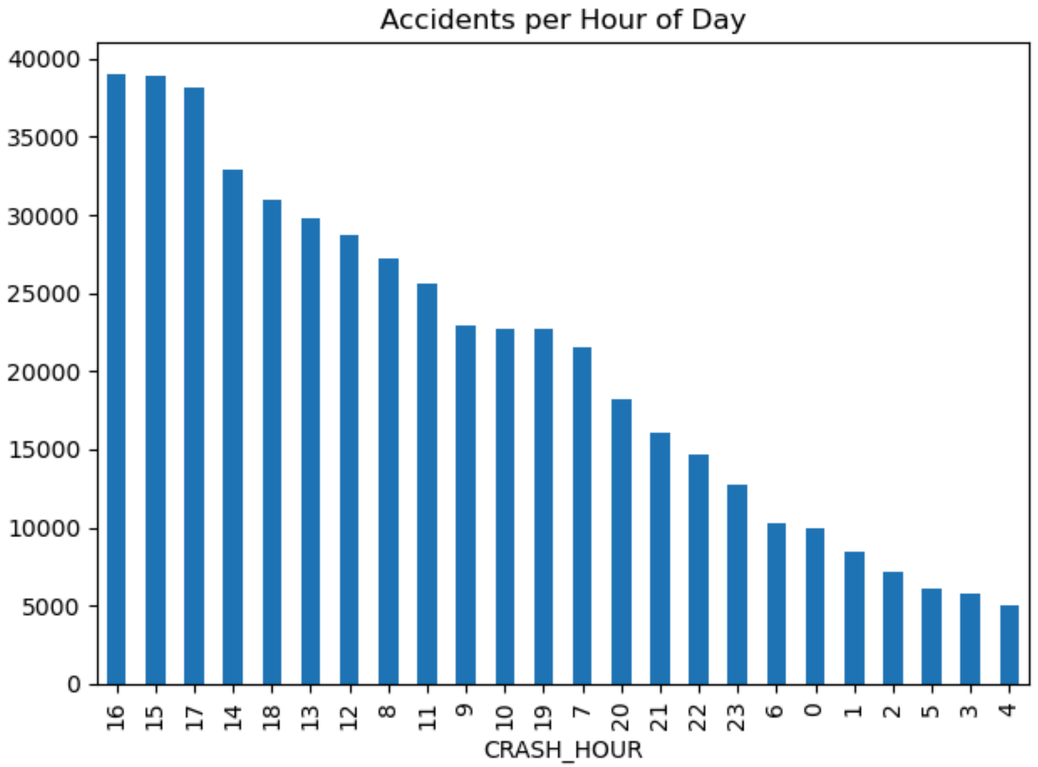


Accidents per Day of the Week

Friday is the most common day for accidents to occur.

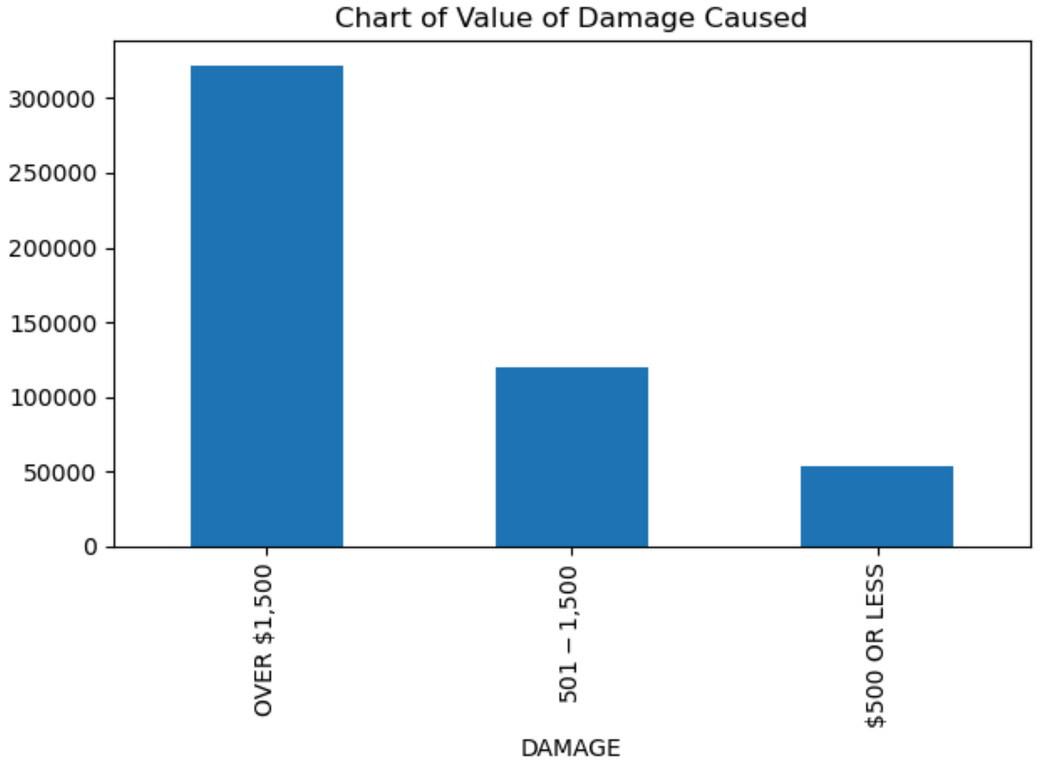
ANALYSIS

Luwate Inc.



Accidents per Hour of the Day

The evening rush hour time seems to be the most common time when accidents occur



Damage Caused

Most accidents, expectedly lead to damage bills of over $1,500

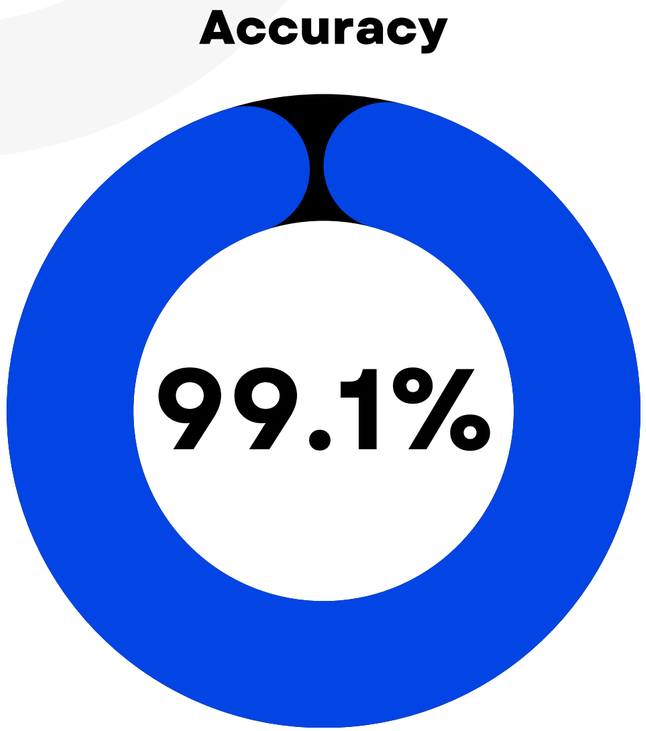
Luwate Inc.

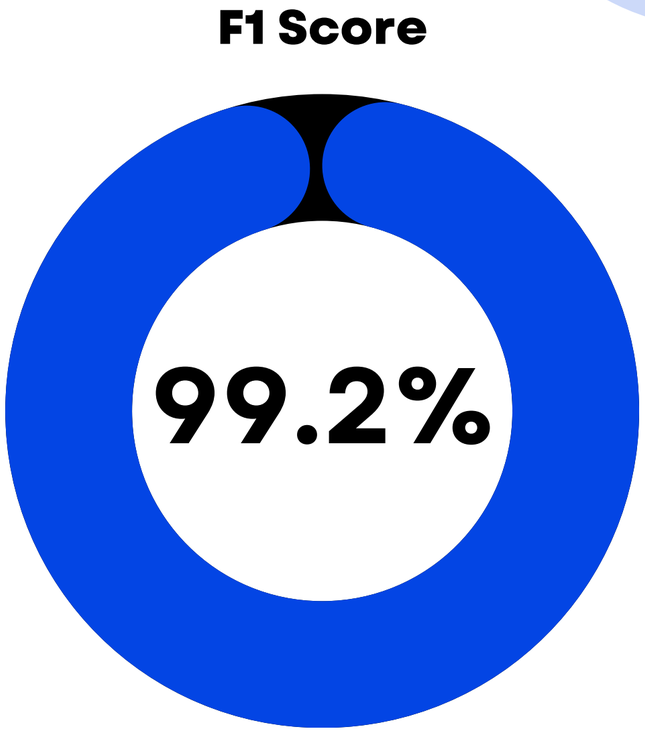
MODELLING

github.com/Luwate

TRAINING

The final model performs extremely well on the train set.



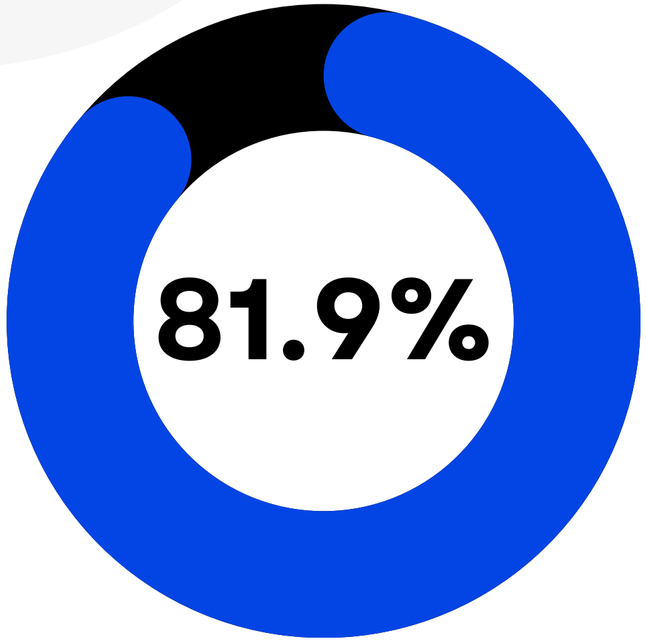


Relatively similar accuracy and F1 scores means that the model is likely not naively predicting the most common class.

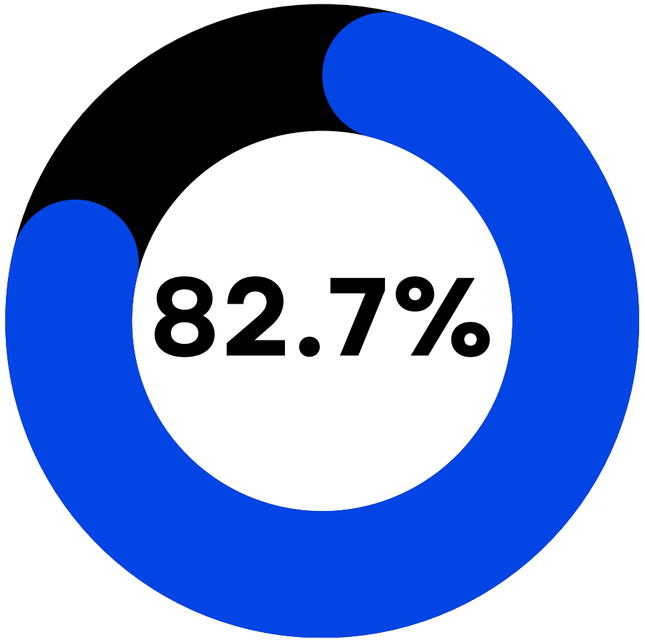
TESTING

The final model performs moderately well on the test set. The drastic drop in performance indicates some level of overfitting.

Accuracy



F1 Score



Relatively similar accuracy and F1 scores means that the model is likely not naively predicting the most common class.

RECOMMENDATIONS

Luwate Inc.

Improvements to Signage

Human error seems to be the overwhelming cause of accidents in the dataset, with failing to yield being the number one cause, improvements should be made to the signage to clearly indicate yield points.

Training

Better training should be provided to drivers in order to reduce the number of accidents caused by human error, which seem to be the overwhelming majority.

Different Modelling Techniques

In order to get more useful models, modelling techniques that build on the work of decision trees such as random forests can be used.

NEXT STEPS

Luwate Inc.

More Advanced Techniques

The model can be improved upon using more advanced machine learning techniques such as Principal Component Analysis.

Different Modelling Techniques

A different classification machine learning model may be used and results can be compared.

Luwate Inc.

THANK YOU

FOR YOUR ATTENTION

December 2024

medium.com/Luwate

github.com/Luwate