# CHICAGO CAR CRASHES

**BELLA SOMWE** 



### INTRODUCTION

For this classification project, the aim is to use machine learning concepts to predicting whether or not a car crash is fatal.

# **OBJECTIVE**

To build a classification model to predict whether a car crash is fatal or not.

#### SPECIFIC OBJECTIVES

- To clean and merge the three datasets(crashes,people,vehicles) so as to make it easy for modelling
- Identify features in the data and use them for engineering a model
- 3. Evaluating our model
- 4. Come up with interpretations upon interpreting our classifier model

### **BUSINESS UNDERSTANDING**

The car industry wants to make their buyers feel safer by being able to identify which crashes are fatal or not so that emergency services can respond quickly to fatal cases and give them prority.

# DATA

Chicago, IL Car Crash Data - From city of Chicago website - 3 data sets:

Crash info

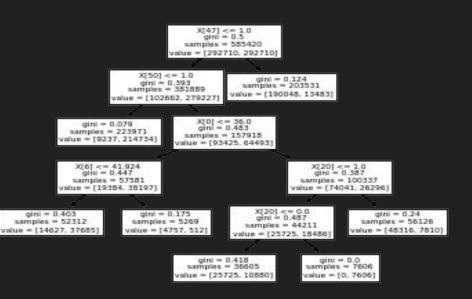
Vehicle info

info on People involved

## THE DECISION TREE

TEST ACCURACY 91.8%

TRAIN ACCURACY 90%



#### RECCOMENDATIONS

The best model for use is the decision tree because it has a relatively high accuracy of 91.8% meaning it is correct 91.8% of the time, however, the precision of 0.0079 indicates that, out of all the positive predictions made by the model, only 0.79% of them were actually positive. This is a very low precision value, which means that the model is making a large number of false positive predictions. This leads to the following reccomendations:

- Consider using different feature selection or engineering techniques to improve the performance of the model.
- Consider adjusting the parameters of the Decision Tree model to optimize its performance.

#### CONCLUSION

In conclusion, the Decision Tree model has a relatively high accuracy, but is not performing well in terms of precision and recall. Further work may be necessary to improve the performance of the model and achieve better results. Still, it would be beneficial to use this model as a safety feature because even though precision is low, better safe than sorry.