Car accident’s severity prediction

Daniel Orozco Venegas

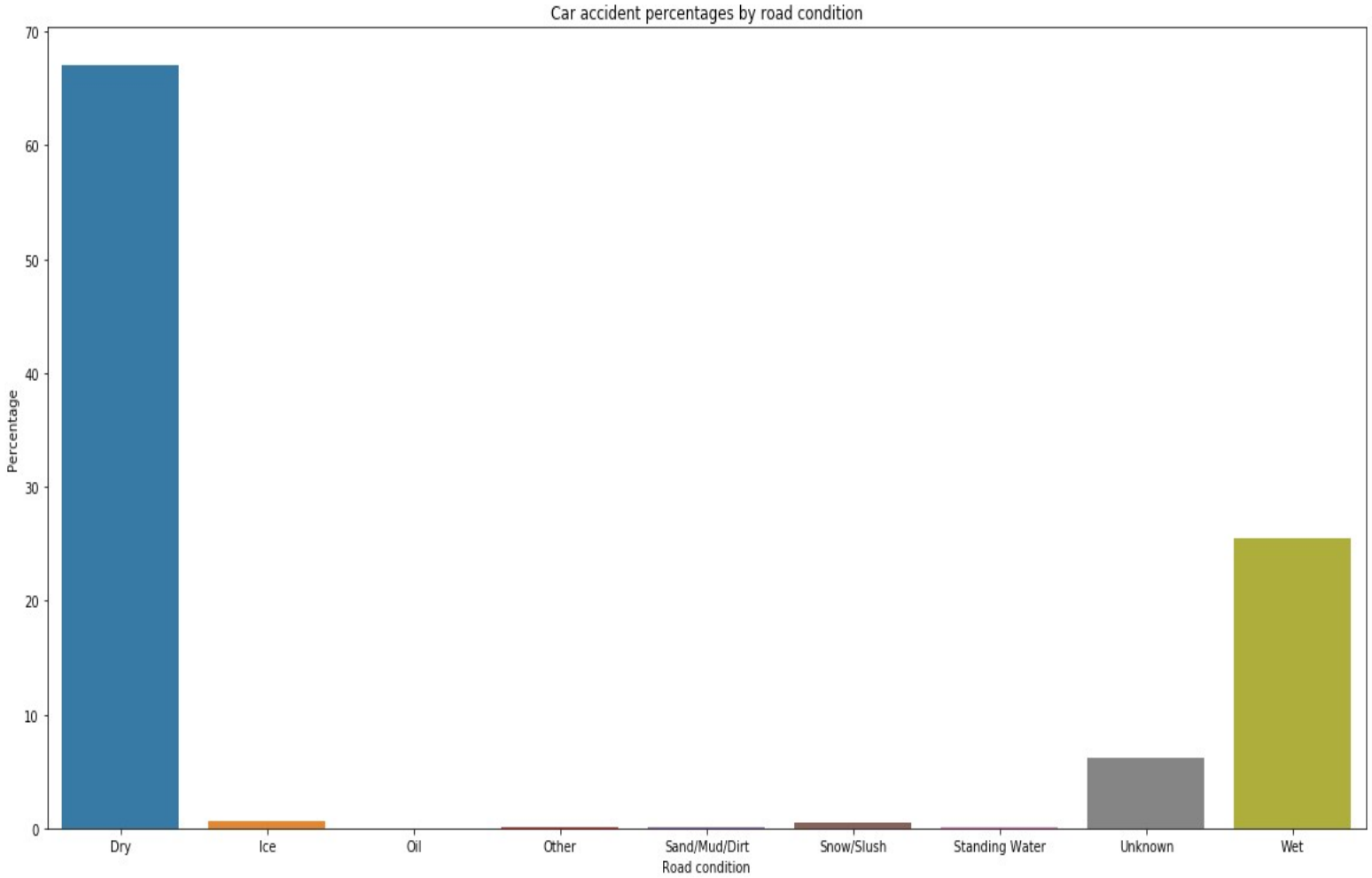
Introduction

- Car accidents impact quality of life.
- Car accidents can be prevented by knowing how likely a severe car accident is.
- Any government could use that knowledge to advice drivers and increase traffic police controls.
- Prediction for Seattle City, but can be implemented in other cities.

Data acquisition and cleaning

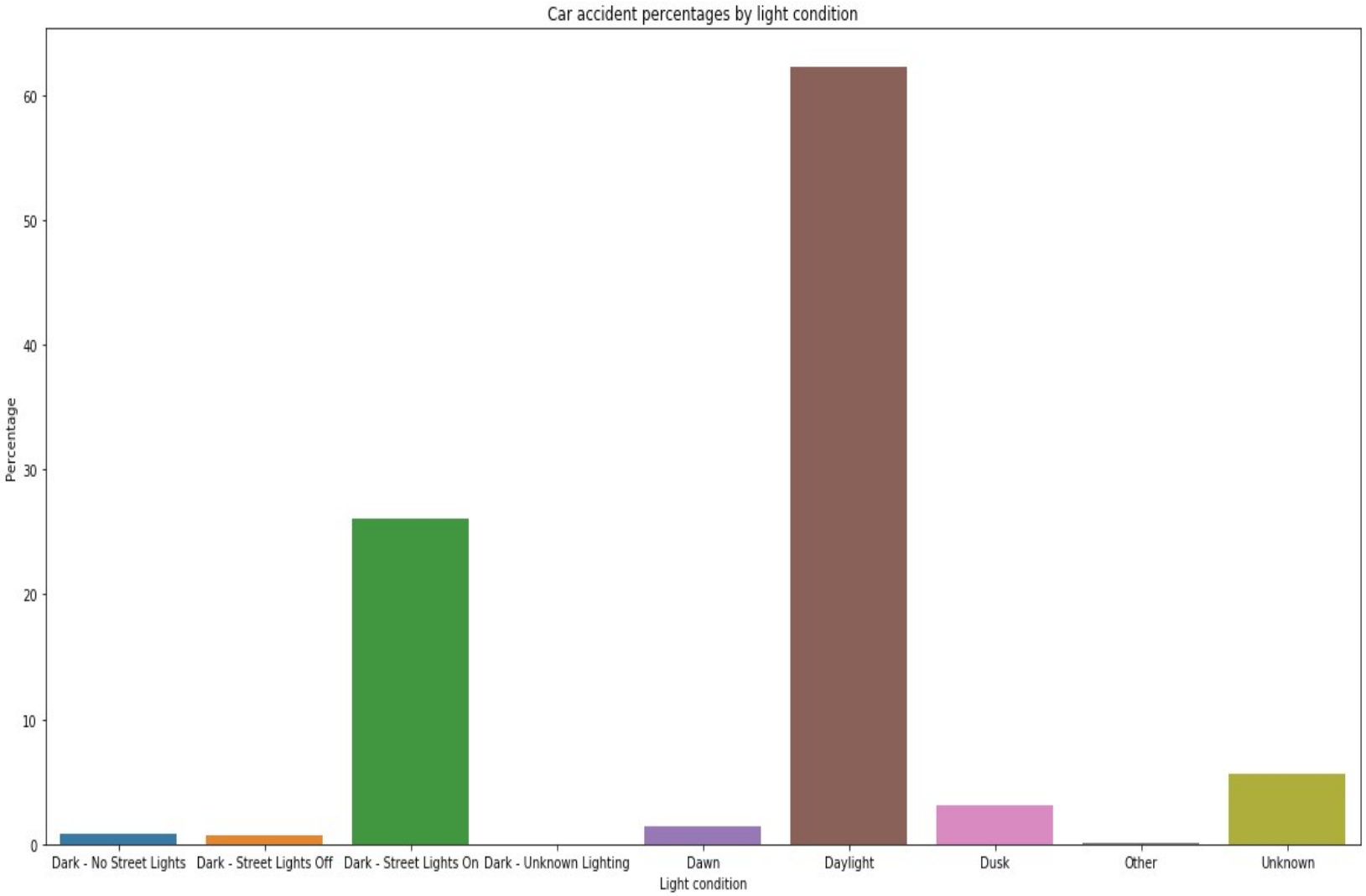
- Car accidents in Seattle City (2004-2020) from Kaggle dataset.
- In total, 194,673 rows and 40 columns in the raw dataset.
- Duplicate, highly similar, and metadata columns were dropped.
- Rows with empty values for some of the selected columns were dropped.
- Cleaned data contains 11 columns and 182895 rows.

Most car accidents happen with dry roads, but wet roads cause more severe accidents



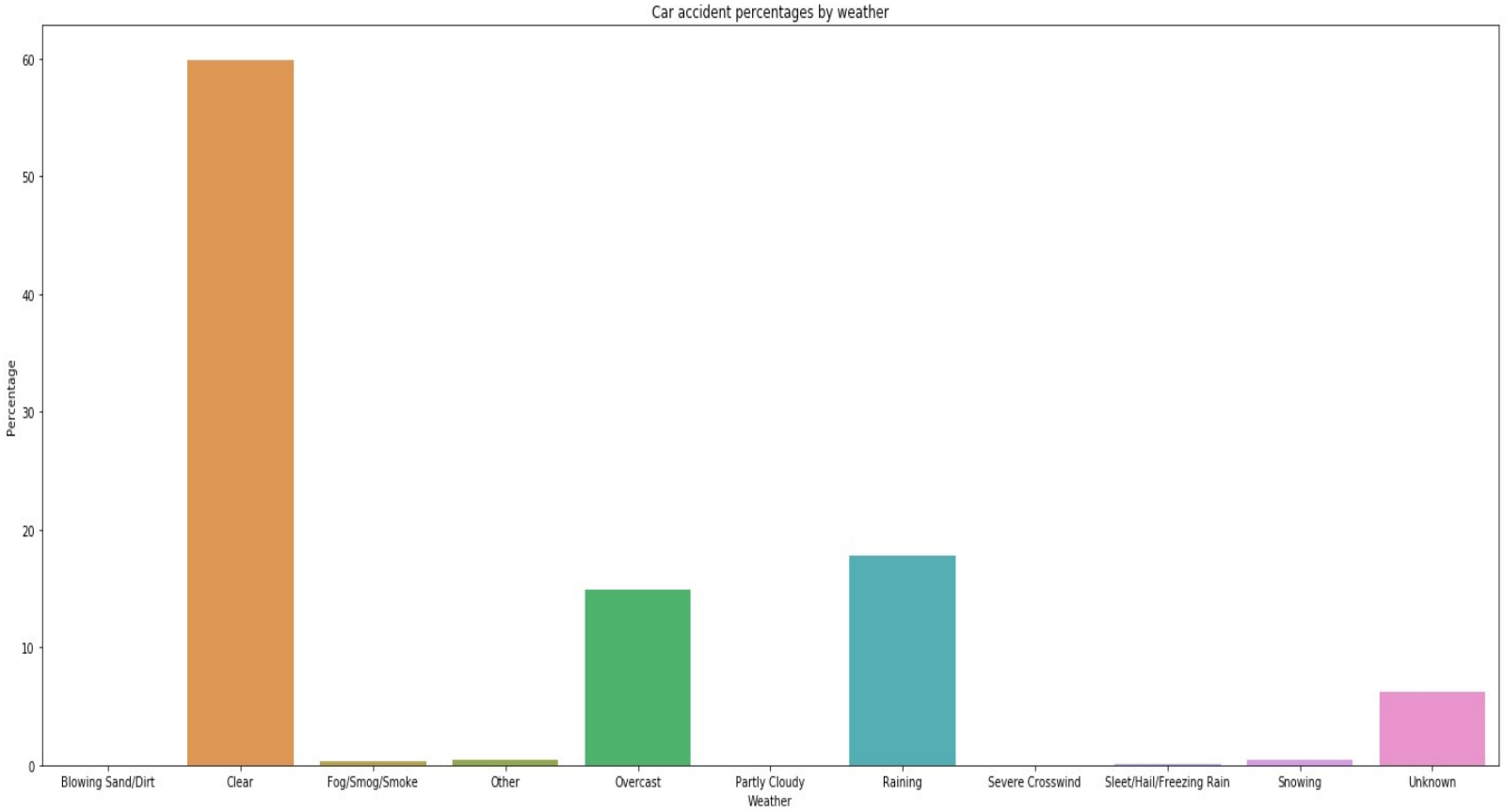
~45% of car accidents with wet roads ended with injured people. Only ~8% with dry roads ended with injured people.

Most car accidents happen with daylight, and many of them are severe accidents



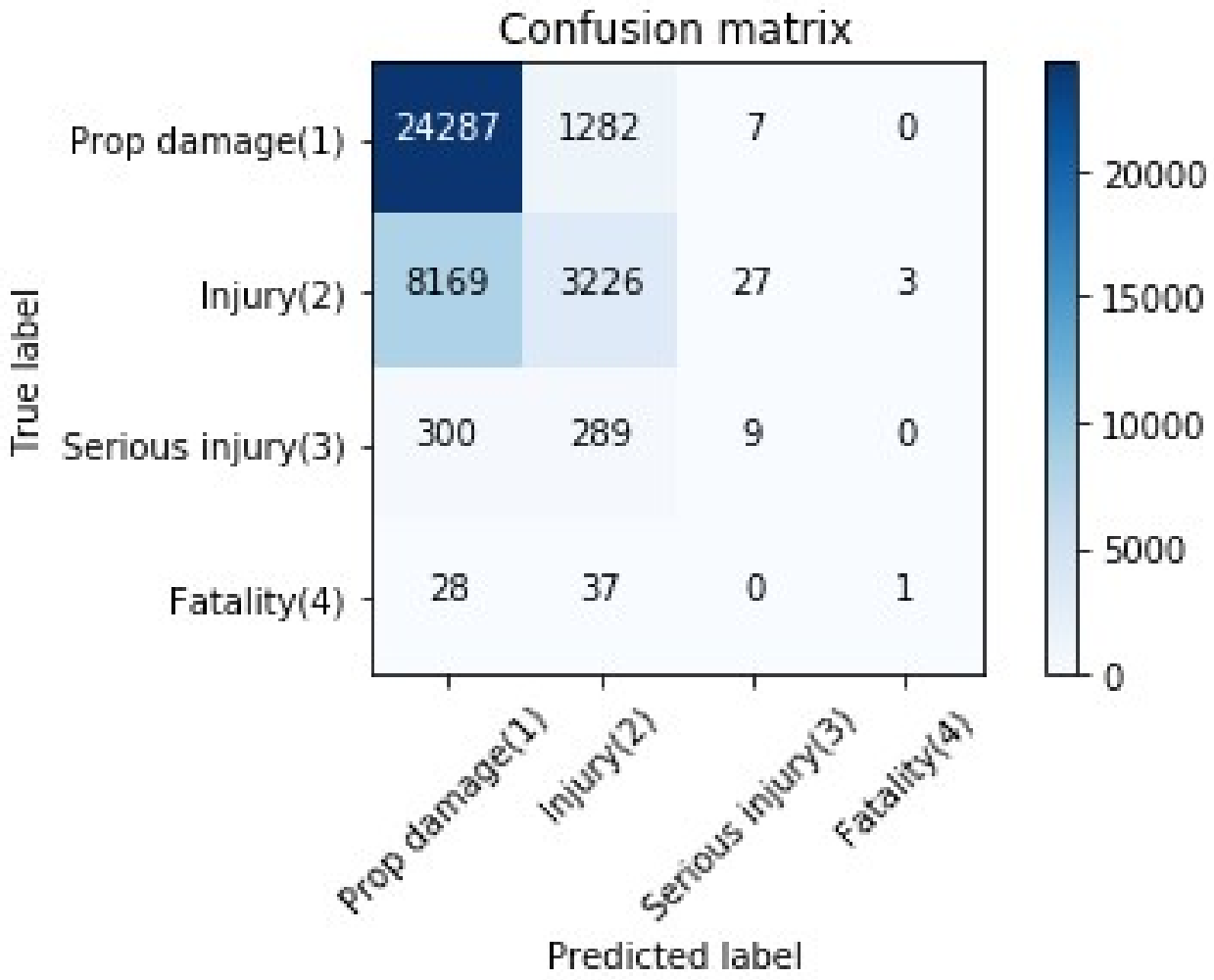
~62% of car accidents with daylight.
~21.5% were under daylight and were severe. Only ~8% with dark light ended with injured people.

Most car accidents happen with clear weather, but overcast and raining weathers cause more severe accidents



~6% of car accidents had raining weather and were severe. ~5% were under overcast weather and were severe.

Random forest model performance



Accuracy score: 73.07%
Confusion matrix:
Property damage: 95.38%
Injury: 28.03%
Serious injury: 1.47%
Fatality: 1.43%

Conclusion and future directions

- Built useful model to predict car accident's severity based on external conditions.
- Accuracy of the model has room for improvement, specially for the highest severities.
- Include speeding as one of the features in the prediction, and more samples for higher severity car accidents.