

Chicago Car Crash Analysis

Predict the primary contributory
cause of a car accident



What are we predicting (binned targets)

1. Sober

- Texting
- Right on red
- Exceeding speed limit
- Following too closely
- Operating vehicle reckless manner
- ETC.

2. Influenced

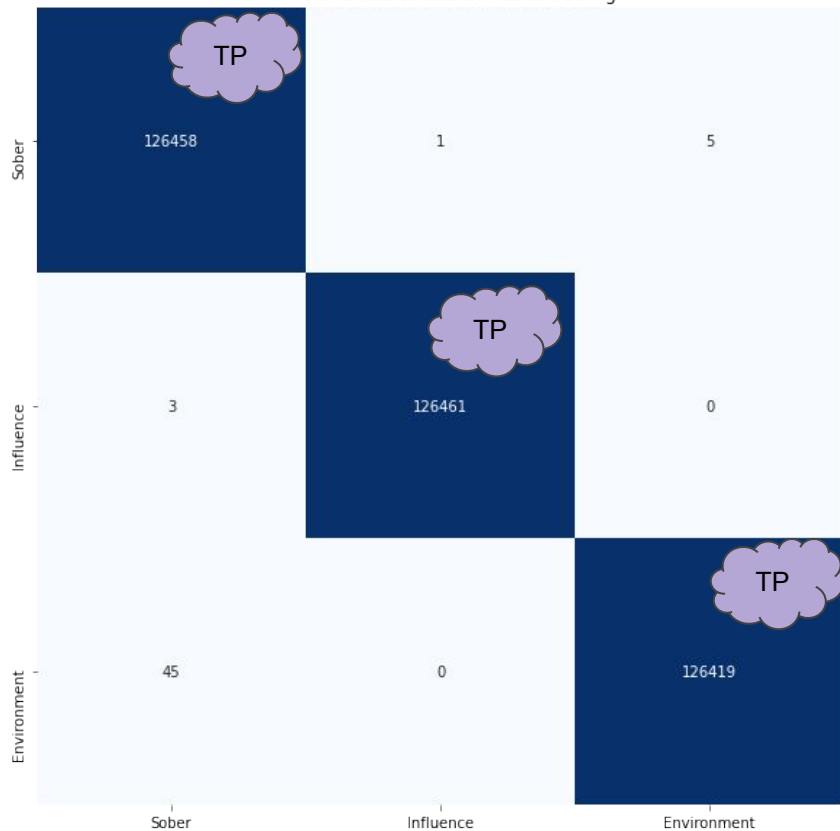
- Had been drinking when arrest not made
- Under the influence of drugs/alcohol during arrest

3. Environment

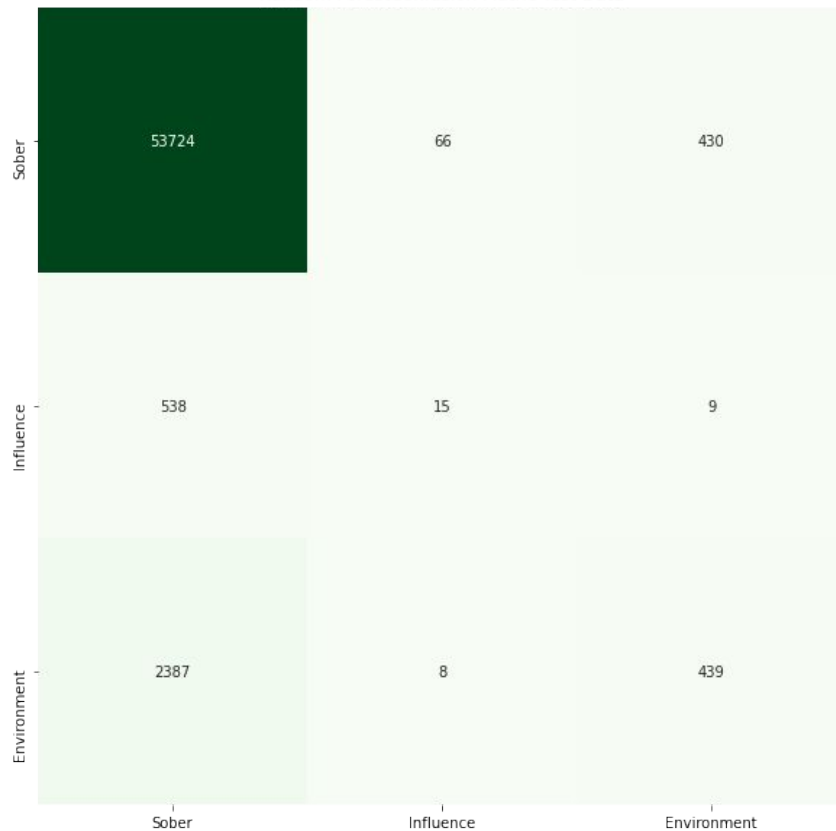
- Animal
- Weather
- Vision obscured
- Road construction/maintenance
- Distraction outside vehicle
- ETC.

Model Confusion Matrices

Random Forest Confusion Matrix Training

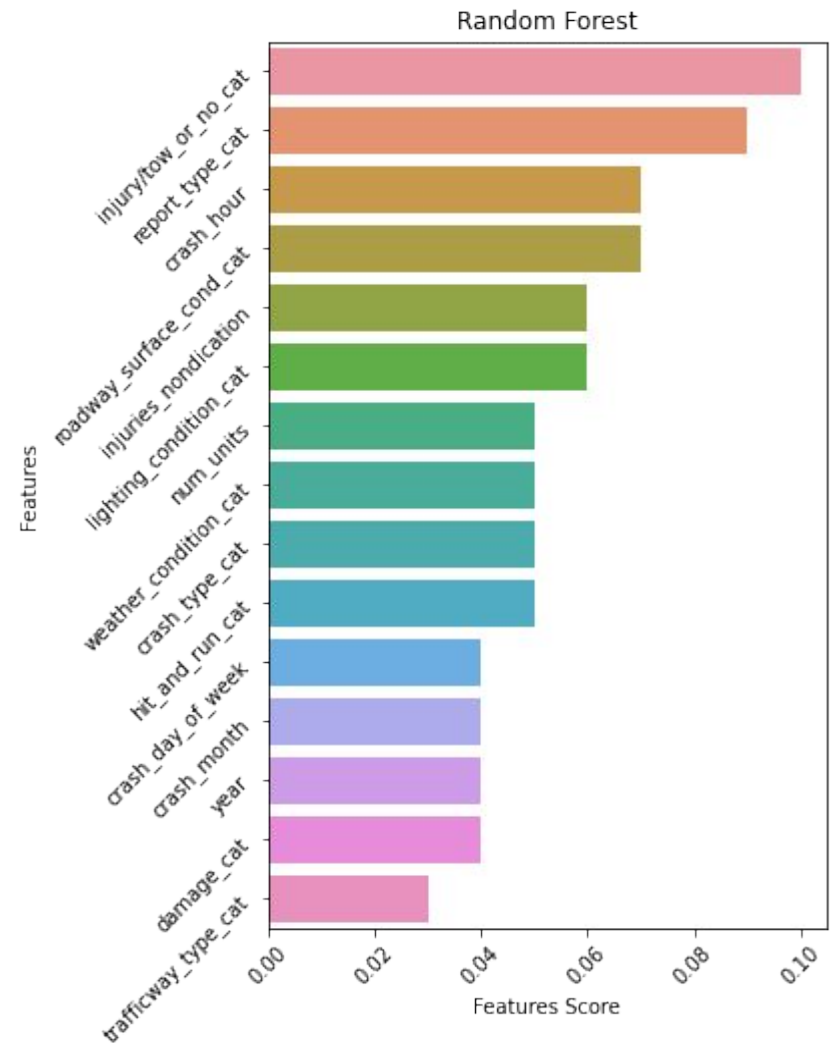


Random Forest Confusion Matrix Test Data



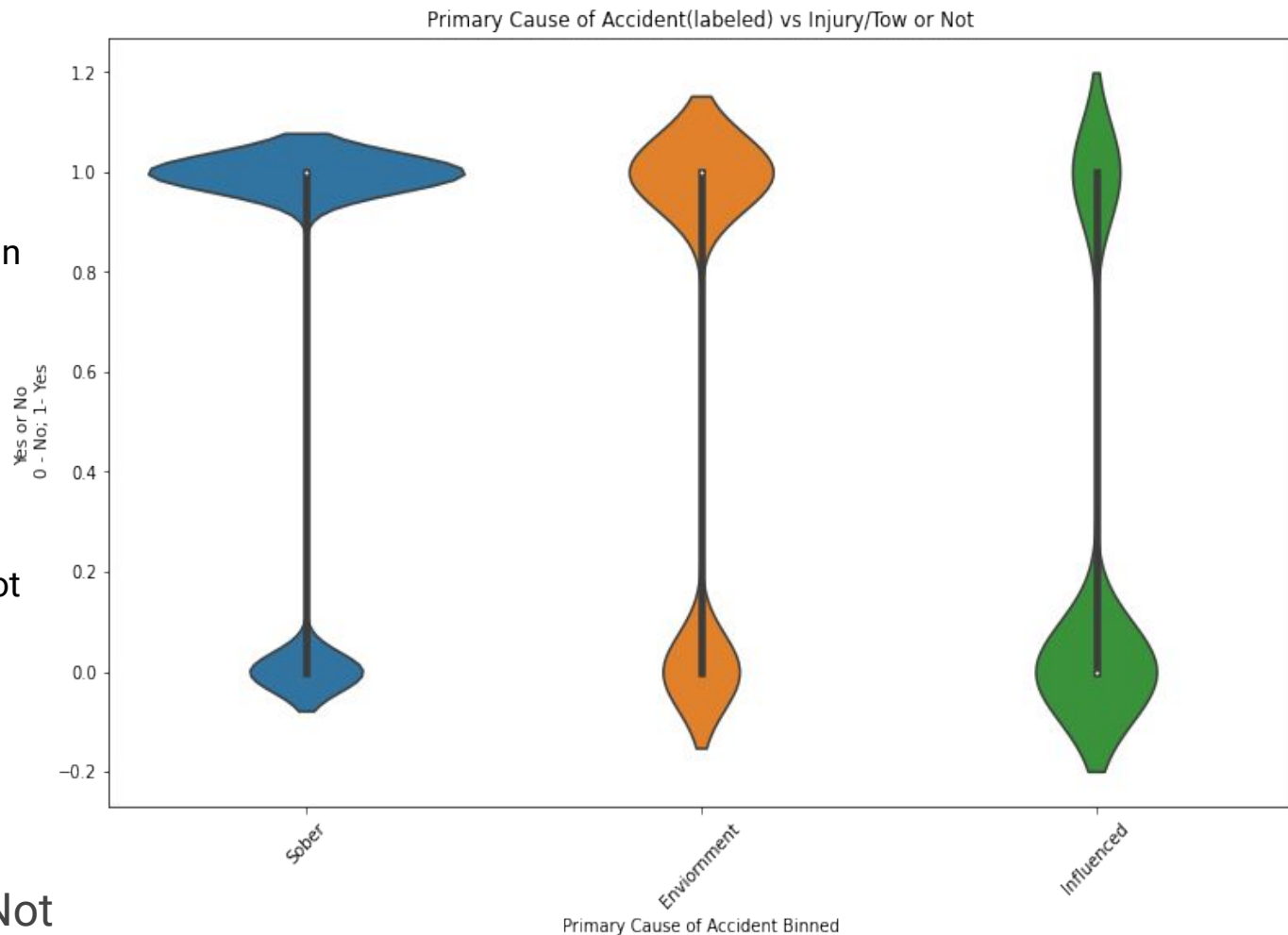
Best Features in Model

- Injury/tow or not
- Report type
- Crash hour
- Roadway surface condition
- Weather condition
- Etc.



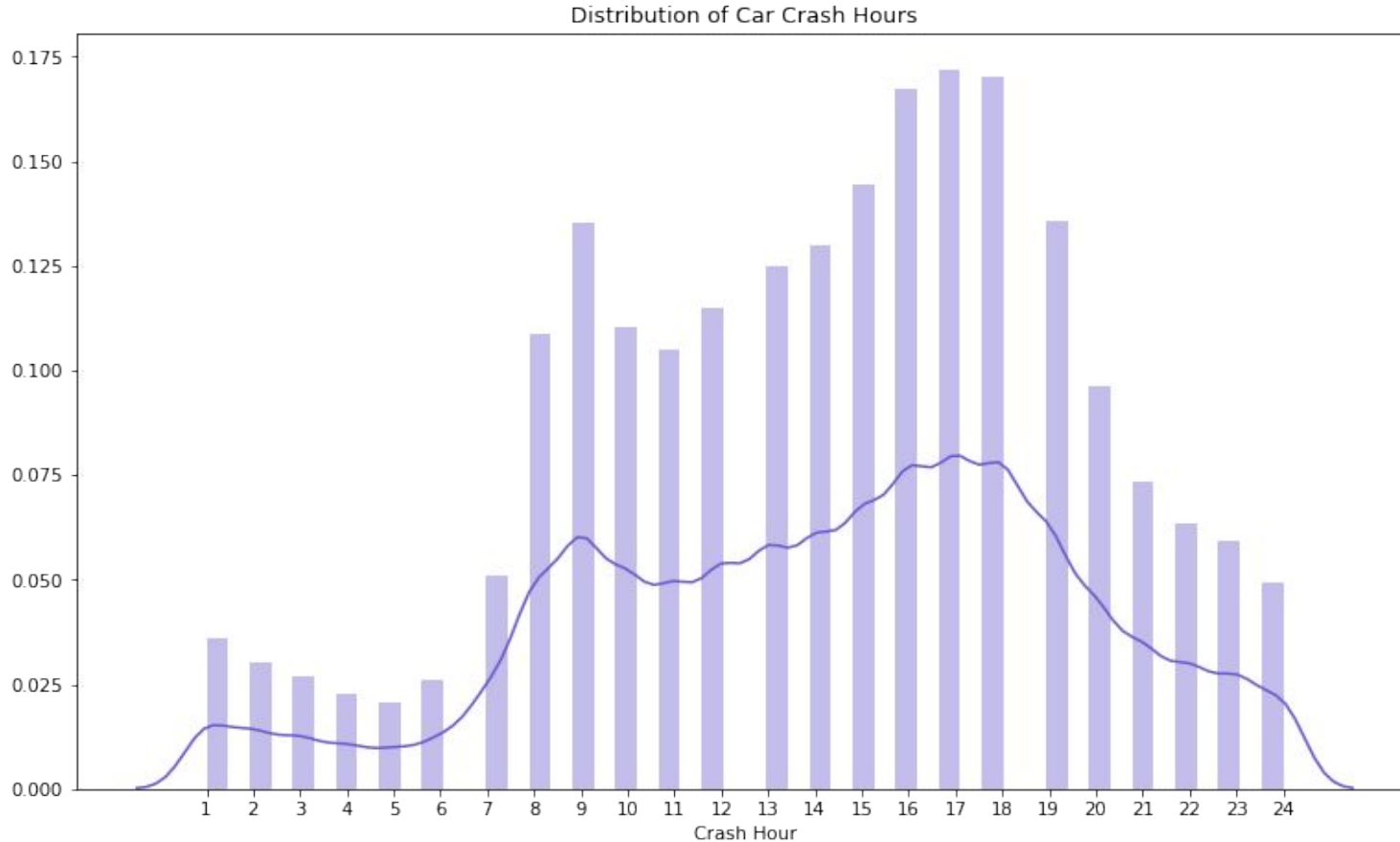
Insights:

- Sober and environment are more likely to sustain an injury or need car to be towed
- Influenced class shows higher distribution as not needing a tow or not injured



Feature: Injury/tow or Not

Feature: Crash Hour



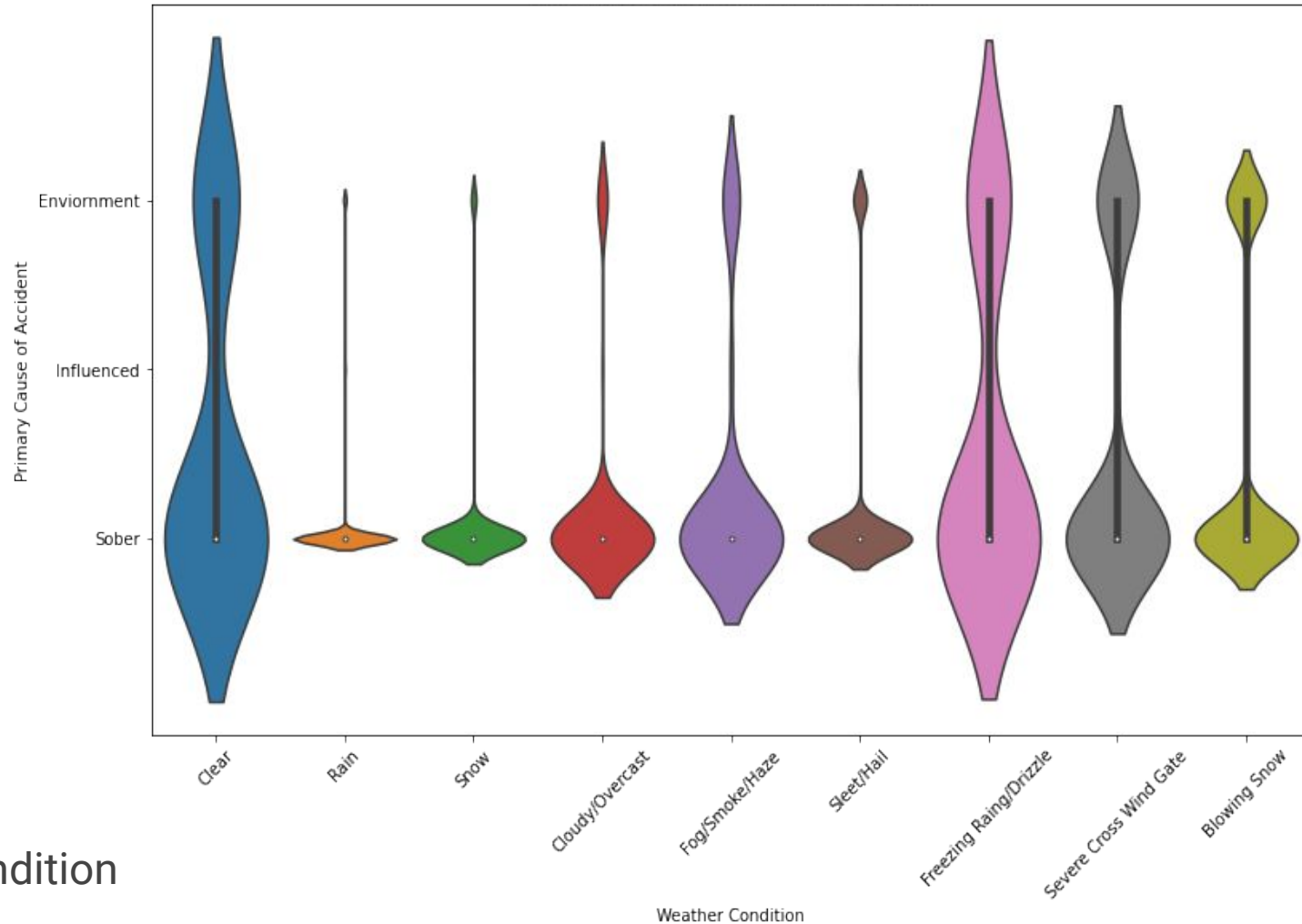
Shown here there is an increase of car crashes between 4 and 6pm;

Which means that most accidents that have happened are around rush hour

Weather Conditions and Cause of Accident

This graph shows the distribution of the different weather conditions and their relationship with the targets

Most accident for all targets occur during clear or raining weather conditions



Feature: Weather condition

Future Work



- Gather more data (with more classes needed) to train models
- Create additional models for the current outcomes to be able to predict exact primary cause of accident

The background is a solid pink color. In the top right corner, there is a decorative pattern of overlapping geometric shapes, including triangles and squares, in various shades of pink and magenta.

Thank You