



# **Analysis of Car Accidents In Chicago**

**By Ari Iwunze**



# Purpose

- ❖ **Primary types of car accidents**
  - What are the main causes of accidents so we can focus on targeted solutions.
- ❖ **Regions with the most car accidents**
  - What areas have the most accidents so more safety and law enforcing officers can be deployed to such areas.
- ❖ **External factors effect on the amount of car crashes and car crashes with injuries**
  - The time of the days effect on car accidents.
  - The weather's effect on car accidents.



# Classification Model

## Important Features:

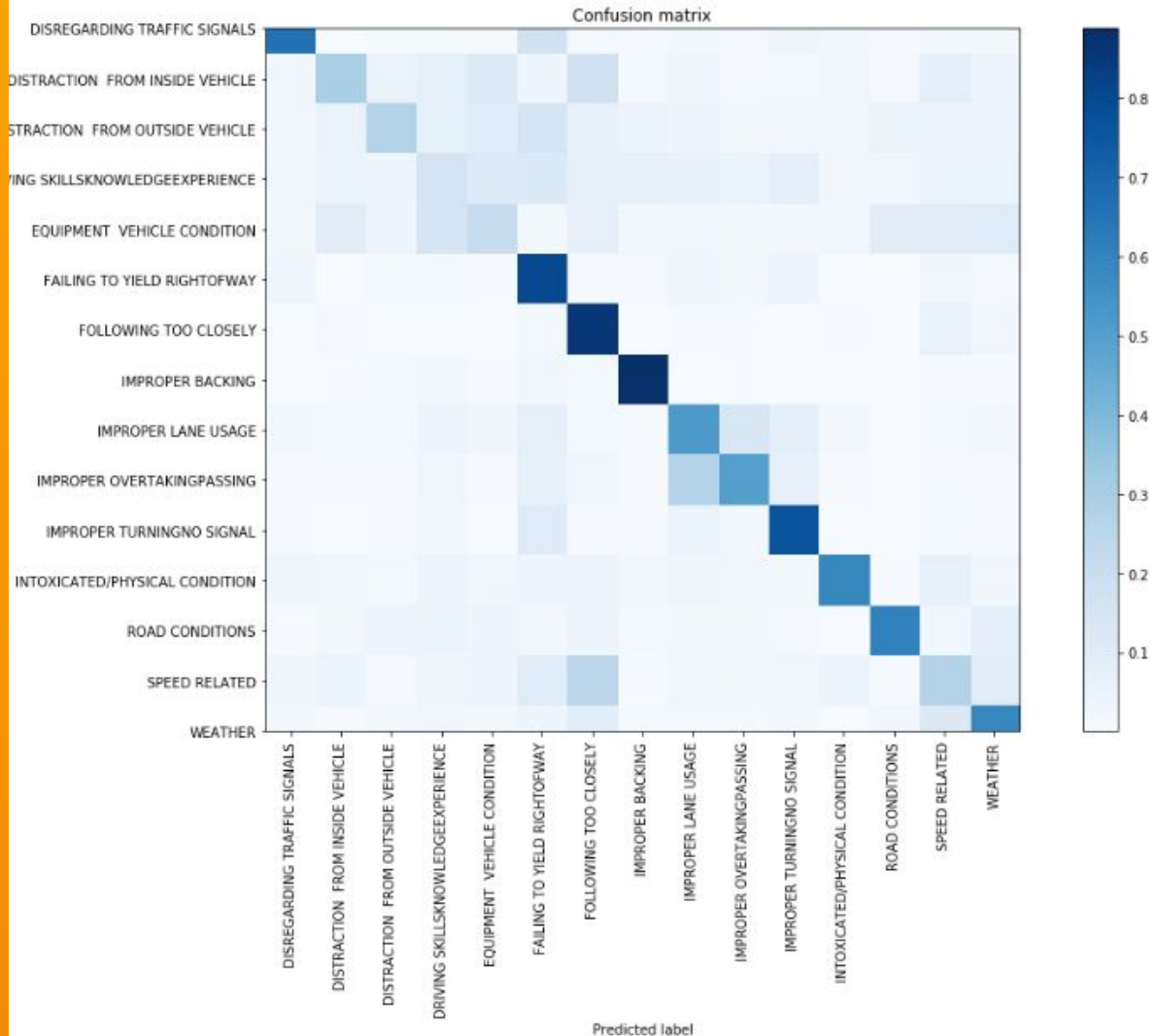
- Driver's Action
- Driver's Vision
- Roadway Surface Condition
- Device Condition
- First Crash Type
- Posted Speed Limit
- Age
- Physical Condition

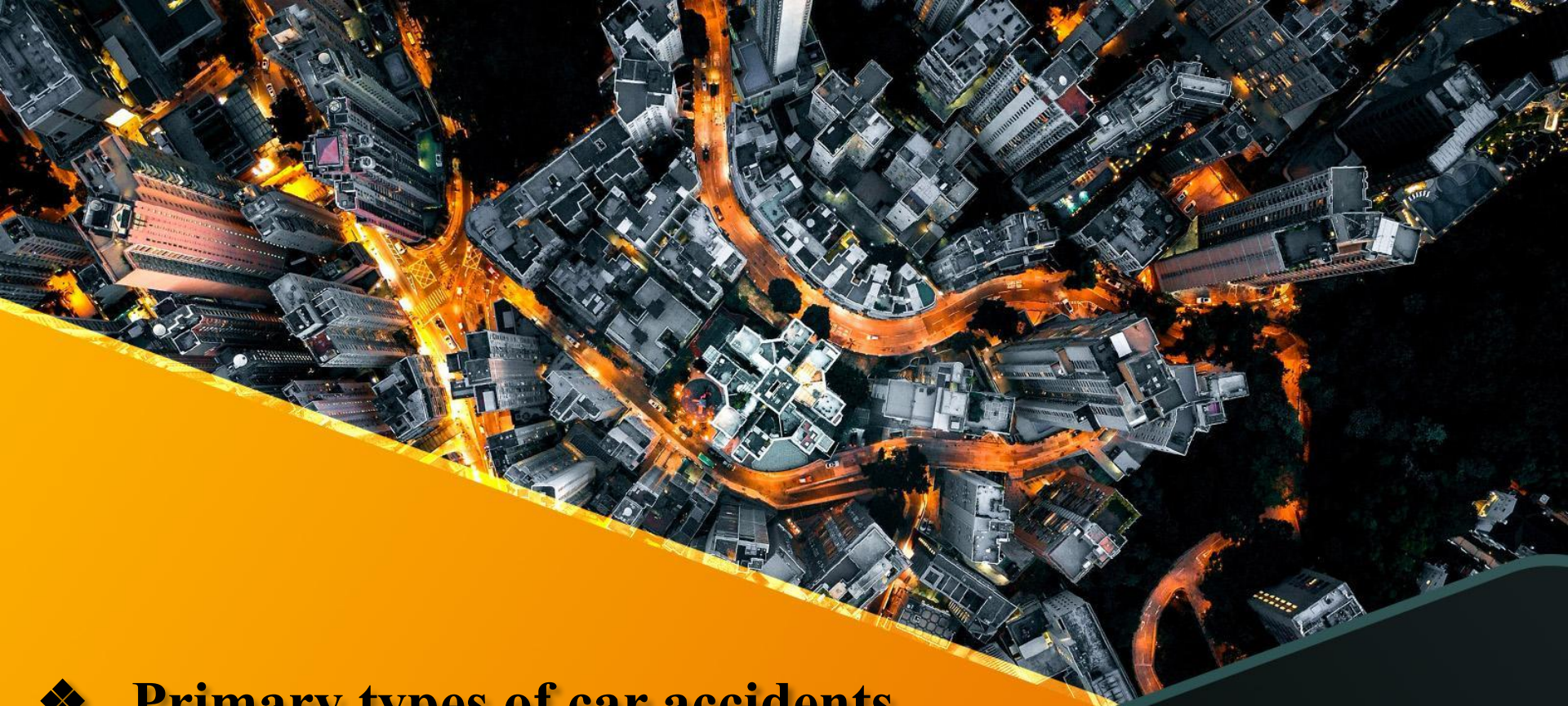
**Accuracy : 64%**

**Average AUC: 0.8**

**Log Loss: 12.5**

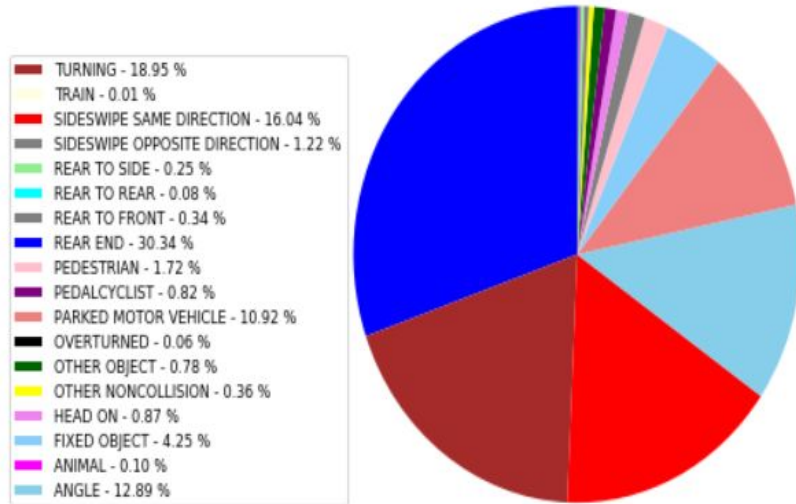
- See Feature importance plot in the Appendix





## ❖ Primary types of car accidents

# Proportion of each car crash type



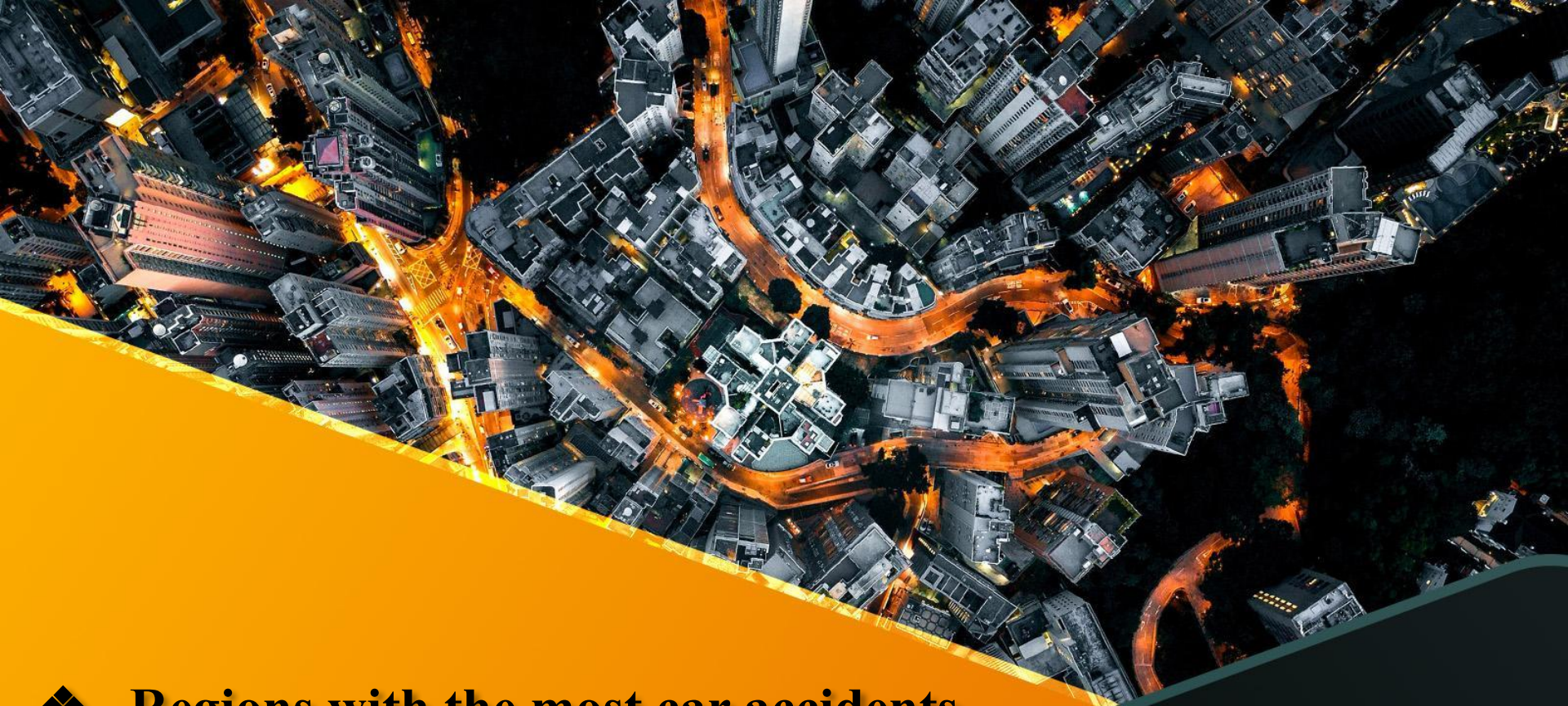
Proportion of crashes by each Crash Type

The most types of car accidents are Rear Ends accidents accounting for 30% of car crashes.

Followed by Sideswipe Same Direction accidents accounting for 17% of car crashes.

The most deadly types of crashes leading in proportions are Turning at 17.26 %. I'll recommend focusing on is accounts for the most fatalities.



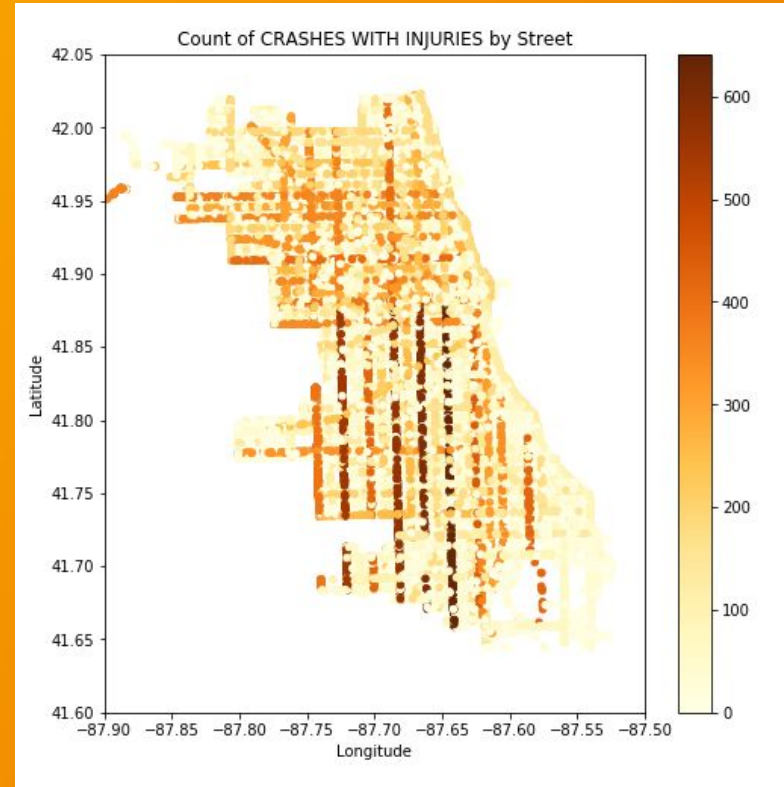


## ❖ Regions with the most car accidents

# Regions with the most car accidents

From the geo-map showing the proportion of injuries caused by car accidents on each street by their longitude and latitude positions.

As we can see from the map, more accidents with injuries occur about south-east from downtown of chicago(South-Central), more especially on the major streets.







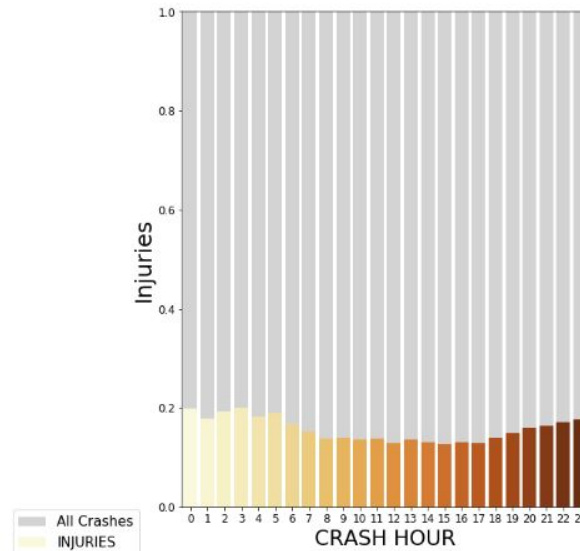
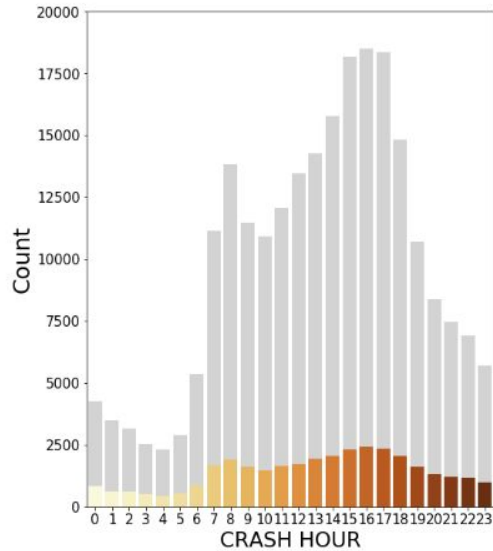
- ❖ **External factors effect on the amount of car crashes and car crashes with injuries**



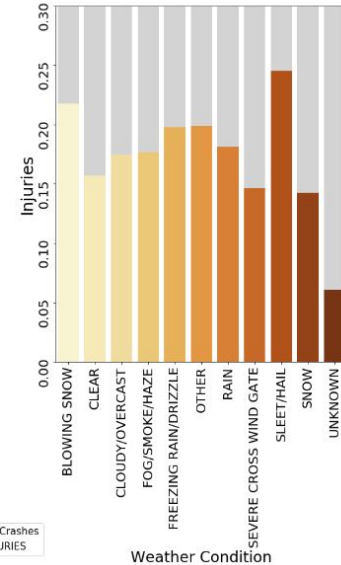
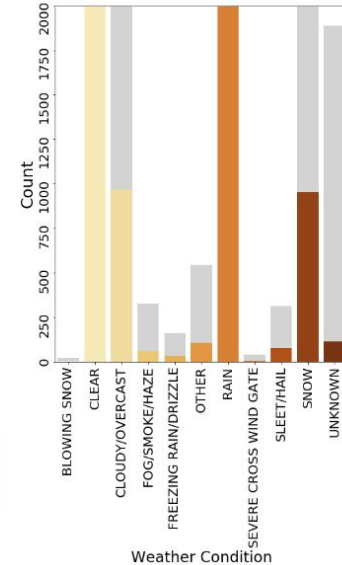
- The time of the days effect on car accidents.
- The weather's effect on car accidents.

From the histogram we can see most accidents occur in clear weather but more accidents with injuries occurs when there's hail or when it snows.

Accident counts and Injuries by CRASH HOUR



Accident counts and Injuries by Weather Condition





## Recommendations

➤ **Roadway surface condition**

More EMT and street sanitation employees should be deployed when there's expected bad weather.

➤ **Posted Speed Limit**

More traffic control measures should be put up on roads with high speed limits such as speed traps

➤ **Device Condition**

It should be made compulsory to go for routine maintenance and check ups.

➤ **Physical Condition**

The city needs to put in place proper laws that will give an establishment the more power to restrict intoxicated customers from driving.



# Conclusion

**Deadliest type of car crash**

**Turning: 17.26%**

17.26 % of crashes leading to fatalities involve some sort of turning so we need better traffic signals to ensure people make better turns.

**Region with the most accidents**

There's a high concentration of accidents with injuries occurring in the the South-Central part of the city.

**Time**

There's a spike in the amount of accidents between 4pm and 8pm.



# Future Work

## ➤ Road Type Division

Segregate the different types of streets/roads to understand the unique properties of accidents that occurs in each

## ➤ More Data

Gather more data regarding:

- ❖ If a driver was on the phone.
- ❖ If a driver exceeded the posted speed limit.
- ❖ If a driver has a good amount of driving experience.

## ➤ Region Division

Deeper analysis on the primary causes of accidents in the North, South, East West and Central regions of the city



**THANK YOU**

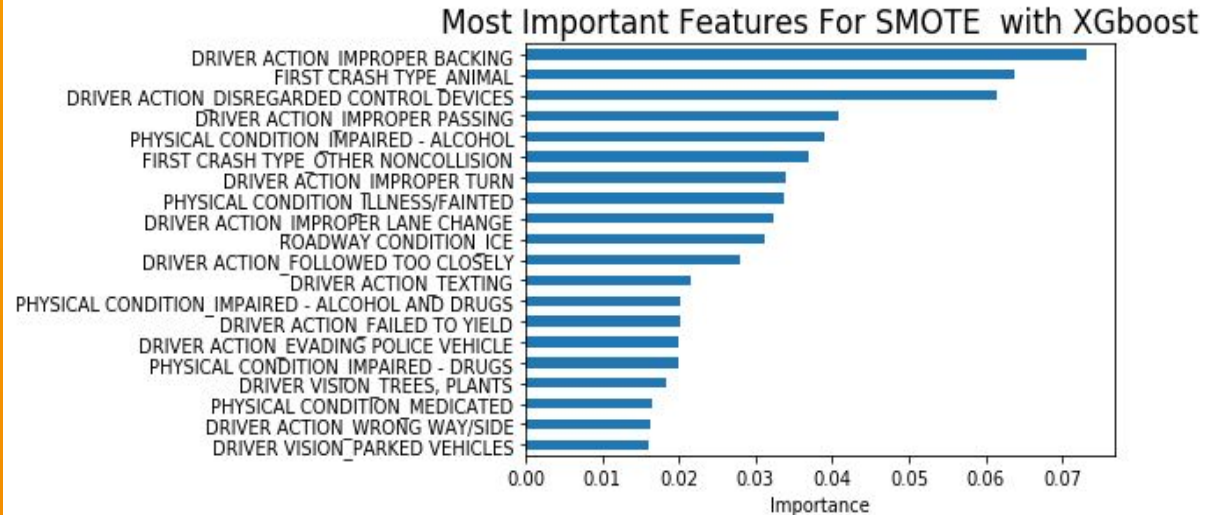
# Appendix

## Features Importance

- The AUC score is **0.8** and this is the average probability that the model can classify each class correctly. This will be the main metric that we'll focus on as we don't want to incorrectly classify the cause of an accident and risk wrongfully blaming a driver for an accident.
- From the models feature importance plot below, we can see the most influential features are:
  - **Driver's Action - Improper Backing.**
  - **First Crash Type - Animal.**
  - **Driver's Action - Disregarded Control Devices.**

- From the confusion matrix we saw prior, the Model has a some difficulty distinguishing accidents caused by the:

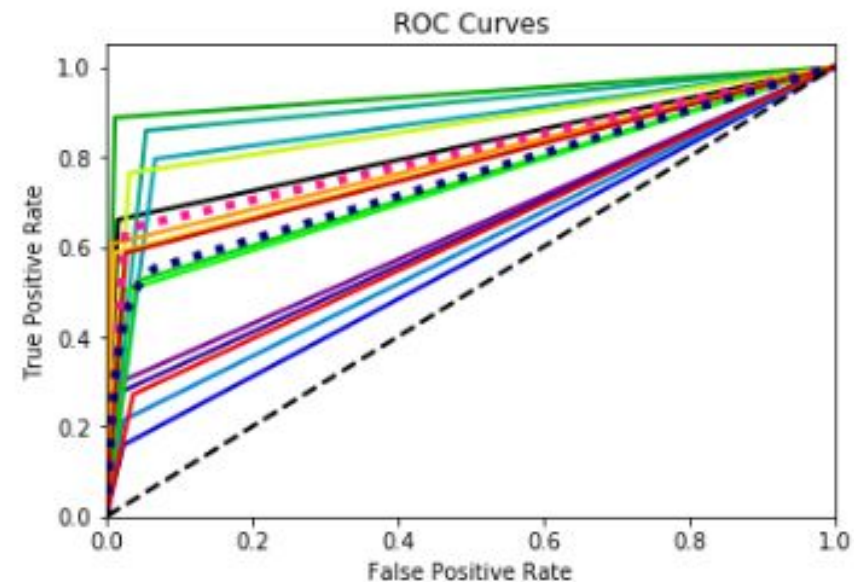
- **Vehicle Cond.**
- **Years of Experience.**
- **Outside Distraction.**
- **Inside Distraction.**







Log Loss: 12.455094861362022  
AUC: 0.8068150367903191  
Precision Score: 0.6393880686752622  
Recall Score: 0.6393880686752622  
Accuracy Score: 0.6393880686752622  
F1 Score: 0.6393880686752622



- ROC curve of class DISREGARDING TRAFFIC SIGNALS (area = 0.82)
- ROC curve of class DISTRACTION FROM INSIDE VEHICLE (area = 0.64)
- ROC curve of class DISTRACTION FROM OUTSIDE VEHICLE (area = 0.63)
- ROC curve of class DRIVING SKILLSKNOWLEDGEEXPERIENCE (area = 0.57)
- ROC curve of class EQUIPMENT VEHICLE CONDITION (area = 0.60)
- ROC curve of class FAILING TO YIELD RIGHTOFWAY (area = 0.86)
- ROC curve of class FOLLOWING TOO CLOSELY (area = 0.90)
- ROC curve of class IMPROPER BACKING (area = 0.94)
- ROC curve of class IMPROPER LANE USAGE (area = 0.74)
- ROC curve of class IMPROPER OVERTAKINGPASSING (area = 0.74)
- ROC curve of class IMPROPER TURNINGNO SIGNAL (area = 0.87)
- ROC curve of class INTOXICATED/PHYSICAL CONDITION (area = 0.79)
- ROC curve of class ROAD CONDITIONS (area = 0.80)
- ROC curve of class SPEED RELATED (area = 0.62)
- ROC curve of class WEATHER (area = 0.78)
- micro-average ROC curve (area = 0.81)
- macro-average ROC curve (area = 0.75)