



PREDICTION THE SEVERITY OF AN ACCIDENT (CASE STUDY)

Konda Lavanya

Introduction

- Every year the lives of approximately 1.35 million people are cut short as a result of a road traffic crash. Between 20 and 50 million more people suffer non-fatal injuries, with many incurring a disability as a result of their injury.
- Road traffic injuries cause considerable economic losses to individuals, their families, and to nations as a whole.
- It would be great if real-time conditions can be provided to estimate the trip safeness.
- In this way, it can be decided beforehand if the driver will take the risk, based on reliable information.

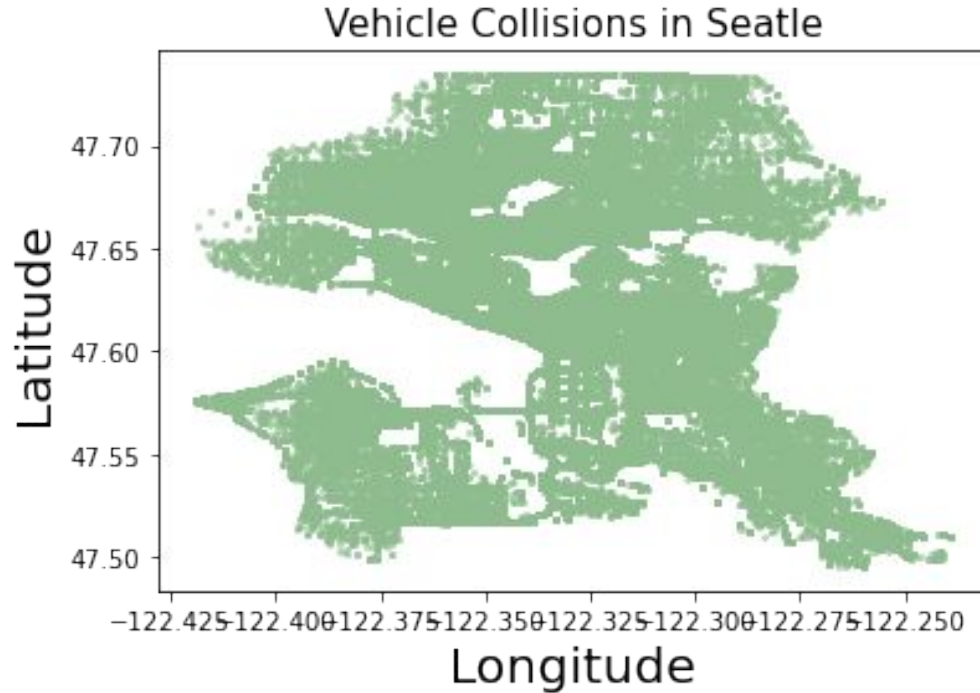
Business Understanding

- Predicting crash injury severity is a crucial constituent of reducing the consequences of traffic crashes.
- We can use the Machine learning models to Predict the severity of the Accidents based on the factors like the type of Road travelling, Location, number of Vehicles, Number of People on Road, weather any many more. These models Helps Road Users, Insurance Companies, Health Care providers, Government etc.
- This help user to have a better understanding of Road Conditions, their impacts, helps to take initiatives to reduce the Accidents and to build New Infrastructure.

Data Understanding

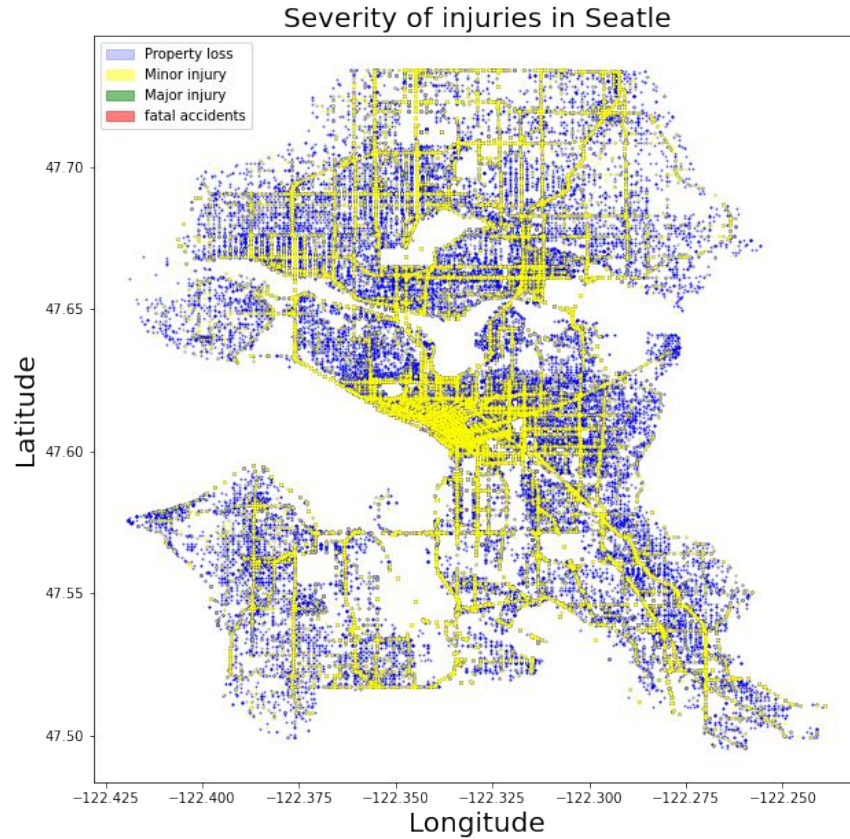
FEATURE	DESCRIPTION
X	Latitude of the Location of the incident
Y	Longitude of the Location of the incident
ADDRTYPE	Where the collision took place Block / Intersection
COLLISIONTYPE	Type of Collision - Right turn,Left turn,Cycles etc
SEVERITYCODE	How severe the injury is (our target)
JUNCTIONTYPE	Kind of junction type - Intersection,Mid-Block etc
UNDERINFL	Was driver drunk and driving
WEATHER	The weather condition
ROADCOND	Condition of the road - Wet,Dry,Ice,Standing water etc
LIGHTCOND	Light Conditions - Daylight,dusk,Street lights On/Off etc
SPEEDING	If driver was speeding or not

Lets us plot and try to understand the Vehicle Collisions in Seattle city.



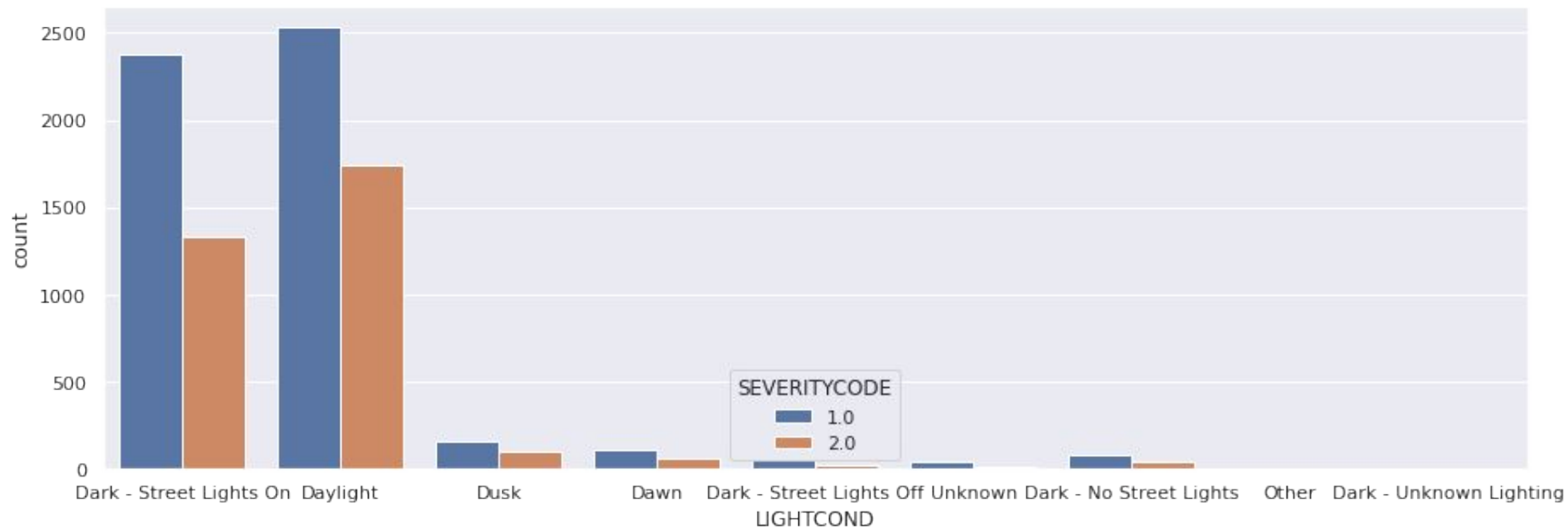
- The plot indicates that the accidents are taking places All over Seattle

Severity of the collisions on a Scatter plot

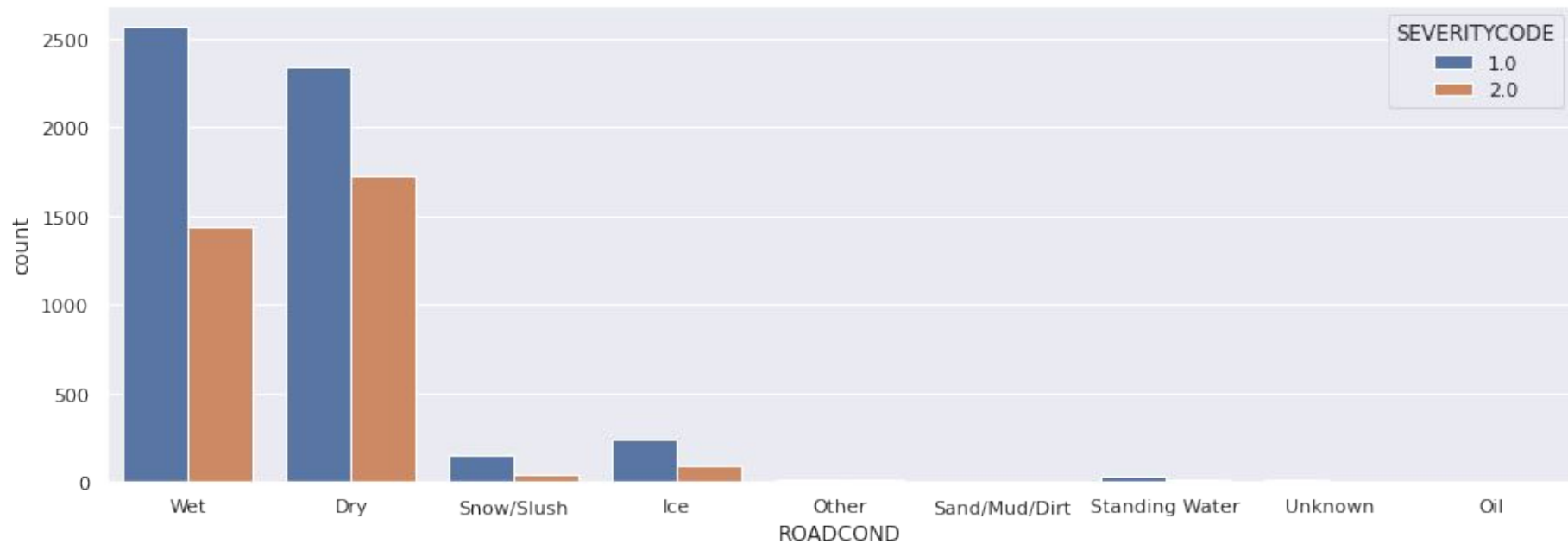


Data Exploration

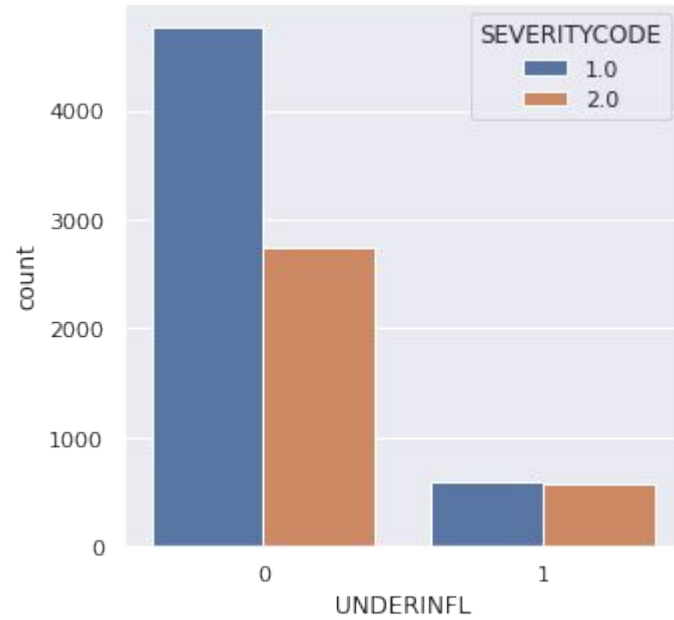
Light condition vs SEVERITYCODE.



Road condition vs SEVERITYCODE.



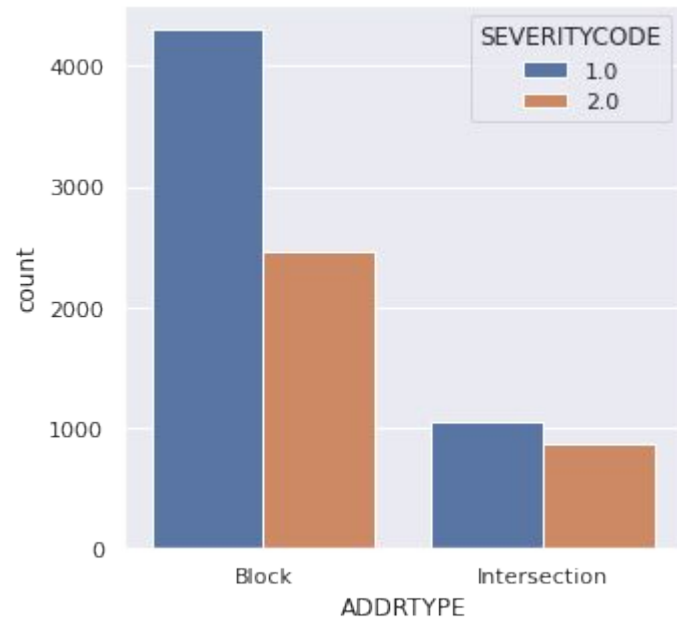
UNDERINFL vs SEVERITYCODE.



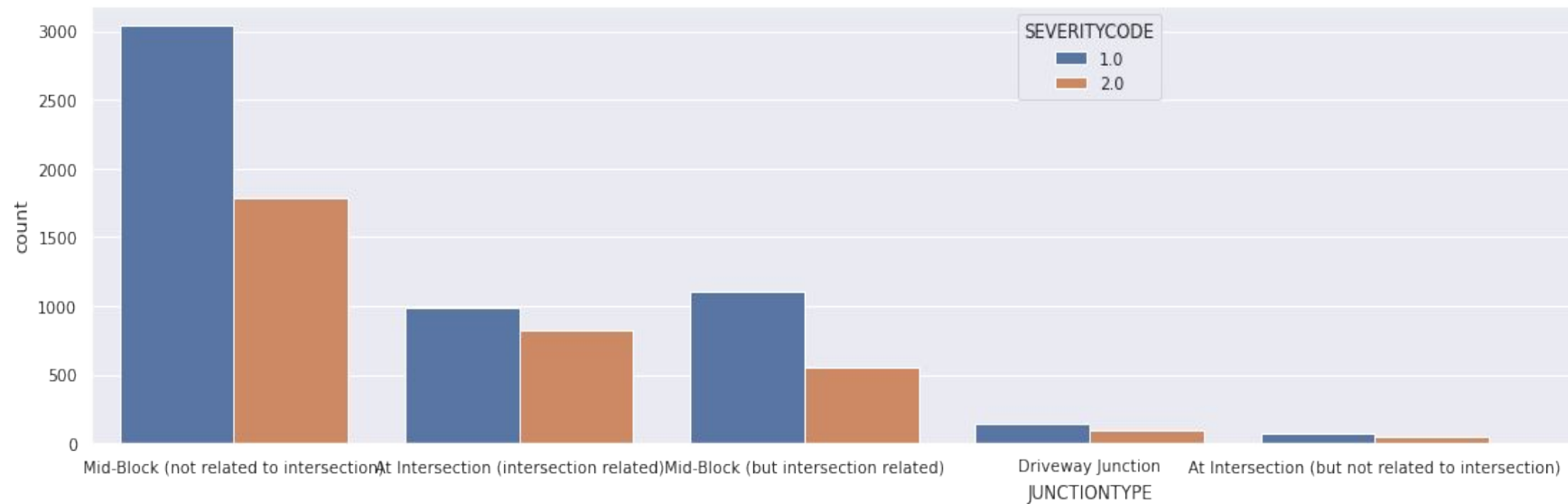
SPEEDING vs SEVERITYCODE



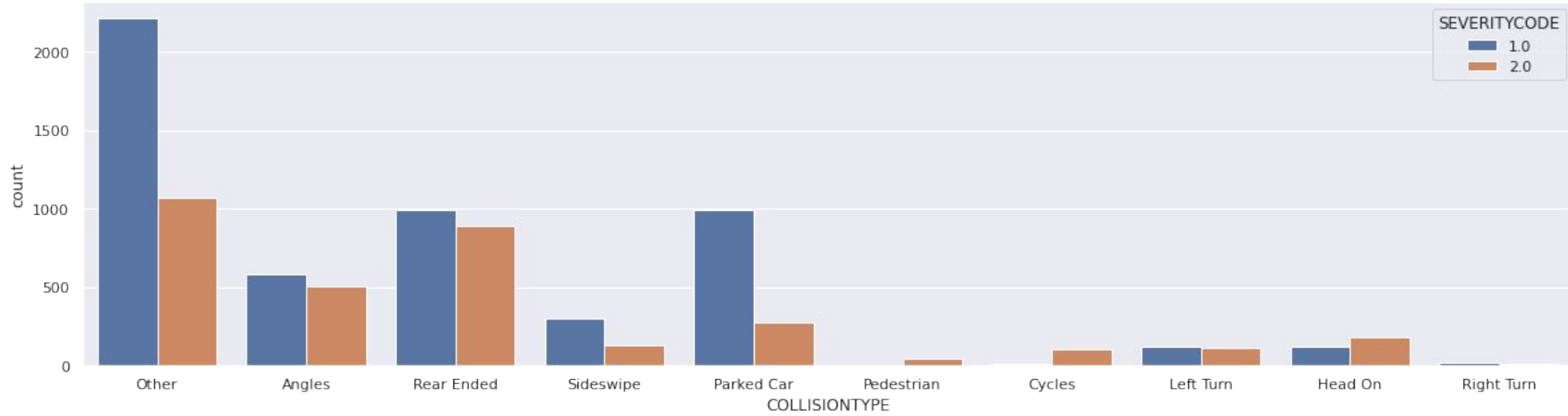
ADDRTYPE vs SEVERITYCODE



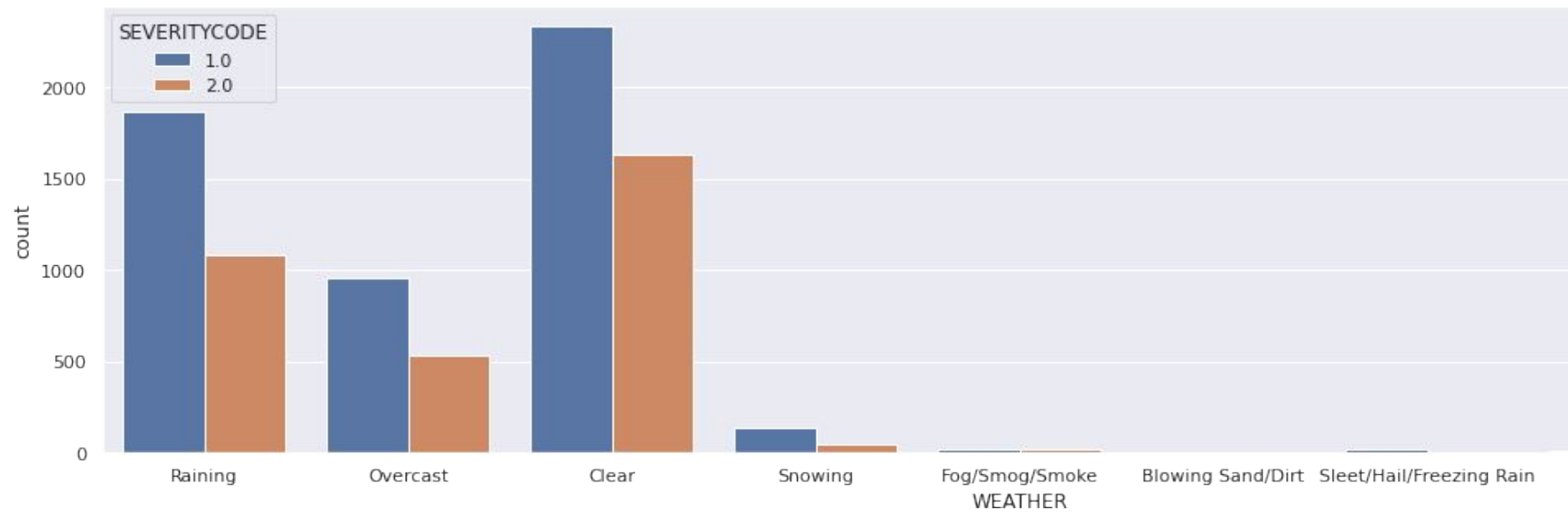
JUNCTIONTYPE vs SEVERITYCODE.



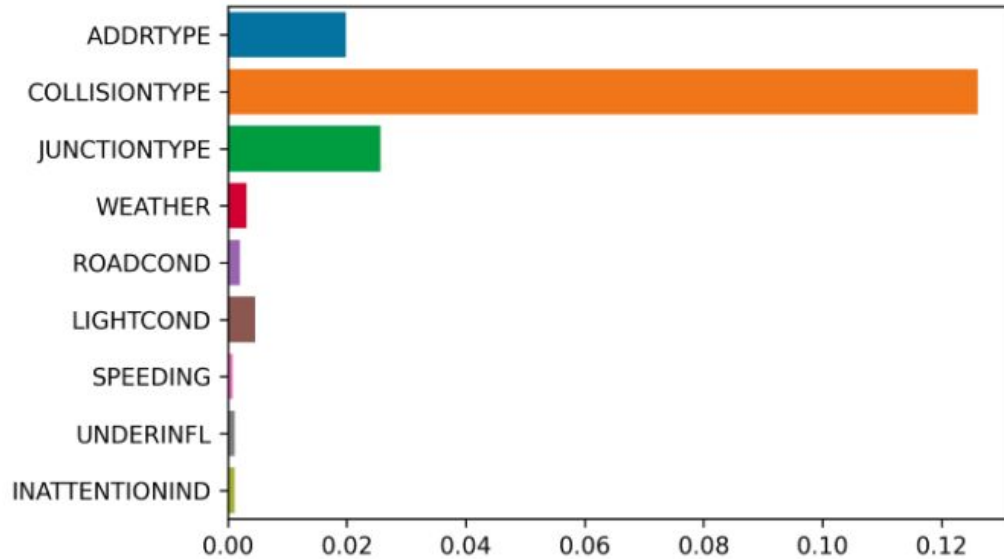
COLLISIONTYPE vs SEVERITYCODE.



WEATHER vs SEVERITYCODE.



Feature Selection:



After Mutual Information Feature Selection from the Mutual Information Importance Values I choose to work with
"ADDRTYPE","COLLISIONTYPE","JUNCTIONTYPE","UNDERINFL","WEATHER","ROADCOND","LIGHTCOND","SPEEDING","SEVERITYCODE"

Data Preparation:

- Dropping all the rows with values Nan / Unknown
- Changing the Values Y to 1 and N to 0
- Since the property Loss and Minor injuries are common So let's drop the rows with SEVERITYCODE - 0,2b,3.
- As they are Categorical Values we need to insert dummy values, i.e perform One-hot Encoding
- After checking the Number of rows with SEVERITYCODE 1,2. As we are required to make a balanced dataset we are dropping a few random rows with SEVERITYCODE 1

Modeling

Results of 5 Classifiers used are:

Classifier	Accuracy	Jaccard Score	f1 Score
KNN	0.55	0.43	0.60
SVM	0.62	0.43	0.60
Decision Tree	0.57	0.40	0.61
Logistic regression	0.62	0.45	0.62
Random Forest Classifier	0.61	0.43	0.61

Conclusions

- Collisions which does not involve personal injuries are twice as frequent as the ones involving damaged people.
- Accidents involving cycles or pedestrians are severe and involves injuries.
- The riskier car collisions are the ones that hit the car from the rear end.
- Intersection collisions are one of the most common types of crash
- Left turns are also risky maneuvers which should also be avoided if the road users want to be safe.
- Extremely dangerous weather and road conditions do not produce a quite significant accident rate, such as snow and ice. However ,caution have to be taken with rainy weather and wet roads, since after clear days and dry roads, these are the following conditions in order of importance.

Thank You!