Capstone Project

Coursera : Data Science

Introduction

Key facts from World Health Organization

- Approximately 1.35 million people die each year as a result of road traffic crashes.

- Road traffic crashes cost most countries 3% of their gross domestic product.

- More than half of all road traffic deaths are among vulnerable road users: pedestrians, cyclists, and motorcyclists.

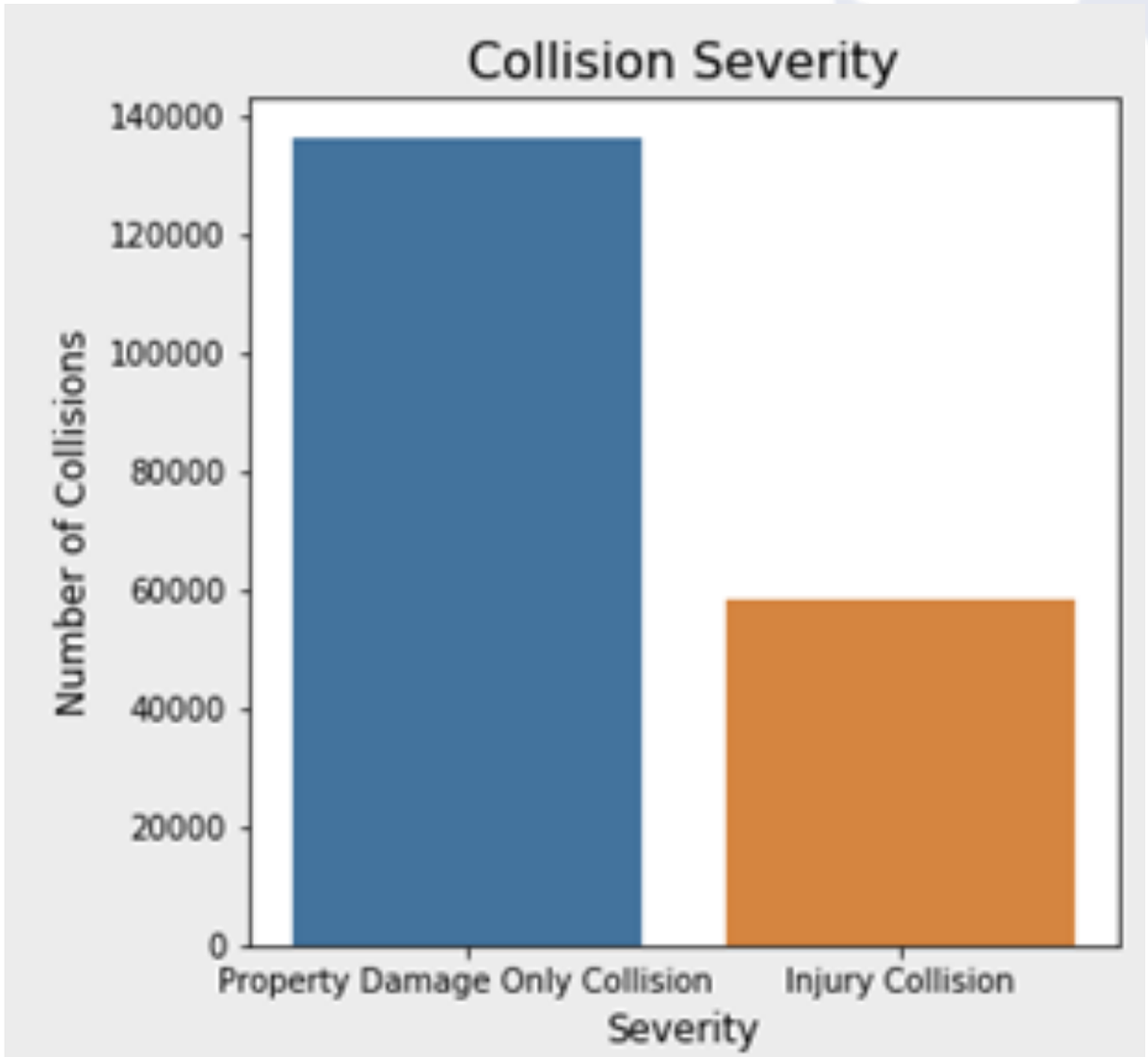
Introduction

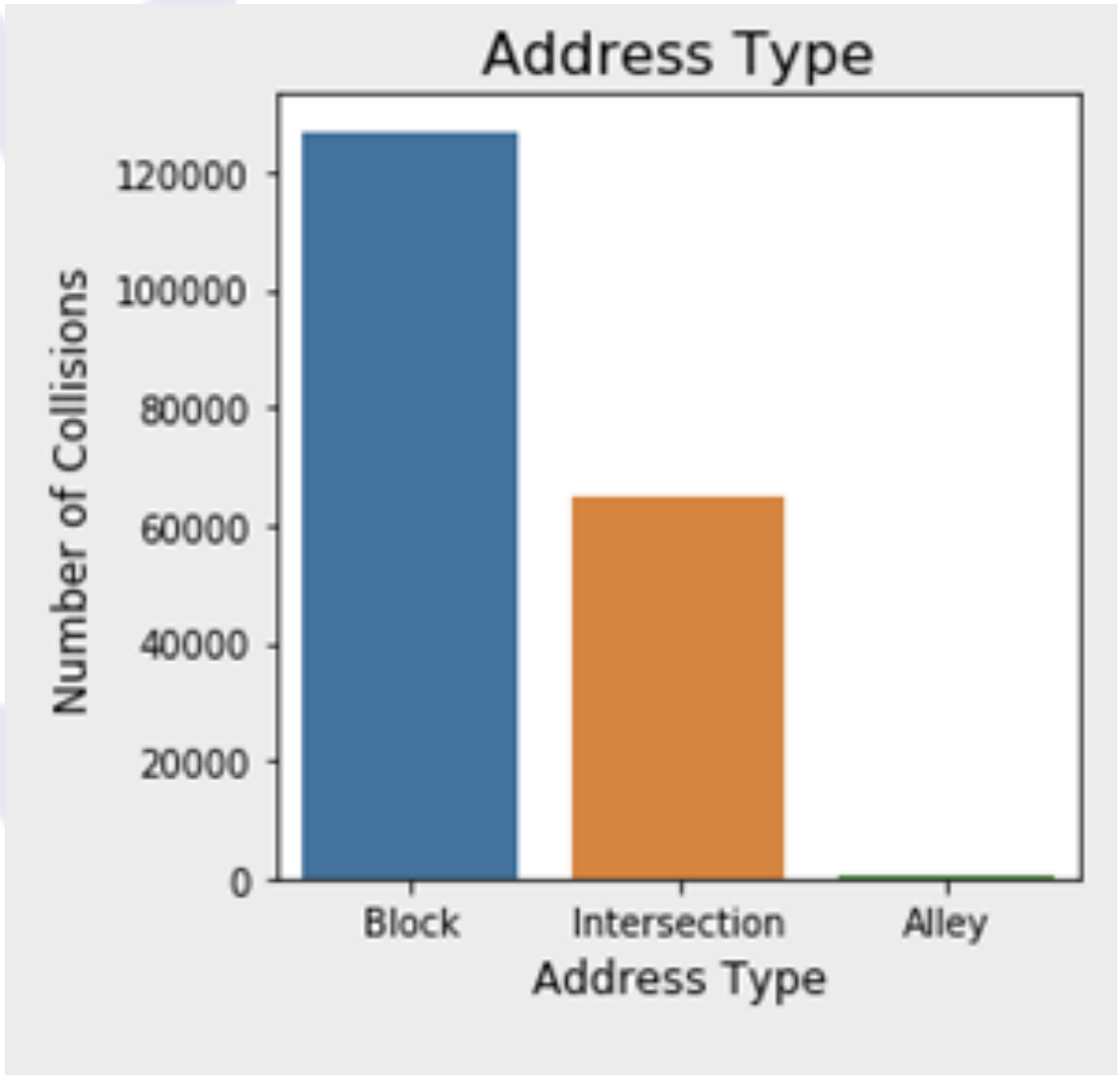
Effects and Audience of Collision

- Injured and dead
- State and Public Property
- Family and People Around
- Economy and Social Problem.

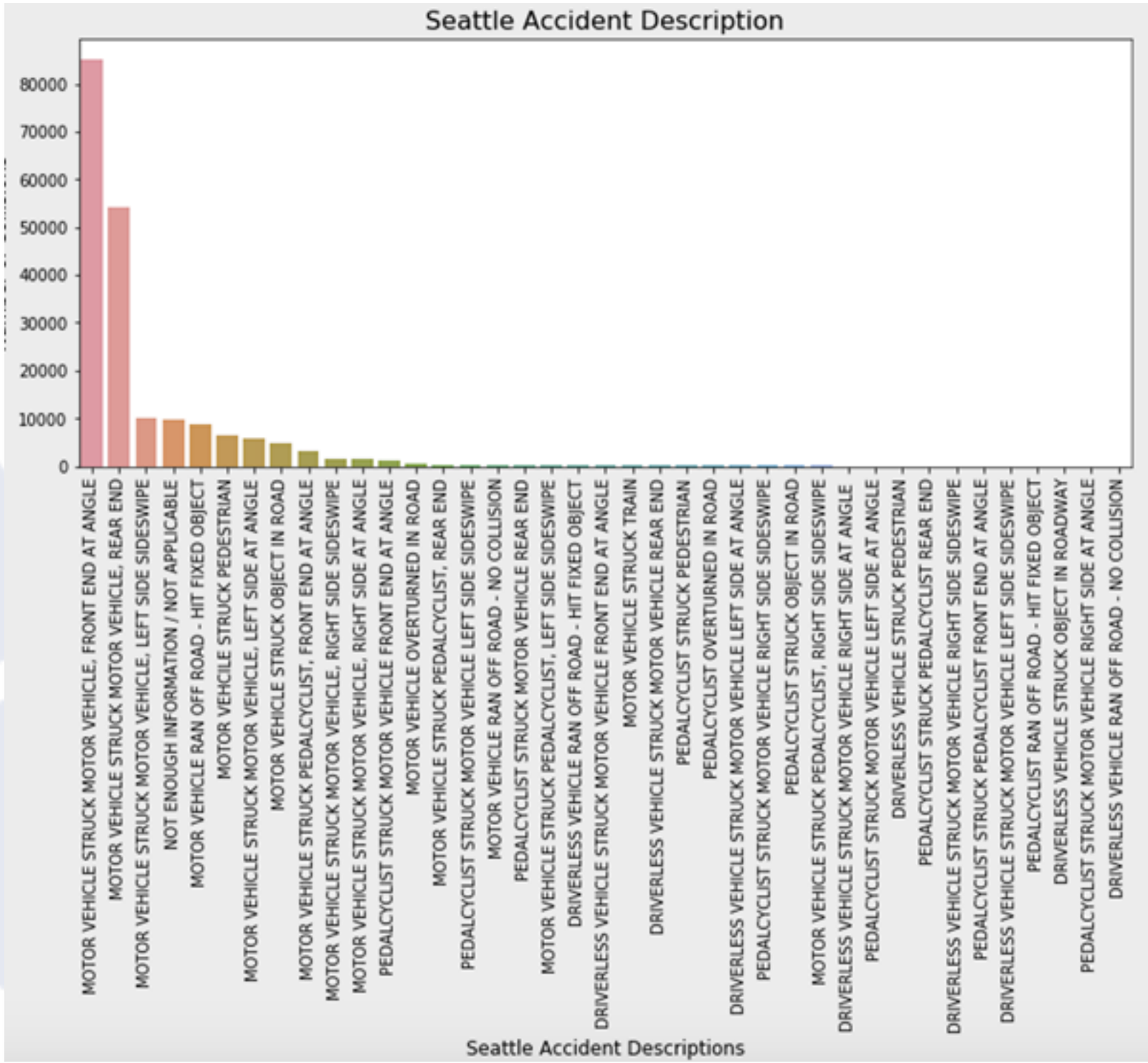
Data and Understanding

The data used is the Seattle's Department of Transportation and recorded by Traffic Records

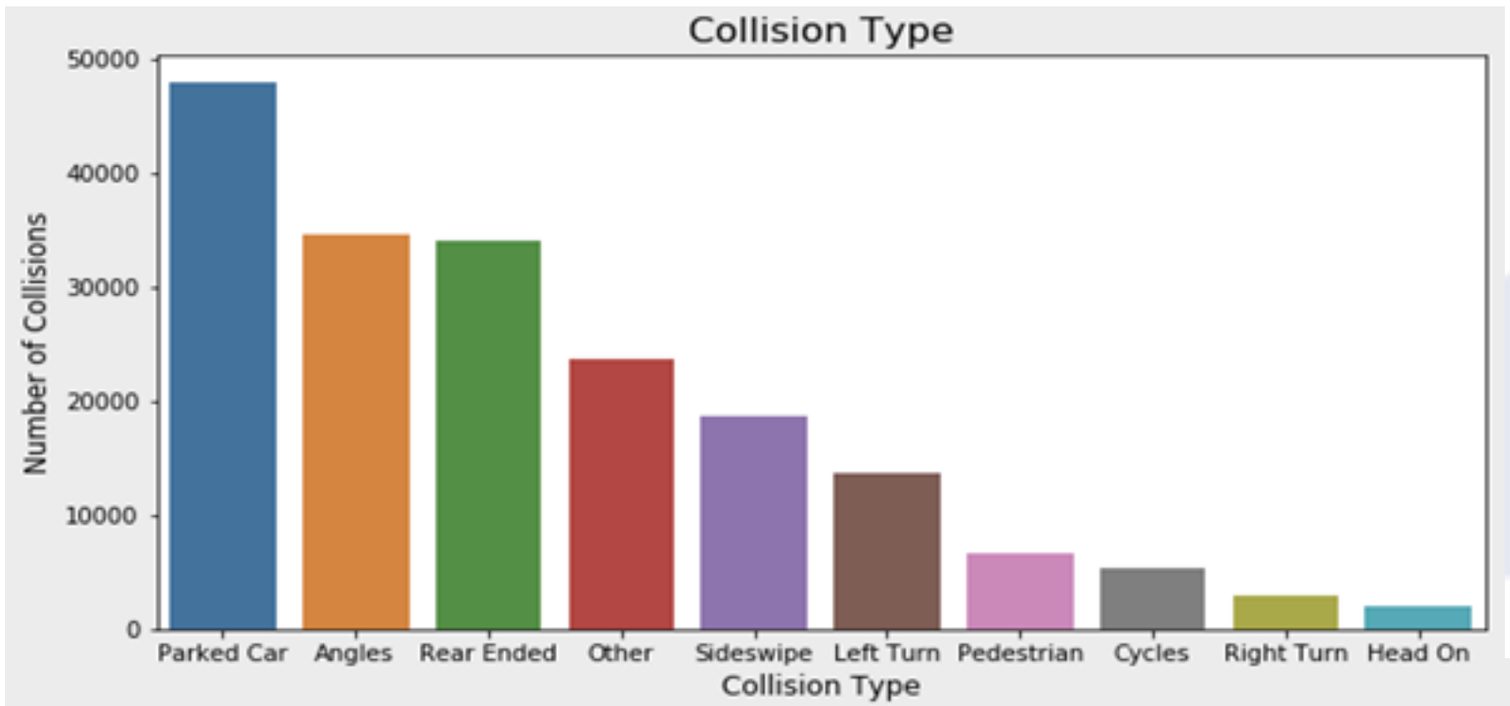


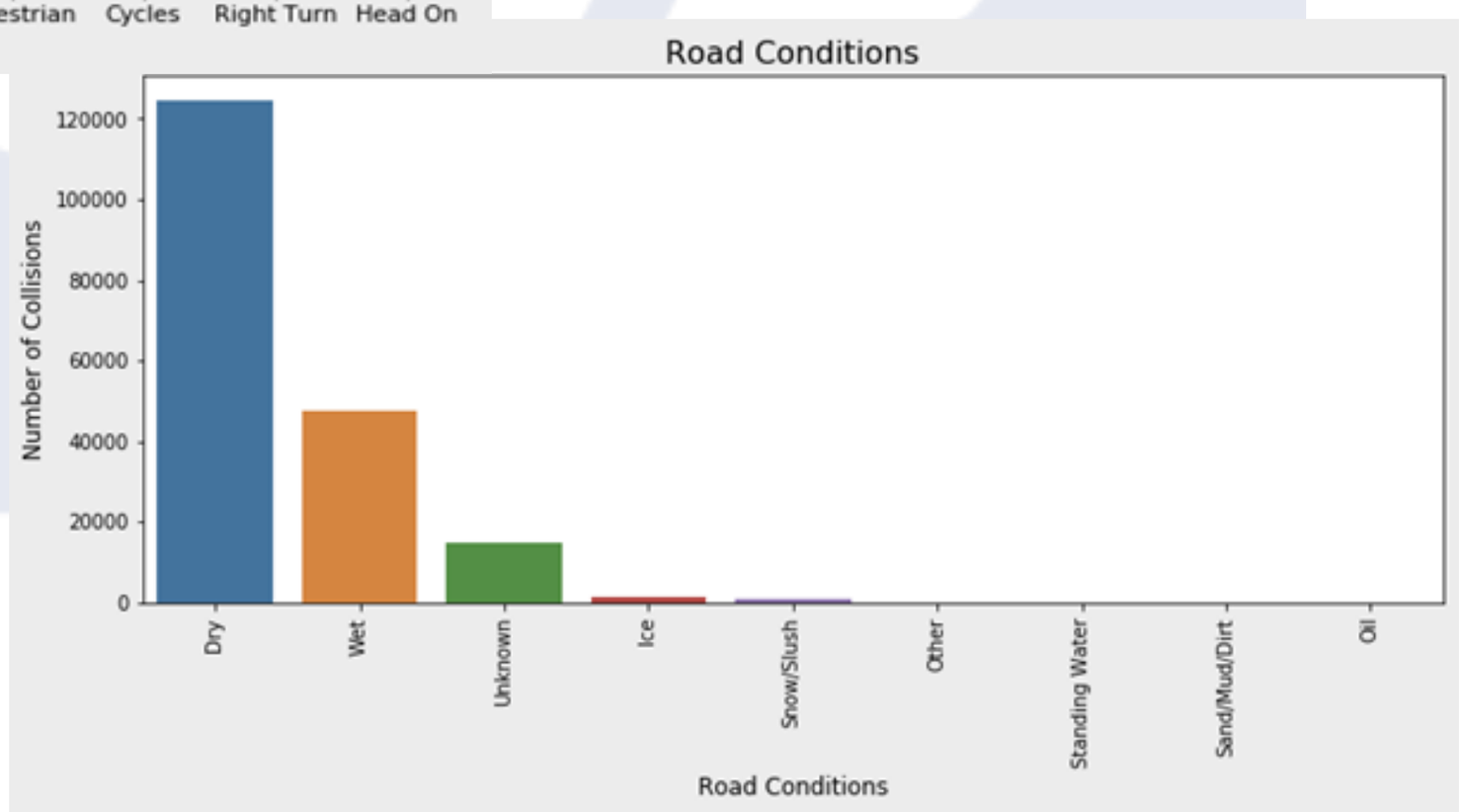


Data and Understanding

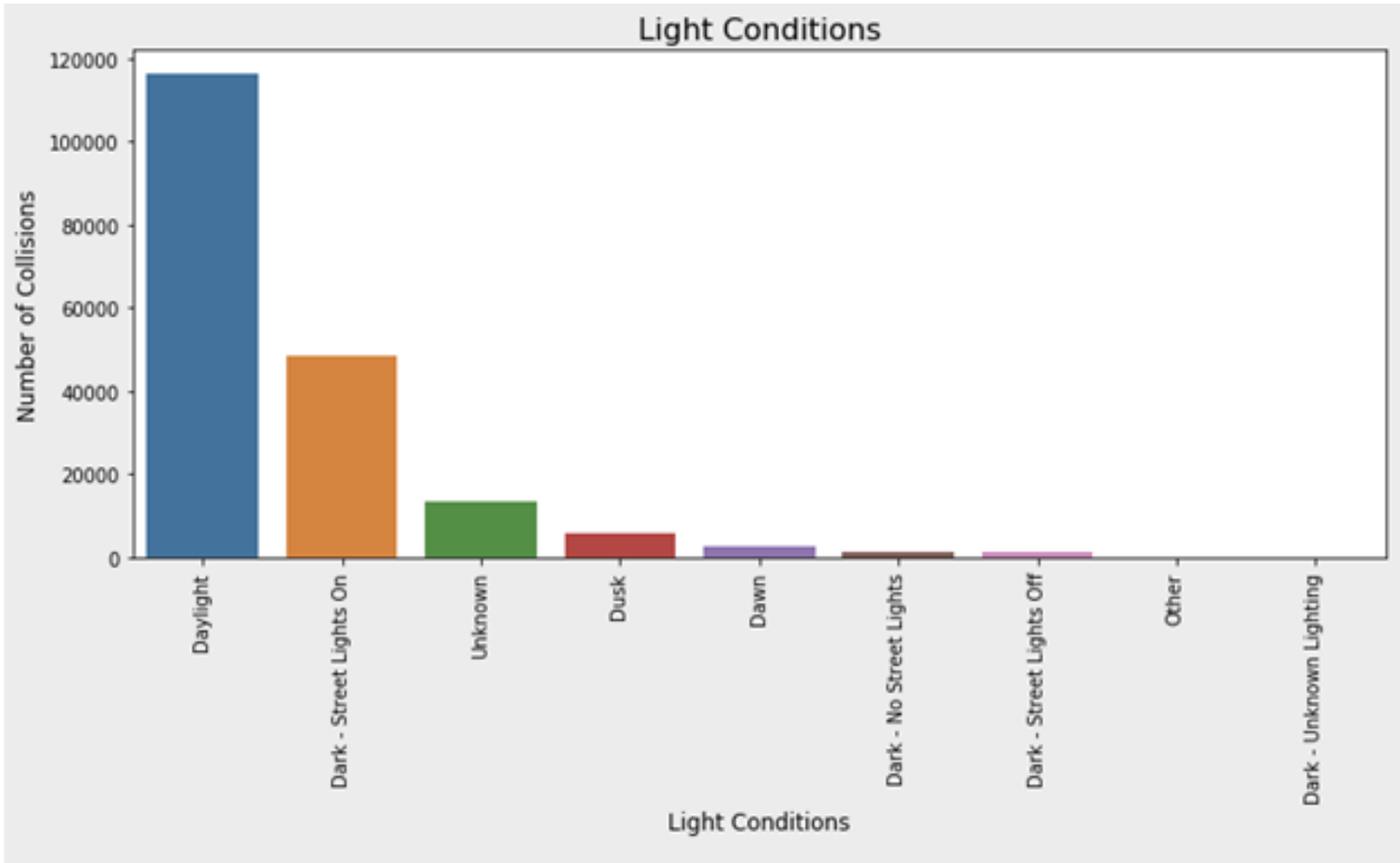


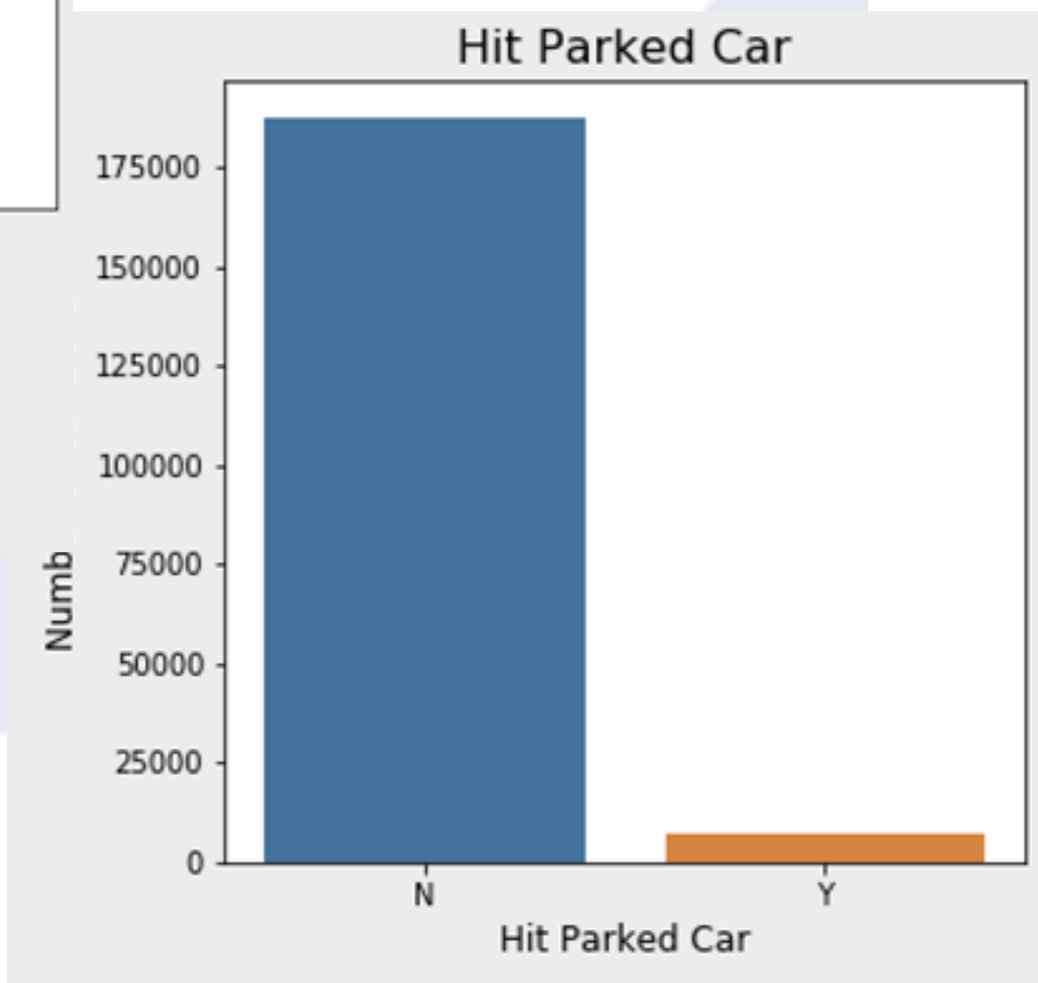
Data and Understanding





Data and Understanding





Methodology

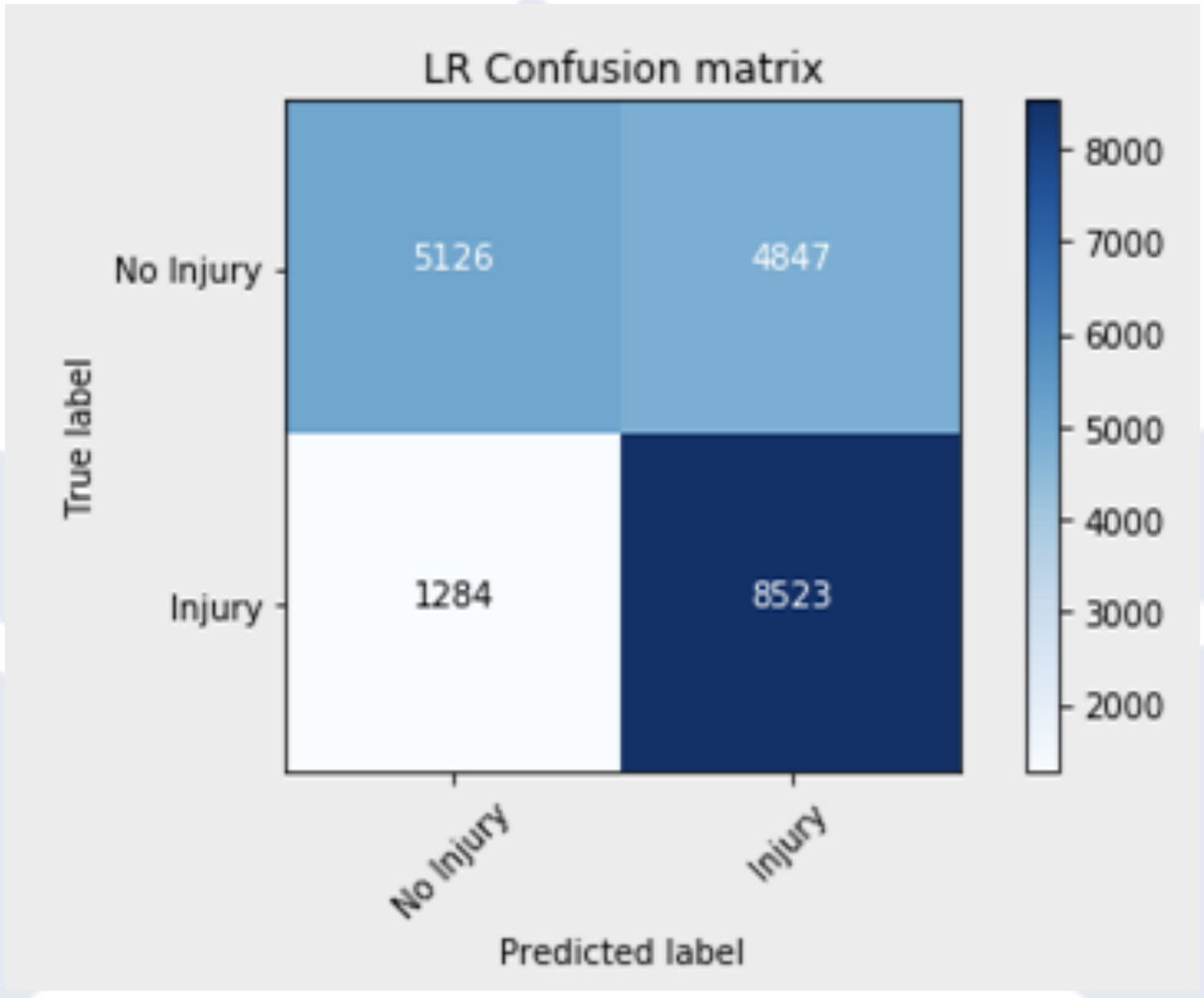
Methodology

Data preparation
Created subset of collision environment data. (ADDRTYPE, COLLISIONTYPE, WEATHER, ROADCOND, LIGHTCOND and JUNCTIONTYPE)
Insert value to missing data.
Clear or change data if necessary.

Modeling
Run Logistic Regression model
Run K-Nerrest model
Run Decision Tree model
Run Random Forest Classifier model
Run Support Vector Machine Classifier medel

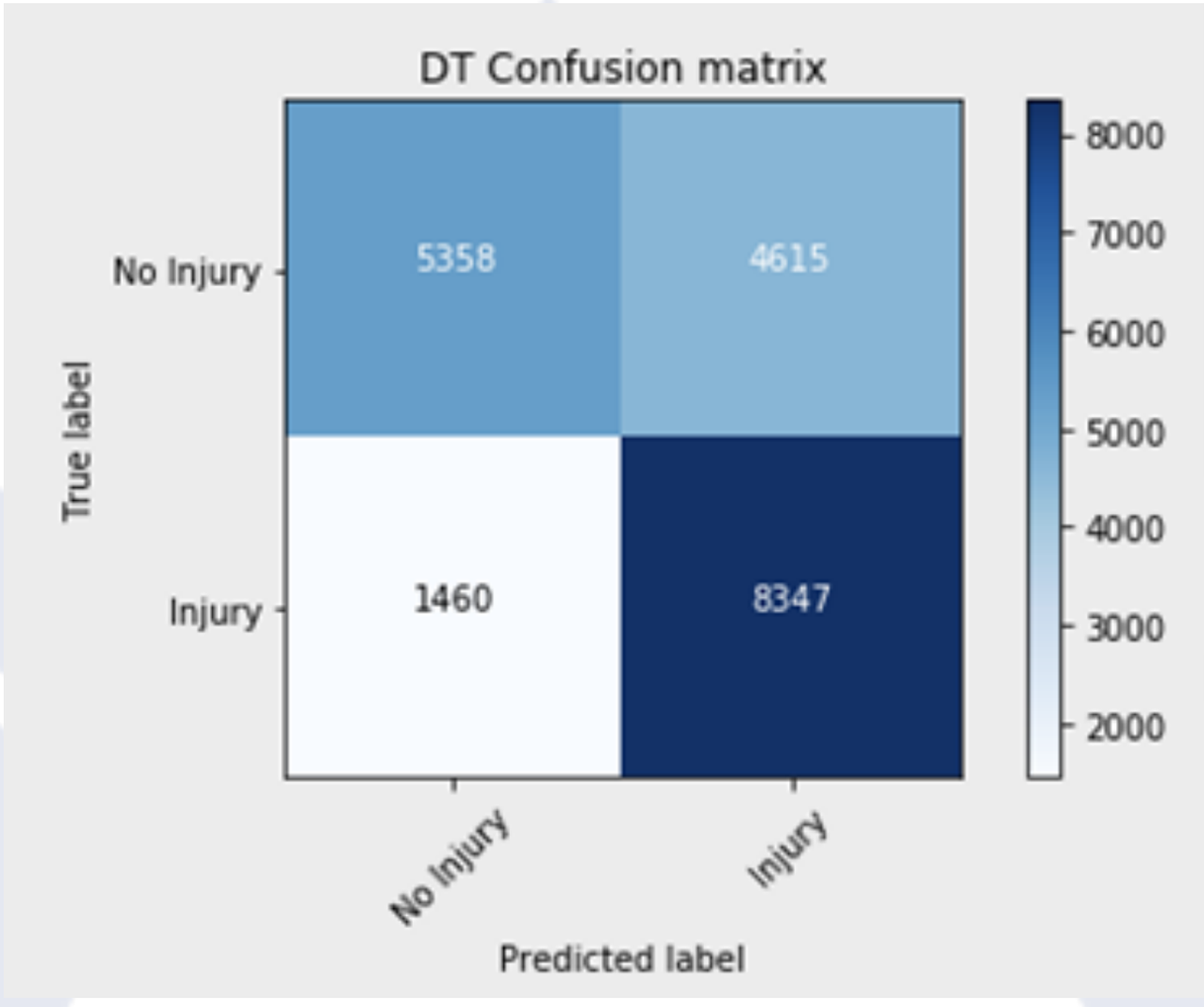
Methodology

Logistic Regression



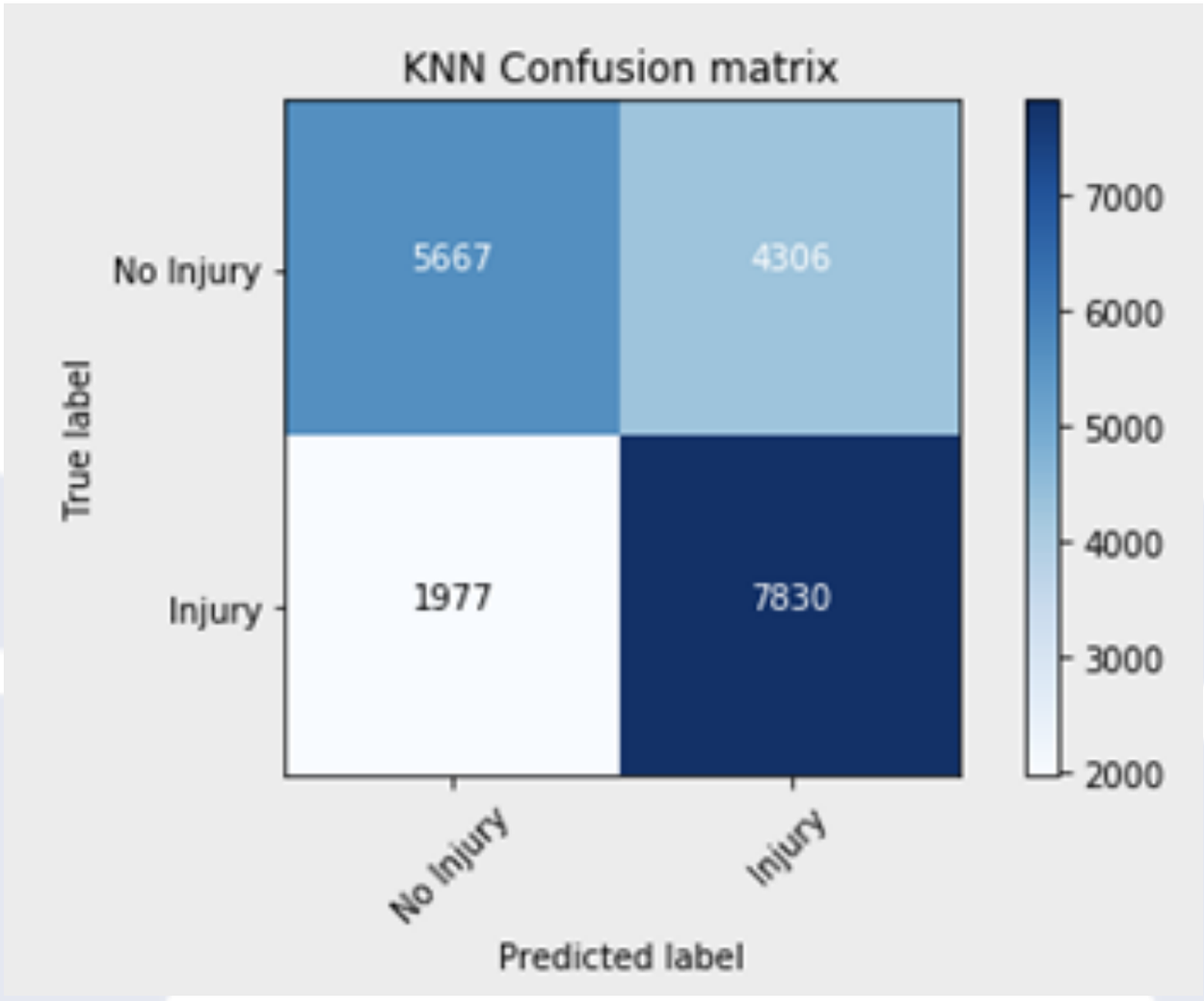
Methodology

Decision Tree



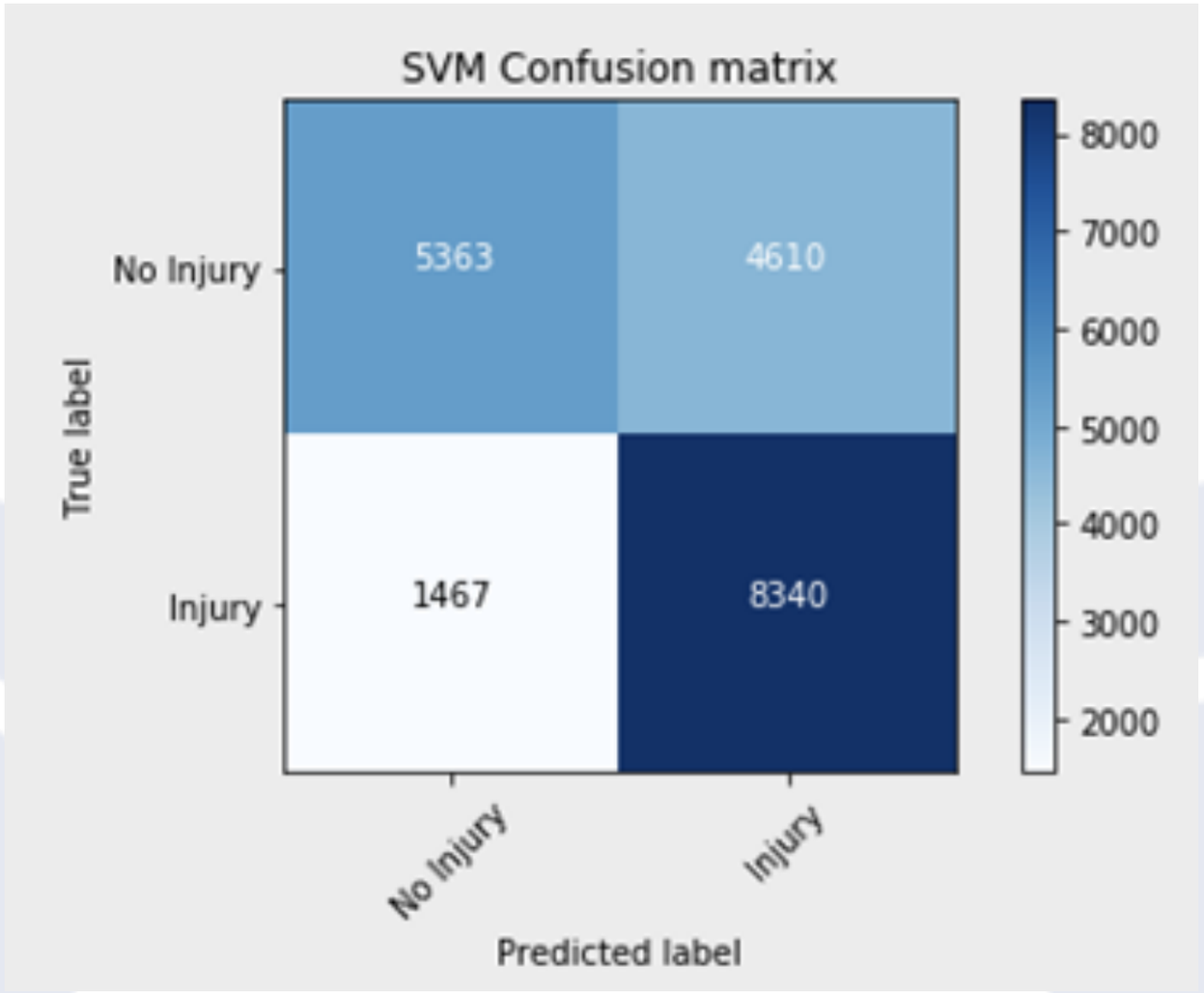
Methodology

K-Nearest Neighbors (KNN)



Methodology

upport Vector Machine



Results

After Run Model

Algorithm	Jaccard	F1-score
Logistic Regression	0.6900	0.6802
Decision Tree	0.6929	0.6853
K-Nearest Neighbors	0.6882	0.6876
Support Vector Machine	0.6900	0.6802

Discussion

Important features of collisions is Park Car. However Jaccard and F1-Score are not over 0.7. In my opinoin, the study features in this model don’t cover every factor of collisions. We Know that the collisions factor are not only the environment. Car Performance, Driver ability and habit also the maim factor of collisions.

Conclusion

This study can't cover every factor of collisions and need more study. however this study should aware the goverment and people to find out the strategy to reduce the number of collisions. such as:

* Alert people in area or environment that more opportunity to have collisions and accident.
* Set car performant standard.
* Check or test ability of drivers.
* change dangerous environment.
* Set emergency team to clear the dangerous factor.