**Introduction/Business Problem:**

Over 37000 people die in road crashes each year in addition to 4.35 million injured

or disabled. Los Angeles a popular destination among tourists has faced

accidents costing to 20 billion USD per year and has been on a constant rise

where steps are taken to predict and halt such incidents. But severity of road

accidents has been one of the least studied area though they have a bigger impact

on the casualty rate. Recently machine and statistical learning has gone

hand in hand predicting solutions to mysterious questions. With advancements

in technology, the process of predicting an event based on a given data set after

training has led to important findings. This project aims to answer questions

related to severity of road accidents in Los Angeles relating to weather and

other contributing factors and take appropriate measures to predict the severity

of road disasters. The data “A Countrywide Traffic Accident data set" by

Sobhan needed for predictive analysis is obtained from smoosavi.org, an online

data set repository. Each accident record comprises of a multitude of inherent

and contextual characteristics such as place, time, description of the natural

language, climate, time and interest points. This project seeks to contribute

the knowledge of accidents motorway accidents by creating models such as logistic

regression and other classification modules.

The expected outcome from this predictive analysis is to predict the level of severity for an accident namely “Low" and “High" based on previous accidents with predictors of weather and

other factors. Here, severity refers to the fatality of accidents with the underlying

assumptions that more the traffic delay higher the severity. The coordinates

available in the data set can be helpful in predicting further instances of accident

at a given locality and various environmental factors influencing them. Predictions

based on nearby traffic signal to considerations to enhance the predicted

model. The predicted data with “high severity" can be helpful in setting up hospitals

or amenity stations near the accident hot spots and thereby significantly

reduce the fatality rates.