

Week-2

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. These losses arise from the cost of treatment as well as lost productivity for those killed or disabled by their injuries, and for family members who need to take time off work or school to care for the injured. Road traffic crashes cost most countries 3% of their gross domestic product.

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. The number of traffic crashes and their victims has been a rising trend globally due to increases in population and motorization. Different factors involved in traffic crashes have a substantial effect on each other, thus making it difficult to individually consider any of the parameters when explaining the severity 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 Set

I worked on Seattle GeoData to Predict Accident Severity in Seattle. provided by the Traffic Records Group in the SDOT Traffic Management Division from Seattle, WA. It includes all collisions provided by the Seattle Police Department and recorded by the Traffic Record, displayed at the intersection or mid-block of a segment from 2004 to the present.

The Dataset Consists of 40 features which are

X,Y,OBJECTID,INCKEY,COLDKEY,REPORTNO,STATUS,ADDRTYPE,INTKEY,LOCATION,

EXCEPTRSNCODE,EXCEPTRSNDESC,SEVERITYCODE,SEVERITYDESC,COLLISIONTYPE,

PERSONCOUNT,PEDCOUNT,PEDCYLCOUNT,VEHCOUNT,INJURIES,SERIOUSINJURIES,

FATALITIES,INCDATE,INCDTTM,JUNCTIONTYPE,SDOT_COLCODE,SDOT_COLDESC,

INATTENTIONIND,UNDERINFL,WEATHER,ROADCOND,LIGHTCOND,PEDROWNOUTGRNT,SDOTCOLNUM,SPEEDING,ST_COLCODE,ST_COLDESC,SEGLANEKEY,CROSSWALKKEY,HITPARKEDCAR

Description of few features is as below:

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