

Capstone Project: Full Report: Car Accident Severity

Introduction

Road Accident is the most undesirable and unexpected thing to occur to a road user, though they happen quite often. Unfortunately, we can see a minatory rise of road accidents conspicuously highroad accidents over the past few years. It has a massive impact on society as well as in the economy as there is an immense cost of fatalities and injuries. according to WHO, the economic cost of road accidents to a developing country like us is 2-3% of GDP, which is a significant loss for a country like ours. Moreover, reducing this loss has become a great matter of concern for our country now.

Data Description

For the accurate prediction of the severity of accidents, a considerable number of traffic accident records with full information is required to train by using the proposed approaches. In this research work, the authors have collected a dataset from the Traffic Bureau that consists of total 194873 traffic accidents record from the year 2004-PRESENT.

The entire dataset will split into two parts- Training Dataset and Test Dataset. 70% of the whole dataset has been chosen randomly by using a python library as a training data set and the remaining 30% has been used as our test dataset. We have used the 70-30 ratio for splitting dataset because of its proven accuracy.

A. Methodology

As a database, I used GitHub repository in my study. My master data which has the main components **Severity Code, Weather Elements, Light Condition, Junction Type, Road Condition and Collission Type**. For classification problems, matplotlib is extensively used the supervised algorithm. The primary perspective of this algorithm is predicting the value of the desired variable by learning decision rules deduced from the features of the data and create a model of that.

```
In [38]: data=df[['SEVERITYCODE', 'COLLISIONTYPE', 'JUNCTIONTYPE', 'WEATHER', 'ROADCOND', 'LIGHTCOND',]]
data.head()
```

Out[38]:

	SEVERITYCODE	COLLISIONTYPE	JUNCTIONTYPE	WEATHER	ROADCOND	LIGHTCOND
0	2	Angles	At Intersection (intersection related)	Overcast	Wet	Daylight
1	1	Sideswipe	Mid-Block (not related to intersection)	Raining	Wet	Dark - Street Lights On
2	1	Parked Car	Mid-Block (not related to intersection)	Overcast	Dry	Daylight
3	1	Other	Mid-Block (not related to intersection)	Clear	Dry	Daylight
4	2	Angles	At Intersection (intersection related)	Raining	Wet	Daylight

Exploring Each Variable

Severity Code

```
In [45]: data['SEVERITYCODE'].value_counts().rename_axis('Values').reset_index(name='counts')
```

Out[45]:

	unique_values	counts
0	1	136485
1	2	58188

Collitsion Type

```
In [46]: data['COLLISIONTYPE'].value_counts
```

Out[46]:

	Values	counts
0	Parked Car	47987
1	Angles	34674
2	Rear Ended	34090
3	Other	23703
4	Sideswipe	18609
5	Left Turn	13703
6	Pedestrian	6608
7	Cycles	5415
8	Right Turn	2956
9	Head On	2024

Junction Type

Out[47]:

	JUNCTIONTYPE	counts
0	Mid-Block (not related to intersection)	89800
1	At Intersection (intersection related)	62810
2	Mid-Block (but intersection related)	22790
3	Driveway Junction	10671
4	At Intersection (but not related to intersection)	2098
5	Ramp Junction	166
6	Unknown	9

Weather

Out[49]:

	WEATHER	counts
0	Clear	111135
1	Raining	33145
2	Overcast	27714
3	Unknown	15091
4	Snowing	907
5	Other	832
6	Fog/Smog/Smoke	569
7	Sleet/Hail/Freezing Rain	113
8	Blowing Sand/Dirt	56
9	Severe Crosswind	25
10	Partly Cloudy	5

Road Condition

Out[50]:

ROADCOND		counts
0	Dry	124510
1	Wet	47474
2	Unknown	15078
3	Ice	1209
4	Snow/Slush	1004
5	Other	132
6	Standing Water	115
7	Sand/Mud/Dirt	75
8	Oil	64

Light Condition

In [51]: data['LIGHTCOND'].value_counts()

Out[51]:

LIGHTCOND		counts
0	Daylight	116137
1	Dark - Street Lights On	48507
2	Unknown	13473
3	Dusk	5902
4	Dawn	2502
5	Dark - No Street Lights	1537
6	Dark - Street Lights Off	1199
7	Other	235
8	Dark - Unknown Lighting	11

Discussion

We observe that most of the accidents in our dataset are Fatal and value for the other three classes is very low. For that reason, in second experiment, we merge Grievous, Simple Injury, Motor Collision these three accident severity classes into one class.

Therefore, we have attained the performances of the proposed approaches for two accident severity classes (Fatal / Grievous).