Car Accident Severity Project

IBM PROFESSIONAL DATA SCIENCE

Introduction/Business Problem

The idea of this project is <u>to analyze the severity of an accident in the United States of America.</u> We're trying to engineer a model to predict the severity vehicle accidents throughout the States. Millions of people die everyday due to accidents. This project could help solve this problem to a big extent.

Problem Statement: What is the magnitude of severity for an accident that occurs in USA?

Data Introduction

This is a countrywide car accident dataset, which covers 49 states of the USA. The dataset contains the driving conditions, number of people and vehicles involved in crash, and the severity of crash.

The data can be found at : https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv

Some entries were missing crucial data that were required. Some colomns were filled with "Unknown" in the number of people injured. To rectify, i dropped the entire row as a mean of the crashes wouldn't have been accurate.

Methodology – Exploratory Data Analysis

Speeding

The proportion of L2 severity is higher when the driver speeds.

Road Conditions

The proportion of L2 severity is higher when the road condition is bad

This shows that Road Conditions and Speeding certainly have a big effect on accident severity.

Machine Learning Predictive Models

Algorithms considered for classification were KNN algorithm, Decision Tree Classifier, Logistic Regression.

	Algorithm	F1-score	Accuracy
0	KNN	0.591378	0.696751
1	Decision Tree	0.576051	0.699679
2	LogisticRegression	0.576051	0.699679

Results

The coefficients have also come out to be positive which shows the imperative effect on the severity of accidents

	Intercept	Coef:SPEEDING	Coef:ROADCOND
0	-0.853729	0.067702	0.068295

As the coefficients are positive, I conclude that the 2 conditions (speeding and road conditions) have an effect of increasing accident severity.

Conclusion

- The objectives of the investigation were met

- Road traffic injuries can be prevented. Effective interventions include designing safer infrastructure and incorporating road safety features

- This model provides empirical evidence against speeding and road conditions