Applied Data Science Capstone: Predicting car accident severity

Introduction

Background: The demand for vehicles rises consistently. Consequently, so does the number of vehicles on the road and the probability of involvement in traffic jams or car accidents. Traffic accidents result not only in loss of human lives but have a huge impact on economy as well. According to the Michigan Traffic Crash Decade-At-A-Glance study[1], there were over 314.000 traffic accidents in US in 2017, which translate into costs around 230 billion dollars per year. Approximately 1.3 million people die yearly as a result of road traffic crashes.[2]

Problem: By collecting and analysing relevant data on car accidents, we aim to answer the following questions:

1. To what extent can weather conditions help us predict the severity of car accidents?
2. Are the total number of people involved in the collision or the number pedestrians relevant in predicting car accident severity?
3. How do road conditions or light conditions influence the car accident severity?
4. Is driving under the influence of drugs or alcohol important in predicting car accident severity?

Interest: Car drivers and vulnerable road users (pedestrians, cyclists, motorcyclists) would or should be interested in an equal manner by the result of this inquiry. Having measured the impact of the abovementioned outer and inner conditions and knowing the results, one might act differently. Having established the (probable) paramount influence of drugs or alcohol in causing car accidents of great severity, authorities might decide to impose higher taxes on such vices.

---

[1] Michigan State Police, Michigan Traffic Crash Decade-At-A-Glance, 2018.
[2] https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries