

# Car Accident Severity

## Capstone Project

### Week 1\_Assignment

#### Report \_Section: Introduction / Business Problem

##### 1. Introduction / Business Problem

###### 1.1. Background

Road traffic injuries (RTIs) are a major public health problem. The annual global status reports on road safety, launched by the World Health Organization (WHO), highlights that the number of road traffic deaths has exceeded one million in recent years. That is over 3000 people dying on the world's roads every day. [1]

To combat this major problem, the volume of research in the areas of accident analysis and prediction has been increasing over the past few decades. [2] The analytical data mining solutions can be largely employed to determine and predict related influential factors and thus to explain RTIs severity.

Among the analytical data mining solutions, supervised machine learning (ML), has become a popular scientific method to predict the severity of accidents. The reasons for this popularity, can be referred to the capacity present in ML to identify the existing patterns in the data and make predictions via the establishment and evaluation of diverse algorithms. Moreover, the ability of MLs to handle large amounts of data is an additional asset for this purpose, as the data on road traffic accidents are often sparse and largely extended.

###### 1.2. Objective

The objective of this project is to predict the car accident severity by training and evaluating supervised machine learning algorithms with the help of a shared dataset including the record of collisions provided by SDOT Traffic Management Division, Traffic Records Group (timeframe: 2004 to present) for the city of Seattle, a city on the west coast of the United States.

As the approach to reach this objective, the cross-industry standard process for data mining (CRISP-DM) (Figure 1) will be implemented. During the coursework of this procedure, the raw

data will be accordingly well understood and prepared before being fed for the predictive modeling analysis in the next steps.

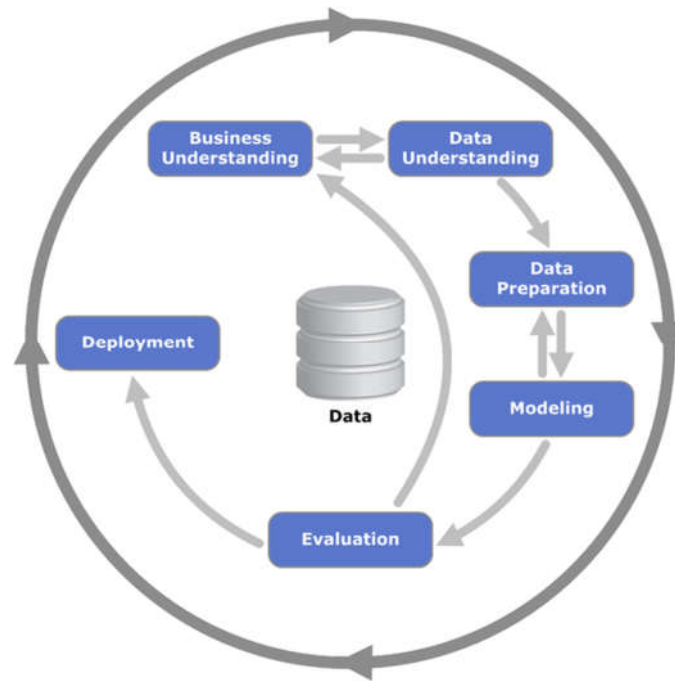


Figure 1, Cross-industry Standard Process for Data Mining (CRISP-DM) [3]

### 1.3. Interest

The report of this project can be targeted to stakeholders, who are involved with road traffic injuries incl. road administrators, traffic control authorities, and emergency road services in order to help them predict the car accident severities and improve the road users' safety margins.

### References:

- [1] World Health Organization: <http://www.who.int>
- [2] Accident Risk Prediction based on Heterogeneous Sparse Data: New Dataset and Insights, S. Moosavi et al., SIGSPATIAL '19, November 5–8, 2019, Chicago, IL, USA
- [3] Cross-industry Standard Process for Data Mining (CRISP-DM), <https://en.wikipedia.org>