

## Predicting Car Accident Severity Based on Traffic Conditions

### Introduction

Automobile accidents are a responsible for over \$800 billion [1] and over 30,000 casualties [2] in the United States each year. Identifying factors which most contribute to accidents could help better formulate policies which could reduce loss of life and save billions of dollars. In this paper, we will build a predictive supervised machine learning model to predict the potential severity of an automobile accident based on existing traffic conditions.

### Data

For this paper, we will examine the data included in the example dataset provided by the Coursera Applied Data Science Capstone course. This dataset contains 194,673 accident observations and records 38 attributes for each accident. We will look to build a feature set useful for accurately predicting how severe an automobile accident will be given existing traffic conditions. The following features will be used to predict automobile accident severity:

Feature	Description
JUNCTIONTYPE	The type of junction where the accident occurred
INATTENTIONIND	Indicates cause of accident related to inattention
UNDERINFL	Whether or not the driver was under the influence
WEATHER	The weather conditions when the accident occurred
ROADCOND	The road conditions when the accident occurred
LIGHTCOND	The lighting conditions when the accident occurred
SPEEDING	Whether or not speeding was a cause of the accident

The target column SEVERITYCODE will be predicted based on these predictor variables.

### References:

[1] <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812013>

[2] <https://www-fars.nhtsa.dot.gov/Main/index.aspx>