Introduction

Traffic accidents are a major cause of death globally. More than 38K people die every year in crashes in US roads. The fatality rate is 12.4 deaths per 100,00 inhabitants[1]. If your home is metropolis city chances are that you have heard, witnessed, or even involved in one. If we can predict the traffic accidents or their intensities, it can potentially save many lives.
Even though some of the accidents happen because of the careless nature of the people involved, which would be impossible to predict. But most of the remaining accidents are influenced by many quantifiable factors like weather conditions, car types, driving speeds, road structures and many others.

Imagine the possibilities where the predicted high risk for accidents environment can be used. For instance, the areas where under a certain condition is expected to have high risk of accidents the emergency services can be in close range to that to reduce the response time and which may be potentially save lives. Similarly, the highway patrols can impart some mechanism during certain conditions to reduce the speed of the passing vehicles which can very likely reduce the risk of the accident to happen in the first place. Another situation where this knowledge could be used is the department of roads construction. They can study more into road nature and how improving the way its build so that it can adapt to weather change with no added risk factor for the people commuting on it.

In light of all the mentioned instances and many more possibilities where this predictive analysis can be very crucial understanding what factors are having the most influence on the severity of the accident and how we could possibly save lives eventually.

Given a dataset of all those quantifiable features and the intensity of the accidents can potentially help in building a predictive model. Fortunately, several such public datasets are available, like the one being used here.

[1] Based on analysis of data from US Department of Transportation in 2019