CAR ACCIDENT SEVERITY PREDICTOR

**Introduction/Business Problem**

Background

Traffic accidents cause serious threat to the human life worldwide. In order to take necessary actions to control this ever-growing problem extensive research has been carried out into the prediction of traffic accidents in both developed and developing countries using various statistical techniques. A number of factors contribute to the risk of car accidents. Some of the important ones are road conditions, weather, type of junction and light conditions. These conditions also influence the severity of the accidents. In this paper, we aim to predict the severity of car accidents through machine learning models.

Business Problem  
There are several factors that can contribute to an accident. In this study, I would delve into the three main factors namely the weather condition, the road condition and the light conditions during the collision to predict the severity of car accidents so that drivers of vehicles would be given a notice as to the best conditions to drive in.

Interest  
The SDOT Traffic Management Division can use a car accident severity predictor in making policies on road safety. This makes them a good target audience for this project.

The results of the project would be particularly helpful for them to plan and improve the road conditions in the city and placement of street lights. They would also be curious to know the influence of other factors like weather and visibility conditions on the occurrence of accidents and the level of severity. Also, the model can be generalized to other areas and cities aside Seattle helping drivers to understand the conditions to which driving can be risky and take precaution. This would go a long way to prevent accidents from occurring.

2. Data Source

The dataset used in this project includes all collisions from Seattle from 2004 to present. This dataset is provided by SPD and recorded by Traffic Records. The dataset has 38 columns and 194,673 entries. The target variable from the dataset is “SEVERITYCODE”. It is made up of numbers, which correspond to different levels of severity:

0: no/negligible chance  
1: Chance of car damage  
2: Chance of injury and/or car damage  
3: Chance of serious injury and car damage

4: Fatal and car damage

Out of the 37 independent features, the ones that would be of interest in this study are the following:

“WEATHER”: A description of the weather conditions during the time of the collision.

“ROADCOND”: The condition of the road during the collision.

“LIGHTCOND”: The light conditions during the collision.

These attributes were selected because they directly affect the severity of road accidents and also they had a large number of entries that can make a reasonable dataset unlike the other independent features.