**Identification of Car accident severity using Machine learning Model**

A description of the data and how it will be used to solve the problem

**2. Data acquisition and cleaning**

**2.1 Data sources:**

I have used data which is available through coursera.

**2.2 Data cleaning:**

Data contains number of fields are repetitive and some fields are completely blank. For this, we need to preprocess our data.

**2.3 Feature selection:**

**In feature selection, I have selected following field which makes impact on car accident.**

car\_df.columns

Index(['X', 'Y', 'ADDRTYPE', 'LOCATION', 'SEVERITYCODE', 'COLLISIONTYPE',

'PERSONCOUNT', 'PEDCOUNT', 'PEDCYLCOUNT', 'VEHCOUNT', 'INCDATE',

'INCDTTM', 'JUNCTIONTYPE', 'SDOT\_COLCODE', 'SDOT\_COLDESC',

'INATTENTIONIND', 'UNDERINFL', 'WEATHER', 'ROADCOND', 'LIGHTCOND',

'PEDROWNOTGRNT', 'SDOTCOLNUM', 'SPEEDING', 'ST\_COLCODE', 'ST\_COLDESC','HITPARKEDCAR'],dtype='object')