**CAPSTONE PROJECT REPORT(Vehicle Collision)**

* 1. **Introduction**
  2. **Background**

These days vehicle accidents are getting very serious issue globally leading to deaths and serious injuries.Here we are concerned abour car collisions and will try to understand the factors on which collisions happens or occurs.

* 1. **Problem**

Problem occurs with business understanding of the above mentioned issue. We need some solution to reduce or eventually stop this collisions. In an effort to reduce the frequency of car collisions in a community, an algorithim must be developed to predict the severity of an accident given the current weather, road and visibility conditions. When conditions are bad, this model will alert drivers to remind them to be more careful.

* 1. **Intrest**

Road Transport Corporation, Drivers & Highway Authorities.

* 1. **Data**

**2.1 Data Sources:**

Data was given by Coursera only. Used that only.

**2.2 Data Understanding:**

Our predictor or target variable will be 'SEVERITYCODE' because it is used to measure the severity of an accident from 0 to 5 within the dataset. Attributes used to weigh the severity of an accident are 'WEATHER', 'ROADCOND' and 'LIGHTCOND'.

Severity codes are as follows:

0 : Little to no Probability (Clear Conditions)

1 : Very Low Probablility - Chance or Property Damage

2 : Low Probability - Chance of Injury

3 : Mild Probability - Chance of Serious Injury

4 : High Probability - Chance of Fatality

In it's original form, this data is not fit for analysis. For one, there are many columns that we will not use for this model. Also, most of the features are of type object, where they should be numerical type.

We must use label encoding to covert the features to our desired data type. With the new columns, we can now use this data in our analysis and ML models.