1. **A description of the problem and a discussion of the background. (15 marks)**

Road vehicle accidents are a global problem that effect lives in a couple of ways like increase in traffic, causing harm to the people involved in the accident, damage to vehicles etc. Accidents are caused by various factors ranging from poor road infrastructure and management, unenforced traffic laws, to unsafe road user behaviour. On average, road crashes cost countries 3% of their gross domestic product. Data also shows that the average severe accident cost has a very wide range due to the random nature of the event from $250,000 to $1 not including injury, damage to fleet reputation or long term impacts associated with higher insurance premiums making the financial benefits of preventing such incidence obvious. Therefore, it is advantageous to predict how severe an accident’s effect is. This information can be used in making data driven decisions.

Considering that our problem is a binary classification one, we will use an F1 score to evaluate the performance of the different models that we will train. For each model, the value for the global score will, in turn, be the average of the F1 scores for each class. So, we will be using a *one vs all*approach. However, we will also take a close look at *accuracy*, *precision* and *recall*; since they all provide valuable insights into understanding how the model is performing.

Finally, we will explain and give details on what are some possible next steps to further improve the overall performance and interpretability of the chosen model.