Predict severity of an accident

Marcelo Quiroga

September, 2020

1. Introduction

Traffic accidents are a recurring problem in every corner of the world. Perhaps the biggest problem is the fatality caused by these accidents because the material damage can be compensated but the human lives are irreplaceable, even if life is not at risk, the fact of being seriously injured remains a very big problem.

The first problem is the fatality of the injured, the second problem is the possibility of being seriously injured. Another problem no less important are the costs in repairing damages, costs that insurers assume, costs that governments assume and a long list of affected.

Governments in each country should analyze these types of studies to issue laws that try to reduce the risk of traffic accidents, the frequent driver who goes to their work every day should know the prevention recommendations that come from these reports because they aim to take care of their lives and the lives of their families.

2. Data

The data set used is based on Seattle city records comprising from 2004 to 2020, there are more than 194 thousand records consisting of different variables totaling 37. These variables include data external to the driver and mobility, such as: weather status, date and time, road conditions, light conditions, among others. Also included are variables own or internal to the driver: alcohol, speed, driver responsibility.

The objective will be to predict the severity of accidents using the variable 'severity', the possible values for this variable are two: injury and prop damage; this represents a binary analysis. Within the dataset you can see 30% of lesions and the remaining 70% prop

damage. This creates the 'imbalance issue' problem and for this reason the relevant recommendations will be used[1].

Recommendations from other similar reports to address this analysis will be taken as an additional reference[2].

[1] How to Handle Imbalanced Classes in Machine Learning - https://elitedatascience.com/imbalanced-classes
[2] High-Resolution Road Vehicle Collision Prediction for the City of Montreal - https://arxiv.org/pdf/1905.08770.pdf