Predicting Crash Severity and Fatalities for NZ Road Accidents

Problem Statement

Car crashes and road accidents could be considered an old topic. Yet, with the progress of the technology involved and the capabilities of these sophisticated machines, it is ever more important to have tools and means available to mitigate their occurrences, as well as their implications and consequences for the people involved. Thanks to the advancement of technological and analytical tools in the last two decades we are now able to better understand how crashes happen. This enables the transport, security and emergency agencies all around the world to have different (predictive) models for quickly analyzing crashes when they happen and dispatch an appropriate response swiftly. Many attempts have been taken by many professionals, scholars and government agencies to provide produce these models; each with different goals, ways of measuring success and precision of their results

Predicting the severity of a car crash is no easy task. And even when possible, precision levels will vary significantly depending on the data available and how well the system or model has been defined. However, if the dataset's features are clearly defined and if there's a thorough description of how this data is collected we have much better chances of arriving at a usable model. In the dataset I'll use data associated with car crashes come in a hybrid mode; meaning we have both categorical and numerical features. This allows us to treat the problem from a mathematical approach and to use performance metrics such as R2 , ROC curves, precision, accuracy and F Scores.