1. INTRODUCTION

Driving is one of the most common part of life for many people. Although it is hard, if not impossible, to predict the happening of an accident, it is easier to predict the severity of an accident. Therefore, an accurate prediction would help drivers to make a decision when they know the level of severity of the accident happened in front of them. Having an accurate prediction is very important for a GPS service provider, such as Google Map, to inform the driver and to choose a suitable route for the driver based on the estimation of the severity.

To meet this demand, this project focuses on the prediction on the severity of an accident based on the data provided by SPD. To achieve a good result, the project chooses several advanced machine learning techniques to build the model so that the model will help us to answer the business question: based on the information we have, whether the accident is an injury (with severity code as 2) or a property damage (with severity code as 1)?