MSc Project report

Prediction of causes of accidents using Logistic regression algorithm

Sivakumaran Kathamutthu

2018

A report submitted as part of the requirements for the degree of

MSc in Computing: Software Technology at Robert Gordon University, Aberdeen, Scotland

Abstract

The predictive analytics is an analytic technic which used to predict of unknown future or current events of a specific domain. Predictive analytics is such important where it can generate much valuable information of any organization to make business decision for future work. Today there are many incidents happens around us which perhaps couldn't think of how it could have been happened as the related party who really engaged or impacted in to the incident might not be able to reach or some time it's kind of incident where its mortal hence couldn't survive of their life. One of such incident I could think of is, road motor vehicle accident happens and its cause is still hidden. But in many developed countries uses some high technic road camera to tackle these kinds of information of road motor vehicle accident or any other accident, but still it's not practical to install such camera to all over the country in any country in the world.

This research provides a comprehensive road motor vehicle accident cause predictive model with high accuracy and reliable as it will be trained with more than six hundred thousand of large already collected dataset before it predicts any given accident information to predict the cause. This solution will be used supervised machine leaning algorithm is Logistic regression algorithm which has been identified specially by its performance, minimal error rate in cross validating of each attribute in accident data attributes and its cause.

This solution comprised with four component such as

1. Amazon cloud S3 bucket to upload road vehicle accident data set as it when collected

2. Bigdata spark cluster to process the dataset uploaded to S3 and store processed data to database

3. Predictive engine which utilized the processed data in database to predict the given accident cause

4. User interface application where user to submit unknown accident cause information.

Also, it has been identified the enhancement of this solution to bring this to next level where to make in to mobile application which could accessible quick and easy to anytime by any situation and any location to quickly predict any unknown situation come of road motor vehicle accident.