# The System of Risk Assessment and Prevention of Road Accidents

Extended Abstract


## ABSTRACT

The research presents how to predict the probability of a traffic accident on one or another part of the road using machine learning algorithms. Machine learning techniques to model traffic accident data records can help to understand the characteristics of drivers' behavior, roadway condition and weather condition that were causally connected with different injury severity.



## KEYWORDS

Traffic accident, prediction, AdaBoosted Decision Trees


## 1 INTRODUCTION

The costs of fatalities and injuries due to traffic accidents have a great impact on the society. In recent years, researchers have paid increasing attention to determining factors that significantly affect severity of driver injuries caused by traffic accidents [1]. There are some conventional ways to prevent them: installation of signs on technically dangerous parts of the road; prevention of emergency weather services by the Ministry of Emergency Situations; collection and synthesis of accident statistics. However, there is a lack of modern and innovative ways to prevent road traffic accidents. There is a necessity to improve this area. The relevance of the project is well-timed notification for drivers about the possibility of an emergency on a particular section of the road. In addition, it would help for optimization of the work of the services that ensure the normal functioning of the road traffic (traffic police, road services).The main goal of the project is to create a service that predicts the probability of a traffic accident on one or another part of the road. The main objectives of the project are collection and analysis of the database; classification training based on classical machine learning algorithms.

## 2 AN OVERVIEW OF ALGORITHMS AHD RESULTS

### 2.1 Business Logic

The business logic of the service is as follows (Figure 1). Various data, such as weather conditions, data from the on-board computer, an existing database, data from the GPS-navigator are input. The risk of a road traffic accident will be assessed based on this data. At the output, we can get both a library with algorithms and a service.

### 2.2 Data Collection and Aggregation

This research used data from the Russian data source web portal [2]. The initial dataset for the study contained traffic accident records from 2015 to 2018.

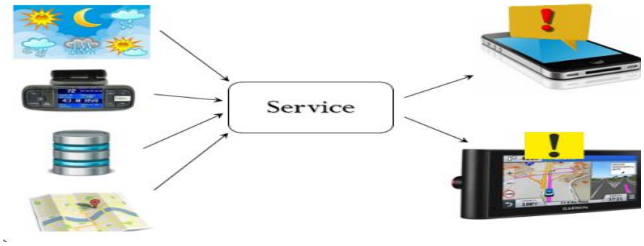


**Figure 1 Business logic.**

### 2.3 Machine Learning Paradigms

Machine learning techniques, which are used: Linear Regression, Random Forest Classifier, Gradient Boost Classifier, Ada Boost Classifier and Decision Trees. The best method for prediction was the Multi-class AdaBoosted Decision Trees. Similar to AdaBoost in the two-class case, Multi-class AdaBoosted Decision Trees algorithm combines weak classifiers and only requires the performance of each weak classifier be better than random guessing (rather than 1/2). In addition, similar to the original AdaBoost algorithm for the two-class classification, it fits a forward stagewise additive model. This algorithm is highly competitive with the best currently available multi-class classification methods, in terms of both practical performance and computational cost [3]. It predicts the most likely type of traffic accident among the nine possible: getting out of the road, hitting a pedestrian, hitting a bicycle rider, a rollover, hitting a motionless vehicle, hitting an animal, head-on crash, hitting an obstacle, another kind of crash.

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.00 | 0.00 | 0.00 | 0 |
| 1 | 0.99 | 0.88 | 0.93 | 83 |
| 2 | 0.88 | 0.37 | 0.52 | 19 |
| 3 | 0.18 | 0.67 | 0.29 | 3 |
| 4 | 0.00 | 0.00 | 0.00 | 1 |
| 5 | 0.80 | 0.67 | 0.73 | 18 |
| 8 | 0.00 | 0.00 | 0.00 | 0 |
| 9 | 0.00 | 0.00 | 0.00 | 0 |
| 91 | 0.30 | 1.00 | 0.46 | 3 |
| avg / total | 0.90 | 0.76 | 0.81 | 127 |

**Figure 2 Multi-class AdaBoosted Decision Trees.**

As seen, the model was able to predict the type of crash occurrence with an accuracy of approximately 76%.

## REFERENCES


[1] Miao Chong, Ajith Abraham, Marcin Paprzycki. 2004. Traffic Accident Analysis Using Machine Learning Paradigms. Informatica 29 (2005) 89–98.

[2] https://xn--80abhddbmm5bieahtk5n.xn--p1ai

[3] Ji Zhu, Saharon Rosset, Hui Zou, Trevor Hastie. January 12, 2006. Multi-class AdaBoost. https://web.stanfo.edu/