

A NOVEL MACHINE LEARNING ALGORITHMS TO ANALYZE THE TRAFFIC ACCIDENTS

Author 1: Internal guide details
K.Srilatha Reddy M.Tech(CSE)
Assistant professor,
Department of Computer Science Engineering,
Teegala Krishna Reddy Engineering College,
Meerpet, Hyderabad,Telangana,India

Author 2: Mohammed
Rizwan
M. Tech (CSE)
Roll No 16R91D5815,
Teegala Krishna Reddy Engineering College,
Meerpet, Hyderabad,Telangana,India

ABSTRACT:

There have been numerous studies on traffic accidents and their severity, particularly in relation to weather conditions and road geometry. In these studies, traditional statistical methods have been employed, such as linear regression, logistic regression, and negative binomial regression modeling, which are the most common linear and non-linear regression analysis methods. In this research, machine learning architecture was applied to this problem using the random forest, artificial neural network, and decision tree techniques to ascertain the strengths and weaknesses of these methods. We agree that the measures must be based on objective and clinical studies of the sources of injuries and the severity of accidents to reap the destination discount's maximum possible results with limited assets in the budget. This document summarizes the four systems study models' overall performance implemented to model the severity of injuries that occurred at some point in site visitor injuries. We consider neural networks learned in hybrids by gaining knowledge of processes, steering machines, selection bushes, and asynchronous hybrid model on choice wood and neural networks. The experiment's effects disclose that the

different device study models considered, the hybrid tree selection technique, outperformed the individual techniques.

Keywords: *Accidents, Traffic management, ANN, CNN, location.*

1. INTRODUCTION

More than 3000 people die in road crashes every day worldwide, including approximately 1000 people in the United States. Such accidents cost \$518 billion globally and \$230.6 billion per year in the United States. In particular, many young people are affected by traffic accidents, as most traffic accidents are caused by the 15–29 age group (Association for Safe International Road Travel Data). Hence, substantial efforts have been made to address the main factors that cause traffic accidents, such as speeding, distracted driving, driver fatigue, road conditions, use of mobile phones, and poor weather conditions. Regarding bad road conditions due to the weather, previous studies have revealed that there is a strong correlation between road friction coefficient and traffic accident risk. The friction

coefficient can be influenced by road conditions such as ice or wetness caused by snow or rainfall. Research has revealed that single-vehicle accidents are mostly caused by wet weather [3]. When a layer of water builds between the tire of a vehicle and the road surface, hydroplaning occurs. If all wheels experience hydroplaning simultaneously, the driver can no longer control the vehicle. It has been shown that hydroplaning occurs at vehicle speeds of 80 km/h on thick water films when the water depth exceeds 2.5 mm. According to the Korean Transportation Safety Authority, the total number of fatalities in road traffic accidents in the country has decreased, but rain-related road deaths increased from 430 in 2013 to 463 in 2016. In particular, in Seoul, the traffic fatality rates on rainy days were the highest among all cities [7]. Traffic accidents are unpredictable

and can happen anywhere and at any time. However, drivers can at least be provided with useful information to help avoid accidents or reduce their probability. The forecasting of traffic accidents and the identification of related factors under various conditions are important for preventing accidents and reducing their frequency. For these reasons, traffic accident prediction models have been developed to reveal the significant factors that affect traffic accidents so that traffic safety can be improved by controlling and/or improving these factors [8]. The rest of this paper is organized as follows. Next, we provide literature reviews on traffic accident forecasting and machine learning applications. The methodologies of the proposed random forest, artificial neural networks (ANN), and decision tree algorithms are introduced, and the data sets employed are described. The modeling results are presented as the estimated values (importance) of factors influencing accidents as

well as interpretations of the model evaluations including the mean square error (MSE) and root mean square error (RMSE). Finally, a discussion is provided and conclusions are drawn.

2. RELATED STUDY

The Twitter information is both noisy and unstructured. An effective text mining method is necessary to extract the useful accident-related information from tweets. In this study, we employ and compare two deep learning methods: Deep Neural Network (DNN) and Long Short-Term Memory (LSTM), in training and classifying the accident-related tweets. Unlike classifiers such as logistic regression, Support Vector Machines (SVMs) or Artificial Neural Network (ANN) with a single hidden layer, deep learning does not seek direct functional relationships between the input features and the output classification results. Instead, it is a set of machine learning algorithms that attempt to learn in multiple levels, corresponding to different levels of

abstraction (Deng and Yu, 2014). The training process of DNN is divided into multiple layers, and the output result is expressed as a composition of layers, where the higher level features are the composition of lower-level features, giving the potential of modeling complex data with fewer units than a similarly performing shallow network (Bengio, 2009). Our efforts can be detailed by three major contributions: First, we propose a systematic feature selection process in extracting both the individual and paired token features from social media. We unveil the language customs of the tweet users in describing the traffic accident detection. Second, we validate the effectiveness of the deep learning approach in classifying the social media data. The results show that deep learning outperforms other prevailing data mining methods. Third, the advantages and disadvantages of tweets in accident detection are verified and fully discussed by comparing tweets with

both accident log from state Department of Transportation and traffic data from thousands of loop detectors.

We only study the tweets which have explicit indications of traffic accidents and do not include the traffic congestion, construction works, etc. because the latter may not necessarily indicate traffic accidents. The types of the traffic accidents in this paper can be summarized into 3 categories: “collision”, “disabled vehicle” and “vehicle on fire”. The raw tweet data need to be preprocessed to constitute the database that can be used for further analysis. The first step is to extract the candidate tweets that possibly describe the on-site traffic accidents. Usually, these candidate tweets should contain one or more keywords such as “accident” or “crash” that are accident-related. However, there has been no consensus on such a vocabulary of the accident-related words. Thus, we turn to the traditional news media and collect about 100 articles of

news that discuss the traffic accident. In all these articles, we select the words that appear the most frequently. The frequency of a word is the times that a specific word appears in these articles.

3. PROPOSED SYSTEM

Models are created using accident data that can help understand features of many capabilities, such as driver behaviour, road conditions, lighting conditions, weather conditions, etc. It can help customers calculate useful protection measures to avoid injury. It can illustrate the statistical approach based on guided graphs by evaluating two possibilities based on out-of-sample projections. The model is achieved to be aware of statistically intrinsic factors that might anticipate the probability of accidents and injuries that could be used to implement and reduce the hazard component. Here, a road-matching analysis is done by reading some records using certain inquiries that apply to the analysis. Inquiries such as the maximum time the force

is at risk and the injury fractures that arise in rural and urban areas and various regions. What is the trend in the number of casualties occurring each year, accidents in excessive frequency areas having more casualties, etc.? This data can be accessed using a Microsoft Excel sheet, and it can obtain the specific answer. This assessment aims to highlight the most important information in a traffic accident and predict it. It can check this technique's results in the next section of the file.

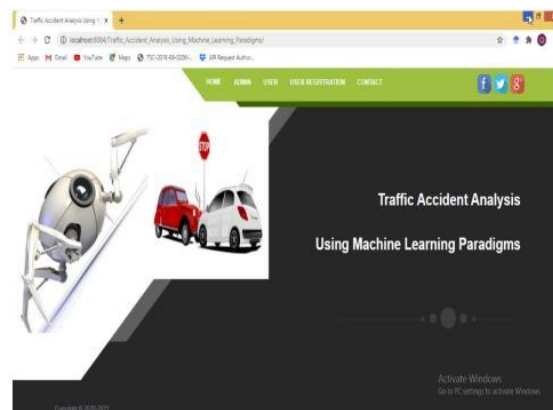


Fig.3.1. Home page.

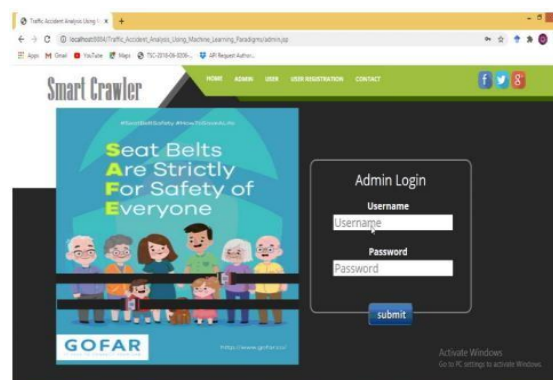
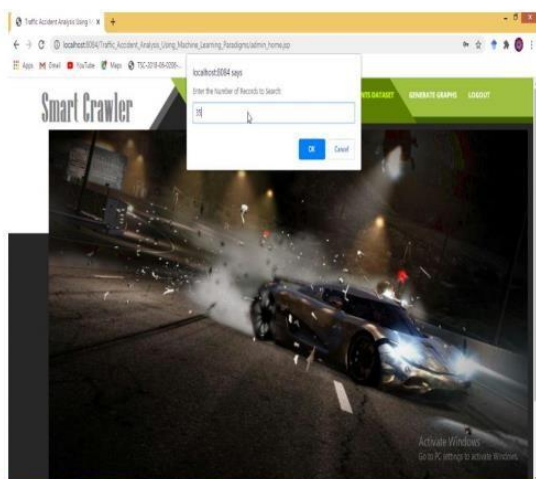


Fig.3.2. Admin login.**Fig.3.3. Enter dataset record number.**

As compared to traditional media, social media blogs are short, brief with few editorial review. Some of the words may have grammatical errors. Also, some tweets posted by the influential users do not provide valid location information. Thus, to ensure both the accuracy and sample size, certain rules are further implemented:

- Include the words that are relevant to accidents but apparently misspelled or personally modified including “accident”, “incident”, etc.
- Include other variations of accident-related words such as the word pairs that have a hyphen in word pair such as roll-over, etc.

- Exclude the tweets posted by influential users including web media, transportation authorities or Department of Transportation.

Note that some public tweets may mention the names of influential users or hashtags and should not be deleted. Also, there are some tweets posted by the personal Twitter account of the reporters working for the authorities. They may still give valid location and time information and should not be deleted.

State	accidents	killed	injury	vehicles	Accident charge	Over Speed
Andhra Pradesh	11028	2290	12707	208	120	80
Assam	85	26	148	426	88	63
Bihar	2277	762	1897	862	200	67
Chhattisgarh	2714	1800	1954	805	78	70
Chandigarh	209	91	327	12018	1	65
Goa	6483	2020	8954	845	140	72
Haryana	67	43	124	14818	0	67

Fig.3.4. Year wise accident report.

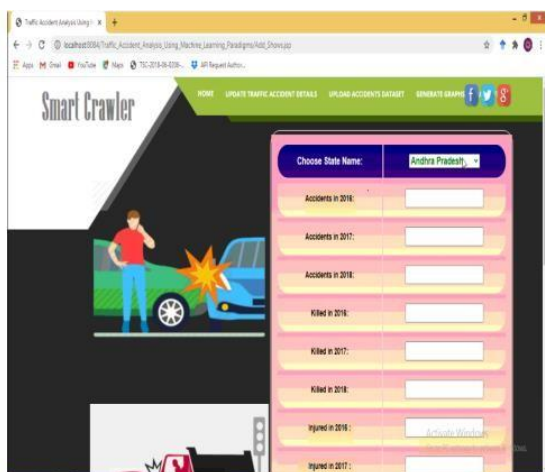


Fig.3.5. Update traffic accidents details.

4. CONCLUSION

In this document, a deep learning-based Traffic accident prediction Convolutional neural network (TAP-CNN) methodology is suggested for visitors to predict the severity of fate's development. Unlike previous strategies that only consider the superficial structure of visitor matchmaking, the proposed technique correctly funds the representation of the latent visitor match characteristic, which consists of the overall function and deeper characteristic associations. Matching information for website visitors. The overall performance of the proposed TASP-CNN model was assessed using traffic matching information over an eight to 12-month period

(2009-2016) from Leeds City Council and altered compared to NBC, KNN, LR, DT, GB, and SVC, if Conv1D, NN, and LSTM-RNN are all common. Effects show that the proposed TASP-CNN model has become a model outperforming the competing fashions.

REFERENCES

- [1] Abdel-Aty and Abdelwahab, H, 2003, "Analysis and prediction of fatal traffic accidents resulting from corner collisions including the impact of vehicles"
- [2] Abdelwahab and Abdel-Aty, "Development of artificial neural network models to predict the severity of driver injuries in traffic accidents at signalized intersections", Paper No. 01- 2234.
- [3] Bedard, M. J., & Hireds, J. P, "The independent contribution of driver, accident and vehicle characteristics to driver deaths", pp. 717-727, 2002.
- [4] Buzeman, and Lovsund, P, 1998, "Safety of vehicle occupants in

frontal collisions: a study of vehicle mass parameters, collision speed, and inherent vehicle protection”, pp. 713-722.

“Injury Trends of Passenger Car Drivers In the USA”, pp. 541-557.

[5] Rose, G and Dia, H, 1997, “Development and evaluation of highway accident detection models for the neural network using field data”, pp. 313-331.

[6] Evanco, W. M., 1999, “The potential impact of rural emergency systems on auto accident deaths”, pp. 455-462.

[7] Hand, D., Mannila, H, & Smyth, P., Principles of Data Mining. The MIT Press, 2001.

[8] Kim and L., Richardson, 1995, “Personal and behavioral predictors of car accidents and injury severity”, pp. 469-481.

[9] Kweon, Y. J., & Kockelman, D. M., 2003, “The overall risk of injury for different drivers - a combination of patterns of exposure, frequency and severity”, pp. 441-450.

[10] Martin & Pilkey, W. D., 2000,