Road Accident Prediction using Machine Learning

Asghar Pasha[1],Vijayalakhshmi[2],MD Atique[3], MD Hussain[4], Harsh narnot[5],Bipin[6]

[1]Asghar Pasha Dept. of Information Science and Engineering, Hkbk College of Engineering, Karnataka, India
[2]Vijayalakhshmi Dept. of Information Science and Engineering, Hkbk College of Engineering, Karnataka, India

---***---

Abstract - *The road has been transformed into a complex building in design and management areas due to the increase in the number of vehicles. This situation has identified the problem of road accidents, contributed to public health and the country's economy, and conducted studies on the solution to this problem. Big data integration has been expanded due to technological advances and data retention at a lower cost. The emergence of the need for data retrieval from this large data scale has found a cornerstone of the data mine. In this study, the allocation of the most relevant machine classification strategy for road accident measurement by data mining is intended.*

1. INTRODUCTION

Road accidents have been a major cause for concern across the Indian subcontinent. In 2019 alone, the country reported over 151 thousand fatalities due to road accidents. Each year, about three to five percent of the country's GDP was invested in road accidents. Notably, while India has about one percent of the world's vehicle population, it also accounted for about six percent of the global road traffic incidents. Almost 70 percent of the accidents involved young Indians.. Studies on traffic have executed that road accident and the death- laceration ratio will increase. Design and control of traffic by advanced systems come in view as the important need. The assumption on the risks in traffic and the regulations and interventions at the end of these assumptions will reduce road accidents. An assumption system that will be prepared with available data and new risks will be advantageous.

2. Literature Survey

2.1 Vision-based traffic accident detection using sparse spatio-temporal features and weighted extreme learning machine

Author: Yuanlong Yu1, Miaoxing Xu1,Jason Gu

Publication: The Institution of Engineering and Technology 2019 (www.ietdl.org)

Abstract: Vision-based traffic accident detection is one of the challenging tasks in intelligent transportation systems due to the multi-modalities of traffic accidents. The first challenging issue is about how to learn robust and discriminative spatio-temporal feature representations. Since few training samples of traffic accidents can be collected, sparse coding techniques can be used for small data case. However, most sparse coding algorithms which use norm regularisation may not achieve enough sparsity. The second challenging issue is about the sample imbalance between traffic accidents and normal traffic such that detector would like to favour normal traffic. This study proposes a traffic accident detection method, including a self-tuning iterative hard thresholding (ST-IHT) algorithm for learning sparse spatio-temporal features and a weighted extreme learning machine (W-ELM) for detection. The ST- IHT algorithm can improve the sparsity of encoded features by solving an norm regularisation. The W-ELM can put more focus on traffic accident samples. Meanwhile, a two-point search strategy is proposed to adaptively find a candidate value of Lipschitz coefficients to improve the tuning precision. Experimental results in our collected dataset have shown that this proposed traffic accident detection algorithm outperforms other state-of-the-art methods in terms of the feature's sparsity and detection performance.

2.2 Design and implementation of an eye blinking detector system for automobile accident prevention

Author:Tariq Jamil ,Iftikhar Uddin Mohammed , Medhat H. Awadalla.

Publication:SoutheastCon 2016

Abstract:Ever increasing number of fatal traffic accidents around the world can be significantly reduced if modern technology is incorporated within the automobile to assess the physical condition of the driver at regular intervals during the movement of the vehicle and preventive measures are automatically taken for the safety of all concerned entities, both within the vehicle and outside the vehicle. In this paper, design of an eye blinking detector system is presented which can monitor the physical state of the driver at regular intervals during his/her driving and, if needed, can raise an audible alarm within the vehicle to alert the.

2.3 Principal Component Analysis of Fatal Traffic Accidents Based on Vehicle Condition Factors

Author:Tang Youming ,Zhong Deliang , Zha Xinyu

Publication:2018 11th International Conference on Intelligent Computation Technology and Automation (ICICTA)

Abstract:In order to find out several pivotal influencing factors of fatal traffic accidents, the number of fatal injuries recorded in the FARS database of the National Highway Traffic Safety Administration of the United States from 2010 to 2016 was calculated. The principal component analysis (PCA) method of multivariate statistical analysis is used to analyze the traffic conditions, and several pivotal influencing factors of fatal traffic accidents are obtained. The results show that tire wear, rim damage, exhaust system failure and coupling failure are the most important factors.

2.4 A Study on Road Accidents in Abu Dhabi Implementing a Vehicle Telematics System to Reduce Cost, Risk and Improve Safety

Author:Omar Kassem Khalil

Publication*:* 2017 10th International Conference on Devel- opments in eSystems Engineering (DeSE)

Abstract:Road accident study in the Emirate of Abu Dhabi is imperative as it will allow the government to improve its transport system and realize the vision of Abu Dhabi Department of Transport to contribute to quality of life, economic growth and environmental sustainability of the Emirate of Abu Dhabi. Despite sustainable and continuous reduction annually, road accidents still remain a serious phenomenon in the Emirate of Abu Dhabi. Road accidents killed 5,564 people in the UAE over the past 6 years, an average of more than two each day, with Abu Dhabi being the main victim [8]. A further examination of the severity index, Police records have shown that the deaths during the years 2006-2011 were a result of nearly 56,700 accidents, which also injured 63,406 people [8]. Research has shown that over 80% of road accidents are related to human factors [6]. Bad driving habits are dangerous and can lead to loss of life. This study will propose possible method of adopting Telematics System for Abu Dhabi Department of Transport to enhance its road safety programs. Telematics System can help reduce fatalities and injuries, thus improving road safety and efficiency of driving performance.

3. Software and Languages used

3.1 Python

Python is an interpreted, high-degree, and general-cause Programming language. Created by guido van Rossum and first released in 1991, python has a layout philosophy that emphasizes code readability, substantially using vast whitespace. It's far the most used programming language currently. It gives constructs that allow clear programming on each small and large scale. The apriori and K means algorithm inside the machine is carried out in Spyder and the algorithm is written in python language.

3.2 Python GUI - tkinter

Python offers multiple options for growing GUI (graphical person interface). Out of all the GUI strategies, Tkinter is the maximum normally used technique. It is a popular python interface to the Tk GUI toolkit shipped with python. Python with Tkinter is the quickest and easiest manner to create GUI programs.

4. Models

The working of the project is divided into the parts:

4.1 Dataset Collection

In machine learning, a dataset is simply a collection of data pieces that can be analyzed and predicted by a computer. Data collection should be uniform, and able to be understood by machines that don't see things the same way that humans do. Hence, it is crucial that after collecting data, it is cleaned, completed, and then annotated with meaningful tags readable by computers.

4.2 Data Pre Processing

Data pre processing is an information mining method that is used to convert the raw facts into a useful and efficient format. It is the essential step while developing a system learning model. While developing a system mastering mission, it isn't always constantly a case that we come across easy and formatted facts. And at the same time as doing any operation with statistics, it's far obligatory to ease it and put it in a formatted way. So for this, we use statistics pre processing task.

4.3 Dataset Splitting.

The cause is that after the dataset is break up into train and take a look at sets, there will not be sufficient statistics inside the schooling dataset for the model to analyze an powerful mapping of inputs to outputs. There will also not be enough information inside the check set to correctly compare the version overall performance. The split is finished via 80% as education set and 20% as test set.

Train set:

The educate set would incorporate the facts in an effort to be fed into the version. In easy terms, our version would analyze from these facts. As an instance, a regression model would use the examples in this record to discover gradients to be able to reduce the value characteristic. Then these gradients will be used to lessen the value and are expecting information efficaciously.

Test set:

The test set contains the facts on which we check the educated and verified model. It tells us how green our ordinary version is and the way likely is it going to predict something which does now not make us feel. There are a plethora of assessment metrics (like precision, do not forget, accuracy, and so forth.) that can be used to measure the performance of our model.

4.4 Data Visualization

Data visualization is the presentation of facts in a graphical layout. It facilitates people to understand the importance of information by way of summarizing and providing a large number of statistics in an easy and smooth-to-recognize layout and facilitates speak facts simply and correctly. To communicate statistics honestly and efficaciously, records visualization uses statistical pictures, plots, information pics, and different tools. Numerical facts can be encoded using dots, traces, or bars, to visually talk a quantitative message.

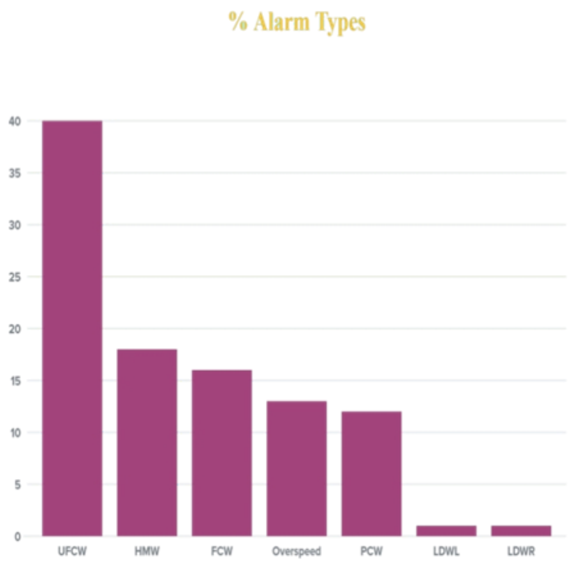


Fig:Alarm Types

- FORWARD COLLISION WARNINGS (FCW)

A FCW alerts drivers of an imminent rear-end collision with a car, truck, or motorcycle.

- URBAN FORWARD COLLISION WARNINGS (UFCW)

UFCW provides an alert before a possible low-speed collision with the vehicle in front, thus assisting the driver at a low speed in densely heavy traffic. This is usually applicable when driving under approx 30 kmph.

- HEADWAY MONITORING WARNING (HMW)

The headway monitoring warning (HMW) helps drivers maintain a safe following distance from the vehicle ahead of them by providing visual and audible alerts if the distance becomes unsafe. Active above 30 kmph, this sensor generates alarm and displays the amount of time, in seconds, to the vehicle in front when that time becomes 2.5 seconds or less.

- LANE DEPARTURE WARNINGS (LDW)

The LDW provides an alert when the vehicle unintentionally departs from the driving lane without using the turn signals. If the turn signals are used when changing lanes, an alert is not generated. Usually active above 55 kmph, LDW might not work well if lanes are unmarked or poorly marked.

This is further classified into: (a) LDWL, for lane departures towards left lane and (b) LDWR, for the same towards right lane.

- PEDESTRIANS AND CYCLIST DETECTION AND COLLISION WARNING (PCW).

The PCW notifies the driver of a pedestrian or cyclist in the danger zone and alerts drivers of an imminent collision with a pedestrian or cyclist. PCW works well when vehicle is below 50 kmph.

- OVERSPEEDING

Detects and classifies various visible speed limit signs and provides visual indication when the vehicle's speed exceeds the posted speed limit.

4.5 Model building

Since the data is unlabeled, we opt for an unsupervised learning approach using clustering techniques. We have tried different clustering algorithms but we present the ones which gave us the most suitable labels as per our intuition. While the different variation of K-Means exists, we used the Expectation-Maximization method. In order to get better convergence of our data points to the centroid, we also increased the number of iterations from a default of 300 to 2000.

5. Working System

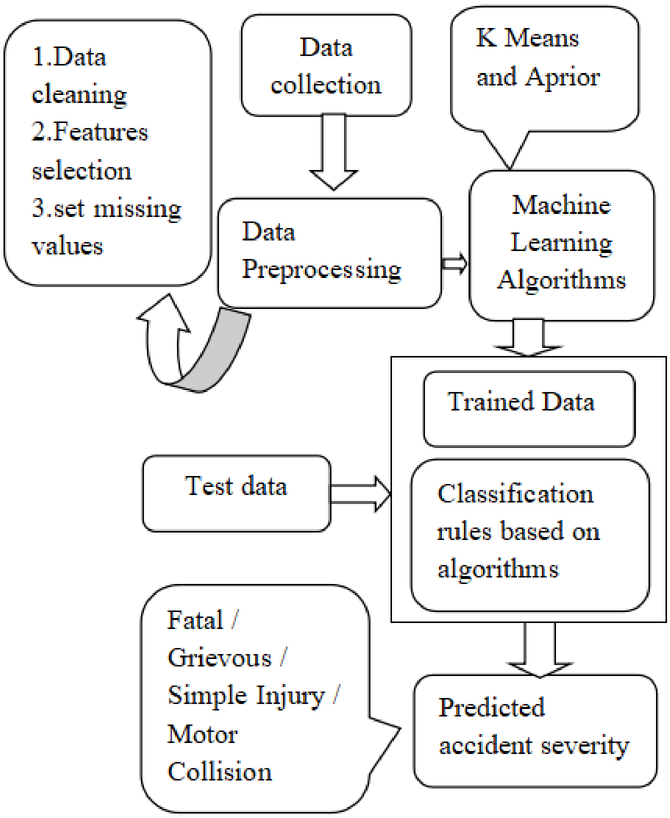


Fig 1:System Architecture

Models are created using accident data records which can help to understand the characteristics of many features like drivers' behavior, roadway conditions, light condition, weather conditions and so on. This can help the users to compute the safety measures which is useful to avoid accidents. It can be illustrated how statistical method based on directed graphs, by comparing two scenarios based on out-of-sample forecasts. the model is performed to identify statistically significant factors which can be able to predict the probabilities of crashes and injury that can be used to perform a risk factor and reduce it. The datasets of Bangalore and Ahmedabad contain information such as coordinates, speed, and the location name besides the vehicular movements as observed. Given that anyone on this planet uses roads, it comes with intuition that additional information could be beneficial to build a robust prediction model. To do so, we aim to expand our dataset by adding more attributes such as weather conditions, density of potholes and the infamous accident hotspots of the city. Due to the absence of labels, we propose unsupervised learning techniques such as clustering to map different data points to relevant clusters. The clusters correspond to the likelihood of an accident taking place after profiling the drivers with the relevant available features. We propose to use different machine learning algorithms on the datasets and thus come up with the best performing ones to therefore predict accident hotspots.

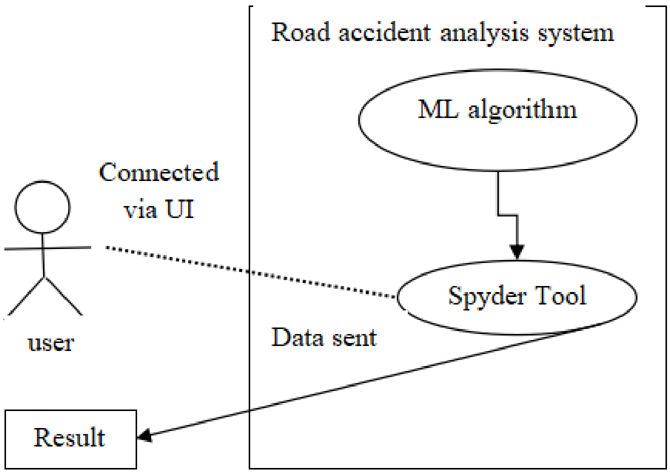


Fig 2:Use case

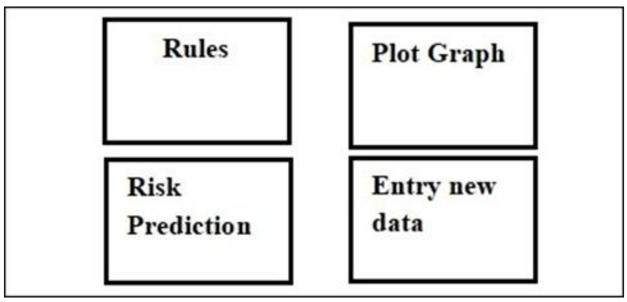


Fig3:Road Accident Prediction System

6. CONCLUSION

Road Accidents area unit caused by varied factors. By researching all the analysis papers it is ended that the Road Accident cases area unit vastly plagued by the factors like sort of vehicles, the age of the motive force, the age of the vehicle, the weather, the road structure than on. therefore we've to build an associate application which supplies economical predictions of road accidents supported by the higher than mentioned factors.

7. References

[1].An IOT Based Smart System for Accident Prevention and Detection (ICCUBEA) 978-1-5386-5257-2/18

[2].Towards Deep Learning based Traffic Accident Analysis 978- 1-7281-3783-4/20/

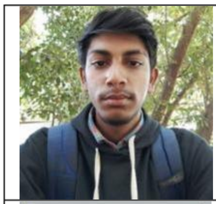
[3].MACHINE LEARNING BASED ANALYSIS FOR ROAD ACCIDENT PREDICTION , February 2020 (ISSN: 2394 – 6598)

[4].Performance Comparison of Machine Learning and Deep Learning While Classifying Driver's Cognitive State SBN No:978- 1-5386-8492-4/18

[5].Survey on Analysis and Prediction of Road Traffic Accident Severity Levels using Data Mining Techniques E-ISSN 2277 – 4106, P-ISSN 2347 – 516

BIOGRAPHIES

	ASGHAR PASHA B.E,M.TECH DEPARTMENT OF ISE,HKBKCE Bangalore 20 years Teaching Experience
	VIJAYALAKSHMI SA B.E,MTECH DEPARTMENT OF ISE,HKBKCE Bangalore 12 years Teaching Experience
	BIPIN LAMICHHANE 8th SEM ISE,1HK17IS015 HKBKCE,Bangalore
	HARSH NARNOT 8th SEM ISE,1HK17IS026 HKBKCE,Bangalore

	MD ATIQUE BOKHARI V 8th SEM ISE,1HK17IS033 HKBKCE,Bangalore
	MOHAMMED HUSSAIN 8th SEM ISE,1HK17IS040 HKBKCE,Bangalore