Prediction of accident prone-zone and alert messages to the hospitals using machine learning and internet of things

Accepted 22nd May 2020

ABSTRACT

There are many inventions in automobile industries to design and build safety measures for automobiles, but traffic accidents are unavoidable. There are a huge number of accidents prevailing in all urban and rural areas. Patterns involved with different circumstances can be detected by developing an accurate prediction model that will be capable of automatic separation of various accidental scenarios. This cluster will be useful to prevent accidents and develop safety measures. In this study, we predicted the accident zone and intimate the driver before crossing it. If in case the accident has occurred, the driver's latitude and longitude will be sent to the close-by hospital. These can be achieved using GNB classification and Internet of Things.

Key words: Machine learning, prediction, internet of things, GNB.

INTRODUCTION

Traffic accidents have huge impact on society, as they pose great cost of fatalities and injuries. In recent years, there is an increase in attention to determine the significant effect of the severity of driver’s injuries caused by road accidents. Road accidents in India claimed over 0.15 million lives in the country in 2018, with over-speeding of vehicles being the biggest reason for casualties, as reported by the government. The ministry of road transport and highways on Tuesday issued a report on Road accidents in India in 2018, which showed that road accidents last year increased by 0.46% as compared with 2017. “A total of 4,67,044 road accidents have been reported by States and Union Territories (UTs) in the calendar year 2018 (https://sites.ndtv.com/roadsafety/important-feature-to-you-in-your-car-5/). Accurate and comprehensive accident records are the basis of accident analysis. The effective use of accident records depends on some factors, such as the accuracy of the data, record retention, and data analysis. There are many approaches applied to this scenario to study this problem. In this system, the zone is predicted and shown to the driver before some distance with a caution message. If he/she is met with an accident, an alert will be sent to the driver and if he/she did not reply the message within threshold time, the driver's location will be sent to the close-by hospital and the hospital would prepare the first aid for the patient.

By the above facts, this research aims to introduce a hybrid model to provide service for the people who met with an accident. The Recommendation is performed by the mixture of machine learning and Internet of Things (IoT). The rest of the study is organized as follows: Literature on road accident rescue system; the proposed model for improving the performance of the proposed system is presented; examination of the proposed model and the results of the proposed model; and finally, conclusion based on the results of our model.

LITERATURE REVIEW

Here, we present related studies on accident rescue, including machine learning techniques and IoT used with different approaches. Kim et al. (2018) found that the speed of the vehicle or the number of vehicles has a greater effect on the road using the sensor big data measured from the atmospheric measurement vehicle and the local road traffic...
data. Chen Xu et al. (2018) classified the vehicular data into location, and then presented a real-world big data application in social-based vehicular networks according to the data characteristics. Lin, (2017) proposed a Density-Based Spatial Clustering of Applications with Noise (GDBSCAN) to analyze the distribution of taxi demand in the city with the support of GPS data, so that the demand can be found. A social-based localization algorithm (SBL) with the help of overlapping and hierarchical social clustering model (OHSC) uses location prediction to aid global positioning in-vehicle networks by analyzing the potential social relationships between vehicles (Tseng et al., 2018).

A SVM-based real-time highway traffic congestion prediction (SRHTCP) model uses the spout and bolt components in Apache Storm to predict road speeds for the next period by exploring streaming traffic and weather data (Chen and Chen, 2016). Hamzah et al. (2016) developed the intelligent network recommendation system; the study recommends that vehicle accesses the network using its analytic framework which analyzes the VANET-Big Data in real-time. A safe trajectory selection method is used for autonomous vehicles on cloud-based Internet vehicles, using VANET-Big Data mining and analyzing real-life accident data (Hamzah et al., 2016). Cao et al. (2016) analyze large-scale real-time accident data obtained from the Florida Department of Transportation (FDOT)-District 4 and predicts traffic congestion by continuously updating the estimated time of arrival (ETA).

A traffic control and guidance method is based on large data traffic prediction with the real-time routing algorithm of vehicle dynamic network. These works demonstrate road safety data collection and Real-time analysis is feasible (Taghi et al., 2010). However, most of these methods focus on analyzing the causes of traffic. Accidents from existing traffic data fail to obtain accident Prediction models based on original and multi-dimensional traffic data with universal application value. The accuracy of these accident prediction models is generally lower and their solution cannot adapt to new computing architectures (Taghi et al., 2010). Therefore, this study proposes the vehicle accident risk prediction model based on Gaussian Naïve Bayesian algorithm to reduce the accident count in the future. At present, in numerous applications, smart phones are used as moving traffic probes and sensors (Tang et al., 2005; Rose, 2006; Prashanth et al., 2008), for investigating the road environment, accident detection, and traffic congestion. These are the great features of an Intelligent Transport System (ITS) that looks for reducing traffic congestion and improving traffic safety (Whitney and Pisano, 1995).

Whitney and Pisano, (1995) proposed an automatic accident detection method. But there was a chance of wrong alarming and also could not give any assurance of accident occurrence. Extending the earlier study, Syedul et al. (2012) proposed a GPS, GSM and GPRS Technology-based accident detection system that can also report to the system. But, that involves a huge cost. Saiprasert and Pattara-Atikom (2013) advocated a system, where in case of unsafe driving, it could report to the system. Moreover, tracking the (Lee, 2011) developed a system that can detect an accident on the highway. The system, pressure can be actual fear, sends the information to the police station. But their pressure sensor was not able to give a precise reading.

Detection from Arm Processor system came into existence and used ARM7 (LPC 2129) processor and also GPS, GSM, RF transmission modules attached to ambulance (Venkatesh et al., 2015). It is based on an approach to make way for an ambulance using the Internet of things (IoT) based intelligent traffic light control system where the ambulance driver accesses the device through GPS. The driver traces the location and sends the updates to traffic control management, and GPS is connected to the ARM processor. It uses a GSM module to send a message to traffic management and get a message from the receiver side. The communication takes place with high security as information is encrypted.

The management side uses to pass data and control to a personal computer (PC). PC controls the traffic light (Singh, 2012) that focuses on the running emergency vehicles and provides a clear path. But there involved a lot of complexity from the perspective of the driver. FPGA Technology-based rescue system (Bhagya and Savitha, 2014) was developed to detect the position of any transportation and send an automated message to the number. This device could notify the vehicle owner to remove traffic congestion. FPGA controls and coordinates all parts used in the system. With the help of the sensor, the system could tell the exact position of the vehicle. This system is not cost-effective also. After an extensive literature review, it can be concluded that most of the systems used multiple sensors for detecting accidents which increases the cost of the project. But in this proposed model, only one vibration sensor was used which reduces the cost of multiple sensors and complexities of interfacing exist to develop open-source software. It is used for open-standards and services for interactive computation across dozens of programming languages. It is a web application that allows you to create and share documents, and code live is its major advantage. It can be used for data cleaning and transformation, numerical simulation, statistical modeling, machine learning and much more. We used run model,

(i). to predict the accident prone-zone and to intimate the driver before crossing it.
(ii). In case the accident has occurred, the driver's latitude and longitude will be sent to the close-by hospital using machine learning and IoT.

PROPOSED METHODOLOGY

In this section, we present our model for sending messages to the close-by medical centre when people meet with an
accident. This model uses machine learning and Internet of Things techniques. Figure 1 shows the different stages of the proposed model. (i) Preprocessing, (ii) Model Creation (iii) Prediction (iv) Recommendation using IoT.

**Preprocessing**

Raw data requires, cleaning, integration and filling of missing values in it. The missing values are taken as zero and label encoding scheme is applied for all the attributes to make them as same data type. Label encoder is a technique to transform features values into numerical values.

**Model creation**

In supervised learning algorithm, there are many models used for classification of sample in order to predict the label. Gaussian Naïve Bayesian algorithm is familiar one based on the bayes conditional probability. It has many features like easy to build, understand, faster prediction and scalability. GNB model predicts the class label based on the Gaussian probability density function. The Algorithm 1 depicts the flow of our proposed model.

**Algorithm 1: Recommendation Algorithm**

*Step 1: Start the Process*
*Step 2: Get the input dataset*
*Step 3: Perform the Label Encoding*
*Step 4: Model Creation*

*Step 5: Predict the path and check, whether it is safe or not*

*Step 6: If Accident occurred*

(i) Set threshold time $T_i$
(ii) If the Response time $(R_t)> T_t$ then send the message to the nearest hospital

*Step 7: Stop the Process*

**Prediction**

Prediction is the process of forecasting the dependent variable using the independent variables. There are different prediction measures like MAE, RMSE, and MSE. MAE is a familiar one. It is defined as the difference between the prediction of a rating of user $u$ on item $i(p_{u,i})$ and the real rating of user $u$ on item $i(r_{u,i})$. The calculation of MAE is shown in Equation 1.

$$MAE = \frac{1}{N} \sum_{u,i} |p_{u,i} - r_{u,i}|$$

**Recommendation using IoT**

The information to the close-by hospital is developed by integrating IoT module with the machine learning algorithm. The various components needed for this integration is shown in Figure 2. This IoT module has, GSM, GPS, Arduino Nano, Arduino Nano Shield, Vibration Sensor and Buzzer.
GSM

GSM (Global System for Mobile communication) is a cellular technology used for transmitting mobile voice and data services. It has been a particularly successful cellular phone technology for a variety of reasons including the ability to roam worldwide with the certainty of being able to operate on GSM networks. One of the main advantages of the GSM standard is the ability to roam and switch carriers by using individual mobile units even though the partner networks are located in their destination. It is an open and digital cellular technology used for transmitting mobile voice and data services operate at the 850MHz, 900MHz, 1800MHz and 1900 MHz frequency bands. GSM system was developed as a digital system using time division multiple access (TDMA) technique for communication purposes.

GPS

GPS stands for Global Positioning System by which anyone can always obtain the position information anywhere in the world. GPS receivers use a constellation of satellites and ground stations to compute position and time almost anywhere on earth. With this information and some mathematical background, a ground-based receiver or GPS module can calculate its position and time.

Arduino Nano

Arduino Nano is a small, complete, and breadboard-friendly board based on the ATmega328P (Arduino Nano 3. x). It has more or less than the same function of the Arduino Due/Arduino Uno but in a different package. It lacks only a DC power jack and works with a Mini-B USB cable instead of a standard one.

Arduino Nano shield

The Arduino Nano shield acts as an expansion board for the Arduino Due/Arduino Nano microcontrollers. There are several different options for power input. Also, the sensor shield adopts the Gravity interface that each pin out includes 5V and GND pins for easy connection to Gravity sensors or servos.

Vibration sensor

Vibration sensors are sensors for measuring, displaying, and analyzing linear velocity, displacement, and proximity, or acceleration. Therefore, vibration analysis is used as a tool to determine equipment condition as well as the specific location and type of problems.

Buzzer

A buzzer or beeper is an audio signaling device, which may be mechanical, electromechanical, or piezoelectric (piezo for short). Typical uses of buzzers and beepers include...
Table 1: Hardware components list.

<table>
<thead>
<tr>
<th>S/N</th>
<th>Name</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GSM Module</td>
<td>SIM 900A</td>
</tr>
<tr>
<td>2</td>
<td>GPS</td>
<td>NEO-6M</td>
</tr>
<tr>
<td>3</td>
<td>Arduino Nano</td>
<td>ATmega328P</td>
</tr>
<tr>
<td>4</td>
<td>Vibration Sensor</td>
<td>SW420</td>
</tr>
<tr>
<td>5</td>
<td>Buzzer</td>
<td>Mini piezo buzzer 12v</td>
</tr>
</tbody>
</table>

Table 2: Prediction Performance of Model.

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNB</td>
<td>0.1341</td>
</tr>
<tr>
<td>SVM</td>
<td>0.2859</td>
</tr>
</tbody>
</table>

alarm devices, timers, and confirmation of user input such as a mouse click or keystroke. The specification of all the above components used in our experiment is given in Table 1.

EXPERIMENT AND RESULTS

Our proposed model is implemented in windows platform with the use of python programming. The dataset used for the prediction of accident prone zone is "accident.csv" (https://www.kaggle.com/jrmistry/killed-or-seriously-injured-ksi). It consists of 56 features and 12558 instances and features which are depending on the driver and where he/she met with the accident, the latitude and longitude of that place and what type of vehicle he/she was driving. This dataset is explored and the basic information are gathered and listed in the following Figure 3. The dataset is divided into training and testing data set in the ratio of 80% and 20% respectively. The Gaussian Naïve Bayesian model is trained using the training set. The test dataset is applied into the training and the label are predicted. To measure the performance of the model, the Mean Absolute Error (MAE) is calculated. The same experiment is repeated several times and the result is shown in Table 2 for two different classifiers.

The next phase of our paper is to activate IoT to send the messages to the nearby medical centre. We are using Arduino nano and Arduino nano shield for our experiment. The vibration sensor detects the engine if there is any damage or abnormal change in the engine. Once the detected vibration reaches the threshold value, the signal is triggered, which means that the accident has happened. The buzzer starts to alarm for a certain period of threshold value (10 seconds). The driver is in the conscious state and if he is safe he has to press the switch which is kept near the steering within ten seconds. The accident has happened and the driver is not conscious then he can't able to press the switch so the buzzer rings for ten seconds and after that, it sends the message to the hospitals so that the ambulance service can easily track and rescue him from risk as soon as possible. We are receiving the message with
location and it will be transmitted to the close by medical centre. Our experiment showed that our proposed model clearly responding for the input given and sending the messages. The output of the message is show in Figure 4. The message received by hospital has, the message given by the driver and location of the accident occurred.

CONCLUSION

Though we have different road rescue systems, we still need an improvement to avoid road accidents. We have implemented a model with machine learning and IoT components to recommend the alert message to the vehicles. The model is constructed using Gaussian naive Bayesian classification and the trained model is evaluated by prediction measures. The message alert system to the hospital is implemented using IoT components. The results show that our proposed model is outperformed. This recommender system gives alert messages to the hospital based on our proposed algorithm.

REFERENCES

https://www.kaggle.com/jrmistry/killed-or-seriously-injured-ksi, last accessed on 10.01.2020