International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Prediction of Accident Prone-Zone and Alert Messages to the Hospitals Using Machine Learning and Internet of Things

Sathiya Devi S[1], Parthasarathy G[2], Mohamed Asif L[3], Ned Curtis Dylan P[4], Parasuraman B[5], Vinoth R[6]

[1]*Assistant Professor, Anna University, BIT Campus, Tiruchirappalli, India*
[2] *Research Scholar, Anna University, Chennai, India*
[3,4,5,6] *SRM TRP Engineering College, Tiruchirappalli, India*

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 project, we have 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 by using GNB classification and Internet of Things.*

Key Words: **Machine Learning, Prediction, Internet of Things, GNB**

1. INTRODUCTION

There is a huge impact on society due to traffic accidents where there is a great cost of fatalities and injuries. In recent years, there is an increase in the attention to determine the significantly affect the severity of the driver's injuries which is caused due to 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, a government report said. 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 to 2017. "A total of 4,67,044 road accidents have been reported by States and Union Territories (UTs) in the calendar year 2018[22]. Accurate and comprehensive accident records are the basis of accident analysis. The effective use of accident records depends on some factors, like 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 didn't reply for 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 paper is organized as follows: Section 2 illustrates the literature of the road accident rescue system. The proposed model for improving the performance of the proposed system is presented in section 3 and Section 4 examines the proposed model and the results of the proposed model. Finally, we conclude the results of our model in section 5.

2. LITERATURE REVIEW

In this section, we present related works on the accident rescue, including machine learning techniques and IoT used with different approaches.

C.J. Kim et. al [1] found 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 [2] classified the vehicular data into location, and then present a real-world big data application in social-based vehicular networks according to the data characteristics.

K. Lin [4] proposed a Density-Based Spatial Clustering of Applications with Noise (GD-DBSCAN) 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) that uses location prediction to aid global positioning in-vehicle networks by analyzing the potential social relationships between vehicles[5].

A SVM-based real-time highway traffic congestion prediction (SRHTCP) model and uses the spout and bolt components in Apache Storm to predict road speeds for the next period by exploring streaming traffic and weather data[6]. Hamzah Al et. al [7] develops the intelligent network recommendation system, that paper recommends vehicles access the network by using its analytic framework which analyzing VANET-Big Data in real-time.

A safe trajectory selection method for autonomous vehicles on cloud-based Internet vehicles, using VANET-Big Data mining and analyzing real-life accident data[8]. Cao Xin et. al [9] analyzes 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 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 [10]. However, most of these methods focus on analyzing the causes of traffic. Accidents from existing traffic data, failing 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 [10]. Therefore, this paper proposes the vehicle accident risk prediction model based on Gaussian Naïve Bayesian algorithm to reduce the accident count in the future.

Nowadays, in numerous applications, smart phones are used as moving traffic probes and sensors [11-13], 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 [14].

Whitney and Pisano [15] proposed an automatic accident detection method. But there was a chance of wrong alarming and also couldn't give any assurance of accident occurrence. Extending the earlier study, M. Syedul Amin et al. [16] 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 et al. [17] advocated a system using where in case of unsafe driving it could report to the system. Moreover tracking the [18] developed a system that can detect an accident on the highway. The system, pressure and that can be actual fear and send 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 that used ARM7 (LPC 2129) processor and also GPS, GSM, RF transmission modules attached to ambulance [19]. 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 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 [20] 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 [21] was developed that detects a 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 computing across dozens of programming languages. It is a web application that allows you to create and share documents and code live is a very big advantage of it. It can be used for data cleaning and transformation, numerical simulation, statistical modeling, machine learning and much more. We used to run model,

-To predict the accident prone-zone and to intimate the driver before crossing it.

-If incase the accident has occurred, the driver's latitude and longitude will be sent to the close-by hospital using machine learning and IoT.

3. PROPOSED METHODOLOGY

In this section, we present our model for sending messages to the close-by medical centre when people met 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.

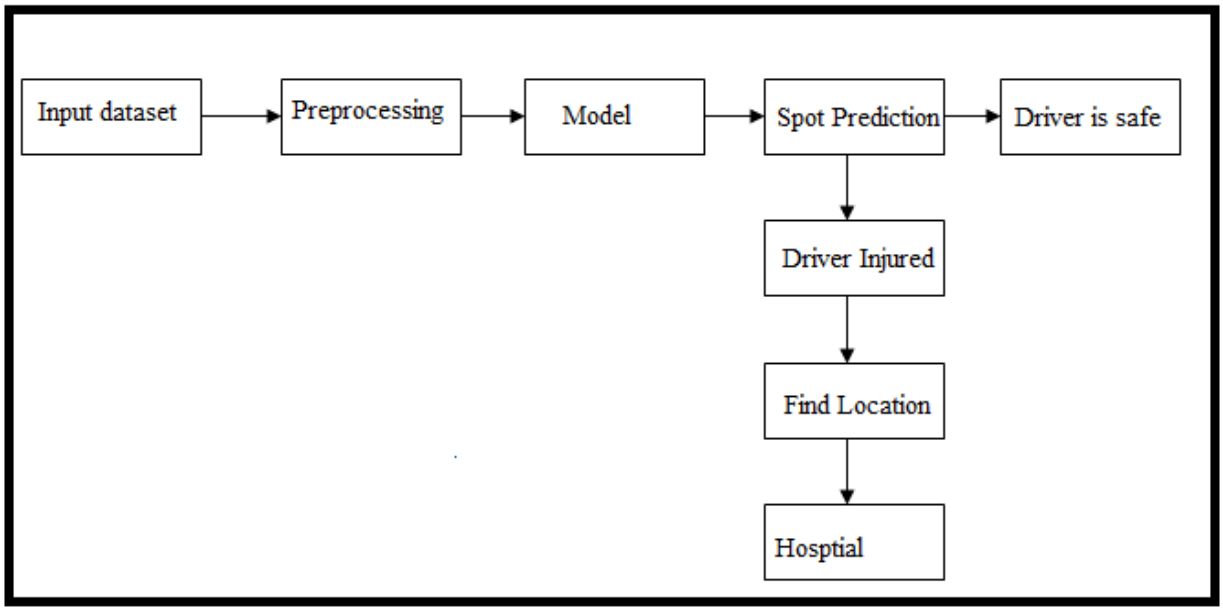


Fig. 1 Proposed Model

3.1 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.

3.2 Model Creation

In supervised learning algorithm, there many models used for classification of sample in order to predict the label. Gaussian Naïve Bayesian algorithm is familiar one based on the bays 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_t
		(ii)If the Response time (R_t)> T_t, then send the message to the nearest hospital
Step 7	:	*Stop the Process*

3.3 Prediction

Prediction is the process of forecasting the dependent variable using the independent variables. There are different prediction measures like MAE, RMSE, and MSE and 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.

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

3.4 Recommendation using IoT

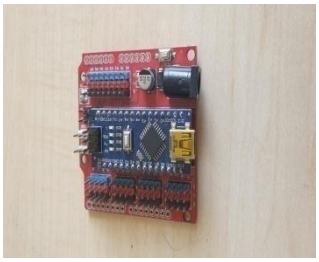
The information to the close-by hospital is developed by integrating IoT module with the machine learning algorithm. The various components needed for this integration is shown in Fig. 2. This IoT module has, GSM, GPS, Arduino Nano, Arduino Nano Shield, Vibration Sensor and Buzzer.

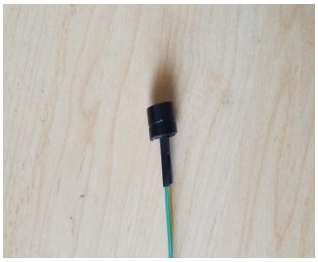
3.4.1 GSM

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 1900MHz frequency bands. GSM system was developed as a digital system using time division multiple access (TDMA) technique for communication purposes.

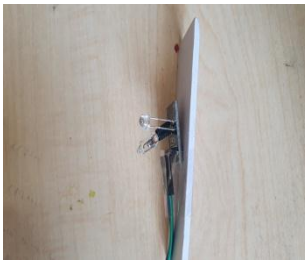
3.4.2 GPS

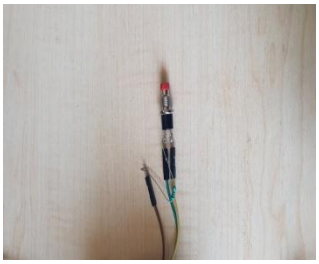
GPS stands for Global Positioning System by which anyone can always obtain the position information anywhere in the world. GPS receivers use a constellation of satellites and ground stations to compute position and time almost anywhere on earth. With this information and some mathematical background,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


Arduino Nano


Buzzer


GSM


Vibrator Sensor


Push Down Button

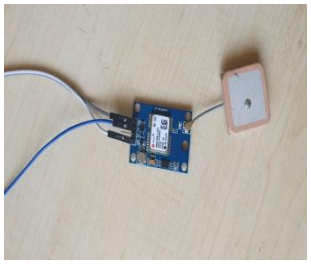

GPS

Fig.2 IoT Components

3.4.3 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 Duemilanove 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.

3.4.4 Arduino Nano Shield

The Arduino Nano shield acts as an expansion board for the Arduino nano and DFRduino 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.

3.4.5 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.

3.4.6 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 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.

Table 1. Hardware components List

Sl.No.	Name	Specification
1	GSM Module	SIM 900A
2	GPS	NEO-6M
3	Arduino Nano	ATmega328P
4	Vibration Sensor	SW420
5	Buzzer	Mini piezo buzzer 12v

4. EXPERIMENT AND RESULTS

Our proposed model is implemened in windows platform with the use of python programming. The dataset used for the prediction of accident prone zone is “accident.csv”[23]. 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.

	ACCNUM	YEAR	...	DISABILITY	FATAL
count	1.255700e+04	12557.000000	...	12557.000000	12557.000000
unique	NaN	NaN	...	NaN	NaN
top	NaN	NaN	...	NaN	NaN
freq	NaN	NaN	...	NaN	NaN
mean	1.576768e+09	2011.547822	...	0.028351	0.136657
std	2.541023e+09	3.104151	...	0.165979	0.343498
min	1.284070e+05	2007.000000	...	0.000000	0.000000
25%	1.108321e+06	2009.000000	...	0.000000	0.000000
50%	1.276036e+06	2011.000000	...	0.000000	0.000000
75%	4.002384e+09	2014.000000	...	0.000000	0.000000
max	7.003292e+09	2017.000000	...	1.000000	1.000000

Fig. 3. Basic information about dataset

Feature importance is a familiar technique to understand the need of feature in the decision. We have implemented feature importance technique and graph is drawn for all the features in the dataset, which is shown in Figure 4.

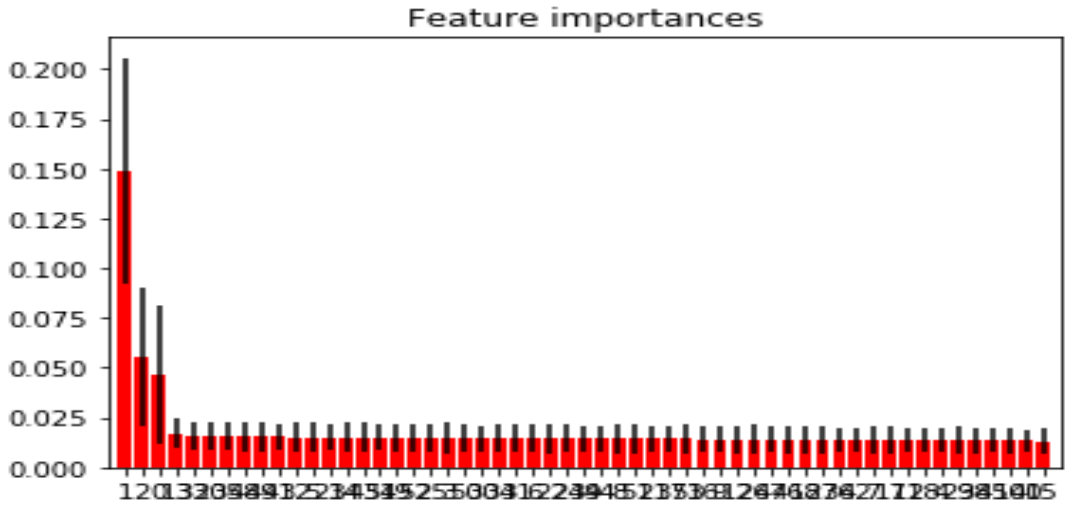


Fig. 4 Feature importance in dataset

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.

Table 2. Prediction Performance of Model

Model	MAE
GNB	0.1341

The next phase of our paper, to activate IoT to send the messages to the nearby medical center. We are using Arduino nano and 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, then the signal is triggered, which means that the accident is happened. The buzzer is started 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 is 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 n rescue him from risk as soon as possible. We are receiving the message with location, 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 5.

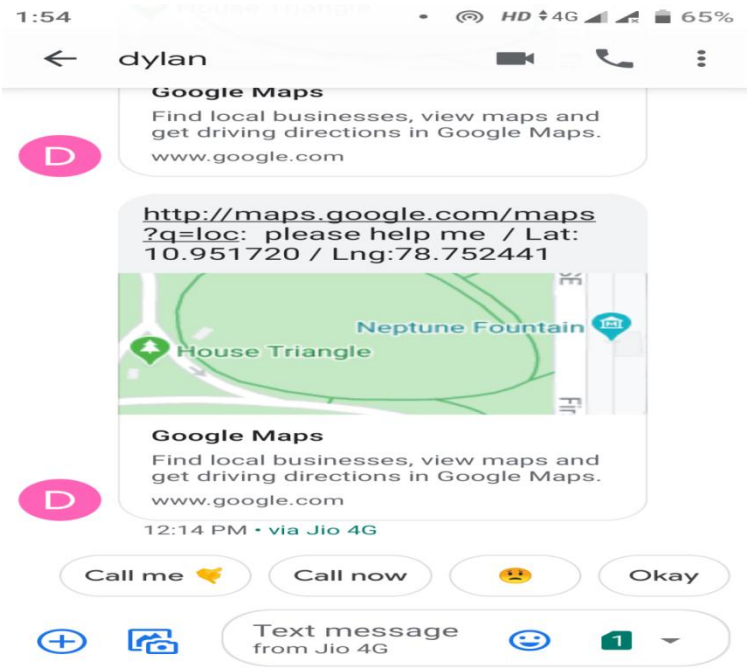


Fig. 5 Format of output message sent to the hospital

The message received by hospital has, the message given by the driver and location of the accident occurred.

5. 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 naïve 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

[1] C.J. Kim, S.H. Park,G. Ha, "Correlation analysis between vehicular and PM using sensor big data", in IEEE Int. Big Data Smart (BigComp), Shanghai, China, Jan. 2018, pp. 644_648.

[2] Chen Xu and Zhenyu Zhou. "Vehicular content delivery: A big data perspective", IEEE Wireless, vol. 25, no. 1, pp. 90_97, Feb. 2018.

[3] L. C. Chen, Y. Wang, "Exploiting taxi demand based on vehicular big data analytics", in IEEE 84th (VTC), Montreal, QC, Canada, Sep. 2016, pp. 1_5.

[4] K. Lin, "Localization based on social big data analysis in the vehicular networks", IEEE Trans. Ind. Inform., vol. 13, no. 4, pp. 1932_1940, Aug. 2017.

[5] T. Yang, H.-C. and L.-D. Chou, "Congestion prediction with big data for real-time highway" .IEEE Access, vol. 6, pp. 57311_57323, 2018.

[6] Y. X. Chen, C. Chen, " Big data analysis supporting vehicular network access recommendation," in IEEE Int. (ICC), Malaysia, May 2016, pp. 1_6.

[7] Hamzah Al Najada and Imad Mahgoub ,"Autonomous vehicles safe-optimal trajectory selection based on big data analysis and user preferences", in IEEE 7th Ubiquitous, Electron. Mobile (UEMCON), New York, NY, USA, Oct. 2016, pp. 1_6.

[8] Hamzah Al Najada and Imad Mahgoub , "Anticipation and alert system of congestion and accidents in VANET using big data analysis for intelligent transportation systems", in IEEE Series (SSCI), Athens, Greece, Dec. 2016, pp. 1_8.

[9] Cao Xin, Chu Na , and Bai Yeshuai, “Analysis on key technologies of prediction and path guidance in intelligent transportation,'' in Int. , Big Data Smart City (ICITBS), Changsha, China, Dec. 2016, pp. 5_8.

[10] Taghi M. Khoshgoftaar ,Jason Van Hulse and Amri Napolitano, “ A hybrid approach to alleviating class imbalance,'' IEEE Trans. , Man, A,, Humans, vol. 40, no. 1, pp. 185_197, Jan. 2010.

[11] Tang Shuming ,Gong Xiaoyan and Wang Feiyue , "Traffic-incident detection-algorithm based on regression," IEEE Transactions on Intelligent Transportation Systems, vol. 6, pp. 38-42, 2005.

[12] G. Rose, "Mobile phones as traffic probes: practices, prospects ," Transport Reviews,vol. 26, pp. 275-291, 2006.

[13] Prashanth Mohan, Venkata N. Padmanabhan and Ramachandran Ramjee “Nericell: Rich Monitoring of Road and Traffic Conditions”, in Proceedings of the 6th ACM conference on Embedded network sensor , 2008, pp. 323-336.

[14] D. A. Whitney , “AutoAlert: Automated Acoustic Detection of Incidents: IDEA Program”, Transportation Research Board,National Research Council., 1995.

[15] Jungkeun Yoon, Brian Noble, and Mingyan Liu , “Surface street traffic estimation”, in Proceedings of the 5th international conference on Mobile systems, and services, 2007, pp. 220-232.

[16] M. Syedul Amin, J. Jalil and M. B. I. Reaz, "Accident detection and reporting system using GPS, GPRS and GSM technology," 2012 International Conference on Informatics, Electronics & Vision (ICIEV), Dhaka, 2012, pp. 640-643.

[17] Saiprasert, Chalermpol & Pattara-atikom, Wasan. (2013). Smartphone Enabled Dangerous Driving Report System. 1231-1237. 10.1109/HICSS.2013.484.

[18] Lee, In. (2011). An accident detection system on highway using vehicle tracking trace. 2011 International Conference on ICT Convergence, ICTC 2011. 716-721. 10.1109/ICTC.2011.6082684.

[19] Venkatesh, H & Perur, Shrivatsa. (2015). An Approach to Make Way for Intelligent Ambulance Using IoT. 3. 2348-6988.

[20] H. Singh , "Intelligent traffic lights based on RFID," International Journal of Computing and Business Research,vol.2012, pp. 1-10, 2012.

[21] Mrs.Bhagya Lakshmi, .Savitha Hiremath, “FPGA Based Vehicle Tracking and Accident Warning using GPS”, International Journal of Scientific & Engineering Research, Volume 5, Issue 2, February-2014.

[22] https://sites.ndtv.com/roadsafety/important-feature-to-you-in-your-car-5/ last accessed on 01.12.2019

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