Big Data Analytics on Road Traffic Accidents in India: A Critical Statistical Review on Safety Measures and Awareness on Accidents

S.Bharathiraja[1], Vaishnavi.P[2]

[1]"Department of Computer Applications, Anna University BIT-campus, Trichy - India"

[2]"Assistant professor, Department of Computer Applications, Anna University BIT-campus, Trichy - India".

Corresponding Email Id: vaishmk@gmail.com

ABSTRACT

In the recent decades, the need for the two wheel vehicles, four wheel vehicles are increased due to travel long distance from a place to another for their stay, office work etc. In the due course the traffic increases leads to endanger to life threat. A critical data analysis is required for the traffic based accidents to find the cause for the accidents, identifying the casualties, remedial measures for the cause of incident. The solution is provide an data analytics based on road accidents is performed with the data available in the repository. In this statistical analysis, it focuses on four aspects on population status, road accident statistics, reason for road accidents, awareness on road accidents. A review report on safety measures and road traffic accident situation in India is taken as base line study for this research. The first part of report provides a comprehensive analysis of vehicle population in India. The second part of the report is pointed out the analysis based on the number of road accidents in the past few years. These statistics indicate the number of fatalities, number of person injured (Grevious hurt/minor injury) and the cause vehicle for road accidents in India. These are declined to continue for many years unless the new policies are implemented. The third part of the report gives a brief insights various reason for the road accidents in India. In future extensive study on the research is useful for safety measures and awareness to the people.

Keywords: Data Analytics, RStudio, RTA, Safety Measures.

1.INTRODUCTION

There are many vehicles driving on the roadway every day in India, and traffic accidents could happen at any time. Some accident involves fatality that causes severe cause of injury to the people in the accident. In the analysis phase it includes human population, vehicle population (motorcycle, car, lorry etc) and is calculated for total number of accident in the particular year, and the reason for the accidents. A comprehensive analysis finds out the number of person killed/year and injured for (Grevious/minor) in a road traffic accidents. In this attempt to analyse the road accident the statistical data from the year 1990 to 2016 is taken for analysis available as repository. Table1 represents the total vehicle population in India. Table2 Classify the number of accident in the year. The results are analysed using R tool for retrieving the accurate pattern of data on road accidents.

When studying road traffic fatal accident data, the analysis of data is critical to find out variables that are closely related to fatal accidents. we apply the statistical analysis and data mining algorithms for this study. Many attributes like weather, surface condition, light condition, alcohol consumption were investigated for fatal rate identification. In the recent studies the driving rules were made based on statistics, association rules, classification and clustering techniques.

A survey on detection of road safety aspects based on national highway of India were done and we had studied main sources of traffic crash data published by the National Crime Record Bureau (Ministry of Home Affairs) and the annual publication of the Ministry of Road Transport & Highways titled Road Accidents in India. The basic information for both reports come from all the police station in the country reported to them. The Global Burden of Disease (GBD) study estimates that there were 264,000 (95%CI: 214,000-321,000) deaths in India in 2013 almost twice the deaths reported by traffic police**.** If the safety measures 80% of road accidents are attributed to the human error of momentary judgment (primarily drivers and vulnerable road user.

A Complete statistical analysis of data for RTA for the year 2009 from case dairies and police records were studied. Road Traffic Accidents (RTAs) have turned out to be a huge global public health and development problem killing almost 1.2 million people a year and injuring or disabling between 20-50 million people more. The analysis is based on the factors on rush & negligence 95.38% of Total RTAs. Out of which 60% of the accidents were recorded during daytime (6 AM to 6 PM). The probability of most accidents are between 12 PM to 6 PM (38.46%). The highest numbers of accidents (32.30%) were observed in the heavy rainy season during the months of July – September [3]. Investigation of driver behaviour near accident black spots were studied. The big data analysis survey report using geo telematics data were used for this analysis. World wide Hundreds and thousands of people lose their lives in car accidents and road disasters every year. In Spain, nearly 177 million euro and about 2 million lost working days are estimated cost of yearly traffic accidents. The figure of 177 million doesn't take into account insurance and compensation, medical costs and health resources, legal defences, surveys, or decreased productivity in a company. The major causes of accidents in human failures are because of 80 to 90 percentage of distractions, speeding, alcohol consumption. The failure are because of seat belt maneuvers such as, harsh cornering, harsh acceleration, harsh breaking. The black spots are analysed using number of victims and number of predominant accident type.

Survey Analysis:

A brief study at all the cities based on the analysis of traffic safety situation in India.[3]. It identifies the area in which the total harm caused by crashes. In a country where 34 births happening every minute and over 10 deaths happening every minute, people really don't care to mourn for the dead. Of the various reasons for a person to die, road accidents are the major one. It is pointed out in analysis that fatality rate have increased during the past few years [5]. In India during the year 2015, a road accident survey reveals 1,374 accidents were happened and 400 deaths were reported. During the road accidents, it took place every day 57 accidents and loss of 17 lives on an average every hour in the country. These statistics indicate that number of fatalities in India is not likely to start to decline for many years to come unless new policies are implemented. Second part of the paper gives a brief insight on the various reasons for road accidents in India.[5]

According to the government of India, ministry of road transport & highways transport research wing. [4] During the year 2015, overloaded vehicles caused 77,116 accidents and 25,199 road accidental deaths. It constituted a share of 15.4 per cent and 17.2 percent respectively in total road accidents and fatalities in the country. If it is analysis for percentage of share in total number accident, number of person injured, number of killed, accident severity, vehicle for accident, causes of accidents, classification age person killed(male/female)etc. Moreover the data are represented as a percent of road traffic accident analysis and generate reports for road traffic accidents.

A complete research studies of road traffic injuries were done leading to cause of death under 30 years is high, causing over a million deaths every year [5]. A road traffic injuries are caused based on exposure to the different reasons and the appropriate for specific reasons are the avoiding Helmets, violation of traffic rules, excessive vehicle permission have an important role in death and injury.

METHOD:

The proposed research work is based on the real time data set for road traffic accident (RTA) available as repository. In this method, a data mining and analysis technique is adapted to analyse the data based on the critical/non critical state. Preprocessing measures are performed before analysis of data to remove the anomalies. The statistical analysis shows that the number of accidents in a selected city during the year and the causes of road traffic accident. Thus, we could able to find the causes of accident records based on the age and gender records, the type of vehicle and the accident severity in the city.

PROCEDURE:

Accident dataset: A road traffic accident Data are collected based on systematic approach to gather and measure the information from a variety of sources and repository to get a complete and accurate picture of an area of interest. The various attributes in the dataset includes total accident, injuries, vehicle etc.
Data pre-processing: In the RTA data set, a data mining technique is adapted that involves the raw data transformed to a specific format. In the dataset, the consistency of the variable are to be checked to retrieve the accurate results. In order to achieve the dataset are to be relieved from the factors based on data inconsistency, incomplete values, data missing, and is likely to contain many errors.
Attribute selection measures: A selection of an attribute in the road traffic accident dataset heuristic for selecting and splitting criteria that "best" separates a given year for data partition. **classification:** Classification in data mining methodology aims at constructing a model (classifier) from a training data set that can be used to classify records of unknown class labels if any in the existing accident dataset.
Visualisation of patterns: The data are after the classification, it is able to predict the pattern hidden in the data that helps the analyst to identify the significance of occurrence, causes of the incident and its remedial measures in a visual context. we, finally conclude the result of statistics for road traffic accident based on the factors of fatal accident, injuries, severity etc to make an analysis based on the generated report to give the awareness to the people.

RESULT ANALYSIS

The training dataset collected from various repositories for data analytics. The output can be achieved using the IDE (R-studio) software. The report has been generated by the graphical representation of Rstudio.

	row.names	Name.of.City	Total.number.of.Fatal.Accidents...2011	All.Accidents...2011	Persons.Killed...2011	Persons.Injured...2011	Sex.of.Casualty.2011
1	1	Agra	336	1299	632	1046	Male
2	2	Ahmedabad	222	2020	227	2036	Female
3	3	Asansol-Durgapur	229	368	230	185	Male
4	4	Aurangabad	161	719	173	744	Female
5	5	Bengaluru	689	6031	720	4952	Male
6	6	Bhopal	275	3459	299	2970	Male
7	7	Coimbatore	253	1126	259	1058	Female
8	8	Delhi	2007	7281	2065	7226	Female
9	9	Dhanbad	74	138	82	106	Female
10	10	Faridabad	229	803	249	595	Male
11	11	Ghaziabad	495	1109	535	817	Female
12	12	Gwalior	233	2030	248	1626	Male
13	13	Hyderabad	428	2651	441	2528	Female
14	14	Indore	407	4995	425	4053	Male
15	15	Jaipur	387	2002	406	1915	Male
16	16	Jabalpur	239	2906	260	2990	Female
17	17	Jamshedpur	189	447	207	292	Female
18	18	Jodhpur	236	526	264	525	Female
19	19	Kochi	171	1986	182	2008	Male
20	20	Kolkata	396	3133	418	2420	Male
21	21	Kollam	209	1668	222	1811	Female
22	22	Kota	102	616	113	688	Male
23	23	Kannur	52	557	55	833	Male
24	24	Lucknow	493	1270	517	741	Male
25	25	Ludhiana	260	444	294	189	Female

Fig1.0 : Road Traffic Accident data from the repository based on population

22	Kota	102	616	113	688	Male
23	Kannur	52	557	55	833	Male
24	Lucknow	493	1270	517	741	Male
25	Ludhiana	260	444	294	189	Female
26	Madurai	123	685	125	656	Male
27	Mallapuram	290	2694	324	3534	Male
28	Mumbai	539	25471	563	5059	Male
29	Meerut	321	886	384	679	Male
30	Nagpur	227	1231	236	1144	Male
31	Patna	507	1225	507	576	Male
32	Pune	62	261	64	213	Female
33	Rajkot	163	1039	172	1042	Male
34	Surat	240	1223	248	841	Male
35	Srinagar	NA	492	75	511	Female
36	Thrissur	134	1262	141	1343	Male
37	Vadodra	164	1343	172	908	Male
38	Varanasi	73	142	73	69	Male
39	Visakhapatnam	410	1548	414	1286	Female
40	Allahabad	NA	NA	NA	NA	Female
41	Amritsar	NA	NA	NA	NA	Male
42	Chandigarh	NA	NA	NA	NA	Female
43	Chennai	NA	NA	NA	NA	Female
44	Kanpur	NA	NA	NA	NA	Female
45	Khozikode	NA	NA	NA	NA	Male
46	Nashik	NA	NA	NA	NA	Female
47	Raipur	NA	NA	NA	NA	Female
48	Thiruvanthapuram	NA	NA	NA	NA	Male

Fig 2.0 - Data preprocessing to extract the hidden patter based on No of cities and their accident

Measuring road traffic accidents based on age and gender for the specific year.

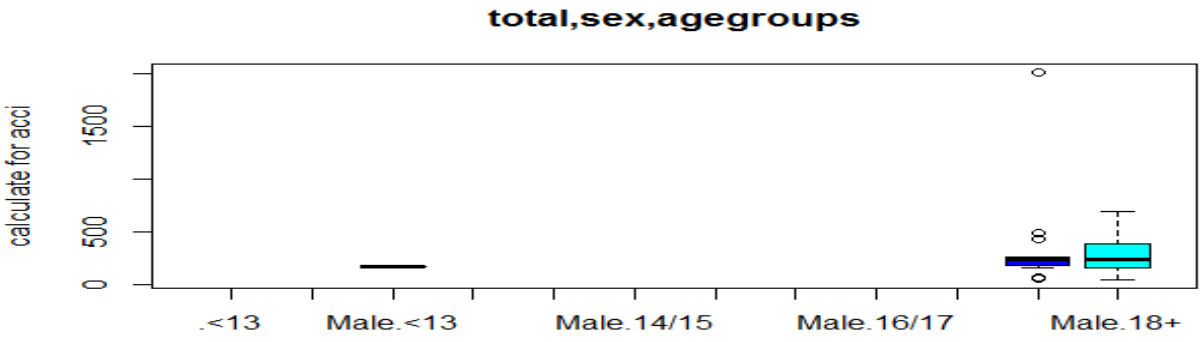


Fig 3.0- Accidents among the age groups based on the category (Male in the year(2011)

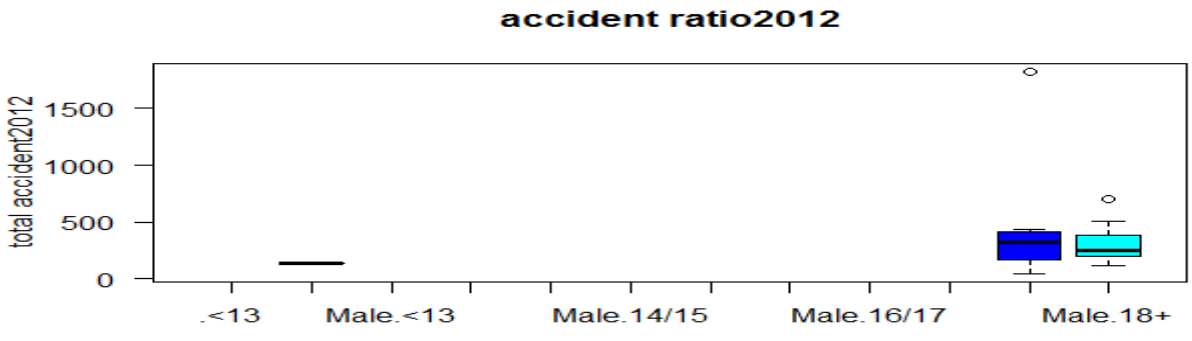
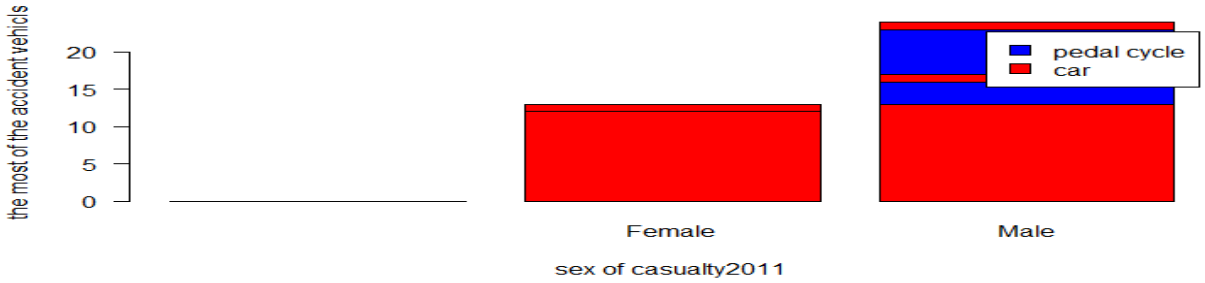


Fig 4.0- Accidents among the age groups based on the category (Male in the year (2012)

In the fig 3.0, & 4.0 shows the status of accident among the gender male/female and the results between the age groups of (<13 to >18+). It is observed that in the age group of(<18yrs) minimum no. of accidents are occurred. In the age group above (>18yrs) it is identified the highest no. of accidents. In the year 2012, there is increase in the no.of accidents at the age of(<18yrs). The reason is the increase of vehicles and the traffic over the years.

Fig 5.0 - Different types(car, pedal cycle) of vehicles and their causes of impact



Predicted accident based on various types of vehicles (Male/Female)

In the above fig5.0, shows the status of various types of vehicles among Male/female group. In the result, there is high impact on the fatal injuries while driving in car when compare to pedal cycle.

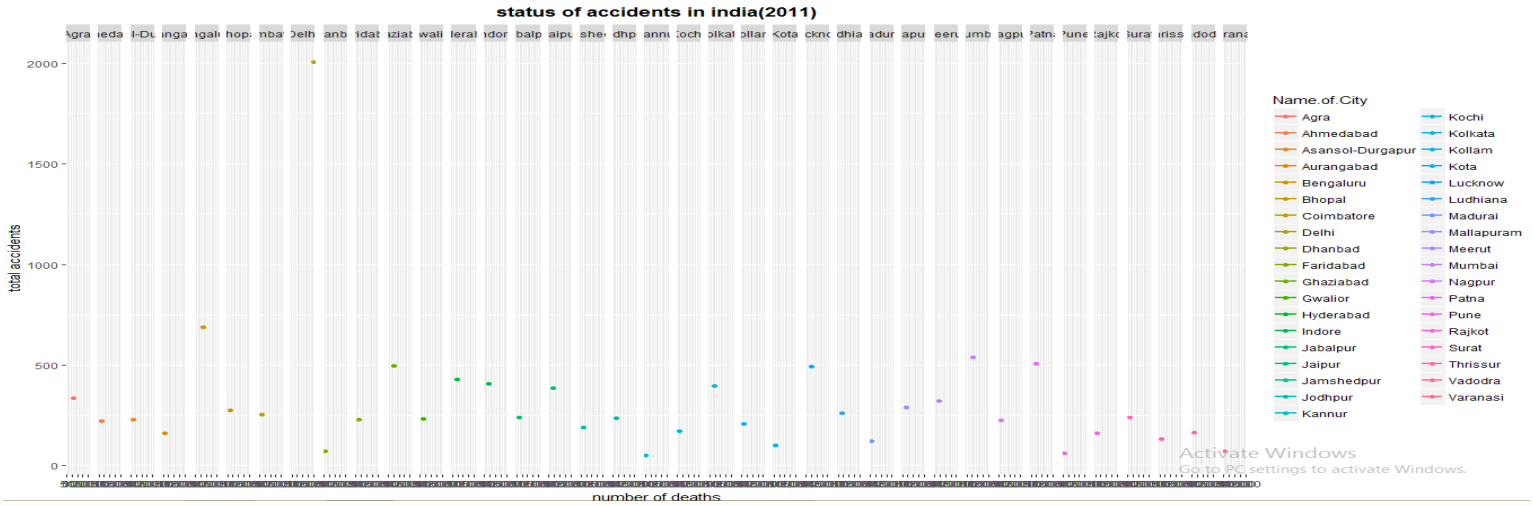


Fig6.0- Total accidents happened among various cities in India (2011)

Also the impact is very high for the male when compare to female. The blue collar shows the results of pedal cycle high for male in the year 2011.

In the result Fig 6.0, the highest accident deaths happened in Delhi (>2000) in the year 2011. Moderate number of accidents in Bangalore (>600). Minimum at the level of less than 100 in Dhanbad, Kannur, Pune.

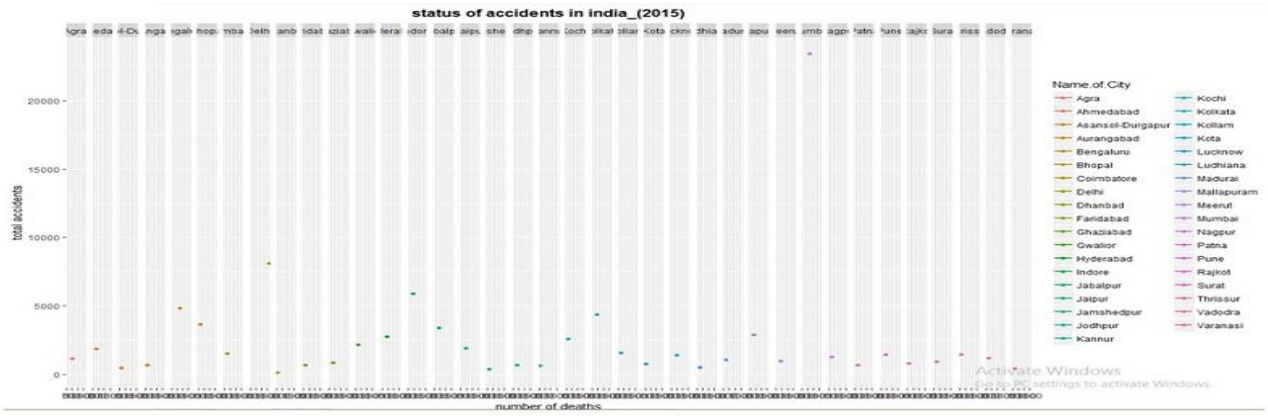


Fig 7.0 - Total accidents happened among various cities in India (2015)

In the Fig7.0, shows the results of accident deaths which is very high when comparing with year 2011. In the scenario of 2015, death of accident in Mumbai(>25,000). Moderate level in Delhi (>5000) and the minimum of deaths in Durgapur, Dhanbad, Jamshedpur, Jodhpur, Kannur, and Varanasi.

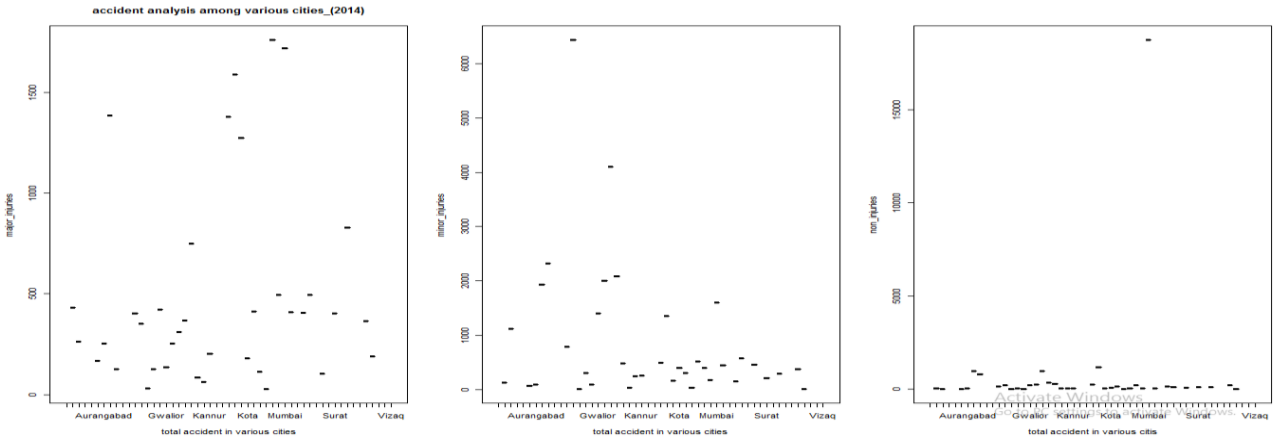


Fig 8.0- Major, Minor and non-injuries accident among various cities in India (2014-2015)

In the above Fig8.0, In the chart1, Kota and Mumbai cities have major injuries, In Chart 2, Gwalior and Kannur cities have next level of Minor injuries. In Chart 3, Mumbai and Surat cities have recorded of non- injuries between the year(2014-2015).

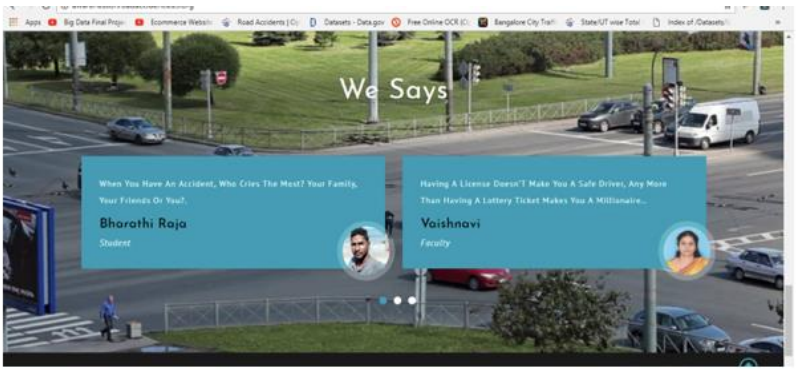


Figure 3: About Page

Fig 9.0- Website Developed based on the accident scenario and the prevention measures.

In the website the statistical information is available based on the Road Traffic Accident occurred between the year 2011 to 2016. The impact study was in detail with the shown predicted graphs. The remedial measures are provided to take the necessary action by the concern authorities to reduce the risk factors based on human error of momentary judgement. Remedial measures to avoid accidents strictly following the traffic rules, measures to be taken by the Govt. of India to increase the public transport, policy to be framed based on global warming.

CONCLUSION

Thus we have studied various statistical methods and tools used to predict road traffic accident in a dataset. The research analysis is based on the selected city in India based on the vehicle population, gender, age groups. In the statistical analysis, the results shows high in fatal injuries over of years due to increase in number of vehicles, avoiding the road rules and human judgement error during the incident. Globally major cause of these accidents are due to violating the traffic rules and the remedial measures are to follow the strict rules by the public, policies needs an modification, if the traffic rules is violated.

REFERENCE

[1].Www.safaribooksonline.com/library/view/getting-started-with/9781449314798/ch01.html
[2]. Ajit goswami* and ripunjoy sonowal division of epidemiology and nutrition, regional medical research centre, (indian council of medical research), n.e. region, dibrugarh, assam, india.
[3] A statistical analysis of road traffic accidents in dibrugarh city assam India published in researchgate.net
[4]. Jorge gonzalez, solution engineer oct20,2017.

[5].www.researchgate.net/profile/sridhar_seshadri2/project/safer-roads-safer-india
[6].Https://www.bestcurrentaffairs.com/report-road-accidents-india-released/
[7].geethabaidepartment of computer science and engineering bangalore.
[8].mariana antunes morgado, department of pediatrics, santa maria hospital, portugal, Tel: 00351918275875; e-mail: morgadoamariana@gmail.com.
[9].government of india ministry of road transport & highways Transport research wing.
[10].liling li, sharad shrestha, gongzhu hu department of computer science central Michigan university, usa (li8l, shres1s, hu1g)@cmich.edu.
[11].pankaj dahiya m.tech. Student, dept. Of civil engg., cbs group of institutions, Jhajjar,haryana.
[8].bhattacharya, s., alberini, a., & cropper, m. L. (2006). *The value of mortality risk Reductions in delhi, india* (publication., from university of maryland and world bank:
[9].george yannis, professorwww.nrso.ntua.gr/geyannis.
[10].Péter holló,habil, d. Sc. Head of dept. For road safety university professor and traffic Engineering széchenyi istváninstitute for transport sciences university,
[11]. Liling li, sharad shrestha, gongzhu hu department of computer science central Michigan university, usa (li8l, shres1s, hu1g)@cmich.edu.
[12]. Ramya v#1department of information science and engineering,dayananda sagar College of engineering bengaluru karnataka 560078. *1ramy.viju@gmail.com.*