International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

Predicting and Preventing Recommender System for Telangana Road Accidents

V. Kakulapati, Srikar Bharade, Nikhil. M

Abstract: ***Today people are suffering with road accidents in world wide. Analyzing these Road accidents are the major challenge in identifying and predicting primary features related with catastrophes. All these features are valuable for anticipatory computes to conquer road mishaps. Integrating various analytics techniques can get better model recognition and avoid road mishaps. As road safety growing quiet apprehension, speedy analytics observes all safety techniques in dynamic to spot malfunction that may signifies road mishaps on identifying key features related with road , mishaps in Telangana state. In our propose work, a framework to analyze the road mishap with classification of accidents and clustering, which analyze mishap data of Telangana stated district wise. The proposed framework describes the recommendation system for predicting road accidents. For this, classify the road accidents into fatal, major and minor. We implemented district wise data into clustering and applying enhanced k-mean algorithm. Further, implemented similarity measures to detecting the places where the severity of accidents happened and also analysing the driver behaviour analysis while accidents occur. The implementation result reveals that the road accident prediction exhibits enhance in certain areas and those areas exists in districts should be the major concern to acquire anticipatory measure to conquer the road mishaps.***

Keywords: traffic, accident, analysis, driver, behavior, factors

I. INTRODUCTION

In World population India is the second largest population. The major revenue is generating and play a key role in the Indian economy is transportation.

Today, the number of motor vehicles is increasing tremendously due to population, advertisement of vehicle loan policies, etc., which is also a major role for the road accidents in India According to surveys and reports there are almost 1.34 lakh victims due to road accidents per year. Predictive analysis of the road is a statistical technique which exhibits the relation between road mishap frequencies and travel circumstances, statistical features, ecological features and behavior analysis of driver.

Manuscript received on February 10, 2020.
Revised Manuscript received on February 20, 2020.
Manuscript published on March 30, 2020.
* Correspondence Author
V.Kakulapati*, Sreenidhi Institute of science and Technology, Yamnampet, Ghatkesar, Hyderabad, Telangana, India.
E-mail: vldms@yahoo.com
Srikar Bharade, Sreenidhi Institute of science and Technology, Yamnampet, Ghatkesar, Hyderabad, Telangana, India.
E-mail: srikar.bharade@gmail.com
Nikhil. M, Sreenidhi Institute of science and Technology, Yamnampet, Ghatkesar, Hyderabad, Telangana, India. E-mail: nikhil96m@gmail.com

Catastrophes are not as unmanageable, nor do they ignore regular learning. Learning of catastrophes based on "macro "level contains of the summary of quantities and rates not related to threat, which no meaning to the public, as well as drivers and passengers [9] .

The rapid and inadvertent urbanization method leads to the incomparable revolution in the scattering of vehicles worldwide. Past decade misery and the victim due to road traffic incidents (RTI) is increasing day by day.

Presently, road mishaps rank 9th consecutively of ailment burden and predicted to be graded 3rd in 2020. According to survey and reports, In India over 70,000 persons is eradicated due to road mishaps per year and this requires to be accepted as an essential public health apprehension.

The reason behind road accidents is inappropriate communications among the vehicles, another highway clients and highway features. This position leads to inappropriate communications may be as a consequence of the complicated inter play of several features for instance, footpath attributes, statistical factors, traffic attributes, highway clients behavior, design of vehicle , behavior analysis of driver and ecological aspects. Consequently, the road mishap is a multifaceted observable fact [10]. The feature of the data gathered in a scrupulous investigation has a hefty effect on the accident predicting methodology. For collecting data, care must be taken or else the consequences of a study will be less significance [11]. Several research works are carried out on road safety in diverse metropolitans, for example, Delhi, Mumbai, Chennai, and Ernakulum. Surveys estimate road accidents all over the world millions of persons die and injured per year [12, 13].

Our proposed system, the impact of road environment can be reduced if its influence is studied. Upgrading the whole road network would be time consuming and large financial resources are required, which may be tough to implement. The quick and cost effective step in enhancing road safety may be identifying accident prone locations and improve them instead of improving the complete road network. In this work we study the causes and consequences of road accidents.

II. RELATED WORK

The road infrastructure has not been developed along with the travel demand due to the deficit of resources. This imbalance is creating the problem. Road accidents are not solely occurring as a result of a single factor like negligence of the driver or deviating traffic laws, but also many other associated factors, for example, road condition, and vehicle condition, the impact of the driver's emotions, environment and combinations of other factors. Hence, the dynamic updating of data is needed which will be implemented in the proposed system.

New methodology to the application of data to enhance traffic control system performance is cyber-physical systems [1]. In this technique they address the function of the control instructions of the data stream, managing, behavioral effectiveness and detectable power and the renovation of the model of traffic control system. The road accident trend Analysis [2] intend to establish the track of development of the number of accident in Ibadan, they utilize time series analysis for forecast and decision making.

With their analysis, enhance in the number of road accident taking place in the prospect.

The research study in the field of road accident analysis in-depth is becoming more important. Identifying the reasons of accidents, GIDAS [3] is capable to bring sufficient data to accomplish an analysis of accidents, which lead to improvement and execution of a tool is known as the accident, provenance analysis method for the compilation of similar provenance data implementing the GIDAS method.

NIS data set of traffic accidents [4] is for the region of Belgium for 1991-2000 and Belgian data "Analysis Form for Traffic Accidents" that should be filled out by law enforcement for each traffic catastrophe that take place with wounded or death fatalities on a public road. The data set contains 340.184 traffic accident records as total.

Analysis of road catastrophe [5] plays a major role in categorizing major features related with the accident and all these assist in taking protective assessment to conquer road mishaps. Several research works has been developing on accident data analysis by utilizing conventional techniques like statistical and data mining methods. These works mainly retrieving key features associated with accidents. Accident is undecided and volatile occurrences occur in any conditions. An accident does not have comparable effects in each and every province of the district. Accidents may be increasing in one district and less impact in some other districts. Time-series models [6] road safety catastrophe forecast for a better analysis of the wide scope of the problem and ability to criticize approaches.

III. FRAMEWORK

A framework to investigate road catastrophe on driver emotions and provide recommendation system for avoid accidents. In this work, establish the relationship between the causes and the consequences with occurrence classification by prominence the active driving circumstances with preliminary speed, pre and post impact speed of vehicles to illustrate the crash scenario.

Driver behavior will help to develop a method providing that highly emotional agitated driver with the intention that can prevent the accident. Investigate the usability of clustering, and visualization methods to the road traffic accident analysis providing the better prediction accident prone areas.

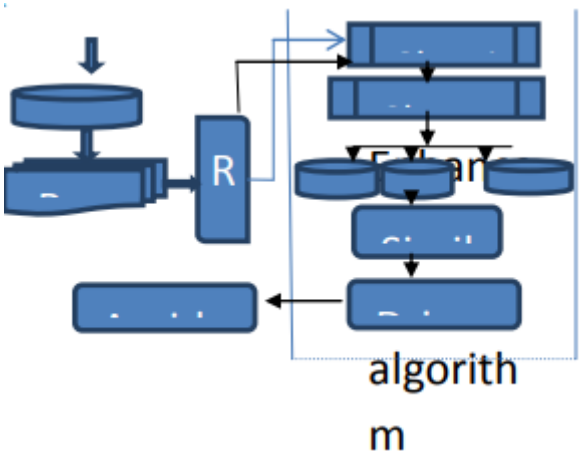

Fig. 1: Our proposed framework

Road accident database: we collected data from different districts of Telangana road accident data. Accident prediction analysis on area wise analyzed in a increased for analysis of trend of road mishap data, demonstrated in Fig. 1. The components of proposed architecture of road accidents as follows:

A .Pre-processing: collected data set is pre-processing for eliminating unnecessary data.

B. Recommender systems: Performance locator can be utilized in both manufacturing sector and in vehicular dynamic circumstances. Such as, in case of vehicular driving, the real-time attainment locator will discover the driver's performance analysis and offer a recommender system, which assist the drivers on vehicle functionalities to utilize while driving. The recommender system provides diverse real-world circumstances for usage of machine learning and video analytics to provide real-time assistance to users.

C. Classification: The road data accident data can be classified as four types: those are fatal, Grievous injuries, minor injuries and persons killed. Accident probability and severity are two conditions that have been set to define the classification factors of an accident: reduce the number of factors, defining the catastrophe, though preserving the valuable data regarding the severity of the accident. The road accident statistics is mainly depends on the permissible results of a catastrophe. Though the vehicle is prepare with an integrated safety system, which assists the driver about its present situation.

Clustering: It generally defining as categorizing of data items into one or more groups. The items in the same group have similar properties when compared with the other group. Diverse clustering algorithms are existing. For our implementation purpose we utilize the enhanced clustering algorithm, which is to partition the data into diverse groups [9]. K-modes grouping, an improved version of K means clustering.

D. Similarity Measures: The road catastrophes might have dissimilar impact for the different category of catastrophes at different places [7]. Furthermore, catastrophes are altering district wise and accident may occur in districts have comparable nature of catastrophes. There are a number of similarity measures [8], for instance, Euclidean distance. The distance between two vectors is calculated by using Euclidean distance similarity measure.

E. Driver behaviour analysis: It models analysis of driver behavior focus on different methods, provide data concerning about driving information. Predicting driver behavior prediction models give information regarding driving nature whether the driving is perfect or not.

IV. EXPERIMENT ANALYSIS

Number of road mishaps is increased 5.31 % from 2014 to 2015. Number of persons eradicates increased 17% from 2014 to 2015. Number of road mishap injuries has increased by 6.1% from 2014 to 2015. Accident severity has increased from 35% from 2014 to 38% in 2015. The road mishap analysis accident data 2015 make public that about 60 accidents and 22 deaths take place every day on Telangana roads. It further reveals that 3 mishaps happen and 1 life is gone per every hour in common on road mishaps in Telangana.

Table 1 : Road Accident Parameters: 2014 and 2015

Parameter	2014	2015	% change over previous year
Total Accidents in Telangana	20,078	21,145	5.31
Total number of Persons Killed in Telangana	6,906	8,047	16.52
Total number of Persons Injured in Telangana	21,636	22,956	6.10
Accident Severity*	34.40	38.05	10.6
* No. of persons killed per 100 accidents			

We collected district wise road accident data of Telangana State and after collecting data, preprocessing the data by removing noisy elements.
After preprocessing data looks like.

Table 2. Normalization of district statistics on various types of accidents

District	Fatal.accidents	Grevious.Injuries	Minor.Injuries	Non.Injuries	Total.accidents	Persons.Killed	Persons.Injured
1	-0.661	-0.51	-0.09	1.7	-0.2	0.7	-0.5
2	0.004	0.01	-0.22	-0.6	-0.2	0.1	-0.1
3	-1.034	-0.14	-0.79	-0.4	-0.9	-1.1	-1
4,	-0.868	-0.52	-0.93	-1.1	-1	-0.9	-1
5	-0.411	-0.52	0.04	-0.3	-0.2	-0.4	0.4
6	-0.147	-0.65	-0.25	0.6	-0.3	-0.2	-0.2
7	0.929	-0.59	0.05	-0.1	0.2	1.1	0.4
8	1.207	-0.66	-0.04	-0.3	0.2	1.2	0.1
9	0.648	1.59	-0.02	-0.4	0.4	0.7	1.1
10	-1.111	-0.86	-1.37	-1.6	-1.5	1.1	-1.7
11	2.051	0.53	2.51	1.3	2.4	1.9	1.9
12	-0.608	2.33	1.11	1.1	1	-0.7	0.5

Table 3 Scaling down normalized values using mean.

Fatal. accidents	Grievous. Injuries	Minor. Injuries	Non. Injuries	Total. accidents	Persons. Killed	Persons. Injured
541.833	134.500	952.416	133.333	1762.083	587.250	1913.000

Table 4: Scaling down normalized values using standard deviation

Fatal. accidents	Grievous. Injuries	Minor. Injuries	Non. Injuries	Total. accidents	Persons. Killed	Persons. Injured
284.304	150.105	570.171	75.051	920.142	305.161	881.967

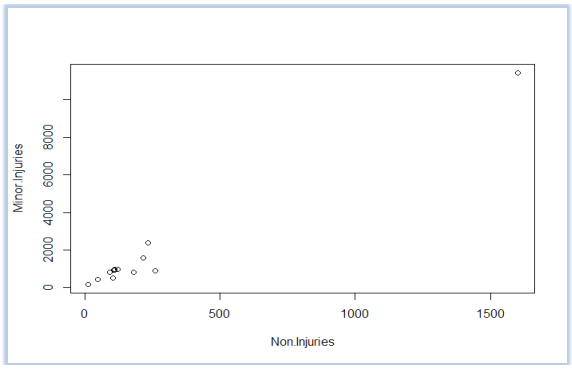

Fig 2: Non-injuries versus Minor.Injuries.

Table 5. The district wise accidental probabilities.

	Persons.Killed	Persons.Injured
[1,]	-0.38845481	-0.34859679
[2,]	-0.25679460	-0.29111041
[3,]	-0.45290771	-0.43118047
[4,]	-0.41930406	-0.42439742
[5,]	-0.33557037	-0.21700555
[6,]	-0.29921232	-0.31027254
[7,]	-0.09097987	-0.20733969
[8,]	-0.07004645	-0.25312531
[9,]	-0.15928893	-0.10542431

```
[10,]  -0.46392530  -0.52733026
[11,]   0.04453649   0.01684024
[12,]  -0.39286184  -0.19496062
[13,]   3.28480977   3.29390314
```

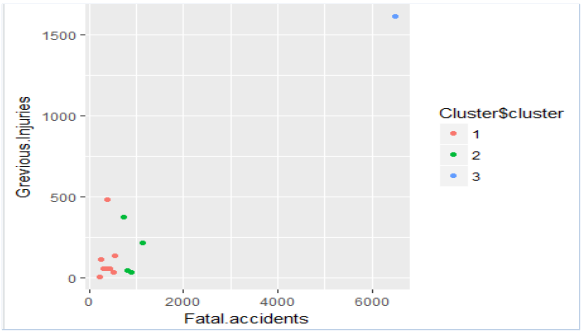

Fig 2. Fatal accidents

	reddy							
11	Cyberabad	1125	214	2381	234	3954	1165	3631
12	Hyderabad	369	484	1587	217	2657	371	2382
	Total	6502	1614	11429	1600	21145	7047	22956

Table 6: Enhanced k-mean clustering algorithm

S.no	**District/ Unit**	**Fatal accidents**	**Grievous Injuries**	**Minor Injuries**	**Non-Injuries**	**Total accidents**	**Persons Killed**	**Perons Injured**
1	Adilabad	354	58	903	260	1575	379	1476
2	Karimnagar	543	136	826	92	1597	618	1815
3	Warangal City	248	113	503	105	969	262	989
4	Warangal Rural	295	56	423	48	822	323	1029
5	Khammam	425	57	974	110	1566	475	2252
6	Nizamabad	500	37	808	182	1527	541	1702
7	Medak	806	46	981	123	1956	919	2309
8	Mahabubnagar	885	35	930	111	1961	957	2039
9	Nalgonda	726	373	941	106	2146	795	2910
10	Ranga	226	5	172	12	415	242	422

Table 7: The cluster means for members for complete linkage

```
//assign each point to its nearest cluster
a. For i = 1 to n
b. For j = 1 to k
c. Compute squared Euclidean distance d2(xi, mj);
d. endfor
e. Find the closest centroid mj to xi;
f. mj = mj+xi; nj = nj+1;
g. MSE = MSE + d2(xi, mj);
h. Clusterid[i] = number of the closest centroid;
i. Pointdis[i] = Euclidean distance to the closest
   centroid;
j. endfor
k. For j = 1 to k
l. mj = mj/nj;
m. endfor
```

Table 8 : Telangana District wise road accident data.

S. No	**Fatal. Accidents**	**Grevious. Injuries**	**Minor. Injuries**	**Non. Injuries**	**Total. accidents**	**Persons.Killed**	**Persons.Injured**
1	0.153591	-0.48633	-0.08 55	0.17322	-0.07 07	0.1996	0.02173
2	0.697023	1.482294	1.19 949	0.6 973	1.25 7324	0.6218	1.20337
3	-1.00421	-0.50964	-1.02 849	-1.0 437	-1.11 586	-1.021	-1.2468

Retrieval Number: F7687038620/2020©BEIESP
DOI:10.35940/ijrte.F7687.038620
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

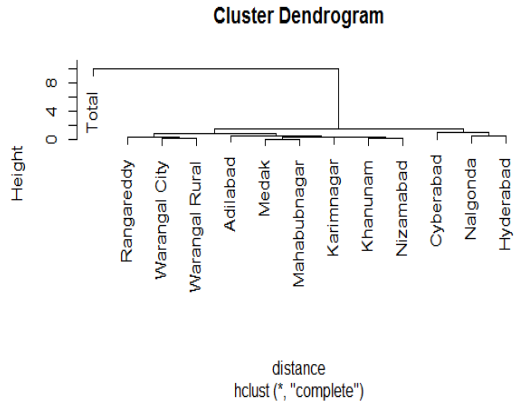

Fig 4. District wise accident data clustering.

Table 9 : accident data mean, median and stand deviation.

	Fatal. accidents	Grevious.Injuries	Minor. Injuries	Non. Injuries	Total. accidents	Persons.Killed	Persons. Injured
min	226	5	172	12	415	242	422
1st qua	354	46	808	105	1527	371	1476
Median	500	58	930	111	1597	541	2039
mean	1000	248.3	1758	246.2	3253	1084	3532
3rd qua	806	214	981	217	2146	919	2382
max	6502	1614	11429	1600	21145	7047	22956

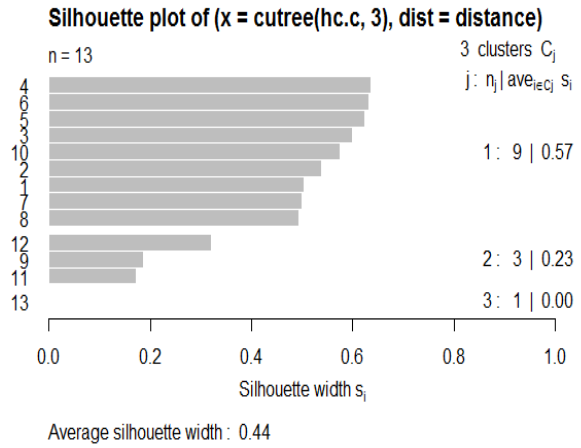

Fig 5. Hierarchical clustering with complete linkage members.

In the cluster summation of squares by cluster:
[1] 0.8860122 0.0000000 0.6494089
(between_SS / total_SS = 98.2 %)

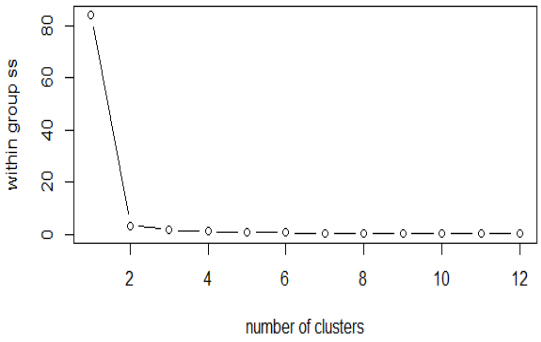

Fig 6: variability or sum of square within the group

V. CONCLUSION

The proposed framework providing better recommendation system for analyzing catastrophe models for diverse types of catastrophes on the road which creates use of K modes grouping algorithm. For implementation analysis, we utilize Telangana road accident data for district wise data and applying similarities between accidents which can occur in diverse districts. For identifying the accident prone sphere in the real time situation in Telangana state and also collect the different catastrophes in diverse kinds of injuries, which predicts the accident prone area and provides information for predicting and preventing the incidents not changed into fatal accidents with the intention of save lives.

VI. FUTURE ENHANCEMENTS

To establish the association between the causes and the consequences with the event classification of an investigated case by highlighting the dynamic driving situation with initial traveling speed, pre-impact and the post-impact speed of involved vehicles to describe the crash scenario. The understanding of the driver's behavior will assist us to develop a system which can easily detect highly emotional agitated driver so that we can prevent the accident. In the future analysis create the recommendation system for reasons and consequences along with occurrence categorization of a study by the importance of the active driving condition with the preliminary traveling rate, pre and post-impact speed of the concerned vehicles to illustrate the catastrophe scenario. Driver behavior analysis will assist to expand a method which can easily perceive extremely emotional disturbed driver with the intention that we can avoid the accident.

REFERENCES

1. ShiJianjun et al., "The Analysis of Traffic Control Cyber-physical Systems", Procedia - Social and Behavioral Sciences Volume 96, 6 November 2013, Pages 2487-2496.
2. Muritala Abdulkabir et al., "Trend Analysis on Road Traffic Accident in Nigeria" Science Invovation , vol 3 no.5., pp. 52-67, 2015, DOI: 10.11648/j.si.20150305.12 .

3. Otte et al., "Accident Causation Factor Analysis of Traffic Accidents on the Example of Elderly Car" Accident Research Unit, Hannover Medical School (1) Germany , 2013.
4. Karolien Geurts " Traffic Accidents Data Set ", http://fimi.ua.ac.be/data/accidents.pdf, 2016.
5. Sachin Kumar et al., "A novel framework to analyze road accident time series data", ournal of Big Data20163:8 https://doi.org/10.1186/s40537-016-0044-5
6. D. Karlis et al., "Time Series Models for Road Safety Accident Prediction", 2008.
7. Kumar S, Toshniwal D. A data mining approach to characterize road accident locations. J Mod Transp. 2016;24(1):62–72. doi:10.1007/s40534-016-0095-5.
8. 20. Warren LT. Clustering of time series data—a survey. Pattern Recogn. 2005;38(11):1857–74.
9. Fraley C, Raftery AE. Model-based clustering, discriminant analysis and density estimation. J Am Stat Assoc. 2002;97(458):611–31.
10. Desai MM, Patel AK (2011) Road Accidents Study Based On Regression Model: A Case Study of Ahmedabad City.
11. Chakra borty S, Roy SK (2005) Traffic accident characteristics of Kolkata. Transport and Communications Bulletin for Asia and the Pacific 74: 75-86.
12. Turner S (1996) Estimating Accidents in a Road network. PhD Thesis, University of Canterbury, New Zealand, pp. 6-7.
13. Ye X, Liu T, Ye Y (2015) Enhanced grain refinement and Microhardness of Ti-Al-V alloy by electropulsing ultrasonic shock. J Alloys and Comp 621: 66-70.

AUTHORS PROFILE

Dr. Kakulapati Vijayalakshmi, received Ph. D in Computer Science & Engineering in the area of Information Retrieval from JNTU Hyderabad. Currently she is working as a Professor in the Department of Information Technology, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India. She is having 26 years of industry and teaching experience. She is a member of various professional bodies like IEEE, ACM, CSTA, LMISTE, LMCSI, IACSIT, FIETE and few more. She is having more than 80+ publications in National, International Journals she is working with evaluating and designing novel search algorithms for web search, summarization. Currently she is working with big data analytics, health informatics, Data science and Internet of Things.

Srikar Bharade, is completed his B. Tech in Information Technology at Sreenidhi Institute of Science and Technology. Currently he is working in Corecompete as Associate consultant. His area or interest is data science and big data analytics.

Mr. Nikhil M., is completed his B. Tech in Information Technology at Sreenidhi Institute of Science and Technology. Currently he is working in Capgemini as senior analyst. His area or interest is data science and big data analytics.