Road Accident Prediction and Classification using Machine Learning

Chirag P
Department of ECE
JSS Science and Technology University,
Mysore, India
chirag.p566@gmail.com

Supreetha M
Department of ECE
JSS Science and Technology University,
Mysore,India
supreetha.manjanna@sjce.ac.in

2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon) | 978-1-6654-9790-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/MysuruCon55714.2022.9972671

***Abstract*—The safety of roads has long been a key concern for the entire world, and everyone has been working to address it. Due to a constant increase in the number of vehicles, traffic has evolved into a challenging structure to manage. Due to this condition, the issue of road accidents has been identified, and studies on potential solutions have been conducted all over the world. In this paper, we have discussed about one such model where we make use of machine learning and data mining concepts to predict and analyze such accidents all over the country. We are using regression and clustering types of Machine Learning algorithms to predict and analyze the accident rate for the year 2022 for all the States and UTs. We are using Linear Regression algorithm for the prediction of the accident rate for the year 2022. Furthermore, we classify all the states and UTs into two clusters where one cluster consists of states with High accident rate and another cluster consists of states with Low accident rate. We are using K-Means clustering algorithm for the classification of the states and UTs into clusters. We are considering many external factors such as Alcohol, Weather, Location, Junction and Vehicle Defect as the main parameters for the prediction and classification. It is high time for the government as well as the citizens of this country to take precautionary measures to avoid such fatal incidents and enhance the road safety and the information obtained from this model will be used to enhance the safety of the travellers.**

Keywords— Road Accident Classification and Prediction, Linear Regression, K-means Clustering

I. Introduction

Road accidents are one of the major global causes of mortality, disability, and hospitalisation, according to the World Health Organization. India accounts for at least one in ten of the global traffic fatalities. In addition to the victims and their families, the economy as a whole bears the cost of road accidents in terms of premature deaths, injuries, disabilities, and lost potential income. 1,31,714 people lost their life in road accidents in 2020, which resulted in 3,66,138 overall injuries and fatalities. Unfortunately, the age group most frequently afflicted by traffic accidents—18 to 45 years old—represents around 70 percent of all accidental fatalities.

Every day, there are many automobiles on the roads, and traffic accidents can occur anywhere, anytime. Everyone tries to prevent accidents in order to be safe, however some incidents do result in fatalities. The traffic accident dataset could be subjected to the data mining technique to uncover some useful information and provide driving advice to the travellers for a safer drive.

Data mining, also referred to as knowledge discovery in data (KDD), is a technique for extracting patterns and other crucial information from massive data sets. Whereas business intelligence focuses on business information and involves data analysis that mainly relies on aggregation. Data mining makes use of a wide range of methods and algorithms to find patterns and relationship in massive amounts of data. It is regarded as one of the most significant information technology tools from earlier decades.

We evaluate data on traffic accidents using regression algorithms, which aid in making predictions about the future based on historical data. On the fatal accident dataset, the K-means clustering technique was used to determine which states are comparable to one another in terms of fatality rates as well as which states are safer or riskier to drive in. For our research, we used datasets on traffic accidents. The datasets from 2013-2021 can be downloaded from data.gov.in. The analysis phase entails gathering data, applying statistics to the prior datasets [2013-2021], and predicting the results for 2022.

For our research, we used datasets on traffic accidents. Downloads of the datasets can be made at data.gov.in. Data sets from 2013 to 2021 were gathered. It provides the exact number of accidents that has occurred in all the states and union territories of India. The analysis phase entails gathering data, applying statistics to the prior datasets, and projecting the outcome for 2022.

II. LITERATURE SURVEY

Asghar Pasha et al, has discussed about allocating the most suitable strategy for classification of road accident measurement utilizing data mining approaches is intended. In order to learn more about the characteristics of elements like driver behavior, road conditions, lighting conditions, weather conditions, etc., they develop models using accident data sets. This will help the users calculate the safety measures that are used to prevent accidents. [1]

Md. Farhan Labib et. al, By using cutting-edge machine learning techniques, this analysis aims to assess traffic incidents and assess their seriousness. There are numerous

978-1-6654-9790-9/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: University of North Texas. Downloaded on June 29,2023 at 00:18:16 UTC from IEEE Xplore. Restrictions apply.

sophisticated machine learning techniques available to study this field. The four most common and advanced supervised learning approaches of machine learning are used by the authorsto analyze traffic accidents because of their accuracy in this field. These methodologies include Decision Tree, K- Nearest Neighbors (KNN), Naive Bayes, and Adaptive Boosting (AdaBoost). By utilizing this process, a mobile application can be developed in the future that will give the user an exact prediction and be extremely advantageous. [2]

MUBARIZ MANZOOR et al, has discussed important factors that are very much strong in correlating with the severity of accidents on highways are discovered by utilizing the Random Forest. The main features that affect severity of accidents involves distance, temperature, wind chill, humidity, visibility, and direction of the windblown. This study also pointed out an togetherness of models by machine learning and deep learning through combining Random Forest and Convolutional Neural Network called RFCNN to foresee the severity of road accidents. They have explained their future scope such as plans to apply the model that is proposed on the multi-domain data-set to make its effectiveness to be proved and how the results can be utilized effectively [3]

Vipul Rana et.al have talked about accidents that are caused by several things. Road accidents are primarily influenced by variables like vehicle kinds, driver age, vehicle age, weather, road structure, and other factors. Therefore, based on the aforementioned parameters, they have created an application that provides effective predicted values for traffic accidents. [4]

YUN-FENG ZHOU et.al has explained about the Unsupervised deep learning framework for traffic accident video prediction based on first-person is built, and a risk scoring evaluation approach is provided, accordingly. [5]

Jayesh Patil et. al, has sought to identify what is driving the majority of the local increase in the number of traffic accidents. K-means algorithm was applied for analysis. [6]

Mubashir Murshed et. al, has detailed a smart system that warns and regulates a vehicle's speed and alerts the appropriate people when an accident happens. Utilizing a distance sensor, this device continuously keeps track of the separation between oncoming traffic and any impediments. When a critical distance approaches, it will warn the driver to regulate speed and automatically slow down. [7]

Niyogisubizo et al has used the machine learning-based algorithms to anticipate the seriousness of crash injuries, examine the most important aspects that contribute to road crashes, and provide recommendations to interested parties. Four classification techniques were used in this work to examine feature importance and forecast accident severity: Random Forest (RF), Multinomial Naive Bayes (MNB), K-Means Clustering (KC), and K-Nearest Neighbors (KNN). [8]

Koteswara Rao Ballamudi et. al, gives an analysis of a number of active enterprises using machine learning to forecast accidents. For prediction, decision trees, native bayes, and KNN algorithms are utilized. [9]

Daniel Santos et. al, has provided a rule generating model and a machine learning hotspot identification method to identify the causes of serious accidents. gathering information and combining datasets such as weather, time, traffic, and road data. An examination of the accident dataset is supported by the rule generating model and addresses the causes of serious traffic accidents. The prediction approach seeks to identify accident hotspots by emphasizing places where accidents are most likely to occur under specific conditions. [10]

Shweta et. al, has talked about the data from Analysis of this kind of data is challenging since road accidents are highly varied in character. The primary task in such data analysis is segmentation. As a result, the research work's suggested K-means clustering method is primarily used for it. Using supervised machine learning, this model's second objective is to extract data, images, and hidden patterns that will be used to help create regulations to pre- vent traffic accidents. Information that is fully meaningful is produced by the combination of segmentation and machine learning algorithms. [11]

Yunzhi Shi et. al, has put out new ideas for statistical and clustering-based feature engineering and data preprocessing. Hyper-parameter optimization (HPO) and the free and open-source AutoGluon module are used to significantly increase the performance of our model's ability to anticipate risk.[12]

Monisha Lakshme Gowda et. al, describes how to control where and when accidents hap- pen in London. When the study's findings are analysed, it becomes clear that ”Ensemble Machine Learning” algorithms can be used to forecast the location and time of traffic acci- dents. Since the classification is based on the cluster location and previous accident time, the suggested approach is appropriate for real-time solutions .

[13] Priyanka et. al, outlines the proposed approach, which focuses on segmenting data on traffic accidents using the k-means clustering approach. Additionally, association rule mining is used to find instances of the entire data set and instances of clusters identified by the k-means clustering technique. Major information is produced by the combined effect of k-means clustering and association rule. [14]

Alyssa Ditcharoen et. al, offer a summary of the elements affecting the severity of traffic accidents and discuss the methods that were often employed in earlier studies, such as logistic regression and power modelling. [15]

III. METHEDOLOGY

A. *Steps Involved in this project*

- Data Preparation: Each model creation starts with data preparation. In the selected attributes, any records with missing values (often represented in the dataset by 0) were eliminated. All numeric values were changed to nominal values in compliance with the data dictionary. Absence of Values: when no data value is kept for the observation.
- Modeling: Prior to applying Regression and grouping correlations between the qualities and the patterns, we first calculated a number of statistics from the dataset to

Authorized licensed use limited to: University of North Texas. Downloaded on June 29,2023 at 00:18:16 UTC from IEEE Xplore. Restrictions apply.

demonstrate the fundamental characteristics or the basic the basic traits of the fatal incidents.

- Result Analysis: The findings of our analysis number of conclusions, including correlations between the variables, clustering of Indian states based on population and fatal accident rates, and classification of locations based on risk of fatal accident likelihood. For these analyses, we used the data analytic tool Highcharts.

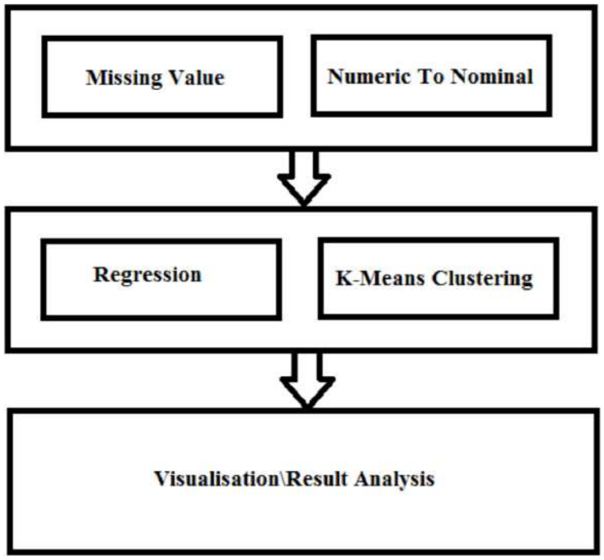


Fig. 1: Proposed Block Diagram

IV. DESIGN AND IMPLEMENTATION

A. Steps for Data Collection

Data collection is a process for compiling and arranging information from a limitless number of sources. Following data collection, the information is organized into Excel sheets (.xlsx) in the order we will handle it. Due to the unstructured nature of the data collected, we grouped the data into pairings of states and years in which the criminal occurrences took place by creating columns and rows in Excel sheets based on our data analysis.

The data sets were internally gathered in secondary form to help with the construction of the predictive model. Secondary data are statistics that were obtained from a record or other public source, such a central government agency, rather than being developed or originated by the investigator himself.

The complete data set is obtained from the website: https://data.gov.in/catalogspath=sector/Transport

As said earlier the data obtained from the above website provides the exact number of accidents that has occurred in all the states and union territories of India.

B. Data Pre-processing

Pre-processing is a data mining technique used to convert unstructured data into a suitable and practical format. Data pre-processing requires the following steps:

1) Data Cleaning**:** The data may contain many irrelevant and missing pieces. Data cleaning is done in order to control this portion. It involves handling noisy, missing, and other types of data.
2) Data Transformation:This transformation is carried outto convert the data into the useful formats needed for the data mining process. It involves the following actions:
   - Normalization: Normalization is carried to obtain data values inside the desired range.
   - Attribute Selection: Brand new attributes are generated from already provided set of attributes.
   - Discretization: It is done to put back the unprocessed values of numerical attributes by interval levels.
   - Concept Hierarchy Generation: According to hierarchy, the attributes are changed from a low level to a high level.
3) Data Reduction: Since mining of data is an approach that is used to manage large chunks of data. It was challenging to assess in these situations while working with a significant amount of data. So we used a data reduction strategy to get away from this. The goal of the reduction was to increase storage efficiency and lower storage and analysis costs.

C. Algorithms used for Prediction and Classification

The algorithm for machine learning is an approach by which the system of AI capabilities performs the processes, normally by foreseeing the values as output from already provided data as input. The important actions of algorithms of machine learning are regression and classification of data.

1) Linear Regression: It is an algorithm of machine learning where it is based on the supervised learning approach. The models of regression is an aimed foreseen value based on the variables which are not dependent. It is utilized for discovering the connection in-between variables and its estimations. Various regression models are different based on – the type of connection between dependent and independent variables they are seeing, and independent variables total numbers that are being utilized.

By using the already provided independent variable, x, and the dependent variable, y, this regression is able to predict the value of y. As a result, this method of regression identifies a linear relationship between the input x and the output y. Thus, it is referred to as linear regression.

To forecast the future values of the dependent variable, a regression algorithm is created to identify the past link between an independent and a dependent variable. Regression uses the historical relationship between variables to forecast how they will behave going forward. Based on the data set gathered for the project, the algorithm employs techniques for linear regression. With the aid of statistical techniques, the linear regression technique aids in forecasting the future behavior of road accidents. To forecast future behavior, the algorithm calculates the mean and variance of the dependent variables and applies the formula Y=b0+b1*x.

Authorized licensed use limited to: University of North Texas. Downloaded on June 29,2023 at 00:18:16 UTC from IEEE Xplore. Restrictions apply.

Functional Hypothesis for Linear Regression:

$$Y = b_0 + b_1 \cdot X \quad (1)$$

when the model training is performed we are provided with: x: training data input, y: data labels, ϑ_1 : *Intercept* and ϑ_2 : *x − Coefficient.*

2

$$J = \frac{1}{n} \sum_{i=1}^{n} (pred_i - y_i)^2 \quad (2)$$

The Root Mean Squared Error in-between the foreseen y_predicted value and y's true value is the Cost function of Linear Regression represented by J.

The main advantage of Linear regression is that it performs extremely good for linearly dividable data but its quite difficult to make assumption of linearity between dependent and independent variables.

2) K-Means Clustering: It is a type of unsupervised learning algorithm, that organizes the not labeled data-set into various clusters is known as K-means Clustering. Here, K depicts the predefined clusters total value that is required to create in the process, so if K=2, it will create two clusters, and if K=3, it will create three clusters, and so on.

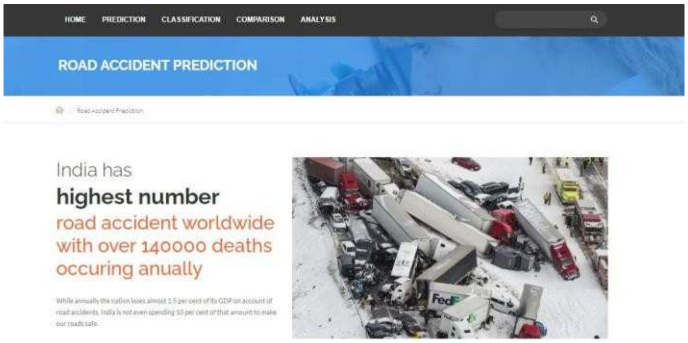
If k is provided, then algorithm K-means could be accomplished by following steps:

- K not empty subsets object partitioning
- Current partition's cluster centroids must be identified
- Every point should be assigned with particular cluster
- Distance computation between every points and allocate points to the cluster where minimum distance is obtained from centroid
- Discovering Centroid of the brand new cluster after reallocating the points
- Calculate the distance again between each data point and the new cluster centres that were discovered.
- Stop if no data point was moved; otherwise, go back to step 5 and repeat.

To Choose the 'k' value in K-means Clustering:

The main advantage of utilizing the K-means is that it is easy to implement but quite difficult in terms of predicting the k-value.

V. RESULTS AND DISCUSSION



India has **highest number** road accident worldwide with over 140000 deaths occuring anually

While annually the nation loses almost 1.5 per cent of its GDP on account of road accidents, India is not even spending 10 per cent of that amount to make our roads safe.

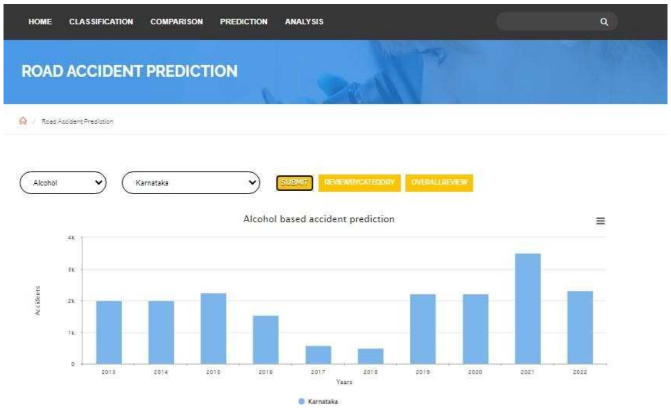
Fig. 2: Home Page

The Figure 2 shows the Home page of the designed system.

A. Prediction

Under this category we can see the predicted results ie number of accidents for the year 2022 considering all the different parameters as mentioned below.

- Prediction based on Alcohol.
- Prediction based on Climate(hot, cloud, heavy rain, light rain, snow).
- Prediction based on type of Location (near bazar, near factory, near school, near temple, others).
- Prediction based on type of Vehicle Defect (Defective-Brakes, Punctured, BaldTyres, Others).
- Prediction based on type of Junction (Tjunction, YJunction, FourArm, Round, RailCross).



Alcohol

Karnataka

Fig. 3: Predicted Results of 2022 based on Alcohol[Karnataka]

The Figure 3 shows the actual accident rate for the year 2013-2021 and predicted accident rate for the year 2022 for the state of Karnataka based on Alcohol.

Authorized licensed use limited to: University of North Texas. Downloaded on June 29,2023 at 00:18:16 UTC from IEEE Xplore. Restrictions apply.

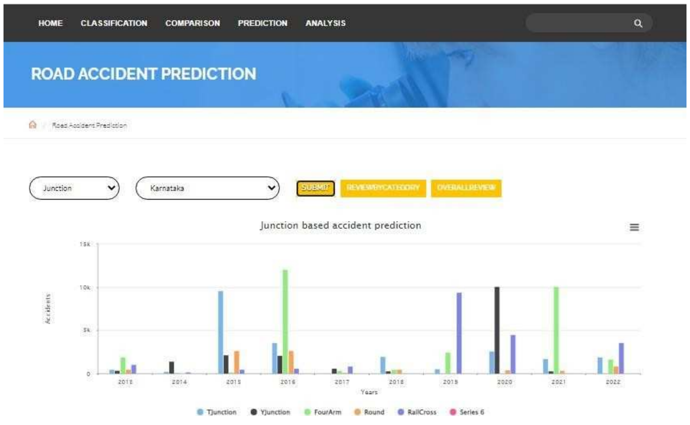


Fig. 4: Predicted Results of 2022 based on Junction[Karnataka]

The Figure 4 shows the actual accident rate for the year 2013-2021 and predicted accident rate for the year 2022 for the state of Karnataka based on Junction [T-Junction, Y-Junction, FourArm, Round, Railcross]

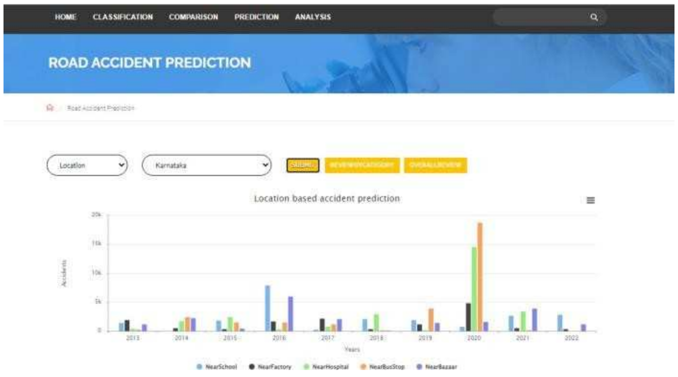


Fig. 5: Predicted Results of 2022 based on Location [Karnataka]

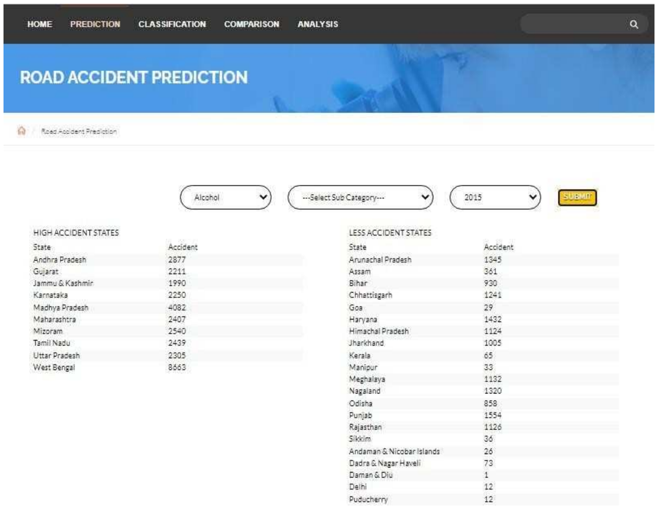
The Figure 5 shows the actual accident rate for the year 2013-2021 and predicted accident rate for the year 2022 for the state of Karnataka considering the location and the main category and following as the sub-categories: near school ,NearFactory , NearHospital, NearBusstop ,NearBazaar.

B. *Classification*

Under this category we can see the classification of states into two clusters of high frequency accident states and Low Frequency accident States for the different categories and Sub-categories for the year 2013-2021.

Parameters considered for classification

- Classification based on Alcohol.
- Classification based on Climate(hot, cloud, heavy rain, light rain, snow).
- Classification based on type of Location (nearbazaar, nearfactory, nearschool, neartemple, others).
- Classification based on type of Vehicle Defect (DefectiveBrakes, Punctured, BaldTyres, Others).
- Classification based on type of Junction (Tjunction, YJunction, FourArm, Round, RailCross).



HIGH ACCIDENT STATES

State	Accident
Andhra Pradesh	2877
Gujarat	2211
Jammu & Kashmir	1990
Karnataka	2250
Madhya Pradesh	4082
Maharashtra	2407
Mizoram	2540
Tamil Nadu	2439
Uttar Pradesh	2305
West Bengal	8663

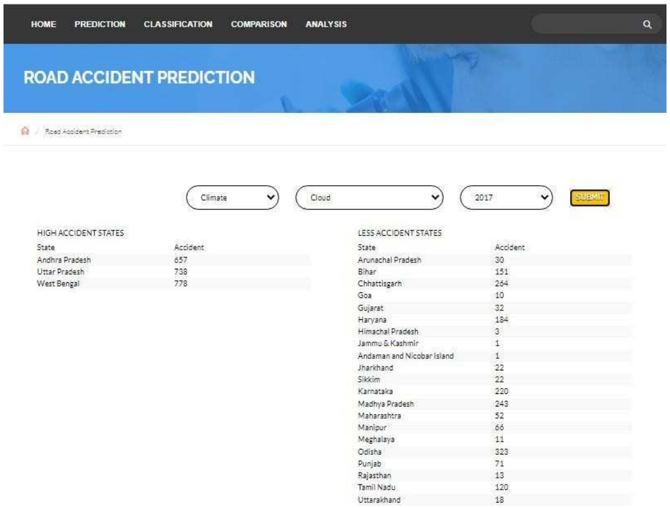
LESS ACCIDENT STATES

State	Accident
Arunachal Pradesh	1345
Assam	361
Bihar	930
Chhattisgarh	1241
Goa	29
Haryana	1432
Himachal Pradesh	1124
Jharkhand	1005
Kerala	65
Manipur	33
Meghalaya	1132
Nagaland	1320
Odisha	858
Punjab	1554
Rajasthan	1126
Sikkim	36
Andaman & Nicobar Islands	26
Dadra & Nagar Haveli	73
Daman & Diu	1
Delhi	12
Puducherry	12

Fig. 6: Classification based on Alcohol[2015]

The Figure 6 shows the classification based on Alcohol for the year 2015.

- Cluster A(high accident states) In cluster A out of 29 states, 10 states had relatively high fatality rate of accident due to alcohol.
- Cluster B(low accident states) That state in cluster B represents safe states with relatively lower fatality rate



HIGH ACCIDENT STATES

State	Accident
Andhra Pradesh	657
Uttar Pradesh	738
West Bengal	778

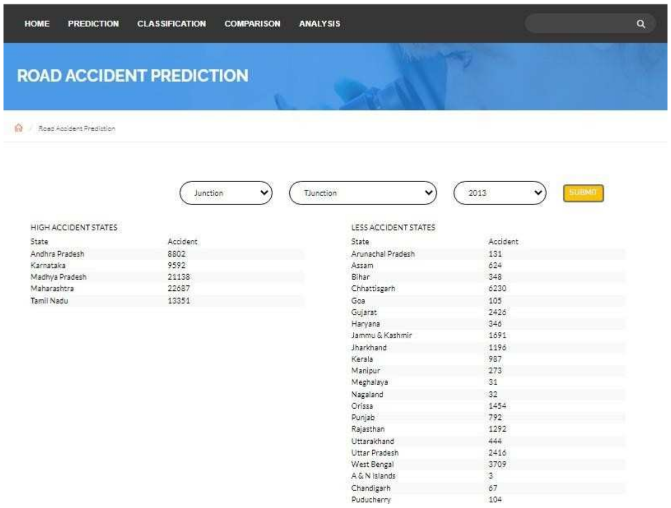
LESS ACCIDENT STATES

State	Accident
Arunachal Pradesh	30
Bihar	151
Chhattisgarh	264
Goa	10
Gujarat	32
Haryana	184
Himachal Pradesh	3
Jammu & Kashmir	1
Andaman and Nicobar Island	1
Jharkhand	22
Sikkim	22
Karnataka	220
Madhya Pradesh	243
Maharashtra	52
Manipur	66
Meghalaya	11
Odisha	323
Punjab	71
Rajasthan	13
Tamil Nadu	120
Uttarakhand	18

Fig. 7: Classification based on Climate(Cloud-2017)

The Figure 7 show the classification based on the Climate(cloud) for the year 2017. The Algorithm show that out of 29 states only 5 states are risky to drive and others are safer.

Authorized licensed use limited to: University of North Texas. Downloaded on June 29,2023 at 00:18:16 UTC from IEEE Xplore. Restrictions apply.



Junction | TJunction | 2013

HIGH ACCIDENT STATES

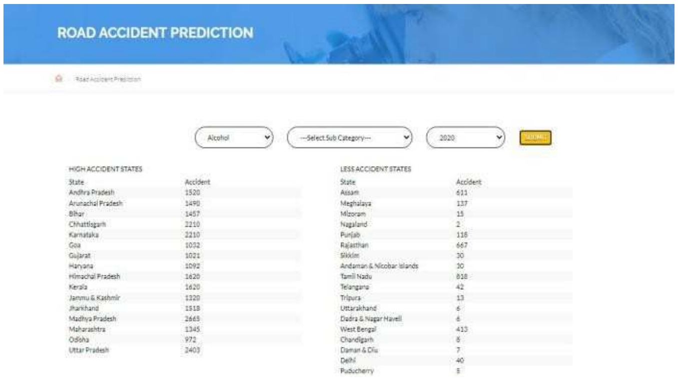
State	Accident
Andhra Pradesh	8802
Karnataka	9592
Madhya Pradesh	21138
Maharashtra	22687
Tamil Nadu	13351

LESS ACCIDENT STATES

State	Accident
Arunachal Pradesh	131
Assam	624
Bihar	348
Chhattisgarh	6230
Goa	105
Gujarat	2426
Haryana	346
Jammu & Kashmir	1691
Jharkhand	1196
Kerala	987
Manipur	273
Meghalaya	31
Nagaland	32
Orissa	1454
Punjab	792
Rajasthan	1292
Uttarakhand	444
Uttar Pradesh	2416
West Bengal	3709
A & N Islands	3
Chandigarh	67
Puducherry	104

Fig. 8: Classification based on Junction(T-Junction)

The Figure 8 shows the fatality of road accident due to Junction(T-Junction) for the year 2013. The figure shows that out of 29 states only 5 states are more risky states to travel, and other states are safety to travelled via T-junction.



Alcohol | ---Select Sub Category--- | 2020

HIGH ACCIDENT STATES

State	Accident
Andhra Pradesh	1520
Arunachal Pradesh	1490
Bihar	1457
Chhattisgarh	2210
Karnataka	2210
Goa	1032
Gujarat	1021
Haryana	1092
Himachal Pradesh	1620
Kerala	1620
Jammu & Kashmir	1320
Jharkhand	1518
Madhya Pradesh	2665
Maharashtra	1345
Odisha	972
Uttar Pradesh	2403

LESS ACCIDENT STATES

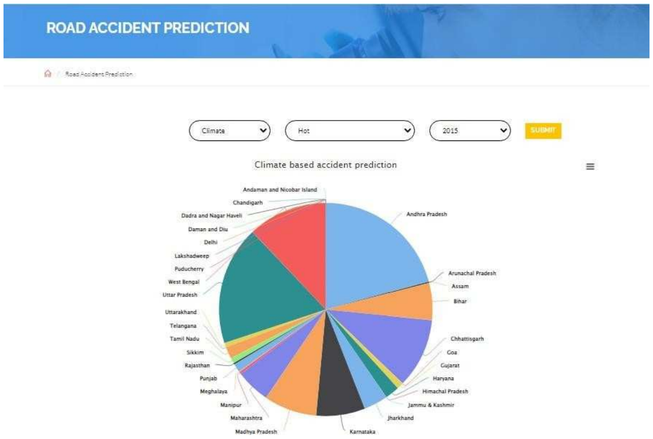
State	Accident
Assam	611
Meghalaya	137
Mizoram	15
Nagaland	2
Punjab	116
Rajasthan	667
Sikkim	30
Andaman & Nicobar Islands	30
Tamil Nadu	818
Telangana	42
Tripura	13
Uttarakhand	6
Dadra & Nagar Haveli	6
West Bengal	413
Chandigarh	8
Daman & Diu	7
Delhi	40
Puducherry	5

Fig. 9: Classification between High Frequency and Low fre-quency States

The figure 9 shows two clusters of state wise road accident rate based on the classification of Alcohol for the year 2020. Cluster A(high accident states)In cluster A out of 29 states, 16 states had relatively high fatality rate of accident due to fog. Cluster B(low accident states)That state in cluster B represents safe states with relatively lower fatality rate. Some states are missing due to accident rate 0 or government has no updated the details.

C. Comparison of rate of accidents of all states and UTs

Under this category we can see the pictorial representation[Pie Chart] of the accident rate of all the states and UTs considering all above parameters from 2013-2021



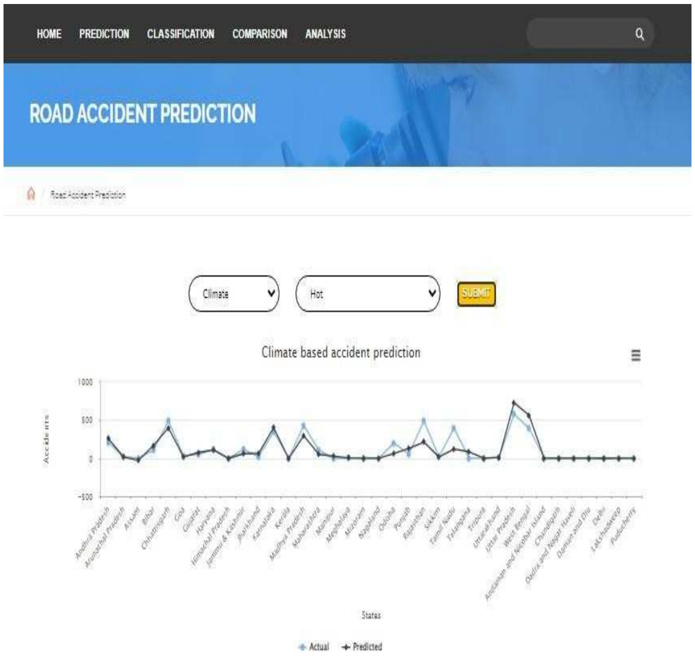
Climate | Hot | 2015

Fig. 10: Comparison of Climate based accident rate for all thestates and UTs in 2015

The Figure 10 shows the pictorial representation of the alcohol based accident rate of all the states and UTs for the year 2015.

D. Analysis

Under this category we can see the comparison between the actual results and predicted results for the year 2021 to check the efficiency of the algorithm.

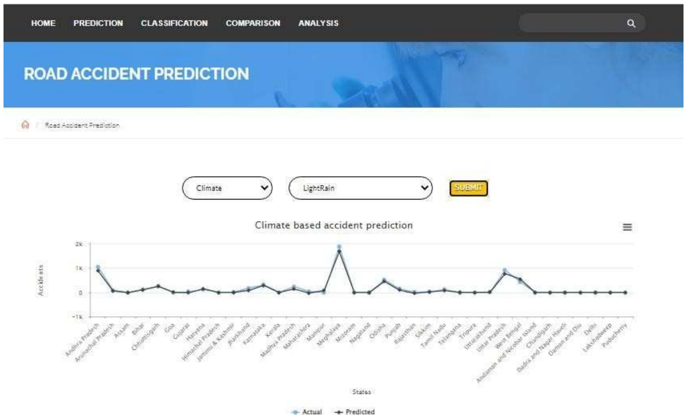


Climate | Hot

Fig. 11: Analysis of accidents due to climate

The Figure 11 shows the comparison between the actual results and predicted results based on climate(Hot).

Authorized licensed use limited to: University of North Texas. Downloaded on June 29,2023 at 00:18:16 UTC from IEEE Xplore. Restrictions apply.



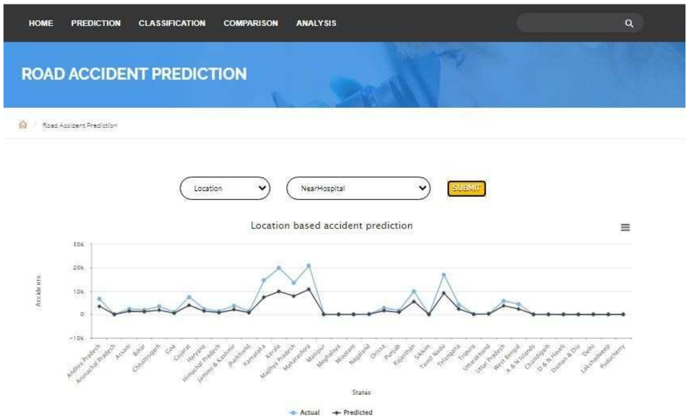
Climate

LightRain

Climate based accident prediction

Fig. 12: Analysis of accidents due to Climate(Light Rain)

The Figure 12 shows the comparison between the actual results and predicted results based on climate(Light Rain).



Location

NearHospital

Location based accident prediction

Fig. 13: Analysis of accidents due to Location(NearHospital)

The Figure 13 shows the comparison between the actual results and predicted results based on Location(NearHospital).

VI. CONCLUSION

In our proposed work we have developed a website using Java. After analysing the data sets of accidents from the previous years ie 2013-2021 we can predict the accident rate for 2022. As discussed in the implementation and results sec- tion we have implemented regression and clustering algorithm for the prediction and classification respectively. Since the previous accidents data obtained is continuous in nature and we have only one dependent and independent variable we decided that Linear Regression algorithm is the best suitable Machine Learning technique for prediction.

From this work, researchers will be able to evaluate the severity of traffic accidents as well as the factors that lead to them. According to statistics, linear regression and categorization, environmental factors like the state of the road, the weather, and the time of day do not significantly affect the fatality rate, however human factors like whether or not a person is drunk and the type of collision will have an effect.

Although no one can anticipate when an accident will happen, the analysis of this data can assist the government and its citizens in taking precautions to ensure their safety. In the future, we'll aim to improve the accuracy of our model and try to use other machine learning techniques to produce more satisfying results with maximum accuracy.

According to the clustering results, some states and regions have a greater fatality rate than others. When travelling in certain dangerous states or places, we need to drive more carefully. Through the assignment completed, we came to the conclusion that there is never enough facts to make a solid decision. Additional tests could be run and more recommendations could be drawn from the data if we could collect more data, including as information on non-fatal accidents, weather, mileage, and other factors. Thus, making the model more efficient.

VII. FUTURE SCOPE

The future scopes of this project are as follows:

- This Web application can be further deployed into android/IOS and make them available to mobile devices so that it can be used by all the users.
- We can use many other effective Machine Learning algorithms to get maximum accuracy.
- In future we can enhance the application by giving more number of historical data and can design the application very efficiently which acts as an encyclopedia for the traffic authorities which gives more information about the traffic accidents that occurs throughout the country.
- We can further improve the performance of the system in terms of operating speed, memory capacity using cloud computing to store the data and process more intensive application.

REFERENCES

[1] Asghar Pasha, Vijayalakhshmi, MD Atique, MD Hussain, Harsh narnot, Bipin ”Road Accident Prediction using Machine Learning ”, 2021 .

[2] Md. Farhan Labib, Ahmed Sady Rifat, Md. Mosabbir Hossain, Amit Kumar Das, Faria Nawrine ”Road Accident Analysis and Prediction of Accident Severity by Using Machine Learning in Bangladesh ”, 2021 .

[3] MUBARIZ MANZOOR, MUHAMMAD UMER, SAIMA SADIQ, ABID ISHAQ,SALEEM ULLAH, HAMZA AHMAD MADNI ”RFCNN: Traffic Accident Severity Prediction Based on Decision Level Fusion of Machine and Deep Learning Model ”, 2021.

[4] Vipul Rana, Hemant Joshi ”Road Accident Prediction using Machine Learning Algorithm”, 2019.

[5] YUN-FENG ZHOU, KAI XIE ”Efficient Traffic Accident Warning Based on Unsupervised Prediction Framework”, 2021

[6] Jayesh Patil, Mandar Prabh ”Road Accident Analysis using Machine Learning ”, 2020.

[7] Mubashir Murshed, Md Sanaullah Chowdhury ”An IoT Based Car Accident Prevention and Detection System with Smart Brake Control”, 2019.

[8] Jovial Niyogisubizo, Evariste Murwanashyaka ”A Comparative Study on Machine Learning-based Approaches for Improving Traffic Accident Severity Prediction”, 2021.

Authorized licensed use limited to: University of North Texas. Downloaded on June 29,2023 at 00:18:16 UTC from IEEE Xplore. Restrictions apply.

[9] Koteswara Rao Ballamudi ”Road Accident Analysis and Prediction using Machine Learning Algorithmic Approaches”, 2021.

[10] Daniel Santos, Jose´ Saias ”Machine Learning Approaches to Traffic Accident Analysis and Hotspot Prediction”, 2021.

[11] Shweta, J Yadav, K Batra, A K Goel, ”A Framework for Analyzing Road Accidents Using Machine Learning Paradigms”, 2021.

[12] Yunzhi Shi, Raj Biswas ”Predicting Road Accident Risk Using Geospa- tial Data and Machine Learning”, 2021.

[13] Monisha Lakshme Gowda ”Traffic Accidents Prediction using Ensemble Machine Learning Approach”, 2020.

[14] Priyanka A. Nandurge, Nagaraj V. Dharwadkar ” Analyzing Road Accident Data using Machine Learning Paradigms”, 2017.

[15] Alyssa Ditcharoen, Bunna Chhour ”Road Traffic Accidents Severity Factors”, 2018.

Authorized licensed use limited to: University of North Texas. Downloaded on June 29,2023 at 00:18:16 UTC from IEEE Xplore. Restrictions apply.