INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Evaluation of Severity Index of Road Casualties on Mangawan to Chakghat Section of NH-30 Using Machine Learning Approach

Alkesh Kumar Khade[1], Dr. Rakesh Mehar[2]

[1]M. Tech Scholar, [2]Assistant Professor, Department of Civil Engineering, Samrat Ashok Technological Institute, Vidisha, M.P. India

Abstract: Each year, over 1.35 million people die in road accidents, and many more are seriously injured, according to the WHO. Predicting the severity of these accidents is vital for improving prevention and response plans. This paper proposes a new method to predict the severity of road accidents using various factors. This research focuses on road accidents at Sohagi Ghat on NH-30. The study looks at accident types and locations along a 3.67 Kilometer section of steep, challenging terrain. A safety audit was conducted on black spots and it revealed that 324 accidents occurred over five years, making this section particularly dangerous. Special safety measures are needed to reduce the risk in this area. The paper aims at forecasting of upcoming highway accidents through moving average approach, exponential smoothing approach and linear forecasting approach using the previous year's data sets. With the help of Kaggle Data Directory, precise accident data for visualization is formed and then visualization of road accidents is performed on Python programming language-based Anaconda Navigator's Jupyter Notebook. While accident severity index of Sohagi Ghat NH-30 is also carried out and some results of machine learning based visualization and some forecasting-based results are driven out and compared with each other. The study calculated accident severity and used Logistic Regression to train a model with 0.85 accuracy value. The precision, recall, and F-1 score of the model were perfect. Injury severity percentages were 10.8% fatal, 48.2% serious, and 41% minor. The Accident Severity Index for the NH-30 Sohagi Ghat location was found to be 0.211.

Keywords: Accident Severity, WHO, Road Safety Audit, Kaggle, Machine Learning, Anaconda Navigator, Severity Index.

I. INTRODUCTION

Road accidents are a significant issue in both developing and developed countries, with developing nations experiencing higher fatalities and serious injuries. Despite implementing road safety regulations, the frequency of incidents remains high, with over 90% caused by driver inattention and careless behaviour. Road traffic accidents (RTAs) pose major health, economic, and social hazards. In India, 2022 saw 1,68,491 road accident fatalities, an increase from 1,53,972 in 2021. The 18 to 45 age group was most affected. India's rapid motorization and urbanization, driven by economic growth, contribute to high accident rates. In 2022, 4,61,312 accidents resulted in 1,68,491 deaths and 4,43,366 injuries, marking increases of 11.9%, 9.4%, and 15.3% respectively from 2021.

According to the NCRB, India's road network spans over 48,65,000 km, contributing 4.8% to the GDP. Understanding road safety's impact on the economy is crucial. Traffic accident fatalities increased from under 20,000 in 1970 to over 1,40,000 by 2016. In 2022, states and UTs reported 4,61,312 road accidents, resulting in 1,68,491 deaths and 4,43,366 injuries. Compared to 2021, traffic accidents increased by 11.9%, fatalities by 9.4%, and injuries by 15.3%. On average, there are 1,264 collisions and 462 fatalities per day, or 53 collisions and 19 deaths per hour nationwide.

The COVID-19 pandemic and the subsequent lockdown, especially between March and April 2020, led to a significant and unexpected decline in road accidents. Although 2019 crash data showed similar patterns, the pandemic in 2020 caused a notable reduction. In 2022, Tamil Nadu had the highest number of vehicle collisions (64,105 or 13.9%), followed by Madhya Pradesh (54,432 or 11.8%). Uttar Pradesh had the most road fatalities (22,595 or 13.4%), followed by Tamil Nadu (17,884 or 10.6%). Section 5 of the study provides detailed state-by-state crash statistics.

Recent road safety research has focused on evaluating the severity of car crashes using two main methods: predictive machine learning (ML) techniques and traditional statistical methods. While traditional models rely on established theories and may struggle with multi-dimensional and nonlinear relationships, Machine Learning algorithms in identifying complex relationships and handling large datasets.

ML approaches are increasingly favoured for their flexibility, accuracy, and ability to predict outcomes based on highway user demographics and historical data, facilitating the development of effective interventions to mitigate road accidents.

II. ABOUT SITE

Sohagi Ghat, located in Rewa district, serves as the administrative headquarters of Teonthar Tehsil in northeastern Madhya Pradesh. The stretch starts at a Y-junction with NH-135 near Mangawan village, passing through various villages like Gangave, Bela, Tikuri, Garh, Kalwari, Katra, Suhagi, and Chandai, terminating at the Madhya Pradesh – Uttar Pradesh Border near Chakghat colony. The research focuses on a critical 3.67 km section from 55.100 km to 58.770 km for further study. The road section of Sohagi Ghat presents challenges due to its steep terrain and varying gradients from -0.40% to 5.99%. Over the past 5 years, 324 accidents, resulting in 35 deaths, have occurred on this stretch. With 9 sharp curves and 4 reverse curves spanning 52 km, including 5 sharp curves and 1 reverse curve in the designated Blackspot area, special safety measures are necessary to mitigate risks.

Figure 1: NH-30 (Mangawan-Chakghat Section) Sohagi Ghat Alignment in Google Map

Table 1: Curve Details of the Site

S. No.	Chainage	Curve Type	Side
1	Km 57+250	Sharp Curve	LHS
2	Km 56+800	Sharp Curve	RHS
3	Km 55+700	Sharp Curve	LHS
4	Km 55+250	Sharp Curve	LHS
5	Km 55+800	Sharp Curve	LHS
6	Km 57+900	S-Curve	-

III.   METHODOLOGY

The model would be developed through accident-related data, which may assist in understanding the features of multiple components, including as the climate, lighting, highway conditions, and behaviour among drivers. This can assist people in evaluating which safety measures are most useful for minimizing accidents. The model can be used to identify risk factors that can be used to reduce risk as well as statistically significant characteristics that can be used to forecast the likelihood of collisions and injuries.
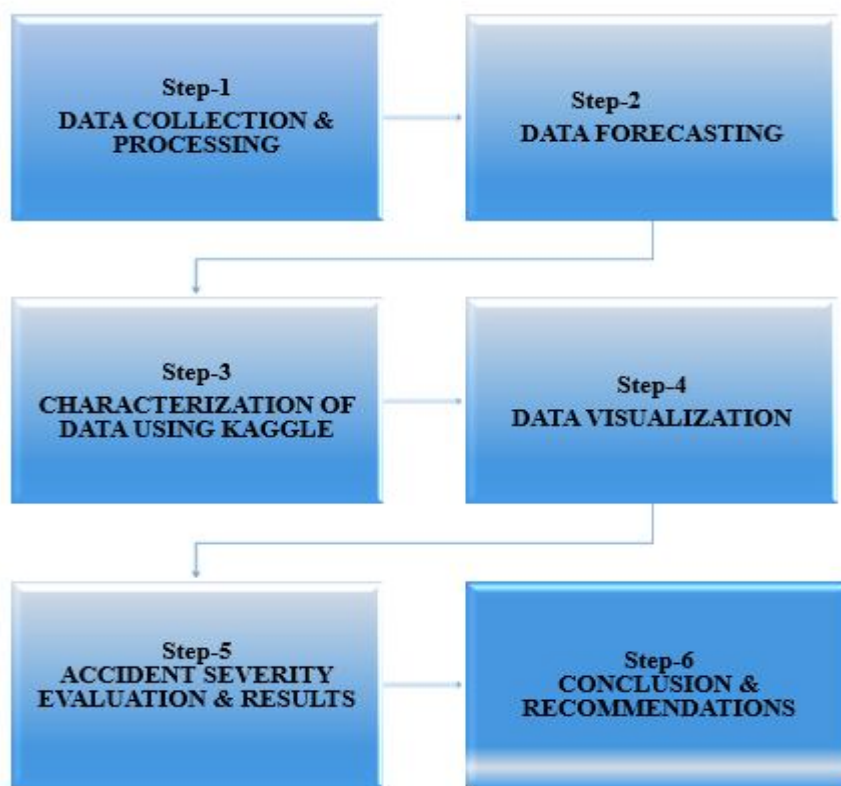

Figure 2: Step by step chart of proposed methodology

For the research work various method options are available but in current study on road accidents linear regression analysis method for data forecast is adopted & for the preparation of model for evaluation of accident severity python tool is used.

A total of 324 accident happened in past 5 years from 2019 to 2023. It is found that there were 35 numbers of casualties occurred on NH-30 near Sohagi Ghat on Mangawan to Chakghat section, also it is recorded that 156 Nos. of people went through hospitalization in the case of serious injuries and 133 Nos. of persons went through minor injuries in which there were no need of hospitalization.

Table 2: Category wise list of injured persons in accident

S. No	Year	Total No. of Accidents	Minor Injury	Hospitalized in Serious Injury	Fatal
1	2019	30	22	8	0
2	2020	62	38	19	5
3	2021	66	39	20	7
4	2022	86	16	50	20
5	2023	80	18	59	3
	Total	324	133	156	35

A. Accident Forecasting by Linear Forecast Approach

This forecasting is based on Microsoft Excel framework, in which past records of the accident is forecasted through some formula and then rounded for making it the countable number. As we can see in table 2 & 3.

Table 3: Forecasted Nos. of Accident by linear forecasting

Years	Total Nos.
2024	124
2025	142
2026	160
2027	178
2028	196
2029	214
2030	232
2031	250
2032	268
2033	286
2034	304
2035	322
2036	340
2037	358
2038	376
2039	394
2040	412
2041	430
2042	448
2043	466
2044	484
2045	502

B. Characterization of data using Kaggle

Kaggle is the world's largest data science community, offering powerful tools and resources to help people achieve their data science goals. It's a website owned by Google where ML engineers and data researchers can access datasets, collaborate with others, and compete in events to solve data mining challenges. Users typically start by downloading datasets in CSV format from Kaggle, which is commonly used for data analysis and manipulation in programs like Microsoft Excel. As we can see in the figure 4.
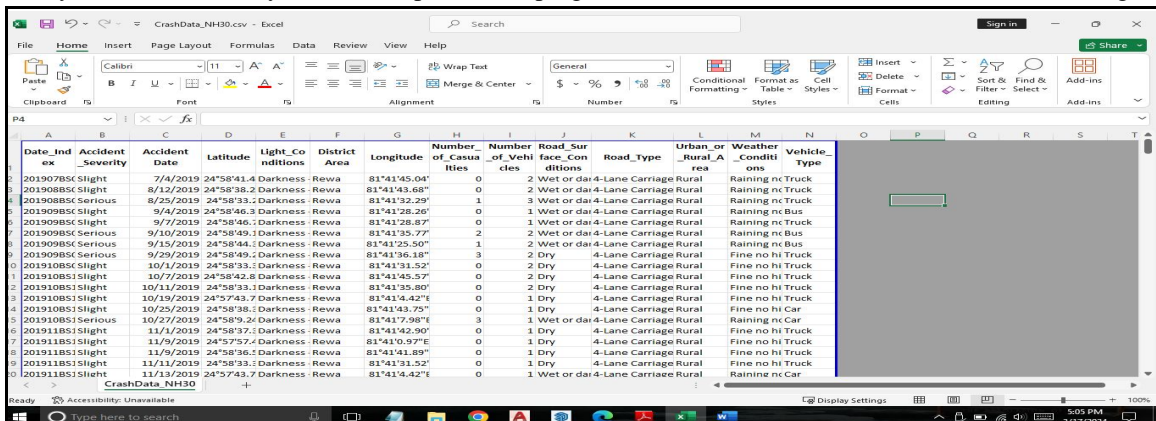

Figure 3: Accident Data categorization in CSV format using Kaggle

C. Visualization of Accident through Machine Learning Approach

Once data collection, forecasting, and arrangement are done using Kaggle, it's time for data visualization. The data solely focuses on traffic accidents and is collected from the project's consultant. Python-based Anaconda Navigator's Jupyter Notebook is then used to prepare data models using various algorithms. Different algorithms represent different types of models, and the type of data prepared determines the visualizations on the screen. Machine learning involves working with vast datasets and algorithms, starting with Interactive Python Notebook (ipynb) files that store code, equations, visualizations, and narrative texts. These documents are prepared along with "Comma Separated Value" format and depict the road accident model.

1) First step is to Open Anaconda Navigator in system.
2) Then Select the Jupyter Notebook Version 6.4.8 on system.
3) After opening the Jupyter Notebook prepared dataset of the project is loaded in the directory.
4) When inputs are given in notebook then it visualizes the output given below. See figure no. 4
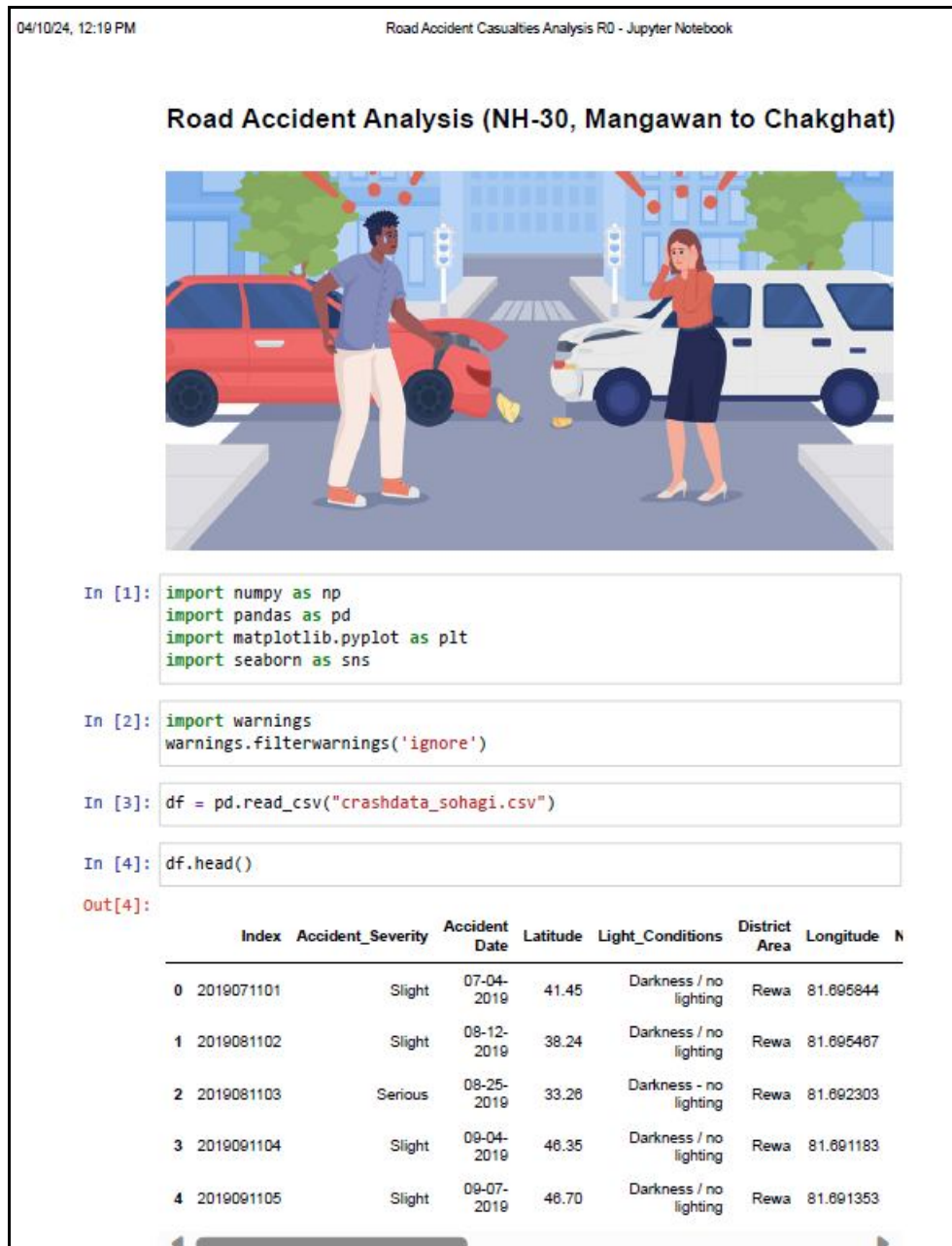

Figure 4: Importing a csv file into Data frame

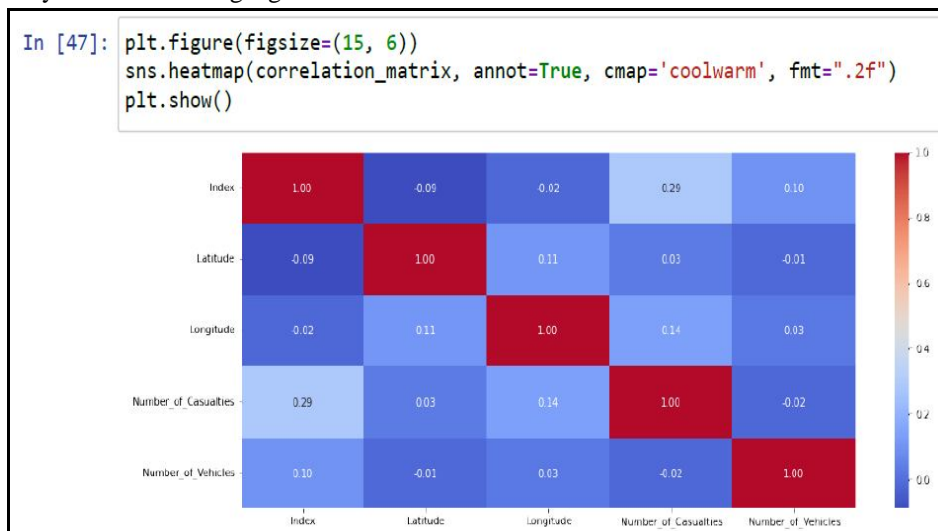5) After applying so many machine learning algorithms the Correlation Matrix of the accident data is formed.

Figure 5: Visualization of Correlation matrix on data frame

6) Then model is trained by Logistic Regression algorithm. It can be seen in figure 6.

Figure 6: Data set training by Logistic Regression using Pandas Library

7) In the end of the visualization of the above given accident dataset accuracy score, Precision, F-1 Score and Recall value is obtained. Shown in Figure 8.

IV. RESULT & ANALYSIS

1) *Accident Severity Index Calculation through Rational methodology for formulation of a road safety improvement program in India: -*

Crashes are unplanned incidents resulting in harm or damage, while calamities are unexpected occurrences that may or may not cause harm. Incidents may occur due to any fault, but not all qualify as accidents. Severity is determined based on the most severe injuries:

- Fatal: If at least one person dies in the collision.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

- Serious or Grievous: Injuries requiring hospitalization and police station reporting.
- Minor: Injuries that don't need hospitalization or police reporting.

For the estimation study of Accident Severity Index (ASI), Weightages for the fatal crashes is taken as 0.658, while for the Grievous Injury or Serious Injury weightage taken is 0.299 and for minor injury weightage taken is 0. 024.Based on fatal accidents, the equation can be termed as: -

Where,

Nf= No. of fatal accidents at the location;

Wf = Weightage assisted to fatal accident;

Ng = No. of grievous accident or serious injured persons;

Wg = Weightage assisted to grievous accident or serious injured person.

Table 3: Weightage classification for Crashes

S. No.	Injury Categorization	Weightage**
1	Fatal/Casualties/Death	0.658
2	Grievous Injury / Serious Injury	0.299
3	Minor Injury	0.024
4	No Injury	0.019

(Source: Taken from a rational methodology for formulation of a road safety improvement program in India by Dr. Rakesh Mehar)

$$ASI_F = N_f * W_f + N_g * W_g,$$
$$ASI = 35*0.658+156*0.299$$
$$ASI = 69.674$$

Thus, Final ASI of each location will be

$$ASI = \frac{1}{2} *\{ASI\ (n) + ASI\ (p)\}$$
$$ASI = 0.21$$

Where, **ASI (n)** = Nos. of fatal/grievous accidents; **ASI (p)** = Nos. of fatal/grievous persons involved in accident.

2) Accident Severity through Machine Learning: Following are some results shown below:

- Severity in pie chart on a scale of 100%, out of them injured persons belongs to Fatal or Casualty 10.8%, Grievous or Serious Injury 48.2%, and Minor injuries 41.0% respectively.
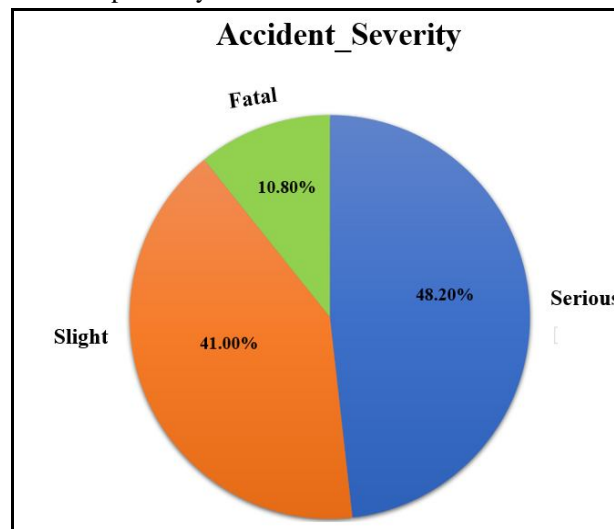

Figure 7: Severity through Machine Learning in Pie Chart

- Accuracy score predicted during machine learning algorithm is 0.85, Support Value is 65, Precision value, recall value and F-1 Score is 1 respectively.

```
Accuracy: 0.85
Confusion Matrix:
[[27  0  0]
 [ 0 28  3]
 [ 0  7  0]]
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        27
           1       0.80      0.90      0.85        31
           2       0.00      0.00      0.00         7

    accuracy                           0.85        65
   macro avg       0.60      0.63      0.62        65
weighted avg       0.80      0.85      0.82        65
```

Figure 8: Accuracy, Precision Value, Recall value, F-1 Score & Support Value through Machine Learning

V.   CONCLUSION

Starting from process-by-process data collection to result evaluation several steps are involved in road accident prediction model; We used machine learning framework. In machine learning, we have a plenty of techniques of machine learning for better data visualization of road accident and road accident severity. In general, we graphically visualize the data sets by the using so many algorithms.

It will be essential to develop the ability to identify situations in which serious injuries are most likely to happen in order to establish techniques that will enhance road user safety. Based on data gathered between 2019 and 2023, this analysis revealed specific factors that have a major influence on the level of severity of accidents/ fatalities caused by vehicular accidents in Sohagi Ghat on NH-30. The most important factors were the types of vehicles, the weight of the vehicle, Geometrics of the accident spot, accident's geographic coordinates, weathering conditions, and the permissible speed. We applied the simplest form of machine learning. Some of the conclusions are given as follows: -.

Logistic Regression, As an outcome of this research work in terms of accident severity percentage which is further classified in three categories, for fatal category the percentage of severity of accident is 10.8%, for the serious/ grievous injured category the percentage of severity of accident is 48.2% and for the normal injury which do not required to be hospitalized has the percentage of 41.0%, while in the other way if we calculate the road accident severity index then our result will be 0.21.

When it came to predicting the severity of injuries occurred accidents at Sohagi Ghat on NH-30, the Logistic Regression showed the accuracy with 0.85, which is quite effective to highlight the accidents.

Accident Forecasting for the upcoming 20 years is studied using Linear Forecasting approach. In which accident trend is rising year by year. While comparing the forecast rate of accidents for year **2045** to base year **2019** the accident rate is rises to **11.16 times**.

VI.   RECOMMENDATIONS FOR THE FUTURE WORK

1) To improve future outcomes and reduce fatalities, a safety program involving the Gram Panchayat and local administration should be organized to raise road safety awareness. Public consultations should include local residents. Additionally, schools and Anganwadi centers should educate children about road safety.

2) Research shows that various models visualize road accident crash severity. The logistic regression model relates predictor variables to binary outcomes, while the Random Forest (RF) model uses averages or votes for aggregated probabilities. Naïve Bayes (NB) predicts based on class probabilities, K-Nearest Neighbors (KNN) uses the majority class of nearest neighbors, and Neural Networks (NNN) mimic the brain's structure with layers of interconnected neurons and use SOFTMAX for classification. Probabilistic models yield varying results, whereas deterministic models like SVM, Linear Regression, Decision Trees, GBM, Rule-Based Models, and LDA do not rely on probability calculations. Deterministic models should be trained for better incident visualization based on data set accuracy. Further study of this research needs to be implemented through different machine learning model techniques.

3) In this thesis, we incorporate numerous road user parameters, but including additional characteristics such as age, sex, experience, behavior, and health data would enhance model visualization. Similarly, expanding datasets to include road characteristics like maintenance, conditions, and speed limits, along with ecological factors, would improve model training and visualization.

4) To reduce road accident injuries, this research recommends trimming the ghat section to a minimum width of 7.0m towards the hillside, per IRC-73:2018 standards. This includes a 3.5m lane, 1.20m for an RCC drain, and 2.30m for a utility corridor during the execution of the site in reconstruction.

a) Hill Trimming: Trimming the hillside will increase Stopping Sight Distance (SSD) and Intermediate Sight Distance (ISD), reducing accident frequency.

b) Road Furniture & Safety Signage: Install traffic signage, Chevron signs, gantries, cautionary signs, rumble strips, road markings, raised pavement markers, delineators, anti-glare sheets, and solar high mast lighting on steep curves for improved visibility.

c) Passing Places: Provide passing places on the project section as per IRC-SP:48-1998 and IRC:52-2019 for vehicle breakdowns and to facilitate smoother traffic flow, marked with dashed lines for safe overtaking.

REFERENCES

[1] Amir Elalouf et al (2023). "Developing machine-learning-based models to diminish the severity of injuries sustained by pedestrians in road traffic incidents.", https://doi.org/10.1016/j.heliyon.2023.e21371

[2] Aamir Ali Khan et al. (2023, February). "Road Accident Analysis and Safety Audit: A Case Study of MDR-132", https://doi.org/10.26552/com.C.2023.026

[3] Antoine Gatera et al. (2023, May). "Comparison of random forest and support vector machine regression models for forecasting road accidents", https://doi.org/10.1016/j.sciaf.2023.e01739

[4] Balasubramani S. et al. (2023, December). "DDDS: Driver decision support system based on the driver behaviour prediction to avoid accidents in intelligent transport system", https://doi.org/10.1016/j.ijcce.2023.12.001

[5] Highway engineering by S.K. Khanna, C.E.G. Justo & A. Veeraragavan. (10th revised Edition) (2017, January)

[6] https://en.wikipedia.org/wiki/Machine_learning

[7] https://www.kaggle.com/datasets/nezukokamaado/road-accident-casualties-dataset

[8] https://towardsdatascience.com/regularization-in-machine-learning-76441ddcf99a

[9] https://intellipaat.com/community/368/how-to-interpret-loss-and-accuracy-for-a-machine-learning-model

[10] IRC 37 - 2018 for Design of Flexible Pavements (Fourth Revision) (without the CD)

[11] IRC 58 – 2015 Guidelines for the Design of Plain Jointed Rigid Pavements for Highways (Fourth Revision) (without CD). Published by: Indian Road Congress.

[12] IRC: SP:48-1998 Hill Road Manual, Published by: Indian Road Congress.

[13] IRC: SP:84-2019. Manual of Specifications and Standards for Four Laning of Highways (Second Revision). Published by: Indian Road Congress.

[14] IRC:35-2015 - Code of Practice for Road Markings. (Second Revision). Published by: Indian Road Congress.

[15] IRC:52-2019 Guidelines for the Alignment Survey and Geometric Design of Hill Roads (Third Revision) Indian Roads Congress 2019. Published by: Indian Road Congress

[16] IRC:67-2022- Code of Practice for Road Signs (Third Revision). Published by: Indian Road Congress

[17] Izuchukwu Chukwuma Obasi et al. (2023, July). "Evaluating the effectiveness of machine learning techniques in forecasting the severity of traffic accidents.", https://doi.org/10.1016/j.heliyon.2023.e18812

[18] Md Adilur Rahim (2021, May), "A deep learning-based traffic crash severity prediction framework", https://doi.org/10.1016/j.aap.2021.106090

[19] Mireille Megnidio- Tchoukouegno et al. (2023, January), "Machine Learning for Road Traffic Accident Improvement and Environmental Resource Management in the Transportation Sector", https://doi.org/10.3390/su15032014

[20] Nicola Tamascelli et al. (2022, June), "Learning from major accidents: A machine learning approach", https://doi.org/10.1016/j.compchemeng.2022.107786

[21] Official Website of Public Works Department, Uttar Pradesh, Available from https://uppwd.gov.in/site/writereaddata/siteContent/20190424201227766620.pdf

[22] Praveen Maghelal et al. (2023, October). "Severity of vehicle-to-vehicle accidents in the UAE: An exploratory analysis using machine learning algorithms" https://doi.org/10.1016/j.heliyon.2023.e20694

[23] Pedestrian Safety, A road safety manual for decision-makers and practitioners, in: Pedestrian Safety: a Road Safety Manual for Decision-Makers and Practitioners, second ed., World Health Organization, Geneva, 2023. https://apps.who.int/iris/bitstream/handle/10665/367419/9789240072497-eng.pdf?sequence=1

[24] Philippe Barbosa Silva et al. (2020, October). "Machine learning applied to road safety modeling.", https://doi.org/10.1016/j.jtte.2020.07.004

[25] Piotr Gorzelanczyk (2023, March). "Using neural networks to forecast the number of road accidents in Poland taking into account weather conditions." https://doi.org/10.1016/j.rineng.2023.100981

[26] Rakesh Mehar et al. (2013), "A Systematic Approach for Formulation of a Road Safety Improvement Program in India", doi: 10.1016/j.sbspro.2013.11.199

[27] Shuli Wang et al. (2024, March). "Geographically weighted machine learning for modeling spatial heterogeneity in traffic crash frequency and determinants in US." https://doi.org/10.1016/j.aap.2024.107528

[28] Shubham Pandey et al/ (2022, May). "Identification and Prioritization of Black Spots in Hilly Road Segment Using Accident Severity Index Method", (http://creativecommons.org/licenses/by/4.0/)

[29] Slava Birfir et al. (2023, April). "Building machine learning models for reducing the severity of bicyclist road traffic injuries" https://doi.org/10.1016/j.treng.2023.100179

[30] Specifications for roads and bridges by MORTH (Ministry of Road Transport and Highways) - includes details of road construction and maintenance. All highways' projects in India are following MORTH Specifications for construction work. https://archive.org/details/gov.in.mort.250.2013

[31] Stefan Candefjord et al. (2021, September), "On Scene Injury Severity Prediction (OSISP) machine learning algorithms for motor vehicle crash occupants in US", https://doi.org/10.1016/j.jth.2021.101124

[32] Suleyman Ceven et al. (2024, March). "Traffic accident severity prediction with ensemble learning methods." https://doi.org/10.1016/j.compeleceng.2024.109101

[33] The United Arab Emirates Government Portal Information and Services: Road Safety, United Arab Emirates, Abu Dhabi, 2022.

[34] Venkataraman Shankar et al. (1995, June), "Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies". https://doi.org/10.1016/0001-4575(94)00078-Z

[35] World Health Organization, Road Traffic Injuries, Fact Sheet, 2018. http://www.who.int/en/news-room/fact-sheets/detail/roadtraffic-injuries. (Accessed 20 March 2022).

[36] World Health Organization, World health statistics. http://apps.who.int/iris/bitstream/handle/10665/272596/9789241565585-eng.pdf?ua=1, 2018. (Accessed 20 March 2022).

[37] World Health Organization, Global Status Report on Road Safety, 2018.
https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/.(Accessed 20 March 2022), 2018c.

[38] World Health Organization, Road traffic injuries. http://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries, 2018. (Accessed 17 January 2022).

[39] World Health Organization, The Department of Social Determinants of Health, Safety and Mobility Unit, https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/global-status-report-on-road-safety-2023

[40] Xian Wen et al. (2021, July), "Applications of machine learning methods in traffic crash severity modelling: current status and future directions", https://doi.org/10.1080/01441647.2021.1954108

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ◎ (24*7 Support on Whatsapp)