Proceedings of the 8th International Conference on Computing Methodologies and Communication (ICCMC-2025)
IEEE Xplore Part Number: CFP25K25-ART; ISBN: 979-8-3315-1211-8

2025 8th International Conference on Computing Methodologies and Communication (ICCMC) | 979-8-3315-1211-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICCMC65190.2025.11140749

# Road Accident Analysis and Severity Prediction by using Adaptive Regularized Extreme Learning based Models

1. Dr.N.K Anushkannan, Professor & Head, Department of ECE, Kathir College of Engineering, Coimbatore, Tamilnadu, India, anushkannan@kathir.ac.in

2. Prof. Raja G. V, Assistant Professor, Department of Electronics and Communication Engineering, Sri Sairam College of Engineering, Bengaluru, Karnataka, India, gvraja.ece@sairamce.edu.in

3. Jayalakshmi V, Assistant Professor, Department of Science and Humanities (English), RMK College of Engineering and Technology, Chennai, Tamil Nadu, India, jayalakshmi@rmkcet.ac.in

4. Dr. A. Rajavel, Department of Electrical and Electronics Engineering, Kamaraj College of Engineering and Technology, Virudhunagar, India, rajavelaiyamperumal@gmail.com

5. Dr. V. Navin Ganesh, Assistant Professor (Selection Grade), Department of Civil Engineering, PSG Institute of Technology and Applied Research, Coimbatore, Tamil Nadu, India, navinganesh@psgitech.ac.in

6. Dr Dileep Pulugu, Associate Professor, Department of Computer Science and Engineering, St Peter's Engineering College, Hyderabad, Telangana, India, dileep.p505@gmail.com

***Abstract* – Traffic collisions represent a significant public health issue, resulting in extensive human suffering and considerable economic and social burdens. These occurrences not only generate substantial medical and rehabilitation costs but also lead to production declines, property damage, and elevated insurance premiums. Furthermore, the enduring consequences on victims, families, and communities exacerbate their effects. Precise study of traffic accidents and prediction of their severity are crucial for alleviating these expenses. This study underscores the significance of data preprocessing, particularly normalisation, to improve dataset quality. Key variables were identified by clustering, chi-square tests, Cramer's V, and predictor importance, subsequently organised into effective groups for efficient analysis. Conventional ELM sometimes use L2 regularisation to mitigate overfitting; yet, the 2-RELM model's dependence on manually chosen regularisation parameters is ineffective. To resolve this, it present ARegELM, an adaptive model that substitutes the static regularisation factor with a dynamic function, facilitating automatic selection. The ARegELM model demonstrated an impressive accuracy of 99.28% in forecasting accident severity. These results indicate that ARegELM improved predictive performance and model usability, hence facilitating better informed decision-making in traffic safety and accident prevention.**

***Keywords*—Road Traffic Accident (RTA),**

## I. INTRODUCTION

Analysis of traffic accidents has recently attracted a lot of interest from academics who are trying to pin down the causes of these collisions. Nevertheless, the vast majority of research methodologies rely on statistical records or conduct basic surveys using questionnaires or interviews. However, adopting such rudimentary methods will not lead to a superior and error-free solution. The most perplexing aspect is how these conventional methods of study struggle to discern behavioural aspects in road accidents. Because accidents are often sudden and unexpected, making it hard to observe them closely[1]. Various national and international organisations have developed strategies, programs, and regulations to lessen the likelihood of accidents. There is currently no real-time warning system that can advise users on the likelihood of an accident, despite several efforts to make roadways safer. Road signs that inform drivers of potential dangers, such as those that indicate high collision zones, locations that are particularly slippery in the rain, speed limits, etc., are common in many nations across the globe. One essential element to improving road safety is the forecast of accident rates on road linkages, in addition to road signs. In order for transportation regulations to be effective, the prediction model investigates the connections between different types of crash injuries and variables such driver behaviour, vehicle attributes, road geometry, road-environmental conditions, and accident causes. The frequency of crashes and their severity can be modelled using data on vehicular accidents[2].

But traffic accidents are unpredictable and spatially heterogeneous, so the elements that contribute to their severity could differ depending on where they happen. For accurate crash prediction and the efficient implementation of appropriate countermeasures, it is essential to conduct a local investigation of the risk variables[3]. One of the most common issues in public health, road traffic accidents cause a lot of suffering for people and have a major financial and social effect on society as a whole. Damage to property, lost productivity, and higher insurance premiums are additional expenditures that arise as a result of these incidents, in addition to the obvious medical and rehabilitation expenses. This load is made further heavier by the cumulative impact on victims, their families, and communities in the long run. In order to lessen the social effects of traffic accidents, accurate accident severity prediction becomes crucial. Ethical considerations about the use of personal information, model complexity, data availability and quality, and other similar issues are obstacles to using ML and data analysis to forecast the severity of accidents. Regardless, there are significant ramifications for traffic and public safety management in this area. Because traffic accidents are complex and ever-changing, current models that account

for factors such as human behaviour, road conditions, weather, and vehicle types do not yet provide adequate results.

How the rest of the paper is structured is as follows. Include all relevant work in Section II; It presented ARegELM in Section III. In Section IV, the results and comments are presented. The paper was ended and future areas were considered in Section V.

## II. Literature Survey

Recent efforts have explored the possibility of applying DL models to the investigation of traffic accidents. The goal of this project is to develop a model for the real-time risk assessment of accidents using data collected from mobile phones and accident records. Using satellite photos and accident records, a CNN was trained to create a map showing the likelihood of traffic accidents[4].But, it's important to note that neither of these studies provides more than a time-invariant accident risk map. In terms of safety-related real-time accident prediction, this is inadequate[5]. An LSTM model was utilised in a recent study to foretell the likelihood of future traffic accidents. Nevertheless, the spatial heterogeneity of the data is not addressed in this work; it is merely a basic model application to data on traffic accidents. Additionally, only urban regions are considered for the study[6]. In contrast, they offer some fresh suggestions in our study to deal with the geographical heterogeneity problems that arise from our data. On top of an LSTM neural network, they apply these concepts. This is the first study that are aware of that uses a DL model to tackle the problem of traffic accident predictions with geographical heterogeneity[7]. A lot of people are now concentrating on real-time traffic accident prediction due to advancements in ML. Lv used the k-nearest neighbour method to forecast traffic accidents and selected feature variables according to the Euclidean metric[8]. After amassing massive amounts of data on traffic accidents on Seoul's highways, Park developed a prediction pipeline using k-means clustering and logistic regression. They have built a Stack denoise Autoencoder using Japanese human movement data to determine the real-time traffic risk. The failure to account for important variables like traffic, weather, and air quality is a drawback of these studies[9]. The model's ability to make predictions may be impaired in the absence of certain details.

ML techniques, such as ANN, SVM, and DL models, have been utilised to address the shortcomings of statistical methodologies. These methods are capable of processing large amounts of multidimensional data and have found applications in data analytics for a variety of traffic safety issues[10]. Furthermore, ML has been regarded as a suitable mathematical model for traffic safety due to its adaptability in modelling, capacity to learn and generalise, and strong prediction abilities. For a long time, ANN and BNN models have been used to examine traffic safety issues instead of the more popular Poisson or NB regression models[11]. This is because these models assume a predefined relationship between the dependent and independent variables, which can lead to inaccurate estimation if broken. The outcome variables are predicted differently by ANN and BNN models, despite the fact that both use multilevel networks[12]. Accurately predicting the impact of traffic accidents requires. In order to evaluate RF analysis approaches, MNL, NNC, SVM, and others. For more serious accidents, NNC outperforms RF, SVM, and MNL in terms of total prediction performance. Using a LSTM model, they can forecast the likelihood of traffic accidents occurring in a certain area within a specified time frame[13].

The regularisation factor typically requires continual adjustment based on varying tasks. Motivated by the significance of the regularisation factor and the associated efforts, they presented an ARegELM by substituting the regularisation factor with a regularisation function. The goals of the aforementioned four studies are to enhance ARegELM performance and to pick the regularisation factor adaptively.

## III. Proposed System

Worldwide, the number of people hurt or killed in car accidents is steadily climbing. The major elements that impact the probability of road traffic fatalities have been the subject to study for several decades. Worldwide, RTA rank high among the most prevalent causes of injury and death, affecting individuals of all ages. Accidents involving motor vehicles are unpredictable, haphazard occurrences that can happen anywhere and at any moment. The enormous societal costs associated with road accidents have motivated researchers to seek out new techniques to learn more about what causes accidents.

The data set has been compiled from handwritten records of road traffic accidents from 2017 to 2020. All sensitive information has been omitted during data encoding, resulting in 32 characteristics and 12,316 instances of the accident [14].

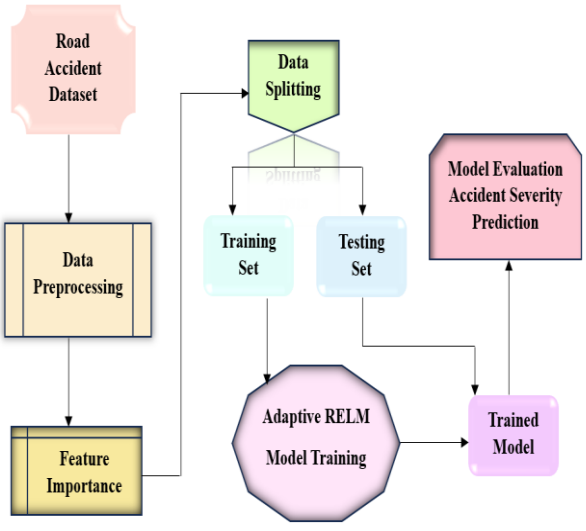


Fig. 1. The Block Diagram for Proposed Model

The classifiers and datasets used for road accident and severity prediction will be covered in this section. Figure 1 shows the proposed data and work flow technique for this study [15].

## V. CONCLUSION

Accidents involving motor vehicles have recently emerged as one of the world's leading killers, ranking tenth among all causes of mortality. It has grown into a big issue in our nation as a result of the staggering annual number of traffic accidents. It is deeply disturbing and unacceptable to permit citizens to lose their lives in traffic accidents. Therefore, a thorough evaluation is necessary to deal with this overloaded scenario. Injuries and fatalities caused by accidents are going to be a worldwide issue. Since the beginning of the motor age, about a century ago, roadway security has been a major issue. Preprocessing employed normalisation to enhance dataset quality. Specifically, the most significant variables are identified using techniques such as clustering, chi-square, Cramer's V, and predictor significance, followed by an ideal grouping of the variables to streamline the analysis. In basic ELM, it is customary to incorporate a penalty term of 2 to mitigate overfitting. To automatically determine an appropriate regularisation factor, we introduced ARegELM by substituting the regularisation factor with a function. The accuracy of the Proposed ARegELM model grew to 99.28%.

REFERENCES

[1] M. F. Labib, A. S. Rifat, M. M. Hossain, A. K. Das, and F. Nawrine, "Road Accident Analysis and Prediction of Accident Severity by Using Machine Learning in Bangladesh," *2019 7th Int. Conf. Smart Comput. Commun. ICSCC 2019*, no. June, pp. 1–5, 2019, doi: 10.1109/ICSCC.2019.8843640.

[2] S. Ahmed, M. A. Hossain, M. M. I. Bhuiyan, and S. K. Ray, "A Comparative Study of Machine Learning Algorithms to Predict Road Accident Severity," *Proc. - 2021 20th Int. Conf. Ubiquitous Comput. Commun. 2021 20th Int. Conf. Comput. Inf. Technol. 2021 4th Int. Conf. Data Sci. Comput. Intell. 20*, pp. 390–397, 2021, doi: 10.1109/IUCC-CIT-DSCI-SmartCNS55181.2021.00069.

[3] A. Jamal *et al.*, "Injury severity prediction of traffic crashes with ensemble machine learning techniques: a comparative study," *Int. J. Inj. Contr. Saf. Promot.*, vol. 28, no. 4, pp. 408–427, 2021, doi: 10.1080/17457300.2021.1928233.

[4] N. Fiorentini and M. Losa, "Handling Imbalanced Data in Road Crash Severity Prediction by Machine Learning Algorithms," *Infrastructures*, vol. 5, no. 7, 2020, doi: 10.3390/infrastructures5070061.

[5] F. Davoudi Kakhki, S. A. Freeman, and G. A. Mosher, "Evaluating machine learning performance in predicting injury severity in agribusiness industries," *Saf. Sci.*, vol. 117, no. April, pp. 257–262, 2019, doi: 10.1016/j.ssci.2019.04.026.

[6] K. Kodepogu, V. B. Manjeti, and A. B. Siriki, "Machine Learning for Road Accident Severity Prediction," *Mechatronics Intell. Transp. Syst.*, vol. 2, no. 4, pp. 211–226, 2023, doi: 10.56578/mits020403.

[7] Z. Yuan, X. Zhou, and T. Yang, "Hetero-ConvLSTM: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 984–992, 2018, doi: 10.1145/3219819.3219922.

[8] Y. Wang, H. Zhai, X. Cao, and X. Geng, "Cause Analysis and Accident Classification of Road Traffic Accidents Based on Complex Networks," *Appl. Sci.*, vol. 13, no. 23, 2023, doi: 10.3390/app132312963.

[9] S. Pourroostaei Ardakani *et al.*, "Road Car Accident Prediction Using a Machine-Learning-Enabled Data Analysis," *Sustain.*, vol. 15, no. 7, 2023, doi: 10.3390/su15075939.

[10] V. K. R. Ballamudi, "Road Accident Analysis and Prediction using Machine Learning Algorithmic Approaches," *Asian J. Humanit. Art Lit.*, vol. 6, no. 2, pp. 185–192, 2019, doi: 10.18034/ajhal.v6i2.529.

[11] Z. Zhang, Q. He, J. Gao, and M. Ni, "A deep learning approach for detecting traffic accidents from social media data," *Transp. Res. Part C Emerg. Technol.*, vol. 86, pp. 580–596, 2018, doi: 10.1016/j.trc.2017.11.027.

[12] C. Dong, C. Shao, J. Li, and Z. Xiong, "An Improved Deep Learning Model for Traffic Crash Prediction," *J. Adv. Transp.*, vol. 2018, 2018, doi: 10.1155/2018/3869106.

[13] D. Santos, J. Saias, P. Quaresma, and V. B. Nogueira, "Machine learning approaches to traffic accident analysis and hotspot prediction," *Computers*, vol. 10, no. 12, 2021, doi: 10.3390/computers10120157.

[14] Zhang, Chaohui, Yijing Li, and Tian Li. "A road traffic accidents prediction model for traffic service robot." Library hi tech 40, no. 4 (2022): 1031-1048.

[15] M. Manzoor *et al.*, "RFCNN: Traffic Accident Severity Prediction Based on Decision Level Fusion of Machine and Deep Learning Model," *IEEE Access*, vol. 9, pp. 128359–128371, 2021, doi: 10.1109/ACCESS.2021.3112546.

[16] S. Seid and Pooja, "Road accident data analysis: Data preprocessing for better model building," *J. Comput. Theor. Nanosci.*, vol. 16, no. 9, pp. 4019–4027, 2019, doi: 10.1166/jctn.2019.8288.

[17] J. Y. Lee, Y. G. Yoon, T. K. Oh, S. Park, and S. Il Ryu, "A study on data pre-processing and accident prediction modelling for occupational accident analysis in the construction industry," *Appl. Sci.*, vol. 10, no. 21, pp. 1–23, 2020, doi: 10.3390/app10217949.

[18] Y. Zhang, Q. Wu, and J. Hu, "An Adaptive Learning Algorithm for Regularized Extreme Learning Machine," *IEEE Access*, vol. 9, pp. 20736–20745, 2021, doi: 10.1109/ACCESS.2021.3054483.