International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

OPTIMAL AMBULANCE POSITIONING FOR ROAD ACCIDENTS WITH DEEP EMBEDDED CLUSTERING

[1]MOGULOJI VYSHNAVI,[2]KUDUMULA VIDHYADHAR REDDY,[3]SHAIK SANA FATHIMA,[4]NALLA POOJA,[5]B.GNANESWARI

[1,2,3,4]Students, Department of computer Science And Engineering, Malla Reddy Engineering College (Autonomous),Hyderabad Telangana, India 500100

[5]Assistant Professor, Department of computer Science And Engineering, Malla Reddy Engineering College (Autonomous),Hyderabad Telangana, India 500100

ABSTRACT

Road accidents continue to be a major global concern due to the high number of injuries and fatalities they cause. Timely medical intervention is critical in such cases, and pre-positioning ambulances—rather than dispatching them on demand—can significantly reduce response time and improve outcomes. This study proposes a deep-embedded clustering-based framework to predict optimal ambulance placement locations. Recognizing that geographical factors and patterns play a vital role in the frequency of road accidents, the model incorporates these elements to enhance predictive accuracy. A deep learning-based technique, Cat2Vec, is used to preserve spatial and contextual patterns during model training for real-time performance. The framework is evaluated against traditional clustering methods such as K-means, Gaussian Mixture Models (GMM), and Agglomerative Clustering. Additionally, a novel scoring function is introduced to assess algorithm performance based on real-time response distance and time. The proposed system demonstrates superior results, achieving 95% accuracy through k-fold cross-validation and a novel distance score of 7.581, outperforming conventional methods.

Keywords: Road Accidents, Ambulance Prepositioning, Deep Learning, Healthcare Services, Optimal Location Prediction, Cat2Vec, K-means, GMM, Agglomerative Clustering, Novel Scoring Function, K-Fold Cross-Validation

I.INTRODUCTION

Road accidents are a leading cause of death globally, significantly impacting both children and adults. These incidents result in severe economic and personal losses for individuals, families, and entire nations. Each year, approximately 1.3 million people lose their lives in road crashes, while an additional 20 to 50 million suffer non-fatal injuries—many of which lead to long-term or permanent disabilities. The continuous growth in the number of vehicles, especially in densely populated urban regions, has contributed to an increase in the frequency and severity of road accidents. Without proactive and data-driven interventions, road traffic injuries are projected to become the fifth leading cause of death worldwide by 2030. Despite the alarming statistics, road safety often receives limited attention, and there remains a shortage of systematic methods to address this issue effectively.

Low- and middle-income countries bear the brunt of this global crisis, with over 90% of road traffic fatalities occurring in such regions. Kenya is a compelling case, experiencing over a thousand road crash-related deaths annually. On average, seven out of thirty-five daily casualties result in fatalities. Most of the affected population falls within the 15 to 59 age range—those considered to be the nation's economically productive demographic—thereby hampering national growth and development. As a lower-middle-income country, Kenya has seen an increase in regional trade over the past decade, further contributing to road congestion. According to the National Transport and Safety Authority (NTSA), 2019 witnessed 5,186 minor injuries, 6,938 serious injuries, and 3,572 fatalities due to road traffic accidents.

Timely and accurate emergency response plays a critical role in mitigating the consequences of road accidents. Key measures include rapid medical intervention, effective communication with emergency personnel, and the use of predictive data analytics to identify accident-prone zones. One of the major challenges lies in the delayed arrival of ambulances, which can prove fatal during emergencies. In large urban centers with complex traffic patterns and geographical layouts, strategically pre-positioning ambulances rather than dispatching them on demand can drastically improve response times and potentially save lives. However, this process requires intelligent decision-making systems that can dynamically suggest optimal ambulance locations throughout the day.

To address this need, the present study introduces a novel deep learning-based approach using Deep Embedded Clustering with Autoencoder (DEC-AE) for determining optimal ambulance positioning across urban areas. Unlike traditional clustering techniques, the DEC-AE framework effectively combines deep feature learning and clustering through a joint optimization process. This approach enables the model to reconstruct input data through learned latent features, capturing the most critical dimensions for accurate clustering. The method also supports adaptive clustering by dynamically determining the optimal number of clusters based on data distribution. This enhances the precision of cluster formation and ensures that ambulance locations are suggested based on actual city-wide patterns.

The study utilizes real-time data from Nairobi, Kenya, including road accident reports, road segment information, and weather conditions. Exploratory Data Analysis (EDA) is conducted to identify key features contributing to accident occurrences. To retain meaningful categorical relationships during data preprocessing, a deep embedding technique called Cat2Vec is employed. The effectiveness of the proposed ambulance positioning strategy is validated using a custom-designed Distance Scoring Algorithm, which calculates the proximity between predicted ambulance locations and actual crash sites. Additionally, the performance of the proposed model is benchmarked against traditional clustering algorithms such as K-means, Gaussian Mixture Models (GMM), and Agglomerative Clustering using multiple evaluation metrics.

In summary, this study addresses four major components: (1) performing EDA to identify accident-related features and risk patterns in Nairobi; (2) implementing a deep clustering approach (DEC-AE) to determine optimal ambulance positions while preserving data relationships through Cat2Vec embeddings; (3) introducing a novel Distance Scoring Algorithm to validate predicted locations; and (4) evaluating the framework against baseline clustering algorithms with and without feature selection. This integrated approach offers a promising solution to enhance emergency response efficiency and road safety in urban environments.

II.LITERATURE SURVEY

The Global Status Report on Road Safety (2015) by the World Health Organization (WHO) presents a comprehensive assessment of road safety across 180 countries. This report highlights that over 1.25 million people die annually due to road traffic injuries, establishing them as a leading cause of death worldwide. It emphasizes the disproportionate burden faced by low- and middle-income countries, where 90% of global road deaths occur despite having only 54% of the world's vehicles. The WHO also stresses the significance of road safety legislation, enforcement, vehicle standards, and post-crash response. This report is crucial for shaping international and national policy frameworks. For this study, it provides foundational data on global trends and the urgency of reducing emergency response times through improved ambulance deployment strategies. The findings underscore the importance of predictive models and strategic planning in minimizing fatalities caused by delays in emergency medical care.

Sivakumar and Krishnaraj (2012) explore the role of alcohol-impaired driving in India's road accident statistics, focusing on the challenges in enforcement and prevention. They argue that despite stringent laws, the lack of regular monitoring, low awareness, and inadequate resources lead to high instances of drunk driving. The study presents data showing a significant proportion of accidents in India are attributable to alcohol consumption, especially during weekends and late-night hours. The authors propose solutions such as increased public awareness campaigns, stricter implementation of breathalyzer tests, and dedicated night patrols. This research contributes to a deeper understanding of accident causation, which is essential for developing accurate predictive models. Within the context of ambulance positioning, understanding the temporal and causal patterns of accidents, such as those influenced by alcohol, enables more strategic deployment of emergency responders in high-risk zones during vulnerable time windows.

Baguley's 2001 report highlights the critical importance of comprehensive road accident data systems in traffic safety planning. It emphasizes that systematic and accurate data collection is fundamental for diagnosing road safety issues, identifying accident hotspots, and formulating effective intervention strategies. The paper discusses the role of data in policymaking, road design improvements, and targeted law enforcement. It advocates for standardized data collection methodologies and better utilization of technological tools in capturing accident details. For this study, Baguley's insights validate the need for robust datasets that reflect the spatial and temporal distribution of accidents. Reliable data supports the proposed deep learning models by enhancing the quality of training inputs. Furthermore, the emphasis on data systems aligns with the study's use of Exploratory Data Analysis and predictive clustering to determine optimal ambulance positions, thus reinforcing the significance of accurate and accessible accident records for actionable insights.

Odero, Khayesi, and Heda (2003) present an in-depth analysis of the road traffic injury crisis in Kenya, detailing its magnitude, causative factors, and current intervention strategies. The study identifies human behavior, inadequate infrastructure, weak enforcement of traffic laws, and insufficient emergency response mechanisms as key contributors to high fatality rates. It also points out the disproportionate impact on economically active age groups, which significantly affects national productivity. The paper highlights the urgent need for government-led reforms, public health planning, and community-based interventions. This reference is particularly relevant as it contextualizes the focus of this study in Nairobi, providing localized insights into accident trends and systemic gaps. The findings support the rationale for improving ambulance positioning and reducing response times through predictive analytics. Moreover, it reinforces the necessity of integrating data-driven models within the broader framework of road safety planning and emergency medical infrastructure in Kenya.

Ferreira and colleagues (2017) offer a comprehensive survey of localization and positioning technologies used by emergency responders. The paper classifies existing systems into GPS-based, inertial, radio frequency, and hybrid models, and evaluates their performance in different operational environments. It emphasizes the importance of real-time localization to improve response efficiency during emergencies, especially in urban settings where time is critical. The survey also discusses technical challenges such as signal loss in dense environments and system latency. For this study, this reference provides critical background on the role of positioning systems in enhancing ambulance pre-positioning strategies. It supports the integration of location-aware models, such as the Deep Embedded Clustering (DEC-AE), which considers spatial data for predicting optimal ambulance placements. The technological insights from this paper guide the development of location-based evaluation metrics and validate the use of spatial analytics in optimizing emergency medical responses.

Maghfiroh, Hossain, and Hanaoka (2018) examine ambulance pre-positioning strategies aimed at minimizing emergency response times in Dhaka, Bangladesh—a city facing challenges similar to Nairobi. Using location-allocation models and simulation techniques, the study demonstrates how strategically placing ambulances across the city can reduce overall response time and improve survival rates. The authors consider various demand scenarios and constraints such as ambulance availability and road network limitations. Their findings confirm that predictive positioning significantly outperforms static placement strategies. This reference provides a methodological foundation for the present study, which applies a deep learning approach to achieve similar objectives in Nairobi. By addressing the complexities of urban traffic and emergency demand prediction, the study supports the use of advanced algorithms like DEC-AE for dynamic ambulance deployment. The contextual similarities between Dhaka and Nairobi also enhance the relevance of this work in validating model effectiveness in comparable socio-economic settings.

Kauffmann et al. (2022) present a cutting-edge approach that bridges clustering techniques with explainable neural networks to improve the interpretability of machine learning models. The paper introduces methods for understanding cluster formation through neural representations, making the black-box nature of deep learning more transparent. It demonstrates how integrating neural networks with clustering algorithms can lead to more meaningful feature extraction and better clustering performance. This reference directly informs the methodology used in the present study, which combines Deep Embedded Clustering (DEC) with feature representation learning through autoencoders. The emphasis on interpretable clustering aligns with the goal of identifying clear, explainable patterns in accident-prone regions for optimal ambulance deployment. Additionally, this work supports the development of the novel distance scoring algorithm by providing insights into latent feature spaces and cluster validation techniques. Thus, Kauffmann's research strengthens both the technical foundation and the explainability component of the proposed system.

III.PROPOSED METHODOLOGY

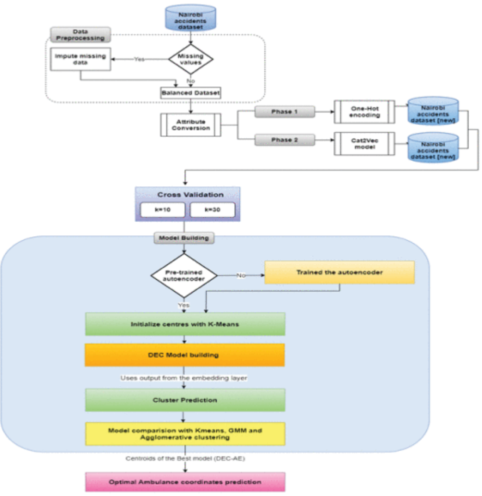
The proposed framework utilizes a real-time accident dataset comprising road traffic incidents, road segment attributes, and weather conditions specific to Nairobi, Kenya. Initially, **Exploratory Data Analysis (EDA)** is conducted to uncover influential features and discern patterns that contribute to road accidents across various locations in the city. This analysis aids in selecting the most relevant attributes for model training and clustering.

To effectively preserve the semantic relationships within categorical variables during preprocessing, a deep learning-based embedding technique called **Cat2Vec** is employed. This approach transforms categorical data into dense vector representations, enabling better feature learning and significantly enhancing clustering performance.

Subsequently, a **Deep Embedded Clustering (DEC)** model is applied to identify optimal ambulance positioning locations. By integrating Cat2Vec embeddings into the DEC framework, the model captures complex, non-linear patterns in the data, resulting in more accurate and interpretable clusters representing high-risk zones.

To evaluate the effectiveness of the predicted ambulance locations, a **novel Distance Scoring Algorithm** is introduced. This algorithm quantifies the distance between each actual crash site and its nearest predicted ambulance location, offering a concrete metric for measuring the practical utility of the model.

Finally, the performance of the proposed system is assessed both with and without feature selection techniques and benchmarked against traditional clustering methods. Multiple clustering evaluation metrics are used to validate the superiority and robustness of the DEC-based approach in identifying optimal ambulance deployment zones in Nairobi.



IV.CONCLUSION

Over the past two decades, the strategies for identifying accident hotspots and determining optimal locations for emergency medical services have significantly advanced, playing a critical role in enhancing traffic safety and emergency response systems. This study focused on predicting the most effective ambulance positioning locations in Nairobi, Kenya, using road accident data from 2018 to 2019.

To achieve this, categorical attributes were transformed into numerical embeddings using the Cat2Vec model, preserving complex relationships in the data during preprocessing. Following feature selection, a clustering-based methodology was implemented using Deep Embedded Clustering (DEC), alongside traditional clustering algorithms such as K-Means, Gaussian Mixture Model (GMM), and Agglomerative Clustering to identify five optimal clusters. These clusters, represented by their centroids, served as the recommended ambulance deployment sites. Performance was evaluated using metrics such as Silhouette Score, Calinski-Harabasz Index, Davies-Bouldin Score, and V-measure, complemented by a custom Distance Scoring Algorithm to quantify proximity between crash sites and predicted ambulance locations. Among all the models, the DEC-AE model integrated with Cat2Vec embeddings delivered the most promising results, achieving a 95% accuracy in k-fold cross-validation and a Distance Score of 7.581, outperforming standard algorithms in minimizing response distances. The consistently superior clustering performance highlights the robustness and practical effectiveness of the DEC-AE approach in capturing hidden data patterns and guiding strategic ambulance placement.

This study offers valuable insights for urban planners and policymakers, enabling data-driven decisions to enhance emergency response infrastructure and ultimately save lives.

V.REFERENCES

[1] World Health Organization, Global Status Report on Road Safety, Geneva, Switzerland, 2015.

[2] T. Sivakumar and R. Krishnaraj, "Road traffic accidents due to drunken driving in India – Challenges in prevention," International Journal of Research in Management and Technology, vol. 2, no. 4, p. 1, 2012.

[3] C. Baguley, "The importance of a road accident data system and its utilization," Technical Report, pp. 1–2, Nov. 2001.

[4] W. Odero, M. Khayesi, and P. M. Heda, "Road traffic injuries in Kenya: Magnitude, causes and status of intervention," Injury Control and Safety Promotion, vol. 10, no. 1–2, pp. 53–61, Apr. 2003.

[5] A. F. G. G. Ferreira, D. M. A. Fernandes, A. P. Catarino, and J. L. Monteiro, "Localization and positioning systems for emergency responders: A survey," IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2836–2870, 4th Quarter, 2017.

[6] M. F. N. Maghfiroh, M. Hossain, and S. Hanaoka, "Minimising emergency response time of ambulances through pre-positioning in Dhaka city, Bangladesh," International Journal of Logistics Research and Applications, vol. 21, no. 1, pp. 53–71, Jan. 2018.

[7] J. Kauffmann, M. Esders, L. Ruff, G. Montavon, W. Samek, and K. Müller, "From clustering to cluster explanations via neural networks," IEEE Transactions on Neural Networks and Learning Systems, early access, Jul. 7, 2022. doi: [To be added].

[8] K. Assi, S. M. Rahman, U. Mansoor, and N. Ratrout, "Predicting crash injury severity with machine learning algorithm synergized with clustering technique: A promising protocol," International Journal of Environmental Research and Public Health, vol. 17, no. 15, p. 5497, Jul. 2020.

[9] A. J. Ghandour, H. Hammoud, and S. Al-Hajj, "Analyzing factors associated with fatal road crashes: A machine learning approach," International Journal of Environmental Research and Public Health, vol. 17, no. 11, p. 4111, Jun. 2020.

[10] P. Tiwari, H. Dao, and N. G. Nguyen, "Performance evaluation of lazy, decision tree classifier and multilayer perceptron on traffic accident analysis," Informatica, vol. 41, no. 1, pp. 39–46, 2017.

[11] T. A. Granberg and H. T. N. Nguyen, "Simulation-based prediction of the near-future emergency medical services system state," in Proceedings of the Winter Simulation Conference (WSC), Dec. 2018, pp. 2542–2553.

[12] C. Boutsidis, P. Drineas, and M. W. Mahoney, "Unsupervised feature selection for the k-means clustering problem," in Proceedings of NIPS, 2009, pp. 1–9.

[13] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 4, pp. 491–502, Apr. 2005.

[14] S. Alelyani, J. Tang, and H. Liu, "Feature selection for clustering: A review," in Data Clustering: Algorithms and Applications, 1st ed., Taylor & Francis, 2013, p. 32.

[15] C. M. Bishop, Pattern Recognition and Machine Learning, New York, NY, USA: Springer, 2006.