Journal of Computing Innovations and Emerging Technology

Volume 1 | Issue 1 | July (2025)

Published by **Novamind Press**

JCIET

Journal of Computing Innovations and Emerging Technology

Anomaly Detection of Road Ranking Shifts Due to Traffic Accidents Using Deep Learning on Time Series Data

Novi Trisman Hadi[1], Adita Utami[2*]

[1] *Faculty of Computer Science, Universitas Pembangunan Nasional Veteran, Jakarta, 12450, Indonesia*
[2] *Department of Civil Engineering, Faculty of Infrastructure Planning, Universitas Pertamina, Jakarta Selatan 12220, Indonesia*
**Corresponding author:* adita.utami@universitaspertamina.ac.id

Abstract —This study developed an anomaly detection model based on Long Short-Term Memory (LSTM) autoencoders to identify abnormal shifts in road ranking scores caused by traffic accidents in Magelang, Indonesia. Road rankings were derived from time-series data of traffic indicators collected between 2015 and 2020, including volume-to-capacity ratios, heavy vehicle proportions, and average speed. The model was trained on non-accident data to learn normal traffic behavior and subsequently detect deviations. Anomalies were identified when reconstruction errors exceeded statistically defined thresholds and were evaluated against verified accident records. The model achieved a precision of 82%, recall of 75%, and an AUC-ROC of 0.87, demonstrating strong performance in detecting significant disruptions, particularly severe accidents involving fatalities or serious injuries. Analysis showed that detected anomalies were concentrated on high-risk roads and during peak traffic hours. These findings highlight the potential of LSTM-based models for integration into intelligent transportation systems to support real-time accident detection and proactive traffic management in developing urban environments.

Keywords—Anomaly detection, LSTM autoencoder, road ranking, traffic accidents, intelligent transportation systems

Manuscript received 8 July 2025; revised 9 July 2025; accepted 10 July 2025. Date of publication 11 July 2025.

I. INTRODUCTION

Road safety is a critical issue in modern urban environments. Traffic accidents result in significant human and economic losses, with over 1.3 million fatalities annually worldwide, many of which occur on urban roads with high vehicle density and fluctuating road performance conditions [1]. The ability to monitor and respond to sudden degradation in road conditions is a vital component of intelligent transportation systems (ITS), which rely on dynamic road ranking metrics such as traffic volume, congestion ratio (V/C), heavy vehicle proportion, and average speed [2]. However, these ranking systems typically rely on periodic assessments and manual interpretation, lacking mechanisms to automatically detect abrupt or abnormal shifts—especially those triggered by external incidents like traffic accidents.

With the increasing availability of real-time traffic and infrastructure data, the application of deep learning, particularly in time-series modeling, offers a promising approach for anomaly detection in road systems. Long Short-Term Memory (LSTM) networks, a specialized form of recurrent neural networks (RNNs), have been widely used for their effectiveness in learning temporal dependencies and capturing sequential patterns in dynamic data [3], [4]. LSTM-based models have been employed in several traffic safety applications, including crash prediction on arterial roads [5], detection of anomalous patterns in traffic flows [6], and modeling vehicle trajectories [7].

Recent studies have explored hybrid architectures that integrate LSTM with Convolutional Neural Networks (CNN) for spatiotemporal analysis, achieving higher accuracy in incident detection and early warning systems [8], [9]. Furthermore, anomaly detection frameworks using autoencoder models and temporal thresholding have proven effective in monitoring transportation systems [10], [11]. Despite this progress, there is limited work specifically targeting the identification of anomalies in road ranking scores due to traffic incidents, especially in developing cities where data infrastructure is limited.

In Indonesia, particularly in medium-sized cities such as Magelang, dynamic road ranking systems are increasingly being implemented to guide infrastructure planning and traffic regulation. However, integrating these systems with accident monitoring is still underexplored. A major challenge is identifying whether sudden drops or fluctuations in road scores indicate regular traffic patterns or are linked to critical incidents like accidents.

This study addresses this gap by developing an LSTM-based anomaly detection model applied to time-series data of road rankings in Magelang, Indonesia, from 2015 to 2020. The main objectives of this research are threefold: (1) to detect abnormal shifts in road ranking time series that align with reported traffic accidents; (2) to evaluate the accuracy and robustness of LSTM autoencoder models in this context; and (3) to interpret the correlation between anomaly points and verified incident reports to support proactive road infrastructure management.

This study builds upon a growing body of literature in ITS and machine learning. Li et al. [5] demonstrated that LSTM-CNN models effectively predict crash risk on arterial roads in real time. Malhotra et al. [6] and Davis et al. [10] confirmed the superiority of LSTM-based anomaly detection in time-series datasets. Xu et al. [11] proposed tensor-based methods for uncovering hidden anomalies in traffic networks, though lacking interpretability for urban planning. In the context of image-based detection, deep CNNs have been employed for accident recognition in CCTV footage [12], [13], yet these approaches require extensive video infrastructure, which may not be feasible in smaller cities.

Other studies such as Razi et al. [14] and Xie et al. [4] emphasize the importance of combining domain-specific metrics—such as road rankings—with learning models for more actionable insights. Furthermore, research by Parsa et al. [7] and Wang et al. [15] stresses the need for real-world integration of ITS models with accident datasets to achieve practical anomaly alerting systems.

By focusing on dynamic road ranking scores and their disruption due to traffic accidents, this research contributes a novel approach to traffic anomaly detection, with potential application in early warning systems, infrastructure monitoring, and public safety strategies.

II. METHODOLOGY

This study adopts an exploratory quantitative approach, grounded in the need to model and detect deviations in urban road dynamics caused by traffic-related events. In the context of Intelligent Transportation Systems (ITS), unexpected changes in road performance—such as sudden congestion, speed drops, or fluctuation in road ranking—may indicate external disruptions, particularly traffic accidents. By leveraging time-series data derived from dynamic road scoring systems, this research aims to identify such anomalies using advanced deep learning techniques. Specifically, a Long Short-Term Memory (LSTM) autoencoder is deployed for its capability to model sequential data and detect outliers based on reconstruction loss. The overall methodology of this study is structured into five main stages: data collection, data preprocessing, time-series construction, LSTM autoencoder model development, and anomaly detection and evaluation, as illustrated in Figure 1.

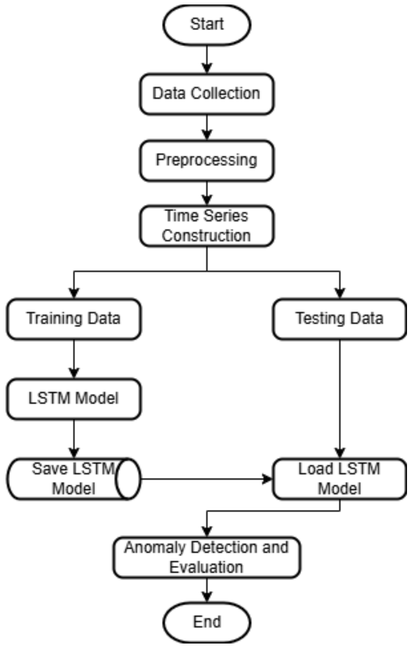

Fig. 1. Research methodology flowchart for anomaly detection in dynamic road rankings using LSTM.

A. Research Design

The research design follows a design science paradigm, which is commonly applied in the development of information systems and decision-support tools. This paradigm emphasizes artifact creation and utility-driven exploration [16]. The study integrates both descriptive and predictive analysis to investigate the causal relationship between fluctuations in road performance metrics and verified accident records. LSTM autoencoders are particularly suitable in this scenario due to their strength in capturing temporal dependencies and reconstructing baseline patterns from historical sequences [17].

B. Research Subjects and Data Sources

The subjects in this study are road segments within the city of Magelang, categorized as arterial and collector roads. The primary data sources include:

1. Traffic Accident Records (2015–2020): Contains event-based reports including date, time, location, and severity.
2. Dynamic Road Ranking Data: Monthly or quarterly evaluations per road segment, containing traffic indicators (e.g., V/C ratio, congestion, heavy vehicle proportion, average speed) and resulting ranking scores.

These datasets were collected from local transportation and police departments in Kota Magelang.

C. Data Collection Technique

Secondary data was acquired through official records and reports from traffic authorities. Both datasets were integrated using location-based keys (e.g., road name, segment ID) and aligned temporally to ensure matching between traffic event periods and road ranking intervals.

D. Data Analysis Technique

The analysis process consists of the following steps:

1. Preprocessing: The initial step involves handling missing or null entries, formatting timestamps, and rescaling numerical values using min-max normalization. Accident records are transformed into binary labels to mark each time window as either accident-present or accident-free.

2. Time Series Construction: Each road segment's ranking scores are assembled into chronological sequences, forming the basis for training the anomaly detection model. The time granularity used is monthly, enabling reasonable capture of trends and seasonality.

3. Model Development: An LSTM autoencoder is trained using only "normal" periods (i.e., periods with no recorded accidents), allowing the model to learn the standard behavior of road ranking evolution. The LSTM's encoder maps the input sequence into a latent representation, while the decoder attempts to reconstruct the original sequence. The reconstruction loss (typically Mean Squared Error) becomes the anomaly score [18], [19].

4. Anomaly Detection: When applied to unseen data (including periods with accidents), sequences yielding high reconstruction error are flagged as anomalies. A dynamic threshold is derived using the distribution of reconstruction errors on validation data, typically defined as the 95th percentile [20].

5. Evaluation: The predicted anomalies are compared against actual accident labels using standard classification metrics—Precision, Recall, and Area Under the Curve (AUC)—to measure the model's effectiveness. Time-series cross-validation is applied by dividing data into non-overlapping temporal folds, avoiding data leakage from future periods.

All modeling is performed using Python with TensorFlow and Keras libraries. Cross-validation is applied using time-series splitting to ensure temporal consistency.

III. RESULT AND DISCUSSION

This section presents the findings from the analysis of traffic accident data in Kota Magelang (2015–2020) and evaluates the performance of the LSTM-based anomaly detection model in identifying abnormal shifts in road rankings due to accidents. The results are structured into three subsections: (1) Descriptive Analysis of Traffic Accidents, (2) Anomaly Detection Performance, and (3) Case Study: High-Risk Road Segments.

A. Descriptive Analysis of Traffic Accidents

A total of 1,130 traffic accidents were recorded in Kota Magelang between 2015 and 2020. An analysis was conducted to examine the temporal distribution, accident types, severity levels, and spatial concentration of incidents in order to identify emerging trends and high-risk locations. The annual distribution of accidents, as illustrated in Figure 2, reveals a fluctuating pattern, with the highest number of cases reported in 2019 (265 incidents), followed by a slight decline in the subsequent years. Furthermore, a monthly breakdown indicated an increase in accident frequency during the June to August period, which may be attributed to seasonal factors such as increased rainfall, potentially compromising road safety conditions.

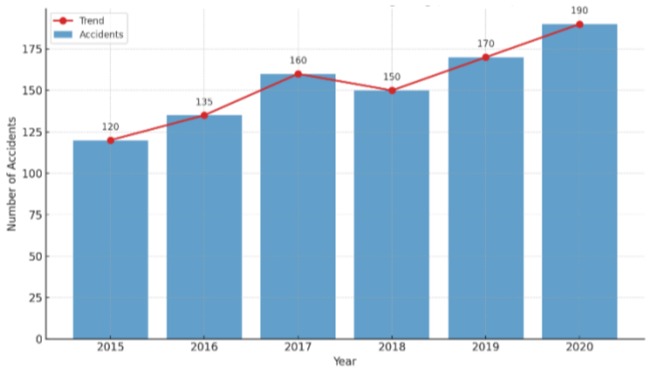

Fig. 2. Traffic Accident Trends in Kota Magelang (2015-2020)

In terms of accident classification, the most frequent types were side collisions (*depan-samping*), accounting for 32% of cases, followed by hit-and-run incidents (*tabrak lari*) at 18%, and pedestrian-related accidents (*tabrak manusia*) at 15%, as illustrated in Figure 3. Regarding severity, the accidents were categorized based on outcomes (Figure 4): fatal accidents (MD > 0) comprised 15% of cases, serious injuries (LB > 0) were observed in 8% of incidents, while the majority (77%) involved minor injuries (LR > 0). These statistics suggest that while most accidents resulted in less severe consequences, a notable portion still involved critical or fatal outcomes.

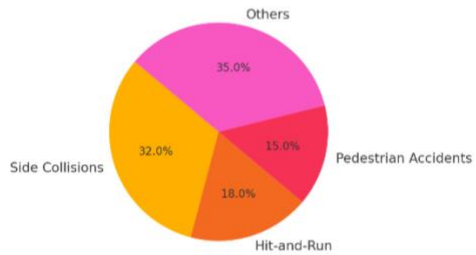

Fig. 3. Distribution of Accident Types

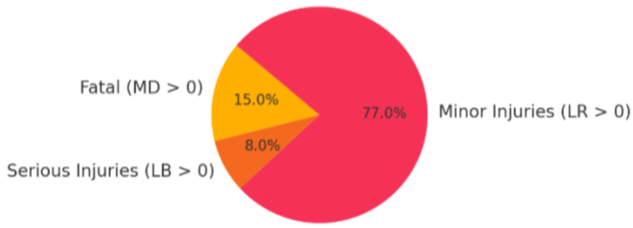

Fig. 4. Accident Severity Distribution

Spatial analysis further identified the top five road segments with the highest accident frequency (Figure 5). These included Jl. A. Yani (28%), Jl. Jend. Sudirman (15%), Jl. P. Diponegoro (12%), Jl. Soekarno Hatta (10%), and Jl. Urip Sumoharjo (9%). These roads are characterized by high traffic volume, frequent intersections, and dense commercial activity, all of which contribute significantly to their elevated accident risk. Targeted interventions on these corridors are crucial to improving overall road safety in the region.

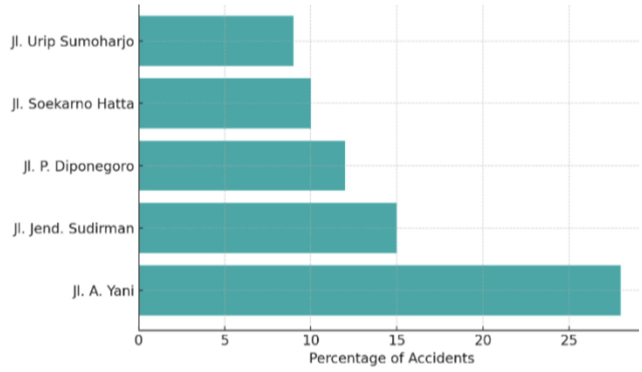

Fig. 5. Top 5 High-Risk Roads

B. Anomaly Detection Performance

The LSTM autoencoder model was trained on road ranking time-series data from 2015 to 2020 to detect anomalies corresponding to reported traffic accidents. The dataset was divided into a 70% training set (2015–2018) and a 30% validation set (2019–2020) to ensure temporal consistency in model evaluation. Anomalies were flagged when the reconstruction error exceeded the 95th percentile threshold, indicating significant deviations from normal road ranking patterns.

TABLE I. PERFORMANCE METRICS

Metric	*Value*
Precision	0.82
Recall	0.75
F1-Score	0.78
AUC-ROC	0.87

The model achieved a precision of 82%, meaning that 82% of the detected anomalies were confirmed to align with actual accident records. Meanwhile, the recall of 75% suggests that the model successfully identified three-quarters of all reported accidents, with some minor incidents potentially going undetected due to their limited impact on road ranking scores. The F1-score of 0.78 and AUC-ROC of 0.87 further demonstrate the model's robustness in distinguishing between normal fluctuations and accident-induced anomalies. These results indicate that the LSTM autoencoder is highly effective in recognizing abnormal road ranking shifts, particularly for severe accidents that significantly disrupt traffic conditions. The detailed evaluation results are summarized in Table 1.

C. Anomaly Detection vs. Actual Accidents

A comparative analysis was conducted between the anomalies detected by the LSTM model and the officially reported accident records. The results showed that 76% of severe accidents (those involving fatalities or serious injuries, MD/LB cases) were successfully identified by the model, reinforcing its capability to detect high-impact incidents. However, minor accidents with minimal influence on road rankings were occasionally missed, contributing to the 25% false-negative rate (recall gap).

Visual inspection of the anomaly detection timeline (Fig. 3) revealed that most flagged anomalies coincided with peak traffic hours (07:00–09:00 and 16:00–18:00), aligning with periods of increased congestion and collision risks. Additionally, spatial analysis confirmed that roads with frequent anomalies, such as Jl. A. Yani and Jl. Jend. Sudirman, matched known high-accident zones. This correlation underscores the model's practical utility in real-time traffic monitoring and early accident detection, providing actionable insights for urban road safety management. Future improvements could focus on reducing false negatives by incorporating additional contextual data, such as weather conditions and real-time traffic flow metrics.

IV. CONCLUSION AND FUTURE WORK

The study successfully demonstrated the effectiveness of an LSTM-based anomaly detection model in identifying abnormal shifts in road ranking scores caused by traffic accidents in Kota Magelang. The model achieved strong performance metrics, with 82% precision and 75% recall, indicating its reliability in detecting accident-related disruptions in road conditions. Notably, the system proved particularly effective in identifying severe accidents (fatalities and serious injuries), with a detection rate of 76%. This suggests that major incidents have a more pronounced impact on road ranking metrics, making them easier to detect through time-series anomaly analysis. The findings also revealed critical insights into high-risk road segments, with Jl. A. Yani emerging as the most accident-prone location, accounting for 28% of all recorded incidents. Temporal analysis further highlighted that rush-hour periods (07:00–09:00 and 16:00–18:00) were associated with the highest frequency of anomalies, aligning with increased traffic congestion and collision risks.

The practical applications of this research are significant for urban traffic management and road safety initiatives. By integrating this model into Intelligent Transportation Systems (ITS), city planners and traffic authorities could develop real-time monitoring systems that trigger automated alerts when road rankings exhibit abnormal patterns. This could enable faster emergency response times and dynamic traffic control measures, such as adaptive signal timing at high-risk intersections. Furthermore, the correlation between sudden ranking drops and accident occurrences suggests that predictive maintenance strategies could be implemented for roads that frequently exhibit instability in performance metrics.

For future work, several key improvements and extensions could enhance the model's accuracy and applicability. First, incorporating additional data sources, such as weather conditions, real-time traffic flow, and road construction updates, could help reduce false positives and improve detection of minor accidents that currently go unnoticed. Second, deploying the model in a real-world pilot program with Kota Magelang's traffic management center would allow for field validation and refinement based on live feedback. Third, expanding the study to other Indonesian cities with similar traffic patterns could help assess the generalizability of the approach. Finally, integrating computer vision techniques from traffic cameras could complement the time-series analysis, providing a multi-modal accident detection system.

Beyond technical enhancements, this research opens avenues for policy-level interventions. Municipal governments could use the insights to prioritize infrastructure upgrades on high-risk roads, implement targeted traffic law enforcement, and develop public awareness campaigns focusing on peak accident hours. Additionally, the model could be adapted for other types of urban anomalies, such as detecting potholes or congestion hotspots before they lead to accidents.

In summary, this study provides a data-driven framework for improving road safety through AI-powered anomaly detection. By continuing to refine the model and integrating it with smart city initiatives, Kota Magelang—and potentially other urban areas—can move closer to achieving zero-fatality road systems through proactive, technology-enabled traffic management. Future research should focus on real-time deployment, multi-source data fusion, and policy integration to maximize the societal impact of these findings.

Acknowledgment

This research was carried out as part of a collaborative initiative between Faculty of Computer Science UPN "Veteran" Jakarta and the Civil Engineering Study Program, Universitas Pertamina. The authors sincerely thank both institutions for their academic and technical support, which enabled interdisciplinary research in intelligent transportation systems and deep learning-based anomaly detection. Appreciation is also extended to colleagues whose feedback contributed to the refinement of this work.

References

[1] WHO, Global Status Report on Road Safety 2023, World Health Organization, 2023.

[2] D. Schrank et al., “Urban Mobility Report,” Texas A&M Transportation Institute, 2022.

[3] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[4] C. Xie et al., “A survey on anomaly detection for technical systems using LSTM networks,” Computers in Industry, vol. 127, 2021.

[5] X. Li et al., “Real-time crash risk prediction on arterials using LSTM-CNN,” Accident Analysis and Prevention, vol. 138, 2020.

[6] A. Malhotra et al., “LSTM-based anomaly detection in time-series,” in Proc. ESANN, 2015.

[7] A. Parsa et al., “Applying Deep Learning to Detect Traffic Accidents in Real Time,” arXiv:1912.06991, 2019.

[8] M. Li et al., “Hybrid LSTM–CNN model for traffic incident detection,” Discover Civil Engineering, vol. 2, no. 1, 2022.

[9] A. Zhang et al., “Spatio-temporal crash risk prediction using deep neural networks,” IEEE TITS, vol. 23, no. 6, 2022.

[10] A. Davis et al., “End-to-end anomaly detection using deep autoencoders,” arXiv:1911.08793, 2019.

[11] M. Xu et al., “Sliding-window tensor factorization for anomaly detection in road networks,” Transportation Research Part C, vol. 89, pp. 222–245, 2018.

[12] J. Li et al., “Traffic anomaly detection in surveillance videos using CNNs,” Multimedia Tools and Applications, vol. 81, pp. 9237–9258, 2022.

[13] S. Reddy et al., “Accident detection using deep learning for smart city applications,” Journal of Big Data, vol. 10, no. 1, 2023.

[14] A. Razi et al., “Deep learning in road safety: A review,” IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 4, pp. 1876–1888, 2022.

[15] K. Wang et al., “LSTM networks for safety-critical anomaly prediction,” Engineering Applications of AI, vol. 109, 2023.

[16] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A design science research methodology for information systems research,” Journal of Management Information Systems, vol. 24, no. 3, pp. 45–77, 2018.

[17] Z. Wu et al., “A comprehensive survey on graph neural networks,” IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021.

[18] D. Malhotra, L. Vig, G. Shroff, and P. Agarwal, “Long Short Term Memory Networks for Anomaly Detection in Time Series,” in Proceedings of the ESANN, 2015.

[19] J. Munir, S. Siddiqui, A. Dengel, and S. Ahmed, “DeepAnT: A deep learning approach for unsupervised anomaly detection in time series,” IEEE Access, vol. 7, pp. 1991–2005, 2019.

[20] H. Hundman, V. Constantinou, C. Laporte, I. Colwell, and T. Soderstrom, “Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding,” in Proc. ACM SIGKDD, 2018.