International Journal of Leading Research Publication (IJLRP)

E-ISSN: 2582-8010 ● Website: www.ijlrp.com ● Email: editor@ijlrp.com

DAYLENGTH ANALYSIS-BASED ACCIDENT HOTSPOT PREDICTION WITH GIS DATA FOR ROAD SAFETY IMPROVEMENT USING ANFGuIS

Lalitha Reddy, Badam

Software Developer (Full-Time) & Adjunct Professor (Part-Time)
Computer Science Department
Southern New Hampshire University, USA
lalithabadambvsr@gmail.com

Abstract:
An intelligent Geographic Information System (GIS)-based Accident Hotspot Prediction (AHP) in smart cities is necessary for road safety. However, the prevailing works overlooked the twilight duration via latitude of a location, causing improper AHP. Thus, daylength-based density analysis using Inverse Hyperbolic Tangent-Kernel Density Estimation (IHT-KDE) and Adaptive Neuro Fuzzy Gudermannian Interference System (ANFGuIS)-based AHP is proposed. Primarily, the GIS-based road accident data are collected and pre-processed. Then, the weather-based clustering is performed using Dirichlet Phi-Square Density-Based Spatial Clustering of Applications with Noise (DPS-DBSCAN). Next, the feature extraction, twilight and latitude-based daylength analysis, and density analysis are performed. Meanwhile, the road network accident propagation is modeled using a Knowledge Graph (KG), and the accident propagation features are extracted. Further, the ANFGuIS-based accident labeling and prediction is done. Finally, severity-based road safety intervention using Haversed Sine Markov Decision Process (HS-MDP) is carried out. Thus, the proposed system predicts the accident hotspot with an accuracy of 98.87%, outperforming existing works.

Keywords: Accident Prediction, Road Safety, Smart Cities, Geographic Information System (GIS), Machine Learning (ML), Data Analytics, and Road Safety Intervention.

1. INTRODUCTION

The motorization and urbanization in the modern smart cities have increased the traffic density, which has also raised road accidents (Kamh et al., 2025) (Deressa et al., 2025). These accidents occur due to the interaction of human, environmental, spatial, and temporal factors (Yalcin et al., 2026). Hence, the determination of crash hotspots earlier is necessary to improve road and public safety (Berhanu et al., 2024). The integration of ML with GIS data analytics has enabled large-scale AHP (Moreno-ponce et al., 2025). However, most of the existing studies relied on historical accident counts (Yan et al., 2024) and failed to capture the spatial-temporal dependencies. The prevailing works categorized the lighting conditions as day or night (Habib et al., 2023), ignoring the twilight duration across the different geographical locations. As the sunrise and sunset durations vary with latitude, the ignorance of this factor causes spatial bias in AHP (Balawi & Tenekeci, 2024). Further, the environmental factors, the influence of accident location, and predictive measures should be analyzed for improving road safety (Wang et al., 2022) (Ma et al., 2021). Therefore, a novel ANFGuIS-based accident hotspot labeling and prediction is proposed.

1.1 Research Overview

In this section, the research gap, problem statement, significance, and contribution of the proposed system are detailed.

Research Gap: Prevailing works considered only the categorical lighting conditions, such as day or night. These models failed to consider the geographical variation in twilight duration across different latitudes, causing spatial bias in prediction accuracy.

Problem Statement: The limitations of the prevailing AHP include,

- The existing (Hazaymeh et al., 2022) ignored the accident propagation effects, reducing the reliability regarding AHP.
- Prevailing (Alsahfi, 2024) analyzed the road accidents without considering environmental conditions like rain, fog, and mist, resulting in inaccurate predictions.
- Traditional (Islam et al., 2022) depended on manual observation, lacking standardized preventive actions for road safety management.

Significance and Scope of the Study: The significance of this research is to develop a spatial-aware adaptive AHP system that supports road safety planning in smart cities. The scope includes the analysis of geographic, temporal, and environmental factors for reliable AHP and safety intervention.

Research Aim and Objectives: The main aim of this research is to develop an ML-based AHP system that integrates GIS for road safety improvement. The key contributions are,

- ❖ To improve the AHP, the daylength analysis, IHT-KDE-based density analysis, and ANFGuIS-based accident hotspot labeling and prediction are performed.
- ❖ The KG-based road network accident propagation modeling is performed to capture the accident propagation relationship.
- ❖ To handle the environmental variations, the weather-based clustering of road conditions using DPS-DBSCAN is implemented.
- ❖ For effective decision-making regarding road safety improvement, the HS-MDP-based road safety intervention is carried out.

The rest of the paper is arranged as follows: Section 2 presents related works, Section 3 describes the proposed methodology, Section 4 discusses experimental results and performance evaluation, and Section 5 concludes the paper with practical implications and future research directions.

2. LITERATURE SURVEY

(Hazaymeh et al., 2022) determined spatiotemporal analysis of traffic accident hotspots. Here, the traffic accident GIS data was collected. Next, the Spatial Autocorrelation method was used to identify the spatial distribution patterns. Then, the number of accidents was aggregated, and the clustering model predicted the hotspots precisely. On the other hand, the propagation of the accidents was not taken into account, causing ineffective AHP.

(Alsahfi, 2024) envisaged road traffic accident analysis via GIS. The accidental data from major cities were collected. Further, the spatiotemporal analysis was performed. Later, the severity was determined by grouping the data. Finally, the hotspots were accurately detected based on space-time cube analysis, population density analysis, and the Getis-Ord Gi* method. Yet, the environmental conditions were not considered, leading to inaccurate hotspot identification.

(Islam et al., 2022) developed road crash severity and hotspot prediction models. Primarily, the crash data from cities was gathered. Next, the severity of the crash was determined using the Random Forest method. Afterwards, the target-specific model interpreted the cause of the crash. Finally, the cluster zones of the hotspot were evaluated effectively using the Getis-Ord Gi* model. On the contrary, this model lacked intervention regarding road safety.

(Aldalain et al., 2023) implemented road accident hotspot identification using Structural Equation Modeling (SEM). The accident data was collected from the traffic department. Further, the road accident location was identified, and the hotspots were mapped using the Getis-Ord Gi* technique. Later, the topographical survey using the total station was analyzed. At last, based on the accident rate and SEM, the hotspots were predicted accurately. However, the geographic variation in lighting conditions was not analyzed, causing biased prediction.

(Santos et al., 2021) presented an ML approach for traffic accident analysis with hotspot prediction. Initially, the traffic accident data with various variables were collected. Then, based on the similarity between features, the clustering was performed. Afterwards, the features were extracted and selected. Finally, based on the rule technique, the hotspots were detected efficiently. But, the intensity of the hotspot for decision-making could not be estimated.

3. PROPOSED METHODOLOGY REGARDING ACCIDENT HOTSPOT PREDICTION VIA GIS DATA FOR ROAD SAFETY IMPROVEMENT

In this section, the twilight and latitude-based daylength and density analysis are performed for effective AHP. The structural diagram of the proposed framework is depicted in Figure 1.

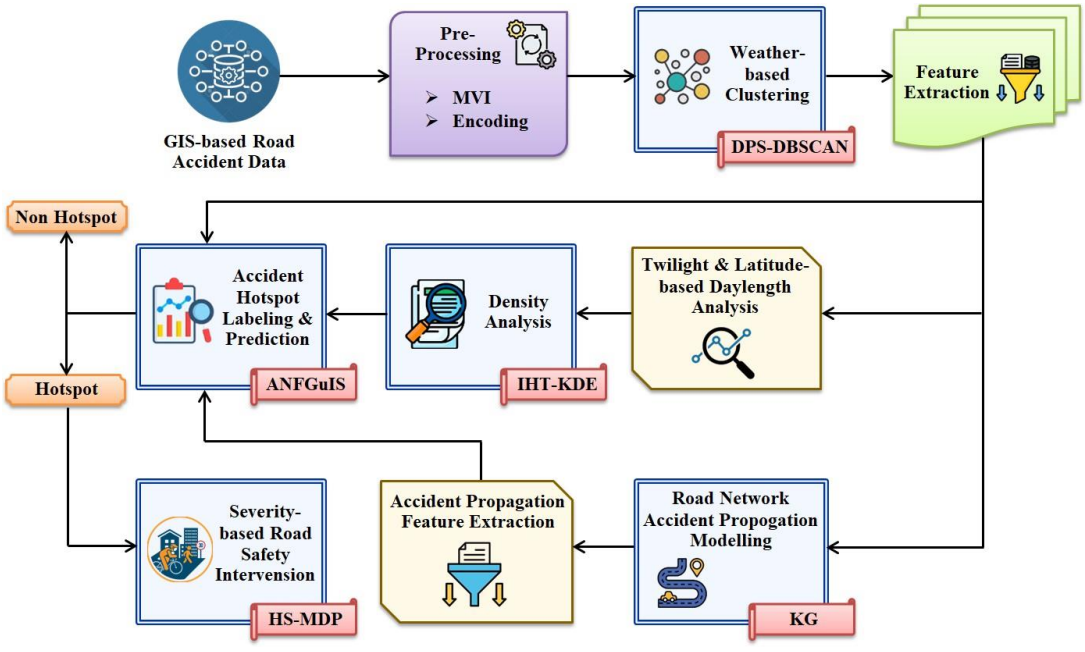

Figure 1: Architecture of the Proposed System

The step-by-step procedure of the proposed AHP framework for road safety improvement is explained below.

3.1 GIS-based Road Accident Data

The proposed framework starts with the collection of road accident data M that consists of GIS data like latitude and longitude of the accident spot, location details, and so on.

$M = M^1, M^2, M^3, M^4, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, M^g$

Where, g denotes the number of M and M^g represents the g^{th} collected data.(1)

3.2 Pre-Processing

Here, M is pre-processed regarding Missing Value Imputation (MVI) and encoding.

Missing Value Imputation: Primarily, the missing data in M is imputed using the mean of the neighbouring data in M. Thus, the MVI data is represented as M'.

Encoding: Afterward, the categorical data in M' are converted into numerical form using one-hot encoding. The encoded output, which is the pre-processed data, is given as G. Thus, the ML models could analyze the data G effectively.

3.3 Weather-based Clustering and Feature Extraction

Afterward, for analyzing the accident pattern regarding the weather condition, G is clustered using DPS-DBSCAN. Here, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN), which does not require cluster numbers, is chosen for weather-based clustering. However, improper selection of the epsilon and minimum points (minpts) parameters degrades the clustering performance. Hence, the epsilon and minpts in DBSCAN are evaluated using Phi-Square (PS) distance and Dirichlet function, respectively.

At first, the epsilon parameter ε, which helps in searching the nearby points during clustering, is calculated using PS distance. This PS distance captures the distributional differences among the attributes.

$$\sqrt{\sum \left(G^1 - G^2\right)^2 / \left(G^1 + G^2 + \alpha\right)} \quad (2)$$

Where, G^1, G^2 signify the data points in G and represents the constant value. Next, the minpts, which determines the minimum number of data needed for clustering within, is estimated using the Dirichlet function. Here, the Dirichlet function allows the algorithm to adapt to varying data concentrations.

$$\lim_{1}^{\infty} \cos G \left(\right)^{2n}$$

(3)

Here, is the mathematical constant and n denotes the total number of inputs. Now, by utilizing the weather condition $G^{\prime}$ in G, the core points L needed for clustering the data are evaluated.

$$L \left(G^{\prime} \right) G \left(\right), \left(\right)$$

(4)

Finally, the weather-based clustering C is performed until all the data points are clustered regarding and .

$$C \left(\right) L G^{\prime} \left(\right)$$

(5)

The pseudo-code for DPS-DBSCAN is described below.

Pseudo-code for DPS-DBSCAN

Input: Pre-processed data G

Output: Weather-based clustering C

Begin
Initialize , n
For G
Evaluate PS distance-based epsilon
Find Dirichlet function-based minpts
Identify core points **While** $G \; L$

$$\sqrt{\sum \left(G^1 - G^2 \right)^2 / \left(G^1 + G^2 + \alpha \right)}$$
$$= \lim_{1}^{\infty} \left[\cos \left(G \times \pi \alpha \right)^{2n} \right]$$

$$L \left(G^{\prime} \right) G \left(\right), \left(\right)$$

End Perform weather-based clustering $C \left(\right) L G^{\prime} \left(\right)$
End while
End for
Obtain C

Then, the features E, such as accident Identification (ID), severity, start time, end time, distance, starting latitude, starting longitude, ending latitude, ending longitude, street, city, country, state, zipcode, timezone, weather timestamp, temperature, win speed, amenity, bump, crossing, junction, roundabout, station, traffic signal, turning loop, sunrise and sunset, civil twilight, nautical twilight, astronomical twilight, and so on, are extracted from C.

3.4 Twilight & Latitude-based Daylength Analysis

Further, based on latitude □E_1□, day of the year □E_2□, and twilight coefficient □p□, the twilight and latitude-based daylength analysis is performed. Here, the values of □p□ regarding the sun horizon are represented as,

□0.8333
□6.0 $p =$ |
□
□12.0
□□18.0
for sunrise/sunset for civil twilight
for nautical twilight
for astronomical twilight

(6)

Next, the daylength □K□is evaluated based on □E_1□, □E_2□, earth's revolution angle □□□, and sun declination □y□as follows,

K □ 24 □ □□24□□ { π
□ arccos
(7) [sin(p)− ⟨sin(E1)×sin(y)⟩] } |
cos(E1)× cos(y)]| |

y □ arcsin □0.39795 cos□□□□
□ □ 0.2163108□ □2□ arctan □0.9671396tan□0.00860□ E □186 □□□
(8) | 2 |

(9)

Thus, daylength □K□ is determined for effective AHP. Now, □K□is normalized to a weight value □K^{wt}□ for the present time □t□.

K^{wt} □1□ □K□t□/ max□K□□

This □K^{wt}□is utilized for density analysis.

3.5 Density Analysis(10)

Then, the density of the accident location is analyzed using IHT-KDE. Here, the Kernel Density Estimation (KDE) is chosen for density analysis. KDE provides a continuous evaluation of accident incidence distribution over the specified place and reveals the spatial concentration pattern present in the traffic incidents. Yet, the conventional KDE causes computational overhead during the evaluation of the kernel function. Therefore, the Inverse Hyperbolic Tangent (IHT) function, which highlights the nearby data points while reducing the influence of distant observations, is used to determine the kernel function □□ □.

□□ □1□1 □1□ K^{wt} □□□□

□ □1□ □arctan □ 2 log□ 1□ K^{wt} □□□(11) – ———
□ □ □ □□□□

Now, the density □D□of the accident location is estimated regarding □□ □, □K□, □K^{wt}□, and smoothing factor □m□ as,

$$D \square \frac{1}{\square \square K^{wt} \square}^{\square} \square \tan^{\square 1}\square \frac{K}{m} \square\square^{\square 1} \square \square K \square\square \quad (12)$$

Thus, the density of the accident location is evaluated.

3.6 Road Network Accident Propagation Modeling

Meanwhile, to analyze the road network accident propagation, the KG is constructed for the features □E□. The KG captures the influence of one accident over another regarding nodes and edges. Here, the entities/nodes □J□ denote accident locations, latitude, longitude, weather conditions, time, traffic signal, crossing, roundabout, and lighting conditions. Further, the edges □I□ represent the accident that occurred simultaneously, nearby the accident, before the accident, spatial closeness between accidents, and accident time interval. Hence, the KG □□ □is modeled as, □ □ $E \square J, I \square$ (13)

Hence, the accident propagation in the road network is determined.

3.7 Accident Propagation Feature Extraction

Further, the accident propagation features □T□, such as node centrality, number of nodes, number of edges, in-degree, out-degree, neighbour accident count, and accident frequency, are extracted from □□ □.

3.8 Accident Hotspot Labeling & Prediction

In this phase, the accident hotspots are labeled and predicted using ANFGuIS. Here, the Adaptive Neuro Fuzzy Interference System (ANFIS), which handles the uncertain data effectively, is selected for accident hotspot labeling and prediction. However, the conventional membership function in ANFIS could not capture the nonlinear patterns. Hence, the Gudermannian membership function is utilized in ANFIS. The pictorial diagram of the proposed ANFGuIS is illustrated in Figure 2.

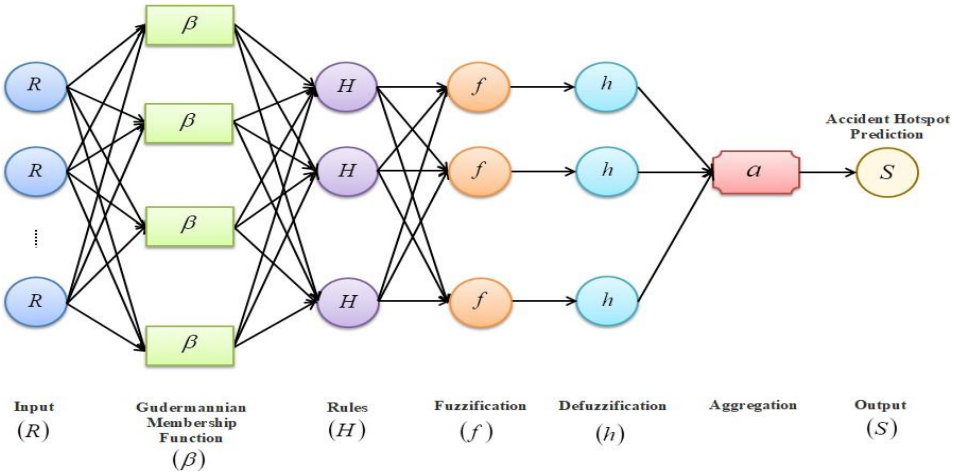

Figure 2: Structure of the Proposed ANFGuIS Model

The inputs, such as features T, E and density D, are represented as R. Initially, the membership function that helps in analyzing the degree of relationship between the inputs is equated using the Gudermannian membership function β. This β provides smooth nonlinear mapping and captures the gradual transition between the data.

$$\beta = \arctan(\sin(D)) \quad (14)$$

In the next layer, the rules H for labeling the hotspot are set based on D and if-then condition.

$$H = \begin{cases} if \\ if\ D \geq 90 \text{ percentile}\ then\ S^1 \\ D < 90 \text{ percentile}\ then\ S^2 \end{cases} \quad (15)$$

The condition states that if D is greater than or equal to 90 percentile, then the accident location is said to be a hotspot S^1. If D is less than 90 percentile, then the location is said to be non-hotspot S^2. Now, the crisp R is converted to fuzzy data f regarding H and β.

$$f = \beta(R, H, \beta) \quad (16)$$

Where, β denotes the weight value. Next, f is defuzzified based on β. Thus, the defuzzified data h is represented as,

h □ □□□ □ f □/ □ □(17)

At last, □h□is aggregated to achieve the AHP output □S □. The aggregation takes place regarding the aggregation factor □a□.

S □ □a □ h□ for all□a □ □h□

(18)

Hence, the accident hotspots are predicted effectively. The pseudo-code for ANFGuIS is stated below.

Input: Resultant Factor □R□
Output: AHP Output □S □

Begin Pseudo-code for ANFGuIS

Initialize □□ □, □D□
For □R□
Evaluate Gudermannian membership function **While** □D□
Set if-then rules □H □
If □D □ 90 percentile□
Then □S^1 □
Else if □D □ 90 percentile□
Then □S^2 □

□ □ arctan□sin□D□□

End if **End** while **Fuzzification**

f □□ □R□□H, □ □□ **Defuzzification** h □ □□□ □ f □/ □ □
Aggregation-based AHP S □ □a □ h□ for all□a □ □h□
End for
End

After predicting the accident hotspots, the intervention regarding road safety is performed.

3.9 Severity-based Road Safety Intervention

Finally, to improve road safety, the severity-based road safety intervention is performed using HS-MDP. Here, the Markov Decision Process (MDP), which learns the actions optimally, is chosen for road safety intervention. However, the increase in the number of states and intervention choices leads to computational complexity. Therefore, the Haversed Sine (HS) function is introduced as a transition-regulation mechanism in MDP.

Primarily, the decision-making is defined based on a set of states □Q□, a set of actions □A□, transition probability □z□, reward function □k □, and discount factor □i□. Here, each □Q□in □S^1 □represents accident severity as low □Q^1 □, moderate □Q^2 □, high □Q^3 □, or critical □Q^4 □. Further, the action □A□is set as,

$$A(S^1) = \begin{cases} \text{for } Q^1 \text{ no action } A^1 \\ \text{for } Q^2 \text{ warning to drivers } A^2 \\ \text{for } Q^3 \text{ traffic control } A^3 \\ \text{for } Q^4 \text{ immediate patrol action } A^4 \end{cases} \quad (19)$$

Also, z is estimated based on the HS function, which provides a smooth confined representation of transition influence.

$$z = \sin^2(S^1 Q / 2) \quad (20)$$

The reward function is updated based on z, A, and Q to intervene in necessary preventive actions.

$$k = u z + Q, A i \quad (21)$$

Here, k represents the updated reward function and u is the weight parameter. Finally, the road safety intervention output B is evaluated based on A, z, k, and i.

$$B = \arg\max A(S^1, z, k, i) \quad (22)$$

Therefore, respective interventions in the accidental hotspot are taken for road safety improvement. Hence, the proposed system precisely detects the accidental hotspot location and improves road safety regarding GIS data. The performance analysis of the proposed techniques is explained below.

4. RESULTS AND DISCUSSION

In this section, the performance analysis of the proposed AHP is carried out and compared with prevailing works to prove the effectiveness of the proposed system. The implementation is carried out on the Python platform.

4.1 Dataset Description

For AHP, the "US Accidents: A Countrywide Traffic Accident Dataset" is taken. This dataset consists of a total of 7728394 data. From that, 80% (6182715) and 20% (1545679) of the data are used for training and testing, respectively.

4.2 Performance Validation

Here, the performance of the proposed models is evaluated and compared with the baseline models.

Table 1: Comparative Analysis of IHT-KDE

Methods	MAE	RMSE	Log-Likelihood	AIC
Proposed IHT-KDE	0.078	0.103	-1125.48	2371.48
KDE	0.281	0.352	-1593.59	3147.49
HDE	1.934	2.371	-1948.58	3904.83
GMM	3.642	4.472	-2357.11	4824.43
MSA	4.821	5.379	-2784.48	5721.84

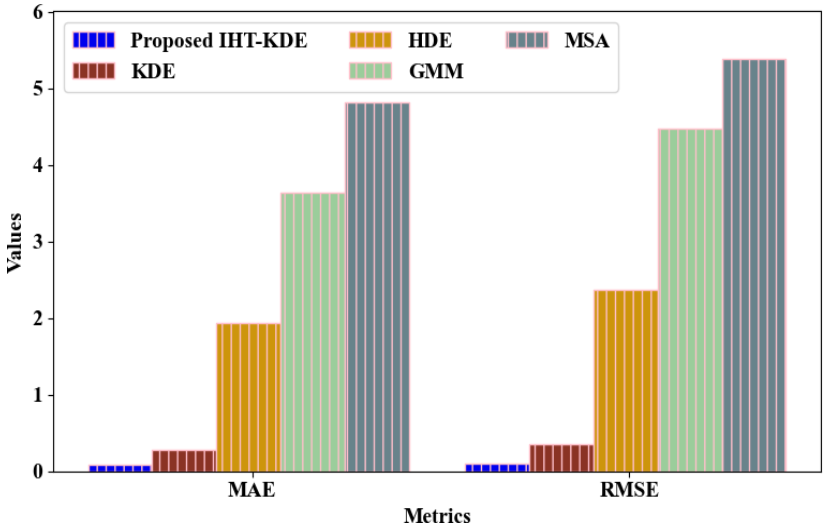

Figure 3: Graphical Comparison regarding Density Analysis

The comparison of the proposed and existing models regarding daylength-based density analysis is shown in Table 1 and Figure 3. The proposed IHT-KDE attained a Mean Absolute Error (MAE) of 0.078, Root Mean Squared Error (RMSE) of 0.103, log-likelihood of -1125.48, and Akaike Information Criterion (AIC) of 2371.48. However, the prevailing KDE, Histogram Density Estimation (HDE), Gaussian Mixture Model (GMM), and Mean Shift Algorithm (MSA) obtained an average MAE, RMSE, log-likelihood, and AIC of 2.67, 3.144, -2170.94, and 4399.65, respectively. Thus, the evaluation of the kernel using the IHT function enhanced the performance of the proposed density analysis model.

Table 2: Comparative Analysis of ANFGuIS

Techniques	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)
Proposed ANFGuIS	98.57	98.43	98.88	98.66
ANFIS	95.42	95.08	95.91	95.49
DLNN	93.76	93.21	94.18	93.69
RBFNN	91.33	90.74	92.05	91.39
MLP	89.64	88.97	90.21	89.59

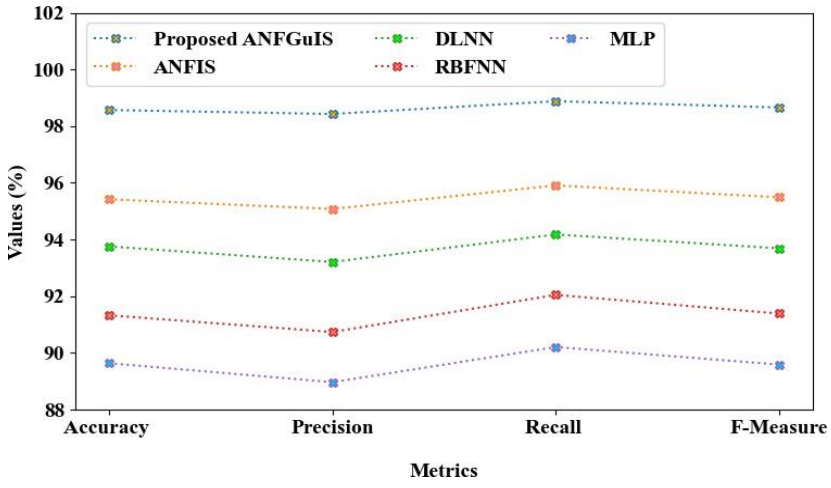

Figure 4: Comparison regarding AHP

As per Table 2 and Figure 4, the proposed AHP model labeled and predicted the accident hotspot with an accuracy of 98.87%, precision of 98.43%, recall of 98.88%, and F-Measure of 98.66%. Yet, the traditional ANFIS, Deep Learning Neural Network (DLNN), Radial Basis Function Neural Network (RBFNN), and Multi-Layer Perceptron (MLP) achieved an average accuracy, precision, recall, and F-Measure of 92.54%, 92%, 93.09%, and 92.54%, respectively. Hence, in the proposed ANFGuIS, the twilight and latitude-based daylength analysis and density investigation improved the AHP.

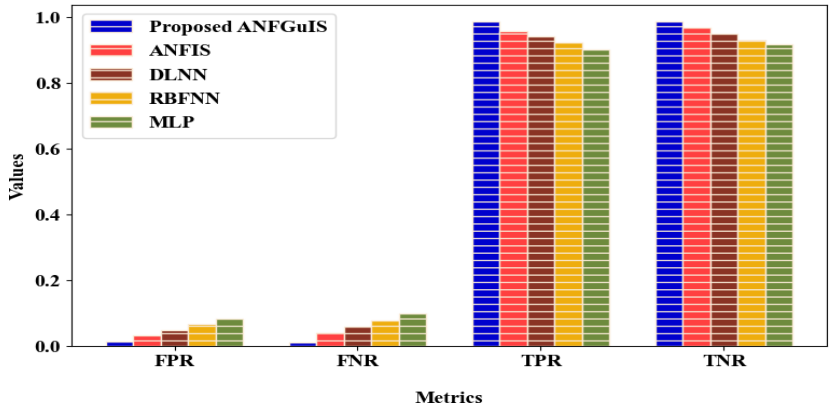

Figure 5: Graphical Comparison of ANFGuIS

The proposed ANFGuIS utilized the Gudermannian membership function for determining the degree of relationship between the input data. Thus, the proposed ANFGuIS predicted the accident hotspot with a False Positive Rate (FPR) of 0.0124, False Negative Rate (FNR) of 0.0115, True Positive Rate (TPR) of 0.9885, and True Negative Rate (TNR) of 0.9876. However, the existing ANFIS, DLNN, RBFNN, and MLP achieved FPRs of 0.0311, 0.0489, 0.0678, and 0.0821, FNRs of 0.0411, 0.0585, 0.0765, and 0.0981, TPRs of 0.9589, 0.9415, 0.9235, and 0.9019, and TNRs of
0.9689, 0.9511, 0.9322, and 0.9179, respectively. Therefore, the proposed model performed better than existing techniques.

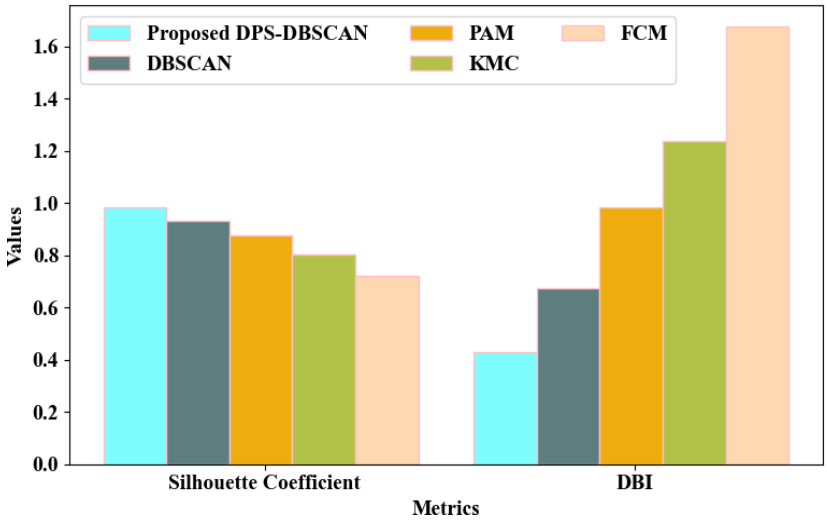

Figure 6: Comparative Analysis of DPS-DBSCAN

The proposed clustering model utilized PS distance and Dirichlet function for epsilon and minpts evaluations, respectively. As depicted in Figure 6, the proposed DPS-DBSCAN attained a Silhouette coefficient of 0.9842 and Davies-Bouldin Index (DBI) of 0.4273. Nevertheless, the conventional DBSCAN, Partition Around Medoid (PAM), K-Means Clustering (KMC), and Fuzzy C-Means (FCM) obtained Silhouette coefficients of 0.9321, 0.8748, 0.8042, and 0.7214 and
DBIs of 0.6742, 0.9841, 1.2362, and 1.6749, respectively. Thus, the proposed DPS-DBSCAN outperformed the baseline models.

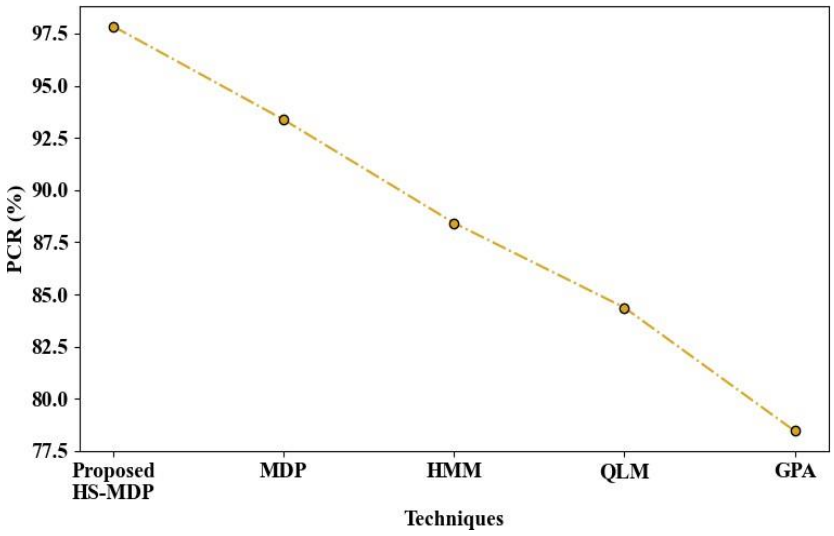

Figure 7: Comparative Analysis of HS-MDP

As illustrated in Figure 7, the proposed HS-MDP and the existing MDP, Hidden Markov Model (HMM), Q-Learning Model (QLM), and Greedy Policy Algorithm (GPA) intervened in road safety with Policy Convergence Rates (PCRs) of 97.84%, 93.37%, 88.42%, 84.37%, and 78.46%, respectively. Hence, due to the usage of the HS function to determine the transition-regulation, the performance of the proposed road safety intervention method was enhanced.

Table 3: Existing Works' Comparison

Study	Methods	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)
Proposed Work	ANFGuIS	98.87	98.43	98.88	98.66
(Jing et al., 2025)	Sequential Gated Recurrent Unit Attention Network (Seq-GRUAttNet)	90.21	91.42	90.16	91.7
(Amorim et al., 2023)	Artificial Neural Network (ANN)	83	84	83	82
(Guo et al., 2024)	Sequential Feature Recurrent Attention Network (SFRAN)	79.5	82.3	86	84.2
(Ulu et al., 2024)	Support Vector Machine (SVM)	87.5	87.6	87.2	87.3

(Khosravi et al., 2024)	K-Nearest Neighbor (KNN)	71	75	72	-

Table 3 describes the comparison of the proposed work and prevailing works regarding AHP. When contrasted with the existing works, the proposed ANFGuIS effectively detected the accident hotspot with higher accuracy, precision, recall, and F-Measure values due to the integration of daylength analysis, density analysis, and accident progression modeling. On the other hand, the traditional works either didn't consider the weather conditions or didn't analyze the twilight conditions regarding the latitude of the accident location. This led to poor performance of the prevailing works when compared to the proposed work. Hence, the enhanced performance of the proposed system was proved.

5. CONCLUSION

In the proposed system, the AHP was performed regarding GIS data. The road accident data was collected and pre-processed. Then, the weather-based clustering was carried out using DPS-DBSCAN with a Silhouette coefficient of 0.9842. Next, the feature extraction and daylength analysis were carried out. Afterward, the daylength-based density was analyzed using IHT-KDE, achieving an MAE of 0.078. Meanwhile, the accident progression was modeled using KG, and the respective features were extracted. Thereafter, the accident hotspots were labeled and predicted using ANFGuIS with an accuracy of 98.87%. Finally, the road safety intervention was performed using HS-MDP, which obtained a PCR of 97.84%. Thus, the proposed framework precisely predicted the accidental hotspot for road safety improvement.

Practical Implication: The proposed framework can be utilized by the transportation department for prioritizing the high-risk zones. It can effectively optimize the timing for traffic signal control in smart cities during the shorter daylight periods.

Future Scope: In the future, the real-time traffic sensor data will also be integrated to enable an effective dynamic hotspot detection system.

REFERENCES:

1. Dataset Link: https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents
2. Aldalain, S. A., Sukor, N. S. A., Obaidat, M. T., & Manan, T. S. B. A. (2023). Road Accident Hotspots on Jordan ' s Highway Based on Geometric Designs Using Structural Equation Modeling. *Applied Sciences*, *13*, 1–21. https://doi.org/10.3390/app13148095
3. Alsahfi, T. (2024). Spatial and Temporal Analysis of Road Traffic Accidents in Major Californian Cities Using a Geographic Information System. *International Journal of Geo-Information*, *13*, 1–22. https://doi.org/10.3390/ijgi13050157
4. Amorim, B. de S. P., Firmino, A. A., Baptista, C. de S., Júnior, G. B., Paiva, A. C. de, & Junior,
5. F. E. de A. (2023). A Machine Learning Approach for Classifying Road Accident Hotspots.
6. *International Journal of Geo-Information*, *12*, 1–17. https://doi.org/10.3390/ijgi12060227
7. Balawi, M., & Tenekeci, G. (2024). Time series traffic collision analysis of London hotspots : Patterns , predictions and prevention strategies. *Heliyon*, *10*(4), 1–29. https://doi.org/10.1016/j.heliyon.2024.e25710
8. Berhanu, Y., Schroder, D., Wodajo, B. T., & Alemayehu, E. (2024). Machine learning for predictions of road traffic accidents and spatial network analysis for safe routing on accident and congestion-prone road networks. *Results in Engineering*, *23*, 1–13. https://doi.org/10.1016/j.rineng.2024.102737
9. Deressa, B. F., Habtegiogis, K. A., Endashaw, D. K., Al-Ramadan, B. M., & Al-Ahmadi, H. M.

(2025). A Systematic Review of GIS-Driven Road Traffic Accident Evaluation. *Vehicle*, *7*, 1–23. https://doi.org/10.3390/vehicles7040161
10. Guo, W., Xu, C., & Jin, S. (2024). Fusion of satellite and street view data for urban traffic accident hotspot identification. *International Journal of Applied Earth Observation and Geoinformation*, *130*, 1–15. https://doi.org/10.1016/j.jag.2024.103853
11. Habib, M. F., Bridgelall, R., Motuba, D., & Rahman, B. (2023). Exploring the Robustness of Alternative Cluster Detection and the Threshold Distance Method for Crash Hot Spot Analysis : A Study on Vulnerable Road Users. *Safety*, *9*, 1–21. https://doi.org/10.3390/safety9030057
12. Hazaymeh, K., Almagbile, A., & Alomari, A. H. (2022). Spatiotemporal Analysis of Traffic Accidents Hotspots Based on Geospatial Techniques. *International Journal of Geo-Information*, *11*, 1–18. https://doi.org/10.3390/ijgi11040260
13. Islam, K., Reza, I., Gazder, U., Akter, R., Arifuzzaman, & Rahman, M. M. (2022). Predicting Road Crash Severity Using Classifier Models and Crash Hotspots. *Applied Sciences*, *12*, 1–22. https://doi.org/10.3390/app122211354
14. Jing, J., Guo, W., Bai, C., & Jin, S. (2025). Highway Accident Hotspot Identification Based on the Fusion of Remote Sensing Imagery and Traffic Flow Information. *Big Data and Cognitive Computing*, *9*, 1–21. https://doi.org/10.3390/bdcc9110283
15. Kamh, H., Alyami, S. H., Khattak, A., Alyami, M., & Almujibah, H. (2025). Exploring Road Traffic Accidents Hotspots Using Clustering Algorithms and GIS-Based Spatial Analysis. *IEEE Access*, *13*, 60944–60954. https://doi.org/10.1109/ACCESS.2024.3443245
16. Khosravi, Y., Hosseinali, F., & Adresi, M. (2024). Identifying accident prone areas and factors influencing the severity of crashes using machine learning and spatial analyses. *Scientific Reports*, *14*, 1–19. https://doi.org/10.1038/s41598-024-81121-7 1
17. Ma, Q., Huang, G., & Tang, X. (2021). GIS - based analysis of spatial – temporal correlations of urban traffic accidents. *European Transport Research Review*, *13*, 1–11.
18. https://doi.org/10.1186/s12544-021-00509-y
19. Moreno-ponce, L. A., Pérez-zuriaga, A. M., & García, A. (2025). Predictive Models and GIS for Road Safety : Application to a Segment of the Chone – Flavio Alfaro Road. *Sustainability*, *17*, 1–24. https://doi.org/10.3390/su17115032
20. Santos, D., Quaresma, J. S. P., & Nogueira, V. B. (2021). Machine Learning Approaches to Traffic Accident Analysis and Hotspot Prediction. *Computers*, *10*, 1–15. https://doi.org/10.3390/computers10120157
21. Ulu, M., Kilic, E., & Türkan, Y. S. (2024). Prediction of Traffic Incident Locations with a Geohash-Based Model Using Machine Learning Algorithms. *Applied Sciences*, *14*, 1–21. https://doi.org/10.3390/app14020725
22. Wang, Z., Lu, Y., Zou, Z., Ma, Y., & Wang, T. (2022). Applying OHSA to Detect Road Accident Blackspots. *International Journal of Environmental Research and Public Health*, *19*, 1–15. https://doi.org/10.3390/ijerph192416970
23. Yalcin, M. A., Kofteci, S., San, B. T., & Burgan, H. I. (2026). A GIS-Based Approach to Analyzing Traffic Accidents and Their Spatial and Temporal Distribution : A Case Study of the Antalya City Center. *International Journal of Geo-Information*, *15*, 1–18. https://doi.org/10.3390/ijgi15010019
24. Yan, R., Hu, L., Li, J., & Lin, N. (2024). Accident Severity Analysis of Traffic Accident Hot Spot Areas in Changsha City Considering Built Environment. *Sustainability*, *16*, 1–15. https://doi.org/10.3390/su16073054