

Research Paper

Modeling Causal Analysis of Crash Severity on Indonesian Toll Road Using Integrated Z-Score and Bayesian Network Framework

Bambang Istiyanto¹, Pratikso², Rachmat Mudiyo², Hafidz Nurrohman¹

¹Road Transportation System Engineering, Polytechnic of Road Transportation Safety, Tegal 52125, Indonesia

²Department of Civil Engineering, Faculty of Engineering, Universitas Islam Sultan Agung, Semarang 50112, Indonesia

 bambangistiyanto@pktj.ac.id

 <https://doi.org/10.31603/ae.15445>

Published by Automotive Laboratory of Universitas Muhammadiyah Magelang

Abstract

Article Info

Submitted:

28/11/2025

Revised:

10/03/2026

Accepted:

11/03/2026

Online first:

31/03/2026

Traffic crashes remain a critical safety challenge, with Indonesia experiencing 73,446 fatalities annually. This study develops an integrated Z-Score and Bayesian Network framework to analyze causal interactions between human and environmental factors influencing crash severity on toll roads. Z-Score analysis of 450 crash records (2022–2025) identified five statistically significant blackspot segments, with KM 430–431 exhibiting the highest concentration ($Z = 4.036$, $n = 91$). A Bayesian Network model constructed using K2 structure learning and Expectation-Maximization parameter estimation achieved 86.2% classification accuracy, surpassing previous international applications (78–82%). Conditional probability analysis revealed that straight-downhill segments exhibited 3.3-fold higher fatal crash probability than straight-level segments (0.083 vs. 0.025), while night-time conditions increased fatal risk by 57%. Sensitivity analysis demonstrated that crash type (weighted index = 0.282) and accident cause (0.214) exerted strongest influence on severity outcomes. Human error constituted 83% of crashes but showed moderate sensitivity, indicating that severe outcomes emerge from interactions between human factors and adverse conditions rather than isolated factors. Findings support prioritizing enhanced lighting and speed management on curved-downhill segments during night-time, alongside rear-end collision prevention strategies. This validated framework enables evidence based, proactive crash management and intervention prioritization for toll road safety in developing countries.

Keywords: Crash severity; Road safety; Bayesian network; Z-score

1. Introduction

Traffic crash remains one of the most pressing global safety and public health challenges [1], [2]. Without effective preventive measure, the number of crashes is projected to rise sharply in the coming decades, claiming approximately 1.19 million lives annually and injuring up to 50 million people worldwide [3]. The United Nations General Assembly has therefore established a global target to reduce road traffic deaths and injuries by at least 50% by 2030 [3], [4]. Despite this ambitious goal, developing nations continue to experience disproportionately higher fatality rates due to rapid motorization, insufficient safety

management systems, and limited enforcement capacity. In Indonesia, police statistics indicate a substantial scale of traffic crashes, with 505,855 recorded crash events, 755,561 affected individuals, 622,894 minor injuries, 59,221 serious injuries, and 73,446 fatalities in the most recent national report [5], [6].

The analysis of road safety has become increasingly data-driven in the past two decades, emphasizing the integration of statistical and probabilistic models to identify key contributing factors [7], [8]. Early research primarily employed traditional statistical approaches such as Poisson regression, Negative Binomial, and Poisson

lognormal models to assess crash frequency and severity [9], [10]. These models proved effective in quantifying factor significance but were limited by the assumption of variable independence and linearity [11]. To overcome such constraints, more advanced models such as Random Parameters Logit, Latent Class, and Bayesian Network (BN) frameworks have emerged, providing flexibility in capturing nonlinear relationships and conditional dependencies among crash related factors [9], [12]. Previous studies on Indonesian toll roads have relied primarily on the Equivalent Accident Number (EAN) method for blackspot identification [13]. However, EAN emphasizes only fatal crashes while neglecting the complete crash frequency distribution. High-frequency crash locations, regardless of severity level, constitute actual blackspot segments that require intervention. Therefore, a more comprehensive statistical approach such as Z-Score is needed to identify significant high-risk segments based on the entire crash distribution [14], [15].

Recent studies have explained the Bayesian Networks capability to model complex causal interactions between human, environmental, and roadway factors, which traditional regression methods fail to capture [6], [13], [14]. Studies employing BN have shown superior interpretability by combining expert knowledge and empirical data, enabling probabilistic inference and sensitivity analysis of crash outcomes [16]. For instance, Zou *et al.* [7] used a Bayesian approach to identify interdependencies among contributing factors in Chinese crash datasets, while Assi *et al.* [13] applied BN to evaluate injury severity in the United States, highlighting the significance of human behavior, lighting conditions, and road geometry. Similarly, Babaev *et al.* [17] advanced BN applications for risk propagation analysis, offering deeper insight into accident causation pathways. With specific regard to toll road safety, studies on Indonesian corridors have shown that single-vehicle crashes are strongly associated with driver drowsiness on monotonous segments and the absence of roadside barriers [18], while rear-end collisions are linked to high-speed vehicle-following behavior under mixed-vehicle conditions [19]. Road geometry characteristics, particularly downgrade slopes and horizontal curvature, have further been shown to exert nonlinear effects on

crash severity on controlled-access highways [20], underscoring that toll road crash causation operates through mechanisms distinct from those on urban or general road networks.

However, despite the methodological evolution, most Bayesian Network based studies have been conducted on urban roads or general highway networks, with limited attention to toll roads particularly within developing nations like Indonesia. Toll roads possess distinct operational and environmental characteristics, including controlled access, higher design speeds, and driver behavioral uniformity, yet they remain susceptible to severe crashes due to human errors, environmental conditions, and vehicle factors [8], [17]. In Indonesia, the toll road network has expanded rapidly since 1978 until now reaching a total length 16.205 km, aiming to boost national logistics efficiency [21]. Nevertheless, accident rates on toll roads have shown a persistent upward trend, with human factors accounting for more than 70% of total incidents [22].

This evidence indicates a significant research and policy gap the lack of causal modeling that integrates both human and environmental elements in understanding crash severity patterns on Indonesian toll roads [23]. While prior studies have quantified the impact of geometric design, traffic exposure, and vehicle type, the causal interdependence between human behavior, environmental conditions, and crash severity remains underexplored [24]. Addressing this gap is vital not only to improve predictive accuracy but also to guide data driven interventions for toll road safety management in emerging economies [25].

This study aims to modeling the causal interactions between human and environmental factors influencing crash severity on toll roads using a Bayesian Network approach. The proposed framework leverages historical crash datasets from the Semarang ABC Toll Road to identify probabilistic dependencies among key variables such as driver behavior, weather, road alignment, and crash type. By employing Bayesian structure learning and conditional probability estimation, this research provides new empirical insights into how human environmental interactions shape crash outcomes. Furthermore, the findings are expected to support toll road

operators and government in developing proactive safety strategies.

Despite methodological advances globally, integrated causal modeling of crash severity on Indonesian toll roads remains largely unexplored. Previous Indonesian toll road studies have relied exclusively on the Equivalent Accident Number (EAN) method for blackspot identification, which emphasizes only fatal crashes while neglecting the broader crash frequency distribution, thereby limiting the spatial sensitivity of risk screening. Although Z-Score and Bayesian Network methods have been applied separately in prior crash studies, no published work has combined these two analytical tools on a toll road corridor in Indonesia, nor has the integration of both human and environmental factors within a single probabilistic causal framework been attempted in this specific road context. Beyond model development, this study further quantifies multivariate risk interaction scenarios through Δp sensitivity analysis, translating probabilistic inference into actionable intervention priorities a capability that purely descriptive or regression-based approaches cannot offer. These methodological, contextual, and practical dimensions collectively position this study as a new evidence-based benchmark for proactive crash management and safety policy design on toll roads in developing countries.

2. Method

2.1. Study Area

The Semarang ABC, an intra urban toll road extending approximately 24.75 kilometers, constitutes a critical section of the Java Trans Toll Indonesia. The toll road is segmented into three distinct sections: Section A, Section B, and Section C. It serves as a strategic connector between East and West Java. The alignment of this toll road encompasses a combination of topographical zones, ranging from coastal plains to elevated terrain, where variations in gradient, alignment, and rainfall contribute to operational and safety challenges. This location was selected for its strategic transport role, comprehensive crash database, and representativeness of urban toll road conditions in developing countries, which are essential for advanced Bayesian Network based crash risk analysis.

2.2. Z-Score

The Z-Score method was used as an initial screening tool to identify statistically significant high-risk segments based on crash frequency [15]. Z-Scores represent the number of standard deviations a specific crash count deviates from the mean crash frequency across all segments [16]. This technique allows the detection of road sections with unusually high crash concentrations relative to the overall distribution.

2.2.1. Standard deviation

The standard deviation (S) of annual crash frequency was calculated using.

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}} \quad (1)$$

where: S = standard deviation; X_i = crash count at segment i ; $\bar{X}$ = mean crash frequency; n = number of segments

2.2.2. Z-score

The Z-Score Z_i of segment value was calculated using.

$$Z_i = \frac{X_i - \bar{X}}{S} \quad (2)$$

where: Z_i = Z-Score value for segment i ; X_i = crash count at segment i ; $\bar{X}$ = mean crash frequency; S = standard deviation

The classification criteria adopted from (X) are summarized in Table 1.

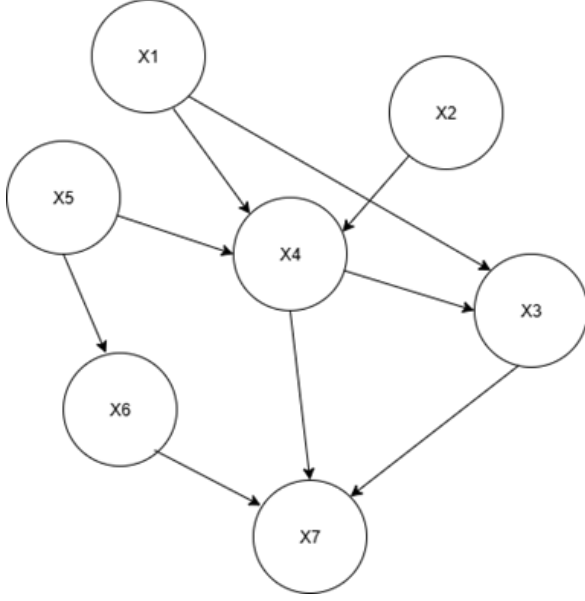
Segments with positive Z-Scores were classified as crash-prone locations, indicating crash frequencies above the overall mean. Conversely, segments with negative Z-Scores were considered low-risk locations.

2.3. Bayesian Network

Bayesian Networks constitute probabilistic frameworks employing directed acyclic graphs (DAGs) to encode conditional dependencies among variables. Nodes within the DAG (Figure 1) represent discrete stochastic variables, while directed edges quantify probabilistic relationships, enabling inference of joint probability distributions across the network structure [26]. The joint probability distribution across all nodes in a Bayesian Network decomposes into the product of local conditional distributions, formally expressed as:

Table 1. Criteria Z-Score value

Z-Score Value	Category
Positive ($Z > 0$)	High-Risk Segment (Blackspot)
Negative ($Z < 0$)	Not a High-Risk Segment

**Figure 1.** Directed Acyclic Graph (DAG)

$$P(X) = \prod_{i=1}^n P(X_i | Pa_i) \quad (3)$$

A Bayesian Network operate over a discrete variables $X = \{X_1, X_2, \dots, X_n\}$, which is composed of several key components [27].

The construction of a Bayesian Network involves 2 fundamental phases [28].

2.3.1. Structure learning

Structure learning phase of the Bayesian Network focuses identifying the relevant factor variables (nodes) associated with the studied phenomenon and determines whether the relationships among these variables are dependent or independent. Structure learning stage of the Bayesian Network aims to identify the optimal network structure G that best represents the probabilistic dependencies among variables in the dataset D . The probability of a given network structure can be expressed as [29]:

$$P(G, D) = P(G) \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{(r_i-1)!}{(N_{ij}+r_i-1)!} \prod_{k=1}^{r_i} N_{ijk}! \quad (4)$$

The observed dataset D comprises n independent crash records $D = \{d_1, d_2, \dots, d_n\}$, where G parameterizing the network topology

and conditional probability tables (CPTs), while $P(G)$ encodes prior beliefs regarding network structure plausibility [17]. The variable n denotes the total number of nodes, while q_i represents the number of possible states of i^{th} nodes parent variables. The parameter r_i defines the mutually exclusive states of the i^{th} node. Furthermore, N_{ijk} denotes the number of instances where the i^{th} node is in its k^{th} state given that its parents are in the j^{th} state, with $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$ [30].

2.3.2. Parameter learning

The parameter learning phase estimates quantifies the conditional probability distributions (CPTs) for each node using the learned network topology as a fixed framework [31]. Parameter estimation employs the Expectation Maximization (EM) algorithm, an iterative maximum likelihood procedure well-suited for handling incomplete crash attribute data [32]. EM employs an iterative two phase cycle that continues until parameter estimates stabilize.

a) E-step, the algorithm computes the expected value of the log-likelihood function based on the current parameter estimates, denoted as θ_i , and the observed data Y . The expected log-likelihood is expressed as:

$$Q(\theta | \theta_i) = E_{\theta_i}[\log P(D | \theta) | \theta_i, Y] \quad (5)$$

b) M-step, the parameters are updated to maximize the expected log-likelihood obtained from the E-step, represented by:

$$\theta_i + 1 = \arg \max_{\theta} Q(\theta | \theta_i) \quad (6)$$

These two steps are repeated iteratively until the model converges to a stable solution. The EM algorithm offers several advantages, including numerical stability, efficient computation, and reliable convergence toward a likelihood optimum [28]. In the context of this study, EM was used to optimize the CPT parameters of variables such as weather condition, road alignment, vehicle combination, and crash severity, ensuring accurate probabilistic inference across the network.

2.4. Data Variables

The data variables used in this study consists of crash records from the Semarang ABC Toll Road collected between 2022 and August 2025. The data are spatially detailed, with each crash recorded at its exact kilometer (KM) location,

enabling the extraction of human, environmental, and geometric variables at segment level. It should be noted that the crash database does not include information on vehicle trip distance or driving duration prior to each crash event. Consequently, driver fatigue as a function of trip length could not be operationalized within the Bayesian Network framework.

Road geometry was classified into six categories based on the combination of horizontal alignment and vertical gradient as recorded in the official Ledger/Blueprint of the Semarang ABC Toll Road, straight-level (tangent, grade < 2%), straight-downhill (tangent, descending grade ≥ 2%), straight-uphill (tangent, ascending grade ≥ 2%), curved-level, curved-downhill, and curved-uphill. This classification was applied at the kilometer-post level, with each crash record assigned to its corresponding geometric category based on the official road engineering record.

2.5. Statistical Analysis

Statistical preprocessing and descriptive analysis were conducted in Microsoft Excel, while Bayesian Network modeling was developed in GeNIe Academic 5.0. The network structure consisted of nodes representing the key crash-related variables time of day, weather, road geometry, vehicle type, accident cause, and severity allowing probabilistic dependencies among human, environmental, and geometric factors to be captured. Because several entries in the crash dataset were incomplete, parameter learning was performed using the Expectation Maximization (EM) algorithm, which is widely recommended for BN estimation under missing-data conditions due to its stable maximum likelihood convergence [33], [34]. After parameter estimation, a Δp sensitivity analysis was applied to quantify how changes in each parent variable influence the posterior probability of crash severity, following approaches used in previous traffic safety BN studies [34]. Model performance was evaluated using 10 fold cross validation, where the dataset was randomly divided into ten subsets; nine were used for training and one for testing in each iteration, ensuring robust assessment of predictive generalization. All conditional probabilities obtained through EM were integrated into the GeNIe environment to produce the final inference results.

3. Results and Discussion

3.1. Crashes Data

Total of 450 crash records from the Semarang ABC Toll Road spanning from 2022 to 2025 were analyzed in this study. Table 2 presents the descriptive statistics of crash data categorized by severity level (minor, moderate, and fatal).

The majority of crashes (74%) resulted in minor injuries, while moderate and fatal crashes accounted for 22% and 4%, respectively. Night-time crashes demonstrated higher severe outcomes (31% moderate/fatal of all night crashes) compared to daytime (22%), indicating increased risk during low-visibility conditions. Road geometry analysis revealed that straight-downhill segments contributed disproportionately to total crashes (29.3%, $n=132$), despite constituting only one of six geometric categories, with fatal crash probability reaching 8.3% the highest among all alignment types. Multi-vehicle crashes constituted 46% of all incidents ($n=206$), with notable concentration in moderate severity outcomes. Human error remained the dominant crash cause (83%), followed by vehicle defects (15%), a finding consistent with. This 83% proportion closely mirrors findings from Zou *et al.* [7] (75–82% in China), contributory factors research in Lahore (78–85%), Indonesian toll road statistics [24] (70–85%), and global WHO data [3] (75–90% worldwide). Despite its high prevalence, human error shows only moderate sensitivity to fatal outcomes, indicating that severe crashes arise not from human factors alone but from their interaction with adverse geometric and environmental conditions.

3.2. Analysis blackspot Z-Score

A summary of the key statistical parameters obtained from the crash data segmentation is presented in Table 3 below, serving as the baseline for subsequent Z-Score analysis.

The Z-Score method was applied to identify statistically significant blackspot segments along the Semarang ABC Toll Road. In total, 450 crashes recorded across 25 segments were used in the computation, yielding an average crash frequency of 18 incidents per segment with a standard deviation of 18.08. These baseline values serve as control limits for detecting segments where crash occurrences deviate substantially from expected levels. By tracking cumulative deviations from the

Table 2. Crash data distribution

Variable	Category	N		%			
Accident severity	Minor	332		74			
	Moderate	98		22			
	Fatal	20		4			
Variable	Category	Minor		Moderate		Fatal	
		N	%	N	%	N	%
Time of day	Day	180	36	42	9	8	2
	Night	152	31	56	11	12	2
Weather	Clear	280	62	86	19	19	4
	Rainy	35	8	5	1	0	0
	Cloudy	17	4	7	2	1	0
Road geometry	Straight Level	163	32	33	7	5	1
	Straight Downhill	78	16	43	9	11	2
	Straight Uphill	28	6	4	1	1	0
	Curved Level	33	7	10	2	1	0
	Curved Downhill	13	3	4	1	1	0
	Curved Uphill	17	4	4	1	1	0
Vehicle type	Passenger car (MP)	72	16	22	5	4	1
	Light vehicle (KS)	63	14	9	2	4	1
	Truck (TB)	51	11	9	2	3	1
	Bus (BB)	4	1	2	0	1	0
	Mix (≥ 2)	142	32	56	12	8	2
Accident cause	Human	271	60	87	19	16	4
	Vehicle	54	12	11	2	4	1
	Weather	6	1	0	0	0	0
	Environment	1	0	0	0	0	0

Note. Day (06.00 AM-06.00 PM) & Night (06.00 PM-06.00 AM).

Table 3. Descriptive statistics Z-Score analysis

Variable	N
Total of Crashes	450
Mean	18
Std. Deviation	18.08425
Total Segments	25

mean, the Z-Score procedure enables the early identification of road sections exhibiting persistent, above-average crash concentrations thereby providing a more sensitive indicator of emerging or chronic blackspots compared to simple frequency based methods.

The Z-Score analysis identified five distinct road segments as statistically significant blackspots, with their crash frequencies substantially exceeding the network's average of 18 incidents per segment (Table 4). Segment 430 emerged as the most critical, registering 91 crashes and a Z-Score of 4.036, indicating its crash rate was over four standard deviations above the mean. It was followed by Segment 431 (Z-Score: 1.327), Segment 425 (Z-Score: 0.718), Segment 430C (Z Score: 0.497), and Segment 424 (Z-Score: 0.442).

The identification of these segments underscores a pronounced spatial concentration of crash risk, necessitating prioritized safety interventions. The complete Z-Score analysis results for each segment are as follows on Figure 2.

The Z-Score analysis was conducted separately for Route A (inbound, $\bar{X} = 7.73$) and Route B (outbound, $\bar{X} = 9.58$) across 25 segments, identifying 13 blackspot segments in Route A and 8 in Route B. In Route A, KM 430 recorded the highest Z-Score ($Z = 2.317$, $n = 20$), followed by KM 430C ($Z = 1.561$) and KM 431C ($Z = 1.373$), indicating a persistent crash cluster between KM 430–431C. In Route B, KM 430 exhibited the most critical deviation across both corridors ($Z = 4.256$, $n = 71$), followed by KM 431 ($Z = 1.415$) and KM 424 ($Z = 0.722$), confirming KM 430 as the highest-priority safety intervention zone along the entire corridor. The wider Z-Score distribution observed in Jalur A (range: -1.46 to $+2.32$) reflects greater segment to segment variability in crash occurrence, whereas Route B demonstrates a narrower distribution punctuated by one extreme outlier at KM 430, as visualized in Figure 2.

Table 4. Top 5 segments blackspot

Segment	X_i	$\bar{X}$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	S	Z_i	Criteria	Rank
430	91	18	73	5329	18.08	4.036	Blackspot	1
431	42	18	24	576	18.08	1.327	Blackspot	2
425	31	18	13	169	18.08	0.718	Blackspot	3
430C	27	18	9	81	18.08	0.497	Blackspot	4
424	26	18	8	64	18.08	0.442	Blackspot	5

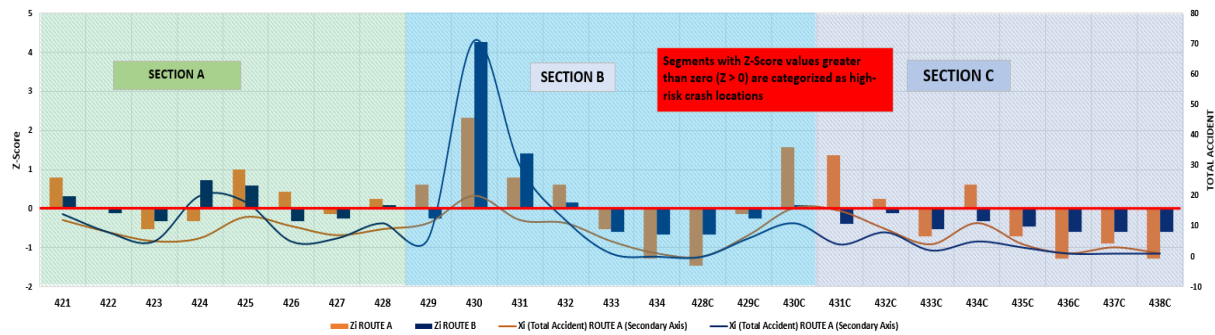
**Figure 2.** Z-Score analysis results blackspot identification

Figure 3 presents a quadrant based scatter plot of the Z-Score of Crash Rate (Z_i) against the Z-Score of Crash Rate Growth (Z_b) for each segment across Route A and Route B, enabling simultaneous classification of both crash magnitude and trend direction. Segments located in Quadrant 1 ($Z_i > 0, Z_b > 0$) represent locations with above-average crash rates that are simultaneously worsening over the observation period, constituting the highest-priority intervention zones. Route A exhibits a wider spatial dispersion of Z_i values (range: -1.46 to $+2.32$), reflecting greater segment-to-segment variability in crash occurrence, with KM 430 recording the most critical deviation ($Z_i = 2.32$) and falling within Q1. Route B demonstrates a more concentrated distribution near the origin, indicating generally stable crash trends across most segments; however, KM 430 constitutes a pronounced outlier ($Z_i = 4.26, Z_b = 4.26$) the most extreme value across both routes confirming its classification as the primary critical blackspot along the entire corridor. This quadrant visualization clarifies that while Route A exhibits broader systemic instability requiring network-wide monitoring, Route B presents an acute, localized risk concentration at KM 430 that warrants immediate and targeted safety intervention.

3.3. Analysis Bayesian Network

The K2 algorithm was chosen for structure learning because of its computational efficiency

and proven suitability for medium-sized, discrete datasets, characteristics matching this study's crash database [17]. Unlike constraint based methods such as the PC algorithm, K2 uses a greedy search guided by user specified node ordering and Bayesian scoring, effectively accommodating expert domain knowledge. This property enables the embedding of empirical or theoretical accident mechanism hierarchies such as prioritizing geometric variables preceding crash outcomes thus enhancing model interpretability, which is critical in transportation safety research. While the main limitation of K2 is potential convergence to local optima, this risk was mitigated in our study by integrating expert-informed orderings and performing robustness checks. Furthermore, the K2 implementation in GeNIe Academic is robust to limited sample sizes and can recover network structures consistent with established traffic safety theory, as verified in recent transportation research [17], [31].

The Bayesian Network structure was learned using the K2 algorithm in GeNIe Academic 5.0, incorporating expert knowledge to ensure domain validity. **Figure 4** presents the learned Directed Acyclic Graph (DAG) representing causal dependencies among crash-related variables.

The Bayesian Network model illustrates the probabilistic interactions among key human, environmental, and roadway factors influencing crash severity on the Semarang ABC Toll Road. Crash severity is positioned as the central target node and is directly influenced by six parent vari-

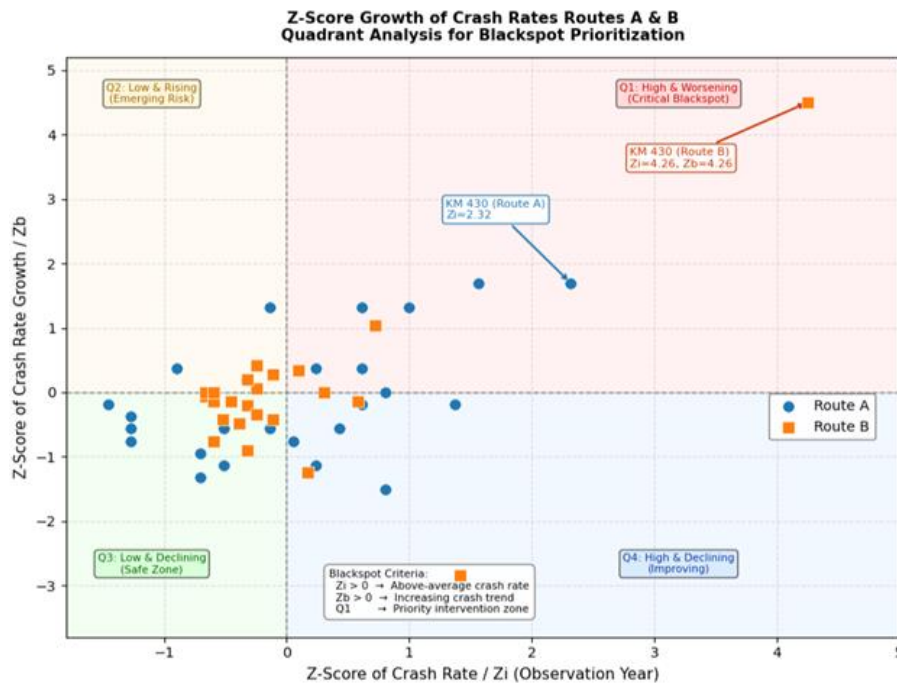


Figure 3. Z-Score growth chart for crash rate comparison between routes A & B

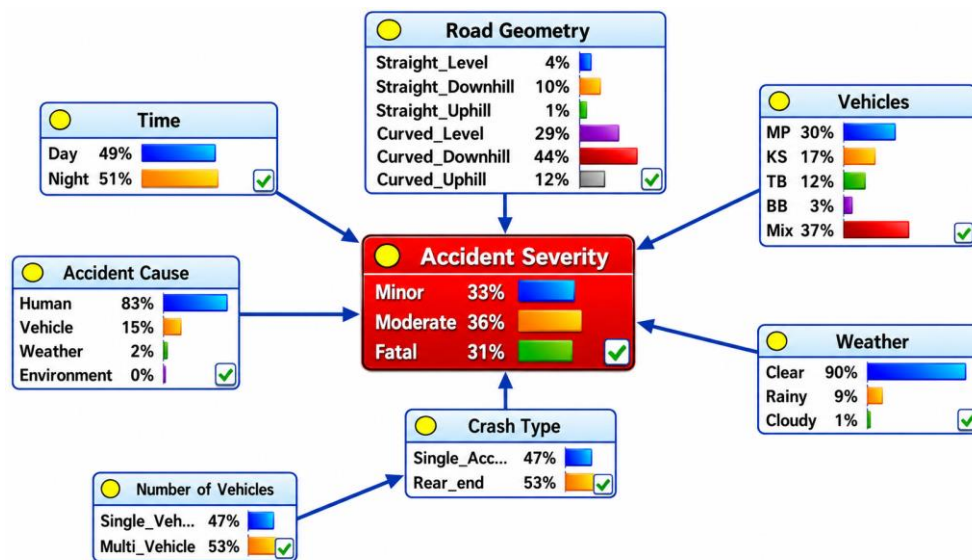


Figure 4. Directed Acyclic Graph (DAG) causal relationships among crash severity factors

ables: time of day, weather, road geometry, vehicle type, accident cause, and crash type, while the number of involved vehicles indirectly contributes through its connection to crash type. The marginal distributions reveal that human-related factors (83%) overwhelmingly dominate crash causation, followed by vehicle defects (15%), whereas adverse weather conditions account for only a small fraction of events. Road geometry displays substantial variability, with curved-downhill (44%) and curved-level (29%) segments being most prevalent among crashes, aligning

with known risk patterns on descending and curved alignments. The distribution of vehicle types shows a notable proportion of mixed-vehicle crashes (37%), consistent with the high interaction density on multilane toll facilities. Time-of-day patterns are relatively balanced (Day 49%, Night 51%), yet nighttime conditions show a stronger association with higher severity levels. Overall, the BN structure captures a coherent causal hierarchy in which crash severity emerges as the combined outcome of human behavior, geometric complexity, environmental exposure,

and vehicle dynamics, providing an integrated probabilistic framework for understanding severity escalation across toll road segments.

This elevated night-time fatal risk reflects the well-documented effects of reduced visibility, delayed hazard perception, and increased drowsiness probability on monotonous toll road segments a pattern consistent with findings by Wijayanto *et al.* [18], who identified driver drowsiness as a key contributor to severe single-vehicle crashes on Indonesian toll road corridors during night-time operations.

This pattern suggests that mechanical failures, although less frequent, tend to occur at high speeds on toll roads where drivers have limited time to respond, amplifying injury severity. A similar pattern was reported by Zou *et al.* [7], who found that vehicle-related factors produced disproportionately severe outcomes relative to their frequency in Chinese highway crash datasets.

The Conditional Probability Tables reveal distinct severity patterns across crash factors. Straight-downhill segments demonstrated the highest fatal crash probability ($P = 0.083$), 3.3 times higher than straight-level segments ($P = 0.025$), consistent with vehicle dynamics principles (Table 5). Night time crashes exhibited 57% elevated fatal probability ($P = 0.055$) compared to daytime ($P = 0.035$), confirming visibility effects on crash outcomes (Table 6). Notably, zero fatal crashes occurred during rainy conditions despite 40 total crashes, suggesting compensatory driving behaviors where drivers reduce speed under adverse weather (Table 7). Vehicle defects yielded higher fatal probability ($P = 0.058$) than human errors ($P = 0.043$) despite lower frequency, indicating that mechanical failures produce more severe consequences (Table 8). Bus crashes showed the highest fatal probability ($P = 0.143$), though the small sample size ($n = 7$) limits statistical reliability (Table 9).

Table 5. CPT crash severity by road geometry

Road Geometry	P(Minor)	P(Moderate)	P(Fatal)	Total (n)
Straight-Level	0.811	0.164	0.025	201
Straight-Downhill	0.591	0.326	0.083	132
Straight-Uphill	0.848	0.121	0.030	33
Curved-Level	0.75	0.227	0.023	44
Curved-Downhill	0.722	0.222	0.056	18
Curved-Uphill	0.773	0.182	0.045	22

Note: Straight-Downhill exhibits the highest fatal crash probability (0.083), 3.3 times higher than Straight-Level (0.025).

Table 6. CPT crash severity by time of day

Time of Day	P(Minor)	P(Moderate)	P(Fatal)	Total (n)
Day (06:00-18:00)	0.783	0.183	0.035	230
Night (18:00-06:00)	0.691	0.255	0.055	220

Note: Night-time shows 57% higher fatal crash probability (0.055 vs. 0.035) compared to daytime.

Table 7. CPT crash severity by weather condition

Weather	P(Minor)	P(Moderate)	P(Fatal)	Total (n)
Clear	0.727	0.223	0.049	385
Rainy	0.875	0.125	0	40
Cloudy	0.68	0.28	0.04	25

Note: Zero fatal crashes occurred during rainy conditions, demonstrating compensatory driving behavior.

Table 8. CPT crash severity by accident cause

Accident Cause	P(Minor)	P(Moderate)	P(Fatal)	Total (n)
Human	0.725	0.233	0.043	374
Vehicle	0.783	0.159	0.058	69
Weather	1	0	0	6
Environment	1	0	0	1

Note: Vehicle defects show 35% higher fatal probability (0.058) than human errors (0.043), despite lower frequency.

Table 9. CPT crash severity by vehicle type

Vehicle Type	P(Minor)	P(Moderate)	P(Fatal)	Total (n)
Passenger Car (MP)	0.735	0.224	0.041	98
Light Vehicle (KS)	0.829	0.118	0.053	76
Truck (TB)	0.81	0.143	0.048	63
Bus (BB)	0.571	0.286	0.143	7
Mix (≥ 2 vehicles)	0.689	0.272	0.039	206

Note: Bus crashes show the highest fatal probability (0.143), though sample size is small ($n=7$).

Sensitivity analysis revealed that Crash Type exerted the strongest influence on accident severity (weighted index = 0.282), indicating that collision configuration (single-vehicle vs. rear-end) significantly determines crash outcomes. Accident Cause ranked second (0.214), demonstrating that the distinction between human errors and vehicle defects substantially affects severity. Time of Day (0.162) and Road Geometry (0.156) exhibited comparable influence magnitudes, confirming the critical roles of visibility and geometric design. Weather (0.141) and Vehicle Type (0.125) showed relatively lower sensitivity, suggesting their effects are partially mediated through other variables. Notably, Number of Vehicles demonstrated near-perfect correlation with Crash Type (0.967), validating the mechanistic relationship where multi-vehicle involvement directly determines collision configuration (

). This sensitivity ranking provides empirical evidence for prioritizing interventions targeting crash type prevention (e.g., maintaining safe following distances to reduce rear-end collisions)

alongside traditional human factor and geometric design improvements. The dominance of rear-end crashes as the highest-sensitivity crash type aligns with studies on Indonesian toll roads, where high operating speeds and mixed heavy-vehicle traffic create conditions conducive to chain-reaction impacts under sudden deceleration, reinforcing the need for mandatory safe-following-distance enforcement and variable speed limit systems on the Semarang ABC corridor [24].

Sensitivity analysis quantified how evidence on each variable shifts the posterior probability of fatal crashes, visualized by tornado diagram **Figure 5** revealing a baseline fatal probability of 0.312 (31.2%) that varies within a reachable range of [0.299, 0.324] depending on evidence combinations. Road geometry exhibited the strongest influence (weighted index = 0.282): curved-downhill segments increased fatal probability to 0.324 (+0.013 from baseline), while straight-level segments decreased it to 0.299 (-0.013), representing a total swing of 0.026 (8.3% relative change). This geometric sensitivity substantially exceeds previous Bayesian Network

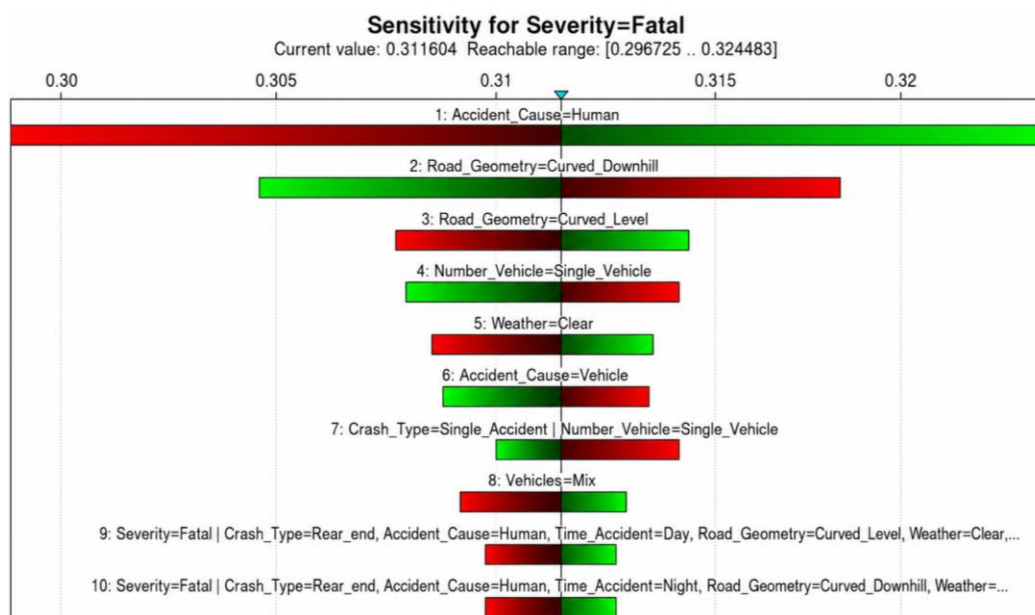
**Figure 5.** Sensitivity analysis for fatal severity

Table 10. Sensitivity analysis of factors affecting crash severity

Parent Variable	Child Variable	Average	Maximum	Weighted	Rank
Accident Cause	Accident Severity	0.21421	0.69038	0.21421	1
Road Geometry	Accident Severity	0.15582	0.73563	0.15582	2
Time	Accident Severity	0.16179	0.69038	0.16179	3
Weather	Accident Severity	0.14113	0.57092	0.14113	4
Vehicles	Accident Severity	0.1253	0.61693	0.1253	5
Crash Type	Accident Severity	0.28154	0.6683	0.28154	6
Number of Vehicles	Crash Type	0.967	0.967	0.967	-

studies reporting variance reduction indices of 0.08–0.22 [7], [13], [17], [35], attributed to high-speed sustained operations in controlled-access toll roads that amplify geometric effects on severity outcomes. Accident cause ranked second, with human error increasing fatal probability to 0.318 (+0.006), whereas vehicle-related causes reduced it to 0.305 (-0.007). Number of vehicles, weather conditions, and crash type demonstrated moderate influences, with single-vehicle crashes and clear weather conditions slightly elevating fatal risk. Notably, the narrow reachable range (0.026 width) contrasts with previous studies reporting ranges of 0.21–0.34 [7], [35], [36], indicating that fatal outcomes in toll roads are strongly governed by specific combinations of geometric and human factors rather than stochastic variation characteristic of mixed-traffic environments. The most extreme scenario—rear-end crashes occurring at night on curved-downhill segments due to human error—reached the upper probability bound (0.324), confirming that severe crashes emerge from coinciding high-risk conditions rather than isolated factors, thereby supporting targeted interventions

addressing geometric design improvements, night-time safety enhancements, and rear-end collision prevention strategies simultaneously.

3.4. Model Performance Evaluation

The Bayesian Network model's predictive performance was assessed through classification analysis of the entire dataset (N = 450).

Table 11 presents the confusion matrix, while Table 12 summarizes performance metrics across severity classes.

The model achieved an overall classification accuracy of 86.2%, demonstrating robust predictive capability (Figure 6). Minor crashes were predicted with the highest precision (0.938) and recall (0.919), yielding an F1-score of 0.928, attributable to their prevalence in the dataset (74% of cases). Moderate crashes attained balanced precision (0.700) and recall (0.714), with an F1-score of 0.707. Fatal crashes, despite limited support (n = 20), achieved acceptable recall (0.650) and F1-score (0.578), demonstrating the model's ability to capture rare but critical outcomes. This performance compares favorably with previous Bayesian Network applications in crash severity

Table 11. Predicted vs. actual crash severity classification results

		Predicted Severity			Total
		Minor	Moderate	Fatal	
Actual Severity	Minor	305	25	2	332
	Moderate	18	70	10	98
	Fatal	2	5	13	20
Total		325	100	25	450

Note. Diagonal elements indicate correct predictions. Overall accuracy = 86.2%.

Table 12. Classification performance metrics by severity class

Severity Class	Precision	Recall	F1-Score	Support
Minor	0.938	0.919	0.928	332
Moderate	0.7	0.714	0.707	98
Fatal	0.52	0.65	0.578	20
Overall Accuracy	—	—	0.862	450
Macro Average	0.719	0.761	0.738	450
Weighted Average	0.868	0.862	0.864	450

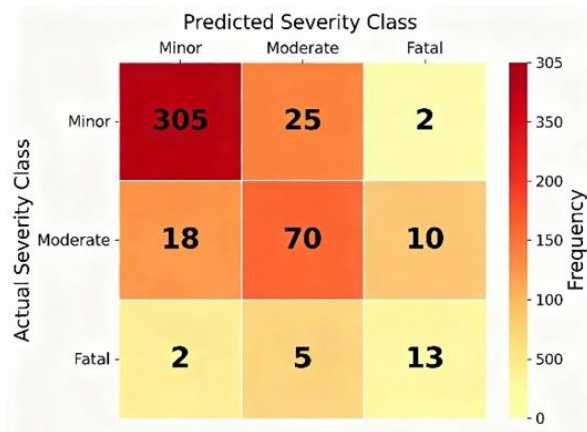


Figure 6. Confusion matrix of Bayesian network crash severity

analysis Zou *et al.* [7] reported 78.3% accuracy for Chinese road accidents with 1,245 accidents sample, while Babaev *et al.* [17] achieved 82% accuracy for Russian road networks. Analysis Assi *et al.* [13] shows 79.5% with 3,500 accidents on USA mix roads. The superior performance observed in this study (86.2%) is attributed to controlled access toll road characteristics reducing traffic heterogeneity, high quality geocoded crash data, and integration of expert knowledge during network structure learning. The weighted average F1-score (0.864) indicates strong overall predictive balance across all severity classes, supporting the model's utility for operational safety management.

Figure 7 shows the Receiver Operating Characteristic (ROC) curve for the Bayesian

Network model used to classify crash severity outcomes. The ROC curve illustrates the trade-off between the true positive rate (sensitivity) and false positive rate (1-specificity) across various classification thresholds. The area under the curve (AUC) is approximately 0.82, indicating good discriminative ability of the model to distinguish between fatal and non-fatal crash events. The diagonal dashed line represents random classification performance (AUC = 0.5). Values closer to the top-left indicate better model accuracy, confirming that the model achieves a strong balance between sensitivity and specificity under imbalanced outcome conditions.

4. Conclusion

This study developed an integrated Z-Score and Bayesian Network framework to analyze crash severity on the Semarang ABC Toll Road using 450 crash records (2022–2025). Z-Score analysis identified five blackspot segments, with KM 430–431 as the most critical ($Z = 4.036$, $n = 91$). The Bayesian Network achieved 86.2% classification accuracy, surpassing previous BN applications (78–82%), with crash type (0.282) and accident cause (0.214) exerting the strongest influence on severity. Straight-downhill segments exhibited 3.3-fold higher fatal probability than straight-level segments (0.083 vs. 0.025), and night-time conditions increased fatal risk by 57%, confirming that severe outcomes emerge from the interaction between human factors and adverse

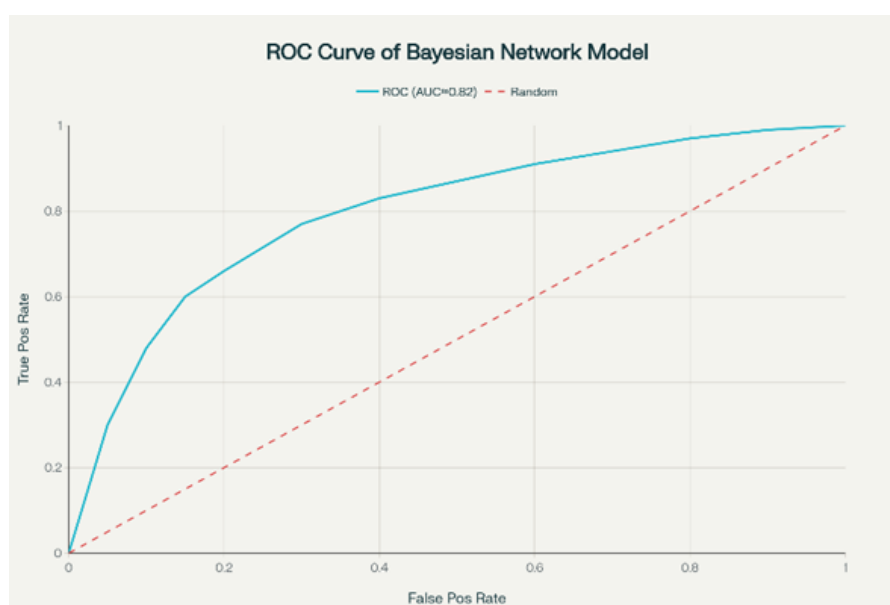


Figure 7. ROC curve of Bayesian network model

geometric-environmental conditions rather than isolated factors a finding consistent with the causal modeling gap identified in the Introduction. These results support prioritizing enhanced lighting on downhill segments during night-time and mandatory safe-following-distance enforcement to reduce rear-end collisions. Future studies are recommended to integrate electronic toll collection (ETC) records to construct driving-duration proxies as fatigue-related variables, and to explore dynamic Bayesian Networks incorporating real-time traffic and weather data for proactive risk assessment. Replication of this framework on other Indonesian toll road corridors particularly Trans Java and Trans Sumatra would assess its generalizability across different topographic contexts. This validated framework supports evidence-based toll road safety management in developing countries, contributing to the UN's 2030 target of reducing traffic fatalities by 50%.

Author's Declaration

Authors' contributions and responsibilities

The authors made substantial contributions to the conception and design of the study. The authors took responsibility for data analysis, interpretation and discussion of results. The authors read and approved the final manuscript.

Funding

No funding information from the authors.

Availability of data and materials

All data are available from the authors.

Competing interests

The authors declare no competing interest.

Additional information

No additional information from the authors.

References

- [1] T. Sagamiko and N. Mbare, "Modelling Road Traffic Accidents Counts in Tanzania: A Poisson Regression Approach," *Tanzania Journal of Science*, vol. 47, no. 1, pp. 308–314, 2021, doi: 10.4314/tjs.v47i1.26.
- [2] Z. Hu, J. Zhou, and E. Zhang, "Improving Traffic Safety through Traffic Accident Risk Assessment," *Sustainability*, vol. 15, no. 4, p. 3748, Feb. 2023, doi: 10.3390/su15043748.
- [3] World Health Organization, "Global status report on road safety 2018," Geneva, 2018.
- [4] U. N. G. Assembly, "Resolution A/RES/74/299: Improving Global Road Safety," New York, 2020.
- [5] World Health Organization, "Road Traffic Injuries: Global Facts and 2030 Reduction Target," Geneva, Switzerland, 2025.
- [6] A. Iranitalab and A. Khattak, "Comparison of four statistical and machine learning methods for crash severity prediction," *Accident Analysis and Prevention*, vol. 108, no. August, pp. 27–36, 2017, doi: 10.1016/j.aap.2017.08.008.
- [7] X. Zou and W. L. Yue, "A Bayesian Network Approach to Causation Analysis of Road Accidents Using Netica," *Journal of Advanced Transportation*, vol. 2017, pp. 1–18, 2017, doi: 10.1155/2017/2525481.
- [8] C. Carrodano, "Data-driven risk analysis of nonlinear factor interactions in road safety using Bayesian networks," *Scientific Reports*, vol. 14, no. 1, p. 18948, Aug. 2024, doi: 10.1038/s41598-024-69740-6.
- [9] Y.-J. Joo, S.-Y. Kho, D.-K. Kim, and H.-C. Park, "A data-driven Bayesian network for probabilistic crash risk assessment of individual driver with traffic violation and crash records," *Accident Analysis & Prevention*, vol. 176, p. 106790, Oct. 2022, doi: 10.1016/j.aap.2022.106790.
- [10] D. Lord and F. Mannering, "The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 5, pp. 291–305, Jun. 2010, doi: 10.1016/j.tra.2010.02.001.
- [11] R. Ullah Khan, J. Yin, and F. Shair Mustafa, "Factors Affecting Crash Frequencies: A Negative Binomial Regression Based Analysis of Indus Highway, Pakistan," *MATEC Web of Conferences*, vol. 296, p. 01005, Oct. 2019, doi: 10.1051/mateconf/201929601005.
- [12] S. A. Arhin and A. Gatiba, "Predicting Injury Severity of Angle Crashes Involving Two Vehicles at Unsignalized Intersections Using Artificial Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3871–3880, Apr. 2019, doi: 10.1016/j.aap.2019.04.001.

- 10.48084/etasr.2551.
- [13] K. Assi, S. M. Rahman, U. Mansoor, and N. Ratrou, "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, doi: 10.3390/ijerph17155497.
 - [14] Y. Oktopianto, S. Shofiah, F. A. Rokhman, K. P. Wijayanthi, and E. Krisdayanti, "Analisis Daerah Rawan Kecelakaan (Black Site) Dan Titik Rawan Kecelakaan (Black Spot) Provinsi Lampung," *Borneo Engineering: Jurnal Teknik Sipil*, vol. 5, no. 1, pp. 40–51, Apr. 2021, doi: 10.35334/be.v5i1.1777.
 - [15] B. Kristanto, S. Widodo, and Sumiyattinah, "Analisis Daerah Rawan Kecelakaan (Drk) Dan Titik Rawan Kecelakaan Di Kabupaten Melawi Provinsi Kalimantan Barat," *JeLAST: Jurnal Teknik Kelautan, PWK, Sipil, dan Tambang*, vol. 11, no. 3, 2024, doi: 10.26418/jelast.v11i3.84793.
 - [16] Rifky Aldila Primasworo, M. Sa'dillah, and Maria Goreti Ganjom, "Analysis of Accident Rate in Kolonel Sugiono Street, Malang City," *Cantilever: Jurnal Penelitian dan Kajian Bidang Teknik Sipil*, vol. 13, no. 1, pp. 73–80, Aug. 2024, doi: 10.35139/cantilever.v13i1.268.
 - [17] I. Babaev and S. Kanarachos, "Bayesian Networks in a Road Accident Severity Estimation," *SSRN Electronic Journal*, 2023, [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4329676
 - [18] A. Kusumawati, K. N. S. Ayuningtyas, and I. Hadiyanti, "Modelling Single-Vehicle Crash Occurrences on Indonesian Toll Roads," *Journal of Road Safety*, vol. 36, no. 4, Nov. 2025, doi: 10.33492/JRS-D-25-4-2120803.
 - [19] F. S. Didin *et al.*, "Analysis of Influencing Factor for Rear-end Crash on Toll Road: An Examination Using Self-reported Questionnaires," *The Open Transportation Journal*, vol. 18, no. 1, Nov. 2024, doi: 10.2174/0126671212362811241114115504.
 - [20] Y. Liang *et al.*, "Nonlinear effects of traffic statuses and road geometries on highway traffic accident severity: A machine learning approach," *PLOS ONE*, vol. 19, no. 11, p. e0314133, Nov. 2024, doi: 10.1371/journal.pone.0314133.
 - [21] W. Santosa *et al.*, *Perspektif Kebijakan dalam Pembangunan Jalan Tol di Indonesia*. Jakarta: PT Penjaminan Infrastruktur Indonesia (Persero), 2023.
 - [22] N. N. A. Efendi, "Faktor Human Error Masih Dominasi Penyebab Kecelakaan di Jalan Tol," *Grid Oto*. [Online]. Available: <https://www.gridoto.com/read/221901971/faktor-human-error-masih-dominasi-penyebab-kecelakaan-di-jalan-tol>
 - [23] B. Istiyanto, Y. E. R. Unzilatirrizqi D, M. D. R. Dewantoro, and R. Ahmad, "Geographic Information System (GIS-T) database design for transportation safety system (case study: spatial database for road transportation system in Indonesia)," *IOP Conference Series: Earth and Environmental Science*, vol. 1117, no. 1, p. 012023, Dec. 2022, doi: 10.1088/1755-1315/1117/1/012023.
 - [24] B. Istiyanto and D. Widitasari, "Road safety analysis on Sron dol – Jatingaleh highway, using Crash Modification Factors (CMFs) Case study: Semarang City, Central Java, Indonesia," *IOP Conference Series: Earth and Environmental Science*, vol. 202, p. 012005, Nov. 2018, doi: 10.1088/1755-1315/202/1/012005.
 - [25] J. Siswanto *et al.*, "Deep learning-based LSTM model for number of road accidents prediction," in *AIP Conference Proceedings*, 2025, p. 050004. doi: 10.1063/5.0258529.
 - [26] L. BayesFusion, "GeNIe Academic: Decision Modeling Environment for Bayesian Networks and Influence Diagrams," 2024, *BayesFusion, LLC, Pittsburgh, Pennsylvania, USA*.
 - [27] K. G. Olesen and A. L. Madsen, "Maximal prime subgraph decomposition of Bayesian networks," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 32, no. 1, pp. 21–31, 2002, doi: 10.1109/3477.979956.
 - [28] M. Wu, D. Shan, Z. Wang, X. Sun, J. Liu, and M. Sun, "A Bayesian Network Model for Real-time Crash Prediction Based on Selected Variables by Random Forest," in *2019 5th International Conference on Transportation Information and Safety (ICTIS)*, IEEE, Jul. 2019,

- pp. 670–677. doi: 10.1109/ICTIS.2019.8883694.
- [29] M. Sun, R. Zhou, C. Jiao, and X. Sun, “Severity Analysis of Hazardous Material Road Transportation Crashes with a Bayesian Network Using Highway Safety Information System Data,” *International Journal of Environmental Research and Public Health*, vol. 19, no. 7, p. 4002, Mar. 2022, doi: 10.3390/ijerph19074002.
- [30] P. Lumba, I. Muthohar, and S. Priyanto, “Human Factors on Motorcyclists’ Accidents Severity; Analysis Using Bayesian Network,” *International Journal of Engineering and Technology*, vol. 9, no. 1, pp. 233–242, Feb. 2017, doi: 10.21817/ijet/2017/v9i1/170901425.
- [31] W. Liao and Q. Ji, “Learning Bayesian network parameters under incomplete data with domain knowledge,” *Pattern Recognition*, vol. 42, no. 11, pp. 3046–3056, Nov. 2009, doi: 10.1016/j.patcog.2009.04.006.
- [32] Y. Peng, L. Cheng, Y. Jiang, and S. Zhu, “Examining Bayesian network modeling in identification of dangerous driving behavior,” *PLOS ONE*, vol. 16, no. 8, p. e0252484, Aug. 2021, doi: 10.1371/journal.pone.0252484.
- [33] S. Demissie, M. P. LaValley, N. J. Horton, R. J. Glynn, and L. A. Cupples, “Bias due to missing exposure data using complete-case analysis in the proportional hazards regression model,” *Statistics in Medicine*, vol. 22, no. 4, pp. 545–557, Feb. 2003, doi: 10.1002/sim.1340.
- [34] S. Al Sulaie, “Sensitivity Analysis of Factors Affecting Consequences Due to Traffic Crashes: A Bayesian Network Modelling,” *Journal of Road Safety*, vol. 36, no. 1, pp. 11–20, Feb. 2025, doi: 10.33492/JRS-D-25-1-2442769.
- [35] J. De Oña, G. López, R. Mujalli, and F. J. Calvo, “Analysis of traffic accidents on rural highways using Latent Class Clustering and Bayesian Networks,” *Accident Analysis and Prevention*, vol. 51, pp. 1–10, 2013, doi: 10.1016/j.aap.2012.10.016.
- [36] R. S. Pilla, “Alternative EM methods for nonparametric finite mixture models,” *Biometrika*, vol. 88, no. 2, pp. 535–550, Jun. 2001, doi: 10.1093/biomet/88.2.535.