Machine Learning for Human Intelligence
2023, Vol. 1, Issue 2

Research Article

Comprehensive Analysis of Traffic Safety and Collision Prediction: A Data-Driven Approach to Transportation Risk Assessment

Asghar Abbas[1] and Usman Humayun[2,⋆]

[1]Department of Chemical Engineering Muhammad Nawaz Sharif University of Engineering and Technology Multan.; Email: engrasgharabbas786@gmail.com
[2]Department of Computer Engineering Faculty of Engineering Bahauddin Zakariya University Multan Pakistan.; Email: usmanhumayun@bzu.edu.pk
⋆Corresponding author: Usman Humayun (usmanhumayun@bzu.edu.pk)

Article History

Academic Editor:
Dr. Muhammad Sajid
Submitted: January 21, 2023
Revised: March 1, 2023
Accepted: September 1, 2023

Keywords:
collision prediction modeling; traffic safety analytics; risk assessment; transportation data analysis; predictive modeling; safety management systems

Abstract

This research presents a comprehensive examination of vehicular transportation safety through the lens of predictive analytics and data-driven methodologies. The study investigates collision patterns, traffic flow dynamics, and risk assessment frameworks within modern transportation systems. Through systematic literature review and bibliometric analysis, we identify two primary research streams: (a) descriptive and predictive modeling approaches that utilize historical data to understand collision patterns and risk factors, and (b) prescriptive optimization methods focused on route selection and hazard mitigation strategies. Our analysis encompasses 856 relevant publications, examining various data collection methodologies, analytical techniques, and modeling approaches used in transportation safety research. The findings reveal significant relationships between driver behavior, environmental conditions, infrastructure characteristics, and collision risk. Key contributing factors include driver fatigue, distracted driving behaviors, roadway geometry, weather conditions, and traffic flow patterns. This study provides a framework for integrating diverse data sources and analytical methods to enhance transportation safety prediction and intervention strategies.

1 Introduction

Transportation safety represents one of the most critical challenges facing modern society, with significant implications for public health, economic stability, and social welfare. Global health statistics indicate that traffic-related events cause roughly 1.4 million deaths each year, making road traffic injuries the eighth greatest cause of mortality globally [1]. The severity of this situation is underscored by forecasts suggesting a possible rise in vehicle-related mortality by 2030 if prevailing patterns persist [1]. In 2017, traffic accidents in the United States resulted in 37,133 fatalities [3], and the economic impact of these events beyond the immediate loss of life. The World Health Organization estimates that highway traffic accidents incur costs equivalent to around 3% of the gross domestic product of most nations [4]. The total economic impact of motor vehicle accidents in the United States surpasses $830 billion each year, including medical expenses, property damage, lost productivity, and additional societal effects [6].

The intricacy of transportation safety issues requires advanced analytical methods capable of efficiently handling large volumes of diverse data sources. Contemporary transportation systems provide incessant flows of data from diverse sensors, monitoring systems, and reporting mechanisms. This data ecosystem comprises extensive crash databases, real-time traffic monitoring systems, meteorological information networks, and infrastructure management systems. The issue is to provide strong analytical frameworks that can integrate various data sources to facilitate evidence-based decision-making in transportation safety management. Modern studies in transportation safety have developed to include two unique but complimentary methodological approaches. The initial stream emphasizes descriptive and predictive modeling methodologies that leverage historical data to discern trends, correlations, and risk factors linked to traffic events. These methodologies generally utilize statistical techniques, machine learning algorithms, and data mining strategies to get meaningful insights from extensive transportation datasets. The second research stream focuses on prescriptive optimization methods intended to enhance operational decision-making, especially in route selection, fleet management, and the transportation of hazardous goods.

This thorough study examines the evident disparity between these two research domains by offering a systematic examination of existing approaches, pinpointing essential research trajectories, and suggesting cohesive frameworks for further exploration. This study conducts a bibliometric analysis of 856 pertinent publications, encompassing peer-reviewed papers, conference proceedings, and technical reports, to provide a systematic evaluation of current transportation safety research. This analysis employs the bibliometric R package to classify and examine the chosen articles through keyword co-occurrence networks and citation patterns. This method facilitates the recognition of conceptual connections within the literature and aids in cultivating a thorough comprehension of prevailing research trends and prospective prospects, as seen in Figure 1.

Figure 1: Bibliography Literature Review

1.1 Research Methodology and Scope

This systematic review is based on a thorough bibliometric study utilizing accepted methodological frameworks. The study includes 856 pertinent documents, comprising published articles, conference papers, and technical reports, chosen according to defined keywords and methodological standards specified in our search approach. The analysis utilizes the bibliometric R package [11] to classify articles and investigate relationships within the transportation safety literature.

The bibliometric approach facilitates the identification of key research themes through keyword co-occurrence analysis and citation network examination. This methodology reveals two primary clusters within the literature: descriptive/predictive modeling approaches and prescriptive optimization methods. Figure 1 illustrates the keyword co-occurrence network, displaying significant linkages corresponding to more than four co-occurrences, with clusters depicted to highlight interconnections among research themes. The network analysis demonstrates clear delineation between prescriptive modeling approaches (represented in the left cluster) and explanatory/predictive modeling methods (shown in the right cluster). This division reflects the fundamental methodological approaches employed in contemporary transportation safety research and provides the structural framework for organizing our comprehensive review.

1.2 Data Collection Frameworks and Methodological Approaches

Transportation safety research relies heavily upon diverse data collection methodologies, each offering unique advantages and limitations for analytical applications. The primary data sources can be categorized into two fundamental approaches: retrospective analysis using police accident reports and prospective data collection through Naturalistic Driving Studies (NDS).

1.2.1 Police Report-Based Data Collection

Traditional transportation safety research has predominantly relied upon police accident reports as the primary data source for retrospective analysis. These reports typically contain comprehensive information regarding incident characteristics, including vehicle counts, injury severity, temporal factors, location details, intersection classifications, roadway types, pedestrian involvement, weather conditions, and road surface conditions [13] [14].

However, this data collection approach presents several inherent limitations that researchers must consider. First, police reports often contain incomplete or estimated information, particularly regarding complex vehicle dynamics and detailed environmental conditions. Second, the accuracy of driver behavior documentation may be compromised due to recall bias or incomplete witness accounts. Third, the denominator problem presents challenges when attempting to establish exposure-based risk calculations, as comprehensive traffic volume data for all roadway segments may not be readily available.

1.2.2 Naturalistic Driving Studies

The emergence of Naturalistic Driving Studies represents a significant advancement in transportation safety data collection methodologies. NDS approaches address many limitations associated with traditional police report analysis by employing continuous data collection systems installed in participating vehicles. These systems capture real-world driving behavior in natural environments, providing researchers with unprecedented access to detailed behavioral and environmental data.

The primary benefits of NDS methodologies encompass: real-time monitoring of driving behavior; extensive gathering of vehicle dynamics, location information, and driver actions; prolonged study durations that encompass diverse driving conditions and scenarios; and objective measurement capabilities that mitigate reporting biases linked to self-reported data or police estimates. The NDS methodology allows researchers to analyze both crash incidents and non-crash situations, thereby aiding in the creation of exposure-based risk models and assisting in the identification of surrogate safety measures that may act as precursors to collision risk.

1.3 Classification of Safety-Critical Events

Research in transportation safety has established advanced classification systems for the identification and categorization of safety-critical incidents. The standardized classification system utilized by more than 45 jurisdictions in the United States offers a uniform framework for data collection and analysis.

This approach differentiates between genuine collisions and Safety-Critical Events (SCE), which operate as proxy indicators for evaluating collision risk. Safety-Critical Events include diverse situations that signify heightened risk levels without necessarily leading to actual collisions. The most well investigated category of SCE pertains to heavy braking events, often characterized as deceleration occurrences over 3.0 m/s^2 [19] [20]. These occurrences offer significant insights into driver behavior patterns and environmental factors that increase accident probability, happening frequently enough to enable comprehensive statistical analysis.

1.3.1 Fatigue and Rest Pattern Analysis

Driver fatigue represents one of the most significant behavioral risk factors affecting transportation safety. Research conducted with commercial truck drivers has revealed strong correlations between fatigue patterns and collision risk [20]. Interview-based studies involving highway truck drivers have identified three primary outcome categories: collision involvement, near-miss events, and self-reported fatigue recognition. as shown in Figure 2. The analysis reveals that fatigue-related risks are influ-

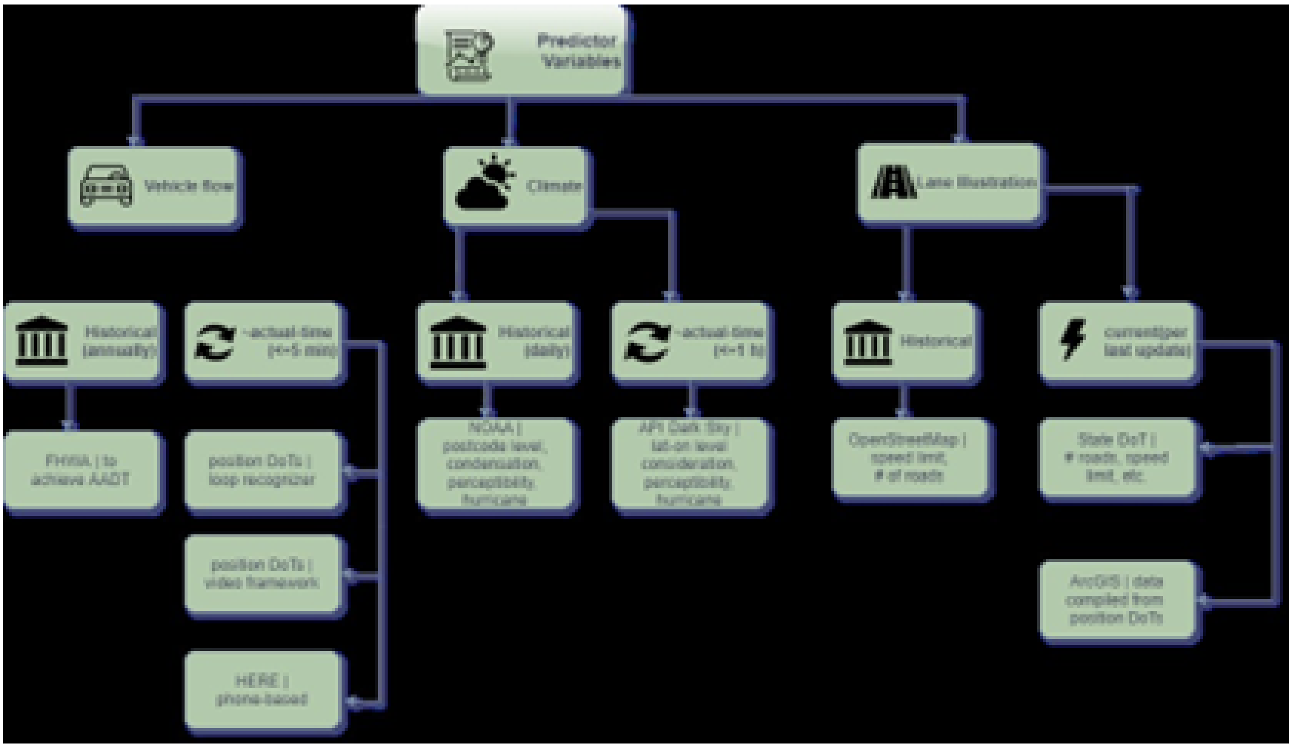

Figure 2: Taxonomy for Predictor Variables

enced by multiple factors, including inefficient trip scheduling, extended duty periods, and inadequate rest opportunities. Logistic regression analysis of these factors demonstrates significant associations between work week fatigue patterns and both collision risk and near-miss event frequency. Particularly notable findings include the relationship between extended work periods exceeding six hours and increased fatigue-related risk indicators. Cross-sectional studies conducted in European contexts have provided additional insights into fatigue-related risk factors among commercial drivers. Italian research examining truck driver safety has identified significant odds ratios for catastrophe exposure, including 2.32 for increased daytime fatigue, 1.45 for sleep deficits, and 1.73 for problematic napping patterns, all excluding the reference value of 1.0.

Advanced fatigue prediction algorithms, such as the Fatigue Avoidance Scheduling Tool (FAST), incorporate regulatory constraints and physiological models to optimize trip scheduling and minimize fatigue-related risks. These systems utilize the Three Process Model of Alertness (TPMA) to predict fatigue levels and support scheduling decisions that comply with regulatory requirements while optimizing safety outcomes.

1.3.2 Distracted Driving Behaviors

The prevalence of distracted driving has emerged as a critical safety concern, with mobile device usage representing a particularly significant risk factor. Global statistics indicate an 11% increase in mobile phone usage while driving over the past decade, with minimal safety differences observed between hands-free and handheld device usage. Research suggests that mobile phone usage increases collision risk by a factor of four, with similar risk elevations observed for both passenger vehicles and commercial trucks. Comprehensive studies analyzing 19,888 safety-critical events among commercial drivers have revealed that 71% of all crashes and 46% of near-crash events involved drivers engaged in non-driving tasks at the time of occurrence. Analysis of 60% of critical events demonstrated some form of driver distraction, highlighting the pervasive nature of this risk factor.

Age-related analysis of distracted driving behaviors reveals significant differences in risk profiles between mature and younger drivers. Studies comparing 109 mature drivers with 42 younger drivers (aged 16.3 to 17.0 years) demonstrate that activities such as texting or phone calls while driving increase injury risk with odds ratios frequently exceeding 3.0.

1.3.3 Weather and Road Conditions

Environmental conditions are crucial in collision risk assessment, as weather-related elements exhibit substantial relationships with crash frequency and severity. Bayesian logistic regression analysis has shown that diminished visibility and inclement weather have multiplicative impacts on collision risk, especially when coupled with difficult roadway layout.

Research on Colorado Interstate 70, employing 30 Remote Traffic Microwave Sensor (RTMS) detectors and six meteorological stations, has yielded comprehensive insights into weather-related risk variables. The study indicates that the likelihood of collisions markedly escalates on steep road segments during inclement weather, with the interplay between geometry and weather generating compounded risk situations.

A thorough investigation of weather influences reveals differing implications across all collision severity classifications. Fatal and severe injury accidents (KA category), property damage crashes (BC category), and minor property damage only (PDO) incidents demonstrate distinct correlations with weather variables, indicating that varying environmental circumstances may uniquely affect crash severity patterns. Roadway Geometry and Infrastructure Infrastructure attributes markedly affect collision risk patterns, with elements such as roadway curvature, lane configuration, shoulder width, and intersection design being pivotal in safety outcomes. Geographic Information Systems (GIS) analysis employing APIs like OpenStreetMap facilitates a thorough evaluation of infrastructure-related risk concerns. The amalgamation of real-time traffic data with infrastructure attributes offers a superior comprehension of dynamic risk elements. The examination of Annual Average Daily Traffic (AADT) trends, alongside roadway geometry data from Federal Highway Administration (FHWA) sources, facilitates the creation of thorough risk assessment models that consider both fixed infrastructure attributes and fluctuating traffic conditions.

2 Analytical Methods and Data Processing Techniques

2.1 Exploratory Data Analysis Frameworks

Effective transportation safety analysis is predicated on robust Exploratory Data Analysis (EDA) approaches that uncover significant patterns from complicated, multi-dimensional datasets. Transportation data has distinct challenges owing to its temporal, geographical, and behavioral intricacies, requiring particular visualization and analytical methodologies.

Transportation safety EDA approaches can be classified into four principal data types: temporal data analysis, geographical configuration analysis, high-dimensional dataset analysis, and multivariate relationship evaluation. Each category necessitates distinct analytical methods and visualization strategies to convey findings and facilitate decision-making processes properly.

2.1.1 Temporal Data Analysis

The temporal analysis of transportation data employs time-series visualization methods to discern trends in traffic flow, collision rates, and behavioral metrics. Line graphs representing transportation variables over time provide fundamental insights into cyclical patterns, trend identification, and anomaly detection. Advanced approaches include ThemeRiver stacked charts for multi-variable temporal analysis and calendar-based visualization methods for identifying periodic patterns.

Circular plotting techniques prove particularly effective for analyzing cyclical data patterns, such as daily and weekly traffic volume variations. These approaches enable researchers to identify distinct travel patterns associated with holidays, special events, and seasonal variations. Statistical decomposition methods, including Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) analysis, support the identification of temporal patterns and facilitate forecasting model development.

2.1.2 Spatial Data Visualization

Spatial analysis of transportation safety data employs geographic visualization techniques to identify location-based risk patterns and infrastructure-related factors. Point-based mapping approaches represent individual incidents or measurements at specific geographic coordinates, enabling the identification of high-risk locations and spatial clustering patterns.

Choropleth mapping techniques aggregate data across geographic regions, facilitating the comparison of safety metrics across different administrative boundaries or infrastructure zones. Line-based visualization methods represent traffic flow patterns and travel routes, supporting the analysis of corridor-level safety performance and identifying optimal routing strategies.

Space-Time-Cube (STC) visualization represents three-dimensional approaches where spatial coordinates form the x and y axes while temporal information comprises the z-axis. This methodology enables simultaneous analysis of spatial and temporal patterns, supporting the identification of emerging risk patterns and seasonal variations in safety performance.

2.2 Dimensionality Reduction Techniques

Transportation safety datasets frequently contain numerous variables representing driver behavior, environmental conditions, vehicle characteristics, and infrastructure attributes. Dimensionality reduction techniques serve three primary objectives: feature selection for improved model performance, attribute extraction for pattern identification, and clustering for data organization and risk stratification.

2.2.1 Feature Selection Methods

Feature selection methods discern the most pertinent variables for predictive modeling, simultaneously decreasing computational complexity and mitigating overfitting. Feature selection procedures are characterized by three basic approaches: filter methods, wrapper methods, and embedding methods.

Filter approaches assess variables autonomously from the prediction model, employing statistical metrics like Pearson correlation or mutual information to prioritize variable significance. These methods provide computational efficiency and simplicity; nevertheless, they may overlook variable interactions or model-specific correlations. Wrapper methods assess variable subsets alongside particular prediction algorithms, employing techniques like evolutionary algorithms and optimization procedures to determine optimal feature combinations. Despite being computationally demanding, wrapper approaches frequently enhance prediction performance by considering variable interactions and model-specific needs.

Embedded approaches incorporate feature selection directly into the model training process, with Random Forest algorithms being the most often utilized methodology in transportation safety literature. These approaches yield changeable significance metrics that inform feature selection while ensuring computational efficiency.

2.2.2 Principal Component Analysis

Principal Component Analysis (PCA) is the primary technique for dimensionality reduction in transportation safety research. PCA converts correlated data into uncorrelated principal components by orthogonal linear transformations, facilitating complexity reduction with minimal information loss. Advanced PCA variations tackle certain data attributes prevalent in transportation datasets, such as Kernel PCA for non-linear associations, Sparse PCA for high-dimensional data, and Robust PCA for datasets with outliers or measurement inaccuracies. The choice of suitable PCA methods is contingent upon the features of the dataset and the analytical goals.

2.2.3 Clustering Techniques

Unsupervised clustering methods group similar data points to maximize within-cluster similarity while minimizing between-cluster similarity. Transportation safety applications utilize clustering for vehicle classification, incident pattern recognition, and environmental condition categorization.

Clustering methodologies encompass partitioning-based approaches (such as k-means), hierarchical methods, density-based techniques, grid-based algorithms, and model-based procedures. The selection of appropriate clustering methods depends upon dataset characteristics, analytical objectives, and computational requirements.

3 Predictive Modeling Approaches

3.1 Statistical Modeling Frameworks

Transportation safety prediction relies heavily upon statistical modeling approaches that can effectively handle the complex, multi-factorial nature of collision risk assessment. The predominant methodological approach utilizes logistic regression analysis for binary outcome prediction (crash vs. non-crash scenarios), with various extensions and modifications to address specific analytical challenges. Case-control study designs represent the most common framework for collision analysis, typically employing sampling ratios ranging from 4:1 to 10:1 (non-crash to crash events) to address the relative rarity of collision events while maintaining statistical power. This approach enables efficient analysis of rare events while providing robust parameter estimation for risk factor identification.

3.1.1 Logistic Regression Analysis

Bayesian logistic regression approaches offer advantages for transportation safety analysis by incorporating prior knowledge and providing uncertainty quantification for parameter estimates. Studies utilizing Bayesian frameworks have demonstrated improved predictive performance and enhanced interpretation capabilities compared to traditional maximum likelihood estimation methods.

Variable selection procedures, including Random Forest-based importance measures and stepwise selection algorithms, support the identification of significant risk factors while managing model complexity. Research conducted using manifestion-control designs with ratios of 4:1 (962 non-crash to 243 crash events) has identified key risk factors including downstream congestion, upstream speed variations, and peak hour traffic conditions. Advanced modeling approaches incorporate spatial and temporal dependencies through hierarchical model structures. Multi-level Bayesian logistic regression models address cross-sectional interaction effects, such as highway merging scenarios, while accounting for random effects associated with different locations or time periods.

3.1.2 Alternative Statistical Approaches

Poisson regression models address collision frequency analysis at the roadway segment level, providing insights into factors influencing crash rates rather than individual crash probability. Integrated modeling approaches combine Poisson and logistic regression frameworks to address both frequency and severity dimensions of collision risk.

Neural network applications utilize multi-layered architectures for pattern recognition in high-dimensional transportation datasets. These approaches demonstrate particular effectiveness for capturing non-linear relationships between risk factors and collision outcomes, though interpretation challenges may limit their application in policy-oriented research.Classification and Regression Trees (CART) offer interpretable modeling frameworks that elucidate hierarchical correlations among risk factors. CART methodologies proficiently discover interaction effects and threshold values for continuous risk factors, facilitating the formulation of pragmatic decision rules for safety management applications.

3.2 Machine Learning Approaches

Advanced machine learning methodologies provide superior capabilities for pattern detection and predictive modeling in transportation safety applications. These methodologies exhibit distinct benefits when addressing extensive datasets characterized by intricate variable interrelations and non-linear trends.

3.2.1 Ensemble Methods

Random Forest algorithms integrate numerous decision trees to deliver robust predictive capabilities while preserving interpretability via changing significance metrics. Gradient boosting methods enhance model performance repeatedly by concentrating on challenging predictions, frequently attaining more predictive accuracy than single-model techniques.

Support Vector Machine (SVM) applications employ kernel approaches to determine optimal decision boundaries in high-dimensional spaces, demonstrating notable efficacy in classification tasks with intricate feature interactions. Support Vector Machine methodologies exhibit robust efficacy in contexts characterized by constrained training datasets or elevated dimensionality of feature spaces.

3.2.2 Deep Learning Applications

Convolutional Neural Networks (CNN) are effective for evaluating spatial patterns in traffic flow and infrastructure data, but Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) architectures are proficient at capturing temporal connections in sequential transportation data.

Deep learning methodologies necessitate meticulous attention to training data prerequisites, computing resource allocation, and interpretability limitations. Although these methods frequently demonstrate enhanced predictive accuracy, their "black box" characteristics may restrict their use in regulatory or policy environments where model transparency is crucial.

4 Integration and Future Directions

4.1 Real-Time Safety Management Systems

The advancement of real-time safety management systems necessitates the creation of analytical frameworks that can handle continuous data streams and deliver actionable insights for operational decision-making [21]. These systems must reconcile analytical complexity with computational performance to facilitate time-sensitive applications. Applications of predictive analytics for real-time risk assessment employ machine learning algorithms trained on historical data to detect heightened risk circumstances and initiate suitable intervention techniques. The efficacy of these systems relies on the accessibility of high-quality real-time data streams and resilient communication networks for transmitting safety information. Future research directions should emphasize the development of integrated analytical frameworks capable of combining multiple data sources and methodological approaches to support comprehensive safety management systems [22]. The advancement of real-time analytics capabilities offers significant potential for proactive safety intervention strategies, though successful implementation requires careful attention to data quality, system integration, and stakeholder coordination requirements.

The ultimate objective of this research stream involves developing practical tools and decision support systems that translate advanced analytical capabilities into improved transportation safety outcomes. This goal requires continued collaboration between researchers, practitioners, and policymakers to ensure that methodological advances contribute meaningfully to enhanced public safety and reduced societal costs associated with transportation incidents [23, 24, 25].

The findings presented in this review provide a foundation for future research initiatives and offer guidance for practitioners seeking to implement evidence-based safety management approaches. Continued advancement in this field requires sustained investment in data collection infrastructure, analytical method development, and practical application frameworks that bridge the gap between research innovations and operational implementation requirements [26, 27].

4.2 Policy and Implementation Considerations

The translation of research findings into effective safety interventions requires consideration of policy frameworks, implementation challenges, and stakeholder coordination requirements. Successful implementation depends upon developing evidence-based recommendations that account for regulatory constraints, resource limitations, and operational feasibility considerations. Future research should emphasize the development of practical tools and decision support systems that facilitate the application of advanced analytical methods in operational transportation safety management contexts [28, 29, 30].

5 Conclusion

This comprehensive analysis of transportation safety research reveals the evolution of analytical methodologies from traditional statistical approaches toward sophisticated data-driven frameworks capable of integrating diverse information sources and supporting evidence-based decision making. The systematic review of 856 publications demonstrates clear delineation between descriptive/predictive modeling approaches and prescriptive optimization methods, each offering unique contributions to transportation safety enhancement. The research identifies critical gaps between these methodological streams and proposes integrated frameworks for future investigation. Key findings emphasize the importance of driver behavior factors, particularly fatigue and distraction, while highlighting the significant influence of environmental conditions and infrastructure characteristics on collision risk patterns. The analytical methods review demonstrates the evolution from basic statistical techniques toward advanced machine learning approaches, with ensemble methods and deep learning applications showing particular promise for complex pattern recognition tasks. However, the implementation of these advanced techniques must balance predictive performance with interpretability requirements and computational feasibility constraints.

References

[1] World Health Organization. (2019). The top 10 causes of death. Retrieved from https://www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death

[2] Insurance Institute for Highway Safety. (2019). Yearly snapshot. Retrieved from https://www.iihs.org/topics/fatality-statistics/detail/yearly-snapshot

[3] National Highway Traffic Safety Administration. (2017). US DOT announces 2017 roadway fatalities down. Retrieved from https://www.nhtsa.gov/press-releases/us-dot-anounces-2017-roadway-fatalities-down

[4] World Health Organization. (2021). Road traffic injuries. Retrieved from https://www.who.int/en/news-room/fact-sheets/detail/road-traffic-injuries

[5] World Bank. (2020). National accounts data, and OECD national accounts data files. Retrieved from https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=US

[6] National Highway Traffic Safety Administration. (2015). Economic and societal impact of motor vehicle crashes. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812013

[7] Ditta, A., Figueroa, O., Galindo, G., & Yie-Pinedo, R. (2019). A review on research in transportation of hazardous materials. *Socio-Economic Planning Sciences*, 68, 100665.

[8] Abkowitz, M., & Cheng, P. D. M. (1988). Developing a risk/cost framework for routing truck movements of hazardous materials. *Accident Analysis & Prevention*, 20(1), 39-51.

[9] Theofilatos, A., & Yannis, G. (2014). A review of the effect of traffic and weather characteristics on road safety. *Accident Analysis & Prevention*, 72, 244-256.

[10] Roshandel, S., Zheng, Z., & Washington, S. (2015). Impact of real-time traffic characteristics on freeway crash occurrence: Systematic review and meta-analysis. *Accident Analysis & Prevention*, 79, 198-211.

[11] Aria, M., & Cuccurullo, C. (2017). bibliometrix: An r-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975.

[12] Garfield, E., & Sher, I. H. (1993). Keywords plus [tm]-algorithmic derivative indexing. *Journal-American Society for Information Science*, 44, 298-298.

[13] Thiese, M. S., Hanowski, R. J., Kales, S. N., Porter, R. J., Moffitt, G., Hu, N., & Hegmann, K. T. (2017). Multiple conditions increase preventable crash risks among truck drivers in a cohort study. *Journal of Occupational and Environmental Medicine*, 59(2), 205.

[14] Newnam, S., Xia, T., Koppel, S., & Collie, A. (2019). Work-related injury and illness among older truck drivers in australia: A population based, retrospective cohort study. *Safety Science*, 112, 189-195.

[15] Dingus, T. A., Hanowski, R. J., & Klauer, S. G. (2011). Estimating crash risk. *Ergonomics in Design*, 19(4), 8-12.

[16] Guo, F. (2019). Statistical methods for naturalistic driving studies. *Annual Review of Statistics and Its Application*, 6, 309-328.

[17] Federal Highway Administration. (2020). Real time system management (FHWA). Retrieved from https://ops.fhwa.dot.gov/511/index.htm

[18] Guo, F., Klauer, S. G., Hankey, J. M., & Dingus, T. A. (2010). Near crashes as crash surrogate for naturalistic driving studies. *Transportation Research Record*, 2147(1), 66-74.

[19] Jansen, R. J., & Wesseling, S. (2018). Harsh braking by truck drivers: A comparison of thresholds and driving contexts using naturalistic driving data. In *Proceedings of the 6th Humanist Conference*, The Hague, The Netherlands (pp. 13-14).

[20] Mollicone, D., Kan, K., Mott, C., Bartels, R., Bruneau, S., van Wollen, M., Sparrow, A. R., & Van Dongen, H. P. (2019). Predicting performance and safety based on driver fatigue. *Accident Analysis & Prevention*, 126, 142-145.

[21] Lovelace, R., Nowosad, J., & Muenchow, J. (2019). *Geocomputation with R*. CRC Press.

[22] Kraak, M.-J. (1999). Visualising spatial distributions. In *Geographical Information Systems: Principles, Techniques, Applications and Management* (pp. 157–173). John Wiley and Sons.

[23] Erdogan, S. (2009). Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *Journal of Safety Research*, *40*(5), 341–351.

[24] Wongsuphasawat, K., Pack, M. L., Filippova, D., VanDaniker, M., & Olea, A. (2009). Visual analytics for transportation incident data sets. *Transportation Research Record*, *2138*(1), 135–145.

[25] Liu, S., Pu, J., Luo, Q., Qu, H., Ni, L. M., & Krishnan, R. (2013). VAIT: A visual analytics system for metropolitan transportation. *IEEE Transactions on Intelligent Transportation Systems*, *14*(4), 1586–1596.

[26] Zeng, W., Fu, C.-W., Arisona, S. M., & Qu, H. (2013). Visualizing interchange patterns in massive movement data. In *Computer Graphics Forum* (Vol. 32, No. 3pt3, pp. 271–280). Wiley Online Library.

[27] Kraak, M.-J. (2003, July). The space-time cube revisited from a geovisualization perspective. In *Proc. 21st International Cartographic Conference* (pp. 1988–1996). Citeseer.

[28] Kapler, T., & Wright, W. (2005). Geotime information visualization. *Information Visualization*, *4*(2), 136–146.

[29] *Traffic accidents including cyclists in (ACT)*. (2015). Retrieved from https://brettromero.com/traffic-accidents-cyclists/

[30] *Visualized as a living circulatory system of traffic 24 hours watch*. (2016). Retrieved from http://metrocosm.com/map-us-traffic/