Investigating Contributing Factors to Fatal Non-Intersection Crashes Involving Elderly Pedestrians Using Association Rule Mining

Ernest Nsong Asiedu *
Department of Civil and Environmental Engineering
Auburn University, Auburn, AL 36849-5337
Tel: 334-559-2053; Email: ena0011@auburn.edu
ORCID ID: 0009-0002-1029-9459

Mohammad Reza Abbaszadeh Lima
Independent Researcher
Ewing, NJ, 08628
Phone: 334-610-9911; Email: mr.abbaszadelima@gmail.com
OCRID ID: 0009-0008-7728-8811

Md Roknuzzaman
Department of Civil and Environmental Engineering
Auburn University, Auburn, AL 36849-5337
Tel: 334-849-9714; Email: mzr0177@auburn.edu
OCRID ID: 0000-0003-4765-774X

*Corresponding Author

ABSTRACT

Fatal pedestrian crashes involving elderly individuals (aged 65+) are a growing public safety concern in the United States. Notably, approximately 80% of fatalities occurring away from intersections underscores the critical need to examine non-intersection crash dynamics. This study investigates contributing factors to such non-intersection fatalities involving elderly pedestrians using Association Rule Mining (ARM) applied to the Fatality Analysis Reporting System (FARS) data from 2019 to 2023. A two-stage analytical framework was employed: first, random forest-based feature selection identified the most influential variables; second, the Apriori algorithm uncovered frequent co-occurring crash patterns. The analysis yielded 26 high-lift rules (lift > 2.6) associated with elderly pedestrians being pronounced dead at the scene. A striking and novel finding was the consistent presence of solo drivers (DRIVER_ALONE = TRUE) in all high-risk scenarios, suggesting that unaccompanied drivers may face heightened risks of missing or failing to respond to elderly pedestrians, particularly in challenging environments. Other common antecedents included roadside pedestrian locations, poor lighting conditions (e.g., "Dark – Not Lighted"), and overnight hours (12 A.M.–6 A.M.). Alcohol involvement, lack of traffic controls, and disabling vehicle damage further contributed to fatal outcomes. These findings highlight the compounded dangers of isolated, low-visibility, and uncontrolled settings, especially when drivers are alone. The study underscores the need for targeted safety interventions such as enhanced roadside lighting, improved infrastructure in non-intersection areas, and time-specific driver awareness efforts. By identifying these high-risk patterns, this research provides actionable insights to inform policy and infrastructure strategies aimed at protecting vulnerable elderly pedestrians.

Keywords: Elderly Pedestrian Safety, Association Rule Mining, FARS Dataset, Non-Intersection Crash

INTRODUCTION

Pedestrian safety has emerged as a critical public health and transportation challenge globally. Traffic crashes account for over 1.2 million deaths and tens of millions of injuries each year, with pedestrians constituting one of the most vulnerable groups. Worldwide, pedestrians represent more than 22% of all road traffic deaths, approximately 270,000 fatalities annually, many of which result in permanent disability or long-term consequences for survivors (*1*).In the United States, the situation is similarly concerning. Pedestrian fatalities rose by 48% from 2009 to 2016, making it the steepest increase among OECD nations during that period(*2*).Alarmingly, recent estimates suggest nearly 6,000 pedestrians die each year in motor vehicle crashes in the U.S., maintaining historically high levels last seen more than 25 years ago(*3*). **Figure 1** shows historical trend of annual average pedestrian fatalities in the U.S.

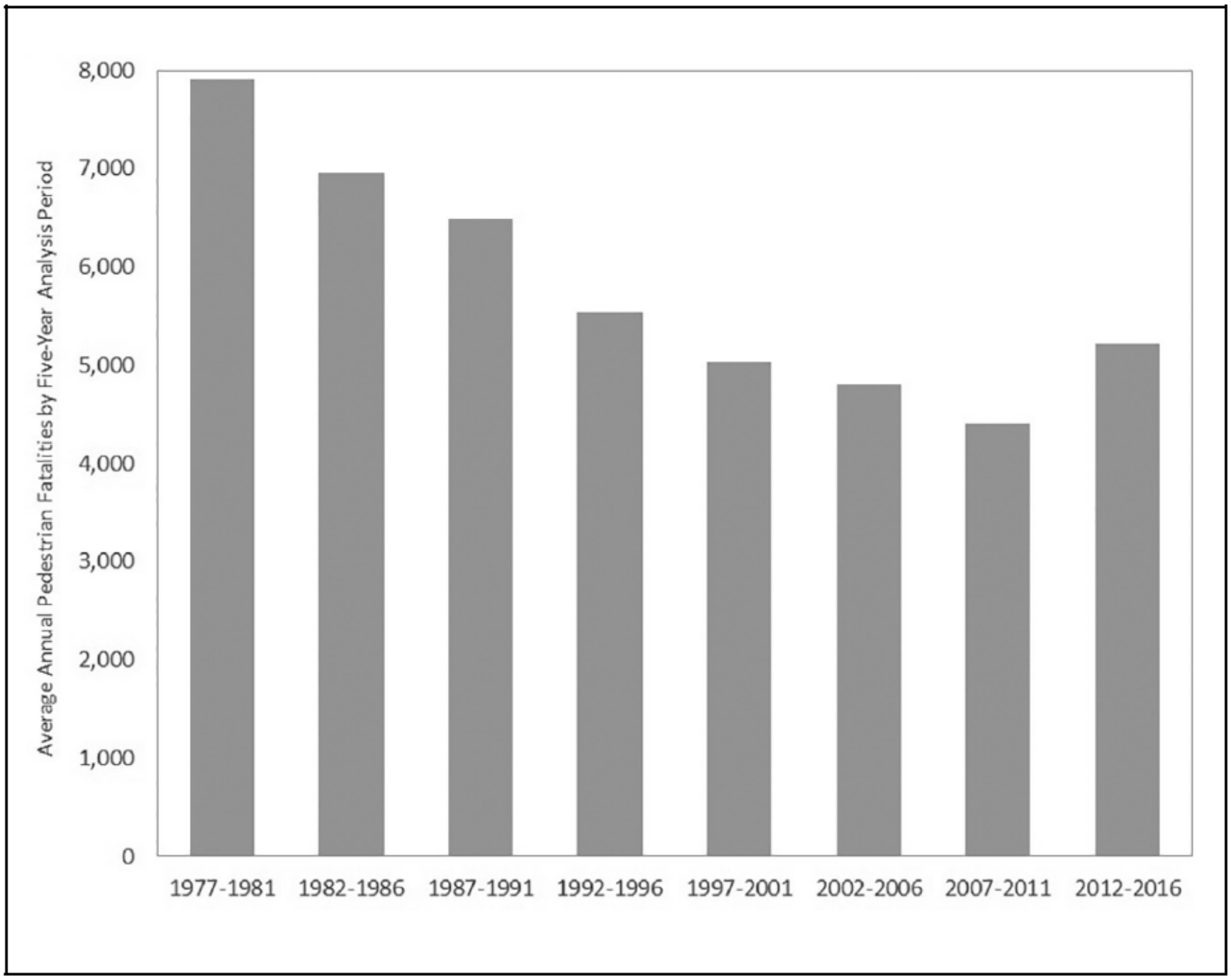

Figure 1 Average annual pedestrian fatalities by 5-year analysis intervals*(2)*

The 2024 NHTSA report estimated 39,345 traffic fatalities in the U.S., a 3.8% drop from 2023, with a similar decline in overall fatality rate. However, about 20% of states still exceed the pre-COVID national average of 1.13 fatalities per 100 million vehicle miles traveled(*4*). This underscores that despite national improvements, pockets of elevated risk persist, often driven by latent or localized factors. Several factors, including age distribution, physical vulnerability, and behavioral tendencies, contribute to the disproportionate representation of pedestrians in traffic fatalities (*5*, *6*), with older pedestrians facing a higher risk of fatality due to slower mobility and reduced reflexes. **Figure 2** illustrates the age distribution of pedestrian deaths per hundred thousand people in the USA, highlighting that historically, older pedestrians have been at greater risk of fatal accidents.

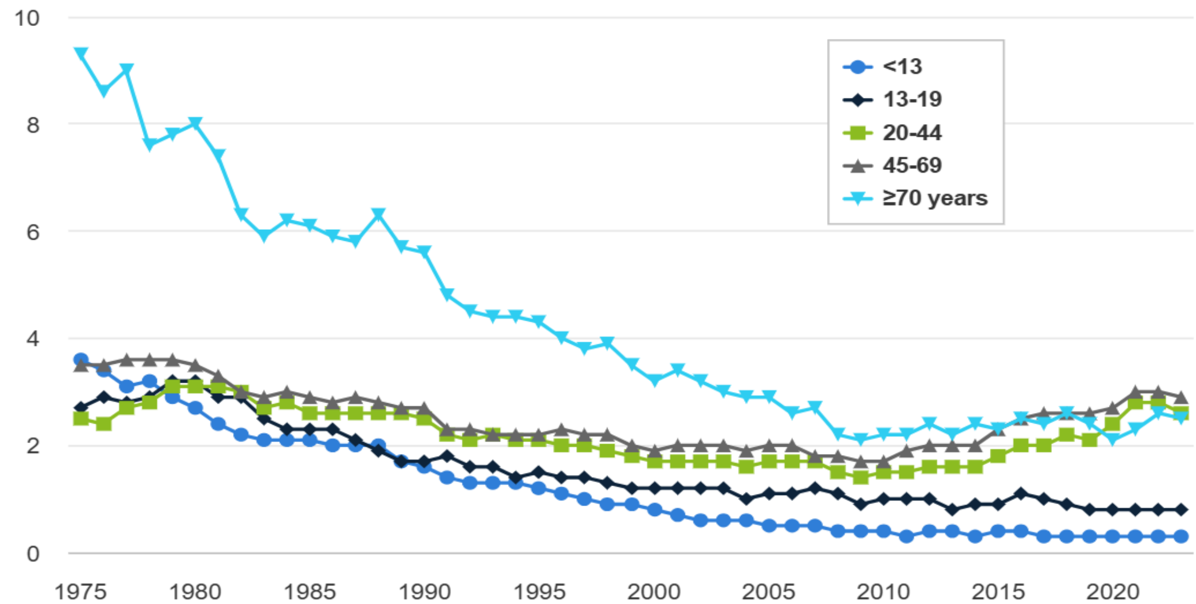

Figure 2 Pedestrian deaths per 100,000 people by age, 1975-2023 *(7)*

Furthermore, most pedestrian fatalities in the U.S. occur away from intersections, approximately 80% according to long-term data (*2*). These non-intersection crashes are particularly hazardous, often taking place under poor lighting conditions, at higher speeds, and in locations lacking pedestrian infrastructure (*8*). Understanding the complex interactions of contributing factors in such crashes is crucial for effective countermeasure development.

Researchers have employed various analytical techniques to identify risk factors associated with pedestrian fatalities, including logistic regression, Bayesian models, decision trees, and data mining methods(*9*, *10*). Several studies have utilized regional or state-level data, such as police crash records, hospital databases, and forensic reports. While insightful, these studies often suffer from limited geographic coverage or lack of consistency in data definitions (*11*).The Fatality Analysis Reporting System (FARS), maintained by the National Highway Traffic Safety Administration, overcomes these limitations by providing a nationally standardized, comprehensive dataset of all fatal crashes on U.S. public roadways. Its extensive coverage, rich variable structure, and consistency make FARS the most authoritative dataset for examining nationwide pedestrian fatality trends (*2*, *5*).

Numerous factors contribute to pedestrian fatality risk, including pedestrian age and gender, vehicle type, driver behavior, lighting conditions, road type, location (urban vs. rural), and time of day (*1*, *6*). Vehicle speed and size (especially large vehicles such as SUVs and trucks) have also been strongly associated with increased fatality risk (*2*). In recent years, distraction (e.g., smartphone use), alcohol involvement, and inadequate lighting have emerged as critical contributors, particularly in nighttime crashes (*3*, *8*).

Association Rule Mining (ARM), a powerful data mining technique, has proven effective in discovering hidden patterns and relationships within large crash datasets. Unlike traditional regression methods, ARM can uncover multiple interacting factors that frequently co-occur with fatal outcomes, making it especially useful for complex phenomena like pedestrian crashes(*9*, *10*). Prior research has demonstrated the utility of ARM in identifying high-risk combinations of driver, pedestrian, and environmental attributes in urban pedestrian crashes (*10*). However, few studies have applied ARM specifically to non-intersection crashes, which have unique dynamics and risk profiles.

This study aims to fill that gap by applying association rule mining to recent FARS data (2019–2023) to explore contributing factors to fatal non-intersection crashes involving elderly pedestrians. It updates prior findings, explores shifting patterns, and highlights emerging risks, offering timely insights to inform targeted safety interventions and support evidence-based transportation policy decisions.

LITERATURE REVIEW

Elderly Pedestrians and Fatality Risk in the U.S.

Fatal pedestrian crashes involving elderly individuals (aged 65+) are a growing public safety concern in the United States. Pedestrians aged 70 and above account for 19% of total fatalities and have the highest per capita fatality rate, averaging 22 deaths per million population (*12*). As life expectancy increases globally, the elderly population continues to grow, including in the U.S., posing challenges across various sectors, particularly transportation(*13*). Older pedestrians are especially vulnerable; for example, a 70-year-old hit at 25 mph faces the same fatality risk as a 30-year-old struck at 35 mph(*14*). Studies have identified roadway design, environmental conditions, vehicle type, and individual characteristics as key contributors to pedestrian fatalities. Geometric features, speed limits, intersection design, and terrain influence crash likelihood. FARS data from 2009–2016 shows a 69% increase in pedestrian deaths on urban arterials, while nighttime fatalities rose 56%, emphasizing poor visibility. This background highlights the high vulnerability of elderly pedestrians and underscores the need to understand the contributing factors in fatal crashes.

Vehicle and Personal Factors Influencing Pedestrian Fatalities

Vehicle type also plays a significant role. Larger vehicles, including SUVs, light trucks, and vans, pose a significantly higher risk of causing severe or fatal injuries compared to smaller passenger cars. his is due to design features like greater body mass, elevated ride height, and more rigid front structures (bumpers, grilles, hoods). Studies show pedestrians struck by light trucks or vans are two to three times more likely to die than those hit by passenger cars. (*16*) Additionally, modern vehicles with higher power-to-weight ratios are often associated with higher impact speeds, worsening injury outcomes. Driver demographics also influence crash outcomes. Male drivers consistently exhibit higher fatality rates than females averaging three times higher per capita. Although both genders have seen annual fatality increases (4.6% for males, 3.7% for females), the gap remains notable. Age matters as well: teen drivers have higher crash rates, but older adult drivers have higher fatal crash rates per mile. For example, impatient drivers attempting to navigate around elderly pedestrians potentially causing fall and collisions (*14*). These vehicle and driver-related characteristics can be the important variables to be incorporated in a comprehensive rule-mining approach to uncover fatal crash patterns specific to elderly pedestrians.

Age-Related Vulnerability in Pedestrian Crashes

Older adult pedestrians are particularly vulnerable in traffic environments due to age-related declines in physical, sensory, perceptual, and cognitive abilities. An observational study using video recordings at multiple urban sites found that older adults were more likely to cross two-way undivided roads in the presence of closer oncoming vehicles and generally adopted less safe crossing strategies compared to younger adults. This supports the view that age-related cognitive and perceptual limitations significantly contribute to crash involvement and fatality risk(*17*). Another study analyzing 4,290 traffic casualties reported significantly higher mortality rates among individuals aged 65 and older pedestrians compared to non-pedestrians. While the definition of "elderly" varies across studies, many use 65 years as the threshold (*18*), others focus on those over 70 years (*19*) and some specifically examine pedestrians aged 75 and above(*20*). Physical challenges such as reduced mobility and visibility further increase risks, especially when walking near reversing vehicles or during low-visibility or winter conditions. Elderly pedestrians often require more time to cross streets, making short signal phases or multilane intersections especially hazardous(*21*). Furthermore, arterial roads characterized by multiple lanes, high speeds, and large traffic volumes pose significant dangers, particularly at non-intersection points where many elderly fatalities occur

(*22*). These behavioral and physiological insights validate the focus on elderly pedestrians and support the decision to study non-intersection locations where age-related impairments compound crash risk.

Non-Intersection Locations as High-Risk Zones

Between 2009 and 2016, studies found that over two-thirds of pedestrian fatalities in urban areas occurred at non-intersection locations. This trend was consistent across all roadway types, where fatalities were more frequent at non-intersections than intersection(*15*). Non-intersections often include zones of high pedestrian activity such as transit stops and bus stations which increase the potential for conflicts between pedestrians and vehicles (*23*).These risks are often exacerbated by geometric roadway features like curves or elevation changes that limit sight distance, as well as the lack of well-designed sidewalks or clearly marked crosswalks (*24*). Several studies have investigated these contributing factors, emphasizing the need for comprehensive design and policy interventions particularly for vulnerable populations such as the elderly. While early research used diverse datasets and methods to identify contributing factors, this review highlights the importance of focusing on non-intersection environments with the aim to identify contextual crash patterns at such locations.

Limitations of Traditional Models and Need for Data-Driven Approaches

A growing body of research has explored contributing factors to fatal pedestrian crashes at non-intersections, revealing complex interactions among driver behavior, environmental conditions, and vehicle type. A mixed logit model using HSIS data (2007–2016) in North Carolina found that severe weather increased the likelihood of fatal pedestrian crashes at non-intersections by 37.3%. The presence of traffic control devices, day of the week, and passenger presence particularly for drivers over age 40years were also significant, with passengers offering a stronger protective effect for male drivers(*25*). Notably, male drivers were more frequently involved in non-intersection fatal crashes, potentially due to aggressive turning behavior. Another study, integrating a mixed logit model with GIS, found that driver inattention had a stronger impact under hot weather conditions (*26*). A study using Poisson regression on FARS and General Estimates System data reported a 50% increase in pedestrian deaths at non-intersections and an 82% increase in SUV involvement in fatal single-vehicle pedestrian crashes over time (*15*). An empirical Bayes data mining approach using FARS (2014–2016) highlighted associations such as backing vehicle incidents involving elderly female pedestrians and poor visibility during dark conditions without proper street lighting(*27*). Additionally, a 2021 FARS study using multivariate logistic regression found that dusk-time and non-intersection crashes were linked to shorter survival times for older pedestrians, while urban crashes resulted in longer survival times than rural ones(*28*). A regression analysis in Louisiana (2010–2016) further revealed pedestrians under the influence of alcohol or drugs were at greater fatality risk(*29*). Finally, Association Rule Mining on Louisiana crash data (2010–2019) linked pedestrian fatalities to high-speed roads (50 mph) under dark, unlit conditions, especially when walking against traffic. Interestingly, elderly pedestrians were more likely to be involved in crashes during daylight with cloudy weather, emphasizing the need to consider multiple contextual factors in developing pedestrian safety strategies (*30*).

Unlike traditional regression models, which often struggle to capture complex or nonlinear interactions, Association Rule Mining (ARM) offers a flexible and powerful alternative for uncovering hidden patterns in large datasets(*31*). ARM also supports interactive visualizations, which enhance interpretation and the practical use of findings(*32*). ARM has been used in a variety of contexts related to pedestrian safety. For example, it has been applied to analyze Fatal pedestrian crashes at intersections (*33*), Pedestrian crashes in urban areas within a specific U.S. state (*10*), Intersection-related crashes in urban areas of developing countries(*34*) (*35*), Fatal collision data across eight U.S. states, incorporating driver, environmental, and roadway characteristics (*36*), and pedestrian midblock crash severity, where the method revealed nuanced associations that might be overlooked by conventional approaches (*31*). These examples highlight ARM's versatility in identifying context-specific risk factors that inform targeted safety interventions. Despite its demonstrated potential, very few studies have applied ARM to the Fatality Analysis Reporting System (FARS) with a focus on elderly pedestrians at non-intersections. This gap presents an opportunity for future research to apply ARM in this context and develop evidence-based strategies to address this vulnerable population's unique risks.

METHODOLOGY

This study employed a two-stage methodological framework to identify contributing factors to fatal non-intersection crashes involving elderly pedestrians (aged 65 and older) in the United States using Fatality Analysis Reporting System (FARS) data from 2016 to 2023. The first stage involved selecting relevant features through a random forest-based feature importance ranking, while the second stage applied association rule mining (ARM) to discover frequent co-occurring crash attributes associated with fatalities. **Figure 3** illustrates the study workflow.

Data Source and Preprocessing

The dataset used in this study was drawn from the national FARS database, which provides detailed information on traffic fatalities, including variables related to crash circumstances, vehicle and driver characteristics, and environmental conditions. The analysis was restricted to records involving pedestrians aged 65 or older fatally injured in non-intersection crashes. Data cleaning and preprocessing steps included merging relevant FARS tables, filtering for eligible cases, removing missing or inconsistent values, and recoding variables to categorical formats suitable for ARM.

Feature Selection Using Random Forest

Given the high dimensionality of FARS data, a feature selection process was applied to reduce noise and improve interpretability. A random forest model was trained using fatal crash outcomes as the target variable, and variable importance was assessed using Mean Decrease in Impurity (MDI), an impurity-based metric that ranks variables based on their contribution to reducing classification error(*37*). To minimize bias associated with MDI, particularly its tendency to favor variables with more categories, only the top-ranked features were retained using a straightforward cutoff rule. This strategy maintains computational efficiency and strong predictive performance without resorting to more intensive recursive or permutation-based methods (*38*).

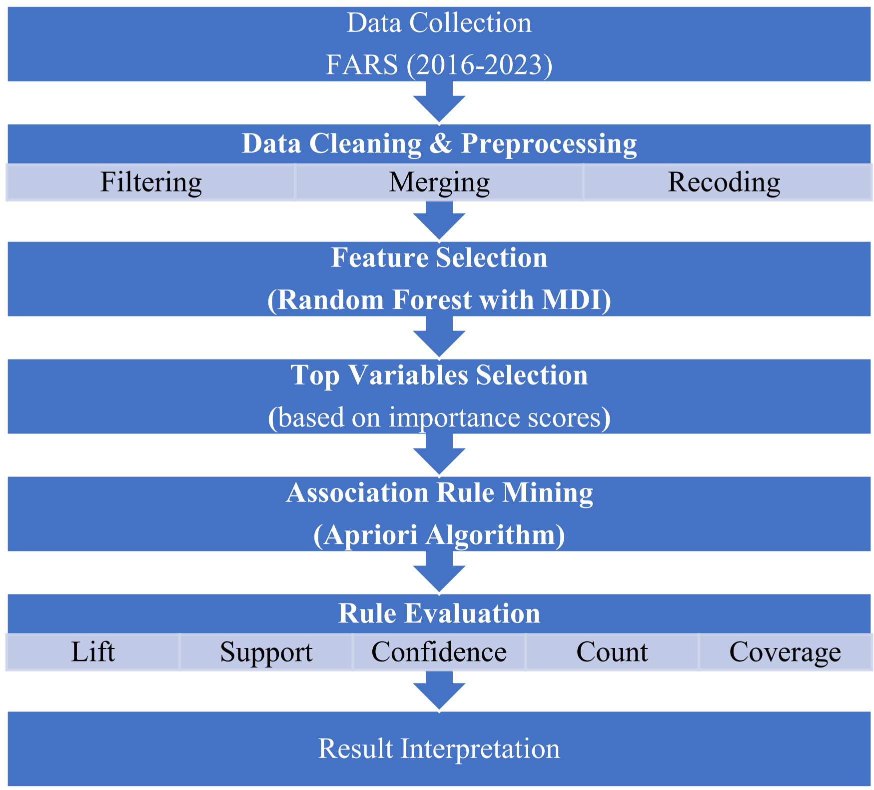

Figure 3 Workflow for the Study Methodology

Association Rule Mining Framework

Association Rule Mining (ARM) was employed to uncover frequent patterns among the selected features that co-occur with elderly pedestrian fatalities. Unlike traditional hypothesis-driven methods, ARM identifies patterns from data without requiring a priori assumptions, making it particularly suitable for exploratory analyses in traffic safety research (*39–41*). This study applies the Apriori algorithm to identify association rules in fatal crashes involving older pedestrians (65 or higher years old) at non-intersection locations, with a specific focus on patterns linked to pedestrian fatalities. In recent years, many investigators have applied ARM as a decision-support tool to extract rules from high-dimensional crash databases that focus on specific categories of variables (*40–43*).

The method used by Hossain et al (*40*) was adopted in this study. Let $R = \{r_1, r_2, r_3, \ldots, r_n\}$ denote the collection of the desired records, where each record contains a subset of attributes drawn from the itemset $X = \{x_1, x_2, x_3, \ldots, x_n\}$. An association rule is written as $U \rightarrow V$ with $U, V \subseteq R$ and $U \cap V = \emptyset$. Here, U is the antecedent (left-hand side or LHS), and V is the consequent (right-hand side or RHS).

Each rule consists of one or more attribute-value pairs that describe frequently co-occurring conditions in fatal pedestrian crashes(*40*). For example, the rule {REL_ROADNAME = On Roadside, TIME_BLOCK = 12 A.M.- 6 A.M., DRIVER_ALONE=TRUE} → {DOANAME = Died at Scene} has U = {REL_ROADNAME = On Roadside, TIME_BLOCK=12 A.M.- 6 A.M., DRIVER_ALONE=TRUE} and V = {DOANAME=Died at Scene}. The length of these rules can vary based on the type of study. These rules capture interdependencies among variables and should not be interpreted as evidence of direct causation (*40*, *44–46*).

Rule assessment draws on fundamental metrics like support (S), confidence (C), lift (L), and count, alongside supplementary indicators like coverage (Cove.), collectively enriching the insight into each rule's robustness and importance within traffic-safety analysis (*45*, *47*). Support measures how often the pattern ($U \rightarrow V$) occurs in the entire dataset, while confidence is the proportion of those occurrences relative to the total number of times U appears (*40*, *41*).The third metric, lift, reflects how frequently the items occur together within the same independent crash event. On the other hand, **count is** the number of records in which the rule's antecedent and consequent appear together. Unlike support, it is not normalized by the total number of transactions; it simply states how many times the full rule occurs. Also, **coverage** (antecedent support) is the frequency with which the rule's antecedent shows up in the dataset, indicating the breadth of situations to which the rule could potentially apply (*48*).The corresponding formulas are as follows:

$$\text{Support }(U) = U' / N$$
$$\text{Support }(V) = V' / N$$
$$\text{Support }(U \rightarrow V) = (U' \cap V') / N$$
$$\text{Confidence }(U \rightarrow V) = \text{Support }(U \rightarrow V) / \text{Support}(U)$$
$$\text{Lift }(U \rightarrow V) = \text{Support }(U \rightarrow V) / [\text{Support}(U) \times \text{Support}(V)]$$

In these formulas, N is the total number of desired records, U′ is the frequency of occurrences containing U, V′ is the frequency containing V, and (U′ ∩ V′) is the frequency containing both U and V.

Lift is pivotal because it shows how much more frequently the antecedent and consequent appear together than would be expected if they were statistically independent (*40*, *49*) A lift greater than 1 indicates a positive correlation between U and V, whereas a value below 1 suggests a negative correlation (*49*, *50*). A lift close to 1 implies that U is independent of the likelihood of V (*40*, *41*, *51*).

Tools and Software

All analyses were conducted in R using the "arules", "dplyr", and "readr" packages for data mining and preprocessing.

RESULTS and DISCUSSION

Dataset Overview and Temporal Trend

The final dataset included 5,314 fatal crashes involving at least one elderly pedestrian (aged 65+) in the U.S. from 2019 to 2023. As shown in **Figure 4**, fatalities increased by 23% over this period, peaking at 1,212 deaths in 2022, with a slight decline in 2023.

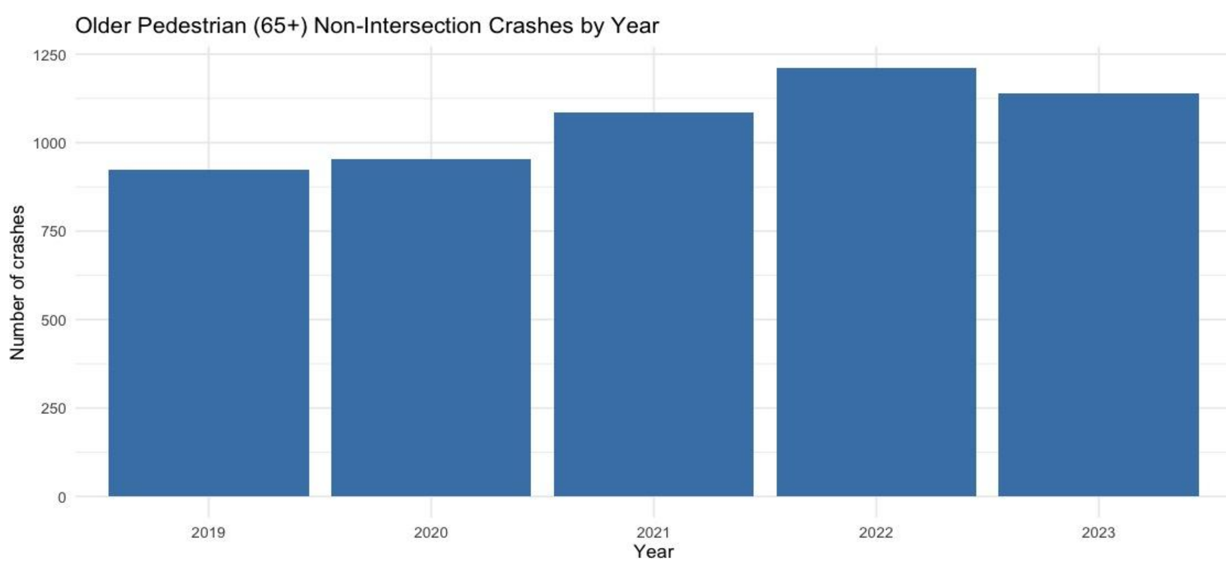

Figure 4 Elderly (65+ years) pedestrian fatality in U.S.A. per year (from 2019 to 2023)

Spatial Distribution of Fatalities

Figure 5 illustrates the geographic distribution of elderly pedestrian fatalities, highlighting stark state-level differences that mirror demographic patterns. Florida (777), California (726), and Texas (510) had the highest fatality counts, aligning with their large elderly populations 10.6%, 8.3%, and 7.1% of the national total, respectively according to the 2023 American Community Survey (*28*). Conversely, Alaska, the Dakotas, Wyoming, and Nebraska each reported fewer than 10 deaths, consistent with their smaller senior populations. While population distribution is a key factor, other contributors include year-round pedestrian activity in warmer states, urban infrastructure, traffic control measures, and access to public transit, all shaping the risk landscape for elderly pedestrians across the U.S.

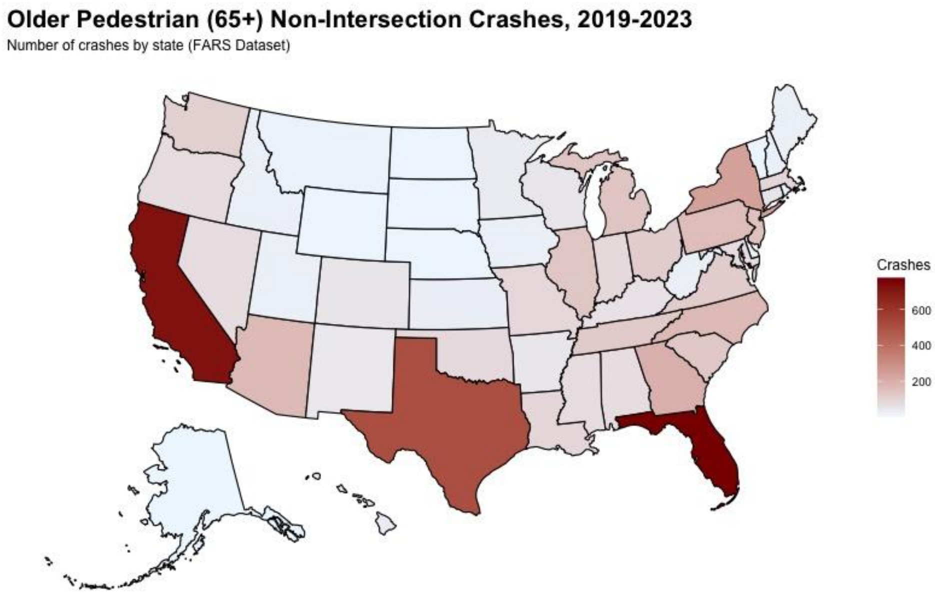

Figure 5 Spatial Distribution of elderly pedestrian fatality

Random Forest Feature Importance Result

To select variables for the analysis, we began with a comprehensive review of previous studies to identify variables that were consistently found to be relevant.(*23*, *27*, *33*, *50*, *52*, *53*) Based on this review, we initially included 30 variables in our research to ensure the ARM model was supplied with meaningful and appropriate inputs. To refine our variable set, we ranked features with the Random Forest algorithm's impurity-based importance (mean decrease in Gini), enabling us to isolate and retain only those variables that most strongly predict the "DOAName" column, which represents the "Died at Scene/En Route" outcome. The results are illustrated in **Figure 6**.

Association Rule Mining (ARM) Setup

To derive rules related to the elderly pedestrians at non-intersection locations, Association Rule Mining (ARM) is applied by setting the "Died at Scene/En Route" column as the consequent (right-hand side) in each rule. To enhance interpretability, the maximum number of items in each rule was limited to five (*39–41*, *43*). Setting appropriate minimum thresholds for support and confidence in ARM is essential to obtain meaningful and insightful results (*40*, *49*). A threshold set too low may yield an overwhelming number of trivial rules, while a threshold set too high could overlook important underlying associations between categories (*40*). As there are no fixed standards for determining the threshold values of support, confidence, and lift, a trial-and-error method, consistent with previous research (*39, 40, 54*) was used. Accordingly, the minimum thresholds were set as follows: confidence at 0.6, support higher than 0.01, and lift at 2. The “DOAName” column, specifically representing “Died at Scene/En Route,” was designated as the consequent in the rules. These thresholds aimed to balance rule relevance and coverage, ensuring that only the most informative and statistically significant patterns were retained.

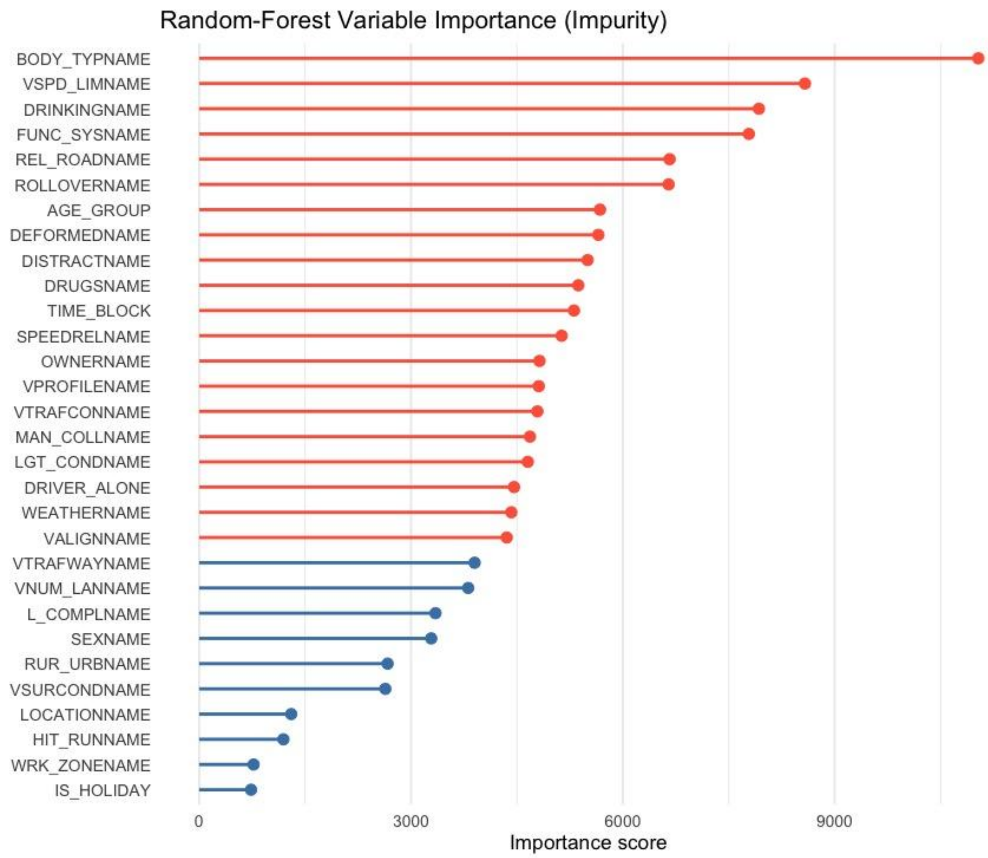

Figure 6 Feature Importance plot based on Random Forest Gini impurity (mean decrease in Gini).

We limited our analysis to the top 20 ranked variables, as the importance of features declines sharply beyond this point, evident in Figure 6, which shows a steep drop in the feature importance plot. This selection ensures that only the most relevant features are included in the ARM analysis to generate meaningful rules. The following set of rules, as shown in **TABLE 1**, was generated based on the criteria outlined earlier in this section and sorted by their lift values. The right-hand side of each rule is the “DOANAME” variable, which specifies the location where the pedestrian was pronounced dead, if applicable. This variable includes four categories: Not Applicable, Died at Scene, Died En Route, and Unknown.

TABLE 1 Rules for elderly pedestrians at non-intersection locations, sorted by lift value.

#	LHS	RHS	Support	Confidence	Coverage	Lift	Count
1	{REL_ROADNAME=On Roadside,DEFORMEDNAME=Disabling Damage,LGT_CONDNAME=Dark - Not Lighted,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0176	0.7140	0.0246	2.8304	8008
2	{REL_ROADNAME=On Roadside,DEFORMEDNAME=Disabling Damage,TIME_BLOCK=12 A.M.,Äì6 A.M.,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0125	0.7123	0.0176	2.8237	5704
3	{REL_ROADNAME=On Roadside,DISTRACTNAME=Not Reported,LGT_CONDNAME=Dark - Not Lighted,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0108	0.7068	0.0153	2.8017	4919
4	{REL_ROADNAME=On Roadside,ROLLOVERNAME=Rollover, Tripped by Object/Vehicle,MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0102	0.7053	0.0145	2.7958	4661
5	{REL_ROADNAME=On Roadside,ROLLOVERNAME=Rollover, Tripped by Object/Vehicle,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0103	0.7038	0.0146	2.7901	4665
6	{REL_ROADNAME=On Roadside,MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport,LGT_CONDNAME=Dark - Not Lighted,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0197	0.7030	0.0280	2.7868	8963
7	{REL_ROADNAME=On Roadside,LGT_CONDNAME=Dark - Not Lighted,DRIVER_ALONE=TRUE,WEATHERNAME=Clear}	{DOANAME=Died at Scene}	0.0136	0.7024	0.0194	2.7844	6195
8	{REL_ROADNAME=On Roadside,VTRAFCONNAME=No Controls,LGT_CONDNAME=Dark - Not Lighted,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0150	0.7010	0.0214	2.7789	6834
9	{REL_ROADNAME=On Roadside,LGT_CONDNAME=Dark - Not Lighted,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0197	0.7009	0.0282	2.7784	8988
10	{REL_ROADNAME=On Roadside,TIME_BLOCK=12 A.M.,Äì6 A.M.,MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0142	0.6983	0.0203	2.7681	6459
11	{REL_ROADNAME=On Roadside,TIME_BLOCK=12 A.M.,Äì6 A.M.,VTRAFCONNAME=No Controls,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0106	0.6967	0.0152	2.7620	4827
12	{REL_ROADNAME=On Roadside,TIME_BLOCK=12 A.M.,Äì6 A.M.,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0142	0.6966	0.0204	2.7615	6470

#	LHS	RHS	Support	Confidence	Coverage	Lift	Count
13	{REL_ROADNAME=On Roadside,OWNERNAME=Driver (in this crash) was Registered Owner,LGT_CONDNAME=Dark - Not Lighted,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0111	0.6951	0.0160	2.7554	5072
14	{REL_ROADNAME=On Roadside,VPROFILENAME=Level,LGT_CONDNAME=Dark - Not Lighted,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0130	0.6943	0.0188	2.7525	5931
15	{ROLLOVERNAME=Rollover, Tripped by Object/Vehicle,DEFORMEDNAME=Disabling Damage,MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0115	0.6902	0.0166	2.7359	5221
16	{ROLLOVERNAME=Rollover, Tripped by Object/Vehicle,MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0128	0.6886	0.0186	2.7298	5816
17	{DRINKINGNAME=Yes (Alcohol Involved),REL_ROADNAME=On Roadside,DEFORMEDNAME=Disabling Damage,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0134	0.6884	0.0195	2.7289	6106
18	{ROLLOVERNAME=Rollover, Tripped by Object/Vehicle,VTRAFCONNAME=No Controls,MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0105	0.6871	0.0153	2.7239	4790
19	{ROLLOVERNAME=Rollover,MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0117	0.6866	0.0170	2.7220	5307
20	{DRINKINGNAME=Yes (Alcohol Involved),REL_ROADNAME=On Roadside,MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0153	0.6803	0.0225	2.6969	6974
21	{DRINKINGNAME=Yes (Alcohol Involved),REL_ROADNAME=On Roadside,VTRAFCONNAME=No Controls,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0115	0.6802	0.0170	2.6965	5252

#	LHS	RHS	Support	Confidence	Coverage	Lift	Count
22	{DRINKINGNAME=Yes (Alcohol Involved),REL_ROADNAME=On Roadside,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0154	0.6792	0.0226	2.6925	6987
23	{DRINKINGNAME=Yes (Alcohol Involved),REL_ROADNAME=On Roadside,DRIVER_ALONE=TRUE,WEATHERNAME=Clear}	{DOANAME=Died at Scene}	0.0110	0.6715	0.0163	2.6620	4986
24	{REL_ROADNAME=On Roadside,AGE_GROUP=Middle Adult (25-45),DEFORMEDNAME=Disabling Damage,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0197	0.6654	0.0296	2.6379	8978
25	{VSPD_LIMNAME=55 MPH,REL_ROADNAME=On Roadside,DEFORMEDNAME=Disabling Damage,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0162	0.6649	0.0244	2.6357	7388
26	{REL_ROADNAME=On Roadside,AGE_GROUP=Middle Adult (25-45),DISTRACTNAME=Not Reported,DRIVER_ALONE=TRUE}	{DOANAME=Died at Scene}	0.0123	0.6611	0.0187	2.6206	5617

Key ARM Findings

Based on **TABLE 1**, the Association Rule Mining (ARM) analysis produced a coherent set of 26 high-lift rules that characterize conditions most strongly associated with pedestrians being pronounced dead at the scene (DOANAME = "Died at Scene"). The lift values of these rules range from approximately 2.60 to 2.83, indicating that the identified combinations of antecedent factors are 2.6 to 2.83 times more likely to result in a fatal outcome than if the antecedents and fatality were statistically independent. Support values, ranging from 0.0106 to 0.0197, suggest that although these combinations are individually infrequent, they occur with sufficient regularity to warrant serious attention. Confidence levels near 0.70 further highlight the predictive reliability of these rules.

As shown in **TABLE 2,** the most frequently occurring factor in the antecedents was DRIVER_ALONE=TRUE, appearing in all 26 rules. This underscores a strong association between solo driving and fatal pedestrian outcomes. The next most common factors were REL_ROADNAME=On Roadside (22 rules) and LGT_CONDNAME=Dark – Not Lighted (8 rules), suggesting that crashes are more likely when drivers are alone, traveling along the roadside, and under poor lighting conditions.

Other notable contributing factors included non-collision events, disabling vehicle damage, rollovers, alcohol involvement, and early morning driving (12 A.M.–6 A.M.). Collectively, these patterns point to a recurring set of high-risk scenarios that are frequently linked to fatal or severe pedestrian crashes.

TABLE 2 Most Frequent Factors in the Antecedents of Rules Linked to Non-Intersection Elderly Pedestrian Crashes

Rank	**Factor that appears in the rule antecedent**	**# of rules**
1	DRIVER_ALONE=TRUE	26
2	REL_ROADNAME=On Roadside	22
3	LGT_CONDNAME=Dark - Not Lighted	8
4	MAN_COLLNAME=The First Harmful Event was Not a Collision with a Motor Vehicle in Transport	8
5	DEFORMEDNAME=Disabling Damage	6
6	ROLLOVERNAME=Rollover, Tripped by Object/Vehicle	5
7	DRINKINGNAME=Yes (Alcohol Involved)	5
8	TIME_BLOCK=12 A.M.–6 A.M.	4
9	VTRAFCONNAME=No Controls	4
10	DISTRACTNAME=Not Reported	2
11	WEATHERNAME=Clear	2
12	AGE_GROUP=Middle Adult (25-45)	2
13	OWNERNAME=Driver (in this crash) was Registered Owner	1
14	VPROFILENAME=Level	1
15	VSPD_LIMNAME=55 MPH	1
16	ROLLOVERNAME=Rollover	1

Discussion of ARM Patterns

Prevalence of Solo Driving and Roadside Crashes

A consistent pattern observed across nearly all high-lift rules is the co-occurrence of **REL_ROADNAME = "On Roadside"** and **DRIVER_ALONE = TRUE**, highlighting an elevated risk when an unaccompanied driver is involved in a crash with a pedestrian located on the roadside. Solo drivers may face increased crash risk due to the absence of passengers who might help them stay alert, recognize hazards, or summon emergency assistance. This lack of support can heighten both the probability of a crash and the severity of resulting injuries especially for older pedestrians, who are more vulnerable to severe outcomes.

Interestingly, this observation contrasts with prior findings that associate peer passengers with increased risky driving behaviors (*55–57*). In the context of elderly pedestrian fatalities, driver aggression may play a lesser role than pedestrian characteristics (e.g., slower mobility) and environmental factors such as poor visibility. The recurring presence of **REL_ROADNAME = On Roadside** suggests that many elderly pedestrians were struck while walking outside the primary roadway, and that drivers, especially those alone may have failed to notice them in time, possibly due to inattention, drowsy or distraction.

Significance of Crash Location and Environmental Conditions

The variable **REL_ROADNAME** appeared in 22 out of 26 rules, emphasizing the importance of spatial context in pedestrian fatalities. This variable reflects the pedestrian's position relative to the trafficway and includes categories such as On Roadway, On Shoulder, On Median, On Roadside, and Outside Trafficway. Its consistent presence underscores the need to address risks associated with off-road pedestrian activity. The strongest rule identified in **TABLE 1** includes **LGT_CONDNAME = "Dark – Not Lighted"**, **DEFORMEDNAME = "Disabling Damage"**, and **ROLLOVERNAME = "Rollover, Tripped by Object/Vehicle"**, portraying a common scenario of severe, late-night crashes in poorly lit environments. These conditions often result in high-impact collisions that disable vehicles or cause rollovers, contributing to on-scene fatalities. This aligns with the findings of a previous study (*29*) suggesting that the current road system, designed primarily for young, healthy users, should be redesigned to better accommodate older, more vulnerable pedestrians.

Behavioral and Temporal Risk Factors

Several rules incorporate behavioral and temporal variables, particularly **TIME_BLOCK = "12 A.M.–6 A.M."** and **DRINKINGNAME = "Yes (Alcohol Involved)"**, suggesting compounded risks during late-night hours. These periods are associated with reduced pedestrian visibility, increased driver fatigue, elevated speeds, and higher rates of substance use. Older pedestrians, already slower and less visible, are especially at risk during these hours.

Another recurring factor, **VTRAFCONNAME = "No Controls"**, appeared in four rules. This variable reflects a lack of traffic control devices such as signs or signals, conditions typical of non-intersection locations. The absence of controls may elevate the risk of pedestrian fatalities by reducing driver awareness or predictability. Likewise, **MAN_COLLNAME = "First Harmful Event Was Not a Collision with a Motor Vehicle in Transport"**, present in eight rules, indicates that many fatal pedestrian crashes did not begin with a vehicle-to-vehicle collision. This suggests that the pedestrian was likely the initial point of impact, highlighting their direct vulnerability.

Additional Notable Factors

Beyond the primary risk factors, several secondary elements emerged with notable associations:

- **TIME_BLOCK = 12 A.M.–6 A.M.**, appearing in four rules, reinforces the dangers of overnight hours. This timeframe combines the lowest pedestrian visibility with behavioral risks such as fatigue and intoxication, making it particularly perilous for older individuals.
- **AGE_GROUP = Middle Adult (25–45)** appeared in two rules. While this age group is generally considered to be composed of experienced drivers, its presence in fatal crash patterns, particularly in roadside scenarios suggests that experience alone does not eliminate crash risk. Contextual factors such as environment and visibility remain critical.

- **WEATHERNAME = Clear**, found in two rules, indicates that fatal crashes occurred even under favorable weather conditions. This suggests that non-environmental factors such as driver behavior or lack of infrastructure may play more significant roles in crash causation.
- **VPROFILENAME = Level**, which appeared in one rule, describes a flat roadway profile. The implication here is that crashes with fatal outcomes occurred even in the absence of challenging terrain, further emphasizing the impact of human behavior and design limitations.

Summary of Findings

The Association Rule Mining analysis reveals a distinct cluster of high-risk scenarios contributing to elderly pedestrian fatalities at non-intersection locations. The most salient patterns involve isolated drivers, roadside locations, poor lighting conditions, and late-night hours, all of which compound the vulnerabilities of older pedestrians. These findings provide actionable insights that can inform targeted interventions, such as enhanced lighting, roadside infrastructure modifications, and increased driver awareness campaigns in areas with limited controls.

CONCLUSION

This study applied Association Rule Mining (ARM) to nationwide FARS data (2019–2023) to uncover contributing factors in fatal non-intersection crashes involving elderly pedestrians. The analysis revealed a consistent set of high-lift rules pointing to recurring high-risk conditions. Key patterns included solo drivers, roadside crash locations, low-light environments, and late-night hours that significantly elevate fatality risk for elderly pedestrians. The frequent co-occurrence of variables such as **DRIVER_ALONE = TRUE**, **REL_ROADNAME = On Roadside**, and **LGT_CONDNAME = Dark – Not Lighted** emphasizes the compounded dangers of driving alone at night in poorly lit, uncontrolled environments. Behavioral and temporal risk factors, including **alcohol involvement** and **overnight driving (12 A.M.–6 A.M.)**, further contributed to the severity of outcomes. Notably, fatal crashes occurred even in favorable weather and terrain, suggesting that human behavior and design shortcomings are more influential than environmental conditions.

These results highlight the need for targeted safety measures, better lighting, clearer signage, and improved driver awareness to protect older pedestrians. Identifying high-risk patterns can guide evidence-based policies and targeted infrastructure upgrades to reduce fatal incidents at non-intersections

ACKNOWLEDGMENTS

The authors utilized Grammarly for grammar checking and language refinement to improve the clarity and readability of this paper.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: N. Asiedu, M.R.A. Lima, M. Roknuzzaman; data collection: M.R.A. Lima**,** N. Asiedu; analysis and interpretation of results: N. Asiedu, M.R.A. Lima; draft manuscript preparation: N. Asiedu, M.R.A. Lima, M. Roknuzzaman; Writing – review & editing: M. Roknuzzaman; All authors reviewed the results and approved the final version of the manuscript.

DECLARATION OF CONFLICTING INTERESTS

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The authors disclosed no financial support for the research, authorship, and/or publication of this article.

REFERENCES

1. Hashemi Nazari, S., J. Hasani, A. Khorshidi, S. Erfanpoor, and B. Nazparvar. Comparison of Risk Factors for Pedestrian Fatality in Urban and Suburban Traffic Accidents. *Archives of Trauma Research*, Vol. 7, No. 2, 2018, p. 39. https://doi.org/10.4103/atr.atr_7_18.
2. Schneider, R. J. United States Pedestrian Fatality Trends, 1977 to 2016. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2674, No. 9, 2020, pp. 1069–1083. https://doi.org/10.1177/0361198120933636.
3. Retting, R., and S. S. Consulting. Governors Highway Safety Association. 2017.
4. NHTSA Estimates 39,345 Traffic Fatalities in 2024 | NHTSA. https://www.nhtsa.gov/press-releases/nhtsa-2023-traffic-fatalities-2024-estimates. Accessed Jun. 24, 2025.
5. Sanders, R., and R. Schneider. An Exploration of Pedestrian Fatalities by Race in the United States. *Transportation Research Part D: Transport and Environment*, Vol. 107, 2022, p. 103298. https://doi.org/10.1016/j.trd.2022.103298.
6. Martínez-Ruiz, V., M. Valenzuela-Martínez, P. Lardelli-Claret, D. Molina-Soberanes, E. Roldán, and E. Jiménez-Mejías. Factors Related to the Risk of Pedestrian Fatality after a Crash in Spain, 1993–2013. *Journal of Transport & Health*, Vol. 12, 2019, pp. 279–289. https://doi.org/10.1016/j.jth.2019.02.008.
7. Fatality Facts 2023: Pedestrians. *IIHS-HLDI crash testing and highway safety*. https://www.iihs.org/research-areas/fatality-statistics/detail/pedestrians. Accessed Jul. 18, 2025.
8. Hossain, A., X. Sun, R. Thapa, Md. M. Hossain, and S. Das. Exploring Association of Contributing Factors to Pedestrian Fatal and Severe Injury Crashes under Dark-No-Streetlight Condition. *IATSS Research*, Vol. 47, No. 2, 2023, pp. 214–224. https://doi.org/10.1016/j.iatssr.2023.03.002.
9. Besharati, M. M., and A. Tavakoli Kashani. Which Set of Factors Contribute to Increase the Likelihood of Pedestrian Fatality in Road Crashes? *International Journal of Injury Control and Safety Promotion*, Vol. 25, No. 3, 2018, pp. 247–256. https://doi.org/10.1080/17457300.2017.1363781.
10. Das, S., A. Dutta, R. Avelar, K. Dixon, X. Sun, and M. Jalayer. Supervised Association Rules Mining on Pedestrian Crashes in Urban Areas: Identifying Patterns for Appropriate Countermeasures. *International Journal of Urban Sciences*, Vol. 23, No. 1, 2019, pp. 30–48. https://doi.org/10.1080/12265934.2018.1431146.
11. Noland, R. B., J. A. Sinclair, N. J. Klein, and C. Brown. How Good Is Pedestrian Fatality Data? Journal of Transport & Health, Elsevier, 7, 2017, pp. 3–9.
12. *2022 Traffic Safety Facts Older Population*. Publication DOT HS 813 616. 2024.
13. How Does the Aging of the Population Affect Our Fiscal Health? Peterson Foundation.
14. aaafoundation. Impact Speed and a Pedestrian's Risk of Severe Injury or Death - AAA Foundation for Traffic Safety. 2011.
15. Hu, W., and J. B. Cicchino. An Examination of the Increases in Pedestrian Motor-Vehicle Crash Fatalities during 2009–2016. *Journal of Safety Research*, Vol. 67, 2018, pp. 37–44. https://doi.org/10.1016/j.jsr.2018.09.009.
16. Lefler, D. E., and H. C. Gabler. The Fatality and Injury Risk of Light Truck Impacts with Pedestrians in the United States. *Accident Analysis & Prevention*, Vol. 36, No. 2, 2004, pp. 295–304. https://doi.org/10.1016/S0001-4575(03)00007-1.
17. Oxley: Differences in Traffic Judgements between... - Google Scholar. https://scholar.google.com/scholar_lookup?title=Differences%20in%20traffic%20judgements%20between%20young%20and%20old%20adult%20pedestrians&publication_year=1997&author=J.%20Oxley&author=B.%20Fildes&author=E.%20Ihsen&author=J.%20Charlton&author=R.%20Day. Accessed Jun. 24, 2025.
18. Zhai, X., H. Huang, N. N. Sze, Z. Song, and K. K. Hon. Diagnostic Analysis of the Effects of Weather Condition on Pedestrian Crash Severity. *Accident Analysis & Prevention*, Vol. 122, 2019, pp. 318–324. https://doi.org/10.1016/j.aap.2018.10.017.

19. Salon, D., and A. McIntyre. Determinants of Pedestrian and Bicyclist Crash Severity by Party at Fault in San Francisco, CA. *Accident Analysis & Prevention*, Vol. 110, 2018, pp. 149–160. https://doi.org/10.1016/j.aap.2017.11.007.
20. Sasidharan, L., and M. Menéndez. Partial Proportional Odds Model—An Alternate Choice for Analyzing Pedestrian Crash Injury Severities. *Accident Analysis & Prevention*, Vol. 72, 2014, pp. 330–340. https://doi.org/10.1016/j.aap.2014.07.025.
21. Duim, E., M. L. Lebrão, and J. L. F. Antunes. Walking Speed of Older People and Pedestrian Crossing Time. *Journal of Transport & Health*, Vol. 5, 2017, pp. 70–76. https://doi.org/10.1016/j.jth.2017.02.001.
22. Zegeer: Pedestrian Safety Strategic Plan: Recommendations... - Google Scholar. https://scholar.google.com/scholar_lookup?title=Pedestrian%20safety%20strategic%20plan%3A%20recommendations%20for%20research%20and%20product%20development%20&publication_year=2010&author=C.%20Zegeer&author=D.%20Nabors&author=D.%20Gelinne&author=N.%20Lefler&author=M.%20Bushnell. Accessed Jun. 24, 2025.
23. Kim, D. The Transportation Safety of Elderly Pedestrians: Modeling Contributing Factors to Elderly Pedestrian Collisions. Accident Analysis & Prevention, Elsevier, 131, 2019, pp. 268–274.
24. Abou-Senna, H., E. Radwan, and A. Mohamed. Investigating the Correlation between Sidewalks and Pedestrian Safety. *Accident Analysis & Prevention*, Vol. 166, 2022, p. 106548. https://doi.org/10.1016/j.aap.2021.106548.
25. Qiu, B., and W. (David) Fan. Mixed Logit Models for Examining Pedestrian Injury Severities at Intersection and Non-Intersection Locations. *Journal of Transportation Safety & Security*, Vol. 14, No. 8, 2022, pp. 1333–1357. https://doi.org/10.1080/19439962.2021.1923101.
26. Geyer, J. A., and D. R. Ragland. Vehicle Occupancy and Crash Risk. *Transportation Research Record*, Vol. 1908, No. 1, 2005, pp. 187–194. https://doi.org/10.1177/0361198105190800123.
27. Das, S., A. Bibeka, X. Sun, H. "Tracy" Zhou, and M. Jalayer. Elderly Pedestrian Fatal Crash-Related Contributing Factors: Applying Empirical Bayes Geometric Mean Method. *Transportation Research Record*, Vol. 2673, No. 8, 2019, pp. 254–263. https://doi.org/10.1177/0361198119841570.
28. Phosri, Y. *Factors Associated with Pedestrian Traffic Fatalities among Older Adults in Urban and Rural Environments*. The Ohio State University, 2024.
29. Das, S., S. Ashraf, A. Dutta, and L.-N. Tran. Pedestrians under Influence (PUI) Crashes: Patterns from Correspondence Regression Analysis. *Journal of Safety Research*, Vol. 75, 2020, pp. 14–23. https://doi.org/10.1016/j.jsr.2020.07.001.
30. Hossain, A., X. Sun, R. Thapa, and J. Codjoe. Applying Association Rules Mining to Investigate Pedestrian Fatal and Injury Crash Patterns Under Different Lighting Conditions. *Transportation Research Record*, Vol. 2676, No. 6, 2022, pp. 659–672. https://doi.org/10.1177/03611981221076120.
31. Chakraborty, R., S. Das, and M. N. Khan. Uncovering Pedestrian Midblock Crash Severity Patterns Using Association Rules Mining. *Transportmetrica A: Transport Science*, Vol. 0, No. 0, pp. 1–48. https://doi.org/10.1080/23249935.2024.2379948.
32. Du, L., F. Huang, H. Lu, S. Chen, and Q. Guo. An Association Rule Mining-Based Modeling Framework for Characterizing Urban Road Traffic Accidents. *Sustainability*, Vol. 16, No. 23, 2024, p. 10597. https://doi.org/10.3390/su162310597.
33. Das, S., R. Tamakloe, H. Zubaidi, I. Obaid, and A. Alnedawi. Fatal Pedestrian Crashes at Intersections: Trend Mining Using Association Rules. *Accident Analysis & Prevention*, Vol. 160, 2021, p. 106306. https://doi.org/10.1016/j.aap.2021.106306.
34. Sivasankaran, S. K., P. Natarajan, and V. Balasubramanian. Identifying Patterns of Pedestrian Crashes in Urban Metropolitan Roads in India Using Association Rule Mining. *Transportation Research Procedia*, Vol. 48, 2020, pp. 3496–3507. https://doi.org/10.1016/j.trpro.2020.08.102.

35. Shahin, M., S. Saeidi, S. A. Shah, M. Kaushik, R. Sharma, S. A. Peious, and D. Draheim. Cluster-Based Association Rule Mining for an Intersection Accident Dataset. Presented at the 2021 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube), 2021.
36. Momeni Kho, S., P. Pahlavani, and B. Bigdeli. Classification and Association Rule Mining of Road Collisions for Analyzing the Fatal Severity, a Case Study. *Journal of Transport & Health*, Vol. 23, 2021, p. 101278. https://doi.org/10.1016/j.jth.2021.101278.
37. Khoda Bakhshi, A., and M. M. Ahmed. Real-Time Crash Prediction for a Long Low-Traffic Volume Corridor Using Corrected-Impurity Importance and Semi-Parametric Generalized Additive Model. *Journal of Transportation Safety & Security*, Vol. 14, No. 7, 2022, pp. 1165–1200. https://doi.org/10.1080/19439962.2021.1898069.
38. Demir, S., and E. K. Sahin. Random Forest Importance-Based Feature Ranking and Subset Selection for Slope Stability Assessment Using the Ranger Implementation. *Avrupa Bilim ve Teknoloji Dergisi*, No. 48, 2023, pp. 23–28.
39. Pande, A., and M. Abdel-Aty. Market Basket Analysis of Crash Data from Large Jurisdictions and Its Potential as a Decision Support Tool. Safety science, Elsevier, 47, 1, 2009, pp. 145–154.
40. Hossain, M. M., M. R. Abbaszadeh Lima, and H. Zhou. Severity Analysis of Secondary Crashes on High-Speed Roadways: Pattern Recognition Using Association Rule Mining. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2678, No. 8, 2024, pp. 919–931. https://doi.org/10.1177/03611981231223194.
41. Abbaszadeh Lima, M. R., M. M. Hossain, H. Zhou, and Y. Song. Data Mining Approach to Explore the Contributing Factors to Fatal Wrong-Way Crashes by Local and Non-Local Drivers. Future Transportation, MDPI, 4, 3, 2024, pp. 985–999.
42. Hossain, A., X. Sun, S. Islam, S. Alam, and M. M. Hossain. Identifying Roadway Departure Crash Patterns on Rural Two-Lane Highways under Different Lighting Conditions: Association Knowledge Using Data Mining Approach. Journal of safety research, Elsevier, 85, 2023, pp. 52–65.
43. Hossain, M. M., H. Zhou, and S. Das. Data Mining Approach to Explore Emergency Vehicle Crash Patterns: A Comparative Study of Crash Severity in Emergency and Non-Emergency Response Modes. Accident Analysis & Prevention, Elsevier, 191, 2023, p. 107217.
44. Kumbhare, T. A., and S. V. Chobe. An Overview of Association Rule Mining Algorithms. *International Journal of Computer Science and Information Technologies*, Vol. 5, No. 1, 2014, pp. 927–930.
45. Hossain, M. M., H. Zhou, M. A. Rahman, S. Das, and X. Sun. Cellphone-Distracted Crashes of Novice Teen Drivers: Understanding Associations of Contributing Factors for Crash Severity Levels and Cellphone Usage Types. *Traffic Injury Prevention*, Vol. 23, No. 7, 2022, pp. 390–397. https://doi.org/10.1080/15389588.2022.2097667.
46. Dol, S. M., and P. M. Jawandhiya. Classification Technique and Its Combination with Clustering and Association Rule Mining in Educational Data Mining—A Survey. Engineering Applications of Artificial Intelligence, Elsevier, 122, 2023, p. 106071.
47. Gu, C., J. Xu, C. Gao, M. Mu, G. E, and Y. Ma. Multivariate Analysis of Roadway Multi-Fatality Crashes Using Association Rules Mining and Rules Graph Structures: A Case Study in China. PLoS One, Public Library of Science San Francisco, CA USA, 17, 10, 2022, p. e0276817.
48. Kim, T. Y., K. Han, and S. O. Hwang. A New Association Rule Mining Based on Coverage and Exclusion for Network Intrusion Detection. *Journal of Internet of Things and Convergence*, Vol. 9, No. 1, 2023, pp. 77–87.
49. Kong, X., S. Das, K. Jha, and Y. Zhang. Understanding Speeding Behavior from Naturalistic Driving Data: Applying Classification Based Association Rule Mining. Accident Analysis & Prevention, Elsevier, 144, 2020, p. 105620.

50. Samerei, S. A., K. Aghabayk, N. Shiwakoti, and S. Karimi. Modelling Bus-Pedestrian Crash Severity in the State of Victoria, Australia. *International Journal of Injury Control and Safety Promotion*, Vol. 28, No. 2, 2021, pp. 233–242. https://doi.org/10.1080/17457300.2021.1907597.
51. Hossain, M. M., X. Sun, E. Mitran, and M. A. Rahman. Investigating Fatal and Injury Crash Patterns of Teen Drivers with Unsupervised Learning Algorithms. IATSS research, Elsevier, 45, 4, 2021, pp. 561–573.
52. Stimpson, J. P., F. A. Wilson, and R. L. Muelleman. Fatalities of Pedestrians, Bicycle Riders, and Motorists Due to Distracted Driving Motor Vehicle Crashes in the U.S., 2005–2010. *Public Health Reports*®, Vol. 128, No. 6, 2013, pp. 436–442. https://doi.org/10.1177/003335491312800603.
53. Pour-Rouholamin, M., and H. Zhou. Investigating the Risk Factors Associated with Pedestrian Injury Severity in Illinois. Journal of safety research, Elsevier, 57, 2016, pp. 9–17.
54. Agrawal, R., T. Imieliński, and A. Swami. Mining Association Rules between Sets of Items in Large Databases. Presented at the SIGMOD/PODS93: Joint ACM SIGMOD International Conference on Management of Data and ACM SIGMOD, Washington D.C. USA, 1993.
55. Curry, A. E., J. H. Mirman, M. J. Kallan, F. K. Winston, and D. R. Durbin. Peer Passengers: How Do They Affect Teen Crashes? Journal of Adolescent Health, Elsevier, 50, 6, 2012, pp. 588–594.
56. Ouimet, M. C., A. K. Pradhan, A. Brooks-Russell, J. P. Ehsani, D. Berbiche, and B. G. Simons-Morton. Young Drivers and Their Passengers: A Systematic Review of Epidemiological Studies on Crash Risk. Journal of Adolescent Health, Elsevier, 57, 1, 2015, pp. S24–S35.
57. Lidbe, A., E. K. Adanu, E. Tedla, and S. Jones. Role of Passengers in Single-Vehicle Drunk-Driving Crashes: An Injury-Severity Analysis. Safety, MDPI, 6, 2, 2020, p. 30.