Nighttime Pedestrian Safety in Different Communities: Application of Artificial Intelligence Techniques

Sheikh M. Usman,[a] Asad J. Khattak,[a,*] A. Latif Patwary[a]

[a] Department of Civil & Environmental Engineering, The University of Tennessee, Knoxville, United States.

[*] Corresponding author.

Email addresses: susman1@vols.utk.edu (S.M. Usman),
akhattak@utk.edu (A.J. Khattak)
apatwary@vols.utk.edu (A.L. Patwary)

Nighttime Pedestrian Safety in Different Communities: Application of Artificial Intelligence Techniques

Abstract

Introduction: Pedestrian safety is a growing concern in the United States transportation sector, with around 7500 pedestrian crash fatalities reported in the US in 2021. Pedestrians are at an even higher risk of crashes at night due to limited visibility and alcohol impairment of the drivers or pedestrians. The United States Department of Transportation (USDOT) has integrated six unique indicators (Economy, Health, Language Proficiency, Resilience, Environmental, and Transportation Access) at the census tract level to promote safety and achieve its Vision Zero goal. This study integrates these six indicators with pedestrian crash data from 2016 to 2019 in North Carolina. The pedestrian crash data are extracted from police reports using the Pedestrian and Bicyclist Crash Analysis Tool, which provides high-quality, detailed crash-type descriptors, resulting in a unique and comprehensive pedestrian crash database. *Methods:* The study applies rigorous methods for analysis, including the inference-based ordered logit model, to quantify key correlates of nighttime pedestrian crash injury severity in communities. To improve forecasting of pedestrian crashes and the resulting injury severity for planning purposes, an Artificial Intelligence (AI) based heterogeneous ensemble method, "Stacking," is applied with an Ordered Logit model and machine-learning techniques, Gradient Boosting, Decision Tree, and Random Forest as the base learners. *Results:* The stacked model integrates the predictive advantages of the base learner models and yields better predictive accuracy (78.85%) than the best-performing base learner (73.56%). The model results reveal associations between the Economy and Transportation Access indicators, roads without lights, pedestrian crossing violations, alcohol impairment, and nighttime pedestrian crash injury severity. *Conclusions and Practical Applications:* The study findings and the application of AI methods can assist safety practitioners in implementing targeted interventions to improve roadway infrastructure and overall road user safety.

Keywords: Roadway infrastructure; Injury severity; Heterogeneous ensemble methods; Stacking; Pedestrian.

1. Introduction

In the United States, pedestrian fatalities have risen recently, with an increase of more than 50% from 2009 to 2019, with more than 85% of those pedestrian fatalities occurring at night (FARS, 2019). Around 75% of pedestrian fatalities occurred at night in 2019, while total pedestrian fatalities in traffic crashes were estimated to be about 7500 in 2021, the highest in the last four decades (GHSA, 2021). Rising pedestrian fatalities in recent years highlight the urgent need to address pedestrian safety, particularly in communities with unique crash risks and challenges of built environments e.g., communities including low-income neighborhoods and communities of color, with inadequate pedestrian-supportive infrastructure, such as sidewalks, crosswalks, streetlights, etc., can face higher risks of pedestrian crashes. Pedestrians in some communities can be at higher risk of crashes at night, partly due to the lack of roadway lighting and the alcohol impairment of drivers and pedestrians. Although the United States Department of Transportation (USDOT) is motivated to achieve its ambitious target of zero traffic fatalities and severe injuries by 2050 (Vision Zero Goal), the recent pedestrian crash statistics indicate that considerable efforts are needed to accomplish this goal. The USDOT identifies communities for prioritizing targeted interventions to bridge the gaps in transportation infrastructure and promote a safe, affordable, and modernized transportation system. The USDOT has identified communities marked with a lack of transportation infrastructure, low income, low level of education attainment, adverse impacts of weather, etc., by collecting data for 22 socioeconomic indicators at the census tract level and classifying them into six broad themes. The census tracts are designated as burdened regarding transportation access, economy, health, resilience, proficiency, and environment. They were accessed and used for this study. The availability of this unique data provides new research pathways to explore the relationship between these indicators and pedestrian crash injury severity through statistical modeling and novel ML techniques.

The paper explores the correlates of nighttime crashes in different communities, focusing on pedestrians. This unique objective can help us address long-standing infrastructural, financial, and policy-related differences relevant to transportation safety in these communities. The study sheds light on the correlates of nighttime pedestrian crashes and the resulting injury severity through data-driven modeling.

Specifically, by applying an inference-based Ordered Logit model, this study investigates the association of the six unique indicators and the conventional crash contributing factors (vehicle, driver, pedestrian, and roadway characteristics) with nighttime pedestrian crash injury severity. Furthermore, to accurately forecast pedestrian crashes and the resulting injury severity for effective planning and infrastructure improvements, the study applies a heterogenous ensemble method (HEM) known as "Stacking." The study reflects one of USDOT's top priorities, development of a safe transportation system, and aligns with the Safe Systems Approach by addressing the safe users and safe road infrastructure components. The study findings can assist policymakers and safety practitioners in implementing more effective pedestrian safety countermeasures tailored to meet the specific requirements of different communities.

2. Literature review

Many studies have explored various aspects of nighttime pedestrian safety, including the absence of road lights (Siddiqui et al., 1982), consumption of alcohol or drugs by pedestrians or drivers (Ferenchak & Abadi, 2021; Pour-Rouholamin & Zhou, 2016), pedestrian crashes with large trucks and SUVs (Kim et al., 2008), pedestrian crashes at mid-blocks, straight road segments, multilane roads, and roads with posted speed limits of 45-50 mph (Aziz et al., 2013; Z. Chen & Fan, 2019a; Das et al., 2020; Rashid et al., 2023) resulting in serious and fatal injuries. However, exploring the differences in pedestrians' sociodemographic, socioeconomic, and racial attributes is important to understand their impact on pedestrian crash injury severity in different communities. Studies have found that pedestrian crashes in the US are linked to a higher population density, a greater number of people aged 15-64 years, and an increase in pedestrians who walk in pedestrian-unfriendly areas (Chimba et al., 2018). Another study focused on the racial and ethnic division among pedestrians in the US found that Black and Native American pedestrians are more likely to suffer fatal injuries in nighttime pedestrian-vehicle crashes, with a higher percentage of fatalities, 79% and 83%, respectively, compared to 72% for White pedestrians (Sanders & Schneider, 2022). The higher likelihood of fatal road traffic crashes in communities can be attributed to many factors (Patwary et al., 2024). These communities are marked with lower vehicle ownership and higher non-motorized travel modes, including higher pedestrian activity. A higher proportion of pedestrians in the traffic stream can lead to more dangerous pedestrian-vehicle interactions. The incumbents of certain communities are more likely to experience unsafe traffic conditions with a lack of pedestrian-specific infrastructure, i.e., sidewalks, dedicated pedestrian crossings, and pedestrian signs (Noland et al., 2013), resulting in a higher crash risk for pedestrians. These communities have a dearth of resources and a lack of physical infrastructure than more privileged neighborhoods (Gibbs et al., 2012). The location of schools and drop/pick-up zones in such communities can result in more pedestrian crashes due to a lack of pedestrian-supporting infrastructure and an elevated level of pedestrian activity at those locations. Yu et al., 2022 found that traffic crashes were more prevalent in school zones in neighborhoods with different racial populations. Furthermore, non-English speaking residents in such communities are more vulnerable to traffic crashes due to linguistic barriers and misinterpreting traffic rules (Chen et al., 2012).

Recent international studies on pedestrian safety provide significant insights into the risk factors and countermeasures to reduce pedestrian-related crashes. In Ghana, the absence of median separations and street lighting emerged as critical contributors to pedestrian hit-and-run accidents, highlighting the need for median installations and better lighting to enhance pedestrian safety (Aidoo et al., 2013). A study from Spain highlighted how pedestrian behavior, such as distraction, significantly increases the likelihood of severe injuries, especially in poorly lit, high-speed areas, suggesting that behavioral interventions and infrastructure upgrades are crucial (Febres et al., 2021). In India, the analysis of fatal pedestrian crashes at urban intersections revealed the adverse impact of high-speed vehicles, inadequate pedestrian infrastructure, and spatial factors like population density, further emphasizing the necessity of dedicated pedestrian spaces and traffic-calming measures (Mukherjee & Mitra, 2020). Meanwhile, research in Hong Kong demonstrated that the accessibility of footbridges and underpasses significantly reduces severe pedestrian crashes, reinforcing the importance of well-integrated pedestrian networks in urban planning (Zhu et al., 2023). Lastly, findings from Colombia indicated that while footbridges improve safety, their effectiveness is undermined by pedestrians' reluctance to use them due to perceptions of insecurity and inconvenience, highlighting the need for comprehensive strategies that address both safety infrastructure and user behavior (Oviedo-Trespalacios

& Scott-Parker, 2017). Together, these studies illustrate the multifaceted challenges in pedestrian safety and the importance of integrating engineering, behavioral, and urban planning approaches to mitigate risks globally.

Parametric statistical models were widely used to analyze crash injury severity of vulnerable road users in previous studies, including the Multinomial Logit (Z. Chen & Fan, 2019b; Tay et al., 2011), Ordered Logit and Probit (Eluru et al., 2008; Nasri et al., 2022; Yasmin et al., 2014), Random Parameters logit (Ijaz et al., 2021, 2022), Partial Proportional Odds models (Sasidharan & Menéndez, 2014) and the binary logit model (Usman & Khattak, 2025). However, the predictive accuracy of statistical models is usually lower than ML models due to the inherent restrictions on model specifications in these models. Unlike statistical models, machine learning methods offer more flexibility in model specifications and are better suited for capturing complex nonlinear relationships in crash data modeling (Pan et al., 2017). While ML methods generally offer higher prediction accuracy than statistical models, they explicitly address the bias-variance trade-off, with different approaches minimizing either bias or variance (Ahmad et al., 2023). Hence, relying solely on a single supervised or unsupervised ML method may lead to less accurate predictions.

More robust and diverse HEM can be utilized to enhance the prediction performance of ML methods (Bhatt et al., 2017; Dietterich, 2000; Tewari & Dwivedi, 2020). Heterogeneous ensembles, such as stacking, use various feature selection algorithms (e.g., Decision Tree, Random Forest, Gradient Boosting, etc.), where a stacking meta-learner combines the optimal predictions from the individual models, also called base-learners, acting as a single decision-maker in the second stage (Chali et al., 2014; Elish et al., 2013; Sabzevari et al., 2018). The literature suggests that the ensemble decisions of multiple models, even without an explicit stacked model, yield better predictive accuracy than single ML and statistical models. A study conducted to predict harsh braking incidents in various areas of Athens, Greece, demonstrated this by combining the results of Geographically Weighted Poisson Regression, Bayesian Conditional Autoregressive (CAR) models, and two variants of Extreme Gradient Boosting (XGBoost) models (Ziakopoulos et al., 2022). The ensemble model consisting of unweighted average predictions of harsh breaking events from the four individual models achieved an impressive prediction accuracy of over 87%, highlighting the effectiveness of aggregating insights from diverse modeling techniques. Studies suggest that heterogeneous ensembles outperform conventional statistical models and other ML methods, demonstrating superior prediction performance (Ahmad et al., 2023; Fernández-Alemán et al., 2019; Petrakova et al., 2015).

Given the predictive advantages of the stacking methodology, surprisingly, very few studies have used this approach to analyze problems related to transportation safety (Ahmad et al., 2023; Ghandour et al., 2020; Tang et al., 2019). Ahmad et al. (2023) applied stacking methodology to predict crash frequency in five-lane roadway segments in Tennessee, while Ghandour et al. (2020) and Tang et al. (2019) applied a stacking framework to explore crash injury severity in traffic crashes. Notably, applying heterogeneous ensembles (stacking) to specifically predict nighttime pedestrian crash injury severity is new and novel. Importantly, nighttime pedestrian safety in different communities is an unexplored area of research with immense potential to reveal the unique infrastructure needs and safety threats experienced by pedestrians in these communities. This study addresses the existing gaps in the literature by applying the robust heterogeneous ensemble (stacking) approach to accurately forecast nighttime pedestrian crash injury severity resulting from vehicle-pedestrian crashes in transportation-burdened census tracts in North Carolina. The contribution of this study lies in applying the robust stacking framework to a unique and comprehensive pedestrian crash dataset curated meticulously by integrating the indicators to focus on the infrastructural and financial differences related to pedestrian safety in different communities.

3. Conceptual Framework

The study aims to answer the central research question: What are the infrastructural, financial, and policy-related differences in communities associated with nighttime pedestrian crash severity? The study explores these differences by integrating indicators, defined by the USDOT, with a comprehensive pedestrian crash dataset. Specifically, six themes of indicators —economic, environmental, proficiency, health, resilience, and transportation access—are examined to identify their association with pedestrian crash severity outcomes. This exploration emphasizes the intersection of traffic safety in different communities for pedestrians in nighttime conditions.

Understanding the spatial and demographic patterns of crash severity in these communities is essential for targeted interventions and policies. By integrating indicators with crash data and applying advanced statistical and machine learning methods, this study provides insights into the multifaceted factors contributing to pedestrian injury severity. These insights can inform transportation planning and policymaking and assist in ensuring that the unique challenges communities face are adequately addressed. The conceptual framework is shown in Figure 1.

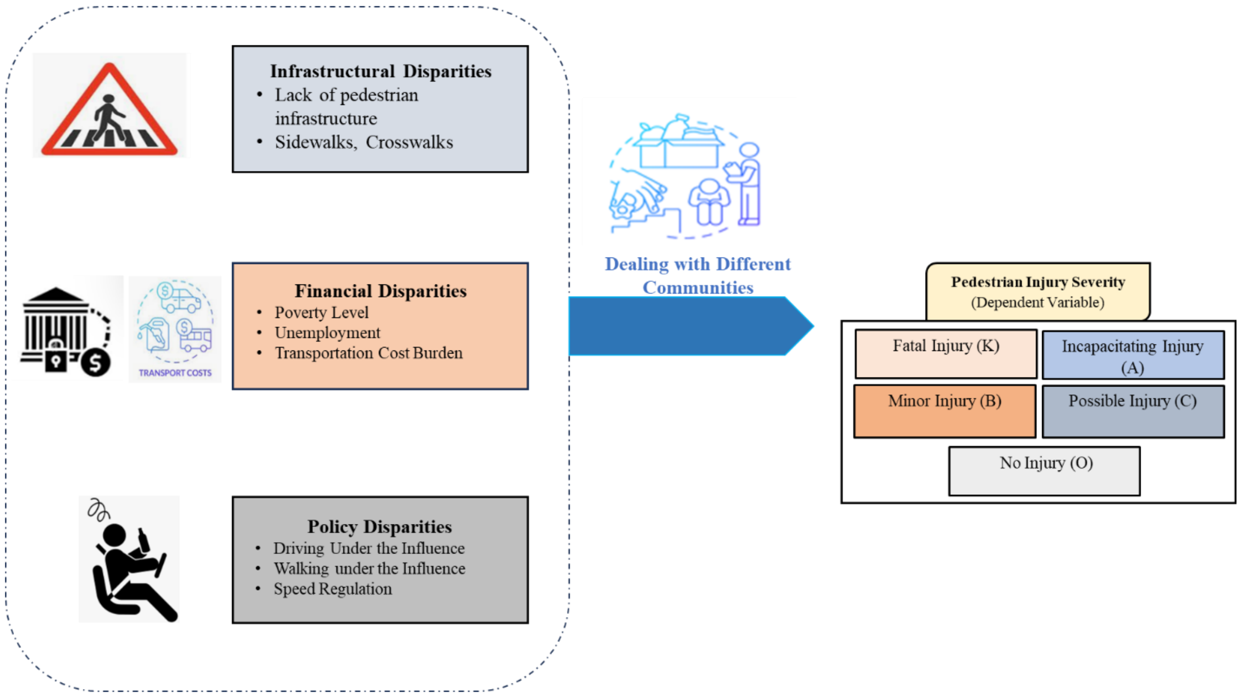


Fig 1. Study Conceptual Framework

4. Data Description

This study integrates two datasets for analysis. Pedestrian crash information for North Carolina from 2016-2019 was obtained from police-reported crash data. The data from this period was selected to ensure homogeneity in pedestrian travel behavior by focusing exclusively on the pre-COVID-19 era. Analyzing pedestrian crashes during COVID-19 could have resulted in obscured findings, potentially painting a falsely safer picture of nighttime pedestrian safety. Such distortions would likely stem from the substantial reduction in pedestrian activity during the pandemic rather than genuine safety improvements. The reason for exploring pedestrian crashes in North Carolina stems from the fact that the state of North Carolina reflects an ample population and sociocultural and geographical diversity. The state's population represents a fusion of ethnic and racial diversity with substantial representation of the Asian American, African American, and Native American racial groups (United States Census Bureau, 2021). This variety in population contributes to the state's multiculturalism. It presents an opportunity for researchers to identify infrastructural, financial, or policy-related differences relevant to transportation safety among the diverse population segments. The extracted police-reported pedestrian crashes are coded by a comprehensive crash-typing tool called the

Pedestrian and Bicycle Crash Analysis Tool (PBCAT). The PBCAT tool codes unique and comprehensive crash descriptors. It provides valuable information related to drivers' and pedestrians' pre-crash positions and unsafe actions. The key variables present in the pedestrian crash database include driver and pedestrian sociodemographic details (e.g., age, gender, and race), driver and pedestrian impairment (due to alcohol or drugs), the position of pedestrians at the time of a crash, roadway factors (roadway types/classes, and configuration), crash types, lighting conditions, speed limit, land use, etc. (independent variables). The injury severity of both pedestrians and drivers is reported on a five-level KABCO injury classification scale. However, given the scope of the present study, only the severity of pedestrian injury is selected as the dependent variable for the analysis.

To identify the unique challenges different communities face, this study integrates the indicators data collected by the US Department of Transportation with the pedestrian crash dataset. These indicators address six broad themes: economy, environment, proficiency, health, resilience, and transportation access. The composition of these indicators from the key relevant attributes is provided in Table 1. For more details on these indicators, please refer to (Patwary et al., 2024; USDOT, 2023c).

Table 1. Composition of Indicators from Key Attributes

Attribute Description	Broad Theme
Total workers 16 or older in a census tract	Transportation
Percentage of non-transit households that have 0 vehicles	Transportation
Percentile percentage of households with no vehicle available estimate	Transportation
Number of transit users 16 and over	Transportation
Five Year average price of gas per state	Transportation
Calculated average number of cars per household	Transportation
Calculated average cost of owning a car	Transportation
Calculated national average annual cost of using transit	Transportation
Calculated average annual cost of transportation	Transportation
Percentile of Mean commute time to work (in minutes)	Transportation
Annual Travel Time in Hours	Transportation
National Walkability Index	Transportation
Average of Transportation Indicator Percentiles (calculated)	Transportation
Percentile percentage of persons aged 65 and older estimate	Health
Adjunct variable - Percentage uninsured in the total civilian noninstitutionalized population estimate, 2014-2018 ACS	Health
Percentile percentage uninsured in the total civilian noninstitutionalized population estimate, 2014-2018 ACS	Health
Percentile percentage of civilian noninstitutionalized population with a disability estimate	Health
Average of Health Indicator Percentiles (calculated)	Health
Percentile Percentage of persons with no high school diploma (age 25+) estimate	Economy
Percentile Percentage of civilian (age 16+) unemployed estimate	Economy
Percentile per capita income estimate	Economy
Percentile Percentage of persons below the poverty estimate	Economy
Percent of Household Units with 30 percent or more income towards housing costs	Economy
Average of Economy Indicator Percentiles (calculated)	Economy
Percentile percentage of persons (age 5+) who speak English "less than well" estimate	Proficiency
Average of Social and Equity Indicator Percentiles (calculated)	Proficiency
Resilience Indicator	Resilience
Percentile for % pre-1960 housing (lead paint indicator)	Environmental
Percentile for Diesel particulate matter level in air	Environmental
Percentile for Air toxics respiratory hazard index	Environmental
Percentile for the PM 2.5 level in air	Environmental
Average of Environmental Indicators	Environmental

Data for these six unique indicators are based on the Center for Disease Control (CDC) Social Vulnerability Index, Census America Community Survey, Housing and Urban Development (HUD) Location Affordability Index, Federal Emergency Management Agency (FEMA) Resilience Analysis & Planning Tool, and FEMA National Risk Index for the year 2018 (Transportation Census Tracts, 2023). Each theme's relevant indicators are averaged to form an aggregated burden indicator. Using the crash location in the pedestrian crash dataset, the indicators of communities are obtained from USDOT's interactive dashboard for such census tracts in North Carolina and included in the pedestrian crash dataset as indicator variables having values of 1 and 0. The census tracts that are burdened in terms of four or more themes are considered overall burdened. A strategic approach was used in ArcGIS software to handle crashes occurring on roadways at the borders of adjacent census tracts. Specifically, a 10-meter buffer zone was created surrounding each census tract. Crashes within these buffer zones were included in the analysis of both adjacent tracts. This approach minimizes the arbitrary assignment of crashes to a single tract, ensures a more comprehensive representation of crash data across census tracts, and is consistent with the previous literature (Liu et al., 2024). A total of 2,420 records of pedestrian crashes occurring at night from 2016 to 2019 in North Carolina were obtained initially. However, some of the crash observations had many missing or unknown values. After removing the crash observations with missing values, the final dataset contained 2329 nighttime pedestrian crash observations. Notably, the dataset includes hit-and-run pedestrian crashes where law enforcement later traced and identified the drivers, allowing for complete driver-related information in these records. Among the 91 crashes removed from the original dataset due to missing or unknown values, 34 (37%) were hit-and-run cases. These records were excluded because they lacked sufficient data across multiple fields to be included in the analysis. It is worthwhile to mention here that the proportion of hit-and-run crashes ranges from 11-13% of the overall annual crash count from 2006-2015 in the US (Benson et al., 2018), and the percentage of hit-and-run crashes typically expected to be reported in crash databases ranges from 10-20% of the total crashes (Liu et al., 2018).

Additionally, pedestrian crashes that occurred in the same daytime period were extracted to provide a comparative analysis (daytime vs nighttime) of pedestrian crashes. The final datasets for pedestrian crashes in the daytime and nighttime consist of 2436 and 2329 observations, respectively, while the overall crash dataset contains 4765 observations of pedestrian crashes from 2016 to 2019. The descriptive statistics of the key study variables in the datasets are presented in Table 2.

Table 2. Results of Descriptive Statistics of Key Variables

Variables	**Daytime Crashes (N = 2436)**		**Nighttime Crashes (N = 2329)**	
Pedestrian Injury Severity (Dependent Variable)	**Frequency**	**Percentage %**	**Frequency**	**Percentage %**
No injury (O)	137	5.62	90	3.86
Possible Injury (C)	1021	41.91	656	28.17
Minor injury (B)	1002	41.13	882	37.87
Incapacitating/Serious Injury (A)	194	7.96	352	15.11
Fatal Injury (K)	82	3.37	349	14.98
Pedestrian Gender				
Male Pedestrians	1398	57.39	1618	69.47
Female Pedestrians	1038	42.61	711	30.53
Pedestrian Age				
Pedestrians aged less than 25 years	612	25.12	534	22.93
Pedestrians aged 25-60 years	1336	54.84	1539	66.08
Pedestrians aged above 60 years	488	20.03	256	10.99
Pedestrian Race				
Asian	47	1.93	33	1.42
Black	1020	41.87	1010	43.37
Hispanic	164	6.73	165	7.08
Native American	6	0.25	42	1.80
White	1199	49.22	1079	46.33
Driver Gender				
Male Drivers	1350	55.42	1468	63.03
Female Drivers	1086	44.58	861	36.97
Driver Age				
Drivers aged less than 25 years	418	17.16	443	19.02
Drivers aged 25-60 years	1419	58.25	1484	63.72
Drivers aged above 60 years	599	24.59	402	17.26
Alcohol/Drug use				
Driver alcohol/drug use	45	1.85	102	4.38
Pedestrian alcohol/drug use	110	4.52	630	27.05
Both driver and pedestrian alcohol/drug use	4	0.16	46	1.98
No alcohol/drug use	2277	93.47	1551	66.59
Crash Location				
Intersection	1030	42.28	540	23.19
Non-intersection	1105	45.36	1535	65.91
Intersection related	301	12.36	254	10.91
Hit & Run Crash				
Yes	68	2.79	91	3.91
No	2368	97.21	2238	96.09
Type of Vehicle				
Commercial Bus	480	19.70	499	21.43
Motorcycle	23	0.94	17	0.73
Passenger Car	1333	54.72	1280	54.96
Truck/Trailer	99	4.06	66	2.83
SUV	501	20.57	467	20.05
Lighting Conditions				
Daylight	2436	100	-	-
Dark Roadways with Lights	-	-	1116	47.92
Dark Roadways without lights	-	-	1213	52.08

Roadway Functional Classification				
Interstates	58	2.38	54	2.32
Major Arterial	226	9.28	552	23.70
Minor Arterial	245	10.06	260	11.16
Collector Roads	59	2.42	165	7.08
Local Roads	1848	75.86	1298	55.73
Position of Pedestrian				
In Travel Lane	1268	52.05	1750	75.14
At Crosswalk	764	31.36	316	13.57
At intersection	46	1.89	35	1.50
Paved Shoulder	104	4.27	93	3.99
Sidewalk	141	5.79	27	1.16
Unpaved Right of Way	113	4.64	108	4.64

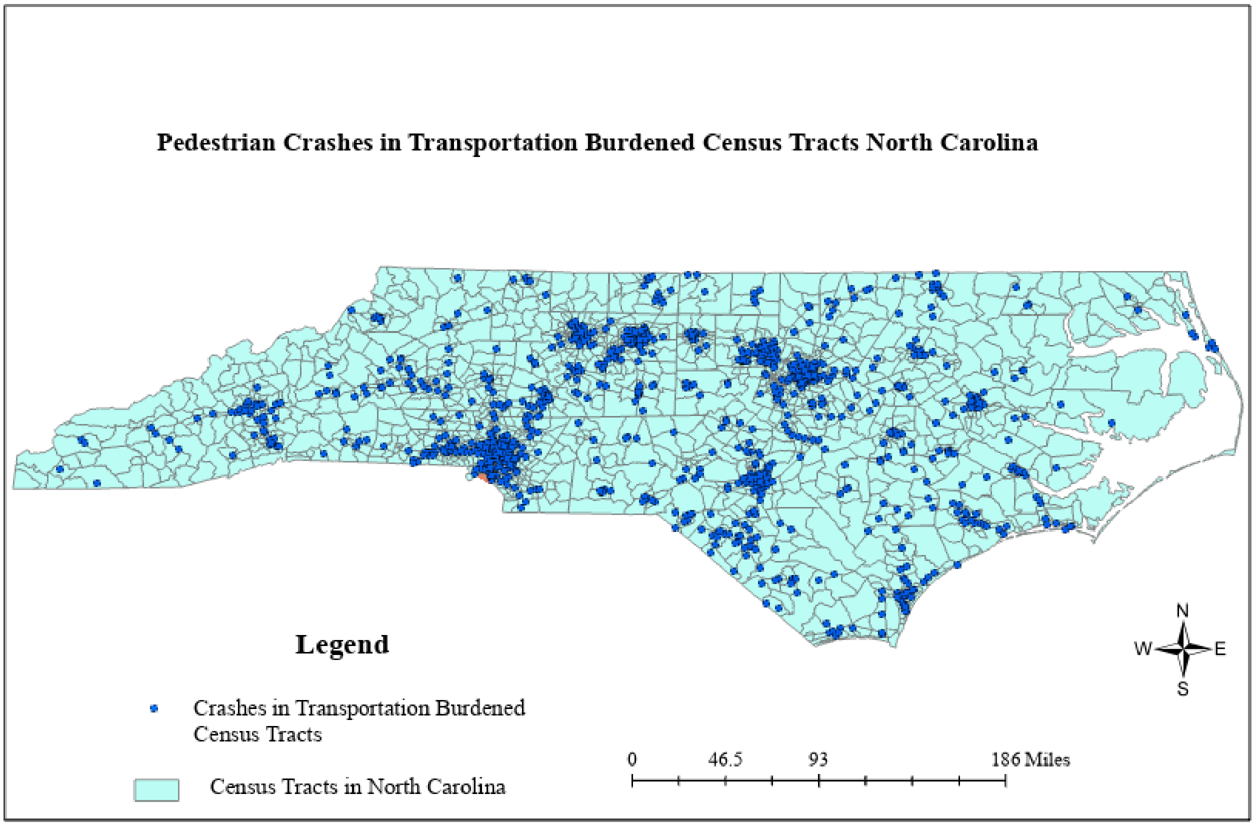
All the crash-related variables used for analysis in this study were categorical. Hence, the reported descriptive statistics in Table 2 include the crash frequencies and percentages associated with the classes of the categorical variables. Of 2,329 crashes at night, 37.87% resulted in minor injuries to pedestrians, while serious and fatal injuries accounted for 15.11% and 14.98% of crashes. Comparatively, during the daytime, 41.13% of crashes resulted in minor injuries, while serious and fatal injuries accounted for 7.96% and 3.37%, respectively. This indicates a higher proportion of serious and fatal injuries in nighttime crashes than in daytime crashes. Male pedestrians mostly experienced crashes at night (69.47%) compared to their female counterparts (30.53%), similar to the pattern observed in daytime crashes, where male pedestrians accounted for 57.39%, and female pedestrians accounted for 42.61% of the crashes. Male drivers were mostly involved in pedestrian crashes at night (63.03)% compared to female drivers (36.97%), which aligns with daytime crashes, where male and female drivers accounted for 55.42% and 44.58% of the crashes, respectively. This could be partially explained by the greater exposure of males to high-risk conditions, such as nighttime travel, which is consistent with trends observed in existing studies. However, it is important to note that the dataset does not include direct pedestrian or driver exposure measures. Further research is needed to examine the relationship between pedestrian and driver exposure and injury severity, given a crash. Among different racial groups, White pedestrians experienced the majority of nighttime crashes (46.33%), followed by Black pedestrians (43.37%), Hispanic pedestrians (7.08%), Native American pedestrians (1.80%), and Asian pedestrians (1.42%). White pedestrians accounted for 49.22%, Black pedestrians accounted for 41.87%, Hispanic pedestrians for 6.73%, Native American pedestrians for 0.25%, and Asian pedestrians for 1.93%

of the crashes in the daytime. According to the 2020 US Census, the general population of North Carolina consists of 61.6% people identified as White, 12.4% as Black, 18.7% as Hispanic, 6% as Asian, and 1.1% as Native American or Alaska Native (United States Census Bureau, 2021). This indicates that Black pedestrians are disproportionately overrepresented in nighttime crashes relative to their population share, while White and Hispanic pedestrians are underrepresented. These differences may warrant further investigation into factors contributing to racial differences in pedestrian crash involvement. Pedestrian crashes in which neither the driver nor the pedestrian was impaired by alcohol/drugs were the most frequent in both nighttime (66.59%) and daytime (93.47%) crashes. However, nighttime crashes showed a significantly higher proportion of crashes involving impaired pedestrians (27.05%) compared to daytime crashes (4.52%). Crashes involving impaired drivers alone also increased from 1.85% during the daytime to 4.38% at nighttime, while crashes involving both impaired drivers and impaired pedestrians rose from 0.16% during the daytime to 1.98% at nighttime. Passenger cars were the most common vehicle type involved in crashes during both timeframes, accounting for 54.96% of nighttime crashes and 54.72% of daytime crashes. Crashes involving SUVs were slightly less frequent at nighttime (20.05%) than at daytime (20.57%), while commercial bus involvement was slightly higher at nighttime (21.43%) compared to daytime (19.70%). A high percentage of crashes was observed on roads without lighting arrangements (51.22%) compared to roads with lights (47.92%) at nighttime. Most crashes during both the day and night occurred on local roads, accounting for 55.73% of nighttime crashes and 75.86% of daytime crashes. 23.70% of nighttime crashes occurred on major arterials while daytime crashes on the same facility had a share of 9.28%. Finally, most pedestrian crashes occurred when pedestrians were in the travel lane during a crash in both timeframes, accounting for 75.14% of nighttime crashes and 52.05% of daytime crashes. This indicates the critical issue of pedestrian crossing violations. In comparison, only 13.57% of nighttime crashes and 31.36% of daytime crashes occurred when pedestrians crossed the road using dedicated crosswalks.

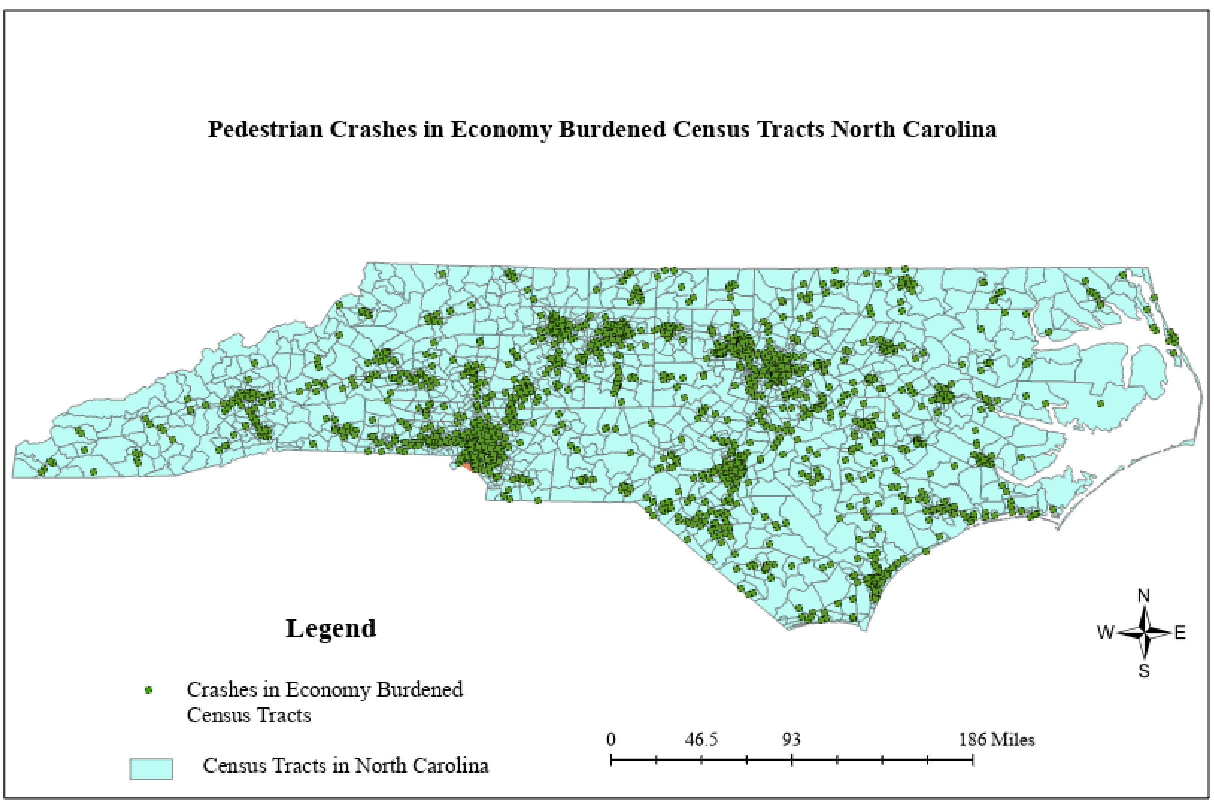
Table 3. Distribution of Nighttime Pedestrian Crashes in Census Tracts in North Carolina

Indicators	Status of Census Tract	Census Tract Count	Census Tract %	Crash Frequency	Crash %	Population	Population %
Transportation Access	Burdened	1304	59.49	1246	53.50	6089332	59.96
	Non-Burdened	888	40.51	1083	46.50	4066292	40.04
Health	Burdened	1414	64.51	1670	71.70	6190114	60.95
	Non-Burdened	778	35.49	659	28.30	3965510	39.05
Economy	Burdened	1158	52.83	2110	90.60	5171833	50.93
	Non-Burdened	1034	47.17	219	9.40	4983791	49.07
Proficiency	Burdened	1027	46.85	1096	47.06	5001488	49.25
	Non-Burdened	1165	53.15	1233	52.94	5154136	50.75
Resilience	Burdened	564	25.73	726	31.17	3069459	30.22
	Non-Burdened	1628	74.27	1603	68.83	7086165	69.78
Environment	Burdened	1062	48.45	1554	66.72	4835734	47.62
	Non-Burdened	1130	51.55	775	33.28	5319890	52.38
Overall	Burdened	817	37.27	1252	53.76	3771145	37.13
	Non-Burdened	1375	62.73	1077	46.24	6384479	62.87
Total		2192	100	2329	100	10155624	100

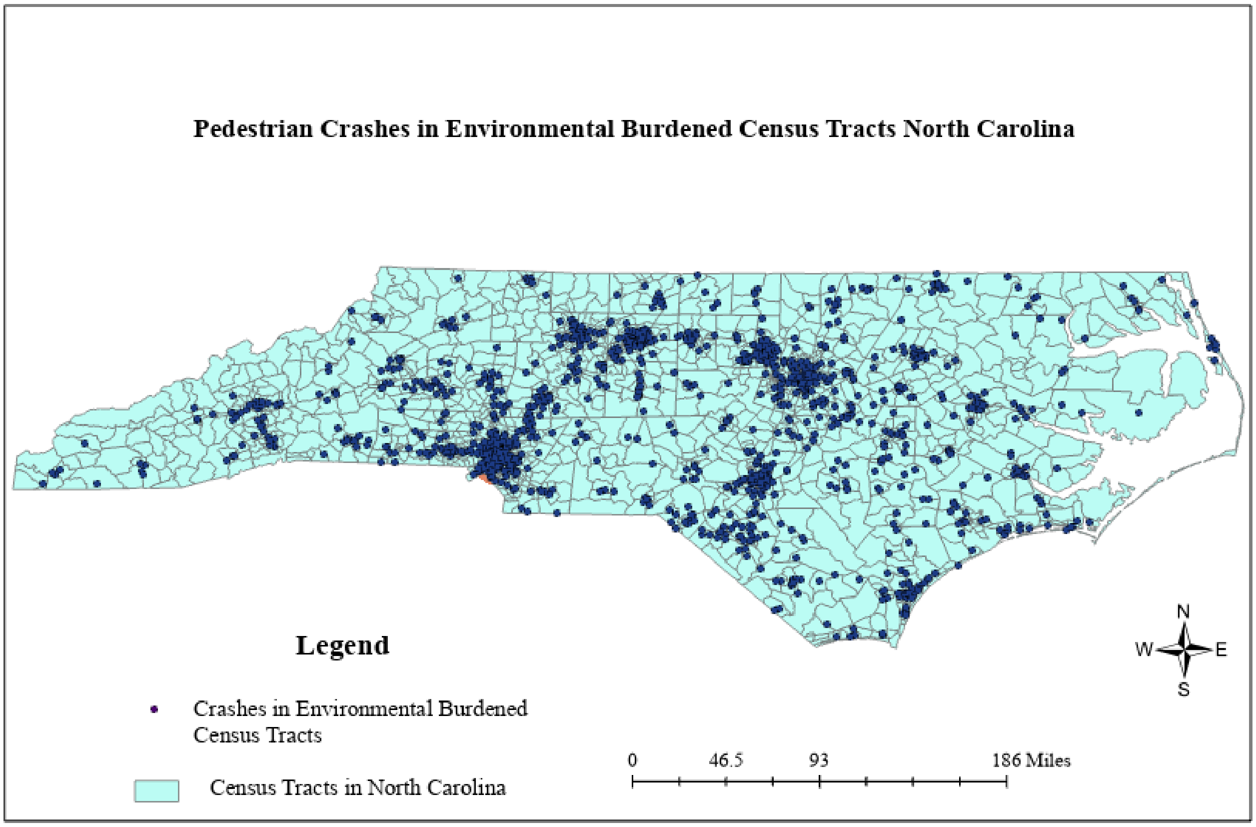
Referring to Table 3, North Carolina has been divided into 2192 census tracts. Among these census tracts, 64.51% are health-burdened, 59.49% are transportation-burdened, while the shares of census tracts burdened in terms of economy, proficiency, environment, and resilience are 52.83%, 46.85%, 48.45%, and 25.73%, respectively. The nighttime pedestrian crash frequency in these census tracts reveals that most of the crashes occurred in economy-burdened census tracts (90.60%), followed by health (71.70%) and environment-burdened (66.72%). Approximately 61% of the state's population resides in health-burdened census tracts. A similar percentage of the population (approximately 60%) lives in transportation-burdened census tracts. The spatial distribution of nighttime pedestrian crashes among the census tracts burdened by each of the themes in North Carolina is illustrated in Figure 2 (a) to (g).


Pedestrian Crashes in Transportation Burdened Census Tracts North Carolina
Legend
Crashes in Transportation Burdened Census Tracts
Census Tracts in North Carolina
0
46.5
93
186 Miles
N
W
E
S

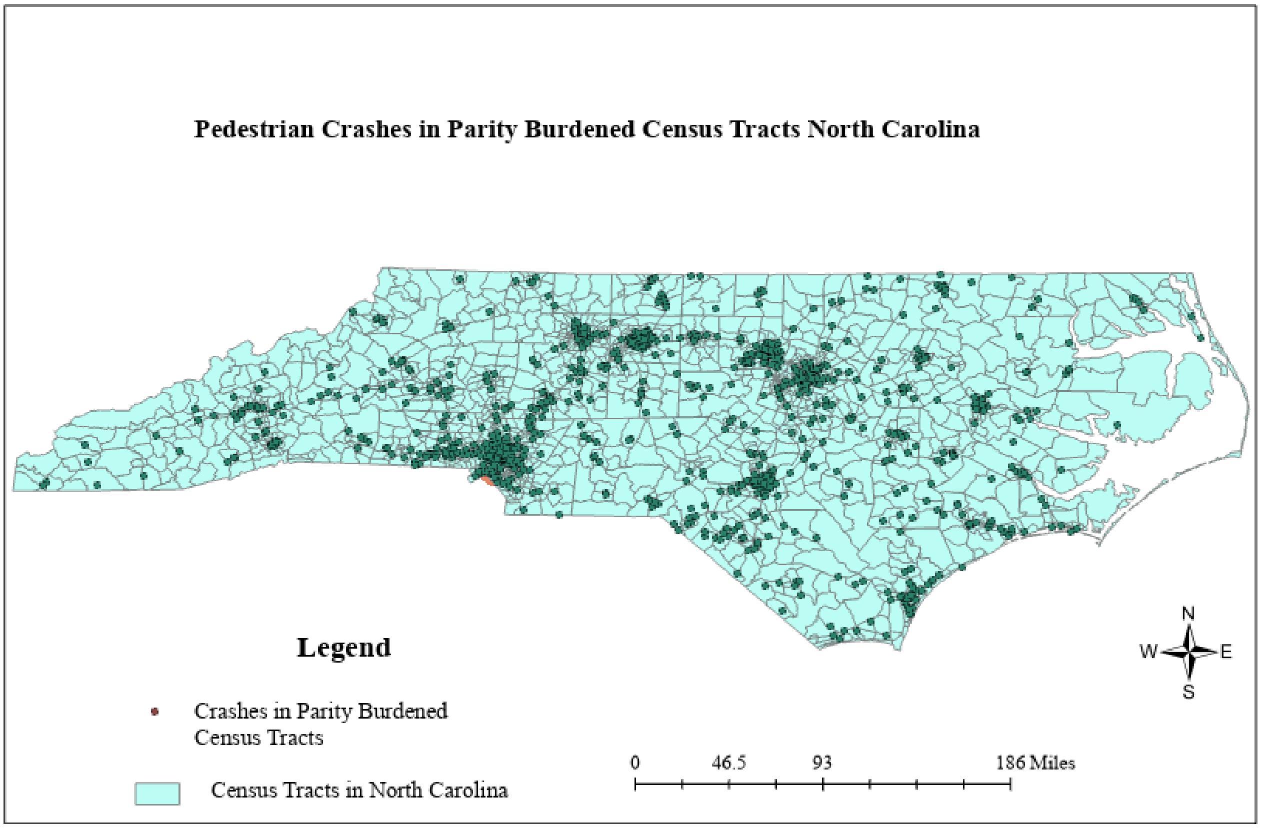
(a)


Pedestrian Crashes in Economy Burdened Census Tracts North Carolina
Legend
Crashes in Economy Burdened Census Tracts
Census Tracts in North Carolina
0
46.5
93
186 Miles
N
W
E
S

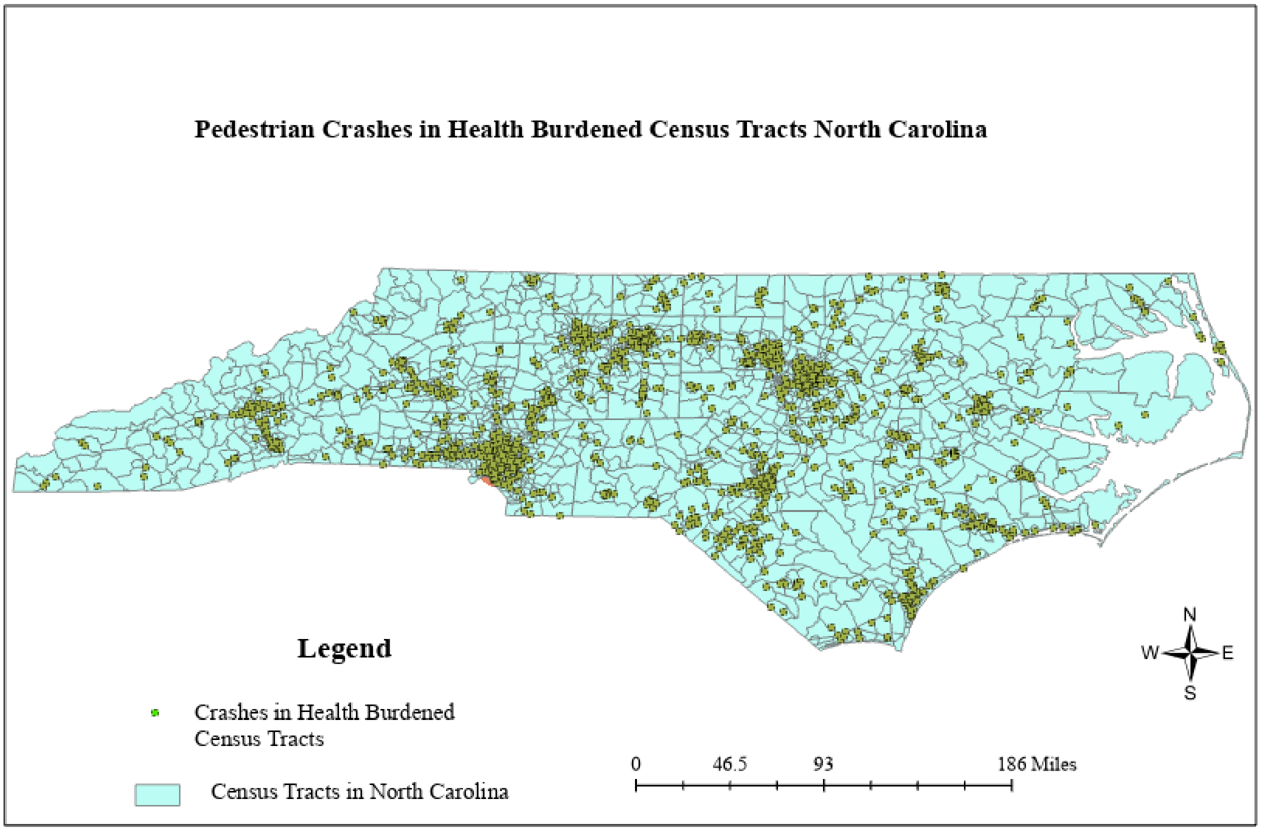
(b)


Pedestrian Crashes in Environmental Burdened Census Tracts North Carolina
Legend
Crashes in Environmental Burdened Census Tracts
Census Tracts in North Carolina
0
46.5
93
186 Miles
N
W
E
S

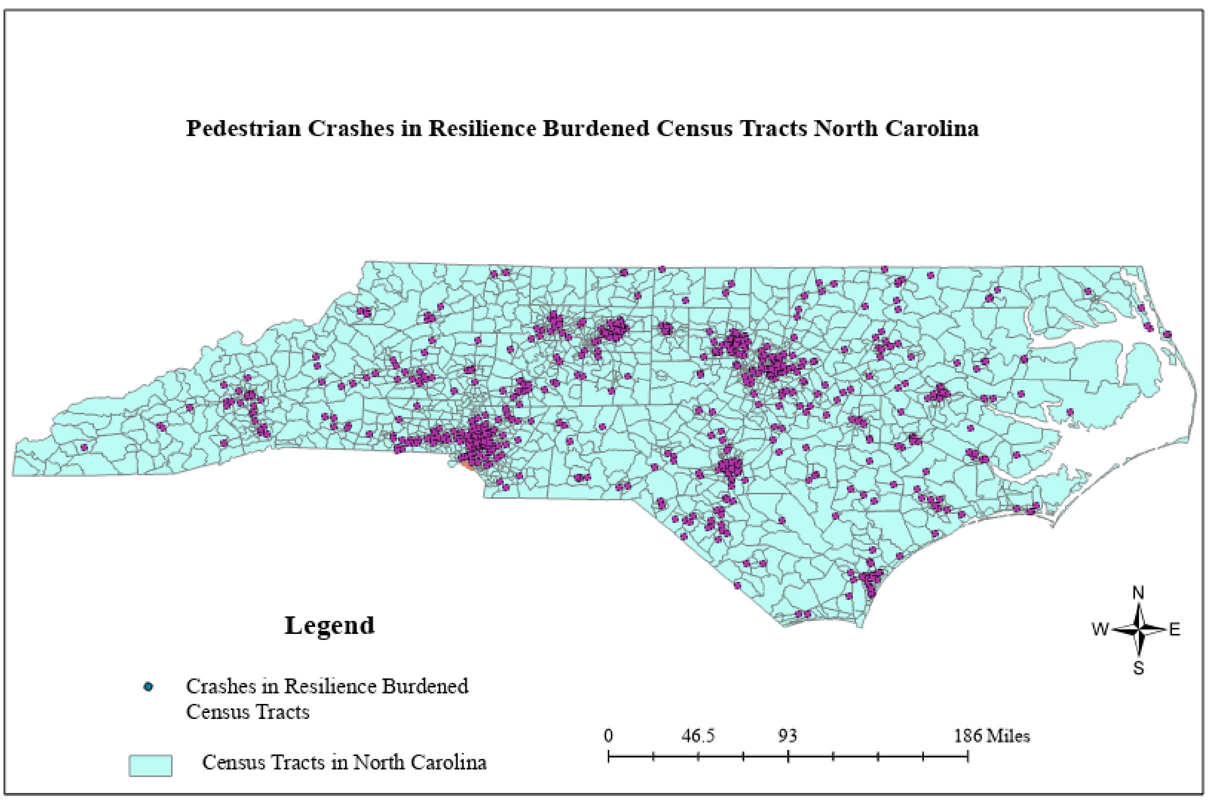
(c)


Pedestrian Crashes in Parity Burdened Census Tracts North Carolina
Legend
Crashes in Parity Burdened Census Tracts
Census Tracts in North Carolina
0
46.5
93
186 Miles
N
W
E
S

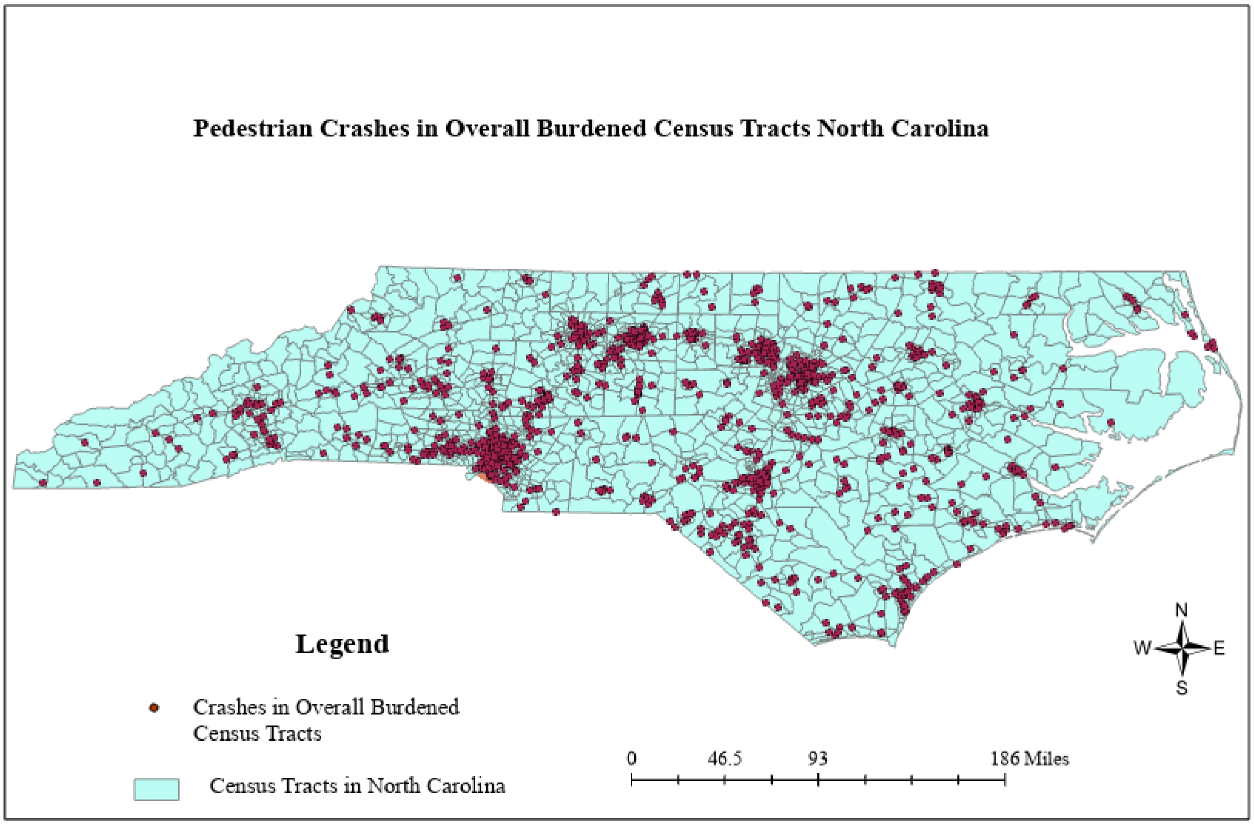
(d)


Pedestrian Crashes in Health Burdened Census Tracts North Carolina
Legend
Crashes in Health Burdened Census Tracts
Census Tracts in North Carolina
0
46.5
93
186 Miles
N
W
E
S

(e)


Pedestrian Crashes in Resilience Burdened Census Tracts North Carolina
Legend
Crashes in Resilience Burdened Census Tracts
Census Tracts in North Carolina
0
46.5
93
186 Miles
N
W
E
S

(f)



(g)

Fig. 2. Spatial Distribution of Nighttime Pedestrian Crashes in Communities - North Carolina

Fig. 2 (a-g) shows that the census tracts constituting the urban regions of Charlotte, Greensboro, Durham, Raleigh, and Fayetteville in North Carolina are mostly designated as burdened regarding the six indicators. A substantial share of the pedestrian crashes at night occurred in these census tracts. Identifying these census tracts is critical to ensure balanced pedestrian-oriented resource allocation and improve overall pedestrian safety.

The explanatory variables were checked for potential correlations with each other before the model estimation step. Either one of the variables found to be highly correlated (Pearson correlation coefficient higher than 0.5 or less than -0.5) was excluded from the model estimation process. Notably, the correlation matrix of the indicators (Table 4) shows that they are not highly correlated, and the values of Pearson's correlation coefficients among the indicators are less than 0.5.

Table 4. Correlation Matrix of Indicators

Indicators	Transportation Access	Health	Economy	Proficiency	Resilience	Environmental
Transportation Access	1.00					
Health	0.41	1.00				
Economy	-0.05	0.03	1.00			
Proficiency	0.23	0.33	0.08	1.00		
Resilience	0.29	0.19	-0.05	0.02	1.00	
Environmental	-0.34	-0.18	0.17	-0.12	-0.41	1.00

5. Methodology

A frequentist inference-based ordered logit model is estimated to explore the correlates of nighttime pedestrian crashes. A stacking approach is adopted for prediction, combining the predictions of four base learners (Ordered Logit, Decision Tree, Random Forest, and Gradient Boosting Model). A description of the stacking framework, the individual models, and performance measures for the models is provided in the subsequent sections.

5.1 Stacking

Stacking represents a HEM that combines the output of multiple classification models using a meta-classifier, also called a super-learner (Wolpert, 1992). Typically, it consists of two layers. Individual classification models are trained on the complete training set in the first layer. In the second layer, the meta-classifier is fitted based on the outputs of the first layer's models. The final result is the output of the meta-classifier. This framework allows the integration of various methods' advantages to enhance classification performance. Figure 3 represents a visual illustration of the study methodology.

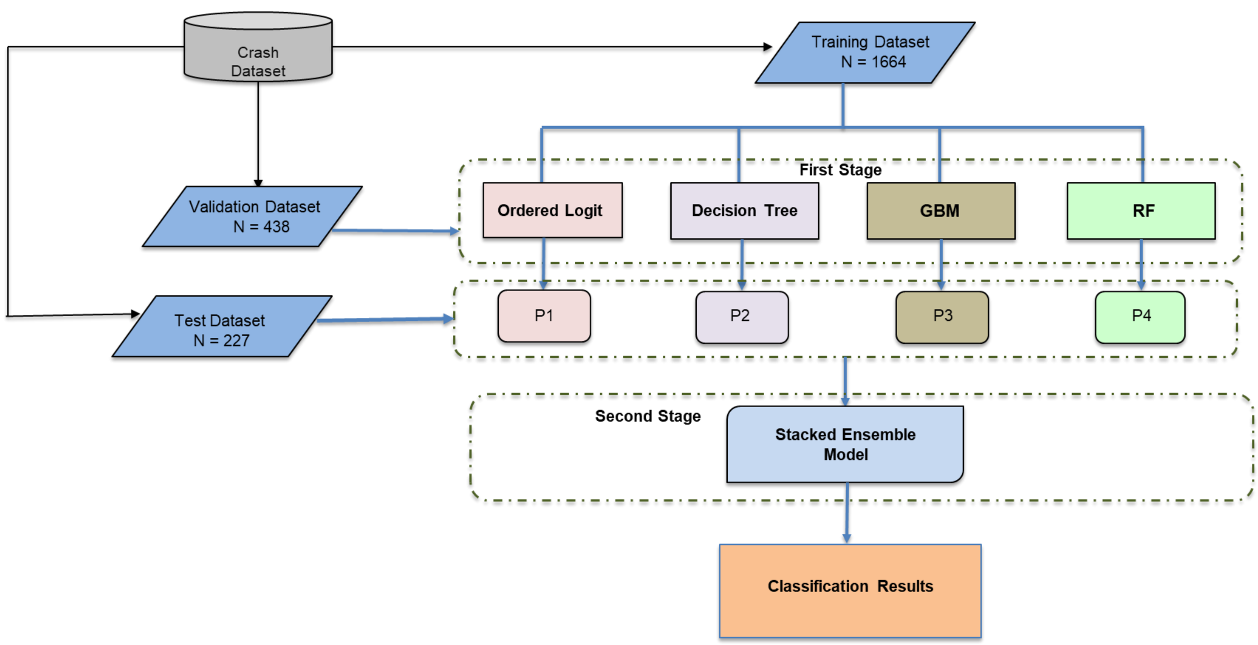


Fig. 3. Visual Illustration of Stacking Methodology

Referring to Figure 3, the nighttime pedestrian crash dataset is split into three segments (i.e., the training dataset N = 1664, the validation dataset N = 438, and the test dataset N = 227). The training dataset estimates the individual base learners, including the inference-based Ordered logit model, decision tree, gradient boosting model (GBM), and random forest (RF). Predictions of pedestrian injury severity are obtained from the base learners denoted as P_1, P_2, P_3, and P_4 using the validation dataset. These predictions act as input data for the stacked ensemble model in the second stage, and the predictions of pedestrian injury severity are obtained from the stacked ensemble model. The base learners use the test data to compare their predictive performance.

The base learners were selected based on the diversity of the ML algorithms, capturing different patterns in the data, as well as linear and nonlinear relationships, and statistical properties such as goodness of fit and significance of the models. Different combinations/subsets of the base learners, including some additional base learner algorithms, i.e., Support Vector Machines and Artificial Neural Networks, were tested, and their respective predictive accuracies and kappa values were obtained. Combining all four base learners (Ordered Logit, Decision Tree, Gradient Boosting, and Random Forest) with Random Forest as the meta-learner yielded optimal results regarding the performance measures reported in the results section.

5.2 Ordered Logit Model

Ordinal scales possess two key characteristics: (1) A distinct arrangement of levels is evident, indicating a clear order. (2) The precise distances between the various levels are not known. Extensive research and numerous established techniques (e.g., Multinomial logit, mixed logit models) exist to effectively and efficiently model categorical data by considering them nominal variables. Nevertheless, disregarding the inherent ordering information may yield dissimilar and less robust outcomes. Conversely, treating an ordered categorical variable as ordinal rather than nominal offers several advantages, including simplicity, straightforward interpretations, increased flexibility, and a closer resemblance to conventional regression analysis (Agresti, 2010; Zheng et al., 2014). The ordered logit model is appropriate for modeling a categorical response variable whose outcomes possess some intrinsic ranking. The formulation of the ordered logit model is provided in **Equation 1**.

$$Logit[P(Y \leq j)] = \alpha_j + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_n x_n = \alpha_j + \beta' X \quad (1)$$

Equation 1 indicates that for different categories j of the response variable Y, the effect of explanatory variables $x_1, x_2, \ldots, x_n$ is captured by a single coefficient β, which is important in interpreting this model.

5.3 Decision tree

The decision tree algorithm readily divides the training data repeatedly into smaller subsets based on minimizing the Gini impurity (Torgo, 1999). Gini impurity quantifies the probability of misclassifying a randomly chosen data point from the node. Mathematically, the Gini impurity of a node with multiple classes (C) is given in **Equation 2**.

$$Gini\ impurity = 1 - \sum (p_i)^2 \quad (2)$$

p_i is the probability of a data point belonging to class i in the node.

A fully grown decision tree is usually prone to overfitting and has poor generalizing capabilities (Kotsiantis, 2013). The tree is pruned to prevent overfitting. Pruning refers to the process of simplifying the tree structure by removing some nodes from the tree to prevent it from capturing random noise.

5.4 Gradient boosting decision tree

Friedman introduced Gradient Boosting (GB) in 2001 (Friedman, 2001). GB is a widely applied method to optimize parameters and improve model performance, achieved by updating parameters in the gradient descent direction of the previous model's loss function. The loss function is used to evaluate model performance, with a smaller value indicating better performance. Gradient Boosting Decision Tree (GBDT) jointly makes decisions by iterating through multiple decision trees. Each decision tree learns the residuals of all previous trees and fits a current decision tree. Iterations continue until a predefined number of trees are built or a specific stopping criterion is met. The final result accumulates all the decision trees generated throughout the iteration.

5.5 Random Forest

Breiman proposed a machine-learning algorithm called Random Forest (RF) in 2001 (Breiman, 2001). It builds an ensemble of decision trees using the bootstrap sampling method, which repeatedly selects training sets from the original data with replacement. Using a bagging technique, random operations select sample subsets to create training sets from the original samples. Bagging is used again to select subsets of features from the entire set. Additionally, the importance of each feature can be ranked based on its contribution to the final decision. Incorporating random operations in the random forest greatly improves its classification performance (Adeel et al., 2024; Khattak et al., 2023; Parmar et al., 2018; Usman et al., 2024).

Table 5 summarizes the study's methods and shows the mathematical formulation and rationale for the classification decision of the models.

Table 5. Mathematical Formulation of Methods Used in the Study

Methods	Mathematical Formulation	Classification Decision Rationale
Ordered Logit Model	$\log\left(\frac{P(Y\leq j)}{P(Y\geq j)}\right) = \alpha_j + \beta' X$, where Y = pedestrian crash injury severity; α and β are the estimated coefficients	Assigns categories based on the highest likelihood of cumulative probabilities.
Decision Tree	Split: Xi ≤ τ, where τ is the feature threshold. Impurity is minimized using the Gini Index: $G = 1 - \sum_{i=1}^{C} p_i^2$, where p_i is the proportion of class i in a node.	Observations are classified by traversing the tree to a terminal node with the majority class.
Gradient Boosting	$f_m(X) = f_{m-1}(X) + \gamma_m h_m(X)$ where γ_m minimizes the loss function L(y, f(X)).	It iteratively improves predictions by minimizing errors from previous models. **Classification decision** Weighted majority vote of weak learners
Random Forest	Y = Argmax $((\sum_{t=1}^{T} h_t(X))$, where $h_t(X)$ is the decision of the t-th tree.	**Classification decision:** Majority Voting. Combines votes from multiple decision trees for robust classification.

Stacking the heterogeneous ensemble models often benefits from combining 4-5 diverse base models (e.g., inferential and prediction-based parametric and non-parametric models) to improve predictions by leveraging the strengths of each model. This is consistent with previous literature, as numerous studies have shown the improved predictive performance of the stacked ensemble model while stacking 4-5 intrinsically different models (Ahmad et al., 2023; Tang et al., 2019). The Ordered Logit model, estimated in our study, provides an inferential framework for understanding the ordinal nature of pedestrian injury severity. At the same time, the Decision Tree, Random Forest, and Gradient Boosting methods capture non-linear relationships and complex interactions in the data. The stacked ensemble model leverages the predictive advantages of the individual models. Hence, the estimation of four different models in our study is justified, considering the improved predictive performance of the stacked ensemble model.

5.6 Performance measures

Four performance measures evaluate the predictive performance of the statistical and ML models using the test dataset as a holdout sample in this study. These measures include Kappa, precision, recall, and F1 score since these measures are more suitable for quantifying the predictive performance of ML models dealing with imbalanced datasets.

5.6.1 Kappa

The Kappa statistic compares the accuracy of the model to the expected accuracy of a random model i.e., the accuracy of a model that predicts classes at random with frequencies proportional to what appears in the data. Values of Kappa close to 1 indicate better models. It can be calculated using **Equation 3**.

$$Kappa\ statistic = \frac{Accuracy_{model} - Accuracy_{random}}{1 - Accuracy_{random}} \tag{3}$$

Where $Accuracy_{model} = \frac{\text{Number of correctly classified observations}}{\text{Total number of observations}}$

And $Accuracy_{random} = \frac{\sum_{i=1}^{j}(Observed\ Outcomes)_i * (Predicted\ Outcomes)_i}{(\text{Total number of observations})^2}$

5.6.2 Precision

Precision for a specific class of the response variable is the ratio of the correctly classified events in a class to the total number of events classified in that class. Precision is obtained using **Equation 4**.

$$Precision = \frac{Number\ of\ correctly\ classified\ events\ in\ a\ class}{Total\ events\ (Correct\ and\ Incorrect)\ classified\ in\ the\ class} \tag{4}$$

5.6.3 Recall

Recall for a specific class of the response variable is the ratio of the correctly classified events in a class (True positives) to the sum of the correctly and incorrectly classified events related to the target class (True positives + False negatives). Recall is obtained from **Equation 5**.

$$Recall = \frac{Number\ of\ correctly\ classified\ events\ in\ a\ class\ (True\ positives)}{True\ positives + False\ negatives} \tag{5}$$

5.6.4 F1 score

The F1 score considers precision and recall in the model's performance evaluation and provides a balanced overall model performance measure. The higher the F1 score of a class, the better the predictive performance of the model specific to that class. The F1 score is calculated using **Equation 6**.

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{6}$$

6. Results and Discussion

6.1 Results of Segmented Model (Daytime vs Nighttime Pedestrian Crashes)

The dataset was segmented into pedestrian crashes occurring during the day and night, and separate models were estimated for the daytime and nighttime pedestrian crashes to highlight the key differences or contrasts in the association of key risk factors influencing pedestrian crash injury severity during the day and night. The model results are presented in Table 6.

Table 6. Results of Segmented Ordered Logit Models (Daytime vs Nighttime Crashes)

Variables	**Overall Model**		**Daytime Model**		**Nighttime Model**	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Transportation Access Indicator	1.46	2.54	1.23	2.44	1.75	3.24
Economy Indicator	0.63	3.62	0.28	2.62	0.84	2.68
Multilane Road	0.34	5.99	0.26	3.23	0.39	4.73
Crash at Intersection	0.18	3.20	0.08	1.54	0.26	3.56
Driver Alcohol Use	0.99	6.06	0.95	3.10	1.29	4.12
Pedestrian Alcohol Use	0.94	11.93	0.78	4.15	1.31	7.99
Both Driver and Pedestrian Alcohol Use	1.96	7.33	1.81	2.16	2.58	5.66
Hit and Run Crash	0.56	2.17	0.34	1.47	0.87	4.30
Pedestrian Age above 60	0.34	4.52	0.46	4.75	0.27	2.25
Rural Locality	0.72	8.79	0.79	5.47	0.58	5.61
Speed limit 60-75 mph	1.08	6.67	0.63	2.03	1.15	5.99
Pedestrian in Travel Lane (Crossing Violation)	0.68	11.28	0.44	5.31	0.83	9.07
Male Driver	0.12	2.54	0.085	2.12	0.15	2.84
SUV	0.13	2.00	0.08	2.29	0.22	2.31
Heavy Vehicle (Truck, Trailer)	0.14	2.35	0.06	2.56	0.18	2.80
Foggy Weather	0.27	2.18	0.06	1.42	0.48	2.04
Thresholds						
μ_1	-2.31		-2.09		-2.04	
μ_2	0.40		0.71		0.54	
μ_3	2.38		2.99		2.34	
μ_4	3.45		4.35		3.35	
Summary Statistics						
N	4765		2436		2329	
LL at null	-6408.99		-2941.24		-3307.97	
LL at convergence	-4761.43		-2345.68		-2305.37	
McFadden's pseudo R^2	0.26		0.20		0.30	
AIC	9562.86		4731.36		4650.74	
BIC	9692.24		4847.32		4765.80	
			Segmentation Likelihood Ratio Test			
Computed LR	220.76					
Df	20					

Critical LR	37.566
Conclusion	Computed LR > Critical LR. Segmentation is justified.

Key differences in the association of the correlates of pedestrian crash injury severity during day and nighttime can be observed in the results presented in Table 6. The transportation access indicator was found significant in both models but with a stronger association in nighttime crashes (coefficient: 1.75, t-stat: 3.24) compared to daytime crashes (coefficient: 1.23, t-stat: 2.44). This suggests higher odds of severe pedestrian crash injury severity at night in the transportation access-burdened census tracts compared to crashes occurring in these census tracts in the daytime. The economy indicator was also found statistically significant in both models, with a higher association in the nighttime model (coefficient = 0.62) compared to that in the daytime model (coefficient = 0.28), indicating higher odds of pedestrian crash injury severity at night in the economy-burdened census tracts compared to daytime. The indicator of pedestrian crashes on multilane roads was significant in both models but was more strongly associated with nighttime crashes (coefficient: 0.39, t-stat: 4.73). Crashes at intersections were found to be significant only in the nighttime model and did not exhibit statistical significance in the daytime model. This indicates that pedestrian crashes at intersections are more likely to result in severe pedestrian crash injuries at nighttime compared to daytime. Driver and pedestrian alcohol use variables were significant in both models, with notably stronger associations in nighttime crashes. For instance, pedestrian alcohol use has a coefficient of 1.31 at night compared to 0.78 during the day, which indicates a higher likelihood of severe pedestrian injury in crashes at night involving alcohol-impaired pedestrians compared to crashes in the daytime. The association of pedestrian crashes in rural localities is more pronounced at night (coefficient: 0.58) than daytime (coefficient: 0.79). Similarly, pedestrian crashes on roadways with posted speed limits of 60–75 mph, involving pedestrians in travel lanes (crossing violations), male drivers, crashes with SUVs and heavier vehicles show significant effects in both models but have a stronger impact on pedestrian crash injuries in nighttime crashes. Pedestrian crashes in foggy weather were also found to be statistically significant only in the nighttime model, while they were insignificant in the daytime model, indicating a higher likelihood of resulting in severe pedestrian crash injuries at nighttime compared to daytime.

Referring to the goodness of fit measures of the models, the models show a better statistical fit with McFadden's pseudo R-squared values of 0.26, 0.20, and 0.30 for the overall model, daytime model, and nighttime model, respectively. The likelihood ratio test was performed to compare the segmented models against the overall model to determine if the segmentation was statistically justified. The null hypothesis in this test is that the overall model provides enough information, suggesting no statistically significant need to estimate segmented models. The chi-square distributed likelihood ratio is calculated as:

$$Computed\ LR = -2[LL_{overall} - LL_{Day} - LL_{Night}] \tag{7}$$

where:

$LL_{overall}$: Log-likelihood at convergence of the overall model

LL_{Day}: Log-likelihood at convergence of the Daytime model

LL_{Night}: Log-likelihood at convergence of the Nighttime model

The computed likelihood ratio is chi-square distributed with degrees of freedom calculated as:

$$Df = P_1 + P_2 - P \tag{8}$$

where, P_1: Number of parameters in the Daytime model

P_2: Number of parameters in the Nighttime model

P: Number of parameters in the overall model

Results for the likelihood ratio test for the segmented models reveal that the computed likelihood ratio is 220.76 with 20 degrees of freedom, significantly higher than the critical chi-square value of 37.566 at a 0.01 significance level. This result leads to rejecting the null hypothesis, indicating that the segmentation is statistically justified.

While this segmented analysis provides insights into differences between daytime and nighttime pedestrian crashes, the primary focus of the study is on nighttime pedestrian crash injury severity. Therefore, the subsequent analysis emphasizes the results derived from the nighttime crash dataset.

6.2 Result of Ordered Logit Model for Nighttime Dataset (Training Data)

The nighttime pedestrian crash dataset was split into three sets in a 70:20:10 proportion. The training dataset, consisting of around 70% of the data (Ntrain = 1664), was used to train the individual base learners, while approximately 20% of the data (N_{val} = 438) was used as a validation dataset to obtain predictions from the trained models. About 10% of the data (N = 227) were used as a test (holdout) sample to evaluate the models' predictive performance. The ordered logit model was selected as the first base learner due to the ordinal nature of the response variable (Pedestrian injury severity). Table 7 presents the results of the ordered logit model. The explanatory variables included in the final model were found statistically significant at a 95% and higher confidence level. The model shows novel associations of indicators with the severity of nighttime pedestrian crash injuries. Results from the model indicate that the Economy and Transportation Access indicators were positively associated with higher levels of pedestrian injury severity. The marginal effects of these indicators suggest that pedestrian crashes at night in the economy- and transportation-burdened census tracts are associated with respective increases in the probability of fatal injury to pedestrians by 0.0057 and 0.0063. While these increases may seem modest, they reflect a meaningful escalation in risk when considered against the baseline chance of fatal pedestrian injuries obtained from the data (i.e., 349/2329 = 14.98%). An increase of 0.0057 or 0.0063 units represents an additional 0.57% or 0.63% increase in the chance of fatal pedestrian injuries, which is meaningful when considering the cumulative impact across all burdened regions or multiple crash events.

This result is intuitive as economic hardship is primarily reflected through variables such as low-income levels, higher housing costs, and higher unemployment ratios (Table 1). Economy-burdened census tracts usually lack pedestrian-specific infrastructure, including dedicated pedestrian crossings, sidewalks, etc. This can lead to unsafe pedestrian actions and result in severe injury crashes. Similarly, transportation-burdened communities are characterized by several factors, including a higher reliance on public transit for daily commutes, limited access to private vehicles, and higher transportation cost burdens (Table 1). These factors may influence exposure to unsafe road environments, potentially increasing the risk of severe injury in pedestrian crashes. Pedestrian crashes involving impairment of only the pedestrian due to alcohol/drug use, only the driver, and both the driver and pedestrian were found highly likely to result in incapacitating and fatal injuries to the pedestrians. The marginal effects of these variables indicate that pedestrian crashes at night involving alcohol-impaired pedestrians are associated with an increase of 0.0845 units in the probability of fatal pedestrian injury. Crashes involving alcohol-impaired drivers are associated with a rise of 0.0905 units in the likelihood of fatal pedestrian injury. The probability of fatal pedestrian injury increases by an even higher amount, i.e., 0.1929 units in pedestrian crashes when both driver and pedestrian are alcohol-impaired. Similarly, the marginal effects show the associations of other explanatory variables with injury severity. Furthermore, positive and statistically significant associations are observed between pedestrian injury severity and pedestrian crashes at night occurring on roads without proper lighting, rural areas, multilane roads (more than two-lane roads), involving male drivers, and elderly pedestrians (above 60 years of age). Previous studies also indicate more severe pedestrian crashes due to the abovementioned factors (Aziz et al., 2013; Haleem et al., 2015; Pour-Rouholamin & Zhou, 2016).

Referring to the model goodness of fit statistics, the model yields a log-likelihood at convergence value of -1823.7 and a McFadden's Pseudo R-squared value of 0.23, which indicates a good model fit. According to the literature, McFadden's pseudo-R-squared values greater than 0.10 indicate meaningful improvement in model fit, while values above 0.20 indicate a good fit (Islam & Jones, 2014; Ulfarsson et al., 2010). The model is overall statistically significant (Prob > $\chi^2(13) = 0.0000$), and the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) values of the model are 3681.54 and 3773.63.

Table 7. Results of Ordered Logit Model for Training Dataset

Variable	Coeff.	t-stat	p-value	Marginal Effects				
				No Injury	**Possible Injury**	**Minor Injury**	**Severe Injury**	**Fatal Injury**
Economy Indicator (1/0)	0.436	4.51	0.000	-0.0017	-0.0077	0.0001	0.0035	0.0057
Transportation Access Indicator (1/0)	0.801	3.43	0.001	-0.0019	-0.0084	0.0002	0.0038	0.0063
Dark Roads without Lights (1/0)	0.618	5.85	0.000	-0.0219	-0.0972	0.0019	0.0444	0.0726
Multilane Roads (1/0)	0.518	5.23	0.000	-0.0183	-0.0814	0.0017	0.0372	0.0608
Driver Alcohol impairment (1/0)	0.770	3.27	0.001	-0.0273	-0.1211	0.0025	0.0553	0.0905
Pedestrian Alcohol impairment (1/0)	0.718	6.77	0.000	-0.0254	-0.1130	0.0023	0.0516	0.0845
Driver and Pedestrian Both Alcohol-impaired (1/0)	1.640	5.12	0.000	-0.0581	-0.2580	0.0053	0.1179	0.1929
Hit & Run Crash (1/0)	1.015	4.33	0.000	-0.0360	-0.1597	0.0033	0.0730	0.1194
Pedestrian Age above 60 (1/0)	0.407	2.83	0.005	-0.0144	-0.0640	0.0013	0.0293	0.0478
Rural Locality (1/0)	0.293	2.29	0.022	-0.0104	-0.0461	0.0009	0.0210	0.0344
Speed limit 60-75 mph (1/0)	0.994	4.55	0.000	-0.0352	-0.1565	0.0032	0.0715	0.1170
Pedestrian Crossing Violations (1/0)	0.667	6.16	0.000	-0.0236	-0.1049	0.0021	0.0479	0.0785
Male Driver (1/0)	0.185	1.97	0.049	-0.0065	-0.0291	0.0005	0.0133	0.0218
Threshold Values								
μ_1				-1.783				
μ_2				0.858				
μ_3				2.618				
μ_4				3.665				
Model Statistics								
Number of Observations				1664				
Log-likelihood at null				-2372.37				
Log-likelihood at convergence				-1823.77				
McFadden's Pseudo R-squared				0.23				
LR χ^2 (13)				1097.20				
Prob > χ^2				0.000				
AIC				3681.54				
BIC				3773.63				

Note: (1/0) means indicator variables containing only 1 and 0 values.

6.3 Prediction-based analysis

This part of the analysis involves obtaining predictions from the trained ordered logit model and three ML models: Decision Tree, Gradient Boosting, and Random Forest for the validation and test datasets. Prediction results for each of the ML models are provided below.

6.3.1 Decision tree

Decision Tree (DT) was the second base learner for classification. The algorithm's operation is based on recursively splitting the dataset into subsets based on the most significant predictor variables, which minimizes impurity at each split. This property makes DT particularly suitable for our analysis, as it effectively captures non-linear relationships and interactions among predictors that influence pedestrian injury severity. Initially, a single decision tree was estimated for the training data. A pruning process was undertaken to improve its generalizability and prevent overfitting—a common issue with complex trees that tend to memorize training data. Pruning is a critical step in decision tree modeling that involves removing sections of the tree that contribute little to the overall prediction accuracy. It aims to simplify the model while retaining its ability to classify observations effectively. The tree was pruned to avoid overfitting and increase its generalizing ability. A decision tree with 20 terminal nodes was found to be the optimal tree with minimum pruning error corresponding to a cost-complexity parameter (cp) value of 0.0047, shown in Figure 4. The cp parameter balances the trade-off between the complexity of the tree and its performance as it penalizes overly complex trees by adding a cost to each additional split, ensuring that only meaningful splits are retained.

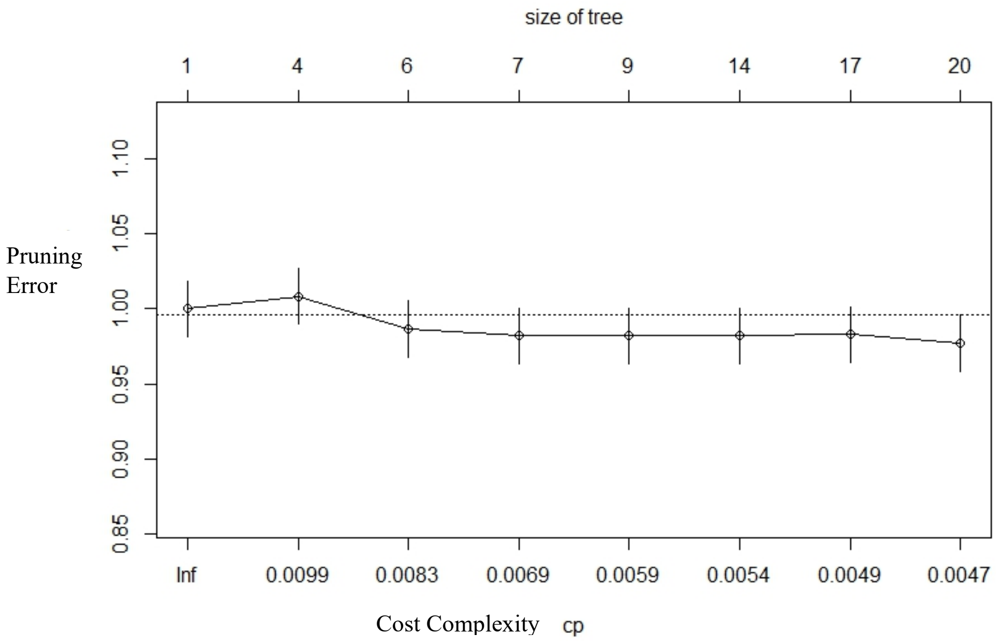


Fig. 4. Selection of Regularization Parameter for Decision Tree Model

Figure 5 presents the classification of pedestrian injury severity by the pruned decision tree. The decision tree contains 20 terminal nodes and uses several key predictor variables such as Local Street, Pedestrian in Travel Lane ("PedinTravelLane"), Intoxicated Drivers ("DrivAlcohol"), Both drivers and pedestrians intoxicated ("BothAlcohol"), Health-Burdened Community ("Health_Disadv"), etc., to classify the pedestrian crash observations in one of the five injury severity classes. These variables reflect the multi-dimensional nature of factors influencing pedestrian crash injury severity, highlighting both behavioral (e.g., intoxication) and environmental (e.g., roadway type) contributors. The number of observations classified in a particular injury severity class is mentioned at each terminal node.

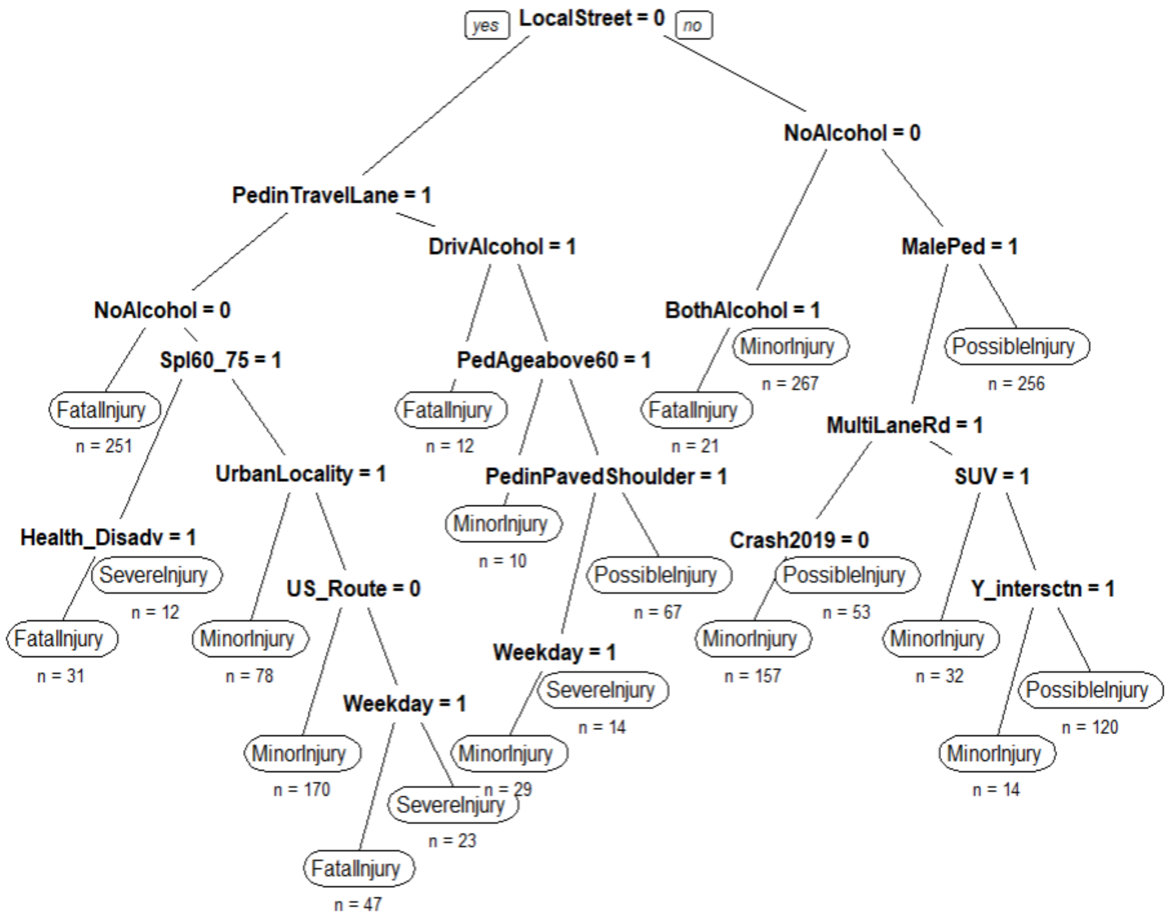


Fig. 5. Pruned Decision Tree Model for Injury Severity

6.3.2 Gradient boosting

This method performs classification using predictions from multiple decision trees, producing more reliable and stable results than a single decision tree. The hyperparameters of the model include shrinkage rate (also called learning rate), number of trees, minimum number of observations in the trees' terminal nodes, and interaction depth. The model hyperparameters were tuned using an extended grid search and 10-fold cross-validation. The 10-fold cross-validation process involves dividing the available data into ten equal subsets. Nine of these subsets are used for training the model, while one subset is retained for testing to evaluate the model's predictive accuracy. This procedure is repeated ten times, with each subset serving as the testing data once. The final estimation is obtained by averaging the results from all ten iterations. A range of values for the shrinkage rate (0.005, 0.01, 0.05, 0.1, 0.5), minimum number of observations in the terminal nodes (3,5,7), and interaction depth (1,3,5,7,10) were considered in the optimization grid search among which the optimal shrinkage rate, the number of trees corresponding to the optimal shrinkage rate, minimum observations at the terminal node, and the interaction depth were found to be 0.005, 1064, 7, and 5, respectively, resulting in a minimum cross-validation error of 1.24. Figure 6 presents the results of the grid search for the first five combinations of the model hyperparameters ranked according to increasing cross-validation error. The final model was estimated using the optimal hyperparameters obtained from the grid search.

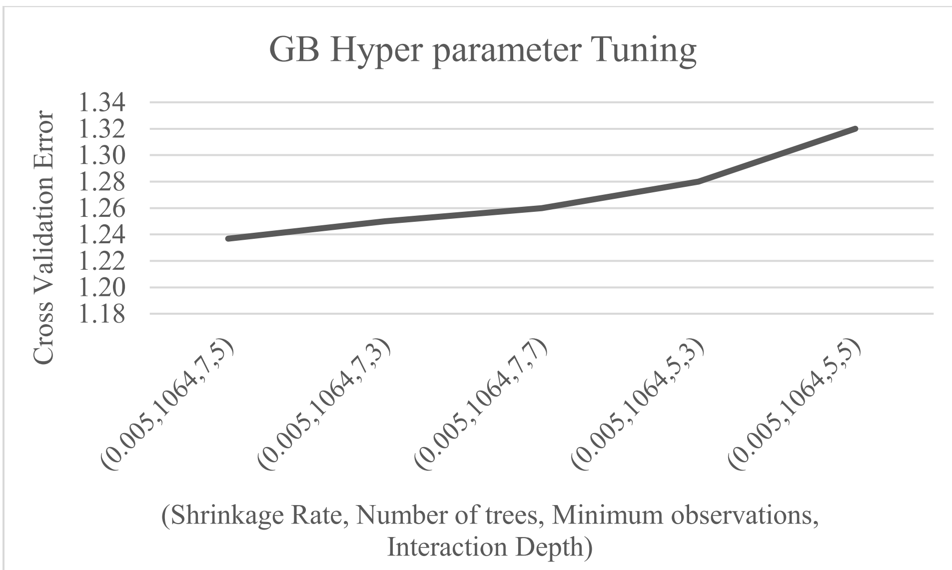


Fig. 6. Results of Tuning GB Model Hyperparameters

6.3.3 Random Forest

Random Forest was used as the final base learner in this study. Like the GB model, hyperparameters were tuned using an extended grid search optimization and 10-fold cross-validation. The model's hyperparameters include the number of trees, the number of variables considered for splitting at each node, the minimum node size, and the sample fraction. A range of values for the number of trees (600, 900, 1200, 1500), number of variables considered for splitting at each node (5, 10, 20, 40), minimum node size (1, 5, 10, 15, 20) and fraction of the sample used (0.6, 0.7, 0.8) were considered in the optimization grid search among which the optimal values for the number of trees, number of variables for splitting each node, minimum node size, and the sample fraction were found to be 900, 5, 5, and 0.7, respectively. The combination resulted in a minimum cross-validation error of 0.775. Figure 7 presents the results of the grid search for the first ten combinations of the model hyperparameters ranked according to increasing cross-validation error.

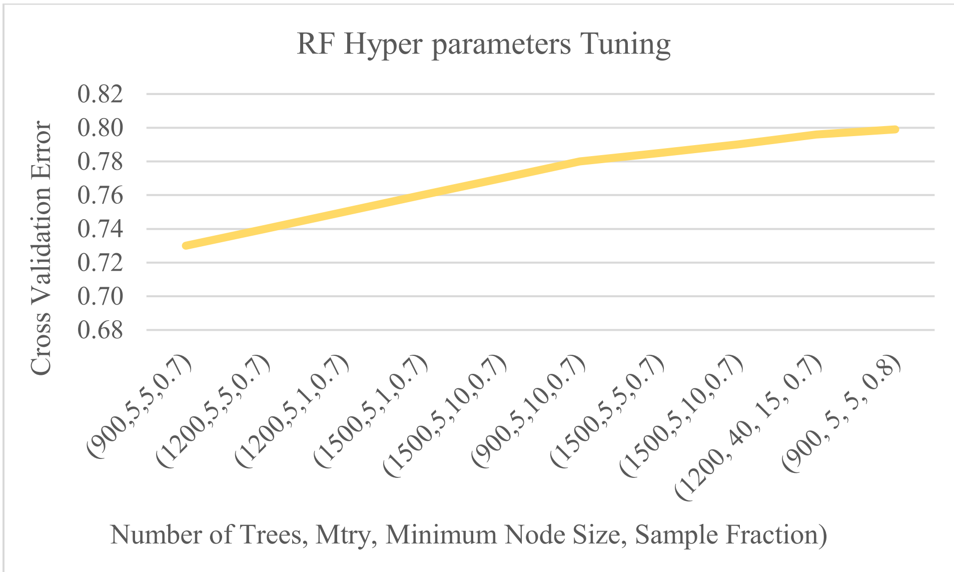


Fig. 7. Results of Tuning RF Model Hyper Parameters

Figures 8 (a, b, c, d) present the variable importance plots of the base learners. The standardized coefficients of the statistically significant variables in the ordered logit model were estimated in STATA software to rank the variable importance. The standardized coefficients indicate the impact on the response variable with a change of 1 standard deviation in the value of the predictor variable and hence provide a basis to assess the importance of each predictor in the model. The importance rankings of predictor variables based on their absolute standardized coefficients are presented in Table 8.

Table 8. Variable Importance based on Standardized Coefficients in Ordered Logit Model

Variables	Coefficients	Standardized Coefficients	Rank
Economy Indicator	0.436	0.113	12
Transportation Access Indicator	0.801	0.367	1
Dark Roads without Lights	0.618	0.309	3
Multilane Roads	0.518	0.259	5
Driver Alcohol impairment	0.770	0.155	9
Pedestrian Alcohol Impairment	0.718	0.321	2
Driver and Pedestrian Both Alcohol-impaired	1.640	0.239	6
Hit & Run Crash	1.015	0.198	8
Pedestrian Age above 60	0.407	0.128	10
Rural Locality	0.293	0.121	11
Speed limit 60-75 mph	0.994	0.222	7
Pedestrian Crossing Violations	0.667	0.290	4
Male Driver	0.185	0.080	13

Figure 8(a) indicates that the "Transportation Access" indicator was found the most important variable in the ordered logit model with the highest value of standardized coefficient (i.e., 0.37) followed by "Pedestrian Alcohol Impairment" and "Dark Roads without Lights" indicator. Figure 8(b) indicates that the "Local Street" indicator was the most important variable in the decision tree followed by the "Urban Locality" indicator. The "Environmental Burden" indicator was important among the burden indicators in the model. Referring to Figure 8(c), most of the important variables in the GB model were statistically significant in the Ordered Logit Model and were also used to grow the decision tree. Hence, the ML models complement the results obtained from the statistical model.

Referring to the variable importance plot for the random forest model (Figure 8(d)), variables indicating different communities (Proficiency and Environmental Indicators) and pedestrian racial attributes (Black and White pedestrians) were found to be more important in the RF model than the other models. These results can give more valuable insights into the distribution of pedestrian crashes across various racial groups in these communities.

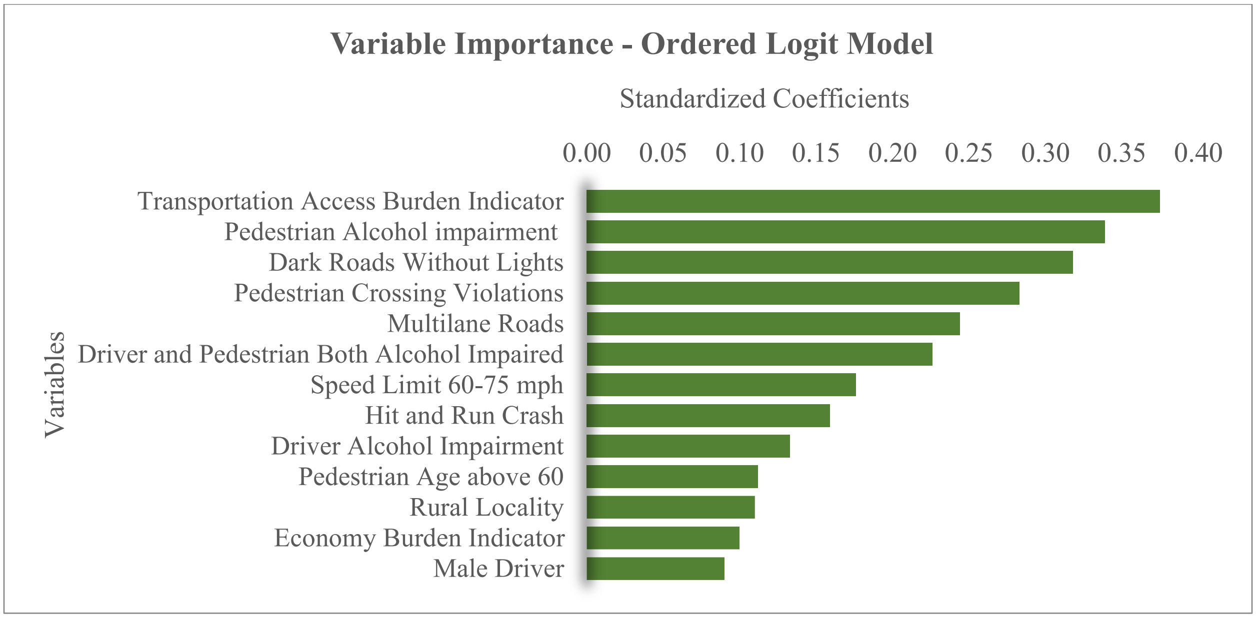


Fig. 8. (a) Variable Importance Plot for Ordered Logit Model

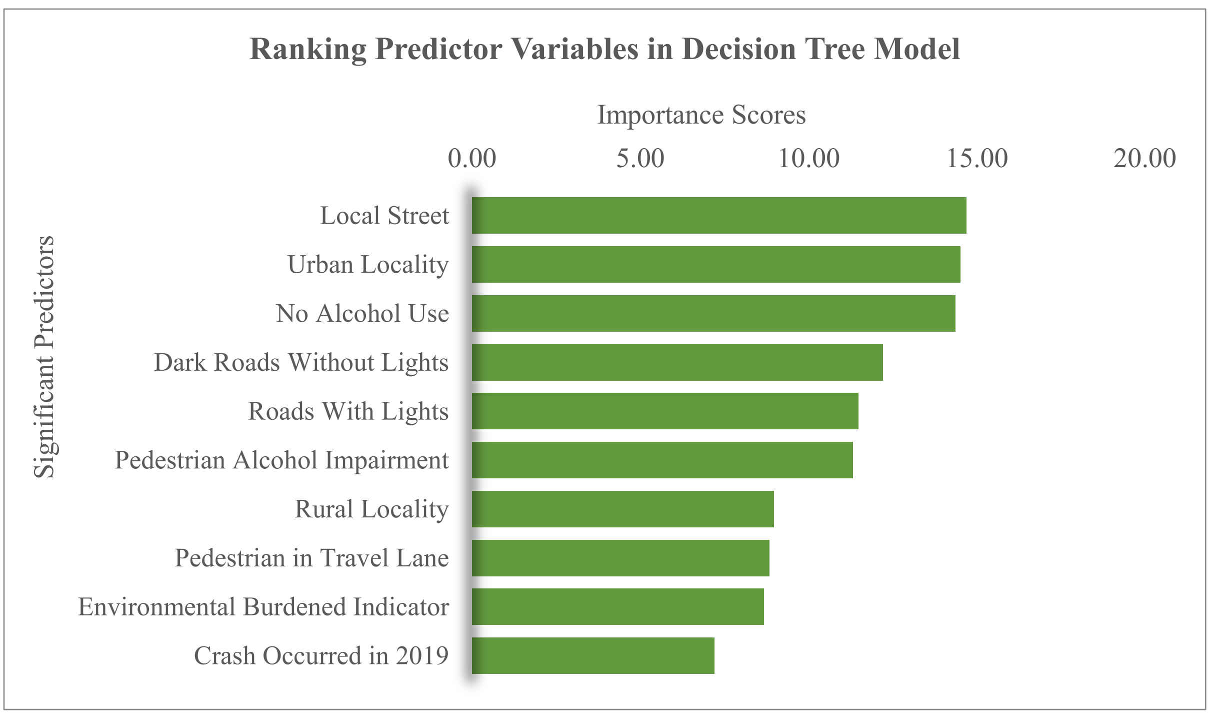


Fig. 8. (b) Variable Importance Plot for Decision Tree

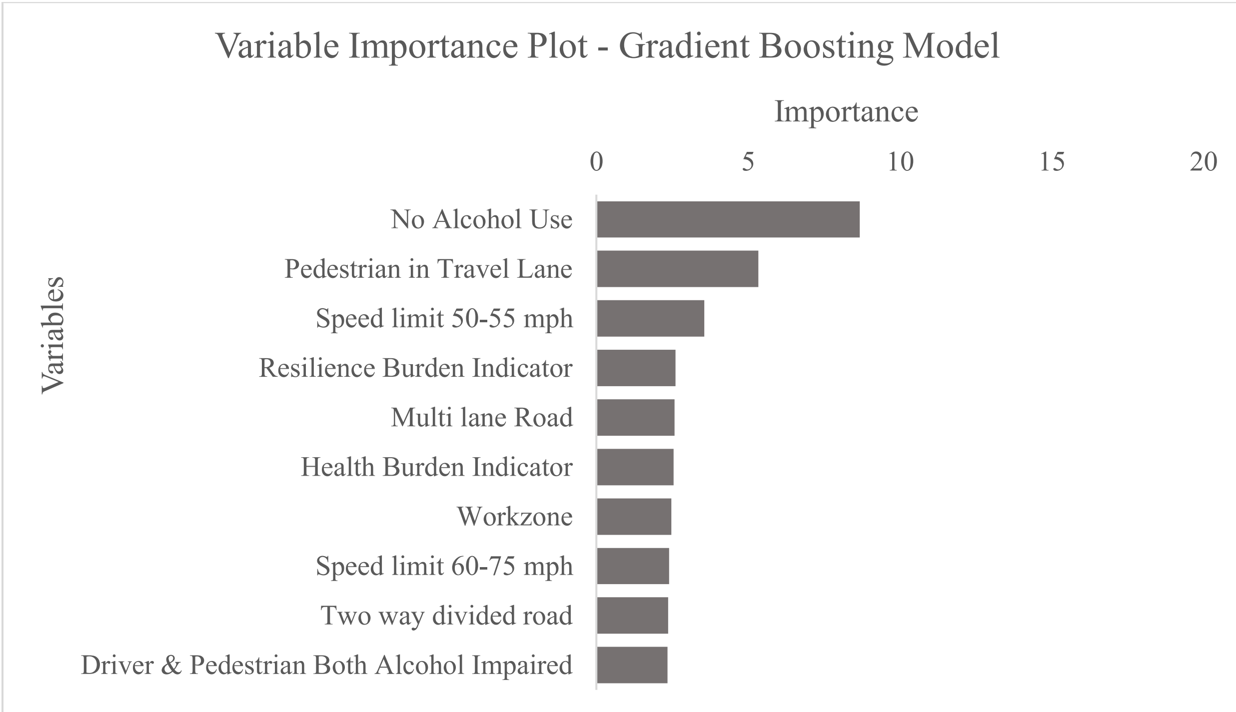


Fig. 8. (c) Variable Importance Plot for Gradient Boosting Model

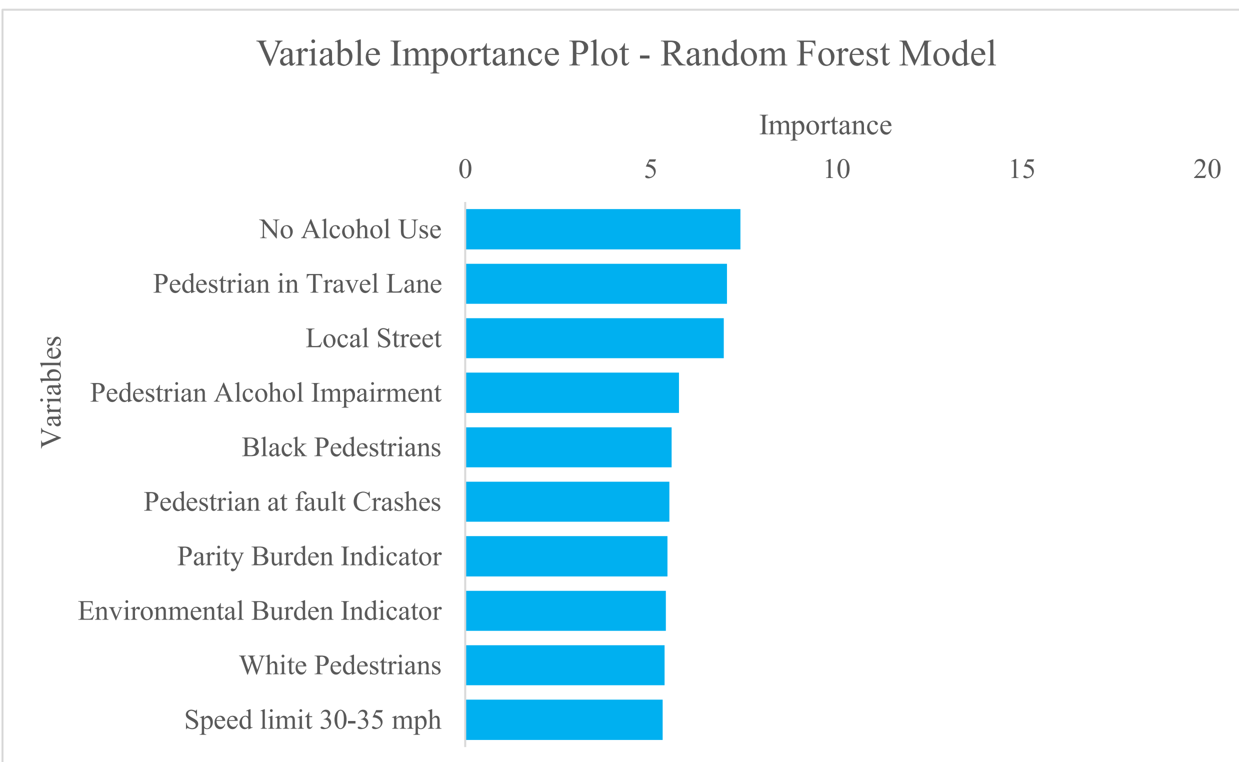


Fig. 8. (d) Variable Importance Plot for Random Forest Model

6.4 Stacking individual base learners and comparison of models' performance

The validation dataset (N_{val} = 438) was provided as input to the trained base learners, and predictions on pedestrian injury severity were obtained from each base learner for the validation dataset. Subsequently, this set of predictions was used as input for the meta-learner in the second layer of the stacking methodology. Furthermore, the predictive performance of the trained base learners was assessed using the test dataset (N_{test} = 227). Confusion matrices, widely used to evaluate the predictive performance of the ML models, were obtained, and performance measures discussed in the previous section were calculated to determine the trained models' predictive ability. Results from the confusion matrices indicate that the ML models' prediction accuracy was higher than the frequentist model, which conforms to the usual norm. The RF model predicted pedestrian injury severity with the highest accuracy (73.56%) among the base learners. The confusion matrices for the base learner models are presented in Tables 9-12.

Table 9. Confusion Matrix for Test Dataset (N_{test} = 227) for Ordered Logit Model

Observed Outcomes	**Predicted Outcomes**					**Total**
	Fatal Injury	Severe Injury	Minor Injury	Possible Injury	No Injury	
Fatal Injury	15	10	7	2	1	**35**
Severe Injury	10	11	8	3	0	**32**
Minor Injury	2	6	70	14	0	**92**
Possible Injury	0	1	12	45	2	**60**
No Injury	0	0	1	4	3	**8**
Total	**27**	**28**	**98**	**68**	**6**	
Performance Measures						
Accuracy			63.44%			
$Accuracy_{random}$			29.08%			
Kappa			0.48			
Precision	0.555	0.393	0.714	0.662	0.500	
Recall	0.428	0.344	0.761	0.750	0.375	
F1 Score	0.484	0.367	0.737	0.703	0.429	

Table 10. Confusion Matrix for Test Dataset (N_{test} = 227) for Decision Tree

Observed Outcomes	**Predicted Outcomes**					**Total**
	Fatal Injury	Severe Injury	Minor Injury	Possible Injury	No Injury	
Fatal Injury	16	10	5	3	1	**35**
Severe Injury	10	14	4	3	1	**32**
Minor Injury	7	18	64	2	1	**92**
Possible Injury	0	1	9	47	3	**60**
No Injury	0	0	1	1	6	**8**
Total	**33**	**43**	**83**	**56**	**12**	
Performance Measures						
Accuracy			64.75%			
$Accuracy_{random}$			26.43%			
Kappa			0.52			
Precision	0.485	0.326	0.771	0.839	0.500	
Recall	0.457	0.438	0.696	0.783	0.750	
F1 Score	0.471	0.373	0.731	0.810	0.600	

Table 11. Confusion Matrix for Test Dataset (N_{test} = 227) for GB Model

Gradient Boosting Model						
Observed Outcomes	**Predicted Outcomes**					**Total**
	Fatal Injury	Severe Injury	Minor Injury	Possible Injury	No Injury	
Fatal Injury	24	4	3	2	2	**35**
Severe Injury	6	20	3	2	1	**32**
Minor Injury	12	5	73	2	0	**92**
Possible Injury	3	4	12	38	3	**60**
No Injury	0	0	1	1	6	**8**
Total	**45**	**33**	**92**	**45**	**12**	
Performance Measures						
Accuracy			70.9%			
$Accuracy_{random}$			26.96%			
Kappa			0.60			
Precision	0.533	0.606	0.794	0.844	0.500	
Recall	0.686	0.625	0.794	0.633	0.750	
F1 Score	0.599	0.615	0.794	0.724	0.600	

Table 12. Confusion Matrix for Test Dataset (N_{test} = 227) for RF Model

Observed Outcomes	**Predicted Outcomes**					**Total**
	Fatal Injury	Severe Injury	Minor Injury	Possible Injury	No Injury	
Fatal Injury	22	5	4	3	1	**35**
Severe Injury	6	23	2	1	0	**32**
Minor Injury	3	12	75	2	0	**92**
Possible Injury	0	2	11	42	5	**60**
No Injury	0	0	1	2	5	**8**
Total	**31**	**42**	**93**	**50**	**11**	
Performance Measures						
Accuracy			73.56%			
$Accuracy_{random}$			27.31%			
Kappa			0.64			
Precision	0.710	0.548	0.806	0.840	0.454	
Recall	0.629	0.719	0.815	0.700	0.625	
F1 Score	0.667	0.622	0.811	0.764	0.526	

The ML models were used as the meta-learner in the stacked ensemble model due to their superior predictive performance than the frequentist model for the test dataset in the first stage. The RF model yielded the highest prediction accuracy when used as a meta-learner in the stacked ensemble model (78.85%) compared to the GB (76.21%) and DT (74.44%) models (Table 13). The confusion matrix for only the stacked ensemble model with RF as the meta-learner is provided in Table 13 for brevity. This result yielded the highest accuracy and kappa value. Table 14 compares the predictive performance of the individual base learners and the stacked models with all possible meta-learners. The kappa value of the stacked model with RF as the meta-learner was the highest (0.71) compared to stacked models with GB (0.69) and DT (0.67) as meta-learners (Table 14). Note that the predictive performance of the stacked ensemble model was superior to that of the individual base learners for each of the ML models used as the meta learner. This result provides substantial evidence of this method's superiority in predictive capabilities over individual ML techniques. The RF model, used as a meta learner, was tuned to 100 trees, two variables used for splitting each node, minimum node size of 3, and sample fraction of 0.8 using extended grid search and 10-fold cross-validation, resulting in a minimum cross-validation error of 0.46. Figure 9 represents the selection of optimal hyperparameters for the RF model in the stacked ensemble model. The variable importance plot for the stacked model, shown in Figure 10, indicates the predictions obtained from the RF base learner for the validation dataset as most important, followed by the predictions obtained from the GB model, while the predictions from the Ordered Logit model were ranked the least important by the model. This result also provides evidence in support of the ML models' superior predictive performance compared to the frequentist models. The stacked models with RF, GB, and DT as meta learners resulted in a percentage increase in the kappa statistic of 10.94%, 7.81%, and 4.69%, respectively, from the highest kappa statistic of RF base learner (0.64) considered as the base.

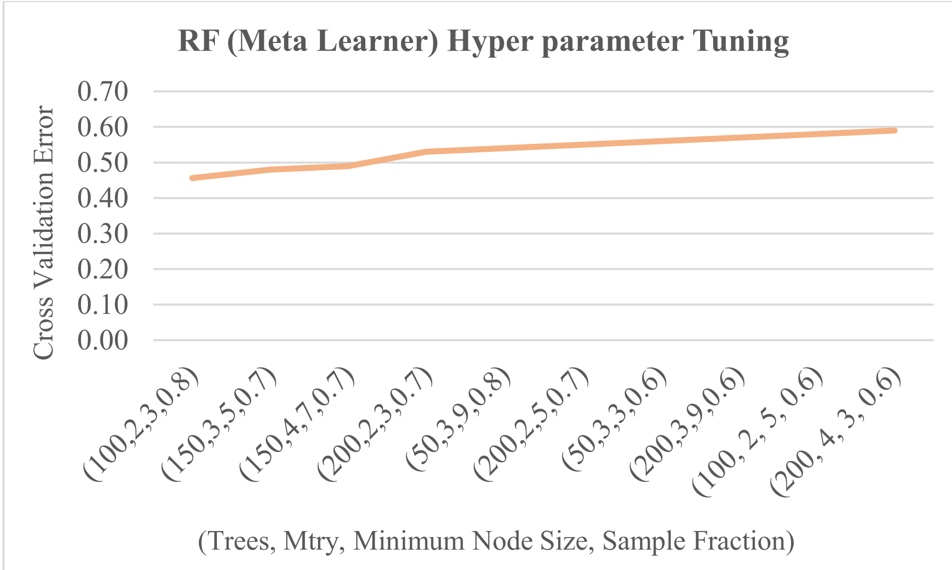


Fig. 9. Selection of Optimal Hyperparameters for RF as Meta Learner

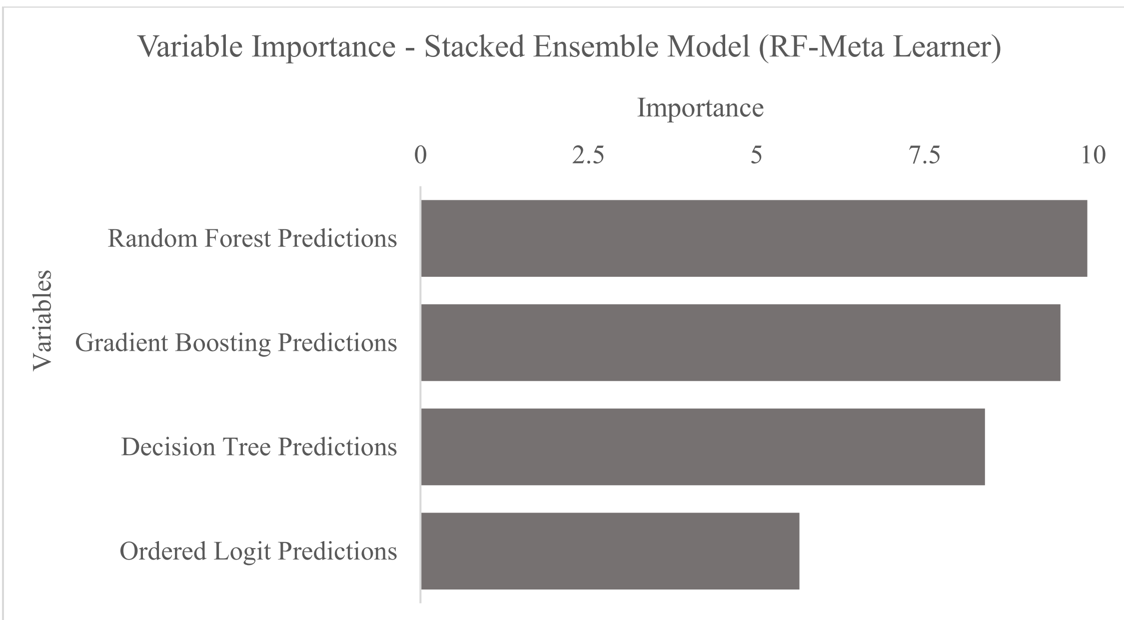


Fig. 10. Variable Importance Plot for Stacked Model with RF as Meta Learner

Table 13. Confusion Matrix for Test Dataset for Stacked Ensemble Model (RF as Meta Learner)

Observed Outcomes	**Predicted Outcomes**				
	Fatal Injury	Severe Injury	Minor Injury	Possible Injury	No Injury
Fatal Injury (35)	28	5	2	0	0
Severe Injury (32)	3	26	2	1	0
Minor Injury (92)	4	11	76	1	0
Possible Injury (60)	0	3	7	44	6
No Injury (8)	0	0	1	2	5
Total	**35**	**45**	**88**	**48**	**11**
Performance Measures					
Accuracy			78.85%		
$Accuracy_{random}$			26.64%		
Kappa			0.71		
Precision	0.800	0.578	0.864	0.917	0.454
Recall	0.800	0.813	0.826	0.733	0.625
F1 Score	0.800	0.675	0.844	0.815	0.526

Note: Values in parentheses next to the injury class indicate the observed frequency of that class.

Table 14. Comparison of Out-of-Sample Prediction Performance of all Models

Models	**Status**	**Kappa**	**% Change in Kappa from the Base**
Random Forest	(Meta learner)	0.71	10.94%
Gradient Boosting	(Meta learner)	0.69	7.81%
Decision Tree	(Meta learner)	0.67	4.69%
Random Forest	(Base learner)	0.64	**(Base)**
Gradient Boosting	(Base learner)	0.60	-6.25%
Decision Tree	(Base learner)	0.52	-18.75%
Ordered Logit	(Base learner)	0.48	-25%

Note: Percentage Change in Kappa = $[(Kappa – Kappa_{Base})/Kappa_{Base}]*100$

7. Limitations

The study findings are derived from pedestrian crashes only in North Carolina. The magnitude of the key correlates of nighttime pedestrian crashes found in the study could vary across different geographical locations due to the varying traffic characteristics, socioeconomic characteristics, and spatial contexts. However, the direction of association of most factors with pedestrian injury severity is generally assumed to be similar to that found in this study. With the availability of similar comprehensive data for other parts of the US, conducting a similar study would be more appropriate, which may lead to more generalizable results. Furthermore, the study methodology (stacking) does not directly provide variable importance of the study's explanatory variables in the second stage, as it aggregates only the predictions obtained from the base learners.

8. Conclusions

This study used a unique pedestrian crash database and applied robust inferential and prediction-based models to analyze nighttime pedestrian crash injury severity in different communities. The results offer significant insights into pedestrian safety in overlooked population segments. By integrating an inferential framework, a community-specific focus, and a data-driven methodology, the study provides a unique approach with the potential to improve nighttime pedestrian safety in specific areas. The study sought to answer the research question: *What are the infrastructural, financial, and policy-related differences in communities that affect nighttime pedestrian crash severity?* The following sub-sections provide the key study findings that highlight the infrastructural, financial, and policy-related differences in communities that influence nighttime pedestrian crash injury severity.

8.1 *Infrastructural Differences*

1. Inadequate lighting on roadways in some communities was identified as a critical factor contributing to nighttime pedestrian crashes. Providing high-intensity streetlights on roadways with substantial pedestrian movements can improve pedestrian visibility and enhance pedestrian safety.

2. The transportation access indicator was a significant indicator of severe pedestrian crash injury. Poor walkability in communities, linked to transportation burdens, such as limited access to safe pedestrian infrastructure, exacerbates pedestrian risks at night.

8.2 *Financial Differences*

1. The Economy Burden Indicator, comprising sub-indicators like poverty levels, low household income, and high unemployment rates, highlights how economic hardship is associated with nighttime pedestrian safety. Limited financial resources in some communities restrict investments in pedestrian-friendly infrastructure and road safety measures.
2. The Transportation Access indicator demonstrates financial challenges through sub-indicators like lack of vehicle access and high dependency on public transit, pushing residents to walk in unsafe environments, especially at night.

8.3 *Policy-Related Differences*

1. The study underscores insufficient policy measures to address alcohol-impaired driving and walking, which was found to be a significant contributor to nighttime pedestrian crashes.

The study implemented a robust heterogeneous ensemble methodology (stacking) that leverages the inferential capabilities of statistical models and the predictive strength of machine learning (ML) techniques. Among the base learner models, Random Forest yielded the highest prediction accuracy and kappa statistic, further enhanced by the stacked ensemble model, yielding a 10.94% increase in the kappa statistic from the one yielded by the best-performing base learner. This combined approach enabled more accurate predictions of pedestrian crash injury severity, facilitating effective safety planning and infrastructure improvements. The study findings can assist safety planners and policymakers in implementing several targeted nighttime pedestrian safety improvements for certain communities, including but not limited to 1) Allocating resources to census tracts with higher pedestrian crash frequencies, 2) Provision of pedestrian-specific infrastructure, such as sidewalks, crosswalks, high-intensity roadway lights at poorly lit locations, 3) Enforcing stricter penalties for alcohol-impaired driving and walking on roadways, 4) Conducting road safety awareness campaigns for pedestrians and drivers, and 5) Ensuring faster emergency medical responses to mitigate crash injury severity.

8.4 *Practical Applications*

The study findings have critical planning and policy implications. The ensemble model developed in the study can aid state DOTs in anticipating future nighttime pedestrian crashes across different communities. This would enable transportation planners and engineers to make informed and balanced decisions regarding resource allocation. The model can also predict crash hotspots where targeted safety interventions could be implemented, such as installing lighting, providing driver and pedestrian warnings, issuing nighttime speed advisories, launching safety education programs, and improving emergency response services.

Funding

This study was partially supported by the USDOT-sponsored Center for Pedestrian and Bicycle Safety (https://www.pedbikesafety.org/) [grant number 69A3552348336] and the Collaborative Sciences Center for Road Safety, CSCRS, www.roadsafety.unc.edu [grant number 69A3551747113].

Acknowledgment

The authors sincerely thank Dr. Daniel Carter, Safety Planning Engineer at the North Carolina Department of Transportation, for sharing the pedestrian crash dataset.

Declaration of Generative AI and AI-assisted technologies in the writing process

While preparing this work, the authors used ChatGPT to enhance the readability of the study. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication’s content.

References

Adeel, M., Khattak, A., Mishra, S., & Thapa, D. (2024). Enhancing work zone crash severity analysis: The role of synthetic minority oversampling technique in balancing minority categories. *Accident Analysis & Prevention*, *208, 107794*. https://doi.org/10.1016/j.aap.2024.107794

Agresti, A. (2010). *Analysis of ordinal categorical data* (Vol. 656). John Wiley & Sons.

Ahmad, N., Wali, B., & Khattak, A. J. (2023). Heterogeneous ensemble learning for enhanced crash forecasts – A frequentist and machine learning based stacking framework. *Journal of Safety Research*, *84*, 418–434. https://doi.org/10.1016/j.jsr.2022.12.005

Aidoo, E. N., Amoh-Gyimah, R., & Ackaah, W. (2013). The effect of road and environmental characteristics on pedestrian hit-and-run accidents in Ghana. *Accident Analysis and Prevention*, *53*, 23–27. https://doi.org/10.1016/j.aap.2012.12.021

Aziz, H. M. A., Ukkusuri, S. V., & Hasan, S. (2013). Exploring the determinants of pedestrian-vehicle crash severity in New York City. *Accident Analysis and Prevention*, *50*, 1298–1309. https://doi.org/10.1016/j.aap.2012.09.034

Benson, A., Arnold, L., Tefft, B., & Horrey, W. (2018). Hit-and-Run Crashes: Prevalence, Contributing Factors and Countermeasures. *AAA Foundation for Traffic Safety, Washington, D.C.* .

Bhatt, S., Cameron, E., Flaxman, S. R., Weiss, D. J., Smith, D. L., & Gething, P. W. (2017). Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. *Journal of the Royal Society Interface*, *14*(134). https://doi.org/10.1098/rsif.2017.0520

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32.

Chali, Y., Hasan, S. A., & Mojahid, M. (2014). Complex Question Answering: Homogeneous or Heterogeneous, Which Ensemble Is Better? *In Proceedings of the International Conference on Applications of Natural Language to Data Bases/ Information Systems*, 160–163. http://www.cs.ualberta.ca/

Chen, C. , Lin, H., & Loo, B. P. (2012). Exploring the Impacts of Safety Culture on Immigrants' Vulnerability in Non-motorized Crashes: A Cross-sectional Study. *Journal of Urban Health*, *89*, 138–152.

Chen, Z., & Fan, W. (David). (2019a). A multinomial logit model of pedestrian-vehicle crash severity in North Carolina. *International Journal of Transportation Science and Technology*, *8*(1), 43–52. https://doi.org/10.1016/j.ijtst.2018.10.001

Chen, Z., & Fan, W. (David). (2019b). A multinomial logit model of pedestrian-vehicle crash severity in North Carolina. *International Journal of Transportation Science and Technology*, *8*(1), 43–52. https://doi.org/10.1016/j.ijtst.2018.10.001

Chimba, D., Musinguzi, A., & Kidando, E. (2018). Associating pedestrian crashes with demographic and socioeconomic factors. *Case Studies on Transport Policy*, *6*(1), 11–16. https://doi.org/10.1016/j.cstp.2018.01.006

Das, S., Ashraf, S., Dutta, A., & Tran, L. N. (2020). Pedestrians under influence (PUI) crashes: Patterns from correspondence regression analysis. *Journal of Safety Research*, *75*, 14–23. https://doi.org/10.1016/j.jsr.2020.07.001

Dietterich, T. G. (2000). An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization. *Machine Learning*, *40*, 139–157.

Elish, M. O., Helmy, T., & Hussain, M. I. (2013). Empirical study of homogeneous and heterogeneous ensemble models for software development effort estimation. *Mathematical Problems in Engineering*, *2013*. https://doi.org/10.1155/2013/312067

Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention*, *40*(3), 1033–1054.

FARS 2019. (2023, May 11). *Fatality Analysis Reporting System. National Highway Traffic Safety Administration (NHTSA).* 2019. https://www-fars.nhtsa.dot.gov/Main/index.aspx

Febres, J., Mariscal, M., Herrera, S., & García-Herrero, S. (2021). Pedestrians' Injury Severity in Traffic Accidents in Spain: A Pedestrian Actions Approach . *Sustainability*, *13*(11), 6439.

Ferenchak, N. N., & Abadi, M. G. (2021). Nighttime pedestrian fatalities: A comprehensive examination of infrastructure, user, vehicle, and situational factors. *Journal of Safety Research*, *79*, 14–25. https://doi.org/10.1016/j.jsr.2021.07.002

Fernández-Alemán, J. L., Carrillo-de-Gea, J. M., Hosni, M., Idri, A., & García-Mateos, G. (2019). Homogeneous and heterogeneous ensemble classification methods in diabetes disease: a review. *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) : Biomedical Engineering Ranging from Wellness to Intensive Care*.

Friedman, J. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, *29*(5), 1189–1232.

Ghandour, A. J., Hammoud, H., & Al-Hajj, S. (2020). Analyzing factors associated with fatal road crashes: A machine learning approach. *International Journal of Environmental Research and Public Health*, *17*(11). https://doi.org/10.3390/ijerph17114111

GHSA. (2021). *2020 PRELIMINARY DATA by State Traffic Fatalities Pedestrian*. www.ghsa.org

Gibbs, K., Slater, S. J., Nicholson, L., Barker, D., & Chaloupka, F. J. (2012). *Income disparities in street features that encourage walking. Bridging the Gap Program.*

Haleem, K., Alluri, P., & Gan, A. (2015). Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis and Prevention*, *81*, 14–23. https://doi.org/10.1016/j.aap.2015.04.025

Ijaz, M., Lan, L., Usman, S. M., Zahid, M., & Jamal, A. (2021). Investigation of factors influencing motorcyclist injury severity using random parameters logit model with heterogeneity in means and variances. *International Journal of Crashworthiness*, *27*(5), 1412–1422. https://doi.org/10.1080/13588265.2021.1959153

Ijaz, M., Liu, L., Almarhabi, Y., Jamal, A., Usman, S. M., & Zahid, M. (2022). Temporal Instability of Factors Affecting Injury Severity in Helmet-Wearing and Non-Helmet-Wearing Motorcycle Crashes: A Random Parameter Approach with Heterogeneity in Means and Variances. *International Journal of Environmental Research and Public Health*, *19*(17). https://doi.org/10.3390/ijerph191710526

Islam, S., & Jones, S. L. (2014). Pedestrian at-fault crashes on rural and urban roadways in Alabama. *Accident Analysis and Prevention*, *72*, 267–276. https://doi.org/10.1016/j.aap.2014.07.003

Khattak, A., Chakraborty, S., Mahdinia, I., & Usman, S. (2023). Safety Enhancement by Detecting Driver Impairment Through Analysis of Real-Time Volatilities. *Collaborative Sciences Center for Road Safety*, 1–75.

Kim, J. K., Ulfarsson, G. F., Shankar, V. N., & Kim, S. (2008). Age and pedestrian injury severity in motor-vehicle crashes: A heteroskedastic logit analysis. *Accident Analysis and Prevention*, *40*(5), 1695–1702. https://doi.org/10.1016/j.aap.2008.06.005

Kotsiantis, S. B. (2013). Decision trees: a recent overview. *Artificial Intelligence Review*, *39*, 261–283.

Liu, J., Das, S., Zhan, F., & Khan, M. (2024). Spatial analysis of geographical disparities in pedestrian safety. *Transport Policy*, *156*, 164–181.

Liu, J., Khattak, A., Chen, C., Wan, D., Ma, J., & Hu, J. (2018). Revisiting Hit-and-Run Crashes: A Geo-Spatial Modeling Method. *Transportation Research Record*, *2672*(38), 81–92.

Mukherjee, D., & Mitra, S. (2020). A comprehensive study on identification of risk factors for fatal pedestrian crashes at urban intersections in a developing country. *Asian Transport Studies*, *6*, 100003.

Nasri, M., Aghabayk, K., Esmaili, A., & Shiwakoti, N. (2022). Using ordered and unordered logistic regressions to investigate risk factors associated with pedestrian crash injury severity in Victoria, Australia. *Journal of Safety Research*, *81*, 78–90. https://doi.org/10.1016/j.jsr.2022.01.008

Noland, R. B., Klein, N. J., & Tulach, N. K. (2013). Do lower income areas have more pedestrian casualties? *Accident Analysis and Prevention*, *59*, 337–345. https://doi.org/10.1016/j.aap.2013.06.009

Oviedo-Trespalacios, O., & Scott-Parker, B. (2017). Footbridge usage in high-traffic flow highways: The intersection of safety and security in pedestrian decision-making. *Transportation Research Part F: Traffic Psychology and Behaviour*, *49*, 177–187.

Pan, G., Fu, L., & Thakali, L. (2017). Development of a global road safety performance function using deep neural networks. *International Journal of Transportation Science and Technology*, *6*(3), 159–173. https://doi.org/10.1016/j.ijtst.2017.07.004

Parmar, A., Katariya, R., & Patel, V. (2018). A Review on Random Forest: An Ensemble Classifier. *International Conference on Intelligent Data Communication Technology and Internet of Things (ICICI)*, 758–763.

Patwary, A. L., Haque, A., Mahdinia, I., & Khattak, A. (2024). Investigating transportation safety in disadvantaged communities by integrating crash and Environmental Justice data. *Accident Analysis & Prevention*, *194*(107366).

Petrakova, A., Affenzeller, M., & Merkurjeva, G. (2015). Heterogeneous versus homogeneous machine learning ensembles. *Information Technology and Management Science*, *18*(1), 135–140.

Pour-Rouholamin, M., & Zhou, H. (2016). Investigating the risk factors associated with pedestrian injury severity in Illinois. *Journal of Safety Research*, *57*, 9–17. https://doi.org/10.1016/j.jsr.2016.03.004

Rashid, H., Shah, S., & Usman; Sheikh. (2023). Factors Influencing Motorists' Injury Severities: An Empirical Assessment of Crashes in District Peshawar, Pakistan. *Global Social Sciences Review, VIII*, 211–224.

Sabzevari, M., Martínez-Muñoz, G., & Suárez, A. (2018). Pooling homogeneous ensembles to build heterogeneous ones. *ArXiv Preprint ArXiv:1802.07877*. http://arxiv.org/abs/1802.07877

Sanders, R. L., & Schneider, R. J. (2022). An exploration of pedestrian fatalities by race in the United States. *Transportation Research Part D: Transport and Environment*, *107*. https://doi.org/10.1016/j.trd.2022.103298

Sasidharan, L., & Menéndez, M. (2014). Partial proportional odds model - An alternate choice for analyzing pedestrian crash injury severities. *Accident Analysis and Prevention*, *72*, 330–340. https://doi.org/10.1016/j.aap.2014.07.025

Siddiqui, N. A., Chu, X., Guttenplan, M., Siddiqui, N. A., & Chu, X. (1982). Crossing Locations, Light Conditions, and Pedestrian Injury Severity. In *Transportation Research Record: Journal of the Transportation Research Board*.

Tang, J., Liang, J., Han, C., Li, Z., & Huang, H. (2019). Crash injury severity analysis using a two-layer Stacking framework. *Accident Analysis and Prevention*, *122*, 226–238. https://doi.org/10.1016/j.aap.2018.10.016

Tay, R., Choi, J., Kattan, L., & Khan, A. (2011). A multinomial logit model of pedestrian-vehicle crash severity. *International Journal of Sustainable Transportation*, *5*(4), 233–249. https://doi.org/10.1080/15568318.2010.497547

Tewari, S., & Dwivedi, U. D. (2020). A comparative study of heterogeneous ensemble methods for the identification of geological lithofacies. . *Journal of Petroleum Exploration and Production Technology*, *10*, 1849–1868.

Torgo, L. (1999). Inductive learning of tree-based regression models. *Universidade Do Porto*.

Transportation Disadvantaged Census Tracts. (2023, May 11). *Historically Disadvantaged Communities*. https://www.transportation.gov/priorities/equity/justice40/transportation-disadvantaged-census-tracts-historically-disadvantaged

Ulfarsson, G., Kim, S., & Booth, K. (2010). Analyzing fault in pedestrian–motor vehicle crashes in North Carolina. *Accident Analysis and Prevention*, *42*(6), 1805–1813.

United States Census Bureau. (2021). NORTH CAROLINA: 2020 Census. *Https://Www.Census.Gov/Library/Stories/State-by-State/North-Carolina-Population-Change-between-Census-Decade.Html*.

USDOT. (2023). SS4A Underserved Communities Census Tracts (Historically Disadvantaged Communities). *Https://Www.Arcgis.Com/Apps/Dashboards/99f9268777ff4218867ceedfabe58a3a*.

Usman, S., & Khattak, A. (2025). Beyond the conventional: Exploring pedestrian safety on interstates with Bayesian and machine learning models. *Journal of Safety Research*, *94*, 141–156. https://doi.org/10.1016/j.jsr.2025.06.018

Usman, S., Khattak, A., Chakraborty, S., Mahdinia, I., & Tavassoli, R. (2024). Detection of distracted driving through the analysis of real-time driver, vehicle, and roadway volatilities. *Journal of Transportation Safety & Security*, 1–22. https://doi.org/10.1080/19439962.2024.2341393

Wolpert, D. (1992). Stacked Generalization. *Neural Networks*, *5*(2), 241–259.

Yasmin, S., Eluru, N., & Ukkusuri, S. V. (2014). Alternative Ordered Response Frameworks for Examining Pedestrian Injury Severity in New York City. *Journal of Transportation Safety and Security*, *6*(4), 275–300. https://doi.org/10.1080/19439962.2013.839590

Yu, C. Y., Zhu, X., & Lee, C. (2022). Income and Racial Disparity and the Role of the Built Environment in Pedestrian Injuries. *Journal of Planning Education and Research*, *42*(2), 136–149. https://doi.org/10.1177/0739456X18807759

Zheng, Z., Liu, Z., Liu, C., & Shiwakoti, N. (2014). Understanding public response to a congestion charge: A random-effects ordered logit approach. *Transportation Research Part A: Policy and Practice*, *70*, 117–134. https://doi.org/10.1016/j.tra.2014.10.016

Zhu, M., Sze, N., Newnam, S., & Zhu, D. (2023). Do footbridge and underpass improve pedestrian safety? A Hong Kong case study using three-dimensional digital map of pedestrian network. *Accident Analysis & Prevention*, *186*, 107064.

Ziakopoulos, A., Vlahogianni, E., Antoniou, C., & Yannis, G. (2022). Spatial predictions of harsh driving events using statistical and machine learning methods. *Safety Science*, *150*(1), 105722.