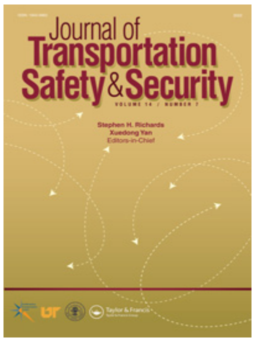

Journal of Transportation Safety & Security

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/utss20

Severity modeling of work zone crashes in New Jersey using machine learning models

Ahmed Sajid Hasan, Md Asif Bin Kabir, Mohammad Jalayer & Subasish Das

To cite this article: Ahmed Sajid Hasan, Md Asif Bin Kabir, Mohammad Jalayer & Subasish Das (2022): Severity modeling of work zone crashes in New Jersey using machine learning models, Journal of Transportation Safety & Security, DOI: 10.1080/19439962.2022.2098442

To link to this article: https://doi.org/10.1080/19439962.2022.2098442

Published online: 18 Jul 2022.

JOURNAL OF TRANSPORTATION SAFETY & SECURITY
https://doi.org/10.1080/19439962.2022.2098442

Severity modeling of work zone crashes in New Jersey using machine learning models

Ahmed Sajid Hasan[a], Md Asif Bin Kabir[b], Mohammad Jalayer[a], and Subasish Das[c]

[a]Civil and Environmental Engineering, Rowan University, Glassboro, New Jersey, USA; [b]Civil Engineering, United International University, Dhaka, Bangladesh; [c]Texas A&M University Transportation Institute College Station, College Station, Texas, USA

ABSTRACT

In the United States, the probability of work zone crashes has increased due to an increase in renovation works by transportation infrastructures. The severity of work zone crashes is associated with multiple contributing factors such as the roadway's geometric design features, temporal variables, environmental conditions, types of vehicles, and driver behaviors. For this study, we acquired and analyzed three years (2016–2018) of work zone crash data from the state of New Jersey. We investigated the performance of several machine learning methods, including Support Vector Machine, Random Forest, Catboost, Light GBM, and XGBoost to predict the type of injury severity resulting from work zone crashes. To evaluate models' performances, some statistical evaluation parameters such as accuracy, precision, and recall scores were calculated. In addition, a sensitivity analysis was conducted to assess the impact of the most influential factors in work zone-related crashes. Random Forest and Catboost outperformed the other models in terms of predicting fatal, major, and minor injuries. According to the sensitivity analysis, crash type and speed limit were the most significantly associated variables with crash severity. The findings of this study are expected to facilitate the identification of appropriate countermeasures for reducing the severity of work zone crashes.

KEYWORDS

Support Vector Machine; Random Forest; Classification and Regression Tree; work zone crash; sensitivity analysis; New Jersey

1. Introduction

Active transportation networks require ongoing road renovation and maintenance. According to the National Highway Traffic Safety Administration (NHTSA), a work zone is "an area where roadwork takes place and may involve lane closures, detours, and moving equipment" (Heneghan, 2018; Work/Construction Zones, 2008). Due to their unique traffic characteristics, these areas are associated with increases in the probability of crashes, injuries, and fatalities of both workers and motorists (Osman, Paleti,

CONTACT Mohammad Jalayer jalayer@rowan.edu Civil and Environmental Engineering, Rowan University, 201 Mullica Hill Road, Glassboro, NJ 08028-1700, USA.

Mishra, & Golias, 2016; Weng, Meng, & Wang, 2013; Weng & Meng, 2011, 2012). In the United States, the number of fatal work zone crashes is increasing every day in comparison with non-work zone crashes (FHWA, 2019; Garber & Woo, 1990; Garber & Zhao, 2002; Hall & Lorenz, 1989). The Federal Highway Administration (FHWA) reported an 11% increase in fatal work zone crashes from 2018 to 2019 (FHWA, 2021). The FHWA also reported a total of 842 deaths in 2019 due to work zone crashes, including 135 construction workers (FHWA, 2021). Specifically, 34 fatal work zone crashes occurred during a recent five-year period from 2014 to 2018 in New Jersey, resulting in the deaths of both pedestrians and workers (National Work Zone Safety Information Clearinghouse, 2019; Yang, Ozbay, Ozturk, & Yildirimoglu, 2013).

To improve the safety of construction workers and road users, the factors associated with the frequency and severity of work zone crashes must be identified. Furthermore, site- and state-specific factors must be investigated to comprehensively understand the causes of work zone crashes. The primary goal of this study is to explore the most critical factors that contribute to the severity of work zone crashes in New Jersey-based on three years of crash data (2016–2018). For this study, we employed three different feature importance techniques—Classification and Regression Tree (CART), univariate selection, and extratree classifiers to identify the most important variables associated with crash severity (by considering the top common factors from each of the techniques according to their rank). Various machine learning algorithms were used following determining the key variables to predict the severity of work zone crashes. Afterward, a sensitivity analysis of the top contributing factors was performed to determine the effects of individual variables on the crash severity predictions of the various models.

This study contributes to the existing literature that analyzes work zone crash severity by conducting a comprehensive review of the state of practice and art in crash severity prediction methods across the USA over the past four decades. To be specific, the list of contributing factors, along with the methods utilized for crash severity prediction, was gathered. In addition, this study overcame the variable selection bias of a single variable selection method by using three different variable selection methods, which is unique compared to the previous studies. Moreover, in this study, the comparison of conventional machine learning techniques with boosting techniques was performed. Furthermore, this study helped find the sensitivity of each of the leading contributing factors that highly impact the crash severity. We expect our findings to provide invaluable information to help engineers, practitioners, and policymakers develop appropriate countermeasures to control the severity of work zone crashes in New Jersey.

2. Literature review

Several studies have been conducted in various states regarding the factors associated with crash severity. Numerous factors (e.g., driver characteristics, crash attributes, roadway features, environmental conditions, and vehicle characteristics) have been found to influence the severity of work zone crashes (Table 1). Driver characteristics have been identified as having a significant impact on the severity of work zone crashes, and the age of the driver (Ghasemzadeh & Ahmed, 2019a, 2019b; Hamzeie, Savolainen, & Gates, 2017; Khattak, Rodriguez, Targa, & Rocha, 2003; Liu, Khattak, & Zhang, 2016; Weng, Zhu, Yan, & Liu, 2016; Weng & Meng, 2011), alcohol ingestion (Garber & Woo, 1990; Hall & Lorenz, 1989; Meng, Weng, & Qu, 2010; Meng & Weng, 2011; See, 2005; Weng et al., 2016), drugged drivers (Ghasemzadeh & Ahmed, 2019a, 2019b), and distracted drivers (Liu et al., 2016) are the major driver attributes that contribute to the severity of these crashes. Previous studies have also found that crash severity in work zones is often influenced by vehicular characteristics such as the number of vehicles involved (Bryden, Andrew, & Fortuniewicz, 1998; Hall & Lorenz, 1989; Khattak & Targa, 2004; Meng et al., 2010; Meng & Weng, 2011; Osman, Paleti, & Mishra, 2018; Qi, Srinivasan, Teng, & Baker, 2005) and the types of vehicles (Daniel, Dixon, & Jared, 2000; Garber & Zhao, 2002; Hall & Lorenz, 1989; Hamzeie et al., 2017; Khattak et al., 2003; Liu et al., 2016; Meng et al., 2010; Nemeth & Migletz, 1978; Weng & Meng, 2011). Numerous studies have also found environmental factors to significantly impact crash severity in the work zone. Specifically, lighting conditions (Daniel et al., 2000; Elghamrawy, 2011; Hall & Lorenz, 1989; Hamzeie et al., 2017; Li & Bai, 2009; Meng & Weng, 2011; Prichard, 2015; See, 2005; Weng & Meng, 2011), road surface conditions (Ahmed et al., 2015; Bryden et al., 1998; Elghamrawy, 2011; Ghasemzadeh & Ahmed, 2019a, 2019b; Hall & Lorenz, 1989; Hamzeie et al., 2017; Meng & Weng, 2011; Weng & Meng, 2011), and weather conditions (Garber & Woo, 1990; Ghasemzadeh & Ahmed, 2019a; Hall & Lorenz, 1989; Liu et al., 2016; Nemeth & Migletz, 1978; Osman et al., 2018; Prichard, 2015; Salem, Genaidy, Wei, & Deshpande, 2006; See, 2005; Weng et al., 2016) have been proven to have a considerable impact on crash severity.

Crash attributes such as the type of crash (Daniel et al., 2000; Garber & Woo, 1990; Garber & Zhao, 2002; Hall & Lorenz, 1989; Khattak & Targa, 2004; Meng et al., 2010; Prichard, 2015; Weng et al., 2016) and the presence of a curve (Hall & Lorenz, 1989; Liu et al., 2016; Meng et al., 2010; See, 2005) have also been shown to contribute significantly to the crash severity of work zones. Roadway features play a critical role in work zone crashes as the geometric design of a roadway directly impacts crash severity. The roadway features that have been found to have a significant impact

Table 1. Summary of previous studies on work zone crashes.

State	Study Period	Environmental	Roadway	Temporal	Vehicle	Driver	Crash	Author (Year of publication)
Alabama	1998–2012	Light conditions, weather conditions, surface condition	Functional class, no of lanes, traffic controls present, work zone crash location, work zone type	Time of day, day of week	–	–	Crash type	(Prichard, 2015)
Arkansas	2000–2005	Light conditions	Functional class,	Time of day	–	Alcohol involved	Curve related	(See, 2005)
Florida	2012–2017	Weather conditions, light conditions, surface condition	Functional class, Shoulder Width, work zone characteristics	Time of day, season	Overall traffic volume, truck volume, vehicle type	Inattentive/ distracted driver	Crash type	(Islam, Alnawmasi, & Mannering, 2020)
Florida, Indiana, New York, North Carolina, Pennsylvania, and Washington	2012–2016	Weather conditions, light conditions, surface condition	Alignment, level of service	–	Vehicle type	Gender, age, seatbelt use	Prior crashes, event type	(Hamzeie et al., 2017)
Georgia	1995–1997	Light conditions	Functional class, objects involved in the crash	Time of day	Vehicle type	–	Crash type	(Daniel et al., 2000)
Illinois	1996–2005	Weather conditions, light conditions, surface condition	traffic controls present, no of lanes, median type, functional class, speed limit	Time of day	–	–	–	(Elghamrawy, 2011)
Iowa	2017–2018	–	Type of roadway, AADT, road curvature, no of lanes, functional class, speed limit, length of work zone, activity of work zone	–	Percentage of truck	–	–	(Kassmeyer, 2020)
Kansas	1998–2004	Weather conditions, light conditions	No. of lanes, functional class	Day of week, time of day	–	–	–	(Li & Bai, 2009)
Michigan	2004–2008	Weather conditions, light conditions, surface condition	No of lanes, speed limit, traffic controls present	–	Total vehicles involved,	Alcohol involved	–	(Meng & Weng, 2011)
Michigan	2004–2008			–		Gender, alcohol involved	Crash type	(Weng et al., 2016)

(*continued*)

Table 1. Continued.

State	Study Period	Environmental	Roadway	Temporal	Vehicle	Driver	Crash	Author (Year of publication)
		Weather conditions, light conditions, surface condition	Speed limit, no of lanes, traffic controls present, crash location		Vehicle age, type of vehicle, airbag, seat position			
Michigan (Southeast)	1999–2008	Weather conditions	–	–	Total vehicles involved, vehicle type	Alcohol involved	Crash type, curve related	(Meng et al., 2010)
Minnesota	2003–2012	Weather conditions	Roadway, access control, alignment, no of lanes, functional class, area type, speed limit, work zone type, location, presence of workers	Day of week, time of day,	Total vehicles involved, Involvement of truck	–	–	(Osman et al., 2018)
New Mexico	1983–1985	Light conditions, weather conditions, surface condition	Grading of the roadway	Day of week, time of day	Total vehicles involved, vehicle type	Alcohol involved, pedestrian involvement	Crash severity, crash type, curve related	(Hall & Lorenz, 1989)
New York	1994–2001	–	Traffic controls present., functional class	–	Total vehicles involved	–	–	(Qi et al., 2005)
New York	1994–1996	Surface conditions	Traffic controls present	–	Total vehicles involved, construction vehicles and equipment	–	–	(Bryden et al., 1998)
North Carolina	2010–2014	Surface condition, weather conditions, light conditions	Speed limit, median type, no of lanes, work zone location, traffic controls present	Day of week, time of day	–	Gender, age	Crash type	(Koilada, Mane, & Pulugurtha, 2020)
North Carolina	2000	Weather conditions, light conditions	Median type, speed limit, traffic controls present	–	Total vehicles involved	–	Crash type	(Khattak & Targa, 2004)
North Carolina	2000	Weather conditions, light conditions	Median type, speed limit, traffic controls present	–	Total vehicles involved, vehicle type	Age	Crash type	(Khattak et al., 2003)
Ohio	1972–1974	Weather conditions, road conditions	Location of the crash, constriction object involved, traffic controls presents	Time of day	Vehicle type	Causes of the crash	Crash severity, crash type	(Nemeth & Migletz, 1978)

(continued)

Table 1. Continued.

State	Study Period	Environmental	Roadway	Temporal	Vehicle	Driver	Crash	Author (Year of publication)
Ohio	2001–2003	Weather conditions, light conditions, surface condition	Location of the crash, traffic controls present, road contour	–	–	–	Crash type	(Salem et al., 2006)
Virginia	1982–1985	Weather conditions	Functional class, speed limit, traffic controls present	–	–	Alcohol involved	Crash severity, crash type	(Garber & Woo, 1990)
Virginia	1996–1999	–	Location of the crash in work zone	Time of day	Vehicle type, total vehicles involved	–	Crash severity, crash type	(Garber & Zhao, 2002)
Virginia	2013	Weather conditions, light conditions	Pavement type, traffic controls present, functional class,	–	Vehicle speed, vehicle type, Total vehicles involved	Age, gender, distracted driver	Crash type, curve related	(Liu et al., 2016)
Washington	2006–2013	Light conditions, weather condition, Surface condition	Temporary traffic control zone, land use, functional class, traffic controls present, crash location	–	Speed of vehicle, total vehicles involved, vehicle age, vehicle type, airbag	Gender, age, drugged driver	Crash severity, crash type	(Ghasemzadeh & Ahmed, 2019a)
Washington	2009–2013	Surface condition, weather conditions, light conditions	Posted speed, functional class, intersection related, work zone type, traffic controls present	Time of day, peak hours	Total vehicles involved	Age, gender, drugged driver	Crash severity, crash type	(Ghasemzadeh & Ahmed, 2019b)
Wisconsin	2009–2012	–	Lane closure type, facility type, temporary traffic control zone	Time of day, day of week	–	–	–	(Cheng, Parker, Ran, & Noyce, 2016)
Washington, Florida, Indiana, North Carolina, Pennsylvania, New York	2011–2013	Weather conditions, light conditions	Speed limit, trip distance	Trip Duration	Vehicle type	Gender	Crash severity	(Ahmed et al., 2015)

on work zone crash severity include the number of lanes (Elghamrawy, 2011; Li & Bai, 2009; Meng & Weng, 2011; Osman et al., 2018; Prichard, 2015; Weng et al., 2016; Weng & Meng, 2011), functional classification of the road (Bryden et al., 1998; Daniel et al., 2000; Elghamrawy, 2011; Garber & Woo, 1990; Ghasemzadeh & Ahmed, 2019b; Li & Bai, 2009; Osman et al., 2018; Qi et al., 2005), traffic control devices (Bryden et al., 1998; Ghasemzadeh & Ahmed, 2019a; Salem et al., 2006), and speed limit (Elghamrawy, 2011; Garber & Woo, 1990; Khattak & Targa, 2004; Meng & Weng, 2011; Osman et al., 2018; Weng et al., 2016; Weng & Meng, 2011). Table 1 summarizes various contributing factors of the work zone crashes in various states across the nation.

In recent years, researchers have applied various statistical methods for crash analysis. These methods include, but are not limited to, the classification tree modeling, the association rule-based approach, hierarchical modeling, regression modeling, mixed generalized ordered response probit modeling, multinomial logistic regression modeling, and the probit classification tree (Liu et al., 2016; Prichard, 2015; Sze & Song, 2019; Weng et al., 2016). Traditional statistical models (e.g., logistic regression, Bayesian logistic regression, and artificial neural networks (ANNs)) have also been used in crash severity analysis (Abdel-Aty, Dilmore, & Dhindsa, 2006; Abdel-Aty, Pande, Lee, Gayah, & Santos, 2007; Abdel-Aty, Uddin, Pande, Abdalla, & Hsia, 2004; Ahmed, Abdel-Aty, & Yu, 2012a, 2012b). Over the past decade, researchers have achieved great success by utilizing various machine learning models for crash severity analysis (Chen, Zhang, Qian, Tarefder, & Tian, 2016a; Gang & Zhuping, 2011; Li, Liu, Wang, & Xu, 2012; Yu & Abdel-Aty, 2013). For instance, Mokhtarimousavi, Anderson, Azizinamini, and Hadi (2019) analyzed work zone crash severity and found the SVM outperformed a mixed logit model. Ashqar, Shaheen, Ashur, and Rakha (2021) demonstrated that Random Forest works better in severity prediction of work zone crashes compared to traditional statistical models such as logistic regression. Researchers have also estimated the sensitivities of the major contributing factors in crash severity to better understand their impact (changes in the prediction accuracy of the model due to changes in individual variables) (Chen, Wang, & Zuylen, 2009; Chen, Zhang, Yang, Milton, & Alcántara, 2016b; Cheu, Xu, Kek, Lim, & Chen, 2006; Gang & Zhuping, 2011; Li et al., 2012; Lv, Tang, Zhao, & Li, 2009). Figure 1 illustrates various methods in the crash severity prediction of work zone crashes. The variables are clustered based on the categories provided in Table 1 including driver characteristics, roadway features, vehicle characteristics, environmental conditions, temporal features, and crash attributes. It is observed that most of the previous studies have used regression models (e.g., linear regression and logistic regression) for the crash severity

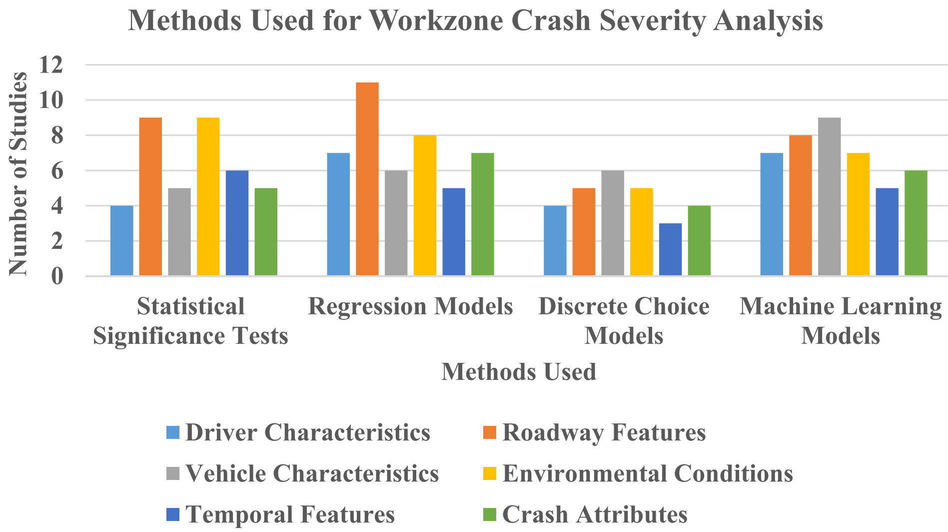

Figure 1. Summary of methods used for work zone crash severity analysis.

prediction, followed by statistical significance tests (e.g., chi-square test and F-test). Several other studies have used discrete choice models (e.g., mixed logit model and ordered probit model) and conventional machine learning models (e.g., hierarchical and association rules, KNN, Support Vector Machine, Random Forest). We note that very few studies have utilized both the conventional machine learning models and the newer boosting techniques, which we address in this paper.

Most recent studies on work zone crashes have utilized discrete choice models (i.e., ordered probit models and logit models) and basic machine learning models (i.e., KNN, SVM, RF) for crash severity prediction. Although boosting techniques were used for crash severity prediction in other types of crashes (Hasan, Kabir, & Jalayer, 2021; Parsa, Movahedi, Taghipour, Derrible, & Mohammadian, 2020; Zheng, Lu, & Lantz, 2018), there has been a relatively scarce work on work zone crash severity analysis using the boosting methods. Nasrollahzadeh, Sofi, and Ravani (2021) used XGBoost technique to identify the factors contributing to the risk of collision at roadside work zones. However, there is a lack of studies in the literature comparing various boosting techniques (XGBoost, LightGBM, Catboost) with conventional machine learning methods, which we addressed in this paper. Therefore, the aim of this study is to predict the severity of work zone crashes using boosting methods and conventional machine learning algorithms (SVM and Random Forest) and compare their performances in predicting work zone crash severity. While most of the previous studies used a single variable selection method (Chang & Chien, 2013; Delen, Tomak, Topuz, & Eryarsoy, 2017; Hossain & Muromachi, 2013; Kashani & Mohaymany, 2011), we systematically identified the important variables using three unique variable selection methods. The results of this paper are expected to significantly assist transportation agencies in identifying the possible contributing factors of crash severity and facilitate the development of effective safety countermeasures to reduce work zone crash frequency and severity.

3. Data

3.1. Data collection

Three years (2016–2018) of work zone crash data in New Jersey, including 8,231 crashes (267 fatal and injury, 1,504 with possible injuries, and 6,460 with no injury), were used in this study. The raw dataset and discarded, incomplete, and erroneous data records were examined due to the raw data containing some missing values and variables unrelated to crash severity. As an example, variables unrelated to crash severity (i.e., "Crash ID" and "Street Name") were eliminated. Finally, 20 independent variables were identified based on the literature review for further data analysis. The selected variables are the most common type of variables used in previous studies, as shown in Table 1. These variables were divided into six categories, including roadway features (i.e., number of lanes, functional class of road, temporary traffic control device, presence of traffic control, and speed limit), temporal features (i.e., day of the week, season, and time of day), driver characteristics (i.e., driver age, distracted driver, drugged driver, and alcohol involvement), environmental conditions (i.e., light conditions, weather conditions, and surface conditions), crash attributes (i.e., curve-related and crash type), and vehicle characteristics (i.e., the total number of vehicles involved and vehicle type).

Figure 2 illustrates the geospatial distribution of the work zone crashes occurring during the three study years in the state of New Jersey. The density of the crashes is mostly situated in the northeastern region of the state. The southwestern region has also experienced a fair density of work zone crashes. The southeastern zone of the state is the least affected by work zone crashes. One of the main reasons for the high frequency of work zone crashes in the northeastern region is the high frequency of traffic in northeastern New Jersey, close to New York. According to the Annual Average Daily Traffic (AADT) data of New Jersey, the busiest roads are located at the northeastern region (New Jersey Department of Transportation, 2018). These roads include several interstate highways, such as I-80, I-78, I-278, and I-80, and also some toll roads, such as I-95 and Garden State Parkway. The southwestern region of the state, which is close to Philadelphia, also has a fair frequency of traffic. This area is also connected with some major interstate highways such as I-295, I-95, and I-

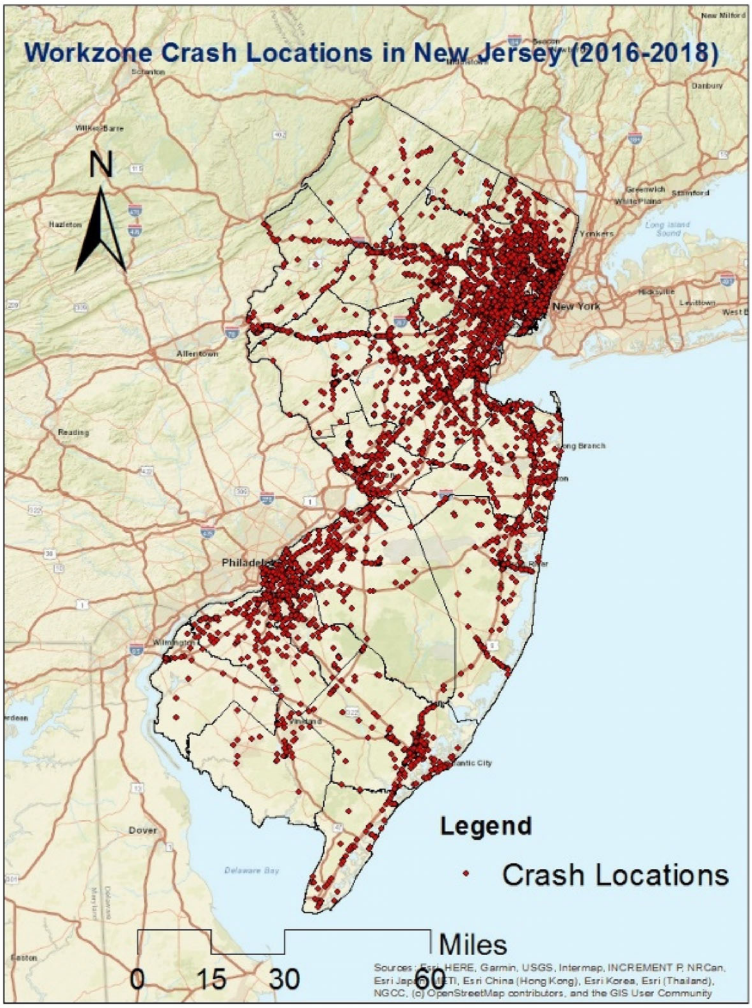

Figure 2. Work zone crash locations in New Jersey (2016–2018).

676, which have considerably higher traffic densities coming from and going to other states.

The work zone crashes in the study period are split into various crash severities per the KABCO definition of crash severity, developed by the National Safety Council (FHWA, 2016). Due to the low frequency of fatal injury (0.16%), suspected serious injury (0.32%), and suspected minor injury (2.76%), we merged all of the injury categories into one severity class and named it 'fatal and injury' (contains 3.24% of the total crashes) (Table 2).

3.2. Descriptive statistics

The descriptive statistics of the study variables related to crash severity are summarized in Table 3. It is observed that 47.76% of crashes were rear-

Table 2. Work zone crash severity in New Jersey (2016–2018).

Crash severity	Definition	Other names	KABCO scale	2016	2017	2018	Total
Fatal injury	The victim is deceased.	Killed	K	6	3	4	13
Suspected serious injury	The victim has a non-fatal injury. Cannot walk, drive or normally continue the activities that they could perform before the motor vehicle crash	Incapacitated	A	8	10	8	26
Suspected minor injury	An evident injury, other than fatal and incapacitating. Injury is visible, such as a lump on head, abrasion, bleeding, or lacerations.	Moderate injury	B	81	65	82	228
Possible injury	A reported or claims of injury that is not fatal, incapacitating or moderate. Injury is not visible to the investigating officer.	Complaint of pain	C	512	464	528	1,504
No apparent injury	Only damage of properties	Property damage only	O	2,272	2,027	2,161	6,460
Total crashes				2,879	2,569	2,783	8,231

ended collisions and 85.55% of the total number of fatal crashes involved wet surface conditions. 44.53% of work zone crashes were attributed to distracted driving. Two-vehicle crashes accounted for 78.08% of all crashes, with 50.94% occurring on two-lane roads. Clear weather accounted for 85.97% of the crashes, while daylight accounted for 72.36%. Surface conditions, traffic signals, and weather conditions all showed skewness in favor of fatal and injury types of crash severity. We noted that compared to other features, the driver characteristics 'alcohol involved' (17.89%) and 'drugged driver' (23.91%) both contributed to higher rates of fatal and injury crashes.

4. Method

4.1. Study design

To analyze the work zone crashes in New Jersey, we have gathered New Jersey crash data for the years 2016–2018. An open-source sci-kit learn python package was used to develop the codes for this study (Pedregosa, 2011). Three different variable importance techniques (i.e., CART, univariate selection, and extratree classifiers) were employed on the crash data to determine the importance of the contributing factors. The categorized data with the important variables were then split into training and testing datasets with a 70/30 ratio. Five different machine learning techniques (Support Vector Machine, Random Forest, LightGBM, XGBoost, and CatBoost) were applied to the dataset, and the performances of the models were evaluated. A sensitivity analysis was performed on the top ten important variables to

Table 3. Distribution of descriptive features in terms of crash severity.

Explanatory variable	Fatal and injury	%	Possible injury	%	No injury	%	Total	Frequency (%)
Total crashes	267	3.24	1,504	18.27	6,460	78.48	8,231	100
Temporal variables								
Season								
Spring	79	3.82	348	16.84	1,640	79.34	2,067	25.11
Summer	79	3.38	450	19.24	1,810	77.38	2,339	28.42
Fall	72	3.23	402	18.02	1,757	78.75	2,231	27.1
Winter	37	2.32	304	19.07	1,253	78.61	1,594	19.37
Day of week								
Weekday	207	3.05	1,228	18.09	5,354	78.86	6,789	82.48
Weekend	60	4.16	276	19.14	1,106	76.7	1,442	17.52
Time of day								
Day	173	2.89	1,042	17.42	4,767	79.69	5,982	72.68
Evening	60	3.38	355	20.02	1,358	76.59	1,773	21.54
Night	34	7.14	107	22.48	335	70.38	476	5.78
Roadway features								
No of lanes								
1	4	6.45	11	17.74	47	75.81	62	0.75
2	134	3.2	755	18.01	3,304	78.8	4,193	50.94
3 or more	129	3.24	738	18.56	3,109	78.19	3,976	48.31
Temporary traffic control zone								
Construction	243	3.2	1,373	18.06	5,986	78.74	7,602	92.36
Maintenance	9	2.51	69	19.27	280	78.21	358	4.35
Utility	8	3.28	56	22.95	180	73.77	244	2.96
Other	7	25.93	6	22.22	14	51.85	27	0.33
Traffic controls present								
Signal/Yield Sign/Other	238	3.21	1,351	18.24	5,819	78.55	7,408	90
None	22	3.71	103	17.37	468	78.92	593	7.2
Unknown	7	3.04	50	21.74	173	75.22	230	2.79
Speed limit (mph)								
<35	22	2.23	149	15.08	817	82.69	988	12
35–45	38	2.51	256	16.89	1,222	80.61	1,516	18.42
45–55	71	3.03	436	18.59	1,838	78.38	2,345	28.49
55–65	136	4.02	663	19.6	2,583	76.37	3,382	41.09
Functional class								
Rural	14	5.76	47	19.34	182	74.9	243	2.95
Urban	253	3.17	1,457	18.24	6,278	78.59	7,988	97.05
Environmental conditions								
Weather conditions								
Adverse	41	3.55	201	17.4	913	79.05	1,155	14.03
Clear	226	3.19	1,303	18.41	5,547	78.39	7,076	85.97
Light conditions								
Daylight	173	2.9	1,043	17.51	4,740	79.58	5,956	72.36
Dawn/Dusk	8	3.14	36	14.12	211	82.75	255	3.1
Dark-not-Lit	16	3.37	104	21.89	355	74.74	475	5.77
Dark-Lit	70	4.53	321	20.78	1,154	74.69	1,545	18.77
Surface condition								
Wet	222	3.15	1,285	18.25	5,535	78.6	7,042	85.55
Dry	45	3.78	219	18.42	925	77.8	1,189	14.45
Vehicle characteristics								
Total vehicles involved								
1	68	6.78	123	12.26	812	80.96	1,003	12.19
2	140	2.18	1,081	16.82	5,206	81	6,427	78.08
3 or more	59	7.37	300	37.45	442	55.18	801	9.73
Vehicle type								
Light vehicles	168	3.16	1,037	19.49	4,116	77.35	5,321	64.65
Heavy vehicles	46	3.81	253	20.98	907	75.21	1,206	14.65
Others	53	3.11	214	12.56	1,437	84.33	1,704	20.7
Crash attributes								
Curve related								

(*continued*)

Table 3. Continued.

Explanatory variable	Fatal and injury	%	Possible injury	%	No injury	%	Total	Frequency (%)
Yes	34	2.89	192	16.33	950	80.78	1,176	14.29
No	233	3.3	1,312	18.6	5,510	78.1	7,055	85.71
Crash type								
Rear-end	109	2.77	979	24.9	2,843	72.32	3,931	47.76
Angle	40	6.23	181	28.19	421	65.58	642	7.8
Sidewise	38	1.49	199	7.81	2,311	90.7	2,548	30.96
Fixed object	61	6.43	127	13.4	760	80.17	948	11.52
Other	19	11.73	18	11.11	125	77.16	162	1.97
Driver characteristics								
Driver age								
Old (65+)	47	3.34	250	17.78	1,109	78.88	1,406	17.08
Middle-aged (21–64)	195	3.23	1,086	18	4,753	78.77	6,034	73.31
Young (16–20)	25	3.16	168	21.24	598	75.6	791	9.61
Drugged driver								
Yes	11	23.91	12	26.09	23	50	46	0.56
No	256	3.13	1,492	18.23	6,437	78.64	8,185	99.44
Distracted driver								
Yes	122	3.33	711	19.4	2,832	77.27	3,665	44.53
No	145	3.18	793	17.37	3,628	79.46	4,566	55.47
Alcohol involved								
Yes	34	17.89	44	23.16	112	58.95	190	2.31
No	233	2.9	1,460	18.16	6,348	78.95	8,041	97.69

evaluate the impact of every individual variable on the severity of crashes. Finally, we suggested safety countermeasures for the work zone crash severity in New Jersey using the findings of this study. Figure 3 illustrates the study design as a flowchart.

4.2. Ranking variable importance and predicting variable selection

Classification problems are associated with insignificant factors that increase the noise level and hamper the model's prediction performance (Chen et al., 2016a). Therefore, the use of variable selection procedures, such as variable importance ranking, is beneficial. Discrete choice models, CART, extratree classifier, univariate selection, random forest, and other techniques can be employed to rank variable importance. Machine learning models use these variable selection techniques to determine the relative importance of variables with respect to the predicted variables (Chen et al., 2015a, 2015b). As the authors of several studies (Ahmed et al., 2012a; Chang & Chen, 2005; Chang & Chien, 2013; Chen et al., 2016b; Delen et al., 2017; Hossain & Muromachi, 2013; Kashani & Mohaymany, 2011; Kuhnert, Do, & McClure, 2000; Lv et al., 2009; Montella, Aria, D'Ambrosio, & Mauriello, 2012) have reported the CART technique, extratree classifier, and univariate selection to be excellent ranking techniques, we applied these three feature selection techniques in our analysis of the relative importance of variables associated with crash severity.

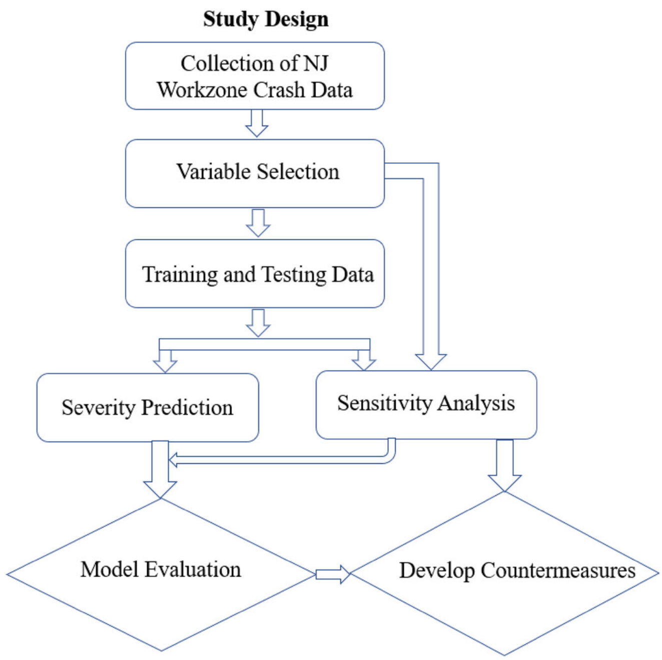

Figure 3. Flow chart of the study design.

4.3. Machine learning models

Five different ML algorithms, including Support Vector Machine, Random forest, XGboost, LightGBM, and Catboost, were deployed in this paper to predict the crash severity of work zone crashes. ML approaches, such as Support Vector Machine and Random Forest, were used in many previous studies in transportation engineering research (Siddiqui, Abdel-Aty, & Huang, 2012; Yu & Abdel-Aty, 2013). Few recent studies on crash severity analysis, on the other hand, have used boosting techniques (Abou Elassad, Mousannif, & Al Moatassime, 2020; Parsa et al., 2020). All of the algorithms utilized in this investigation were explained in detail by Friedman (2001). This section has a brief description of all of the ML models.

4.3.1. Random Forest

Random forest (RF) is a tree-based classifier that employs two distinct machine learning techniques: random feature selection and bagging (Breiman, 2001). Random feature collection creates decision trees immediately, whereas bagging creates each tree separately. Rather than employing all of the features in the decision trees, Random forest selects the features of the subsets at random. For forecasting the output of a new dataset, Random Forest uses the mean value of the outputs from random independent bootstrap training data.

4.3.2. Support vector machine

The Support Vector Machine (SVM) method has been described in detail by previous researchers with respect to crash severity analysis (Ahmed et al., 2012a; Li et al., 2012; Zhang & Xie, 2007). The SVM is capable of classifying both separable and non-separable data (Boser, Guyon, & Vapnik, 1992; Cortes & Vapnik, 1995) The aim of SVM is to maximize the margin or separation among various classes of crash severity. Simultaneously, it makes some errors, which can be adjusted later using a penalty parameter (Ahmadi, Jahangiri, Berardi, & Machiani, 2020). The equations used by the SVM are as follows (Li, Lord, Zhang, & Xie, 2008):

Minimize
b, w, ξ

$$\left(\frac{1}{2}\mathrm{wTw} + \mathrm{C}\sum_{n=1}^{N}\xi n\right) \tag{1}$$

subject to:

$$\mathrm{yn}\ (\mathrm{wT}\phi(\mathrm{xn}) + \mathrm{b}) \geq 1 - \xi n, \mathrm{n} = 1, 2 \ldots \mathrm{N} \tag{2}$$

$$\xi n \geq 0 \text{ for n } = 1, 2 \ldots \text{ N}, \tag{3}$$

where

W	Parameter to define decision boundary between classes
C	Regularization (or penalty) parameter
ξn	Error parameter to denote margin violation
B	Intercept associated with decision boundaries
$\phi(x_n)$	Function to transform data from X space into some Z space
y_n	Target value for the n^{th} observation

We employed multiclass classification to categorize the dataset into more than two classes, and three models were initially developed using two classes each (i.e., no injury vs. fatal and injury, fatal and injury vs. possible injury, or no injury vs. possible injury). For a new observation, these three models generate votes for the different classes, and the predicted class is determined by the class that receives the most votes. The task of maximizing the distance between the points closest to the hyperplane and computation of the coefficient (C) value is performed by introducing a Lagrange multiplier to the SVM classifier with the basic form as shown in Equation 4 (Li et al., 2008):

$$\text{Min max}\left\{\frac{1}{2}\mathrm{w_T w}\Big(+\mathrm{C}\sum\nolimits_{n=1}^{N}\xi n-\sum\nolimits_{n=1}^{N}\alpha n[yn\ \ (wT\phi(xn)+\ \ b)-1+\ \ \xi n]\right.$$
$$\left.-\sum\nolimits_{n=1}^{N}\beta n\ \xi n\right\} \tag{4}$$

where α_n, $\beta_n > 0$ are Lagrange multipliers.

For linearly inseparable data, a kernel function is applied for the non-linear transformation of the linear kernel. The data is transformed into a higher-dimensional space using the kernel function. One widely used kernel function for crash severity analysis is the radial basis function (RBF) (Li et al., 2008). The non-linear transformation of kernels and the RBF kernel are expressed as shown in Equations 5 and 6, respectively (Li et al., 2008):

$K(x_n, x_m) \equiv \phi(x_n)^T\ \phi(x_m)$ (5)

$K(x_n, x_m) = \exp(-\gamma\ ||x_n - x_m||^2)$, $\gamma > 0$ (6)where γ is the kernel parameter. With the RBF as the kernel function, the SVM model has two parameters (C, γ) that must be determined. We carefully determined these parameters (C, γ) using the grid search algorithm (Chang & Lin, 2011). Grid search combs through the iterations to find the best possible combination of these parameters that yields the maximum accuracy. We used the grid-search technique for fitting the model into the training dataset. Linear, RBF, and polynomial kernels with different combinations of C and gamma values were used to find out the best combination of the hyperparameters to fit the model in our dataset.

4.3.3. Boosting methods

The objective of the boosting methods is to improve the prediction performance by combining a set of weak classifiers with a strong classifier. Three different boosting methods, such as XGboost, LightGBM, and Catboost, were used in this study The gradient boosting approach adjusts the losses by regressing the gradient vector function at each iteration (Friedman, 2001). Starting with the weak decision tree, which was used as the foundation decision Tree, a gradient boosting model adjusts the sequence of each decision tree. XGboost is a slow boosting strategy that reduces misclassification errors at each iteration by using sequential model training. LightGBM is a boosting method based on the development of more accurate and complex decision trees leaf-by-leaf. Catboost is a boosting approach that works with both numerical and category input variables. It takes care of the variables during the training period, which saves time on preprocessing.

Table 4. Parameter values used for machine learning algorithms.

Algorithms	Parameters
Support vector machine	C = 1.0, kernel='rbf', degree = 3, gamma='scale', shrinking = True, cache_size = 200, class_weight = None, verbose = False, random_state = 0
Random forest	n_estimators = 100, criterion='gini', random_state = 0, verbose = 0, class_weight = None, max_samples = None, max_depth = None, min_samples_split in_samin_samples_leaf = 1, max_features='auto', max_leaf_nodes = None, ccp_alpha = 0.0, bootstrap='True'
XGBoost	n_estimators = 200, min_samples_split = 2, min_samples_leaf = 1, max_features='auto', max_depth = 50, bootstrap='False'
LightGBM	boosting_type='gbdt', num_leaves = 31, max_depth= −1, learning_rate = 0.1, n_estimators = 100, n_jobs= −1, subsample_for_bin = 200,000, class_weight = None, min_child_weight = 0.001, min_child_samples = 20, subsample = 1.0, colsample_bytree = 1.0, random_state = 0
Catboost	iterations = 500, learning_rate = 0.03, depth = 6, l2_leaf_reg = 3.0, loss_function= 'Logloss', verbose = None, boosting_type = None, random_state = 0

4.4. Model performance evaluation

The performance of the model was determined based on its accuracy and precision (Delen et al., 2017). The accuracy rate, the metric most widely used to evaluate prediction models, is the percentage of correctly classified samples from the total test dataset (Equation 7). Precision is the percentage of correct predictions among the total number of predictions for a class (Equation 8). In this study, we assessed the performance of severity prediction with respect to both accuracy and precision. In addition, we experimented with a variety of parameter values for the machine learning algorithms under consideration and chose the combination that produced the best output results. These variables are listed in Table 4.

$$\text{Accuracy} = \frac{\text{Total number of correct predictions}}{\text{Total number of predictions}} \tag{7}$$

$$\text{Precision} = \frac{\text{Total number of correct predictions for } a \text{ specific class}}{\text{Total number of predictions for } a \text{ specific class}} \tag{8}$$

4.5. Sensitivity analysis (interpretable machine learning)

An important aspect of crash severity analysis is finding meaningful interactions between the explanatory and target variable (Li et al., 2008). Because of the complexity of the black-box machine learning algorithms (like SVM or Random Forest), it is difficult to interpret the impact of the input variables on the output of the model (Stewart, 2020). Interpretable Machine Learning is a technique that enables the machine learning models to explain their behaviors in understandable terms to humans (Doshi-Velez

& Kim, 2017). Researchers have proposed various interpretable Machine Learning techniques for inferring estimates of machine learning algorithms (Aas, Jullum, & Løland, 2021; Fish & Blodgett, 2003; Olden, Joy, & Death, 2004). Interpretable machine learning techniques are classified into global interpretability and local interpretability. Local interpretability explains a small sample of the whole population or works on each prediction. On the contrary, global interpretability describes how the model works with the inspection of model concepts. Global interpretation explains the whole population without taking any samples of the population (Du, Liu, & Hu, 2018; Jiang, Xie, & Ren, 2019; Molnar, 2019). Some of the local explanation methods are Explanation Vectors (Baehrens et al., 2010), Local Interpretable Model-agnostic Explanations or LIME (Ribeiro, Singh, & Guestrin, 2016), and Shapley values (Strumbel & Kononenko, 2010). However, for an overall explanation (global interpretation) of the model, techniques such as permutation feature importance, sensitivity analysis, and partial dependence plots (Molnar, 2019) are used.

Sensitivity analysis is a global interpretation method, which is performed by changing the value of one variable while keeping the values of the other variables constant (Delen, Sharda, & Bessonov, 2006; García-Herrero, Gutiérrez, Herrera, Azimian, & Mariscal, 2020; Parr, Turgutlu, Csiszar, & Howard, 2018; Tang, Liang, Han, Li, & Huang, 2019; Xie, Lord, & Zhang, 2007; Yu & Abdel-Aty, 2013). Thus, the sensitivity of a variable indicates the changes in crash severity prediction by various models. The more change in the proportion of the predicted classes, the more the models are sensitive to the variable. Sensitivity analysis increases the transparency of the model, enhances the interpretability or explainability of the model, and helps make necessary decisions (Guidotti et al., 2018; Saltelli, 2019). This study used sensitivity analysis to infer the impacts of factors that contribute to crash injury severity. The findings of sensitivity analysis would be useful to decision-makers in suggesting countermeasures for the contributing factors of work zone crashes.

5. Results and discussion

5.1. Variable selection

The input variables for machine learning models were selected based on the results of the correlation matrix and the feature importance scores. From the correlation matrix of the initially chosen 20 variables, a strong correlation (i.e., 0.7 or more) was observed between light conditions and time of day and between weather and surface conditions. However, from the feature importance scores, light conditions and weather conditions were found to have a higher score than the time of day and surface

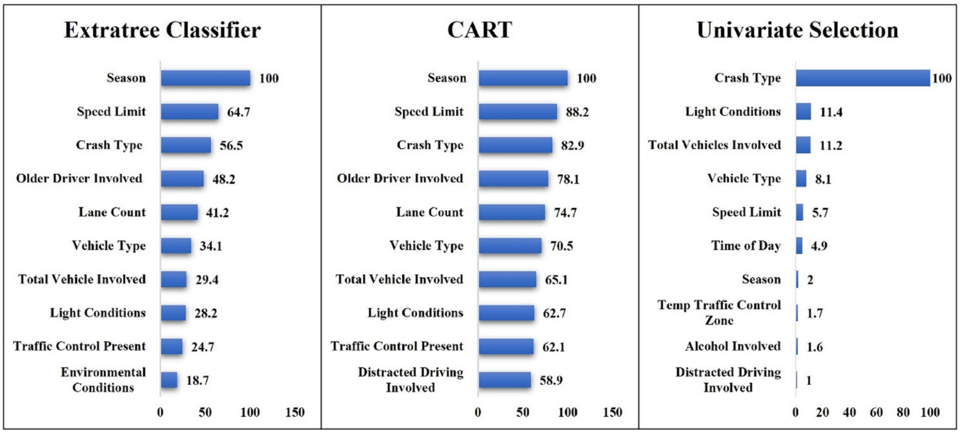

Figure 4. Feature importance of variables.

conditions, respectively. Hence, the time of day and surface conditions were removed from the dataset before training the model to avoid collinearity. After removing these two variables, a total of 18 input variables were selected to train and test the models for crash severity prediction.

The results from the three feature selection techniques are shown in Figure 4. Although the feature scores were obtained for all 20 variables, only the top 10 features are illustrated in Figure 4 for better visualization. It is noteworthy that the feature selection results from univariate selection method are different from the other two methods. For instance, 'crash type' had been identified as the most important feature in univariate selection, whereas 'season' was the top feature for the other two methods. Univariate feature selection is based on chi-squared test values, while the CART and extratree classifiers are based on tree-based decision-making. Because of the similarity in the methods, CART and extratree classifiers have generated similar rankings of the variables. On the other hand, the univariate selection method has different rankings of variables because of it's different approach for ranking the variables. The 'crash type' in univariate selection had a significantly higher value than other variables. Hence, as an individual feature, 'crash type' has the highest impact on the crash severity. One possible reason for this could be the high dependency of the crash severity on the type of crashes. These could be further explained from the exploratory data analysis in Table 3. For instance, angle (6.23%), fixed object (6.43%), and other (11.73%) types of crashes had a reasonably higher proportion of fatal crashes than the rear-end (2.77%) and sidewise (1.49%) crashes. On the contrary, only 65.58% of the crashes were PDO in angle crashes, whereas 90.70% of sidewise crashes were PDO. Because of a wide variety of crash severity proportions for various types of crashes, crash type has a significantly higher impact on the crash severity.

For performing the sensitivity analysis, the variables demonstrating higher feature importance in each of the three feature selection techniques (i.e., CART, extratree classifier, and univariate selection) are considered. Figure 4 shows the importance scores of the top ten variables from each feature selection technique. The feature scores are normalized with the best-performing variable set to 100 and the least important variable score set to 0. The importance scores of the variables in CART and extratree classifier indicate their contributions as key splitters of the regression tree in improving crash severity predictions (Banerjee, Arora, & Murty, 2008). On the other hand, variables with stronger dependency on crash severity have higher importance scores in univariate selection. Six features (crash type, season, speed limit, light conditions, vehicle type, and total number of vehicles involved) were repeated in the top 10 important features by each feature selection method (Figure 4). As a result, these six variables were chosen for performing sensitivity analysis in this research.

5.2. Model performance

Following feature selection, we used 18 variables as inputs to predict crash severity. A training and testing ratio of 70:30 was maintained for the crash severity prediction. Figure 5 shows the performances (accuracy, precision, recall) of all the machine learning models on the training dataset in predicting crash severities of work zone crashes. As shown in Figure 5, the overall accuracy of the training datasets varied from 78% to 90%, with RF having the highest accuracy and SVM having the least. Except for the SVM model, all of the models demonstrated good precision in the prediction of

(a) Predicted class

Actual class	NI	PI	FI	Recall
NI	4499 (78%)	0 (0%)	0 (0%)	100%
PI	1082 (19%)	0 (0%)	0 (0%)	0%
FI	180 (3%)	0 (0%)	0 (0%)	0%
Precision	78%	0%	0%	78% Accuracy

(b) Predicted class

Actual class	NI	PI	FI	Recall
NI	4435 (77%)	62 (1%)	2 (0%)	99%
PI	453 (8%)	626 (11%)	3 (0.1%)	58%
FI	54 (1%)	7 (0.1%)	119 (2%)	66%
Precision	90%	90%	96%	90% Accuracy

(c) Predicted class

Actual class	NI	PI	FI	Recall
NI	4473 (78%)	26 (0.5%)	0 (0%)	99%
PI	773 (13%)	308 (5%)	1 (0%)	28%
FI	105 (2%)	4 (0.1%)	71 (1%)	39%
Precision	84%	91%	99%	84% Accuracy

(d) Predicted class

Actual class	NI	PI	FI	Recall
NI	4481 (78%)	17 (0.3%)	1 (0%)	100%
PI	933 (16%)	149 (3%)	0 (0%)	14%
FI	120 (2%)	3 (0.1%)	57 (1%)	32%
Precision	81%	88%	98%	81% Accuracy

(e) Predicted class

Actual class	NI	PI	FI	Recall
NI	4465 (78%)	34 (1%)	0 (0%)	99%
PI	719 (12%)	363 (6%)	0 (0%)	34%
FI	99 (2%)	3 (0.1%)	78 (1%)	43%
Precision	85%	91%	100%	85% Accuracy

NI = No Injury
PI = Possible Injury
FI = Fatal & Injury

Figure 5. Confusion matrix of classification models of various ML techniques using the training set: (a) support vector machine, (b) random forest, (c) XGBoost, (d) LightGBM, and (e) Catboost.

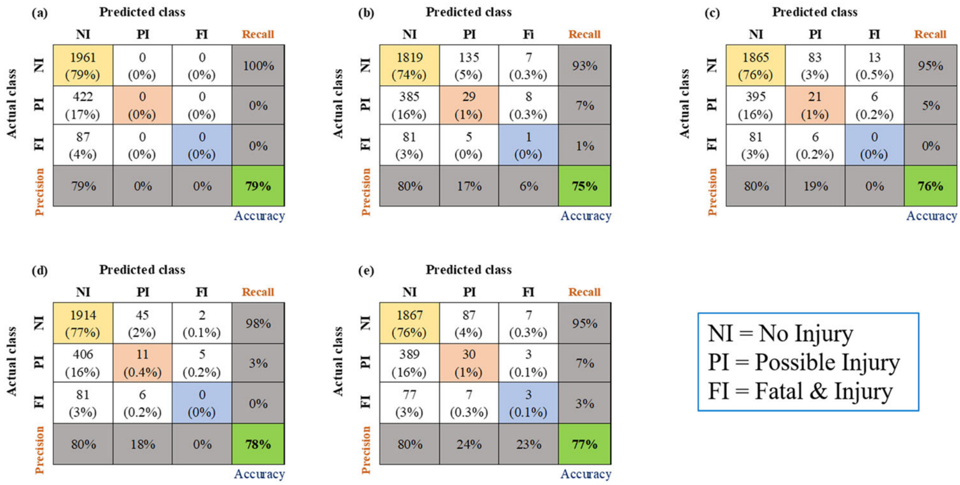

Figure 6. Confusion matrix of classification models (without SMOTE) of various ML techniques using the test set: (a) support vector machine, (b) random forest, (c) XGBoost, (d) LightGBM, and (e) CatBoost.

fatal and injury severity classes. The recall performance of RF was the best, followed by Catboost and XGBoost methods.

Figure 6 demonstrates the performances (accuracy, precision, recall) of all of the machine learning models in predicting crash severities on the test dataset. The overall accuracy of the test dataset ranged from 75% to 79%, with SVM having the highest accuracy and RF having the least. However, a detailed look into the confusion matrix showed that SVM did not accurately predict any crashes with fatal and injury or possible injury. For predicting the possible injury class, the boosting methods showed fair precision compared to Random Forest and SVM, with Catboost having the maximum precision of 24%. Among the five machine learning methods, only RF and Catboost accurately predicted some fatal and injury crashes, with Catboost having a maximum precision of 23%. Hence, comparing the performances of all these models, it is observed that RF and Catboost outperformed Support Vector Machine, XGBoost, and LightGBM. We note that SVM had the least accuracy among all studied models. These findings are consistent with some previous studies, where RF performed better than SVM and other models (logistic regression models, KNN, Decision Tree, and ordered probit model) in crash severity prediction (Ashqar et al., 2021; Zhang, Li, Pu, & Xu, 2018). However, there are also cases where SVM and other boosting methods perform better than RF. For instance, Hasan et al. (2021) found SVM to have higher accuracy in predicting injury crashes than RF and XGBoost. Similarly, Mansoor, Ratrout, Rahman, and Assi (2020) found SVM to perform similarly to AdaBoost in predicting injury crashes. It should be noted that in both studies, SVM performed worse than boosting methods and RF in predicting non-severe (possible injury and no injury) crashes. Nevertheless, the results of this study were just the opposite, where SVM could not predict severe crashes (fatal and injury crashes). These findings suggest that the model performance varies due to the variation of the dataset and the distribution of the severity classes in the dataset.

5.3. Sensitivity analysis of important variables

In previous studies, machine learning models were used to analyze the sensitivity of the significant explanatory variables (Fish & Blodgett, 2003; Li et al., 2008, 2012; Yu & Abdel-Aty, 2013). For the purposes of this study, we used the six variables (crash type, season, speed limit, light conditions, vehicle type, and total vehicles involved) obtained from the feature selection results. A Pearson correlation test was performed on the variables to find the correlation between the factors. Figure 7 shows a similarity matrix of the chosen variables. According to Pearson's correlation test, a correlation less than 0.3 is considered a low correlation while a correlation higher than 0.7 is considered a good correlation (Ratner, 2009). As seen in the

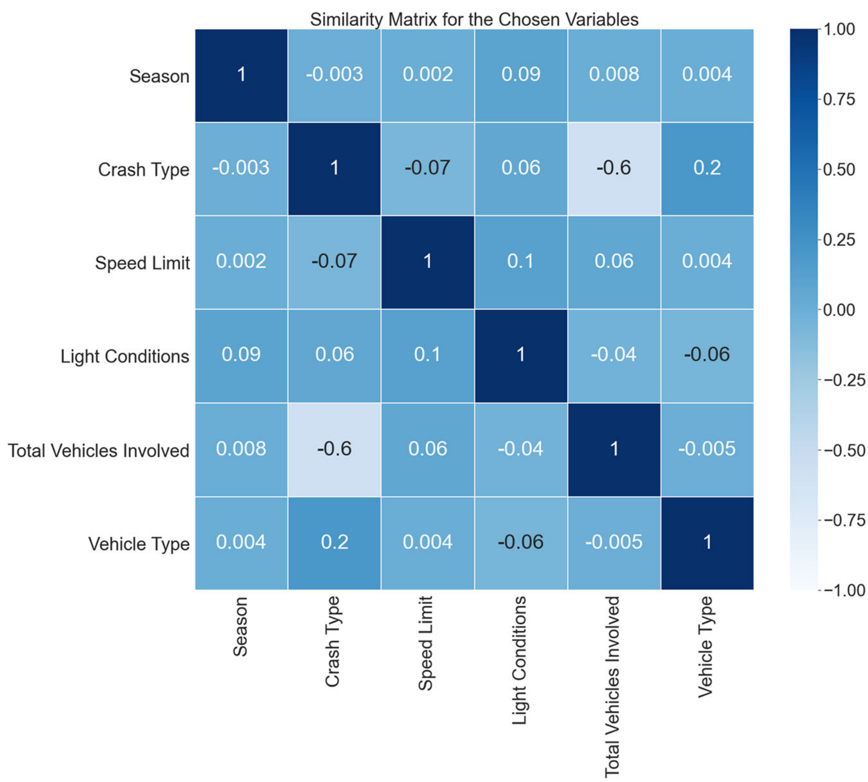

Figure 7. Correlation matrix of chosen variables.

figure, none of the variables had a strong correlation. In addition, five of those variables maintained a low correlation amongst themselves, making it clear that the variables themselves are not highly correlated. Due to their low intercorrelation and high importance scores, these parameters can be used in the sensitivity analysis.

Figure 8 shows the changes in the crash severity predictions of various models with the changes on a single variable. Each explanatory variable was changed by a user-defined amount while other variables remained unchanged. As all the variables in our dataset were converted to dummy variables and the input values for the variables were perturbed from 0 to 1. Afterward, the proportion of each severity level before and after the perturbation of a variable was recorded to calculate the variable impact on crash severity (Zhang et al., 2018).

Figure 8 indicates the sensitivity results obtained for the selected variables. The bars in different colors represent the changes in the predicted proportion of the reported severity classes due to changing one specific variable.

All the models have demonstrated an increase in the proportion of the fatal and injury crashes and a decrease in the no injury crashes due to the angular type of crash. The proportion of no injury crashes decreased (10%–15%), while the proportion of fatal and injury crashes increased (5%–15%) for angular types of crashes. Yang, Ozbay, Ozturk, and Xie (2015) also found that the proportion of angular crashes increased in the activity and termination areas of work zones.

The presence of 3 or more vehicles contributes to an increase in the proportion of both possible injury and fatal and injury crashes. The proportion of possible injury increases by 16% to 24%, while the proportion of the fatal and injury class increases by 1% to 10% due to the involvement of 3 or more vehicles in a work zone-related crash. Some previous studies also observed an increase in crash severity due to an increase in the total number of vehicles involved (Qi et al., 2005).

All of the models have demonstrated a rise in the proportion of the fatal and injury class for the speed limit of 55–65 mph, with SVM showing a maximum increase of the fatal and injury class by 6%. An increase in the speed limit makes it difficult to control and decelerate the vehicle in complex scenarios, which increases the likelihood of crashes. A previous study also found that a speed limit higher than 40 mph influences crash severity (Weng et al., 2016). In addition, according to the FHWA, the number of fatal work zone crashes where speed was a factor increased by nearly 40 percent in the year 2019 (FHWA, 2021).

The dark-lit condition increases the proportion of possible injury from 5% to 10%, while it decreases the proportion of both no injury and fatal

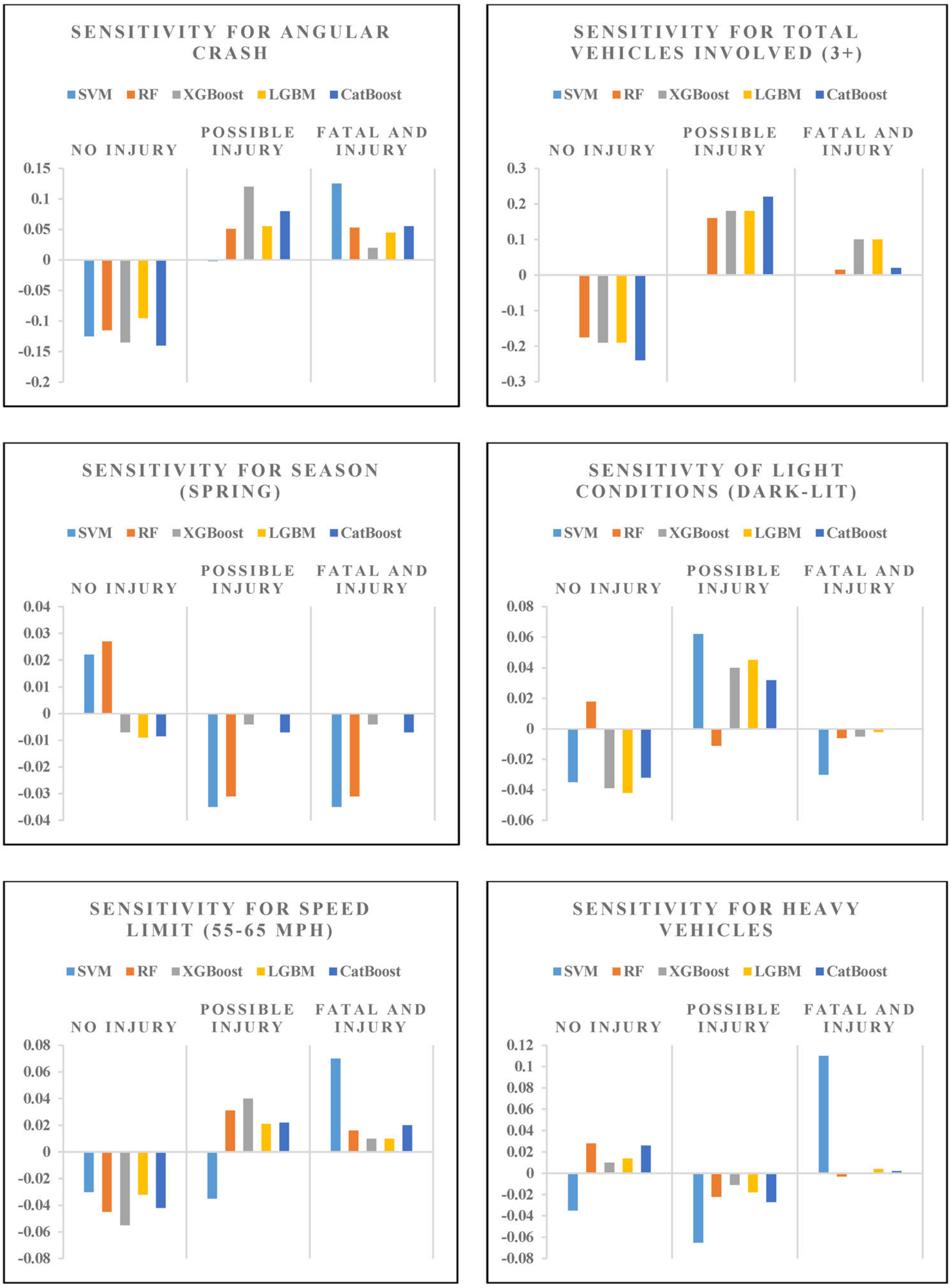

Figure 8. Sensitivity of the top contributing factors in predicting crash severity.

and injury crashes. Previous studies have reported that driving at nighttime or in a limited lighting condition deteriorates drivers' performance (Bai & Li, 2009; Katta, 2013).

The involvement of heavy vehicles demonstrated a decrease in possible injury from 2% to 6%, while it increased fatal and injury crashes from 1%

to 8%. According to Bai and Li (2009), the frequent involvement of heavy vehicles in crashes is a major safety concern in work zone areas. Fatal work zone crashes involving large trucks increased by 43 percent in a recent five–year period (2013–2017), which contributed to the overall increase in fatal work zone crashes across the nation (FHWA, 2019).

6. Conclusion

In this study, we evaluated the performance of the five different machine learning models in predicting the severity of recorded work zone crashes in New Jersey. We used three feature selection models (CART, univariate selection, extratree classifier) to identify significant variables and rank their importance with respect to crash severity. Then, we considered a 70:30 train test ratio to train and test the dataset with different machine learning models. Finally, we conducted a sensitivity analysis of the contributing factors to interpret the behavior of the model when the values of individual variables were changed. The results indicated that the crash type, season, number of vehicles involved, lighting conditions, type of vehicles involved, and the speed limit contributed significantly to the severity of work zone crashes. The prediction performances of Random Forest and Catboost were better than the Support Vector Machine, XGBoost, and LightGBM methods. Although all of the models worked better to predict the no injury class, only RF and Catboost were able to predict the fatal and injuries class. After we compared the performance of the models, our aim was also to determine the sensitivity of the features that have the greatest effect on crash severity, i.e., to identify the individual impacts of the most critical factors. Following a sensitivity analysis of the factors identified as contributing the most to crashes, we found that angular crashes, a speed limit of 55–65 mph, and heavy vehicles' involvement have the greatest sensitivity in increasing the severity of fatal and injury crashes. On the contrary, the spring season and dark-lit conditions show higher sensitivity in reducing the severity of fatal and injury crashes.

The outcomes generated by this model will help transportation safety practitioners develop and implement appropriate countermeasures to reduce the number of work zone crashes. For instance, more attention should be given to roads with higher speed limits during times when lighting conditions are poor or when large vehicles are present. Raising awareness to prevent drivers from speeding would also decrease the likelihood of work zone crashes. Installation of the Speed Display Signs would be a useful countermeasure to speeding in the work zone. Initiatives like pavement markings colored PCMS (Portable Changeable Message Signs) and overhead signs can be introduced to increase the effectiveness of the restrictions on the truck lanes. Technological initiatives such as navigation app alerts, HAAS alerts, and an Advance Traffic Management System (ATMS) could be useful to warn drivers before entering work zones (Macchione, Meehan, & Baker, 2020).

This study has three major contributions. First, this study applied five different machine learning models to classify the injury types of work zone crashes. Second, this study performed a detailed variable importance analysis and conducted a sensitivity analysis to make the models explainable with respect to the key variables. Third, some potential countermeasures were suggested based on the study outcomes.

Like other studies, this study has some inherent limitations. First, this study is limited to using three years of crash data. Second, the dataset used for this study contained some missing values that had to be removed prior to analysis, resulting in a reduction in the total used data. Third, the crash severity classes 'fatal', 'major injury', and 'minor injury' had to be merged as a single crash severity class due to the low proportion of these crashes. Fourth, the current study did not apply undersampling or oversampling to resolve the data imbalance issue. Future studies can focus on work zone crashes with a larger duration to address the aforementionee d limitations.

ORCID

Ahmed Sajid Hasan http://orcid.org/0000-0002-2874-6918
Mohammad Jalayer http://orcid.org/0000-0001-6059-3942
Subasish Das http://orcid.org/0000-0002-1671-2753

References

Aas, K., Jullum, M., & Løland, A. (2019). Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. *Artificial Intelligence*, *298*, 103502.

Abdel-Aty, M., Dilmore, J., & Dhindsa, A. (2006). Evaluation of variable speed limits for real-time freeway safety improvement. *Accident Analysis and Prevention*, *38*(2), 335–345.

Abdel-Aty, M., Pande, A., Lee, C., Gayah, V., & Santos, C. D. (2007). Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. *Journal of Intelligent Transportation Systems*, *11*(3), 107–120. doi:10.1080/15472450701410395

Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, M. F., & Hsia, L. (2004). Predicting freeway crashes from loop detector data by matched case-control logistic regression. *Transportation Research Record: Journal of the Transportation Research Board*, *1897*(1), 88–95. doi:10.3141/1897-12

Abou Elassad, Z. E., Mousannif, H., & Al Moatassime, H. (2020). A real-time crash prediction fusion framework: An imbalance-aware strategy for collision avoidance systems. *Transportation Research Part C: Emerging Technologies*, *118*, 102708. doi:10.1016/j.trc.2020.102708

Ahmadi, A., Jahangiri, A., Berardi, V., & Machiani, S. (2020). Crash severity analysis of rear-end crashes in California using statistical and machine learning classification methods. *Journal of Transportation Safety & Security*, *12*(4), 522–546. doi:10.1080/19439962.2018.1505793

Ahmed, M. M., Abdel-Aty, M., & Yu, R. (2012a). Assessment of interaction of crash occurrence, mountainous freeway geometry, real-time weather, and traffic data. *Transportation Research Record: Journal of the Transportation Research Board*, *2280*(1), 51–59. doi:10.3141/2280-06

Ahmed, M. M., Abdel-Aty, M., & Yu, R. (2012b). Bayesian updating approach for real-time safety evaluation with automatic vehicle identification data. *Transportation Research Record: Journal of the Transportation Research Board*, *2280*(1), 60–67. doi:10.3141/2280-07

Ahmed, M. M., Ghasemzadeh, A., Eldeeb, H., Gaweesh, S., Clapp, J., Ksaibati, K., & Young, R. (2015). *Driver performance and behavior in adverse weather conditions: An investigation using the SHRP2 naturalistic driving study data–phase 1*. Laramie, Wyoming: FHWA.

Ashqar, H. I., Shaheen, Q. H., Ashur, S. A., & Rakha, H. A. (2021). Impact of risk factors on work zone crashes using logistic models and Random Forest. In 2021 IEEE International Intelligent Transportation Systems Conference (ITSC) (pp. 1815-1820). IEEE.

Baehrens, D., Schroeter, T., Harmeling, S., Kawanabe, M., Hansen, K., & Muller, K. (2010). How to explain individual classification decisions. *Journal of Machine Learning Research*, *11*, 1803–1831.

Bai, Y., & Li, Y. (2009). *Reducing Work Zone Crashes by Using Vehicle's Warning Flashes as a Warning Sign (Report No.K-TRAN: KU-07-3)*. University of Kansas Center for Research, Inc., Lawrence, Kansas.

Banerjee, A., Arora, N., & Murty, U. S. (2008). Classification and regression tree (CART) analysis for deriving variable importance of parameters influencing average flexibility of CaMK kinase family. *Electronic Journal of Biology*, *4*(1), 27–33.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992 *A training algorithm for optimal margin classifiers*. Proceedings of the Fifth Annual Workshop on Computational Learning Theory, 130401. USA: ACM. doi:10.1145/130385.130401

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. doi:10.1023/A:1010933404324

Bryden, J., Andrew, L., & Fortuniewicz, J. (1998). Work zone traffic accidents involving traffic control devices, safety features, and construction operations. *Transportation Research Record: Journal of the Transportation Research Board*, *1650*(1), 71–81. doi:10.3141/1650-09

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *2*(3), 27.

Chang, L., & Chen, W. (2005). Data mining of tree-based models to analyze freeway accident frequency. *Journal of Safety Research*, *36*(4), 365–375.

Chang, L. Y., & Chien, J. T. (2013). Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. *Safety Science*, *51*(1), 17–22. doi:10.1016/j.ssci.2012.06.017

Chen, C., Zhang, G., Qian, Z., Tarefder, R. A., & Tian, Z. (2016a). Investigating driver injury severity patterns in rollover crashes using support vector machine models. *Accident Analysis and Prevention*, *90*, 128–139.

Chen, C., Zhang, G., Tarefder, R., Ma, J., Wei, H., & Guan, H. (2015a). A multinomial logit model-Bayesian network hybrid approach for driver injury severity analyses in rear-end crashes. *Accident Analysis and Prevention*, *80*, 76–88.

Chen, C., Zhang, G., Wang, H., Yang, J., Jin, P. J., & Michael Walton, C. (2015b). Bayesian network-based formulation and analysis for toll road utilization supported by traffic information provision. *Transportation Research Part C: Emerging Technologies*, *60*, 339–359. doi:10.1016/j.trc.2015.09.005

Chen, C., Zhang, G., Yang, J., Milton, J. C., & Alcántara, A. D. (2016b). An explanatory analysis of driver injury severity in rear-end crashes using a decision table/Naïve Bayes (DTNB) hybrid classifier. *Accident Analysis and Prevention*, *90*, 95–107.

Chen, S., Wang, W., & Zuylen, H. V. (2009). Construct support vector machine ensemble to detect traffic incident. *Expert Systems with Applications*, *36*(8), 10976–10986. doi:10.1016/j.eswa.2009.02.039

Cheng, Y., Parker, S. T., Ran, B., & Noyce, D. A. (2016). Work zone crash cost prediction with a least median squares linear regression model. *Transportation Research Record: Journal of the Transportation Research Board*, *2555*(1), 38–45. doi:10.3141/2555-05

Cheu, R., Xu, J., Kek, A., Lim, W., & Chen, W. (2006). Forecasting shared-use vehicle trips with neural networks and support vector machines. *Transportation Research Record: Journal of the Transportation Research Board*, *1968*(1), 40–46. doi:10.1177/0361198106196800105

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi:10.1007/BF00994018

Daniel, J., Dixon, K., & Jared, D. (2000). Analysis of fatal crashes in Georgia work zones. *Transportation Research Record: Journal of the Transportation Research Board*, *1715*(1), 18–23. doi:10.3141/1715-03

Delen, D., Sharda, R., & Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis and Prevention*, *38*(3), 434–444.

Delen, D., Tomak, L., Topuz, K., & Eryarsoy, E. (2017). Investigating injury severity risk factors in automobile crashes with predictive analytics and sensitivity analysis methods. *Journal of Transport & Health*, *4*, 118–131. doi:10.1016/j.jth.2017.01.009

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *Stat*, *1050*, 2.

Du, M., Liu, N., Hu, X. (2018). *Techniques for interpretable machine learning*. Technical report. Retrieved from http://arxiv.org/abs/1808.00033

Elghamrawy, T. M. (2011). *Optimizing work zone practices for highway construction projects*. University of Illinois at Urbana-Champaign. Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/24344/Elghamrawy_Tarek.pdf?sequence=1

FHWA. (2016). *KABCO injury classification scale and definitions*. Retrieved January 20, 2021, from https://safety.fhwa.dot.gov/hsip/spm/conversion_tbl/pdfs/kabco_ctable_by_state.pdf

FHWA. (2019). *Work zone safety: Drive like you work here*. Retrieved January 31, 2021, from https://ops.fhwa.dot.gov/publications/fhwahop19027/fhwahop19027.pdf

FHWA. (2021). *National workzone awareness week*. Retrieved from https://ops.fhwa.dot.gov/wz/outreach/nwzaw_factsheet_2021/nwzaw_factsheet_2021.pdf

Fish, K., & Blodgett, J. (2003). A visual method for determining variable importance in an artificial neural network model: An empirical benchmark study. *Journal of Targeting, Measurement and Analysis for Marketing*, *11*(3), 244–254. doi:10.1057/palgrave.jt.5740081

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232.

Gang, R., & Zhuping, Z. (2011). Traffic safety forecasting method by particle swarm optimization and support vector machine. *Expert Systems with Applications*, *38*(8), 10420–10424. doi:10.1016/j.eswa.2011.02.066

Garber, N., & Woo, T.-S. (1990). *Accident characteristics at construction and maintenance zones in urban areas (No. VTRC 90-R12)*. Virginia Transportation Research Council, Charlottesville, Virginia.

Garber, N., & Zhao, M. (2002). *Crash characteristics at work zones*. Charlottesville, Virginia: Virginia Transportation Research Council.

García-Herrero, S., Gutiérrez, J. M., Herrera, S., Azimian, A., & Mariscal, M. A. (2020). Sensitivity analysis of driver's behavior and psychophysical conditions. *Safety Science*, *125*, 104586. doi:10.1016/j.ssci.2019.104586

Ghasemzadeh, A., & Ahmed, M. M. (2019a). Complementary parametric probit regression and nonparametric classification tree modeling approaches to analyze factors affecting severity of work zone weather-related crashes. *Journal of Modern Transportation*, *27*(2), 129–140. doi:10.1007/s40534-018-0178-6

Ghasemzadeh, A., & Ahmed, M. M. (2019b). Exploring factors contributing to injury severity at work zones considering adverse weather conditions. *IATSS Research*, *43*(3), 131–138. doi:10.1016/j.iatssr.2018.11.002

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, *51*(5), 1–42. doi:10.1145/3236009

Hall, J., & Lorenz, V. (1989). Characteristics of construction-zone accidents. *Transportation Research Record*, *1230*, 20–27.

Hamzeie, R., Savolainen, P., & Gates, T. (2017). Driver speed selection and crash risk: Insights from the naturalistic driving study. *Journal of Safety Research*, *63*, 187–194.

Hasan, A. S., Kabir, M. A., & Jalayer, M. (2021). Severity analysis of heavy vehicle crashes using machine learning models: A case study in New Jersey. In International Conference on Transportation and Development 2021, 285–296. doi:10.1061/9780784483534.025

Heneghan, J. (2018, October 12). *Work zone speed limits clarified*. Retrieved January 31, 2021, from https://bordaslaw.com/blog/work-zone-speed-limits-clairified

Hossain, M., & Muromachi, Y. (2013). Understanding crash mechanism on urban expressways using high-resolution traffic data. *Accident Analysis and Prevention*, *57*, 17–29.

Islam, M., Alnawmasi, N., & Mannering, F. (2020). Unobserved heterogeneity and temporal instability in the analysis of work-zone crash-injury severities. *Analytic Methods in Accident Research*, *28*, 100130. doi:10.1016/j.amar.2020.100130

Jiang, L., Xie, Y., Ren, T. (2019). *Modelling highly unbalanced crash injury severity data by ensemble methods and global sensitivity analysis*. Proceedings of the Transportation Research Board 98th Annual Meeting (pp. 13–17). Washington, DC.

Kashani, A. T., & Mohaymany, A. S. (2011). Analysis of the traffic injury severity on two-lane, two-way rural roads based on classification tree models. *Safety Science*, *49*(10), 1314–1320. doi:10.1016/j.ssci.2011.04.019

Kassmeyer, C. (2020). Work zone crash prediction model and characteristics analysis. Master of Science Thesis.

Katta, V. (2013). Development of *crash severity model for predicting risk factors in work zones for* Ohio. Doctoral dissertation, University of Toledo.

Khattak, A., & Targa, F. (2004). Injury severity and total harm in truck-involved work zone crashes. *Transportation Research Record: Journal of the Transportation Research Board*, *1877*(1), 106–116. doi:10.3141/1877-12

Khattak, A., Rodriguez, D., Targa, F., & Rocha, M. (2003). *Understanding the role of truck-driver. Occupational and high-risk roadway factors in truck-involved collisions*. Southeastern Transportation Center, University of Tennessee, Knoxville, Tennessee.

Koilada, K., Mane, A., & Pulugurtha, S. (2020). Odds of work zone crash occurrence and getting involved in advance warning, transition, and activity areas by injury severity. *IATSS Research*, *44*(1), 75–83. doi:10.1016/j.iatssr.2019.07.003

Kuhnert, P. M., Do, K. A., & McClure, R. (2000). Combining non-parametric models with logistic regression: An application to motor vehicle injury data. *Computational Statistics & Data Analysis*, *34*(3), 371–386. doi:10.1016/S0167-9473(99)00099-7

Li, X., Lord, D., Zhang, Y., & Xie, Y. (2008). Predicting motor vehicle crashes using support vector machine models. *Accident Analysis and Prevention*, *40*(4), 1611–1618.

Li, Y., & Bai, Y. (2009). Highway work zone risk factors and their impact on crash severity. *Journal of Transportation Engineering*, *135*(10), 694–701. doi:10.1061/(ASCE)TE.1943-5436.0000055

Li, Z., Liu, P., Wang, W., & Xu, C. (2012). Using support vector machine models for crash injury severity analysis. *Accident Analysis and Prevention*, *45*, 478–486.

Liu, J., Khattak, A., & Zhang, M. (2016). What role do precrash driver actions play in work zone crashes? Application of hierarchical models to crash data. *Transportation Research Record: Journal of the Transportation Research Board*, *2555*(1), 1–11. doi:10.3141/2555-01

Lv, Y., Tang, S., Zhao, H., & Li, S. (2009). *Real-time highway accident prediction based on support vector machines*. In Chinese Control and Decision Conference (pp. 4403–4407). IEEE.

Macchione, T., Meehan, K., & Baker, M. (2020). *A comprehensive management approach to improving work zone safety performance*. 2020 MASITE Annual Meeting.

Mansoor, U., Ratrout, N. T., Rahman, S. M., & Assi, K. (2020). Crash severity prediction using two-layer ensemble machine learning model for proactive emergency management. *IEEE Access*. *8*, 210750–210762. doi:10.1109/ACCESS.2020.3040165

Meng, Q., & Weng, J. (2011). A genetic algorithm approach to assessing work zone casualty risk. *Safety Science*, *49*(8–9), 1283–1288. doi:10.1016/j.ssci.2011.05.001

Meng, Q., Weng, J., & Qu, X. (2010). A probabilistic quantitative risk assessment model for the long-term work zone crashes. *Accident Analysis and Prevention*, *42*(6), 1866–1877. doi:10.1016/j.aap.2010.05.007

Mokhtarimousavi, S., Anderson, J. C., Azizinamini, A., & Hadi, M. (2019). Improved support vector machine models for work zone crash injury severity prediction and analysis. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(11), 680–692. doi:10.1177/0361198119845899

Molnar, C. (2019). *Interpretable machine learning - A guide for making black box models explainable.* Retrieved January 31, 2021, from https://christophm.github.io/interpretable-ml-book/

Montella, A., Aria, M., D'Ambrosio, A., & Mauriello, F. (2012). Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery. *Accident Analysis and Prevention*, *49*, 58–72.

Nasrollahzadeh, A. A., Sofi, A. R., & Ravani, B. (2021). Identifying factors associated with roadside work zone collisions using machine learning techniques. *Accident Analysis and Prevention*, *158*, 106203.

National Work Zone Safety Information Clearinghouse. (2019). *Work zone fatal crashes and fatalities.* Retrieved January 31, 2021, from https://www.workzonesafety.org/crash-information/work-zone-fatal-crashes-fatalities/#new%20jersey

Nemeth, Z. A., & Migletz, D. J. (1978). Accident characteristics before, during, and after safety upgrading projects on Ohio's rural interstate system. *Transportation Research Record*, *672*, 19–23.

New Jersey Department of Transportation. (2018). *Weekday average speed data*. Bureau of Transportation Data and Support Traffic and Technology Section. Retrieved from https://www.state.nj.us/transportation/refdata/roadway/pdf/wim/18_Classdata_spd.pdf

Olden, J. D., Joy, M. K., & Death, R. G. (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*, *178*(3–4), 389–397. doi:10.1016/j.ecolmodel.2004.03.013

Osman, M., Paleti, R., & Mishra, S. (2018). Analysis of passenger-car crash injury severity in different work zone configurations. *Accident Analysis and Prevention*, *111*, 161–172.

Osman, M., Paleti, R., Mishra, S., & Golias, M. (2016). Analysis of injury severity of large truck crashes in work zones. *Accident Analysis and Prevention*, *97*(December), 261–273. doi:10.1016/j.aap.2016.10.020

Parr, T., Turgutlu, K., Csiszar, C., & Howard, J. (2018). *Beware default random forest importances*. Retrieved January 31, 2021, from https://explained.ai/rf-importance/

Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. K. (2020). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accident Analysis and Prevention*, *136*, 105405.

Pedregosa, F. V. (2011). Scikitlearn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Prichard, B. (2015, May 10). *Exploration of crash severity and location in work zone-related crashes in Alabama.* Retrieved from https://etd.auburn.edu/bitstream/handle/10415/4473/Bradley%20Prichard_1-7-15.pdf?sequence=2&isAllowed=y

Qi, Y., Srinivasan, R., Teng, H., & Baker, R. (2005). *Frequency of work zone accidents on construction projects*. University Transportation Research Center, The City College of New York, New York.

Ratner, B. (2009). The correlation coefficient: Its values range between +1/−1, or do they? *Journal of Targeting, Measurement and Analysis for Marketing*, *17*(2), 139–142. doi:10.1057/jt.2009.5

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). *Why should I trust you? Explaining predictions of any classifier*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). ACM.

Salem, O. M., Genaidy, A. M., Wei, H., & Deshpande, N. (2006). *Spatial distribution and characteristics of accident crashes at work zones of interstate freeways in Ohio*. IEEE Intelligent Transportation Systems Conference, 1642–1647.

Saltelli, A. (2019). A short comment on statistical versus mathematical modelling. *Nature Communications*, *10*(1), 1–3. Retrieved from https://www.nature.com/articles/s41467-019-11865-8 doi:10.1038/s41467-019-11865-8

See, C. (2005). *Crash analysis of work zone lane closures with left-hand merge and downstream lane shift.* Kansas City: University of Missouri. Retrieved from https://kuscholarworks.ku.edu/bitstream/handle/1808/4162/umi-ku-2645_1.pdf;jsessionid=C8219F2411B5072F28BE6811117FFA1B?sequence=1

Siddiqui, C., Abdel-Aty, M., & Huang, H. (2012). Aggregate nonparametric safety analysis of traffic zones. *Accident Analysis and Prevention*, *45*, 317–325. doi:10.1016/j.aap.2011.07.019

Stewart, M. (2020). *Guide to interpretable machine learning*. Retrieved March 31, 2021, from https://towardsdatascience.com/guide-to-interpretable-machine-learning-d40e8a64b6cf

Strumbel, E., & Kononenko, I. (2010). An efficient explanation of individual classifications using game theory. *Journal of Machine Learning Research*, *11*, 1–18.

Sze, N. N., & Song, Z. (2019). Factors contributing to injury severity in work zone related crashes in New Zealand. *International Journal of Sustainable Transportation*, *13*(2), 148–154. doi:10.1080/15568318.2018.1452083

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.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer2V Verlag.

Weng, J., & Meng, Q. (2011). Analysis of driver casualty risk for different work zone types. *Accident Analysis and Prevention*, *43*(5), 1811–1817. doi:10.1016/j.aap.2011.04.016

Weng, J., & Meng, Q. (2012). Effects of environment, vehicle and driver characteristics on risky driving behavior at work zones. *Safety Science*, *50*(4), 1034–1042. doi:10.1016/j.ssci.2011.12.005

Weng, J., Meng, Q., & Wang, D. (2013). Tree-based logistic regression approach for work zone casualty risk assessment. *Risk Analysis*, *33*(3), 493–504.

Weng, J., Zhu, J.-Z., Yan, X., & Liu, Z. (2016). Investigation of work zone crash casualty patterns using association rules. *Accident Analysis and Prevention*, *92*, 43–52.

Work/Construction Zones. (2008). Retrieved January 31, 2021, from https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/workzones.pdf

Xie, Y. C., Lord, D., & Zhang, Y. L. (2007). Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis. *Accident Analysis and Prevention*, *39*(5), 922–933.

Yang, H., Ozbay, K., Ozturk, O., & Yildirimoglu, M. (2013). Modeling work zone crash frequency by quantifying measurement errors in work zone length. *Accident Analysis & Prevention*, *55*, 192–201. doi:10.1016/j.aap.2013.02.031

Yang, H., Ozbay, K., Ozturk, O., & Xie, K. (2015). Work zone safety analysis and modeling: A state-of-the-art review. *Traffic Injury Prevention*, *16*(4), 387–396.

Yu, R., & Abdel-Aty, M. (2013). Utilizing support vector machine in real-time crash risk evaluation. *Accident Analysis and Prevention*, *51*, 252–259.

Zhang, J., Li, Z., Pu, Z., & Xu, C. (2018). Comparing prediction performance for crash injury severity among various machine learning and statistical methods. *IEEE Access*, *6*, 60079–60087. doi:10.1109/ACCESS.2018.2874979

Zhang, Y., & Xie, Y. (2007). Forecasting of short-term freeway volume with v-support vector machines. *Transportation Research Record: Journal of the Transportation Research Board*, *2024*(1), 92–99. doi:10.3141/2024-11

Zheng, Z., Lu, P., & Lantz, B. (2018). Commercial truck crash injury severity analysis using gradient boosting data mining model. *Journal of Safety Research*, *65*, 115–124.