safety

MDPI

Article

Application of the Apriori Algorithm for Traffic Crash Analysis in Thailand

Ittirit Mohamad [1], Rattanaporn Kasemsri [2], Vatanavongs Ratanavaraha [3] and Sajjakaj Jomnonkwao [3,*]

1 Program of Energy and Logistics Management Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; ittirit.mohamad@gmail.com
2 School of Civil Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; kasemsri@sut.ac.th
3 School of Transportation Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; vatanavongs@g.sut.ac.th
* Correspondence: sajjakaj@g.sut.ac.th

Abstract: Accidents pose significant obstacles to economic progress and quality of life, especially in developing countries. Thailand faces such challenges and this research seeks to assess the frequency and most common causes of road accidents that lead to fatalities. This study employed the Apriori algorithm to examine the interrelationships among factors contributing to accidents in order to inform policymaking for reducing accident rates, minimizing economic and human losses, and enhancing the effectiveness of the healthcare system. By analyzing road accident data from 2015 to 2020 in Thailand (167,820 accidents causing THB 1.13 billion in damages), this article specifically focuses on the drivers responsible for fatal highway accidents. The findings reveal several interconnected variables that heighten the likelihood of fatalities, such as male gender, exceeding speed limits, riding a motorbike, traveling on straight roads, encountering dry surface conditions, and clear weather. An association rule analysis underscores the increased risk of injury or death in traffic accidents.

Keywords: data mining; predictive analysis; Apriori algorithm; machine learning; associated rule; BDA; crash frequency; crash injury; road safety

check for updates

Citation: Mohamad, I.; Kasemsri, R.; Ratanavaraha, V.; Jomnonkwao, S. Application of the Apriori Algorithm for Traffic Crash Analysis in Thailand. *Safety* **2023**, *9*, 58. https://doi.org/10.3390/safety9030058

Academic Editor: Raphael Grzebieta

Received: 6 July 2023
Revised: 22 August 2023
Accepted: 23 August 2023
Published: 28 August 2023

1. Introduction

Road traffic accidents are a worldwide issue that has been troubling civilization for a long time. Specifically, road accidents in Southeast Asia and Africa have been continuously increasing for at least the last 10 years (2008–2018) [1]. According to WHO data, in 2018, Thailand was ranked number one for road accidents in Asia and number nine in the world. An average of 32.7 Thais per 100,000 die in road accidents every year [1]. Not only has this caused an economic upheaval but it has impacted the country's public health system. Road accidents have also caused the country's limited resources to be used in ways that are harmful to its progress. They negatively impact the country's human resources, resulting in the death or disability of its residents.

In Thailand, examples of road safety policies include law enforcement (e.g., for exceeding speed limits or the consumption of alcohol), road safety programs in educational institutions, the development of advertising media, an increase in the number of training hours required to obtain new drivers' licenses and their renewals, engineering solutions for road safety audits, and research funding. To establish these regulations, predicted data on the number of accidents were used to determine operational budgets [2]. However, the average number of roadway fatalities in Thailand from 2015 to 2020 remained consistent at 32–35% for the sixth year in a row, as shown in Figure 1. The existing policy appears to be ineffective. Learning from every element recorded in a big dataset and starting to predict and minimize accidents before they occur might be the way out. Previous studies have utilized machine learning algorithms to predict injury severity. Some focused on independent factors, such as the environment, drivers, current weather, and road conditions, and even compared performance models, as shown in Table 1.

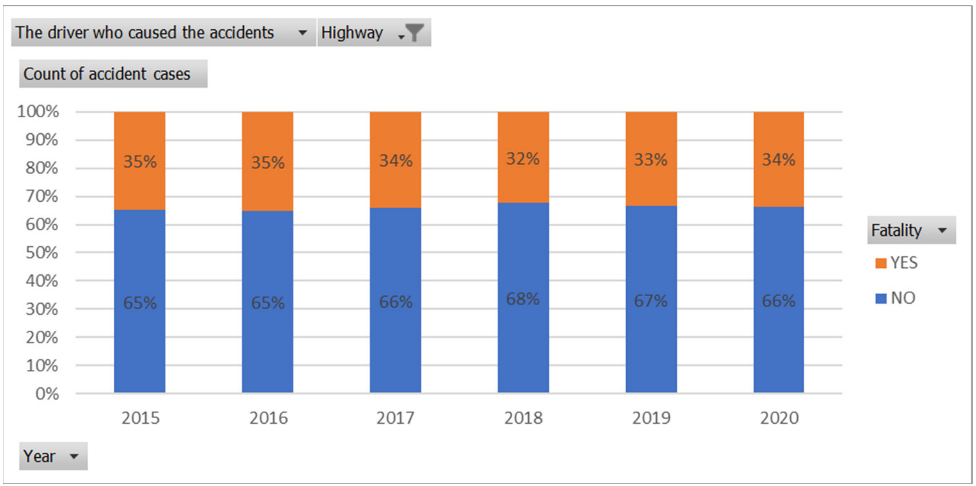

Figure 1. Stacked column chart illustrating highway accidents in Thailand over the years.

Against the backdrop of Thailand's road safety policies and the persistent fatality rate, several pivotal research questions have emerged. Initially, the efficacy of the current policies comes into focus. Can big data analysis enhance road safety predictions? Moreover, the role of machine learning algorithms in accounting for concurrent elements contributing to fatalities warrants exploration. How can the alignment of risk factors be effectively managed? Additionally, what are the implications of accurate accident prediction for road safety planning and policy formulation? Lastly, the role of comprehensive factor analysis in deepening insights into accidents and guiding interventions merits consideration. These inquiries collectively illuminate policy effectiveness, the potential of predictive analytics, the intricacies of accident causation, and avenues for refining road safety strategies.

However, these studies did not consider the conditions of the events for the drivers who were killed. The conditions being discussed included the type of roadway; vehicle type; external factors, such as environmental and weather conditions; and internal factors, e.g., driver behaviors and information, such as gender and age; these can be used to understand which factors interfere with each other or any linkage between them that increases the chances of fatality. According to the Swiss cheese theory [3], if all of the holes (factors) are aligned by chance, an accident will happen and result in death. In contrast, the risk may be decreased by controlling the primary element that has the strongest influence on fatality. For example, a researcher noted that accidents are typically caused by a combination of circumstances rather than by one or two factors. In addition, if the elements were combined, how likely is it that someone would die? However, what happens if a risk factor is reduced? This is why forecasts have appeared in order to simulate situations. However, predicting an accident event is also essential for establishing road safety, budgeting, staffing, and policy planning.

Table 1. Studies of road accidents using data mining and machine learning.

	Methodology														
Author	**Apriori Algorithm**	**Associated Rule**	**Bayesian Logistic**	**Cluster Analysis**	**Decision Tree**	**Deep Learning**	**Gradient Boosting**	**K-means**	**K-Nearest Neighbor**	**Multinomial Logistic Regression**	**Neural Network**	**Naïve Bayes**	**Random Forest**	**Regression on python**	**Support Vector Machine**
Sonal and Suman [4]	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-
Gutierrez-Osorio and Pedraza [5]	-	-	-	-	-	✓	-	-	-	-	✓	-	-	-	-
Abellán, et al. [6]	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-
Al Mamlook, et al. [7]	-	-	✓	✓	✓	-	-	-	✓	-	-	✓	✓	-	✓
Mafi, et al. [8]	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-
Recal and Demirel [9]	-	-	-	-	✓	-	✓	-	-	✓	✓	-	-	-	✓
Bahiru, et al. [10]	-	-	-	-	✓	-	-	-	-	-	-	✓	-	-	-
Cuenca, et al. [11]	-	-	-	-	-	✓	✓	-	-	-	-	✓	-	-	-
Kuşkapan, et al. [12]	-	-	-	-	-	-	-	-	✓	-	-	✓	-	-	✓
Ospina-Mateus, et al. [13]	-	-	-	-	✓	-	-	-	✓	-	✓	✓	✓	-	✓
Kumar and Toshniwal [14]	-	✓	-	-	-	-	-	✓	-	-	-	-	-	-	-
Helen, et al. [15]	-	✓	-	-	-	-	-	✓	-	-	-	-	-	-	-
El Abdallaoui, et al. [16]	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-
Guido, et al. [17]	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	✓
John and Shaiba [18]	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Feng, et al. [19]	-	✓	-	-	-	-	-	-	-	-	✓	-	-	-	-
Bhavsar, et al. [20]	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-
Samerei, et al. [21]	-	✓	-	✓	-	-	-	-	-	-	-	-	-	-	-
John and Shaiba [22]	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Guido, et al. [23]	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-

Earlier research on road traffic accidents also categorized them according to variables in forms that were presumed to be associated with every accident, according to international research.

Age: Zhang and Fan [24] found that accidents are more likely to occur among junior drivers (≤25 years) who have a lack of discipline, are inexperienced with traffic regulations, and have less driving experience. The majority of traffic accidents in Dubai are caused by a lack of space between vehicles, with the youth (≤35 years) being the most commonly involved; the peak hours are late at night and the overwhelming majority of drivers have been discovered to be inebriated [18]. Young (18–24 years old) drivers lack experience in controlling their speed and adjusting well while driving [25]. John and Shaiba [22] found that the majority of alcohol-involving accidents are caused by youths (≤35 years) late at night.

Gender: Ospina-Mateus, et al. [26], Mohamad, et al. [27] observed that men are more likely to be involved in serious accidents than women.

Driver behaviors: When compared to other drivers, intoxicated drivers have a higher accident rate [15]. The most important aspect in predicting the severity of an injury is driving over the speed limit [7].

Driver: Drivers are more likely to be injured or killed in accidents than other passengers [16].

Time: Traveling at night increases the chances of car accidents [28].

Road and light conditions: Chen, et al. [29] observed that road slope and visibility were predictors of driver injuries. Highway intersections are riskier for all accident types. Poor road conditions increase the likelihood of accidents, especially on motorways [30]. The road type, lighting, speed limits, and road surface all play key roles in accident incidence [19]. The majority of fatal injuries occur as a result of aggressive driving, inattentiveness, and speeding. However, compared with other situations, dark or dim roads also play significant roles [31].

Weather conditions: [14] Sonal and Suman [4] observed that external factors, including weather conditions, such as fog, rain, and snow, have greater impacts on road accidents than internal factors, such as the driver.

Types of vehicles: Chen, et al. [32] mentioned that this factor is significant for driver injuries and fatalities in rear-end accidents involving trucks, lighting, wind, and multiple vehicles. The analysis revealed that the most essential and impactful traffic accident elements are speed limit, weather conditions, number of lanes, lighting conditions, and accident timing while gender, age, accident location, and vehicle type have less of an impact on severity [10].

Researchers are continuing to evaluate the literature on road accidents and the factors involved. Though a wide range of research from across the world has been covered, Table 2 concentrates on research from the same region as that in this study.

Table 2. Previous research identifying the factors that determine the severity of driving injuries.

Variables	Findings
(1) Driver Characteristics	
Gender	Decreased injury severity: • male [33–35] Increased injury severity: • female [34,36–38] • male [39–41]
Age	Decreased injury severity: • younger than 25 [34,42] Increased injury severity: • older than 65 [36,38,39,41,43,44] • younger than 25 [40]
Speeding	Increased injury severity: • speeding vehicle [33,37,39,45,46]
Drunk	Increased injury severity: • drunk driving [15,18,33,36,39,41,43,45]
Fatigued	Increased injury severity: • dozed off [41]
Overtaking	Increased injury severity: • improper overtaking [35,47]

Table 2. *Cont.*

Variables	Findings
(2) Vehicle Characteristics	
Vehicle type	Decreased injury severity: • SUV/van [48] • pick-up truck [36,48] • passenger car [49] Increased injury severity: • rollover SUV/van [47] • large truck [40,47,49] • pick-up [40] • motorcycle [24]
(3) External Factors (environmental and road conditions)	
Light status	Decreased injury severity: • darkness without light [33] Increased injury severity: • darkness without light [39,43,47] • daylight [45] • after midnight [43] • nighttime [28,37]
Dry/wet road surface	Decreased injury severity: • wet road [43,50] Increased injury severity: • wet road [35,40] • dry road [24,45]
Weather	Decreased injury severity: • rain [51] Increased injury severity: • rain [31,35,47,52] • fog, rainfall, snowfall [31]
Time	Increased injury severity: • daytime [53] • nighttime [41,48]

2. Data Description and Methodology

2.1. Data Description

According to data collected by a government organization in Thailand, there were a total of 167,820 road accidents reported between 2015 and 2020 PDPM [54]. This study specifically examines the subset of drivers who were responsible for their accidents, totaling 129,015 accidents, of which 95,249 were non-fatal and 33,766 were fatal (24,559 on highways and 9207 on non-highways). To analyze this dataset, this research employed a data analysis technique involving the steps outlined in Figure 2:

- Data cleaning: missing and incompletely captured data were detected and corrected;
- Data validation: the quality of the data was assessed after the dataset was cleaned was validated;
- Data conversion: to facilitate data analysis, the data were partitioned into binary mode;
- Data analysis and interpretation: the research team then conducted an in-depth analysis of the data to uncover insights that would inform their conclusions;

– Data visualization: data visualization techniques were used to create a visual representation of the information and data for enhanced clarity and understanding.

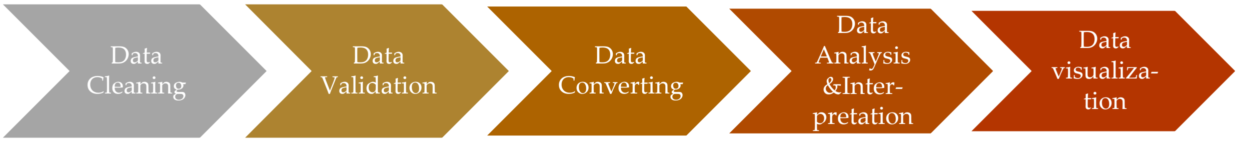

Figure 2. Steps of the data analysis process.

Table 3 comprises data that were classified into four distinct categories, which included fatalities (HW and NHW) and non-fatalities (HW and NHW), in order to identify any correlations between road types and accidents. However, this particular study focused solely on highway fatalities. The decision to prioritize highway accidents is underpinned by their distinctive characteristics compared to incidents occurring on urban streets. Highway accidents often involve higher speeds and longer road stretches, thus potentially amplifying collision severity and consequences and necessitating tailored safety interventions. Moreover, the factors contributing to highway accidents might diverge from those in urban contexts due to highway-specific features, such as the longer distances between exits and fewer intersections, which influence driver behavior and accident causation dynamics. This focus on highway accidents enables an exploration of these unique factors to inform targeted preventive strategies. Additionally, highway accidents involve a distinct demographic of drivers, including long-haul truckers and extended commuters, thus introducing specific risk factors that warrant thorough investigation. The relatively limited presence of pedestrians and cyclists on highways, as opposed to city streets, necessitates an individualized approach to safety analysis that acknowledges the distinct patterns that emerge. Lastly, the impacts of road design, signage, and maintenance on highway safety set highways apart from urban streets; by scrutinizing highway accidents, an evaluation of highway-specific safety measures and infrastructure effectiveness emerges, thus providing insights that are not necessarily applicable to city streets.

Table 3. Division of accidents caused by drivers based on highway vs. non-highway locations.

Road Type	Fatality		
	No	**Yes**	**Grand Total**
Non-Highway	47,136	9207	56,343
Highway	48,113	24,559 *	72,672
Grand Total	95,249	33,766	129,015

* The number of fatalities on highways.

To facilitate data analysis, the research team converted the overall dataset into a binary format, where each accident event was represented as either a "Yes" or "No" value and input into Python-based software. Table 3 presents the data segregated by road type and fatality. Notably, the large number of accidents (24,599, denoted by an asterisk in Table 3) captured our attention and motivated us to explore this further. Table 3 provides an overview of accidents caused by drivers that are categorized by whether they occurred on highways or non-highways.

Table 4 contains data from every event, encompassing 34 attributes related to accident data collection. These attributes include details such as the roadway type, vehicle type, environmental conditions, weather conditions, driver behavior, driver information, and driver status.

Table 4. Total of 34 attributes with a description of the setting.

Group	Attribute Name	Attribute Description
Roadway	Highway	1—Yes
	Dry surface road	1—Yes, 0—Otherwise
	Straightaway	1—Yes, 0—Otherwise
	Obstruction	1—Yes, 0—Otherwise
	Road condition	1—Yes, 0—Otherwise
	Vehicle condition	1—Yes, 0—Otherwise
Vehicle Type	Motorcycle	1—Yes, 0—Otherwise
	Mini-truck/pick-up (4 wheels)	1—Yes, 0—Otherwise
	Sedan	1—Yes, 0—Otherwise
	Light truck (6 wheels)	1—Yes, 0—Otherwise
	Heavy truck (10+ wheels)	1—Yes, 0—Otherwise
	Other type of car	1—Yes, 0—Otherwise
External Factors (Environment and Weather Condition)	Daytime (06.00–18.00)	1—Yes, 0—Otherwise
	Night with light	1—Yes, 0—Otherwise
	Night without light	1—Yes, 0—Otherwise
	Low visibility	1—Yes, 0—Otherwise
	Clear weather	1—Yes, 0—Otherwise
Internal Factors (Driver Behavior)	Drunk	1—Yes, 0—Otherwise
	Over speed limit	1—Yes, 0—Otherwise
	Breaking through traffic lights	1—Yes, 0—Otherwise
	Breaking through traffic signs	1—Yes, 0—Otherwise
	Overtaking	1—Yes, 0—Otherwise
	Using a mobile phone	1—Yes, 0—Otherwise
	Short cut-off	1—Yes, 0—Otherwise
	Drugs	1—Yes, 0—Otherwise
	Driving in the opposite direction	1—Yes, 0—Otherwise
	Dozing off	1—Yes, 0—Otherwise
	Overweight carry	1—Yes, 0—Otherwise
	Cannot conclude	1—Yes, 0—Otherwise
Driver Info	Gender	1— Male, 0—Otherwise
	Youth: 15–35	1—Yes, 0—Otherwise
	Adult: 36–60	1—Yes, 0—Otherwise
	Senior: 61–90+	1—Yes, 0—Otherwise
Driver Status	Fatality (death)	1—Yes

2.2. Methodology

The Apriori algorithm [55] was utilized to conduct frequent itemset mining on a large relational dataset to uncover the most common individual items and extend them to larger itemsets, as long as they appeared frequently enough in the database. The resulting frequent itemsets generated by Apriori could then be used to generate association rules that revealed overall trends. Association rule learning is a machine learning method that employs rules to identify strong relationships between variables within large databases by using various measures of attraction [56]. By detecting correlations and co-occurrences between datasets, association rules are useful for explaining data patterns that may appear to be unrelated, such as those found in relational and transactional databases. This process of applying association rules is known as association rule mining or mining associations. Please refer to Figure 3 for further details.

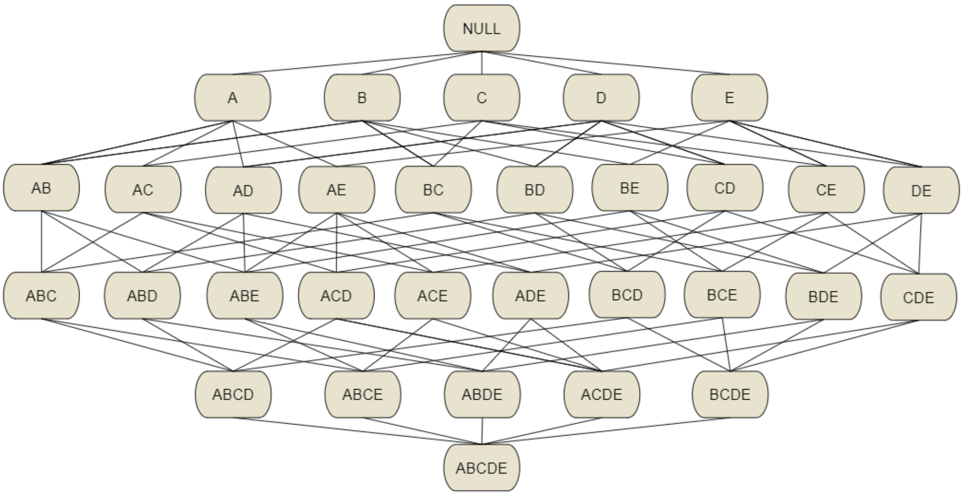

Figure 3. Diagram of associate rule mining.

Rule definition and measurement

An association rule is determined by two factors: support and confidence. The frequency with which a specific rule appears in the database being mined is referred to as the support. The number of times that a particular rule turns out to be true in practice is referred to as the confidence:

Let I = {...} represent a collection of "n" binary characteristics known as items;

Let J = {...} be a set of transactions referred to as a database.

Each transaction in J has a distinct transaction ID and includes a subset of the items in I. A rule is defined as an implication of the type xy, in which x, $y \subseteq$ I if only $x \neq \emptyset$, $y \neq \emptyset$, and $x \cap y = \emptyset$. The sets of objects x and y are referred to as the rule's antecedent and consequent, respectively.

Support (1) is an indicator of how frequently an itemset appears in the dataset:

$$\text{Support}(x) = \frac{Frequentitem(x)}{N(TotalNumberoftransaction)} \tag{1}$$

Confidence (2) is an indication of how often a rule has been found to be true:

$$Confidence[LHS(x) \Rightarrow RHS(y)] = \frac{Support(LHS, RHS)}{Support(LHS)} \tag{2}$$

Lift (3) is the ratio of the observed support to the support expected if x and y are independent:

$$Lift[LHS(x) \Rightarrow RHS(y)] = \frac{Support(LHS, RHS)}{Support(LHS) \times Support(RHS)} \tag{3}$$

A rule may have a significant association in a collection of data because it frequently appears; but, it may occur considerably less frequently when implemented. This would be an example of strong support but low confidence.

The following steps were used to perform associated rule mining:

- The accident transactions were sequenced by event (binary): If there was minimal support, the effectiveness of the accident was measured. If it was >50% (threshold), then others below 50% were removed;
- The frequency itemset from 1 was used to build a new itemset (length: 2). After using the join command, if all were set, the sequencing did not matter;
- The support score was recalculated by using the transaction in 1.1 until the intersection; for example: Transaction {Road wet} = {1,1,1,0,1, 0...}; Transaction {Darkness} = {1,1,1,1,0,0. . .};

Transaction {Road wet, Darkness} = {1,1,1,0,0,0. . .}; If the minimum support was < threshold, it was removed;

- The frequency itemset from 1.2 was used to create a new itemset (length: 3). However, it was noted that the initial item needed to be the same (by using the join command) and only one linkage could join: Transaction {Road wet, Darkness} = {1,1,1,0,0,0. . .}; Transaction {Road wet, Drunk} = {1,1,1,0,1,0. . .}; Transaction {Road wet, Darkness, Drunk} = {1,1,1,0,0,0. . .}; Frequency of all itemsets;
- The following two items or more were considered and then the confidence and lift were calculated.

It is important to note that a rule may have a significant association in a collection of data because it frequently appears; but, it may occur considerably less frequently when implemented. This would be an example of strong support but low confidence.

Considering its ability to identify frequent itemsets, generate association rules, provide insights into complex relationships between factors, scale well with large datasets, and have an easy implementation process, the Apriori algorithm is a valuable tool for accident data analysis. Because of these benefits, it is an effective approach for understanding the factors that contribute to accidents and informing evidence-based decision making in accident prevention and road safety improvement. However, it is computationally demanding, consumes significant memory, generates redundant rules, and has limitations with implicit itemsets.

3. Descriptive Statistics and Results

In order to gain insight into the patterns and distribution of the data, a distribution chart was generated for 72,672 highway accidents over a 24-hour period by using kernel density as a time series for descriptive statistics, as shown in Figure 4. To distinguish between day and night periods, the following values were assigned:

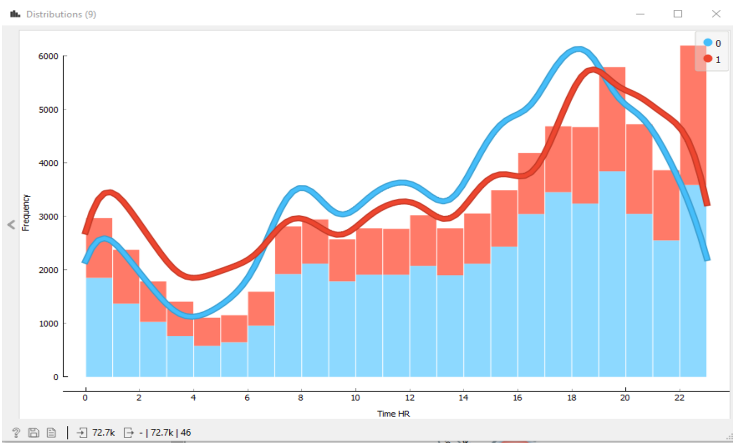

Figure 4. Highway accident distribution plot created by using a 24-hour time series w/kernel density as a line chart.

1—representing fatalities from highway accidents; $\mu = 13.19$, $\sigma = 7.03$;

0—representing non-fatalities from highway accidents; $\mu = 13.57$, $\sigma = 6.37$.

The majority of accidents occurred during the daytime (08.00–18.00) while peaks occurred at 19.00–20.00 and 22.00–23.00; a high fatality rate occurred at night (19.00–07.00).

Subsequently, the researchers focused on frequent itemsets related to fatality to extract rules that highlighted causal relationships, as illustrated in Figure 5. Identifying the co-occurrence of elements assisted in identifying linkages between them, with a minimum support of 50%. As shown in Figure 5, the most commonly occurring itemset in the 2018 dataset was associated with the following items: dry roads (95.98%), clear weather (87.33%), male drivers (86.42%), motorcycles (80.77%), straight roadways (71.99%), and exceeding the speed limit (69.03%).

Frequent Itemsets (2)

Itemsets	Support	%
Dry Surface Road=1	23572	95.98
Cannot Conclude =0	22709	92.47
Break Through Traffic lights=0	23364	95.13
Break Through Traffic Signs=0	23238	94.62
Drive in opposite direction=0	23188	94.42
Overtake=0	22917	93.31
Use Mobile Phone=0	23548	95.88
Drug=0	23569	95.97
Doze off=0	22724	92.53
Overweight Carry=0	23542	95.86
Obstruction=0	23190	94.43
Vehicle condition=0	23329	94.99
Road condition=0	23340	95.04
Light Truck =0	23335	95.02
Heavy Truck =0	23409	95.32
Other Type of car=0	23232	94.6
Sedan=0	22037	89.73
Low visibility=0	21652	88.16
Mini truck/ Pick up =0	21267	86.6
Drunk=0	20805	84.71
Clear Weather=1	21308	86.76
61-90=0	20480	83.39
Gender=1	20364	82.92
Short Cut off=0	19025	77.47
Motorcycle=1	18992	77.33
Night without Light=0	18299	74.51
Straight Way=1	17078	69.54
Night with Light=0	16920	68.9
Over Speed limit=1	16369	66.65
36-60=0	14475	58.94
Clear Weather=1	21448	87.33
Gender=1	21224	86.42
Drug=0	21220	86.4
Use Mobile Phone=0	21201	86.33
Overweight Carry=0	21195	86.3
Cannot Conclude =0	20458	83.3
Break Through Traffic lights=0	21046	85.7
Drive in opposite direction=0	20863	84.95
Break Through Traffic Signs=0	20949	85.3
Vehicle condition=0	21005	85.53
Obstruction=0	20875	85
Overtake=0	20639	84.04
Doze off=0	20454	83.29
Road condition=0	20955	85.33
Drunk=0	18569	75.61
Low visibility=0	19253	78.39
Short Cut off=0	17409	70.89
Motorcycle=1	19665	80.07
61-90=0	17005	69.24
Night without Light=0	16183	65.89
Straight Way=1	17679	71.99
Night with Light=0	14944	60.85
Over Speed limit=1	16952	69.03
36-60=0	12554	51.12

24.6k -

Figure 5. Itemset frequency extraction.

Following the analysis of the frequent itemsets in the highway fatality dataset, a total of 1558 association rules (1377 of which had a lift of ≥1) were generated, meeting the set threshold criteria (support 50%, confidence 95%). This was accomplished by using the Orange 3.30 software [57]; the results are presented in Figure 6. The support distribution had a mean of $\mu = 0.680263$ and a standard deviation of $\sigma = 0.0954974$ while the confidence distribution had a mean of $\mu = 0.972597$ and a standard deviation of $\sigma = 0.0126851$.

In total, 1558 rules were discovered through rule mining and were clustered based on their confidence levels, which are represented by different color shades. The *y*-axis displays the confidence while the *x*-axis represents the support. The results show that Group 1 (confidence: 0.95–0.965) had a majority of rules with the antecedent being male and dry surfaces being the consequence; this is represented by the blue shade. Group 2 (confidence: 0.965–0.98) had a majority of rules with the antecedent being motorcycles and driving over the speed limit and the consequence being dry surface roads; this is represented by the green shade. Group 3 (confidence: 0.98–0.995) had high confidence but low support, with clear weather as the antecedent and dry surface as the consequence; this is represented by the yellow shade. This suggested that these two elements played a significant role in road accident mortality (Figure 7) and it is essential to exercise extreme caution during clear weather on dry surfaces.

Association Rule for Fatality on HW

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent		Consequent
0.551	0.966	0.570	1.404	1.206	0.094	Over Speed limit=1, Mini truck/ Pick up =0, Sedan=0	→	Motorcycle=1
0.577	0.964	0.599	1.337	1.203	0.098	Straight Way=1, Mini truck/ Pick up =0, Sedan=0	→	Motorcycle=1
0.773	0.962	0.803	0.997	1.202	0.130	Dry Surface Road=1, Mini truck/ Pick up =0, Sedan=0	→	Motorcycle=1
0.704	0.962	0.732	1.093	1.201	0.118	Clear Weather=1, Mini truck/ Pick up =0, Sedan=0	→	Motorcycle=1
0.689	0.958	0.719	1.113	1.196	0.113	Gender=1, Mini truck/ Pick up =0, Sedan=0	→	Motorcycle=1
0.548	0.995	0.551	1.742	1.037	0.019	Clear Weather=1, Over Speed limit=1, Mini truck/ Pick up =0	→	Dry Surface Road=1
0.527	0.995	0.529	1.813	1.037	0.019	Clear Weather=1, Road condition=0, 36-60=0	→	Dry Surface Road=1
0.514	0.995	0.517	1.856	1.037	0.018	Clear Weather=1, Doze off=0, 36-60=0	→	Dry Surface Road=1
0.566	0.995	0.569	1.686	1.037	0.020	Clear Weather=1, Over Speed limit=1, Sedan=0	→	Dry Surface Road=1
0.516	0.995	0.519	1.849	1.036	0.018	Clear Weather=1, Overtake=0, 36-60=0	→	Dry Surface Road=1
0.523	0.995	0.526	1.824	1.036	0.018	Clear Weather=1, Obstruction=0, 36-60=0	→	Dry Surface Road=1
0.529	0.995	0.532	1.805	1.036	0.019	Clear Weather=1, 36-60=0, Heavy Truck =0	→	Dry Surface Road=1
0.527	0.995	0.530	1.812	1.036	0.019	Clear Weather=1, Break Through Traffic lights=0, 36-60=0	→	Dry Surface Road=1
0.526	0.995	0.529	1.816	1.036	0.018	Clear Weather=1, Vehicle condition=0, 36-60=0	→	Dry Surface Road=1
0.531	0.995	0.534	1.797	1.036	0.019	Clear Weather=1, 36-60=0	→	Dry Surface Road=1
0.531	0.995	0.534	1.797	1.036	0.019	Clear Weather=1, Drug=0, 36-60=0	→	Dry Surface Road=1
0.531	0.995	0.533	1.799	1.036	0.019	Clear Weather=1, Use Mobile Phone=0, 36-60=0	→	Dry Surface Road=1
0.531	0.995	0.533	1.799	1.036	0.019	Clear Weather=1, Overweight Carry=0, 36-60=0	→	Dry Surface Road=1
0.513	0.995	0.516	1.859	1.036	0.018	Clear Weather=1, Cannot Conclude =0, 36-60=0	→	Dry Surface Road=1
0.527	0.995	0.530	1.810	1.036	0.019	Clear Weather=1, 36-60=0, Light Truck =0	→	Dry Surface Road=1
0.525	0.995	0.528	1.818	1.036	0.018	Clear Weather=1, 36-60=0, Other Type of car=0	→	Dry Surface Road=1
0.524	0.995	0.527	1.821	1.036	0.018	Clear Weather=1, Break Through Traffic Signs=0, 36-60=0	→	Dry Surface Road=1
0.523	0.995	0.525	1.827	1.036	0.018	Clear Weather=1, Drive in opposite direction=0, 36-60=0	→	Dry Surface Road=1
0.604	0.995	0.608	1.580	1.036	0.021	Clear Weather=1, Over Speed limit=1, Heavy Truck =0	→	Dry Surface Road=1
0.694	0.995	0.698	1.376	1.036	0.024	Clear Weather=1, Road condition=0, Motorcycle=1	→	Dry Surface Road=1
0.602	0.995	0.605	1.586	1.036	0.021	Clear Weather=1, Over Speed limit=1, Obstruction=0	→	Dry Surface Road=1
0.511	0.995	0.514	1.868	1.036	0.018	Straight Way=1, Clear Weather=1, Motorcycle=1	→	Dry Surface Road=1
0.593	0.995	0.597	1.609	1.036	0.021	Clear Weather=1, Over Speed limit=1, Overtake=0	→	Dry Surface Road=1
0.604	0.995	0.608	1.580	1.036	0.021	Clear Weather=1, Over Speed limit=1, Road condition=0	→	Dry Surface Road=1
0.601	0.995	0.604	1.589	1.036	0.021	Clear Weather=1, Over Speed limit=1, Break Through Traffic Signs=0	→	Dry Surface Road=1
0.608	0.995	0.611	1.570	1.036	0.021	Clear Weather=1, Over Speed limit=1	→	Dry Surface Road=1
0.608	0.995	0.611	1.570	1.036	0.021	Clear Weather=1, Cannot Conclude =0, Over Speed limit=1	→	Dry Surface Road=1
0.608	0.995	0.611	1.570	1.036	0.021	Clear Weather=1, Over Speed limit=1, Drug=0	→	Dry Surface Road=1
0.608	0.995	0.611	1.571	1.036	0.021	Clear Weather=1, Over Speed limit=1, Overweight Carry=0	→	Dry Surface Road=1
0.608	0.995	0.611	1.571	1.036	0.021	Clear Weather=1, Over Speed limit=1, Use Mobile Phone=0	→	Dry Surface Road=1

24.6k 12.9k | 1558

Figure 6. Exploration of association rule mining: A total of 1558 rules were unearthed.

A hierarchical cluster analysis (HCA) was performed by applying an agglomerative analysis to 1558 rules to group related antecedents into clusters with distances. The Euclidean distance was used as a complete linkage criterion to calculate the distance between the clusters. The resulting dendrogram (Figure 8) shows three clusters for the antecedent. Cluster 1 (C1) contains straightaways, driving over the speed limit, dry surface roads, clear weather, and male gender. Cluster 2 (C2) contains straightaways, driving over the speed limit, clear weather, and male gender. Cluster 3 (C3) contains motorcycles, driving over the speed limit, clear weather, and male gender.

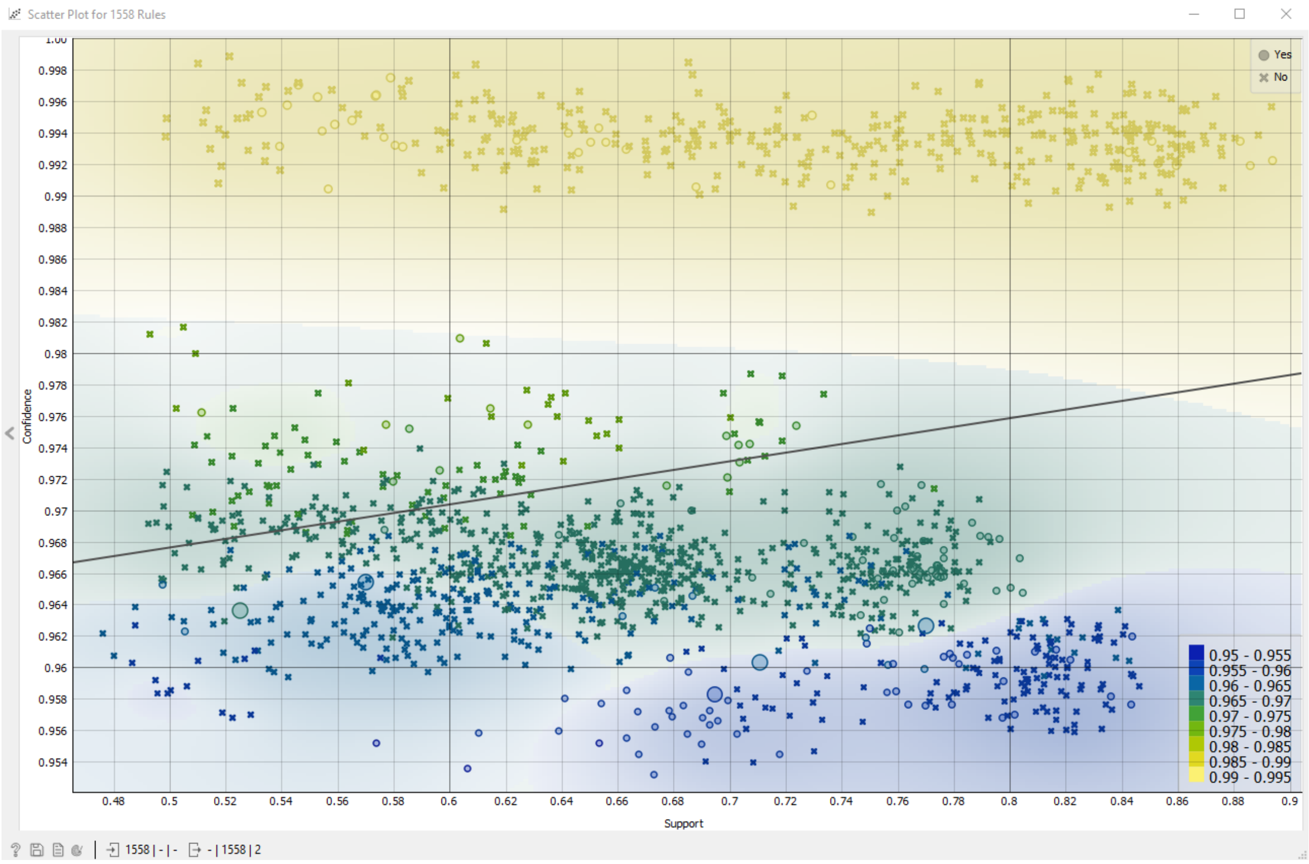

Figure 7. Visualization of 1558 discovered rules through a scatter plot of support vs. confidence.

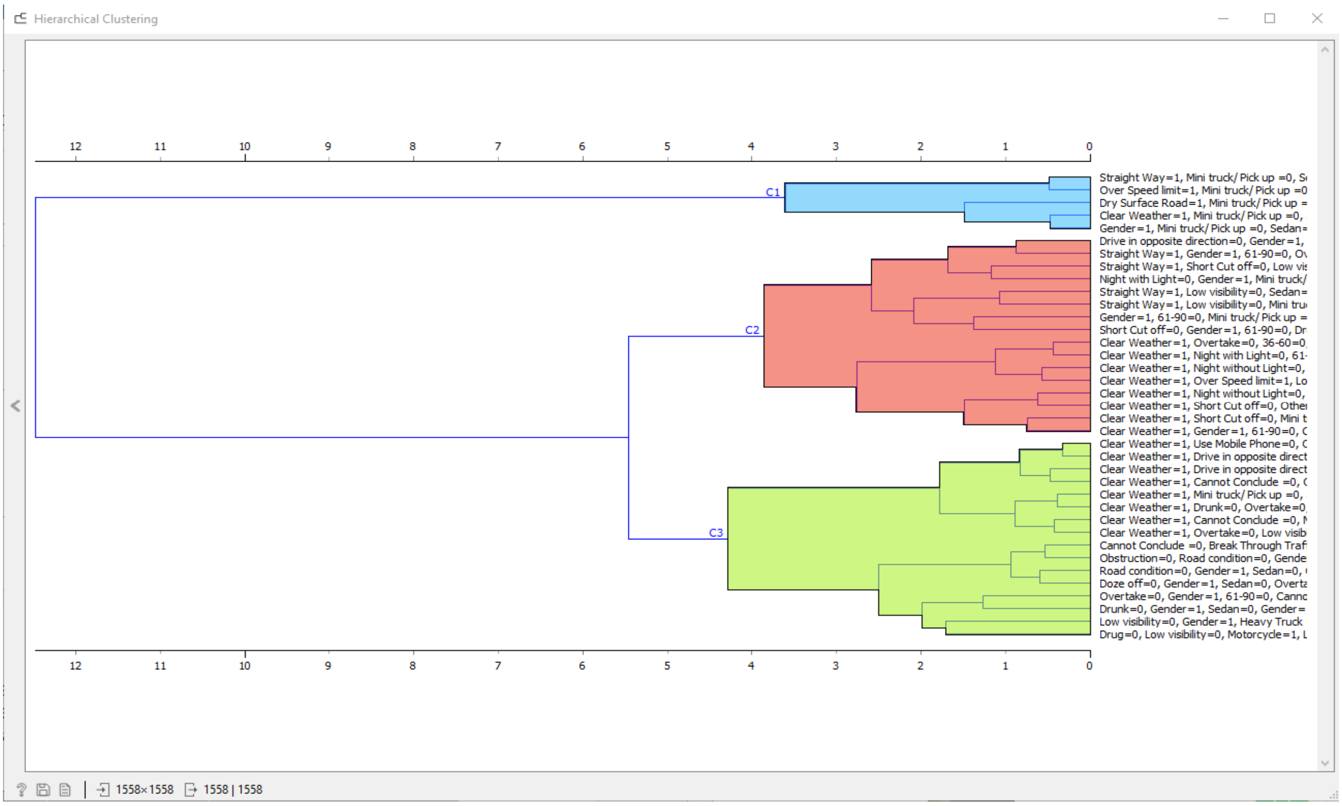

Figure 8. Dendrogram for 1558 rules discovered with their antecedents.

In order to determine the factors leading to motorcycle fatalities, this study utilized association rule mining and hierarchical clustering analysis. The results showed that the majority of motorcycle fatalities occurred on straightaways while driving over the speed limit and clear weather also played a significant role. On the other hand, dry surface roads were found to be a common consequence of the antecedents in clusters C2 and C3. To further explore the relationships between the antecedents and consequences, a set of association rules with a high lift and a wide gap between the support and confidence were identified. The rules are plotted in Figure 9 and presented in Table 5, with a minimum support score of 50%, a confidence threshold of 95%, and a lift threshold of 1. For instance, the rule with the widest gap between the support and confidence involved straightaways, clear weather, and motorcycles as antecedents and dry surface roads as a consequence, with an increase from a support of 0.511 to a confidence of 0.995 by 0.484. Additionally, the rule with the highest lift was associated with different antecedents; but, all of them contained motorcycles.

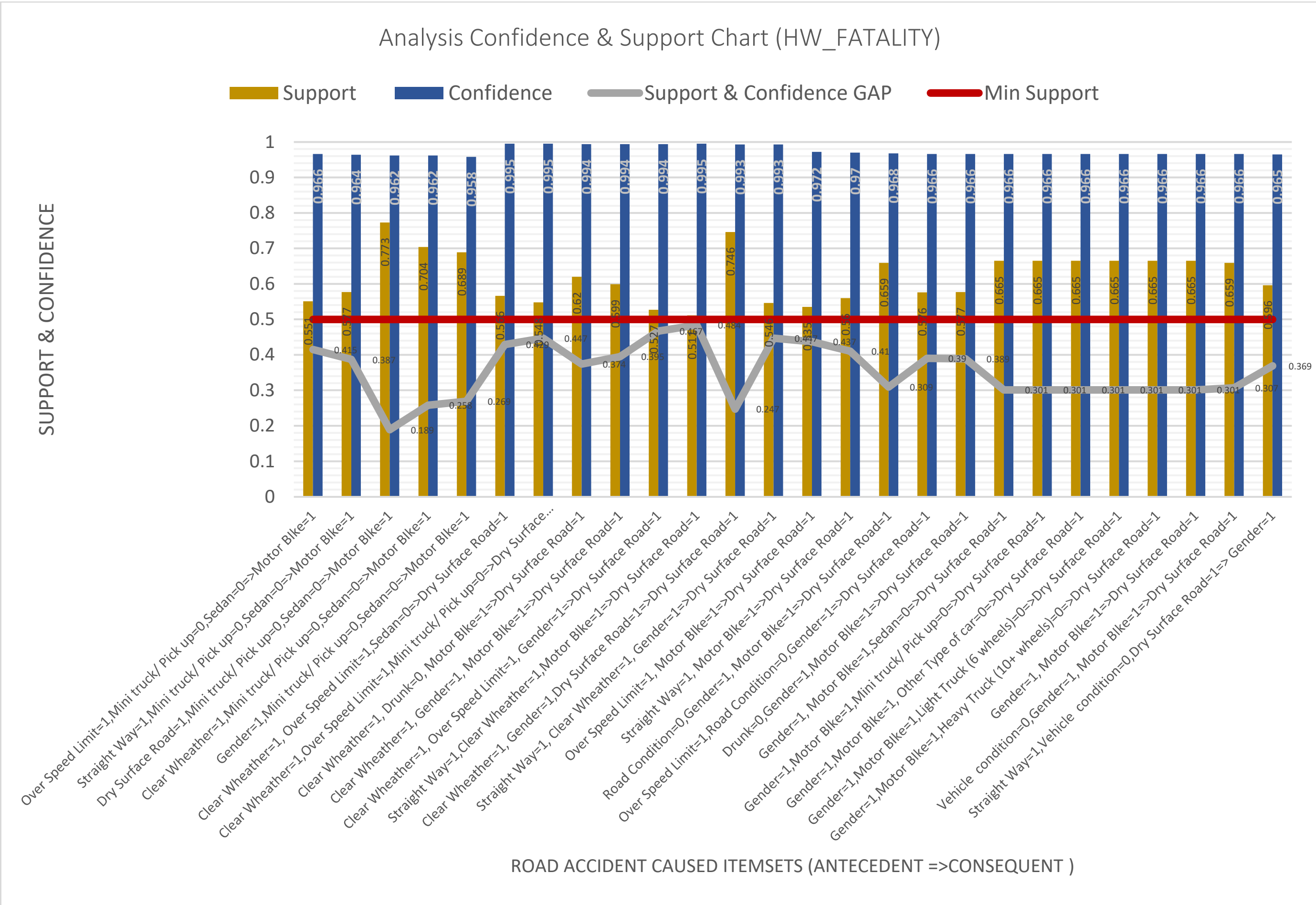

Figure 9. Chart depicting trends in the gap between confidence and support among noteworthy rules.

Table 5. Emphasis on rules with a high lift and substantial gap between support and confidence.

Antecedent_1	Antecedent_2	Antecedent_3	Consequence	Support	Confidence	Lift
Over Speed Limit = 1	Mini-truck/Pick-up = 0	Sedan = 0	Motorcycle = 1	0.551	0.966	1.206
Straightaway = 1	Mini-truck/Pick-up = 0	Sedan = 0	Motorcycle = 1	0.577	0.964	1.203
Dry Surface Road = 1	Mini-truck/Pick-up = 0	Sedan = 0	Motorcycle = 1	0.773	0.962	1.202
Clear Whether = 1	Mini-truck/Pick-up = 0	Sedan = 0	Motorcycle = 1	0.704	0.962	1.201
Gender = 1	Mini-truck/Pick-up = 0	Sedan = 0	Motorcycle = 1	0.689	0.958	1.196

Table 5. *Cont.*

Antecedent_1	Antecedent_2	Antecedent_3	Consequence	Support	Confidence	Lift
Clear Weather = 1	Over Speed Limit = 1	Sedan = 0	Dry Surface Road = 1	0.566	0.995	1.037
Clear Weather = 1	Over Speed Limit = 1	Mini-truck/Pick-up = 0	Dry Surface Road = 1	0.548	0.995	1.037
Clear Whether = 1	Drunk = 0	Motorcycle = 1	Dry Surface Road = 1	0.62	0.994	1.036
Clear Whether = 1	Gender = 1	Motorcycle = 1	Dry Surface Road = 1	0.599	0.994	1.036
Clear Whether = 1	Over Speed Limit = 1	Gender = 1	Dry Surface Road = 1	0.527	0.994	1.036
Straightaway = 1	Clear Weather = 1	Motorcycle = 1	Dry Surface Road = 1	0.511	0.995	1.036
Clear Weather = 1	Gender = 1	Dry Surface Road = 1	Dry Surface Road = 1	0.746	0.993	1.035
Straightaway = 1	Clear Weather = 1	Gender = 1	Dry Surface Road = 1	0.546	0.993	1.035
Over Speed Limit = 1	Motor Bike = 1		Dry Surface Road = 1	0.535	0.972	1.013
Straightaway = 1	Motor Bike = 1		Dry Surface Road = 1	0.56	0.97	1.011
Road Condition = 0	Gender = 1	Motorcycle = 1	Dry Surface Road = 1	0.659	0.968	1.008
Over Speed Limit = 1	Road Condition = 0	Gender = 1	Dry Surface Road = 1	0.576	0.966	1.007
Drunk = 0	Gender = 1	Motorcycle = 1	Dry Surface Road = 1	0.577	0.966	1.006
Gender = 1	Motorcycle = 1	Sedan = 0	Dry Surface Road = 1	0.665	0.966	1.006
Gender = 1	Motorcycle = 1	Mini-truck/Pick-up = 0	Dry Surface Road = 1	0.665	0.966	1.006
Gender = 1	Motorcycle = 1	Other Type of car = 0	Dry Surface Road = 1	0.665	0.966	1.006
Gender = 1	Motorcycle = 1	Light Truck (6 wheels) = 0	Dry Surface Road = 1	0.665	0.966	1.006
Gender = 1	Motorcycle = 1	Heavy Truck (10+ wheels) = 0	Dry Surface Road = 1	0.665	0.966	1.006
Gender = 1	Motorcycle = 1		Dry Surface Road = 1	0.665	0.966	1.006
Vehicle condition = 0	Gender = 1	Motorcycle = 1	Dry Surface Road = 1	0.659	0.966	1.006
Straightaway = 1	Vehicle condition = 0	Gender = 1	Dry Surface Road = 1	0.596	0.965	1.006

4. Discussion

According to this study, the higher risk of accidents on clear straight roads may be due to drivers—particularly male drivers—violating speed limits in good weather conditions. Thailand has over 42 million registered vehicles, with motorcycles accounting for 50% of the total and being responsible for the majority of road fatalities [58], thus potentially contributing to the largest number of fatalities from significant accidents. As Jomnonkwao, Uttra and Ratanavaraha [2] observed, motorcyclists are responsible for the vast majority of road fatalities; meanwhile, prior studies showed different types of cars and motorcycles, such as SUV/vans [47], large trucks [40,47,49], and pick-ups [40], that have rolled over. Additional research on motorcycle riders specifically, as well as other types of road users, may be conducted in the future. Aside from motorcycles, Sonal and Suman [4] observed that external factors, such as weather conditions, including fog, rain, and snow, showed greater impacts on road accidents than internal factors, such as the drivers themselves. Meanwhile, Thailand's climate has no snow or ice and rain is only present for roughly five months of the year. The remainder of the year has clear weather conditions and dry road surfaces in the chilly and summer seasons. Further research on motorcycle riders and other road users is necessary.

In a previous study, highway junctions were identified as the most hazardous locations for accidents [14]. However, this current research highlights that straight roads without junctions pose a significant risk due to drivers often exceeding speed limits. Additionally, previous studies by Bahiru, Kumar Singh and Tessfaw [10] showed that internal factors, such as gender, age, accident location, and vehicle type, have a relatively minor impact on the severity of accidents; although, being male remains a significant contributing factor to fatalities on highways.

This study's findings have important implications for policymakers working to reduce the factors that contribute to highway accidents and raise awareness of risky driving behaviors. Policymakers can develop targeted measures to address these factors and improve road safety by considering the discovered association rules. Implementing laws that control speed limits, specifically on straightaways, is one potential measure. To remind drivers of the safe speed range, light signs indicating the appropriate speed limit can be strategically placed along these road segments. Additionally, warning signs can be used to alert drivers to potential hazards and encourage them to drive cautiously. Installing cameras to monitor driving speeds, especially on motorcycles, can be an effective deterrent for speeding. The Thai government has already taken steps to reduce road accidents by improving infrastructure, enforcing traffic laws, and implementing safety education and awareness programs in schools and workplaces. These initiatives seek to instill a sense of responsibility in individuals and educate them on the importance of following traffic rules and practicing safe driving habits. However, while these efforts are commendable, more comprehensive measures are required to address the underlying causes of accidents. For instance, there is a need to focus on enhancing safety awareness among drivers and promoting responsible behavior on the roads. This could involve targeted awareness campaigns that highlight the potential consequences of reckless driving, as well as educational programs that provide drivers with the necessary skills and knowledge to navigate challenging road conditions.

In the context of this discussion, it is critical to identify the distinguishing features that set Thailand apart from other countries in terms of its high rate of road accidents. This situation is the result of several distinct factors, emphasizing the importance of a tailored and nuanced approach to road safety initiatives.

To begin, in Thailand's road safety landscape, the cultural dimension of alcohol consumption is critical. Thailand's sociocultural norms include widespread acceptance of alcohol consumption, which can inadvertently increase the prevalence of drunk driving incidents. This cultural attitude toward alcohol, combined with limited enforcement and public awareness campaigns, may contribute to an increase in the number of road accidents involving intoxicated drivers, thus distinguishing Thailand from countries with different attitudes toward alcohol consumption and stringent anti-drunk-driving laws. Another unique aspect contributing to Thailand's road safety challenges is the issue of drivers operating vehicles without proper licenses. This phenomenon could be attributed to a number of factors, such as lax enforcement of licensing regulations and difficulties in ensuring compliance due to the country's geographical and administrative complexities. This trend differs from that in regions where stricter licensing regulations are more effectively enforced, resulting in a significant difference in the prevalence of unlicensed drivers contributing to road accidents. The state of the road infrastructure is also an important factor to consider. While Thailand has experienced rapid economic development, modernization of the road infrastructure has not kept pace. Some regions may have cutting-edge road networks while others may be in desperate need of maintenance and improvement. This disparity in road infrastructure quality combined with the varying levels of urbanization and development may contribute to disparities in accident rates across the country, thus distinguishing Thailand from countries with more uniformly developed road networks.

Furthermore, Thailand's urban landscape is unique, with a mix of urban centers, rural areas, and tourist destinations, which contributes to a diverse range of road users and driving conditions. This complex mix of environments can make it difficult to develop comprehensive road safety strategies that effectively address the needs of these various contexts. The coexistence of bustling cities, remote rural settings, and tourist-heavy regions complicates efforts to harmonize road safety measures. To summarize, Thailand's high rate of road accidents can be attributed to a combination of cultural factors, licensing issues, varying road infrastructure quality, and the diversity of its urban landscape. These features distinguish Thailand from other countries and highlight the need for tailored road safety strategies that address these distinguishing features effectively.

In essence, this study advances beyond established knowledge by uncovering hidden nuances in the antecedents of road accidents and their implications. The incorporation of insights from a government official underscores the practical value of our findings and their potential impacts on road safety policies and interventions in Thailand.

5. Conclusions

The association rule analysis performed in this study identified several key factors that significantly increase the likelihood of fatalities in highway accidents. Male drivers, speeding, motorcycles, straight and dry roads, and clear weather were identified as interconnected variables associated with an increased risk of injury or death in traffic accidents. The highest risk was observed for males riding motorcycles at speeds exceeding the speed limit on straight roads in clear weather. The confidence levels in the association rules gradually increased, indicating a stronger relationship between these factors. Notably, the presence of multiple factors increases the likelihood of an accident. Straightaways being identified as a significant contributor to accidents emphasizes the importance of exercising caution at intersections and on curved roads.

The rule pertains to males who ride motorcycles at speeds that exceed the posted speed limit while traveling on straight roads under clear weather conditions. While these findings may appear obvious, this algorithm assists in uncovering and quantifying the relationships between these factors in a systematic manner that goes beyond common sense. In essence, this study provides a structured method for confirming what may be intuitively understood. This method improves our understanding of how these factors interact to increase the likelihood of an accident. As a result, the importance of this study stems from its rigorous methodology, which uncovers and quantifies the complex relationships among the identified variables. Despite the factors' apparent simplicity, this study's strength lies in its ability to reveal the intricate connections between them. This greater understanding can help guide more effective strategies and policies for reducing the risks of traffic accidents.

In conclusion, policymakers should consider the association rules uncovered in this study as a basis for designing and implementing effective strategies for reducing highway accidents. By combining legislative measures, infrastructure improvements, enforcement efforts, and educational initiatives, it is possible to create a safer road environment and foster a culture of responsible driving. Continued efforts and collaboration among relevant stakeholders are vital for addressing the complex factors contributing to accidents and ensuring the wellbeing of road users in Thailand.

6. Limitations and Further Research

This study utilized accident data from the COVID-19 pandemic, during which the government imposed lockdowns and restrictions on travel between provinces. People were also cautious about traveling to isolated areas on their own, which suggests that they did not travel extensively. Therefore, the accident and fatality numbers for 2019–2020 may not accurately reflect the actual figures for the country.

As a related aspect, future research may expand the analysis to include all road types, specific types of vehicles, criminal and medical data, and non-highway data to assist policymakers in selecting the most practical options with solid data support.

Furthermore, the incorporation of autonomous driving technology has emerged as a compelling approach to reducing the occurrence of road accidents as a promising avenue for future research. With advancements in artificial intelligence and automation, self-driving vehicles have the potential to revolutionize road safety by reducing the role of human error, which is a major cause of accidents [59]. Exploring the feasibility, efficacy, and implications of introducing autonomous driving technology within the context of Thailand's unique road safety landscape could provide valuable insights. This line of inquiry could look into infrastructure readiness, regulatory changes, public acceptance, and potential benefits, thus contributing to the region's ongoing discussion about improving road safety [60].

Author Contributions: I.M.: conceptualization, methodology, software, formal analysis, data curation, visualization, and writing—original draft. S.J.: conceptualization, validation, and writing—review and editing. R.K. and V.R.: supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This work and the APC were supported by (i) the Suranaree University of Technology (SUT), (ii) Thailand Science Research and Innovation (TSRI), and (iii) the National Science Research and Innovation Fund (NSRF) (Project: 179277).

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Human Research Ethics Committee of Suranaree University of Technology (EC-65-13).

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request due to restrictions.

Acknowledgments: The authors would like to thank the Department of Disaster Prevention and Mitigation, Ministry of the Interior, for supporting the road traffic crash data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. WHO. World Health Organization: Global Status Report on Road Safety 2018. Available online: https://extranet.who.int/roadsafety/death-on-the-roads/ (accessed on 1 April 2018).
2. Jomnonkwao, S.; Uttra, S.; Ratanavaraha, V. Forecasting Road Traffic Deaths in Thailand: Applications of Time-Series, Curve Estimation, Multiple Linear Regression, and Path Analysis Models. *Sustainability* **2020**, *12*, 395. [CrossRef]
3. Perneger, T.V. The Swiss cheese model of safety incidents: Are there holes in the metaphor? *BMC Health Serv. Res.* **2005**, *5*, 71. [CrossRef] [PubMed]
4. Sonal, S.; Suman, S. A framework for analysis of road accidents. In Proceedings of the 2018 International Conference on Emerging Trends and Innovations in Engineering and Technological Research, ICETIETR 2018, Ernakulam, India, 11–13 July 2018.
5. Gutierrez-Osorio, C.; Pedraza, C. Modern data sources and techniques for analysis and forecast of road accidents: A review. *J. Traffic Transp. Eng. Engl. Ed.* **2020**, *7*, 432–446. [CrossRef]
6. Abellán, J.; López, G.; de Oña, J. Analysis of traffic accident severity using Decision Rules via Decision Trees. *Expert Syst. Appl.* **2013**, *40*, 6047–6054. [CrossRef]
7. Al Mamlook, R.E.; Ali, A.; Hasan, R.A.; Mohamed Kazim, H.A. Machine Learning to Predict the Freeway Traffic Accidents-Based Driving Simulation. In Proceedings of the IEEE National Aerospace Electronics Conference, NAECON, Dayton, OH, USA, 15–19 July 2019; pp. 630–634.
8. Mafi, S.; AbdelRazig, Y.; Doczy, R. Machine Learning Methods to Analyze Injury Severity of Drivers from Different Age and Gender Groups. *Transp. Res. Rec.* **2018**, *2672*, 171–183. [CrossRef]
9. Recal, F.; Demirel, T. Comparison of machine learning methods in predicting binary and multi-class occupational accident severity. *J. Intell. Fuzzy Syst.* **2021**, *40*, 10981–10998. [CrossRef]
10. Bahiru, T.K.; Kumar Singh, D.; Tessfaw, E.A. Comparative Study on Data Mining Classification Algorithms for Predicting Road Traffic Accident Severity. In Proceedings of the International Conference on Inventive Communication and Computational Technologies, ICICCT 2018, Coimbatore, India, 20–21 April 2018; pp. 1655–1660.
11. Cuenca, L.G.; Puertas, E.; Aliane, N.; Andres, J.F. Traffic Accidents Classification and Injury Severity Prediction. In Proceedings of the 2018 3rd IEEE International Conference on Intelligent Transportation Engineering, ICITE 2018, Singapore, 3–5 September 2018; pp. 52–57.
12. Kuşkapan, E.; Çodur, M.Y.; Atalay, A. Speed violation analysis of heavy vehicles on highways using spatial analysis and machine learning algorithms. *Accid. Anal. Prev.* **2021**, *155*, 106098. [CrossRef]
13. Ospina-Mateus, H.; Quintana Jiménez, L.A.; Lopez-Valdes, F.J.; Berrio Garcia, S.; Barrero, L.H.; Sana, S.S. Extraction of decision rules using genetic algorithms and simulated annealing for prediction of severity of traffic accidents by motorcyclists. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 10051–10072. [CrossRef]
14. Kumar, S.; Toshniwal, D. A data mining approach to characterize road accident locations. *J. Mod. Transp.* **2016**, *24*, 62–72. [CrossRef]
15. Helen, W.R.; Almelu, N.; Nivethitha, S. Mining Road Accident Data Based on Diverted Attention of Drivers. In Proceedings of the 2nd International Conference on Intelligent Computing and Control Systems, ICICCS 2018, Madurai, India, 14–15 June 2018; pp. 245–249.
16. El Abdallaoui, H.E.A.; El Fazziki, A.; Ennaji, F.Z.; Sadgal, M. Decision Support System for the Analysis of Traffic Accident Big Data. In Proceedings of the 14th International Conference on Signal Image Technology and Internet Based Systems, SITIS 2018, Las Palmas de Gran Canaria, Spain, 26–29 November 2018; pp. 514–521.
17. Guido, G.; Shaffiee Haghshenas, S.; Shaffiee Haghshenas, S.; Vitale, A.; Astarita, V.; Park, Y.; Geem, Z.W. Evaluation of Contributing Factors Affecting Number of Vehicles Involved in Crashes Using Machine Learning Techniques in Rural Roads of Cosenza, Italy. *Safety* **2022**, *8*, 28. [CrossRef]
18. John, M.; Shaiba, H. Apriori-Based Algorithm for Dubai Road Accident Analysis. *Procedia Comput. Sci.* **2019**, *163*, 218–227. [CrossRef]
19. Feng, M.; Zheng, J.; Ren, J.; Xi, Y. Association Rule Mining for Road Traffic Accident Analysis: A Case Study from UK. In *Advances in Brain Inspired Cognitive Systems*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2020; pp. 520–529.
20. Bhavsar, R.; Amin, A.; Zala, L. Development of Model for Road Crashes and Identification of Accident Spots. *Int. J. Intell. Transp. Syst. Res.* **2021**, *19*, 99–111. [CrossRef]

21. Samerei, S.A.; Aghabayk, K.; Mohammadi, A.; Shiwakoti, N. Data mining approach to model bus crash severity in Australia. *J. Saf. Res.* **2021**, *76*, 73–82. [CrossRef]
22. John, M.; Shaiba, H. Analysis of Road Accidents Using Data Mining Paradigm. *Lect. Notes Data Eng. Commun. Technol.* **2022**, *68*, 215–223. [CrossRef]
23. Guido, G.; Haghshenas, S.S.; Haghshenas, S.S.; Vitale, A.; Astarita, V.; Haghshenas, A.S. Feasibility of Stochastic Models for Evaluation of Potential Factors for Safety: A Case Study in Southern Italy. *Sustainability* **2020**, *12*, 7541. [CrossRef]
24. Zhang, X.F.; Fan, L. A decision tree approach for traffic accident analysis of Saskatchewan highways. In Proceedings of the Canadian Conference on Electrical and Computer Engineering, Regina, SK, Canada, 5–8 May 2013.
25. Bucsuházy, K.; Matuchová, E.; Zůvala, R.; Moravcová, P.; Kostíková, M.; Mikulec, R. Human factors contributing to the road traffic accident occurrence. *Transp. Res. Procedia* **2020**, *45*, 555–561. [CrossRef]
26. Ospina-Mateus, H.; Quintana Jiménez, L.A.; López-Valdés, F.J.; Morales-Londoño, N.; Salas-Navarro, K. Using Data-Mining Techniques for the Prediction of the Severity of Road Crashes in Cartagena, Colombia. *Commun. Comput. Inf. Sci.* **2019**, *1052*, 309–320. [CrossRef]
27. Mohamad, I.; Jomnonkwao, S.; Ratanavaraha, V. Using a decision tree to compare rural versus highway motorcycle fatalities in Thailand. *Case Stud. Transp. Policy* **2022**, *10*, 2165–2174. [CrossRef]
28. Mphela, T. Causes of road accidents in botswana: An econometric model. *J. Transp. Supply Chain Manag.* **2020**, *14*, 1–8. [CrossRef]
29. Chen, C.; Zhang, G.; Yang, J.; Milton, J.C.; Alcántara, A.D. An explanatory analysis of driver injury severity in rear-end crashes using a decision table/Naïve Bayes (DTNB) hybrid classifier. *Accid. Anal. Prev.* **2016**, *90*, 95–107. [CrossRef] [PubMed]
30. Malin, F.; Norros, I.; Innamaa, S. Accident risk of road and weather conditions on different road types. *Accid. Anal. Prev.* **2019**, *122*, 181–188. [CrossRef] [PubMed]
31. Yadav, J.; Batra, K.; Goel, A.K. A Framework for Analyzing Road Accidents Using Machine Learning Paradigms. *J. Phys. Conf. Ser.* **2021**, *1950*, 012072.
32. Chen, C.; Zhang, G.; Tarefder, R.; Ma, J.; Wei, H.; Guan, H. A multinomial logit model-Bayesian network hybrid approach for driver injury severity analyses in rear-end crashes. *Accid. Anal. Prev.* **2015**, *80*, 76–88. [CrossRef] [PubMed]
33. Xie, Y.; Huynh, N. Analysis of driver injury severity in rural single-vehicle crashes. *Accid. Anal. Prev.* **2012**, *47*, 36–44. [CrossRef]
34. Behnood, A.; Mannering, F. The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach. *Anal. Methods Accid. Res.* **2017**, *14*, 41–53. [CrossRef]
35. Li, Z.; Wu, Q.; Ci, Y.; Chen, C.; Chen, X.; Zhang, G. Using latent class analysis and mixed logit model to explore risk factors on driver injury severity in single-vehicle crashes. *Accid. Anal. Prev.* **2019**, *129*, 230–240. [CrossRef]
36. Wu, Q.; Zhang, G.; Zhu, X.; Liu, X.C.; Tarefder, R. Analysis of driver injury severity in single-vehicle crashes on rural and urban roadways. *Accid. Anal. Prev.* **2016**, *94*, 35–45. [CrossRef]
37. Osman, M.; Mishra, S.; Paleti, R. Injury severity analysis of commercially-licensed drivers in single-vehicle crashes: Accounting for unobserved heterogeneity and age group differences. *Accid. Anal. Prev.* **2018**, *118*, 289–300. [CrossRef]
38. Hou, Q.; Huo, X.; Leng, J.; Cheng, Y. Examination of driver injury severity in freeway single-vehicle crashes using a mixed logit model with heterogeneity-in-means. *Phys. A Stat. Mech. Appl.* **2019**, *531*, 121760. [CrossRef]
39. Kim, J.-K.; Ulfarsson, G.F.; Kim, S.; Shankar, V.N. Driver-injury severity in single-vehicle crashes in California: A mixed logit analysis of heterogeneity due to age and gender. *Accid. Anal. Prev.* **2013**, *50*, 1073–1081. [CrossRef] [PubMed]
40. Li, Z.; Chen, C.; Wu, Q.; Zhang, G.; Liu, C.; Prevedouros, P.D.; Ma, D.T. Exploring driver injury severity patterns and causes in low visibility related single-vehicle crashes using a finite mixture random parameters model. *Anal. Methods Accid. Res.* **2018**, *20*, 1–14. [CrossRef]
41. Champahom, T.; Jomnonkwao, S.; Watthanaklang, D.; Karoonsoontawong, A.; Chatpattananan, V.; Ratanavaraha, V. Applying hierarchical logistic models to compare urban and rural roadway modeling of severity of rear-end vehicular crashes. *Accid. Anal. Prev.* **2020**, *141*, 105537. [CrossRef]
42. Li, Z.; Ci, Y.; Chen, C.; Zhang, G.; Wu, Q.; Qian, Z.; Prevedouros, P.D.; Ma, D.T. Investigation of driver injury severities in rural single-vehicle crashes under rain conditions using mixed logit and latent class models. *Accid. Anal. Prev.* **2019**, *124*, 219–229. [CrossRef]
43. Zhou, M.; Chin, H.C. Factors affecting the injury severity of out-of-control single-vehicle crashes in Singapore. *Accid. Anal. Prev.* **2019**, *124*, 104–112. [CrossRef] [PubMed]
44. Wei, F.; Cai, Z.; Liu, P.; Guo, Y.; Li, X.; Li, Q. Exploring Driver Injury Severity in Single-Vehicle Crashes under Foggy Weather and Clear Weather. *J. Adv. Transp.* **2021**, *2021*, 9939800. [CrossRef]
45. Krull, K.A.; Khattak, A.J.; Council, F.M. Injury Effects of Rollovers and Events Sequence in Single-Vehicle Crashes. *Transp. Res. Rec.* **2000**, *1717*, 46–54. [CrossRef]
46. Yu, M.; Zheng, C.; Ma, C. Analysis of injury severity of rear-end crashes in work zones: A random parameters approach with heterogeneity in means and variances. *Anal. Methods Accid. Res.* **2020**, *27*, 100126. [CrossRef]
47. Jafari Anarkooli, A.; Hosseinpour, M.; Kardar, A. Investigation of factors affecting the injury severity of single-vehicle rollover crashes: A random-effects generalized ordered probit model. *Accid. Anal. Prev.* **2017**, *106*, 399–410. [CrossRef]
48. Se, C.; Champahom, T.; Jomnonkwao, S.; Karoonsoontawong, A.; Ratanavaraha, V. Temporal stability of factors influencing driver-injury severities in single-vehicle crashes: A correlated random parameters with heterogeneity in means and variances approach. *Anal. Methods Accid. Res.* **2021**, *32*, 100179. [CrossRef]

49. Huo, X.; Leng, J.; Hou, Q.; Yang, H. A Correlated Random Parameters Model with Heterogeneity in Means to Account for Unobserved Heterogeneity in Crash Frequency Analysis. *Transp. Res. Rec. J. Transp. Res. Board* **2020**, *2674*, 036119812092221. [CrossRef]
50. Yu, H.; Yuan, R.; Li, Z.; Zhang, G.; Ma, D.T. Identifying heterogeneous factors for driver injury severity variations in snow-related rural single-vehicle crashes. *Accid. Anal. Prev.* **2020**, *144*, 105587. [CrossRef] [PubMed]
51. Jung, S.; Qin, X.; Noyce, D.A. Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accid. Anal. Prev.* **2010**, *42*, 213–224. [CrossRef] [PubMed]
52. Jafarzadeh Ghoushchi, S.; Shaffiee Haghshenas, S.; Memarpour Ghiaci, A.; Guido, G.; Vitale, A. Road safety assessment and risks prioritization using an integrated SWARA and MARCOS approach under spherical fuzzy environment. *Neural Comput. Appl.* **2023**, *35*, 4549–4567. [CrossRef] [PubMed]
53. Shaheed, M.S.; Gkritza, K.; Zhang, W.; Hans, Z. A mixed logit analysis of two-vehicle crash seventies involving a motorcycle. *Accid. Anal. Prev.* **2013**, *61*, 119–128. [CrossRef] [PubMed]
54. PDPM. Thailand Department of Public Disaster Prevention and Mitigation. Available online: https://www.disaster.go.th/en/ (accessed on 1 April 2020).
55. Srikant, R.A.a.R. Fast algorithms for mining association rules. In Proceedings of the 20th International Conference on Very Large Data Bases, VLDB, Santiago, Chile, 12–15 September 1994; pp. 487–499.
56. Frawley, W.J.; Piatetsky-Shapiro, G.; Matheus, C.J. Knowledge Discovery in Databases. *AI Mag.* **1992**, *13*, 57. [CrossRef]
57. Demšar, J.; Curk, T.; Erjavec, A.; Gorup, C.; Hočevar, T.; Milutinovič, M.; Možina, M.; Polajnar, M.; Toplak, M.; Starič, A.; et al. Orange: Data mining toolbox in python. *J. Mach. Learn. Res.* **2013**, *14*, 2349–2353.
58. DLT. Department of Land Transportation. Available online: https://www.dlt.go.th/th/public-news/view.php?_did=2806 (accessed on 1 June 2021).
59. Mibileye. Available online: https://www.mobileye.com/ (accessed on 12 August 2023).
60. Yoffie, D.B. Mobileye: The Future of Driverless Cars. In *Harvard Business School Case*; Harvard Business Publishing: Boston, MA, USA, 2014.