Journal of Cyber Security and Risk Auditing Vol.2025, No.2

STAP
Smart Technologies Academic Press

Journal of Cyber Security and Risk Auditing

https://www.jcsra.thestap.com/

Unveiling the Causes of Fatal Road Accidents in Iraq: An Association Rule Mining Approach Using the Apriori Algorithm

Malath Riyadh Alboalebrah[1], Salam Al-augby[1]

[1]*Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq*

ARTICLE INFO

Article History

Received: 27-02-2025
Revised: 15-03-2025
Accepted: 28-03-2025
Published: 02-04-2025

Academic Editor:

Prof. Youakim Badr

Vol.2025, No.2

DOI:
https://doi.org/10.63180/jcsra.thestap.2025.2.1

ABSTRACT

With the increase in fatal accidents in Iraq, they have become a source of concern for both authorities and the public. Therefore, it has become necessary to conduct an analysis of these road accidents. This study aims to provide recommendations to responsible authorities after assessing the frequency of fatal traffic accidents and identifying the most common causes. This will provide actionable insights for decision-makers to formulate laws that allow for the reduction of these accidents and the reduction of human and economic losses. This paper applied data mining algorithms to three years of traffic fatal accident data in Iraq, excluding the Kurdistan Region. The results showed that people without driver's licenses and with primary school certificates were more likely to fail to wear seatbelts, making them a dangerous group. Married individuals aged 36-41 were also associated with fatal accidents. Based on the results, some recommendations were made to reduce these accidents.

Keywords: Data mining, Association rule, Apriori, Iraq, Fatal road accident.

How to cite the article

Riyadh Alboalebrah, M., & Al-augby, S. (2025). Unveiling the Causes of Fatal Road Accidents in Iraq: An Association Rule Mining Approach Using the Apriori Algorithm. Journal of Cyber Security and Risk Auditing, 2025(2), 1–11. https://doi.org/10.63180/jcsra.thestap.2025.2.1

1. Introduction

Traffic accidents are a global crisis, with approximately 1.19 million deaths annually and between 20 and 50 million people are injured and left with disabilities. Their repercussions extend far beyond this, as these accidents burden nations and families economically [1]. The situation in Iraq is also alarming. The Iraqi Ministry of Planning reported that fatal traffic accidents increased by 0.3% in 2023, with the number of deaths reaching 3,000[2]. Despite numerous studies on traffic accidents [3], few have delved into the factors associated with fatal accidents. This study explores the causative factors and the relationships between them, aiming to gain a deeper understanding of their causes. This study uses association rule mining with Apriori algorithms to highlight fatal accidents, a tool used to uncover hidden patterns in datasets [4].

Since traffic safety is an important issue in Iraq, this study will present preventive strategies that can reduce these accidents. Data on fatal traffic accidents in Iraq for three years was analyzed, the results were presented, and proposals were made that would significantly enhance safety measures [5]. Through this data analysis, fatal black spots can be identified more

*Corresponding author. Email: malathr.reayad@student.uokufa.edu.iq

accurately. Using big data analysis methods to uncover hidden rules and values is crucial, as the rule derived from the correlation model between factors reveals more hidden points.

The goal of using association rules instead of other methods [6]is the algorithm's ability to search for correlations between risk attributes. This algorithm can uncover valuable relationships between variables in a large dataset [7]. Therefore, this algorithm will help in better understanding fatal accidents. Since fatal accidents are a global concern, decoding these fatal accidents is of utmost importance. These results will reveal insights that differ from traditional statistical results [3].

The subsequent sections of this work are structured as follows. The literature review is in Sect. 2, followed by methodology in Sect. 3. The experimental results and discussion from implementation are presented in Sect. 4, and Sect. 5 concludes the paper.

2. Literature Review

Association rule mining has effectively uncovered hidden patterns and relationships in traffic accident data. Below are some key points from recent research focusing on fatal traffic accidents:

I. Mohamad et al. [8] 2023 This research utilizes the Apriori algorithm to examine the occurrence and primary factors contributing to fatal road accidents in Thailand. Through an analysis of road accident data spanning from 2015 to 2020, the study centers on drivers culpable for fatal incidents on highways, uncovering a web of interconnected variables that amplify the likelihood of fatalities.

The results underscore an escalated risk associated with traffic accidents linked to male drivers, speeding violations, motorcycle involvement, straight road configurations, dry road surfaces, and clear weather conditions.

R. Tamakloe et al.[9] 2024 With the increasing use of personal vehicles as a means of urban transportation, this study was conducted to analyze personal mobility device data in Seoul for five years, from 2017 to 2021, using a complex analysis and correlation rules. The results revealed three main clusters of accidents, with the incidence of fatal accidents being associated with dry road conditions, male personal mobility device users, and days of the week. Based on these findings, some measures were recommended to enhance personal mobility device safety.

C. Xu et al.[7] 2018 This study analyzed fatal accidents using association rules to investigate the contributing factors. The Apriori algorithm was used to discover association rules with detection rules, a common method applied to 126 serious traffic accidents in China between 2009 and 2013. The results indicated a correlation between driver behavior, vehicle condition, and road geometry. Accidents were also observed to be more frequent on weekends. The study emphasizes the importance of developing specific policies to reduce these serious accidents.

M. Tariq et al.[10] 2022 In this study, the Apriori algorithm, which is often widely used with pattern extraction Association rule mining, was used to extract useful patterns that can be used to support decision-makers. This algorithm was implemented on data from the state of Gujarat in Pakistan for the years 2018 to 2020. The results showed that speeding bicycles are often prone to accidents and are more dangerous at intersections in the afternoon during the middle of the week.

M. Shahin et al.[11] 2022This study analyzed traffic accident data from Isfahan Province, Iran. The data included 576 accidents from 2014. The data were analyzed using data mining techniques (association rules with the Apriori algorithm). The results showed that serious accidents are influenced by a wider range of factors than non-serious accidents. It is important to note that the use of seat belts reduced the severity of injuries, and that serious accidents particularly involved pedestrians. These findings have implications for developing strategies that can contribute to improving traffic safety.

A. Ziakopoulos et al.[12] 2023 The study employed association rule extraction with the Apriori algorithm to uncover underlying patterns among road users involved in crashes in Greece. The data was divided into mainland and island regions. The total dataset analyzed in the study comprised 41,541 serious crashes from 2015 to 2017. The results reveal associations between clear weather, urban road environments, male road users, and the presence of vulnerable road users in high-frequency crash injuries, offering insights into anticipated patterns in serious crash injuries.

From the above, it is clear that association rule mining has the potential to uncover the interconnections between factors that cause accidents [11]. The results of this algorithm can provide strategies to mitigate these accidents. This research highlights the interconnections between various factors, such as weather conditions, driver age, and the lack of a valid driver's license. We will focus on fatal accidents in particular, an important area that requires focus and an understanding of the factors associated with them.

3. Methodology

3.1 Dataset Description

The traffic accident database for Iraq, excluding the Kurdistan region, was sourced from the General Traffic Directorate and encompasses data from 2022 and 2023, supplemented by partial records from 2021. The refined dataset comprises 16 features and 10,050 records following an extensive data-cleaning process. Preliminary analyses suggest that traffic accidents are influenced by a multifaceted array of factors, including human-related variables (e.g., driver age), natural conditions (e.g., state of light), and road-specific attributes (e.g., type of road). The categorical variables and their respective levels are detailed in Table 1.

To facilitate subsequent analyses, the research team prepared the dataset by applying preprocessing steps as outlined below:

- Data Cleaning: This process involved identifying and rectifying incomplete entries and orthographic errors through applied data analysis and cleaning techniques. Tools such as Python libraries—namely Pandas and NumPy—were employed to enhance data integrity. Records with missing attribute values were systematically excluded, reducing the dataset from an initial 13,000 records to 10,050. After the data has been cleaned, the fatal accidents will be separated from the overall accident records and analyzed independently. This process yielded 1,852 fatal incidents.
- Data Conversion: The data was converted from categorical variables to binary variables, for analysis[8].

3.2 Methodology: The Apriori Algorithm and Association Rule Learning

The Apriori algorithm is a popular technique for extracting frequent item sets for mining association rules, and it is beneficial for analyzing large datasets[11]. The goal of this algorithm is to find items that frequently appear together and expand them into larger groups based on a certain threshold number. These frequent item sets help generate association rules, which in turn reveal relationships between data[14]. The Apriori algorithm is an ideal algorithm for uncovering hidden patterns and is widely used in applications such as shopping carts and recommendation systems.

The Association rule mining known as "if/then" statements to find patterns and relationships in large datasets[15][16][11][17]. This method gives out important associations between data elements, providing valuable insights for various applications. Association rule mining is useful in market basket analysis[18], healthcare[19], and traffic accident analysis. The Association rule mining relies on the FP-Growth and Apriori algorithm to identify frequent item sets, which aids in discovering rules[17]. Recently, this type of analysis has been applied in the field of road traffic safety[7]. The Apriori algorithm and the subsequent generation of association rules are explained in these steps:

1. **Determining the Minimum Support Threshold** :The minimum support threshold is a basic criterion that determines the number of times an item or set of items appears in a database to be considered significant. This value is typically expressed as a percentage of the total transactions and serves as a filter for selecting items worthy of further study[9].

2. **Finding Frequent Item sets :**The algorithm scans the entire database to determine how frequently each item occurs. This is known as support. Items that meet the support threshold are the first step toward the next process[17].
3. **Generation of Candidate item:** Based on the elements from the previous step, the algorithm generates larger item sets by systematically combining pairs of frequent item sets. This process adheres to the apriori principle, which states that all subsets of a frequent itemset must be frequent, eliminating unfeasible combinations early and improving computational efficiency. To estimate the scale of this endeavor, a lattice structure can be used to enumerate the full set of possible item sets.

   For example, in a dataset with items I= {a,b,c,d,e}, the item set lattice (as depicted in Figure 1) delineates all conceivable combinations, from singletons to the full set.

   In general, a dataset comprising k items can potentially yield 2^{k-1} itemsets, excluding the null set. Given that k may be considerably large in practical scenarios—such as transactional databases or bioinformatics applications—the search space of item sets expands exponentially, posing a significant computational challenge. The Apriori algorithm mitigates this complexity by restricting candidate generation to combinations of previously identified frequent item sets, ensuring that only promising candidates proceed to the support evaluation phase[20].

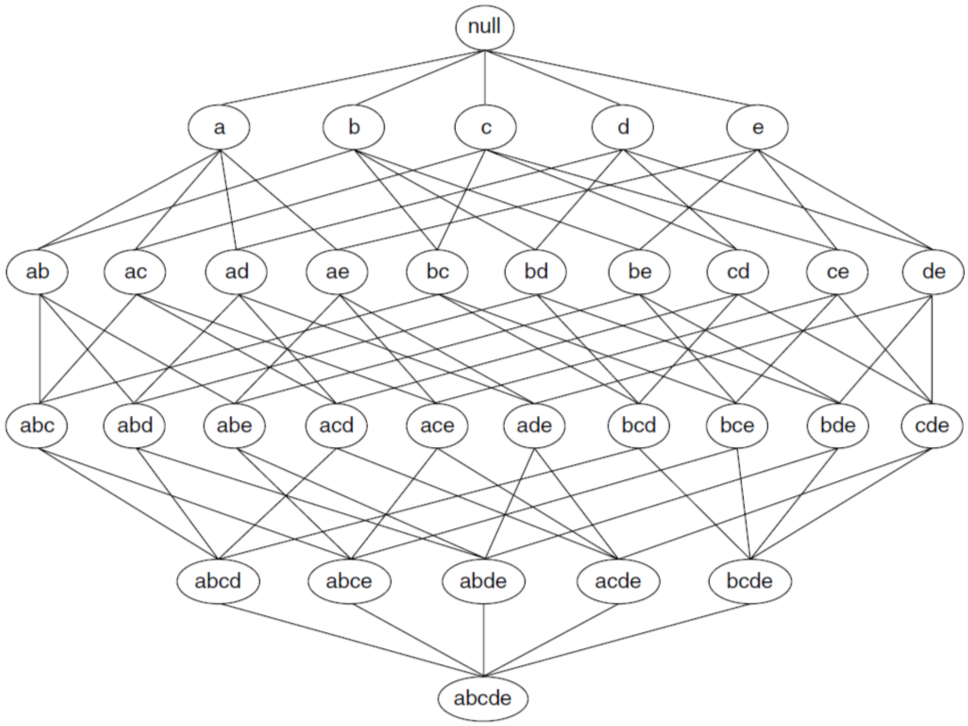


Figure 1. The itemset lattice [23].

4. Support Calculation for Candidate Items : The database is rescanned to calculate the support of each candidate itemset. Those exceeding the minimum support threshold are retained as frequent itemsets, while those falling below it are discarded. This pruning mechanism ensures that only sufficiently prevalent patterns are carried forward.

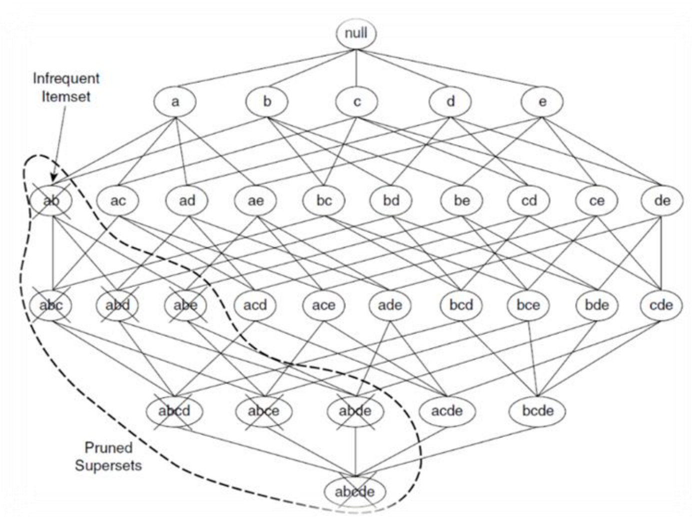


Figure 2. An illustration of spport by using pruning. [23]

5. Iterative Expansion of Item sets

The process iterates by generating increasingly larger item sets—progressing from pairs to triples, quadruples, and beyond—and repeating the candidate generation and support calculation steps. Iteration continues until no new frequent item sets can be identified, signaling the exhaustion of viable combinations under the specified support constraint.

6. Derivation of Association Rules

With the complete set of frequent item sets in hand, the algorithm proceeds to generate association rules of the form X→Y where X (the antecedent) and Y (the consequent) are disjoint subsets of a frequent itemset[9]. The strength of each rule is assessed using two key metrics:

Confidence: is used to determine the frequency of occurrence of a specific accident (Y) in the records containing a certain factor (X), i.e., the conditional probability of Y given X. It is a method for measuring the accuracy of the association rule and indicates the reliability of the rule. Suppose the rate of the number of records containing the factors X and Y to the number of records containing factor X[21], calculated as

$$C = \frac{P(x \cap y)}{P(x)} \qquad \textbf{... (1)}$$

Lift: is a measure used in association rule mining to evaluate the strength of the relationship between two items in a dataset[21]. It is calculated as follows:

$$L = \frac{P(x \cap y)}{P(x)P(y)} \qquad \text{...(2)}$$

Lift measures the impact of the presence of condition X on the likelihood of event Y occurring compared to its independent occurrence. If the lift is greater than 1, it means that the presence of X increases the likelihood of Y. If the lift is less than 1, it means that the presence of X decreases the likelihood of Y. If the lift equals 1, it means that the presence of X has no effect on the likelihood of Y[9].

7. Creation of Association Rules: After identifying frequent itemsets, the algorithm generates association rules. These rules divide the frequent itemset into two parts: antecedent (if) and consequent (then). The strength of these rules is evaluated using metrics such as confidence and lifte, which measure the likelihood and importance of associations[22]. This step is essential for discovering meaningful patterns and relationships within the data, providing practical insights for decision-making.

4. Results and Discussion

Fatal traffic accidents are a major concern for public safety in Iraq. This study looks into data on fatal accidents to determine the main factors that cause them. By using advanced data analysis techniques, we found important patterns and relationships that help us understand why these accidents happen. The data, which comes from the General Traffic Directorate and covers the years 2021 to 2023, gives a detailed look at what influences fatal accidents. Our findings aim to provide practical recommendations to reduce these accidents and improve road safety in Iraq. The result was obtained with a minimum support threshold of 0.162. Setting the minimum support at 16% ensures that no item or combination of items is considered frequent unless it appears in at least 300 accidents (16% of 1852 fatal accidents).

This procedure resulted in 1249 item sets and 14028 rules. Specific criteria were defined for evaluating association rules: Support (S) ≥ 16%, Confidence (C) ≥ 60%, and Lift (L) ≥ 1.2, aligning with criteria used in previous studies.

The analysis yielded 123 association rules. See figure 2

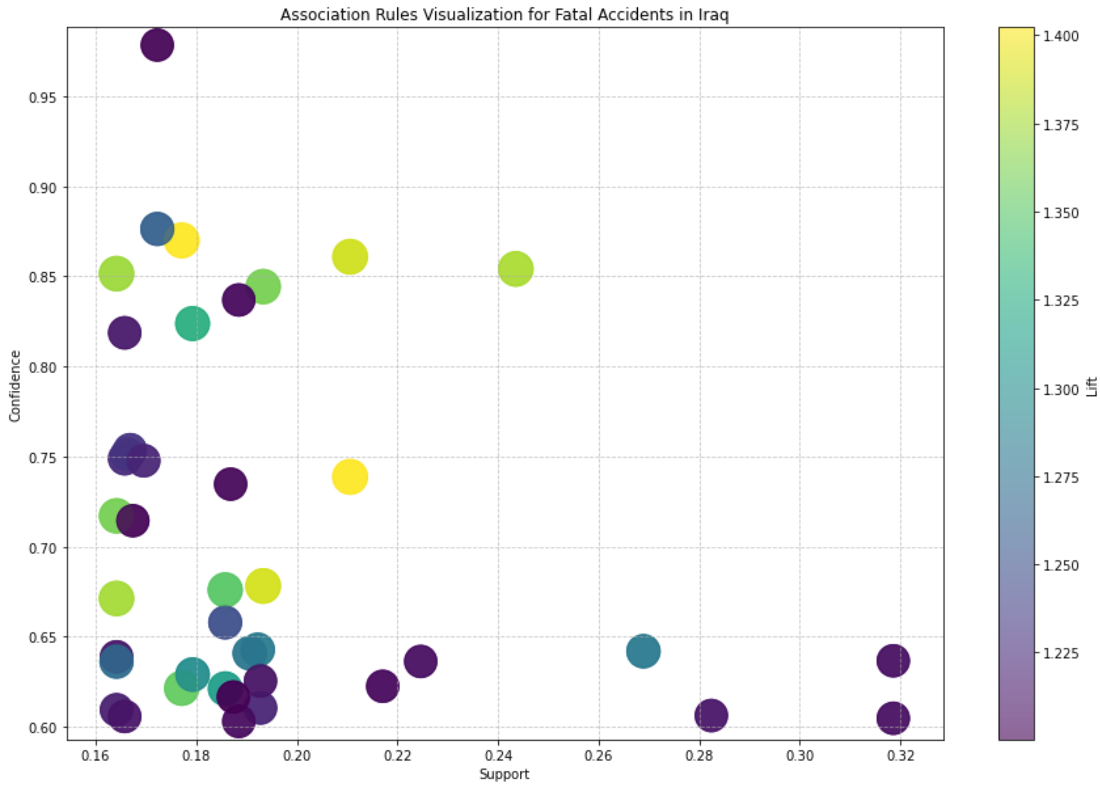


Figure 3. All rules with S=>16%, C=>60%, L=>1.2

The rules with the highest lift values were selected from the 123 rules that emerged after determining the support, confidence, and lift values. The rules were categorized based on the number of items in association rules

Table 2. List of 3-item association rules.

Rule	antecedents	consequents	S	C	L
1	AS=**Fatal** injuries, DB=Traffic violations	NA=Crash	0.24	0.85	1.376
2	AS=**Fatal** injuries, NA=Car rollover	DB=speed and inattention	0.16	0.74	1.22

The first rule (**Rule 1**) indicates a strong association between the presence of traffic violations (DB=Traffic violations) in fatal injury incidents (AS=Fatal injuries) and the occurrence of a crash (NA=Crash). This rule has a support of 0.24, meaning that 24% of the analyzed fatal accidents involved this combination of factors. The confidence of 0.85 suggests that 85% of incidents with traffic violations and fatal injuries resulted in a crash, while the lift of 1.376 indicates that this combination is 37.6% more likely to occur than if the variables were independent. This finding highlights the critical role of traffic violations in exacerbating crash severity, particularly in fatal outcomes.

The second rule (**Rule 2**) associates fatal injuries in car rollover incidents (AS=Fatal injuries, NA=Car rollover) with driver behavior characterized by speed and inattention (DB=Speed and inattention). With a support of 0.16 and a confidence of 0.74, this rule indicates that 16% of fatal accidents involved this pattern, and 74% of fatal rollover incidents were linked to speed and inattention. The lift of 1.22 indicates an approximate 22% increase in the likelihood of the association between 'DB=Speed and inattention' and 'AS=Fatal injuries, NA=Car rollover' compared to a random occurrence, where the difference is calculated as a percentage of the value exceeding 1 (1.22 - 1 = 0.22). This result underscores the combined risk of speeding and inattentiveness in rollover crashes, which are particularly lethal, and supports the need for targeted interventions such as awareness campaigns and speed control measures at high-risk locations.

These findings provide actionable insights for policymakers in Iraq, emphasizing the importance of addressing human factors—such as traffic violations, speeding, and inattention—to reduce the incidence of fatal traffic accidents.

Table 3. List of 4-item association rules.

Rule	antecedents	consequents	S	C	L
1	AR=Out of town, AS=**Fatal** injuries, DB=Traffic violations	NA=Crash	0.17	0.87	1.40
2	AR=Out of town, DB=Traffic violations	AS=**Fatal** injuries, NA=Crash	0.17	0.87	1.40
3	AS=**Fatal** injuries, DB=Traffic violations	SB_=Not in use, NA=Crash	0.21	0.73	1.40
4	DB=Traffic violations	AS=**Fatal** injuries, SB_=Not in use, 'NA=Crash	0.21	0.73	1.40
5	AS=**Fatal** injuries, DB=Traffic violations	MS= Married, NA=Crash	0.19	0.67	1.38
6	DB=Traffic violations, SB_=Not in use,	AS=**Fatal** injuries, NA=Crash	0.21	0.86	1.38
7	MS= Married, DB=Traffic violations	AS=**Fatal** injuries, NA=Crash	0.19	0.84	1.36
8	AS=**Fatal** injuries, DB=Traffic violations	AR=Out of town, NA=Crash	0.17	0.62	1.35
9	CA=moving car, DB=Traffic violations, AS=**Fatal** injuries	A=Crash	0.17	0.82	1.32
10	CA=moving car, DB=Traffic violations	AS=**Fatal** injuries, NA=Crash	0.17	0.82	1.32
11	Certificate=primary, NA=Crash	AS=**Fatal** injuries, LS=NO license	0.19	0.64	1.28
12	AS=**Fatal** injuries, Certificate=primary, SB_=Not in use	LS=NO license	0.26	0.64	1.28
13	Age=36-41, AS=**Fatal** injuries})	CA=moving car, MS= Married	0.17	0.87	1.26
14	AS=**Fatal** injuries, NA=Car rollover	AR=Out of town, CA=moving car,	0.16	0.75	1.23
15	AS=**Fatal** injuries, NA=Car rollover	AR=Out of town, MS= Married	0.16	0.75	1.23
16	Year =2023, SB=Not in use, NA=Crash	AS=**Fatal** injuries, LS=NO license	0.16	0.60	1.21
17	AS=**Fatal** injuries, SB_=Not in use , NA=Crash	LS=NO license	0.31	0.60	1.20
18	AS=**Fatal** injuries, LS=NO license	SB_=Not in use, NA=Crash	0.31	0.63	1.20
19	SB_=Not in use, NA=Crash	AS=**Fatal** injuries, LS=NO license	0.31	0.60	1.20
20	Age=36-41, CA=moving car	AS=**Fatal** injuries, MS= Married	0.17	0.97	1.20

From Table 3, the Association Between Traffic Violations and Fatal Accidents (**Rules 1, 2, 3, 4, 5, 6, 7, 8, 9, 10**):
Pattern: Rules involving "DB=Traffic violations" consistently show a strong association with fatal injuries (AS=Fatal injuries) and crashes (NA=Crash). For example:
Rules 1 and 2: Traffic violations in out-of-town areas (AR=Out of town) are associated with fatal injuries and crashes (S=0.17, C=0.87, L=1.40). The high confidence (0.87) indicates that 87% of incidents involving violations outside of town result in fatal injuries and crashes, with a 40% increased likelihood (Lift=1.40).

Rules 3 and 4: Traffic violations combined with non-use of seatbelts (SB_=Not in use) are linked to fatal injuries and crashes (S=0.21, C=0.73, L=1.40). This suggests that failing to wear a seatbelt exacerbates accident severity when violations are present.

Rules 9 and 10: Traffic violations with a moving car (CA=moving car) are also associated with fatal injuries (S=0.17, C=0.82, L=1.32).

Interpretation: Traffic violations represent a primary risk factor that increases the likelihood of fatal accidents, particularly in out-of-town areas or when seatbelts are not used. The high lift (1.40) confirms that this relationship is not random, highlighting the need for stricter traffic enforcement.

Non-Use of Seatbelts and Lack of Licenses (**Rules 8, 10, 17, 18, 19**):
Pattern: Rules involving "SB_=Not in use" (non-use of seatbelts) and "LS=NO license" (no driving license) demonstrate a strong association with fatal accidents.
Rules 17, 18, and 19: Non-use of seatbelts combined with the absence of a driving license is linked to fatal injuries and crashes (S=0.31, C=0.60-0.63, L=1.20). The high support (0.31) indicates that this pattern is prevalent in 31% of accidents.

Rule 12: Individuals with primary education (Certificate=primary) who do not use seatbelts and lack driving licenses are associated with fatal injuries (S=0.26, C=0.64, L=1.28).

Interpretation: These rules identify a high-risk group: drivers without licenses, particularly those with primary education, tend to neglect seatbelt use, increasing the risk of fatal accidents. The lift (1.20-1.28) confirms a positive association, calling for targeted interventions such as awareness campaigns and stricter penalties.

Social Status and Accidents (**Rules 5, 10, 17, 18, 19**):
Pattern: Married individuals (MS=Married) appear in several rules linked to fatal accidents.
Rules 5 and 10: Married individuals committing traffic violations are associated with fatal injuries (S=0.19, C=0.67-0.84, L=1.36-1.38).

Rules 13 and 20: Married individuals aged 36-41 years with a moving car are linked to fatal injuries (S=0.17, C=0.87-0.97, L=1.20-1.26).

Rule 15: Married individuals involved in rollover accidents (NA=Car rollover) outside urban areas are associated with fatal injuries (S=0.16, C=0.75, L=1.23).

Interpretation: Married individuals, mainly those aged 36-41, may be more prone to fatal accidents due to their driving patterns (e.g., out-of-town travel or violations). The high confidence (up to 0.97 in Rule 20) indicates the strength of this relationship, suggesting the need for awareness programs targeting this demographic group.

Education and Licensing (**Rule 11**):
Pattern: Individuals with primary education (Certificate=primary) involved in crashes are associated with the absence of a driving license (S=0.19, C=0.64, L=1.28).

Interpretation: This rule highlights a link between low education (primary level) and the lack of a driving license, increasing the risk of fatal accidents. This underscores the importance of traffic education and interventions targeting individuals with limited education.

Recent Accidents (**Rule 16**):
Pattern: In 2023, accidents involving non-use of seatbelts and crashes are associated with drivers without licenses (S=0.16, C=0.60, L=1.21).

Interpretation: This rule indicates the persistence of this issue into 2023, reinforcing the need for continued enforcement of regulations and enhanced traffic awareness.

General Findings
Traffic Violations as a Primary Factor: Traffic violations are a recurring factor in rules with high lift values (1.32-1.40), confirming their role in increasing fatal accident risks, especially when combined with non-use of seatbelts or out-of-town driving.

High-Risk Groups: Drivers without licenses, particularly those with primary education, form a high-risk group due to their tendency to avoid seatbelt use, elevating the likelihood of fatal crashes.

Demographic Patterns: Married individuals aged 36-41 years exhibit recurring patterns in fatal accidents, necessitating targeted interventions for this group.

Importance of Seatbelts: The rules emphasize that non-use of seatbelts heightens the risk of fatal accidents, particularly among unlicensed drivers.

Table 4. List of 5&6-item association rules.

Rule	antecedents	consequents	S	C	L
1	AS=**Fatal** injuries, DB=Traffic violations, SB= in use	MS= Married, NA=Crash	0.16	0.67	1.374
2	MS= Married, DB=Traffic violations, SB=Not in use	AS=**Fatal** injuries, NA=Crash	0.16	0.85	1.372
3	AS=**Fatal** injuries, Certificate=primary, SB_=N in use, NA=Crash}})	LS=NO license,	0.18	0.675	1.350
4	Certificate=primary, SB_=Not in use, State of Light=Day	AS=**Fatal** injuries, LS=NO license, })	0.16	0.63	1.27
5	AS=**Fatal** injuries, LS=NO license, Certificate=primary,	SB=Not in use, NA=Crash	0.18	0.65	1.24
6	NA=Crash, CA=moving car, LS= Valid	AS=**Fatal** injuries, AR=Out of town, MS= Married	0.16	0.74	1.22
7	AR=Out of town, CA=moving car, MOT=Priva AS=**Fatal** injuries,	MS= Married, DB=speed and inattention	0.19	0.610	1.22
8	MOT=Private, AS=**Fatal** injuries, MS= Married, State of Light=Day	AR=Out of town, CA=moving car	0.18	0.73	1.20

Association of Traffic Violations, Non-Use of Seatbelts, and Marital Status with Fatal Crashes (**Rules 1 and 2**):
Rules 1 and 2: These rules reveal a significant linkage between fatal injuries (AS=Fatal injuries), traffic violations (DB=Traffic violations), non-use of seatbelts (SB=Not in use), and marital status (MS=Married) with crash incidents (NA=Crash). Rule 8 shows that fatal injuries, traffic violations, and non-use of seatbelts are associated with married individuals involved in crashes (S=0.16, C=0.67, L=1.374), while Rule 9 indicates that married individuals committing traffic violations without seatbelts are highly likely to experience fatal injuries and crashes (S=0.16, C=0.85, L=1.372). The confidence values (0.67 and 0.85) suggest that 67% and 85% of such cases result in the specified outcomes, with lifts of 1.374 and 1.372, indicating a 37.4% and 37.2% increased likelihood compared to random occurrence. Interpretation: These findings highlight the compounded risk posed by traffic violations and non-use of seatbelts among married drivers, emphasizing the need for targeted enforcement and awareness campaigns to promote seatbelt compliance within this demographic.

Impact of Education Level, Seatbelt Non-use, Licensing, and Daytime Conditions on Fatal Accidents (**Rules 3, 4, and 5**):
Rule 3: This rule links fatal injuries, primary education (Certificate=primary), non-use of seatbelts, and crashes with the absence of a driving license (LS=NO license) (S=0.18, C=0.675, L=1.350). The confidence of 0.675 indicates that 67.5% of such cases involve unlicensed drivers, with a 35% increased likelihood (Lift=1.350).

Rule 4: Individuals with primary education, non-use of seatbelts, and daytime conditions (State of Light=Day) are associated with fatal injuries and unlicensed drivers (S=0.16, C=0.63, L=1.27). These accidents during the daytime suggest that good visibility might lead to overconfidence, contributing to negligence in safety measures like seatbelt use, resulting in a 27% elevated risk.

Rule 5: Fatal injuries, the absence of a license, primary education, non-use of seatbelts, and crashes are interconnected (S=0.18, C=0.65, L=1.24), with a 24% increased likelihood.

Interpretation: These rules identify a high-risk cohort with low educational attainment, lack of licensing, and non-use of seatbelts. The presence of daytime conditions in Rule 20 further indicates that accidents during daylight hours may be exacerbated by behavioral factors such as overconfidence or increased traffic density, highlighting the need for safety interventions even in optimal visibility conditions.

Influence of Vehicle Type, Licensing, Driving Behavior, and Daytime Conditions on Fatal Accidents (**Rules 6, 7 and 8**):

Rule 6: Crashes involving moving cars (CA=moving car) with valid licenses (LS=Valid) are associated with fatal injuries, out-of-town locations (AR=Out of town), and married individuals (S=0.16, C=0.74, L=1.22), indicating a 22% increased likelihood.

Rule 7: Out-of-town accidents involving private vehicles (MOT=Private), fatal injuries, married individuals, and speed with inattention (DB=speed and inattention) show a similar pattern (S=0.19, C=0.610, L=1.22).

Rule 8: Private vehicles, fatal injuries, married individuals, and daytime conditions (State of Light=Day) are linked to out-of-town moving car accidents (S=0.18, C=0.73, L=1.20). The daytime setting suggests that such accidents may occur due to increased traffic volume or driver complacency in good lighting conditions, with a 20% increased likelihood.

Interpretation: These rules indicate that married drivers operating private vehicles, particularly under speed and inattention, are at heightened risk of fatal accidents outside town, even with valid licenses. The occurrence of these incidents during the daytime (Rule 8) underscores the potential role of overconfidence or higher traffic density in exacerbating risks, necessitating focused interventions on speed management and driver awareness.

5.Conclusion

This study applied the Association rule mining with Apriori algorithm to the Iraq Traffic Accident Dataset (ITAD). The results revealed important patterns that shed light on the factors contributing to fatal traffic accidents and the extent of their interconnectedness. Among the most important findings was a strong link between traffic violations and failure to wear a seatbelt, which significantly increases the risk of fatal accidents, especially in areas outside of cities. Drivers without licenses, particularly those with only a primary education, were also found to be at greater risk due to their frequent neglect of seatbelt use. Furthermore, married individuals between the ages of 36 and 41 were more frequently associated with fatal accidents, highlighting the need for safety interventions for this group. The presence of diurnal conditions (light condition = day) in certain patterns, such as those involving unlicensed drivers and private vehicle operators, suggests that good visibility can lead to overconfidence, or that heavy traffic can increase risk. Based on these findings, strategies can be proposed to reduce fatal crashes:

1. Tighten traffic violation laws and seatbelt usage, with a particular focus on areas outside a town. and daytime hours to reduce complacency during optimal visibility conditions.
2. Launching education programs to emphasize the importance of seatbelt usage and the importance of obtaining valid driver's licenses, with a particular focus on individuals with limited education to reduce their high-risk behavior.
3. Conducting awareness and training initiatives specifically targeting married individuals between the ages of 36 and 41 to reduce the risk of accidents within this age group by including guidelines on safe driving practices.
4. Implementing speed monitoring and awareness campaigns to address driver negligence, particularly among private vehicle operators, with a focus on daytime driving.
5. Strengthening licensing laws to ensure all drivers have valid licenses and introducing new methods for testing driver licenses to ensure that driver license laws are effective and reduce the incidence of further fatal accidents.

Future research should incorporate data from the Kurdistan region to provide a more comprehensive analysis. In conclusion, this study offers valuable evidence-based insights for policymakers aimed at reducing fatal traffic accidents, emphasizing the critical role of improved legislation, preventive interventions, and context-specific safety measure.

References

[1] “World Health Organization(WHO), Road traffic injuries.” Accessed: Oct. 29, 2024. [Online]. Available: https://www.who.int/publications/i/item/9789241565684

[2] “Republic of Iraq - Ministry of Planning, Increase in traffic accidents.” Accessed: Oct. 29, 2024. [Online]. Available: https://mop.gov.iq/archives/23786

[3] H. H. Joni, A. A. Mohammed, and A. A. Shakir, “Classification of traffic accidents datasets between 2003–2017 in Iraq,” Data Brief, vol. 28, Feb. 2020, doi: 10.1016/j.dib.2019.104902.

[4] J. Han, M. Kamber, and J. Pei, “Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems),” 2011.

[5] S. S. Bhowmick, “Association Rule Mining: A Survey,” 2003. [Online]. Available: https://www.researchgate.net/publication/244235958

[6] P. Abdullah and T. Sipos, "Drivers' Behavior and Traffic Accident Analysis Using Decision Tree Method," Sustainability (Switzerland), vol. 14, no. 18, Sep. 2022, doi: 10.3390/su141811339.

[7] C. Xu, J. Bao, C. Wang, and P. Liu, "Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China," J Safety Res, vol. 67, pp. 65–75, Dec. 2018, doi: 10.1016/j.jsr.2018.09.013.

[8] I. Mohamad, R. Kasemsri, V. Ratanavaraha, and S. Jomnonkwao, "Application of the Apriori Algorithm for Traffic Crash Analysis in Thailand," Safety, vol. 9, no. 3, Sep. 2023, doi: 10.3390/safety9030058.

[9] R. Tamakloe, K. Zhang, A. Hossain, I. Kim, and S. H. Park, "Critical risk factors associated with fatal/severe crash outcomes in personal mobility device rider at-fault crashes: A two-step inter-cluster rule mining technique," Accid Anal Prev, vol. 199, May 2024, doi: 10.1016/j.aap.2024.107527.

[10] M. Tariq, N. Q. Mehmood, and S. Z. Mahfooz, "Discovering associated factors behind road accidents using association rule mining: A case study from Gujarat, Pakistan," World Journal of Advanced Research and Reviews, vol. 15, no. 3, pp. 001–011, Sep. 2022, doi: 10.30574/wjarr.2022.15.3.0885.

[11] M. Shahin, M. R. Heidari Iman, M. Kaushik, R. Sharma, T. Ghasempouri, and D. Draheim, "Exploring Factors in a Crossroad Dataset Using Cluster-Based Association Rule Mining," in Procedia Computer Science, Elsevier B.V., 2022, pp. 231–238. doi: 10.1016/j.procs.2022.03.032.

[12] A. Ziakopoulos, E. Michelaraki, D. Nikolaou, K. Folla, and G. Yannis, "Association Rule Mining for Island and Mainland Road Crash Injuries in Greece," Transportation Research Procedia, vol. 72, pp. 163–170, 2023, doi: https://doi.org/10.1016/j.trpro.2023.11.390.

[13] T. J. Sun, X. F. Huang, F. K. Xie, J. Zhang, X. H. Jiang, and A. Y. Yu, "Road traffic mortality in Zunyi city, China: A 10 – year data analysis (2013–2022)," Chinese Journal of Traumatology, Nov. 2023, doi: 10.1016/J.CJTEE.2023.09.007.

[14] X. Zhang, Y. Bian, X. Zhao, J. Huang, and Z. Liu, "The association of road, traffic, and environmental factors with wrong-way riding behavior on shared e-bikes," J Clean Prod, vol. 470, Sep. 2024, doi: 10.1016/j.jclepro.2024.143294.

[15] T. A. Kumbhare Santosh V Chobe, "An Overview of Association Rule Mining Algorithms." [Online]. Available: www.ijcsit.com

[16] M. H. Santoso, "Application of Association Rule Method Using Apriori Algorithm to Find Sales Patterns Case Study of Indomaret Tanjung Anom," Brilliance: Research of Artificial Intelligence, vol. 1, no. 2, pp. 54–66, Dec. 2021, doi: 10.47709/brilliance.v1i2.1228.

[17] V. Kotu and B. Deshpande, "Association Analysis," in Predictive Analytics and Data Mining, Elsevier, 2015, pp. 195–216. doi: 10.1016/b978-0-12-801460-8.00006-9.

[18] Y. Liu, "Application and practice of data mining techniques," Applied and Computational Engineering, vol. 4, no. 1, pp. 237–241, May 2023, doi: 10.54254/2755-2721/4/20230456.

[19] T. M. Alam et al., "A novel framework for prognostic factors identification of malignant mesothelioma through association rule mining," Biomed Signal Process Control, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102726.

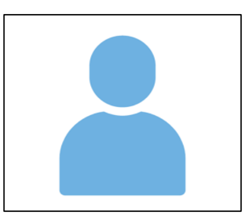
[20] J. Han, J. Pei, and H. Tong, "Data Mining: Concepts and Techniques," 2023.

[21] W. Zhong, J. Yuan, Y. Ren, and L. Du, "Characterization of Urban Road Traffic Accidents based on Data Mining," in 2022 12th International Workshop on Computer Science and Engineering, WCSE 2022, International Workshop on Computer Science and Engineering (WCSE), 2022, pp. 117–125. doi: 10.18178/wcse.2022.06.018.

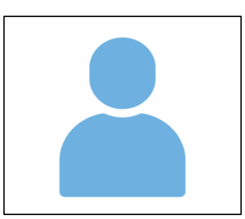
[22] S. K. Sivasankaran, P. Natarajan, and V. Balasubramanian, "Identifying Patterns of Pedestrian Crashes in Urban Metropolitan Roads in India using Association Rule Mining," in Transportation Research Procedia, Elsevier B.V., 2020, pp. 3496–3507. doi: 10.1016/j.trpro.2020.08.102.

[23] P.-Nin. Tan, Michael. Steinbach, and Vipin. Kumar, Introduction to data mining. Pearson, 2018.

Biographies



Malath Riyadh Alboalebrah. Student master in Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq. malathr.reayad@student.uokufa.edu.iq.



Salam Al-augby. He work in as lecturer in Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq. salam.alaugby@uokufa.edu.iq