

## Applications of Data Mining in Mitigating Fire Accidents Based on Association Rules

<https://doi.org/10.3991/ijim.v15i12.22687>

Ibrahim Nasir Mahmood <sup>(✉)</sup>, Hussein Ali Khuodhur Aliedane  
Basra University for Oil and Gas, Basra, Iraq  
ibrahim.mahmood@buog.edu.iq

Mustafa Ali Abuzaraida  
Universiti Utara Malaysia, Sintok, Malaysia

**Abstract**—Due to the increased rate of fire accidents which cause many damages and losses to people souls, material, and property in Basra city. The necessity of analyzing and mining the data of the fire accidents became an urgent need to find a solution. The need increased for a solution that helps to mitigate and reduce the number of accidents. In this paper, data mining techniques and applications including data preprocessing, data cleaning, and data exploration have been applied. Data mining applications is performed to analyze and discover the hidden knowledge in ten years of data (fire accidents happened from 2010 – 2019) which is approximately 20k record of accidents. These data mining techniques along with the association rules algorithm is applied on the dataset. The applied approach and techniques resulted in discovering the patterns and the nature of the fire accidents in Basra city. It also helped to reach to recommendations and resolutions for mitigating the fire accidents and its occurrence rate.

**Keywords**—Data Mining, Association Rules, Fire Accidents Analysis, Data Analysis, Data Cleaning

### 1 Introduction

The Issue of increased rate of fire accidents in Basra city requires a technical approach and research to mitigate and help find solutions to decrease and to avoid the high rate of accidents. Basra is a city located south of Iraq, and is considered as the second largest city in Iraq. It is also known as the center of oil production and refinery. Because of its strategic location on the Arab Gulf, it represents the main port of Iraq and the center of transportation, storage, and trade. So, it is an important city as it serves as the backbone for the national economy with a population of over 2.6 million and area of 19070 km<sup>2</sup>.

The current situation of fire accidents prevention in Basra lacks the proactive strategy and techniques to identify root causes, high impact locations, distribution pattern of response units, and seasonable nature of the accidents. Therefore, the applications

of data mining techniques and algorithms is applied to data of ten years of fire accidents in Basra city. The data size is approximately 20k records which was obtained from the civil defence department at the city and represents the data of fire accidents for the last decade (2010 – 2019).

The Main objective of this research is to determine the reason (root cause) causing the fire accidents to happen in Basra city. In addition, the locations where most of the accidents happen and for which reason. The time / season on which these accidents occur and for which reason and what's the highest rate areas. All this knowledge need to be discovered in order to help mitigate and lower the high rate of fire accidents happening recently in Basra city.

Furthermore, the discovered knowledge will help the related authorities to discover weak points in safety and response plans and apply corrective actions for future resource planning and distribution. Therefore, more lives can be saved, more damage can be avoided, and generating more cost, time, and efforts saving for all the related parties affected by fire accidents in Basra city.

Applying data mining techniques and algorithms on this amount of data helped to discover patterns of the accidents and the root cause for it. In addition, it helped to discover the places (city center and its counties) with high impact with high impact factor causing the accidents. Furthermore, the seasons on which the accidents rate increase and for which factors.

The obtained results supported the recommendations for the authorities to plan the distribution of the safety response units along with other solutions and recommendations presented at the end of the paper.to the related parties and authorities which support proactivity and designing of clear and specific strategy to avoid fire accidents in Basra and therefore, decrease their occurrence.

## **2 Data Mining Approach**

Data mining is an area of computer science where knowledge and patterns can be derived from large amount of data. In the field of data mining, the non-organized, non-structured, and messy data can be a great source for discovering hidden knowledge that can solve prolonged issues and problem. Examples of the issues that can be solved by applying the data mining techniques include cost savings and avoidance, lowering accidents rates, diseases spreads patterns, and etc. [1].

Data Mining involve several approaches and techniques that result in knowledge extraction and trends discovering. It involves several steps and phases. It includes identifying the domain of the problem to be solved, selecting of the target data to be processed, and preprocessing of the data in order to prepare it to be used for knowledge extraction [2]. Its phases also involve extracting the knowledge by applying the data mining techniques, algorithms, interoperation and evaluation of the obtained results by filtering, visualizing or interpreting [2].

Data mining starts with obtaining the dataset (database) which the knowledge needs to be extracted of. Not all the data (attributes / features) in the dataset might be needed during the data mining process. Thus, the obtained dataset will be processed to

extract only the target data that is actual data which the main goal is dealing with. So, some of the attributes / features in the data set can be not included in the target data. The target data then goes through a preprocessing phase which includes data cleaning [3]. The preprocessing phase could involve outlier detection and removal, detection and estimating of missing values, and correcting syntax mistakes in data [4]. Then the obtained structured and clean data can be analysed and explored using various data mining algorithms and techniques.

Data mining approach involve several techniques and algorithms. The selecting of specific technique or algorithm is based on the target purpose and the application type to which it is applied. The data mining techniques include classifying, clustering, sorting and analysing algorithms. For instance, decision trees, association rules, clustering and analysing algorithms [5]. Thus, data mining is a new method to define trends and patterns in datasets which facilities the management of these patterns and trend and improves suggested solutions for prolonged issues [7].

### 3 Methodology

#### 3.1 Dataset details

There is no current online available database for the fire accidents in Iraq. The database of the fire accidents in Basra city is stored as MS Excel file due to the lack of the IT infrastructure in the fire safety department. The process of accessing of these data require several authority and legal approvals and for research purposes only. So, one of the challenges to undergo such a research is the dataset availability.

The obtained dataset contains all the data needed to perform this research. It includes the location, the date, the reason, the safety material used, the type, and severity level of an accident. It also consisted of approximately 20k records of data representing the fire accidents occurred in Basra city from 2010 – 2019 (data of one decade). Each record represents a single fire accident occurred within Basra governorate borders.

The dataset of the fire accidents in Basra city included 8 attributes which are the accident id, accident type (household, vehicle, workshop), accident address, accident date, accident area, accident cause, human damage, and physical damage. The type of attributes included numeric, text, and date values. Table (1) shows a sample of these attributes.

**Table 1.** Sample of the Fire Accidents Dataset

Accident Type	Date	Location	Accident Cause	Damage
Vehicle	01/01/2010	Abu Alkhassib	Electrical Impact	Minor
House	01/01/2010	Abu Alkhassib	Cigarettes	Moderate
Apartment	01/01/2010	City Center	Electrical Impact	Minor
House	03/01/2010	Qurna	Gas Leak	Minor
House	03/01/2010	City Center	Children Messing	Moderate
Market	03/01/2010	Hartha	Electrical Impact	Major

### **3.2 Data mining platform and tools**

The data mining platform selected for this study is Orange platform which is developed by Bioinformatics Lab at University of Ljubljana, Slovenia, in collaboration with the open source community [6]. This platform was chosen because it supports the processing data written in Arabic. It also has powerful tools that utilized in the preprocessing of the data. In addition, it contains the implementation of the needed and required data mining algorithms and techniques. Besides that, it contains powerful tools of interactive data visualization and visual programming [6].

### **3.3 Data preprocessing**

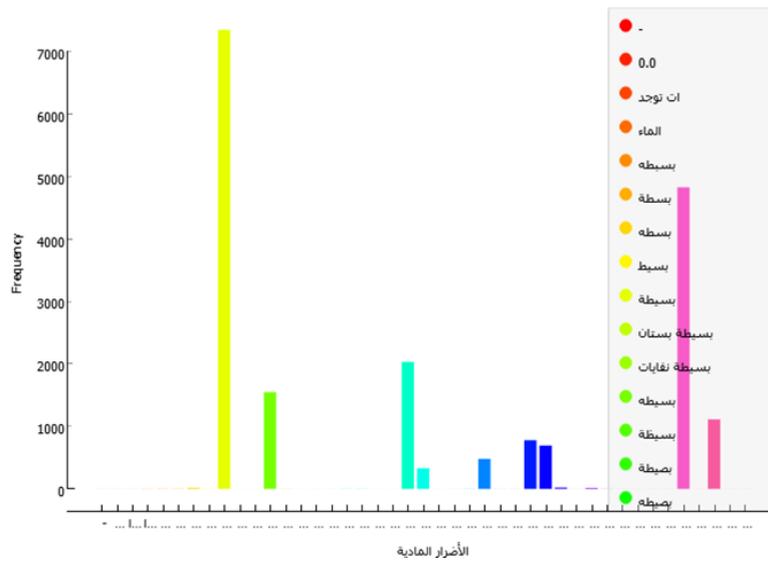
Data preprocessing phase is the most important step in data mining approach. Data preprocessing reserve about 60% of the total time of data mining project [2]. Therefore, it is the most important and effective phase. It can be accomplished by applying several techniques including data cleaning, data reduction, data transforming, and data integration [2]. During data cleaning phase, missing values are filled, noise is removed, and any inconsistencies are solved [2]. Data reduction include removing of repeated instances and discretion of continues attributes [2]. Data transformation is the method to transform text or graphical data into a process able form by generalizing, normalizing, or scaling the data [2].

The database for the fire accidents in Basra city is received in MS Excel format. It contains 20k record covering the accidents occurred in Basra city for ten years 2010 – 2019. In addition, All the data features/attributes are written in Arabic language. The examined dataset looked not clean and non-organized. It contained missing values, syntax errors, duplication due to syntax mistakes, and extra spaces between words. So, the preprocessing of this dataset involved data cleaning activities which such as filling in the missing values, correcting the syntax mistakes which caused duplication in the names of locations, accidents causes, and severity levels. Also, all the extra spaces have been processed and removed as seen in Figure 2. Moreover, the Date attribute has been transformed from its date value to a season value in order to reduce data and make it more logical to be classified and distributed.

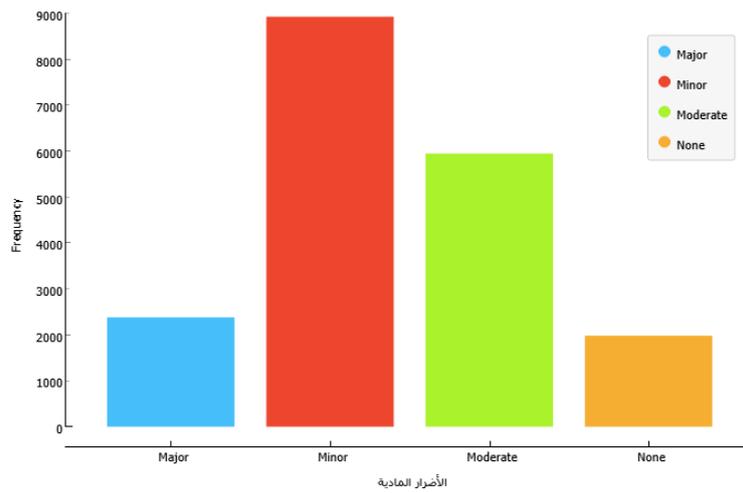
The data preprocessing tasks have been accomplished by utilizing MS Excel formulas and Orange interactive visualization tools. The results of this phase led to well-organized data which can be analysed logically to obtain actual vision and results. Thus, the data preprocessing is the most important step in data mining process. Figure 3 shows the clean data for the physical damage attributes where the translation of the values is (Minor, Moderate, Major, and None).

**Table 2.** Sample of Dirty Data Values

Damage	Cause	Date	Location	Incident ID
Minnor	Electrical Impact	30/06/2016	House	11642
Majr	Lectrcal impact	06/07/2016	Cables	11675
Mnior	Electrical loads	21/07/2016	Cables	11783
Majjor	Ciggarratees	24/07/2016	Generator	11814
Non	Electrical Impact	29/07/2016	House	11842



**Fig. 1.** Damage level data before preprocessing



**Fig. 2.** Damage level data after preprocessing

### 3.4 Data exploration

After cleaning and organizing the fire accidents dataset, the data became more understandable and usable. The data exploration is the process of diving into the data to accomplish basic understanding of the dataset. It is very useful to understand the structure of the dataset, the relationships between the variables within the dataset, the distribution and structure of values, and the overall characteristics of the dataset.

It is also considered as a refinery phase for the data cleaning and preprocessing phase as any of missed faults in the data can be rediscovered and notices during the data exploration process. The visualization tools such as the Box Plot, the Distributions, and the Scatter Plot within Orange platform have been utilized in achieving this phase and the obtained results is discussed in the next section.

### 3.5 Association rules analysis

The association rules analysis is an approach which helps to discover relationship between events. It states that the occurrence of two or more events is related with certain probability. It helps to discover relationships between unrelated data in a dataset. Which means it finds the relationships between objects that frequently occurred or used together [1]. Association rules mining is one of the well accepted used in data mining. It is utilized to find the frequent item in large dataset. It is very important for decision making and problem analysis [8].

The association rules algorithm works by identifying and analysing of co-occurrence in a dataset. It defines the if-then relationships between items in the dataset. It works by discovering all sets of items that have support greater than the minimum support , and it generates rule based on the largest item sets that have a confidence greater than the minimum confidence [1]. Support is a sign of the occurrence frequency of an item in a dataset. Confidence is the number of trues of the if – then statement.

To clarify the idea of the association rules analysis, “let  $I$  be an item set. Let  $D$ , the task-related data, be a set of database transactions where each transaction  $T$  is a nonempty item set such that  $T \subseteq I$ . Each transaction is related to an identifier, called a  $TID$ . Let  $A$  be a set of items. A transaction  $T$  is said to contain  $A$  if  $A \subseteq T$ .

An association rule is an implication of the form  $A \rightarrow B$ , where  $A \subseteq I$ ,  $B \subseteq I$ ,  $A \neq \emptyset$ ,  $B \neq \emptyset$ , and  $A \cap B \neq \emptyset$ . The rule  $A \rightarrow B$  holds in the transaction set  $D$  with **support**  $s$ , where  $s$  is the percentage of transactions in  $D$  that contain  $A \cup B$  (i.e., the *union* of sets  $A$  and  $B$  say, or, both  $A$  and  $B$ ). This is considered to be the probability,  $P(A \cup B)$ . The rule  $A \rightarrow B$  has **confidence**  $c$  in the transaction set  $D$ , where  $c$  is the percentage of transactions in  $D$  containing  $A$  that also contain  $B$ . This is calculated to be the conditional probability,  $P(B|A)$  “ [1] . Equations 1,2, and 3 summarize this process.

$$\text{support}(A \rightarrow B) = P(A \cup B) \tag{1}$$

$$\text{confidence}(A \rightarrow B) = P(B|A) \tag{2}$$

$$\text{confidence}(A \rightarrow B) = P(B|A) = \frac{\text{support}(A \cup B)}{\text{support}(A)} \tag{3}$$

Therefore, it uses support and confidence as the criteria for defining associations and generating rules by analysing data based on frequently occurs and if/then patterns. It works by analysing the item sets that contain two or more items. The association rules summarize the rules that well-stated in the dataset. Strong rules are those which meet both a minimum support threshold (min sup) and a minimum confidence threshold (min conf) [1].

The goal of the association rules algorithm in this study is to extract potential relationships among the fire accidents properties. It serves as complete analysis of huge number of incidents to find the actual large set of association. In this study, it is used to find the relationships between the fire accidents elements such as place, time, and reason of the fire accident.

## **4 Results and Discussion**

The results of the data exploration phase for the dataset of the fire accidents in Basra city for the last ten years, showed that the most reason causing the fire accidents is the electrical impact (59% of the accidents), followed by children messing around (17% of the accidents) as the second most reason, and intended action (5% of the accidents) besides cigarettes (5% of the accidents) as the third most reason causing fire accidents in Basra city.

In addition, the results also showed that most of the accidents happened in the city centre of Basra (36% of the accidents), followed by Abu-Alkhasseb county as second (16% of the accidents), Zubair county as third (12% of the accidents), Qurna county as fourth (7% of the accidents), and Medinah county as fifth place (4% of the accidents) where most of the accidents happened. Furthermore, the accidents pattern observed that most of the accidents happens during Summer and for various reasons. Also, the damage level was discovered to be 46% of the accidents with low severity damage, 31% with medium level of severity, 13% with high severity of damage, and 10% with no damage or losses.

Also, the exploration helped to identify some interrelationships within the fire accidents database. These relationships resulted in some important findings. For instance, the accidents rate increased during Summer 2018 more than all Summer seasons for the past decade. The reason for this unusual increment found to be because of the protesting that took place during Summer 2018. Figures 3,4, and 5 show sample of the results obtained during this phase.

By applying the association rules analysis with support value equal to 2% and confidence value equal to 10% applied to the fire accidents datasets, the results contained 30 rules. The result involved relationships between the location where the accidents occurred, the season on which the accidents occurred, and the reason caused the fire accident. In addition, the support value increased to 3% and the confidence value to 20% which resulted into 16 rules. Furthermore, the value of confidence has increased to 50% in order to generate more general and clear rules and resulted into 12 rules.

The general rule that observes through the interrelationship between the three elements of an accidents mentioned earlier is that most of the accidents happened due to

electrical impact and the highest number of them occurred during summer season. Also, it showed the three places in the city where the most accidents happened. Figures 6,7, and 8 show the results obtained during the association rules analysis.

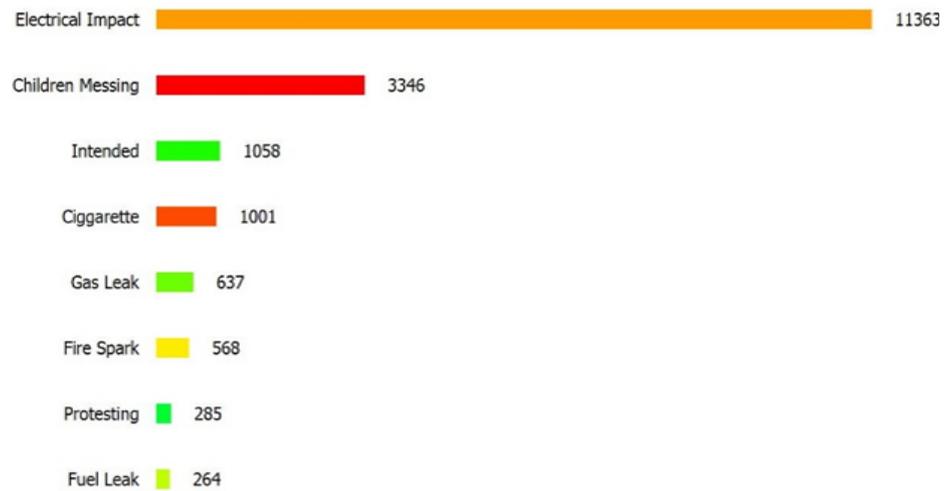


Fig. 3. Distribution of fire accidents causes

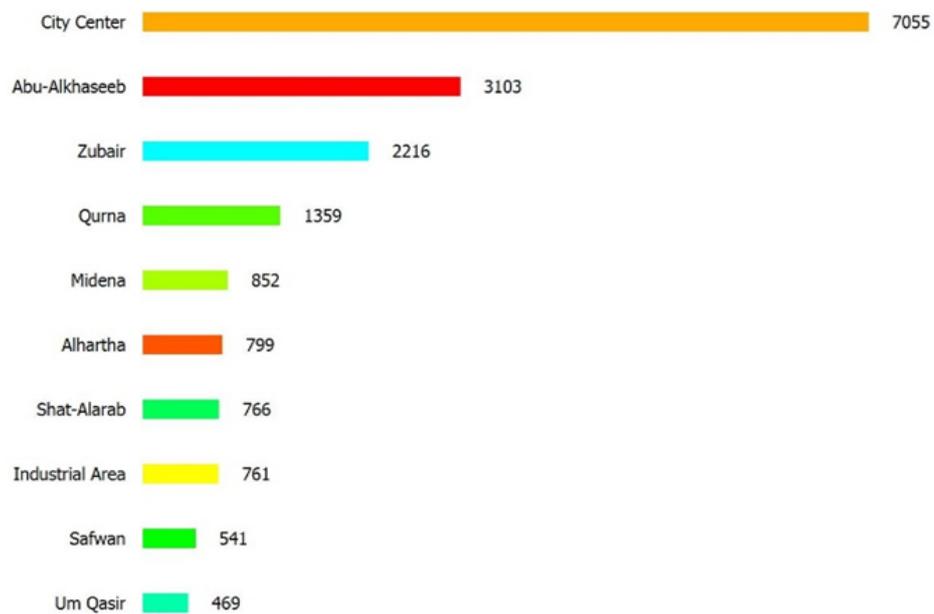


Fig. 4. Distribution of fire accidents locations



Fig. 5. Interrelationships of the data in the dataset

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent	Consequent
0.230	0.390	0.591	0.621	1.063	0.014	Cause=Electrical Impact	→ Location=City Center
0.230	0.628	0.367	1.611	1.063	0.014	Location=City Center	→ Cause=Electrical Impact
0.107	0.181	0.591	0.273	1.122	0.012	Cause=Electrical Impact	→ Location=Abu-Alkhaseeb
0.107	0.663	0.161	3.662	1.122	0.012	Location=Abu-Alkhaseeb	→ Cause=Electrical Impact
0.075	0.126	0.591	0.195	1.098	0.007	Cause=Electrical Impact	→ Location=Zubair
0.075	0.648	0.115	5.128	1.098	0.007	Location=Zubair	→ Cause=Electrical Impact
0.053	0.145	0.367	0.474	0.833	-0.011	Location=City Center	→ Cause=Children Messing
0.053	0.305	0.174	2.108	0.833	-0.011	Cause=Children Messing	→ Location=City Center
0.040	0.608	0.065	9.025	1.030	0.001	Date=Summer 2019	→ Cause=Electrical Impact
0.038	0.453	0.083	7.129	0.767	-0.011	Date=Summer 2018	→ Cause=Electrical Impact
0.037	0.587	0.062	9.461	0.994	-0.000	Date=Summer 2013	→ Cause=Electrical Impact
0.035	0.664	0.052	11.262	1.124	0.004	Date=Summer 2010	→ Cause=Electrical Impact
0.035	0.418	0.083	4.426	1.139	0.004	Date=Summer 2018	→ Location=City Center
0.033	0.570	0.058	10.155	0.965	-0.001	Date=Summer 2017	→ Cause=Electrical Impact
0.033	0.620	0.053	11.064	1.050	0.002	Date=Summer 2015	→ Cause=Electrical Impact
0.032	0.627	0.051	11.666	1.062	0.002	Date=Summer 2012	→ Cause=Electrical Impact
0.031	0.683	0.045	13.167	1.155	0.004	Date=Summer 2016	→ Cause=Electrical Impact
0.031	0.656	0.047	12.682	1.111	0.003	Date=Summer 2011	→ Cause=Electrical Impact
0.030	0.516	0.059	10.083	0.873	-0.004	Date=Summer 2014	→ Cause=Electrical Impact
0.029	0.415	0.071	8.361	0.703	-0.012	Location=Qurna	→ Cause=Electrical Impact
0.027	0.691	0.040	14.834	1.169	0.004	Location=Shat-Alarab	→ Cause=Electrical Impact
0.027	0.419	0.065	5.604	1.141	0.003	Date=Summer 2019	→ Location=City Center

Fig. 6. Association rules with 2% support and 10% confidence

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent	Consequent
0.230	0.390	0.591	0.621	1.063	0.014	Cause=Electrical Impact	→ Location=City Center
0.230	0.628	0.367	1.611	1.063	0.014	Location=City Center	→ Cause=Electrical Impact
0.107	0.663	0.161	3.662	1.122	0.012	Location=Abu-Alkhaseeb	→ Cause=Electrical Impact
0.075	0.648	0.115	5.128	1.098	0.007	Location=Zubair	→ Cause=Electrical Impact
0.053	0.305	0.174	2.108	0.833	-0.011	Cause=Children Messing	→ Location=City Center
0.040	0.608	0.065	9.025	1.030	0.001	Date=Summer 2019	→ Cause=Electrical Impact
0.038	0.453	0.083	7.129	0.767	-0.011	Date=Summer 2018	→ Cause=Electrical Impact
0.037	0.587	0.062	9.461	0.994	-0.000	Date=Summer 2013	→ Cause=Electrical Impact
0.035	0.664	0.052	11.262	1.124	0.004	Date=Summer 2010	→ Cause=Electrical Impact
0.035	0.418	0.083	4.426	1.139	0.004	Date=Summer 2018	→ Location=City Center
0.033	0.570	0.058	10.155	0.965	-0.001	Date=Summer 2017	→ Cause=Electrical Impact
0.033	0.620	0.053	11.064	1.050	0.002	Date=Summer 2015	→ Cause=Electrical Impact
0.032	0.627	0.051	11.666	1.062	0.002	Date=Summer 2012	→ Cause=Electrical Impact
0.031	0.683	0.045	13.167	1.155	0.004	Date=Summer 2016	→ Cause=Electrical Impact
0.031	0.656	0.047	12.682	1.111	0.003	Date=Summer 2011	→ Cause=Electrical Impact
0.030	0.516	0.059	10.083	0.873	-0.004	Date=Summer 2014	→ Cause=Electrical Impact

Fig. 7. Association rules with 3% support and 20% confidence

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent	Consequent
0.230	0.628	0.367	1.611	1.063	0.014	Location=City Center	→ Cause=Electrical Impact
0.107	0.663	0.161	3.662	1.122	0.012	Location=Abu-Alkhaseeb	→ Cause=Electrical Impact
0.075	0.648	0.115	5.128	1.098	0.007	Location=Zubair	→ Cause=Electrical Impact
0.040	0.608	0.065	9.025	1.030	0.001	Date=Summer 2019	→ Cause=Electrical Impact
0.037	0.587	0.062	9.461	0.994	-0.000	Date=Summer 2013	→ Cause=Electrical Impact
0.035	0.664	0.052	11.262	1.124	0.004	Date=Summer 2010	→ Cause=Electrical Impact
0.033	0.570	0.058	10.155	0.965	-0.001	Date=Summer 2017	→ Cause=Electrical Impact
0.033	0.620	0.053	11.064	1.050	0.002	Date=Summer 2015	→ Cause=Electrical Impact
0.032	0.627	0.051	11.666	1.062	0.002	Date=Summer 2012	→ Cause=Electrical Impact
0.031	0.683	0.045	13.167	1.155	0.004	Date=Summer 2016	→ Cause=Electrical Impact
0.031	0.656	0.047	12.682	1.111	0.003	Date=Summer 2011	→ Cause=Electrical Impact
0.030	0.516	0.059	10.083	0.873	-0.004	Date=Summer 2014	→ Cause=Electrical Impact

Fig. 8. Association rules with 3% support and 50% confidence

## 5 Conclusion and Future Work

To conclude this work, it is important to highlight that this paper will open a door for new research involving the study of the fire accidents in Basra city. The results of analysing a decade of data of the fire accidents helped to discover the hidden

knowledge that is very necessary for the related parties. This study helped to find out the most reason causing the fire accidents in Basra city is the electrical impact, which can lead to an inference that most of the electrical power connections and constructions do not adhere to specific installation standards.

In addition, from the association rules analysis, most of the accidents occurred due to electrical impact did happen during summer season, which can draw an inference that the high temperature led to huge load on the nonstandard electrical power connection and installation causing the electrical power impact. Moreover, this research can help the department of civil defense in the city to plan and distribute the required resource in an effective and efficient method.

This study also helped to show the places where the most accidents happened where the city centre as the first place, Abu Alkhsseb as second, and Zuabir as third. Moreover, it showed the damage levels and their distributions through various locations and causes. Almost half of the accidents caused low levels of damage. The seasons on which the accidents happened more or less also were discovered during this research. Summer was the most season during which accidents happened more for all the ten years followed by Winter as the second season.

Finally, this study is the first study involving the analysis of the fire accidents in Basra city and can be followed by other studies and researches in the future.

The future work can include and not limited to studying and analysing the fire accidents in each county individually in order to find the root causes and the resolution that help to reduce the high rate of the fire accidents. Another future work can involve analysing the pattern of the specific place (Household, Workshop, Vehicle, Explosion) where the fire accidents happened. Moreover, a predictive model can be built for the fire accidents in Basra city which is currently work that is being carried out by our team. And lots of other area can utilize the fire accidents dataset to discover knowledge and provide resolutions

## 6 References

- [1] J. Han, J. Pei, and M. Kamber (2011), Data mining: concepts and techniques. Elsevier.
- [2] Acuña E. (2011) Preprocessing in Data Mining. In: Lovric M. (eds) International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-04898-2\\_51](https://doi.org/10.1007/978-3-642-04898-2_51).
- [3] D. Zhang and K. Jiang (2012), Application of Data Mining Techniques in the Analysis of Fire Incidents. *Procedia Eng.*, vol. 43, pp. 250–256. doi: <https://doi.org/10.1016/j.proeng.2012.08.043>.
- [4] M. A. Abuzaraida and A. F. Elramalli (2017), Identifying the suitable reduction technique for mining medical data. in ICIT 2017 - 8th International Conference on Information Technology, Proceedings. <https://doi.org/10.1109/icitech.2017.8080029>.
- [5] K. Solanki, P. Berwal, and S. Dalal, (2016). Analysis of application of data mining techniques in Healthcare, *Int. J. Comput Appl.*, vol. 148, no.2,

[https://www.researchgate.net/publication/306125443\\_Analysis\\_of\\_Application\\_of\\_Data\\_Mining\\_Techniques\\_in\\_Healthcare](https://www.researchgate.net/publication/306125443_Analysis_of_Application_of_Data_Mining_Techniques_in_Healthcare) <https://doi.org/10.5120/ijca2016911011>

- [6] Demšar, J. & Curk, Tomaž & Erjavec, Ales & Gorup, Crt & Hocevar, Tomaz & Milutinovic, Mitar & Možina, Martin & Polajnar, Matija & Toplak, Marko & Staric, Anze & Stajdohar, Miha & Umek, Lan & Žagar, L. & Žbontar, J. & Žitnik, M. & Zupan, Blaz (2013), “Orange: data mining toolbox in Python,” *The Journal of Machine Learning Research.*, vol. 14, no. 1, pp. 2349–2353. [https://www.researchgate.net/publication/285018297\\_Orange\\_Data\\_Mining\\_Toolbox\\_in\\_Python](https://www.researchgate.net/publication/285018297_Orange_Data_Mining_Toolbox_in_Python) [https://doi.org/10.1007/978-3-540-30116-5\\_58](https://doi.org/10.1007/978-3-540-30116-5_58)
- [7] Sebt, M. V., Komijani, E., & Ghasemi, S. S. (2020). Implementing a Data Mining Solution Approach to Identify the Valuable Customers for Facilitating Electronic Banking. *International Journal of Interactive Mobile Technologies (IJIM)*, 14(15), 157. <https://doi.org/10.3991/ijim.v14i15.16127>
- [8] Suhirman, & Mohamad Zain, Jasni & Chiroma, Haruna & Herawan, Tutut. (2014). Data Mining for Education Decision Support: A Review. *International Journal of Emerging Technologies in Learning*. 9. 10.3991/ijet.v9i6.3950, <https://online-journals.org/index.php/i-jet/article/view/3950>

## 7 Authors

**Ibrahim Nasir Mahmood** is a lecturer in computer science field working at Basra University for Oil and Gas, Basra, Iraq. He received his BSc degree in Software Engineering on 2008. He holds Master’s degree in Computer Science from University of Bridgeport, CT, USA. He received his degree with Academic Achievement Award and Honors, He is a member of the UPE International Computing Society. He works at the Department of Chemical and Oil Refinery Engineering in Faculty of Oil and Gas Engineering, Basra University for Oil and Gas. His research interests include data mining, data science, machine learning, IoT, E-learning, and Algorithms.

**Dr. Hussein Ali Khuodhur Aliedane** is a senior lecturer at the Department of Oil and Gas Engineering in Faculty of Oil and Gas Engineering, Basra University for Oil and Gas, Basra Iraq. He holds PhD degree in Computer Science. He currently works as vice dean for scientific affair. His research interests include Data Mining, GIS, Data Science, E-Learning, Digital Image Processing and Algorithms.

**Dr. Mustafa Ali Abuzaraida** is international senior lecturer. He obtained his master degree in (Intelligent System) from University Utara Malaysia (UUM) in 2006 and a Ph.D. in (Computer Science) from International Islamic University Malaysia (IIUM) in 2015. His research interest is on Image processing, Data science, Data mining, Artificial intelligence application, and Natural Language Processing (NLP) mainly on text normalization and sentiment analysis and recognizing online handwritten text. Currently, he is working at School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, Sintok, Kedah, Malaysia as international senior lecturer.

Article submitted 2021-03-16. Resubmitted 2021-04-19. Final acceptance 2021-04-19. Final version published as submitted by the authors.