Peer-reviewed version available at Int. J. Environ. Res. Public Health 2018, 15, 462; doi:10.3390/ijerph15030462

Analysis of occupational accidents in underground and surface mining in Spain using data mining techniques

Lluís Sanmiquel[1*], Marc Bascompta[1], Josep Mª Rossell[2], Hernán Francisco Anticoi [3], Eduard Guash[3]

[1]ICL Chair in Sustainable Mining. lluis.sanmiquel@upc.edu. marc.bascompta@upc.edu

[2]Department of Mathematics. josep.maria.rossell@upc.edu.

[3]Department of Mining Engineering, Industrial and ICT. hernan.anticoi@upc.edu. eduard.guasch@upc.edu.

*Correspondence: lluis.sanmiquel@upc.edu; 34 938777238.

Abstract: An analysis of workplace accidents in the mining sector has been done using the database from the Spanish administration between the period 2005-2015 and applying data mining techniques. Data has been processed by means of the software Weka. Two scenarios were chosen regarding the accidents database, surface and underground mining. The most important variables involved in occupation accidents and their association rules have been determined. These rules are formed by several predictor variables that cause an accident, defining its characteristics and context. This study exposes the 20 most important association rules of the sector, either surface or underground mining, based on statistical confidence levels of each rule obtained by Weka. The outcomes display the most typical immediate causes with the percentage of accident basis of each association rule. The most typical immediate cause is body movement with physical effort or overexertion and type of accident is physical effort or overexertion. On the other hand, the second most important immediate cause and type of accident change in both scenarios. Data mining techniques have been proved as a very powerful tool to find out the root of the accidents, apply corrective measures and verify their effectiveness, either for public or private companies.

Keywords: Data mining, Association rules, Previous Cause, Type of Accident, Overexertion.

---

1. Introduction

The mining industry is an important economic sector for many countries, including Spain, and includes coal, metal and non-metal minerals. Its use is extensive worldwide and it comprises the production of necessary elements all sectors, such us construction, energy, agriculture, medical or electronic. According to the mining statistics of Ministry of Industry, Energy and Digital Agenda from Spain, in 2015, there were 2896 surface and underground mines, having 4540 people directly employed in underground mines, 12702 in quarries and 6755 in mineral processing plants.

Unfortunately, many studies have pointed out the mining sector as one of the most dangerous activities due to its intrinsic characteristics: environmental conditions with a significant degree of humidity and dust, danger of injury from falling rocks or falls on the same and different level, etc. [1-7]. All these factors have influence on the incidence and severity of accidents, both higher than in other economic sectors. Moreover, the incidence index, number of accidents per 100000 workers, was 4.3 times higher in the Spanish mining sector than the overall value of the country in 2015, verifying the danger of mining compared to other economic sectors. Besides, the same index also indicates that underground mining is much more hazardous than surface mining, 4.8 times higher than in surface mining, between 2005 and 2015. When the comparison is done with the mining sector from other countries, the Spanish ratio is 5.5 times higher than in the United States or 16.1 than in Australia according to the data from the corresponding public agencies. However, the Spanish incidence index has been significantly reduced, 2.1 times between 2005 and 2015 [8]. Despite this remarkable

Peer-reviewed version available at Int. J. Environ. Res. Public Health 2018, 15, 462; doi:10.3390/ijerph15030462

evolution, there is still room for improvement and reduce the direct and indirect costs generated by injuries and deaths [9]. Many studies related to high-risk industries have proved that safety culture and organizational performance are key factors to reduce workplace accidents rates [10-14].

The accident analysis in Spanish surface mining between 2005 and 2015 shows that 19.5% of the accidents were caused due to physical overexertion on the musculoskeletal system, with an immediate cause because of body movement, being the most common. Moreover, some studies also indicate a direct relationship between overexertion injuries and age of the employee [15].

The knowledge of the main occupational hazards, their causes and factors is necessary to improve the safety conditions of the employees. Therefore, it will require a tool able to handle large amounts of information, known as data mining, with the idea to study the genesis of occupational accidents, together with the extraction of rules or behavioral patterns of injuries. Weka software is a collection of state-of-the-art machine learning methods and data preprocessing tools composed by all major learning techniques for classification and regression: Bayesian classifiers, decision trees, rule sets, linear regression and nearest-neighbor methods [16-18]. Weka 3.9 software has been used in this case.

Methods based on Bayesian networks are more sensitive in detecting associations among categorical variables than other statistics methods. Hence, this methodology can be very useful to obtain reliable conclusions for the decision-making process with regard to safety issues. Bayesian networks have also been applied in many other scientific fields such as civil engineering [19], geological engineering [20], ecology [21], medicine [22], road traffic safety [23-25], environmental assessment impact [26, 27] business risk and product life-cycle analysis [28]; workplace tasks [29], workplace risk area [30], interrelation between hygienic workplace conditions and occupational accidents [31] or construction and mining accidents [32,33].

The objective of this paper is to extract information from the annual digital database of the Spanish mining accidents between the period 2005-2015. The accidents are classified in function if the place where took the accident is an underground mine or an open-pit mine. The main factors influencing in the typology of occupational accident have been identified in each class, with the idea of improving preventive policies, minimizing risks, injuries and deaths in the mining sector.

2. Materials and Methods

2.1 Study population

A total of 56034 mining accidents are classified in two scenarios: Scenario I, the accident takes place in an underground mine; Scenario II, the accident occurs in a surface mine. 28894 and 27140 instances are collected in Scenario I and Scenario II respectively.

2.2 Variables

Although 54 variables were registered in every accident, only 11 of them have been chosen due to their relevance in the study. The variable *Type of Accident* has been selected as target attribute, while the remaining variables are considered as predictors.

The first step is the detection of possible errors in the database. Numerical and graphic statistical methods are used to verify if they correspond with the different groups where variables are classified. If so, the instances are corrected or removed from the original dataset. The 11 selected variables are the following:

- Age (A): Age of the injured worker, distributed in seven classes: [16-24], [25-29], [30-34], [35-39], [40-44], [45-54], [more than 55].
- Contract (C): Four different types of employment contract. 1=permanent and full-time; 2=permanent and part-time; 3=temporary and full-time; 4=temporary and part-time.
- Day Hour (DH): Hour of the day when the accident happened. Five classes are considered: (0-6], (6-10], (10-14], (14-18], (18-24].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2018 doi:10.20944/preprints201801.0231.v1

Peer-reviewed version available at *Int. J. Environ. Res. Public Health* **2018**, *15*, 462; doi:10.3390/ijerph15030462

- Experience (E): Months of experience of the injured employee, seven groups: [0-12], [13-30], [31-60], [61-120], [121-180], [181-240], [more than 241].
- Physical Activity (PA): Physical activity done by the worker at the time of the accident. Seven categories are considered: 1=machine operations; 2=working with hand tools; 3=driving or being in a conveyance; 4=manipulation of objects; 5=manual handling of loads; 6=performing a movement; 7=others.
- Place (P): Place where the injury occurred. Five groups are considered: 1=treatment plants, workshops and storages; 2=general constructions or demolitions; 3=surface mine; 4=underground mine; 5=other places. This variable is only considered in surface mining.
- Preventive Organization (PO): Preventive organization in the company, organized in six categories: 1=the employer; 2=designated workers; 3=own prevention service; 4=joint prevention service; 5=external prevention service; 6=without prevention service.
- Previous Cause (PC): Previous cause before the accident, grouped in seven categories: 1=electric problem, explosion, fire, overflow, overturn, leak, spill, vaporization or emanation; 2=fracture, slip, fall or collapse; 3=loss of control of the working machinery, total or partial; 4=falls/tumbles of a person; 5=body movement without physical effort; 6=body movement with physical effort or overexertion; 7=other causes.
- Size (S): Number of employees of the mining activity. Groups are: [0-9], [10-19], [20-49], [50-99], [100-499], [more than 500].
- Type of Accident (TA): It explains the cause of the accident. Seven categories are considered: 1=electric contact, fire, contact with hazardous substances, drowning; 2=impact or collision with stationary object; 3=hit or collision with a moving object; 4=contact with a sharp or pointed object; 5=being trapped, crushed or suffering an amputation; 6= physical effort or overexertion; 7=others.
- Work Hour (WH): How many hours the employee worked before the accident. Six groups are established: [0-1], [2-4], [5-8], [9-10], [11-12], [more than 13].

Parameters created will have a terminology such as A1, which means the age of the injured (A) is between 16 and 24 (1). Figure 1 shows the absolute frequency distribution for the variable *Type of Accident*, used as target attribute.

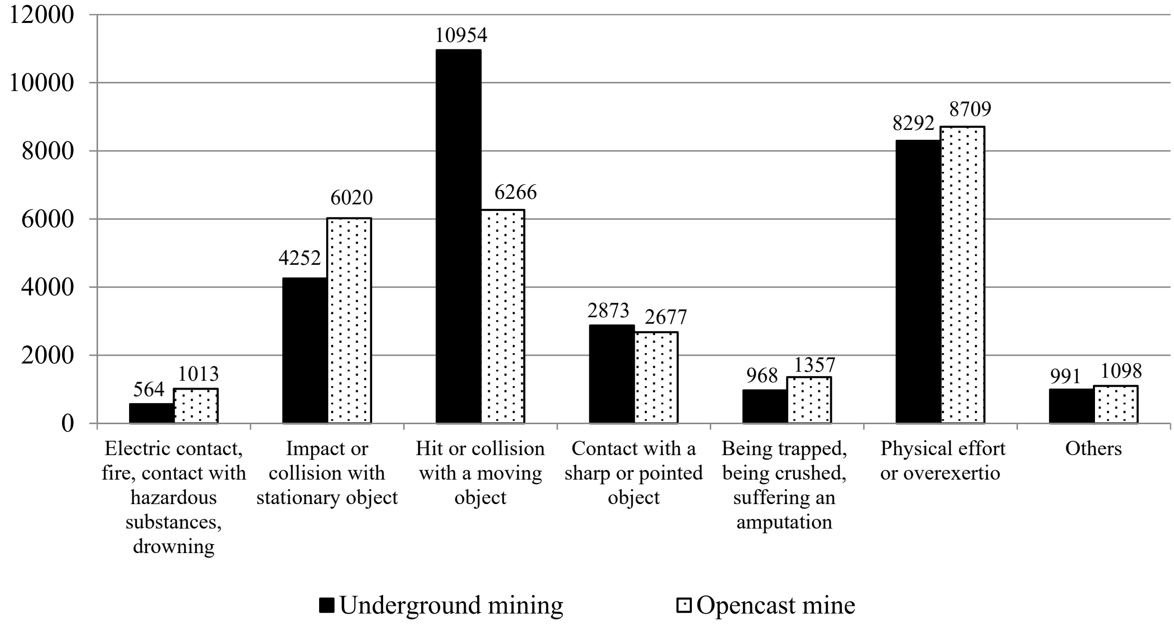

Figure 1. Frequency distribution for the variable Type of Accident.

2.3 Significant predictors

Classification tools are used to perform further analysis on pre-processed data, obtaining a model produced on the full trained data. A ranking of the most significant predictors for the target variable *Type of Accident* is established by means of attribute evaluators and data mining tools. *Select Attributes* of Weka provides methods to identify the variables that are predictive for *Type of Accident*. Some attribute evaluators combined with search methods have been used, Table 1. Two different selection modes called *full training set* and *cross-validation* have been applied, choosing n=10 folds for the *cross-validation* method. In this case, dataset is randomly reordered and split into 10 folds of approximately equal size. A fold is used to test, while the others train the classifier.

Table 1. Attribute evaluators, search methods and selection modes.

Target attribute	Attribute evaluators	Search methods	Selection modes
Type of Accident	ChiSquaredAttributeEval	Ranker	Full training set and 10 cross-validation
	CfsSubsetEval	GreedyStepwise	
		ExhaustiveSearch	
		BestFirst	
	ClassifierSubsetEval	RandomSearch	
	InfoGainAttributeEval	Ranker	

Position of variables in each learning scheme is calculated to determine the final ranking for the best predictors that have influenced in *Type of Accident*. Tables 2 and 3 show the ranking of the predictor variables for Scenarios I and II, respectively. All the accidents in Scenario I belong to the same place, P=4 (underground mine.

Table 2. Ranking of the variables for Scenario I.

Variables	PC	S	PA	E	PO	C	A	WH	DH
Ranking	1	2	3	4	5	6	7	8	9

Table 3. Ranking of the variables for Scenario II.

Variables	PC	PA	P	E	S	A	PO	DH	C	WH
Ranking	1	2	3	4	5	6	7	8	9	10

2.4 Decision trees and association rules

A methodology to produce decision trees is needed. Decision tree learning is a method frequently used in data mining to create a model that predicts the value of a target variable based on several input variables. There are many specific decision-tree algorithms, but one of the most common is the C4.5, a method of classification that generates a decision tree using the training data [34]. In this paper, an improved algorithm from C4.5, the J48 implemented in Weka, has been used for *Type of Accident* as target attribute. Once the J48 algorithm is applied, the *confusion matrix*, commonly named *contingency table*, shows how many instances have been correctly or incorrectly assigned to each class. 77.4% of cases have been properly assigned for Scenario I and 70.6 % for Scenario II, which can be considered as acceptable.

Association rules are obtained to identify relationships between attributes in data, working with the *Predictive Apriori* algorithm, which iteratively reduces the minimum support until it finds the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2018 doi:10.20944/preprints201801.0231.v1

Peer-reviewed version available at *Int. J. Environ. Res. Public Health* **2018**, *15*, 462; doi:10.3390/ijerph15030462

required number of rules with the minimum confidence [35]. The *confidence* or *predictive accuracy* indicates the number of instances for which all the conditions are true, *coverage*, divided by the number of instances for which the conditions in the antecedent are true. The best rules for *Type of Accident* allow extracting information about the main causes of the injuries for Scenarios I and II.

3. Results and discussion

130 and 110 association rules have been obtained for Scenario I and Scenario II, respectively. Tables 4 and 5 expose the 20 best association rules for both scenarios, sorted by their confidence level with 3 or more predictor variables and having the *Type of Accident* (TA) as target variable. Another ranking has also been created, including the predictors variables ordered by the times one variable appears in the corresponding rule, Tables 6 and 7. The number of associations has been chosen with the idea to have an important number of them, but always within an acceptable confidence range.

Table 4. The 20 best association rules for Scenario I.

Predictor Variable 1	Predictor Variable 2	Predictor Variable 3	Predictor Variable 4	Target Variable	Confidence	% Accidents
C1	PC6	PA4	PO3	TA6	0.868	6.21
PC6	PA4	PO3	S5	TA6	0.867	6.30
C1	PO3	PA4	S5	TA6	0.863	7.10
E4	C1	PC6		TA6	0.842	5.05
E4	C1	PC6		TA6	0.828	5.60
C1	PC6	WH2	E4	TA6	0.816	8.37
C1	PC6	PO3		TA6	0.814	13.40
PC6	WH2	PO3		TA6	0.814	7.32
C1	PC6	WH2	PO3	TA6	0.814	7.20
S5	PC2	PA2		TA3	0.812	5.88
C1	PC2	PA4		TA3	0.811	6.52
S5	C1	PC2	PA1	TA3	0.804	5.34
PC2	PA2	WH2		TA3	0.803	5.73
PC2	PA4	PO3		TA3	0.799	5.73
C1	PC6	WH3		TA6	0.799	5.28
C1	PC2	PA4	PO3	TA3	0.799	5.59
S5	PC2	WH3		TA3	0.798	5.13
S5	C1	PC6	PO3	TA6	0.798	8.56
S5	PC6	PO3		TA6	0.797	8.72
S5	C1	PC6		TA6	0.796	9.59

Table 5. The 20 best association rules for Scenario II.

Predictor Variable 1	Predictor Variable 2	Predictor Variable 3	Predictor Variable 4	Target Variable	Confidence	% Accidents
C1	PC6	WH2		TA6	0.811	7.11
C1	PC6	WH2	PO5	TA6	0.809	5.73
PC6	PA4	P1	E1	TA6	0.807	7.65
PC6	PA4	PO5		TA6	0.806	6.38
S3	PC6	P1	WH2	TA6	0.796	5.68
PC6	WH2	PO5		TA6	0.795	9.23
PC6	DH2	PO5		TA6	0.794	6.04
C1	PC6	PO5	E1	TA6	0.792	10.75
PC6	P3	PO5		TA6	0.792	6.35
C1	PC6	P1		TA6	0.786	7.30

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2018 doi:10.20944/preprints201801.0231.v1

Peer-reviewed version available at Int. J. Environ. Res. Public Health 2018, 15, 462; doi:10.3390/ijerph15030462

C1	PC6	P1	PO5	TA6	0.785	5.25
PC6	DH3	PO5		TA6	0.783	5.56
PC6	P1	PO5		TA6	0.771	8.38
S3	PC6	PO5		TA6	0.771	5.17
PC4	PA6	P1	PO5	TA2	0.766	6.25
S3	E1	PC6	PO5	TA6	0.753	5.34
C1	PC4	PO5		TA2	0.750	5.68
S3	PC6	PO5		TA6	0.745	6.40
PC6	WH3	PO5		TA6	0.743	5.73
C1	PA6	PO5		TA2	0.464	6.44

Table 6. Number of times a variable appears in an association rule, Scenario I.

Variables	**PC**	**C**	**PO**	**WH**	**S**	**PA**	**DH**	**A**	**E**
Num. of repetitions	75	57	47	42	39	33	17	15	7
Ranking	1	2	3	4	5	6	7	8	9

Table 7. Number of times a variable appears in an association rule, Scenario II.

Variables	**PO**	**C**	**PC**	**WH**	**P**	**DH**	**PA**	**E**	**S**	**A**
Num. of repetitions	64	42	37	32	25	20	16	13	10	4
Ranking	1	2	3	4	5	6	7	8	9	10

There are some important differences between the ranking from Weka, Tables 2 and 3, and the number of times a variable appears in association rules, Tables 6 and 7; variables involved in more than 45% of all accidents are classified in lower positions by Weka. For instance, C, PO and WH, from Scenario I, have low positions according to Table 2, but they are at the top of Table 6. Besides, C1, PO3 and WH2 are 82.7%, 74.9% and 51.5% of 28894 accidents, respectively. Something similar happens in Scenario II, where the same variables are in low positions in Table 3 and in high positions in Table 7 with an 83.2%, 62.2% and 48.4% of 27140 accidents, respectively. According to the data processed, the mining sector characteristics explain some of these initial outcomes.

a) A high percentage of work contracts in the sector are full-time and permanent, C1.
b) The main type of *Preventive Organization* varies if it is underground or surface mining, due to the average of employees. Almost all surface activities have less than 50 workers, while the majority of underground mining activities have more than 99 employees. This fact leads to internal prevention services in underground mines, PO3, and external prevention services in surface mining, PO5.
c) Variable *Work Hour* (WH) has few influence in the *Type of Accident*. It appears in 6th and 5th position in the top 20 for Scenarios I and II, respectively. Besides, it has been found that the majority of accidents occur between the first 2-4 working hours in 8 out of 11 rules in both scenarios, WH2. This fact could suggest there is a lack of attention due to a meal break, which is usual between the second and third hour of working, decreasing the concentration level when work is restarted.

3.1 Scenario I

Previous Cause (PC), *Size* (S) and *Physical Activity* (PA) variables are in the top of Table 2. Hence, these predictors are the most influential variables in the *Type of Accident*.

Peer-reviewed version available at Int. J. Environ. Res. Public Health 2018, 15, 462; doi:10.3390/ijerph15030462

PC is in 19 out of 20 association rules and the type PC6, body movement with physical effort or overexertion, is in 12 out of 19 rules, generating a TA6 accident, physical effort or overexertion. Meanwhile PC2, fracture, slip, fall or collapse, is in the other 7 rules, with a TA3 accident, hit or collision with a moving object. According to these rules, the accident with a prediction variable PC6 and an exit of TA6 is given in 91.6%; whereas 39.9% have the prediction variable PC2 and an exit TA3. Results suggest that preventive policies in underground mines should be focused on reducing accidents due to physical effort or overexertion, perhaps mechanising some actions as well as increasing load handling training programs. In addition, falling of objects from the ceiling and walls is also important, being necessary to train workers about workplace inspections and control procedures. These statements are in accordance with [36], who studied 320 underground mines and concluded that the three most influencing factors of coal mine accidents are: 1) lack of safety education and training, 2) rules and regulations of safety production responsibility and 3) rules and regulations of supervision and inspection.

Size (S) variable is the second most important according to Table 2. It appears in the best 8 out of 10 association rules with a class S5 in all cases, mining activities between 100-499 employees, being a predictor variable in 57.6% of the accidents. Besides, Table 8 gathers the following information from the database: Percentage of accidents, workers, size of the company and incidence index. The last variable index is indicative of the incidence of accidents among different groups or subpopulations [37], and it is defined as the ratio of percentage of injured workers of a given subpopulation to the percentage of the total workforce represented by this subpopulation.

Table 8. Official data from the Spanish administration in the period 2005-2015.

Variable	<= 49 Workers	50-99 Workers	100-499 Workers	>= 500 Workers
Accidents	12.2%	9.4%	50.5%	27.9%
Workers	4.2%	5.3%	25.8%	64.7%
Incidence Index	2.9	1.8	2.0	0.4

The incidence index is much higher in underground activities with less than 50 employees than in the other groups, the ratio is reduced as the company gets bigger. This fact matches with other studies, either in the mining sector [38],or other industrial activities [39]. Besides, the difference between companies with a workforce of less and more than 500 employees is quite relevant, having a much lower incidence index.

According to Weka, the *Physical Activity* (PA) developed at the moment of the accident is classified as the third most influent variable in Scenario I, appearing in the best 9 out of 20 association rules, Table 4. The predominant activities are: manipulation of objects, PA4, and working with hand tools, PA2. On the other hand, *Experience* (E) is the 4th in the ranking, outstanding the class E4, workers between 61 and 120 months of experience, which is in discordance with previous studies [40,8], where they point out the highest accident incidence in cases with less than 30 months of experience. Further research should be done in this regard. In reference to the *Age* (A) variable, it does not have a significant influence to the *Type of Accidents*, being in the 7th and 8th positions in Tables 2 and 6, respectively. The three best association rules are detailed in the following paragraphs:

1. Employee with a full-time permanent contract in a company with an internal prevention service. The worker suffers and accident while doing a physical activity manipulating objects, with an immediate cause based on body movement with physical effort or overexertion.
2. Employee from a company with an internal prevention service and a size between 100 and 499 workers. He suffers and accident while doing a physical activity manipulating objects, with an immediate cause based on body movement with physical effort or overexertion.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2018 doi:10.20944/preprints201801.0231.v1

Peer-reviewed version available at *Int. J. Environ. Res. Public Health* **2018**, *15*, 462; doi:10.3390/ijerph15030462

3. Worker with a full-time permanent contract in a company with an internal prevention service and a size between 100 and 499 workers. He suffers and accident while doing a physical activity manipulating objects, with an immediate cause based on body movement with physical effort or overexertion.

3.2 Scenario II

Predictor variables PC, PA, E and S are among the most important ones, like in Scenario I, as well as *Place* (P) variable, which is specific for surface mining and it is the third with more influence. On the other hand, variables A, PO, DH, C and WH are at the bottom of the ranking, but A and DH have more relevance than in Scenario I.

PC variable appears in 18 out of 20 association rules and it is a class PC6 in 16 out of these 18, body movement with physical effort or overexertion, with a *Type of Accident* TA6, whereas it is PC4, falls/tumbles of a person, in the other two, with TA2, impact or collision with stationary object, generating 114.4% and 11.9% of the accidents, respectively. Preventive measures similar to Scenario I should be taken for the *Type of Accident* PC6-TA6 and PC4-TA2 due to its important consequences.

Place (P) is the third most influent variable and it appears in 7 out of 20 association rules, Table 5, and 6 of these rules are P1, treatment plants, workshops and storages, which includes 40.5% of all accidents. These results are in accordance with previous studies [14]. Regarding *Experience* (E) variable, it is the 4th with more influence to accidents generation. It appears in 3 out of 20 and in 23.7% of all accidents as association rule with a class E1, less than 12 months of experience, which is in accordance with the current literature [40,8]. However, it is in discrepancy with results from Scenario I. The three best association rules are detailed in the following paragraphs:

1. Worker with a full-time permanent contract suffers an accident between the first 2-4 working hours, with an immediate cause of body movement with physical effort or overexertion.
2. Worker with a full-time permanent contract in a company with an external prevention service. He/she suffers an accident between the first 2-4 working hours and an immediate cause of body movement with physical effort or overexertion.
3. Employee with less than 1 year of experience working in a treatment plant, workshop or storage suffers an accident due to object manipulation, with an immediate cause of body movement with physical effort or overexertion.

4. Conclusions

The use of data mining techniques has been proven as a powerful tool to determine the main factors and rules of occupation accidents in the mining sector from a large database. The usage of Weka could help public and private companies to find out the root of the accidents, apply corrective measures and verify their effectiveness in the time. Regarding both scenarios, some common patterns and differences have been found:

a) Among the five more influent variables, four are coincident: PC, S, PA and E.
b) The most typical accident has an immediate cause of body movement with physical effort or overexertion.
c) The second most important type of accident in Scenario I has an immediate cause of fracture, slip, fall or collapse (PC2), causing a type of accident hit or collision with a moving object (TA3). This type of accident includes ceiling, wall or object collapse. On the other hand, Scenario II displays a second most important type of accident with an immediate cause of falls/tumbles of a person (PC4) and a type of accident impact or collision with stationary object (TA2).
d) The age of the injured workers has little influence to the Type of Accident.

Peer-reviewed version available at Int. J. Environ. Res. Public Health 2018, 15, 462; doi:10.3390/ijerph15030462

e) The size of the company has more influence in the origin of the accident in Scenario I than in Scenario II.

Acknowledgments

We would like to thank the Spanish Ministry of Employment and Social Safety for allowing us the use of the Spanish annual general database of accidents in the mining sector between 2005 and 2015.

Author Contributions: The author Lluís Sanmiquel wrote the main manuscript text and prepared the figures. Marc Bascompta carried out the data collection. Josep Ma Rosell processed the data in Weka software and extracted the best association rules. Hernan Anticoi and Eduard Guash helped the manuscript writing and organization as well as the review of the English text.

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

References

1. Saxen, H.; Adam, J.L.; Duhamel-Henry, N.; Jacquier, B.; Linares, C.; Hull, B.P.; Leigh, J.; Driscoll, T.R.; Mandryk, J. Factors associated with occupational injury severity in the New South Wales underground coal mining industry. Safety Science 1996, 21(3), 191-204.
2. Mitchell, R.J.; Driscoll; T.R.; Harrison, J.E. Traumatic work-related fatalities involving mining in Australia. Safety Science 1998, 29 (2), 107-123.
3. Gyekye, S.A. Causal attributions of Ghanaian industrial workers for accident occurrence: Miners and non-miners perspective. Journal of Safety Research 2003, 34(5), 533-538.
4. Martín, J.E.; Rivas, T.; Matías, J.M.; Taboada, J.; Argüelles, A. A Bayesian network analysis of workplace accidents caused by falls from a height. Safety Science 2009,47(2), 206-214.
5. Sanmiquel, L.; Freijo, M.; Edo, J.; Rossell, J.M. Analysis of work related accidents in the Spanish mining sector from 1982-2006. Journal of Safety Research 2010, 41, 1-7.
6. Yarahmadi, R.; Bagherpour, R.; Khademian, A. Safety risk assessment of Iran's dimension stone quarries (Exploited by diamond wire cutting method). Safety Science 2014, 63, 146-150.
7. Bagherpour, R.; Yarahmadi, R.; Khademian, A. Safety risk assessment of Iran's underground coal mines based on preventive and preparative measures. Human and Ecological Risk Assessment 2015, 21 (8), 2223-2238.
8. Sanmiquel, L.; Rossell, J.M. ; Vintró, C. Study of Spanish mining accidents using data mining techniques. Safety Science 2015, 75, 49-55.
9. Hämäläinen, P.; Takala, J.; Leena, K. Global estimates of occupational accidents. Safety Science 2006, 44, 137-156.
10. Mallick, S.; Mukherjee, K. An empirical study for mines safety management through analysis on potential for accident reduction. International Journal of Management Science 1996, 24(5), 539-550.
11. Torres, N.; Dinis, C. Environmental, health and safety management systems for underground mining. 1st. International Conference on Sustainable Development and Management of the Subsurface. Utrecht, Netherlands, 2003.
12. Zhong-Xue, L.; Jia-Jie, L.; Cui-Ping, L.; Shuang-Yue, L. Overview of the South African mine health and safety standardization and regulation systems. Journal of Coal Science and Engineering 2008, 14 (2), 329-333.
13. Bottani, E.; Monica, L.; Vignali, G. Safety management systems: Performance differences between adopters and non-adopters. Safety Science 2009, 47(2), 155-162.
14. Sanmiquel, L.; Rossell, J.M.; Vintró, C. ; Freijo, M. Influence of occupational safety management on the incidence rate of occupational accidents in the Spanish industrial and ornamental stone mining. WORK 2014, 49, 307-314.
15. Margolis, K.A. Underground coal mining injury: A look at how age and experience related to days lost from work following an injury. Safety Science, 2010, 48(4), 417-421.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 January 2018 doi:10.20944/preprints201801.0231.v1

Peer-reviewed version available at Int. J. Environ. Res. Public Health 2018, 15, 462; doi:10.3390/ijerph15030462

16. Bishop, C.M. Pattern Recognition and Machine Learning (Information Science and Statistics), Springer, 2006.
17. Witten, I.H.; Frank, E.; Hall, M.A. Data Mining: Practical Machine Learning Tools and Techniques (3rd Ed.), Elsevier, USA, 2011.
18. Bouckaert, R.R.; Frank, E.; Hall, M.; Kirkby, R.; Reutemann, P.; Seewald, A.; Scuse, D. Weka Manual for version 3-7-10. Nieznane czasopismo, 2013.
19. Gerassis, S.; Martín, J.E.; García, J.T.; Saavedra, A; Taboada, J. Bayesian decision tool for the analysis of occupational accidents in the construction of embankments. Journal of Construction Engineering and Management 2017, 143(2), 04016093.
20. Rivas, T.; Matías, J.M.; Taboada, J.; Argüelles, A. Application of Bayesian networks to the evaluation of roofing slate quality. Engineering Geology 2007, 94 (1-2), 27-37.
21. Adriaenssens, V. ; Goethals, P.L.M. ; De Pau, C.N. Application of Bayesian belief networks for the prediction of macroinvertebrate taxa in rivers. Annales de Limnologie – International Journal of Limnology 2004, 40(3), 181-191.
22. Antal, P.; Fannes, G.; Timmerman, D.; Moreau,Y.; De Moor, B. Using literature and data to learn Bayesian networks as clinical models of ovarian tumors. Artificial Intelligence in Medicine 2004, 30(3), 257-281.
23. Chang, L.Y.; Chen, W.C. Data mining of tree-based models to analyze freeway accident frequency. Journal of Safety Research 2005, 36, 365-375.
24. Flask, T.; Schneider, W. A Bayesian analysis of multi-level spatial correlation in single vehicle motorcycle crashes in Ohio. Safety Science 2013, 53, 1-10.
25. Tavakoly, A.; Rabieyan, R.; Besharati, M. A data mining approach to investigate the factors influencing the crash severity of motorcycle pillion passengers. Journal of Safety Research 2014, 51, 93-98.
26. Baran, E.; Jantunen, T. Stakeholder consultation for Bayesian decision support systems in environmental management. In: Proceedings of the Regional Conference on Ecological and Environmental Modeling (ECOMOD 2004), 15-16. University Sains Malaysia, Penang, Malaysia, 2004.
27. Matías, J.M.; Rivas, T.; Ordóñez, C.; Taboada, J. Assessing the environmental impact of slate quarrying using Bayesian networks and GIS. In: Proceedings of the Fifth International Conference on Engineering Computational Technology. Las Palmas de Gran Canaria, 2006, 345-346.
28. Zhu, J.Y.; Deshmukh, A. Application of Bayesian decision networks to life cycle engineering in green design and manufacturing. Engineering Applications of Artificial Intelligence 2003, 16, 91-103.
29. Ghasemi, F.; Kalatpour, O.; Moghimbeigi, A.; Mohammadfam, I. Selecting strategies to reduce high-risk unsafe work behaviors using the safety behavior sampling technique and Bayesian network analysis. Journal of Research in Health Sciences 2017, 17(1), 00372.
30. Galán, S.F.; Mosleh, A.; Izquierdo, J.M. Incorporating organizational factors into probabilistic safety assessment of nuclear power plants through canonical probabilistic models. Reliability Engineering and System Safety 2007, 92, 1131-1138.
31. García-Herrero, S.; Mariscal, M.A.; García-Rodríguez, J.; Ritzel, D.O. Working conditions, psychological/physical symptoms and occupational accidents. Bayesian network models. Safety Science 2012, 50, 1760-1774.
32. Matías, J.M.; Rivas, T.; Martín, J.E.; Taboada, J. A machine learning methodology for the analysis of workplace accidents. International Journal of Computer Mathematics 2008, 85(3), 559-578.
33. Rivas, T.; Paz, M.; Martín, J.E.; Matías, J.M.; García, J.F.; Taboada, J. Explaining and predicting workplace accidents using data-mining techniques. Reliability Engineering and System Safety 2011, 96, 739-747.
34. Quinlan, J.R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA, 1993.

35. Scheffer,T. Finding association rules that trade support optimally against confidence. In L. de Raedt, & A. Siebes (Eds.), Proceedings of the Fifth European Conference on Principles of Data Mining and Knowledge Discovery, 424-435, Freiburg, Germany, Berlin: Springer-Verlag, 2001.
36. Zhang, Y.; Shao, W.; Zhang, M.; Li, H.; Yin, S.; Xu, Y. Analysis 320 coal mine accidents using structural equation modeling with unsafe conditions of the rules and regulations as exogenous variables. Accident Analysis and Prevention 2016, 92, 189-201.
37. Butani, S.J. Relative risk analysis of injuries in coal mining by age and experience at present company. Journal of Occupational Accidents 1988, 10(3), 209-216.
38. Hunting, K.L.; Weeks, J. L. Transport injuries in small coal mines: An exploratory analysis. American Journal of Industrial Medicine 1993, 23, 391-406.
39. Saari, J. Pequeñas y medianas empresas. Jornada técnica: La prevención de riesgos laborales en la PYME (In Spanish). Sevilla. Asepeyo, 2005, 11-20.
40. Kecojevic, V.; Komljenovic, D.; Groves, W.; Radomsky, M. An analysis of equipment-related fatal accidents in U.S. mining operations: 1995–2005. Safety Science 2007,45, 864-874.