Application of Bayesian network model in explaining occupational accidents in a steel industry

Conference Paper · November 2017

DOI: 10.1109/ICRCICN.2017.8234531

4 authors:

Sobhan Sarkar
Indian Institute of Technology Kharagpur

Anindra Kumar

Sunil Kumar Mohanpuria
Indian Institute of Technology Kharagpur

Jhareswar Maiti
Indian Institute of Technology Kharagpur

Application of Bayesian Network Model in Explaining Occupational Accidents in a Steel Industry

Sobhan Sarkar
Research Scholar
Department of Industrial & Systems Engineering
IIT Kharagpur
sobhan.sarkar@gmail.com

Anindra Kumar
B.Tech.
Department of Agricultural & Food Engineering
IIT Kharagpur
anindra.726@gmail.com

Sunil Kumar Mohanpuria
B.Tech.
Department of Ocean Engineering & Naval Architecture
IIT Kharagpur
sunilmohanpuria@yahoo.com

J Maiti
Professor
Department of Industrial & Systems Engineering
IIT Kharagpur
jhareswar.maiti@gmail.com

***Abstract*— In the occupational accident analysis, identification of the interrelationships of the factors behind the accidents is very important. To explore the relationships or the impacts of the causal factors on the accidents and to predict the incident outcomes i.e., injury, near miss, and property damage cases, Bayesian Network (BN) model is used in this paper. The proposed model is validated using the data retrieved from an integrated steel manufacturing industry in India using sensitivity analysis. BN performs well in terms of prediction with 88.28% accuracy using 10-fold cross validation. In addition, some important key findings are obtained from the analysis like the factors slip-trip-falls, crane dashing, and the months February and July are found to be the sensitive factors towards incident outcomes in the plant. The proposed model, therefore, has a good potential for explaining accident prediction and causation in manufacturing industry and can be applied in different domains also.**

Keywords—Occupational accidents, Steel industry, Bayesian Network, Sensitivity analysis.

I. Introduction

In today's world, advancements in industrialization have led to the establishment of industries and workplaces. In this highly competitive atmosphere, every industry has encountered a number of problems, one of that is the industrial accidents. Various circumstances and events prompt the occurrence of occupational accidents in the industry. From the prior study, it has been realized that most of these accidents are due to the poor safety standards, non-compliance/availability of the standard operating procedures (SOP), health of workers, poor machine conditions, and so forth. The advancements both in technology and automations in industries have resulted in different types of accidents. Thus, it is required to address the loss of lives caused by these occupational accidents. Various studies have been conducted to investigate the nature of occupational accidents to suggest the ways of preventing them. With proper analysis of past data, these accidents can be controlled if they can be well predicted.

A. Occupational accident statistics

As per the statistics from National Programme for Control and Treatment of Occupational Diseases in 2009, 100 million occupational injuries are reported in India resulting in 0.1 million deaths. In addition, it is also reported that 45,000 fatal injuries (45% of the total deaths due to injuries in the world) and 17 million non-fatal injuries (17% of the world) at work take place in each year [1]. There are various factors present behind an accident to take place. In India, a total of 1433, 1383, and 1417 occupational fatal accidents took place in 2011, 2012, and 2013, respectively [2]. Among these accidents, Indian steel plants share a total of 29, 52, and 31 fatal accidents respectively [2]. Thus, there is an urgent need for the management of steel plant to take initiatives to curve down the number of accidents by improved safety practices at work. Usually, there is a large number of factors responsible for an accident. So, a good prediction model is necessary that can capture the synergistic effect of all these factors and predict the occurrence of accidents/incidents so that preventive measures could be initiated well in advance.

B. Tools and techniques used in analysis

There is an array of different tools and techniques available for explaining of occuptional accidents like dicision tree (DT) [3], SVM [4], rule-based approach [5], Bayesian Network (BN), or other statistical approaches [6], [7], [8]. BN among them, has been found to be very effective. The BN has the ability to provide us graphical inter-relationships. It provides us a very useful way to deal with compound problems because of its ability to combine robust probability methods. It uses the Baye's theorem to generate conditional probabilities. Updating and revision of prior beliefs are the important characteristics in BN model with the supply of new evidence (data). Joint-strategies is one of the main advantage of BN model by which the posterior probability of a variable can be determined by integrating two or more hypotheses. By using Bayesian statistical methods and BN model, over-fitting of data can be avoided. The BN model has been very helpful for predictions. It allows for prediction of uncertainties and risks better than the models that account only for expected values. In the analysis of occupational accidents, BN is widely used for explaining the effect of factors towards incident/ accident outcomes. Due to its capability of prediction, it has been used in diverse areas, such as biotechnology [9], medicine [10], studies of web site usability [11], lifecycle engineering [12], reliability [13], healthcare systems, customer satisfaction surveys [14],

978-1-5386-1931-5/17/$31.00©2017IEEE

geology [15] etc. In addition, it has been successfully applied in safety studies [16], [17], [18], [19]. For example, the work by Ren et al.[16] contributed in offshore safety estimation by proposing a technique to model causal relationship with BN. It is capable of providing calculated numerical values for the occurring of each failure event and also gives graphical inter-relationships. Zhou et al.[17] presented a BN model to construct a probabilistic relational network model among causal factors, personal experience factors and safety climate factors, which exert impact on human safety behavior. Some have applied BN model on road accidents; Helai et al. [18] proposed a Bayesian hierarchical binomial logistic model to know the important variables influencing the vehicle damage and seriousness of driver injury at signalized intersections. De Oa et al. [20] used the BN model to categorize accidents on road by seriousness of the injuries. Martin et al. [21] used BN model to study the workplace accidents happened because of auxiliary equipment (anchore, ladders, scaffolding, and so forth) which can result into accidents. This helps them to identify which situation have the maximum impact on work-place accidents during these activities, like duration of tasks, acquiring of worng working postures, and workers insufficient information of safety regulations. McCabe et al. [22] used BN model to show that the high levels of interpersonal conflict, greater pressure at work, and poor-quality leadership were mainly responsible for work-related accidents. Kim et al. [23] used BN model for safety culture in high-reliability industries such as nuclear power plants and aviation. Gerassis et al. [24] used BN model to quantify the specific causes of different types of accidents in the construction of areas. Mohammadfam et al. [25] aimed to study managing and improving safety behavior of employees by using BN model.

C. Research gap and contribution of the paper

Based on the review of literature (though not mentioned in-depth in this paper), the utility of BN in different domain including occupational accident has been realised. However, the exploration of the interrelationship of factors and their impact on accident in steel industry has rarely been addressed by any privious researchers. Moreover, collecting any kind of accident data from industry is itself a difficult as well as challenging task. Therefore, the present work addresses this issue and attempts to bridge the gap between theory and practices by implementing the recommendations, and findings obtained from BN analysis at industrial level. The aim of the paper is set to find the impact of different factors on accident/incident outcome using BN model. However, the study is limited to steel industry only.

The remainder of the paper is structured as follows: **Section II** discusses the methods used in this study through the flowchart of the proposed research methodology. In **Section III**, a case study from a steel plant is provided. Results from analysis and the discussion have been presented in **Section IV** and finally, conclusions with future scopes are discussed in **Section V**.

II. METHODOLOGY

In this process, after the collection of data, data have been pre-processed. After the pre-processing step, data have been split into 75% and 25% as training and testing, respectively.

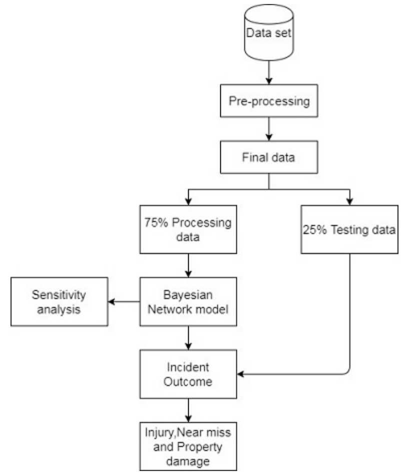

Fig. 1: Proposed research methodological flow chart.

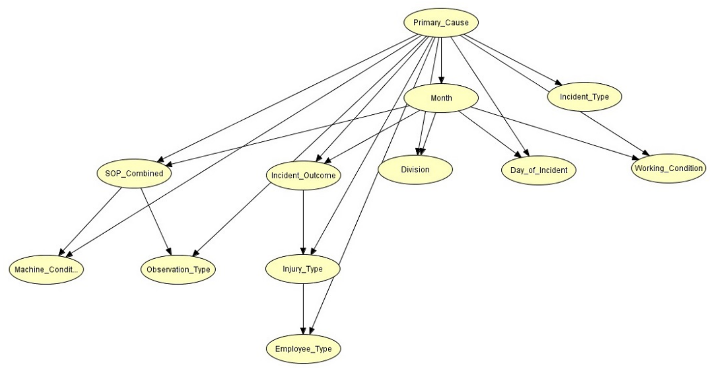

Fig. 2: Structure of BN.

On training set of data, BN model has been built and its performance have been tested through accuracy and sensitivity on testing data. In addition, sensitivity analysis are carried out on different attributes of the data set to predict incident outcome (refer to Fig.1). A short note on BN theory has been provided as working principle of our study.

A. Bayesian Network theory

BN model is the graphical model tool that represents joint probability distribution of a set of random variables with a possible mutual causal relationships. Let's assume a set of random variables say $U = X_1, X_2, X_3, ..., X_n$ where $n \geqslant 0$. BN model for their random variables will be a directed acyclic graph and a set of probability tables. A BN model represents joint probability as Eq.(1) given below.

$$P(X_1, X_2, ..., X_n) = \prod_{i=1}^{n} P(X_i | Pa\{X_i\}) \quad (1)$$

In Eq.(1), $Pa(X_i)$ represents parent set of X_i in BN model, where $i = 1, 2, 3, ..., n$, and n is the number of random variable.

III. Case study

A. Problem statement

In this work, we attempted to fill the research gap by collecting data from a steel plant in India. The plant (under study) has been experiencing accidents at its workplace. The safety model, precautionary measurement, and safety database, despite being present, have not been nevertheless found to be efficient enough to mollify the event of accidents. Thus, the administration of the plant has been looking for some efficacious solutions having much potential to reduce the incidents at workplace as well as to identify the inter-relationship among the factors attributable to the occurrence of accidents.

B. Data set & data description

The data from 2010 to 2013 with 2052 observations have been retrieved from the electronic database of the safety management system of the steel plant in India which are briefly discussed below.

1) *Day of incident (DOI)*- This attribute implies the day on which the incident was occurred There are seven categories in it, all days of a week.
2) *Month of incident (MOI)*- This attribute implies the month in which the incident was occurred There are twelve classes in it, all months of a year.
3) *Division*- This represents the division where the incident was taken place In total, ten divisions were considered, namely Div1, Div2, Div3, Div4, Div6, Div7, Div9, Div10, Div11, and Div12.
4) *Incident outcome*- It is the outcome variable. It has three different classes; (i) injury (I) - when someone gets injured physically from an incident, (ii) near miss (NM) - when someone is narrowly escaped from an incident having full potential to cause injury or damage, and (iii) property damage (PD) - when there is damage to private or public property due to the incident.
5) *Primary Cause*- This attribute refers to the top event that qualifies the incident which has occure. It has 23 classes. They are gas leakage (GL), crane dashing (CD), slip trip fall (STF), electric flash (EF), dashing collision (DC),working at heights (WAH), energy isolation (EI), toxic chemicals (TC), fire explosion (FE), equipment machinery (EM), rail (R), derailment (D), skidding (S), hydraulic pneumatic (HP), process incidents (PI), hot metals (HM), structural integrity (SI), road incidents (RI), run over (RO) and lifting tools and tackles (LTT).
6) *Type of injury (IT)*- This attribute represents the type of injury. It has 12 classes They are claim injury on duty (IOD), claim injury at work (IOW), death, exgratia, fatal, first aid, foreign body, IOW, Injury type not applicable (ITNA), normal, restricted work cases (RWI), serious injury. Note that for NM and property damage cases, type of injury is ITNA.
7) *Working condition (WC)*- This attribute represents the condition of work when the incident took place. It has three categories, group working (GW) representing the condition where people work in groups, single working (SW) representing person working alone, and Others (NAPP) representing situations when no workers were present.
8) *Machine Condition (MC)*- It implies the condition of machine when the accident took place; either machine is in idle condition (MI) or in running condition (MR), or others (Napp) i.e., not related to machine.
9) Observation type (OT) - This attribute denotes the basic causes of incident and has four categories as; (i) unsafe act (UA) representing the person himself is responsible for the cause of incident, (ii) unsafe act and unsafe condition (UAUC) representing the incident occurred due to presence of both the factors, persons fault and hazardous condition, (iii) unsafe act by other (UAO) representing the incident occurred due to the others fault, and (iv) unsafe condition (UC) representing a situation which is likely to cause incidents.
10) *Employee Type (ET)*- This attribute has two classes, namely Employee and Contractor.
11) *Incident type (IT)*- This attribute represents whether an accident happened is due to human behaviour (Beh), or process (Pro) which is non-human fault.
12) *SOP Combined*- It is called the standard operating procedure(SOP). It has six classes.

C. Data pre-processing

Before performing analysis, the data set was pre-processed by removing inconsistencies, missing values, duplicate entries etc. This process consumes nearly 60% of the total modeling time. With the help of this step, data set has been reduced to 1062 data points from 2052 data points. Once it is cleaned or pre-processed, it is ready for building the BN model, which is followed by results and discussion mentioned in the next section.

IV. Results and discussion

To analyse the effect of certain attributes on incident outcomes, BN Model has been used. To build the BN model, we have used preprocessed training data set so as to recognize the relationships existing among attributes considered in the study. This allows for the complicated relationship to be explored visually that may be present in our data set. Fig. 2 shows global BN structure for the training data set, in which primary cause is the root node and is connected to all other eleven nodes. Other nodes are partially interconnected among each other.

For the structure learning of BN model, there are many software packages available. Analytica, Bayes Net Toolbox, Hugin and Netica are reorganizable software that allow users to build, analysis and visualize data efficiently with ease. Friedman et al. [26] compared Tree Augmented Naive Bayes(TAN) with C4.5, Naive Bayes and selective Naive Bayes. Experiments revial that TAN holds significance improvement over the others. There are various algorithms available for structure learning. Here, we have used Netica BN software that uses TAN algorithm. TAN algorithm approximates the interaction between variables by using an imposed tree structure on the Naive-based structure.

Fig. 3 depicts intial belief bars for different nodes of BN, which consist of initial probability values of every state in each

TABLE I: Confusion matrix obtained from BN modeling.

Predicted			Actual
I	NM	PD	
61	2	3	I
0	120	3	NM
0	22	45	PD

TABLE II: Results: Precision, F1 Score and Recall.

	Precision	Recall	F1 Score
I	92.424	100	96.062
NM	97.561	83.33	89.885
PD	67.164	88.235	76.27

node. These probability values of different nodes used to carry out sensitivity analysis by changing variables state probability to 100%. A noteworthy results can be obtained graphically using BN structure.

A. Validation

To calculate the error rate, BN model was tested on test data set. The data set was randomly partitioned from initial data set using k-fold partitioning (here, k=2) that generates training and test data set in the ratio of 75% (n=796) and 25% (n=266), respectively. The error rate is 15.04%, that defines the BN model has that most of its predictions correct for incident outcome. Confusion matrix shown in Table I decribes performence of BN model in terms of accuracy being 88.28%. Thereafter, with the help of confusion matrix we obtained Table II, which contains *Precision*, *Recall* and *F1-Score* (in %).

In addition, Table III shows that how much the network is confident in its beliefs but it is actually wrong. It contains columns with certain conditions like <1%, <10% , >%90 and >99% and these represents confidence percentage every time it makes a decision. It contains ratios implying that the number of times network was wrong over the number of times it makes a confident prediction.

B. Sensitivity analysis

1) *For primary cause*
The results given in Table IV show primary cause that dashing (D) exerts a great influence on incident outcome. Taking the extreme value 71.8%, the probability of property damage is increased by 45.5% from initial value of 26.4%. As a result, injury value goes down to 9.04% from initial value of 26.5%. crane dashing (CD) also shows similar effect on NM cases that decreases by 28%, which were at 47.1% initially.

Sensitivity analysis of primary cause (refer to Table IV) also indicates that slip/trip/fall (STF) cases increase the probability of NM up to 65.7% although the probability of property damage case goes down to 7.29%. Among all primary cause states, road incidents have highest probability rate of 42.8%. Here, from the sensitivity analysis of BN, the factors like CD, STF and road accident are found to be sensitive significantly.

2) *For employee type*
The sensitivity analysis of employee type given in Table V reveals that the node employee type has direct impact on incident outcome. Given the condition for employee type's state *Contractor*, probability of NM increases up to 56.6% which is 9.5% more from initial probability value and for state injury, probability rate gets reduced by 12.2%. While state is employee, the probability of injury rises up to 58.4% which is more than twice of intial probability, well above the rest of variables. Here, Employee is the most sensitive parameter.

3) *For month*
In Table VI for month February, the probability rate of NM decreased by 24.1% and probability rate of

TABLE III: Predicted probability for incident outcomes.

	Predicted Probability			
State	<1%	<10%	>90%	>99%
I	0.00 (0/52)	0.00 (0/189)	0.00 (0/22)	0.00 (0/0)
NM	0.00 (0/11)	0.00 (0/39)	4.00 (2/50)	0.00 (0/0)
PD	0.00 (0/0)	2.02 (2/99)	0.00 (0/2)	0.00 (0/0)
Total	0.00 (0/63)	0.61 (2/327)	2.70 (2/74)	0.00 (0/0)

TABLE IV: Sensitivity analysis of primary cause.

	Incident Outcome		
Primary cause	I	NM	PD
initial	26.5	47.1	26.4
CD	9.04	19.1	71.8
D	15.3	34.3	50.3
DC	24.2	42.7	33.1
EF	19.1	46.5	34.4
EI	25.6	40.4	34
EMB	23	36.2	40.7
FE	15.1	51.7	33.2
GL	22.1	57.8	20.1
HM	28.1	48.5	23.4
HP	23.4	47.7	28.9
LTT	25.8	54.2	20
MA	39	37.1	23.9
MH	34.8	45.6	19.6
OI	26.5	44.3	29.3
PI	36.1	42.8	21.1
R	27	38	35
RI	42.8	42.4	14.7
RO	34	38.1	27.9
S	21.4	56	22.6
SI	23.8	49.3	26.9
STF	27	65.7	7.29
TC	32.1	35.9	32.1
WH	28.4	43.1	28.4

TABLE V: Sensitivity analysis of employee.

	Incident Outcome		
Employee Type	I	NM	PD
initial	26.5	47.1	26.4
Contractor	14.3	56.6	29
Employee	58.4	21.9	19.6

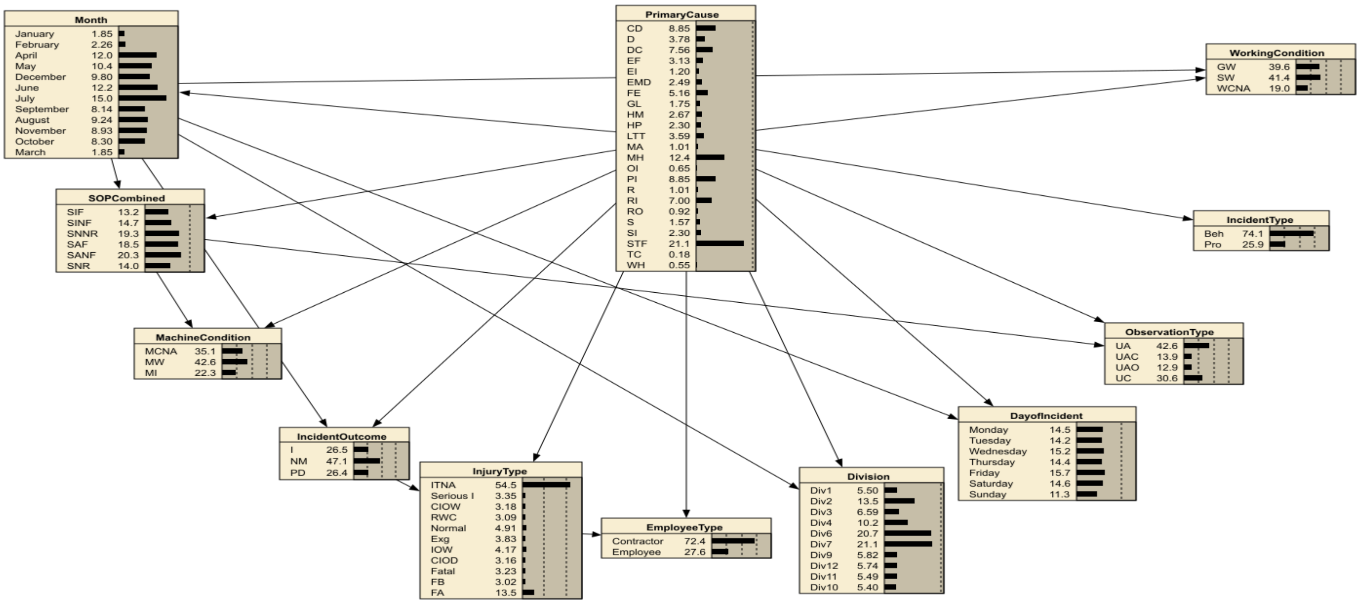

Fig. 3: Belief bars of BN structure.

TABLE VI: Sensitivity analysis of month.

	Incident Outcome		
Month	I	NM	PD
initial	26.5	47.1	26.4
January	27.7	27.7	44.5
February	23	23	54
March	27.7	27.7	44.5
April	20.8	55.5	23.8
May	17.8	60.1	22.1
June	21.7	59.1	19.2
July	17	67.7	15.3
August	22	49.9	28.1
September	32	35.1	32.9
October	38.9	32	29.1
November	41.8	28.3	29.9
December	39.1	26.4	34.5

TABLE VII: Sensitivity analysis of division.

	Incident Outcome		
Division	I	NM	PD
initial	26.5	47.1	26.4
Div1	26.8	43.3	29.9
Div2	26.4	47.7	25.8
Div3	28.9	43.1	28
Div4	30.3	41.1	28.6
Div6	24.5	48.9	26.6
Div7	26.3	55.2	18.6
Div9	25.5	41.4	33.3
Div10	26.8	43.1	30.1
Div11	26.6	42.5	30.9
Div12	25.8	41.6	32.6

TABLE VIII: Sensitivity analysis of machine condition.

	Incident Outcome		
Machine Condition	I	NM	PD
initial	26.5	47.1	26.4
MCNA	27.6	51.1	21.3
MW	24.9	43.3	31.8
MI	27.9	47.9	24.2

property damage increased by 27.6%. Bearing in mind the fact that initial probabilities of these were 47.1% and 26.4%, respectively, these changes are significant.In July, NM cases is increased by 20.6%. Probability rates of injury and property damage cases are decreased by 9.5% and 11.1%, respectively. Both of these are minimum values observed in all the states. Probability rate of injury cases is spiked by 15.3% in November in comparison with initial rate of incidents. Here, the months July and February are found to be the most sensitive towards incident outcomes.

4) *For division*
From the Table VII, it is observed that there is no significant influence in probabilities of incident outcome states except in Div7, which has a 7%-8% change in NM and PD rates, respectively. Other states in Table VII indicates only 2% to 3% change in incident outcome rates. Here, Div7 is the most sensitive parameter.

5) *For machine condition*
In TABLE VIII, there is no significant change in probability of incident outcome observed, except machine condition state machine working (MW) which influences the property damage by 5.4% from initial probability rate, whereas other states only vary by 2% to 3%. Here, MW is the most sensitive parameter.

V. Conclusions

Complex nature of occupational accidents compel us to perform in-depth analysis of different factors to understand how they effect incident outcomes to mitigate the occurrence of accidents at work. Knowing the casual relationship between predictor and response variables enables us to accomplish more effictive strategy against occupational accidents. Therefore, BN model was developed to establish relationship between predictor and response variables. Some of the key findings using BN model are listed below.

- The sensitivity analysis carried out for primary cause CD shows increase in rate of property damage by 45.5%, more than twice of initial probability which is the highest incident rate and decrease in injury and NM cases were observed by 17.46% and 28%, respectively. STF shows increse in NM cases by 18.6%.
- There is an increase in the rate of injury for employees by 31.9% which is more than twice of initial probability.
- Month of February is spiked in rate of property damage by 27.6%. It is twice of initial probability that is the highest incident rate. Decrease in the NM cases is observed as 24.1%. In July, NM cases are increased by 20.6%.
- Among divisions, division 7 (i.e., Div7) has slight impact on probability rate of both NM and property damage by 7% to 8%. Rest of the attributes have very slight impact of about 5% or less on incident outcomes probability rates.

Finally, the following strong direct relationships were observed to incident outcome; CD, employee, July and February. The results from this study indicate that the nodes Day of incident, working condition, SOP combined, observation type and incident type have very less influence on incident outcomes. Therefore, BN is very effective modeling technique for explaining occupational accidents. The future scopes of the present study may include the modeling the workers' response to proximity warnings for hazards, or development of BN-based decision support system for enhancing safety at the working place in industry.

References

[1] N. Wagner, P. Wagner, P. Jayachandran *et al.*, "Distance learning courses in occupational medicine-methods and good practice," *Indian Journal of Occupational and Environmental Medicine*, vol. 9, no. 2, p. 57, 2005.

[2] "www.indiastat.com/table/crimeandlaw/6/ industrialaccidents."

[3] S. Sarkar, A. Patel, S. Madaan, and J. Maiti, "Prediction of occupational accidents using decision tree approach," in *India Conference (INDICON), 2016 IEEE Annual*. IEEE, 2016, pp. 1–6.

[4] S. Sarkar, S. Vinay, V. Pateshwari, and J. Maiti, "Study of optimized svm for incident prediction of a steel plant in india," in *India Conference (INDICON), 2016 IEEE Annual*. IEEE, 2016, pp. 1–6.

[5] S. Sarkar, A. Lohani, and J. Maiti, "Genetic algorithm-based association rule mining approach towards rule generation of occupational accidents," in *International Conference on Computational Intelligence, Communications, and Business Analytics*. Springer, 2017, pp. 517–530.

[6] S. Gautam, J. Maiti, A. Syamsundar, and S. Sarkar, "Segmented point process models for work system safety analysis," *Safety science*, vol. 95, pp. 15–27, 2017.

[7] O. B. Krishna, J. Maiti, P. K. Ray, B. Samanta, S. Mandal, and S. Sarkar, "Measurement and modeling of job stress of electric overhead traveling crane operators," *Safety and health at work*, vol. 6, no. 4, pp. 279–288, 2015.

[8] K. Singh, N. Raj, S. Sahu, R. Behera, S. Sarkar, and J. Maiti, "Modelling safety of gantry crane operations using petri nets," *International journal of injury control and safety promotion*, vol. 24, no. 1, pp. 32–43, 2017.

[9] J. Peterson and R. Kenett, "Modeling opportunities for statisticians supporting quality by design efforts for pharmaceutical development and manufacturing," *Biopharmaceutical Report, American Statistical Association Publication, USA*, 2011.

[10] P. Antal, G. Fannes, D. Timmerman, Y. Moreau, and B. De Moor, "Using literature and data to learn bayesian networks as clinical models of ovarian tumors," *Artificial Intelligence in medicine*, vol. 30, no. 3, pp. 257–281, 2004.

[11] R. Kenett and S. Salini, "New frontiers in survey data analysis," *Quality Progress*, 2009.

[12] J. Zhu and A. Deshmukh, "Application of bayesian decision networks to life cycle engineering in green design and manufacturing," *Engineering Applications of Artificial Intelligence*, vol. 16, no. 2, pp. 91–103, 2003.

[13] H. Langseth and L. Portinale, "Bayesian networks in reliability," *Reliability Engineering & System Safety*, vol. 92, no. 1, pp. 92–108, 2007.

[14] R. S. Kenett and S. Salini, "Modern analysis of customer satisfaction surveys: comparison of models and integrated analysis," *Applied Stochastic Models in Business and Industry*, vol. 27, no. 5, pp. 465–475, 2011.

[15] T. Rivas, J. Matías, J. Taboada, and A. Argüelles, "Application of bayesian networks to the evaluation of roofing slate quality," *Engineering Geology*, vol. 94, no. 1, pp. 27–37, 2007.

[16] J. Ren, I. Jenkinson, J. Wang, D.-L. Xu, and J.-B. Yang, "A methodology to model causal relationships on offshore safety assessment focusing on human and organizational factors," *Journal of Safety Research*, vol. 39, no. 1, pp. 87–100, 2008.

[17] Q. Zhou, D. Fang, and X. Wang, "A method to identify strategies for the improvement of human safety behavior by considering safety climate and personal experience," *Safety Science*, vol. 46, no. 10, pp. 1406–1419, 2008.

[18] H. Huang, H. C. Chin, and M. M. Haque, "Severity of driver injury and vehicle damage in traffic crashes at intersections: a bayesian hierarchical analysis," *Accident Analysis & Prevention*, vol. 40, no. 1, pp. 45–54, 2008.

[19] S. Sarkar, S. Vinay, and J. Maiti, "Text mining based safety risk assessment and prediction of occupational accidents in a steel plant," in *Computational Techniques in Information and Communication Technologies (ICCTICT), 2016 International Conference on*. IEEE, 2016, pp. 439–444.

[20] J. de Oña, R. O. Mujalli, and F. J. Calvo, "Analysis of traffic accident injury severity on spanish rural highways using bayesian networks," *Accident Analysis & Prevention*, vol. 43, no. 1, pp. 402–411, 2011.

[21] J. E. Martin, T. Rivas, J. Matías, J. Taboada, and A. Argüelles, "A bayesian network analysis of workplace accidents caused by falls from a height," *Safety Science*, vol. 47, no. 2, pp. 206–214, 2009.

[22] B. McCabe, C. Loughlin, R. Munteanu, S. Tucker, and A. Lam, "Individual safety and health outcomes in the construction industry," *Canadian Journal of Civil Engineering*, vol. 35, no. 12, pp. 1455–1467, 2008.

[23] Y. G. Kim, S. M. Lee, and P. H. Seong, "A methodology for a quantitative assessment of safety culture in npps based on bayesian networks," *Annals of Nuclear Energy*, vol. 102, pp. 23–36, 2017.

[24] S. Gerassis, J. Martín, J. T. García, A. Saavedra, and J. Taboada, "Bayesian decision tool for the analysis of occupational accidents in the construction of embankments," *Journal of construction engineering and management*, vol. 143, no. 2, p. 04016093, 2016.

[25] I. Mohammadfam, F. Ghasemi, O. Kalatpour, and A. Moghimbeigi, "Constructing a bayesian network model for improving safety behavior of employees at workplaces," *Applied ergonomics*, vol. 58, pp. 35–47, 2017.

[26] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Machine learning*, vol. 29, no. 2, pp. 131–163, 1997.