Prediction of Occupational Accidents in Steel Industry using Bayesian Belief Network

[1]Shivam Singh, [2]Arunjeet Chakraborty, [3]Shatrudhan Pandey, [4]Abhishek Kumar Singh

Department of Production & Industrial Engineering, Birla Institute of Technology, Mesra, Ranchi, India
[1]btech10472.18@bitmesra.ac.in, [2]arunjeet22@gmail.com, [3]er.shatrudhanp@gmail.com, [4]abhishekkumar@bitmesra.ac.in

Abstract

Nowadays, emphasizing health and safety in the process industry, building a proper model for predicting occupational safety and health hazards is essential. In general, accidents in the steel industry are often overlooked because of their sheer numbers compared to other more accident-prone sectors such as Oil and Gas. However, recent studies have shown that the number of casualties in steel industries is increasing, and this issue needs to be taken care of urgently. Some of the factors that lead to occupational hazards in the steel industry are unfavorable working setup, machine conditions, lagging Standard Operating procedures (SOP), and physical, chemical and radiation issues leading to poor health and safety conditions for workers and various property damage. Given the data set retrieved from an integrated steel plant via Incident Data Reporting System (IDRS), multiple parameters can be obtained, which can help us identify whether a casualty will occur. Various ML algorithms can be applied to detect the probability of a loss arising from this data. Each of the ML algorithms has its advantages and disadvantages. Given the complexity of the dataset, Bayesian Belief Network proves to be the ideal fit in terms of accuracy.

Keywords - Process Safety, Steel Industry, Bayesian Belief Network

1. INTRODUCTION

India's steel industry, primarily iron and steel, is the country's largest ferrous metal sector and a key economic component. India is ranked second among all countries in terms of crude steel production. Steel consumption per capita has become an indicator of a country's economic growth as a result of the Indian economy's globalization and the expansion of industries that use steel as raw material, such as the automobile industry, railways, and defence. Steel usage is also rising in the country. The availability of domestic raw resources and low labor wages, compared to the rest of the globe, are the key advantages for Indian steel companies. Steelmaking has always been a risky, and accidents are unavoidable. However, due to enhanced safety system implementation, the world has changed once more. Steel manufacturing is complicated, which means workers are exposed to a variety of hazards, including extreme heat, high pressure, radiation, heights, heavy machinery, and hazardous pollutants, all of which put them at risk of accidents. In the process sector, it is critical to developing a system for recognizing occupational dangers.

It has also been seen that steel industry workers take long-term sick leave more often than in other industries, which more or less can be explained because of constant exposure to hot, humid, noisy, gaseous, and dusty environments, which negatively impacts the production in the steel industry thus affecting various interrelated output in sectors such as railways, housing construction, shipbuilding, etc. [1], [2].

In a recent study, it has been found that the total number of industrial accidents (major /minor) has risen to 30 crores per year, with more than 3 lakh reported deaths [3].

It must also be noted that advancements in technology and automation have resulted in different types of accidents; thus, a proper prediction of hazards taking all critical parameters must be built on an urgent basis.

To identify the parameters necessary for building our model on “Prediction of occupational hazards in steel industry using Bayesian Belief Network,” thorough research has been done. The Quality Management System Standard ISO 9001 has been used by the Steel Industry for many years, along with ISO 14001:2015 for Environmental Management and ISO 45001:2018 for Health and Safety Management. The Bureau of Indian Standards (BIS) has helped our cause identify the parameters necessary for safe operation in an everyday process industry [4].

A knowledge graph is then used to identify, separate, and establish a rough relationship between the key factors responsible for occupational accidents in the steel industry [5], [6]. The knowledge model needs to be known first to understand a knowledge graph. A knowledge model is a collection of interdependent descriptions of concepts, relationships, events, etc., that allows both people and machines to process efficiently—the reports contributing to each other form an extensive network, from which meaningful output can be obtained.

Now, to understand the interrelation between different factors and to predict the possible outcomes of the accident, “Bayesian Belief Network” is used, which is a graphical method used to predict uncertain events using the conditional probability theorem (CPT) [7]. Bayesian Belief Networks are used when the real-world applications are probabilistic and there is a probabilistic connection between various events/nodes. It is mainly used for anomaly detection, diagnostics, classification, and prediction under uncertainty. It is also commonly known as the Bayes network, belief network, decision network, and Bayesian model. A Directed Acyclic Graph (DAG) is used, which helps establish relationships between various factors through a network of nodes and ultimately leads to determining the Joint Probability of each node through the conditional probability theorem (CPT). The probabilistic dependence/independence between each variable is observed through a network of arcs. It is also noted that each node is conditionally independent of its non-descendants, given its immediate parents.

The Tree Augmented Naive Bayes algorithm (TAN algo) is used for structure learning of the model leading to the identification of various measures required for process safety in steel plants. In other words, it is used for finding the approximation of the dependencies between multiple factors.

CART (classification and regression trees) and CHAID (chi-square automated interaction detection) are two decision trees used to make predictions in data mining. These predictions may reveal traits of people who are vulnerable to workplace accidents. The CART and CHAID algorithms can forecast the outcome of workplace accidents in a steel plant. Furthermore, safety authorities can reduce the rate of accidents by using predictions for detecting vulnerable workers in steel plants. However, there are some limitations to these techniques for predicting accidents in steel plants as it has often been reported that the system becomes unstable when the structure gets complicated, and therefore it is recommended that the use of Bayesian networks to predict the outcome of injuries is done in steel industries as it is better suited to capture the complexity.[8].

Other methods such as BN Augmented Naïve Bayes (BANs) and General BNs (GB) can also be used. Still, through various researchers it is found that TAN algo is comparably faster than the above-mentioned unrestricted BN methods [9], [10]. Finally, the model has been validated through Sensitivity analysis.

2. METHODOLOGY

First, various parameters responsible for accidents in the steel industry have been identified and segregated through the help of a knowledge graph. Figure 1. shows us all the essential parameters taken for our model. Then based on our critical parameters, all the datasets were retrieved through an integrated steel plant via Incident Data Reporting System (IDRS) [11]. Data wrangling has been done to obtain our desired output from the vast dataset.

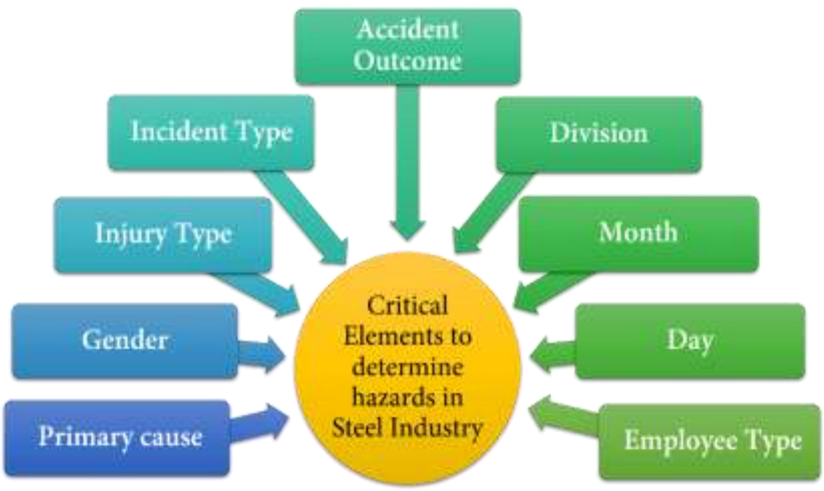

Figure 1. Critical elements to determine hazards in steel industry

Then through the application of TAN algo, an approximation of dependencies between various factors has been determined. This method is used in different fields, like medicine [12], reliability [13], and lifecycle engineering [14].

Finally, Directed Acyclic Graph (DAG) has been obtained as shown in Figure 2. before applying the conditional probability theorem (CPT) responsible for finding the Joint probability of each factor/node.

As shown in Figure 2. various probabilistic inter-dependencies have been generated between the factors through a chain of arcs.

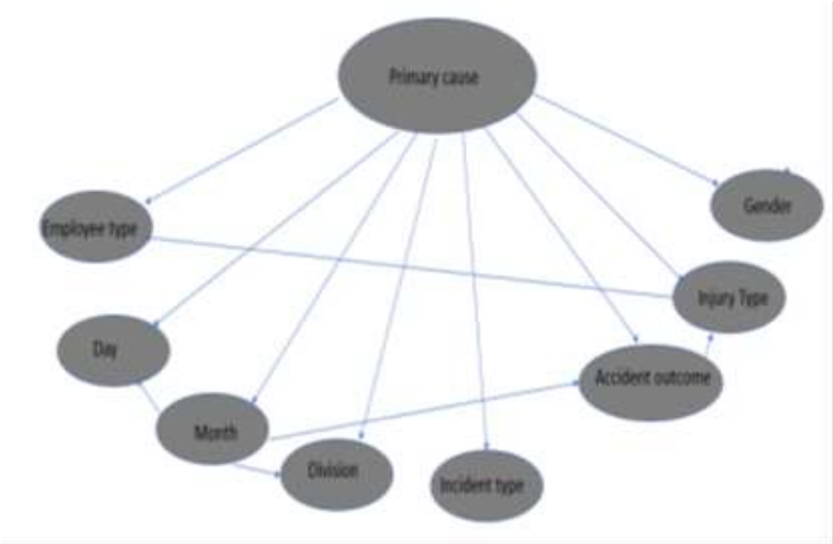

Figure 2. Structure of Bayesian Belief Network

Real possibilities of each node based on the number of parent nodes is to be determined:

- Primary cause has only one possibility because of the absence of any parent node.
- Gender has two possibilities because of 1 parent (Primary cause).
- Injury type has two parent nodes (Primary cause, Accident outcome) and a child node (employee type). So, it has four real possibilities.
- Accident outcome having two parents (Primary cause, Month) and one child (Injury type) has four possibilities.
- Incident type having only one parent (Primary cause) has two possibilities.
- Division having two parents (Primary cause, Month) has four possibilities.
- Month having one parent (Primary cause) and three children (Day, Accident outcome, Division) has two possibilities.
- Day having two parents (Primary cause, Month) has four possibilities.
- Employee type having two parents (Primary cause, Injury type) has four possibilities.

So, there are 27 total conditional possibilities in our model.

Now, using the conditional probability theorem (CPT), the Joint probability of each node is to be found.

To find the joint probability of each node, the formula for conditional probability theorem (CPT). Equation 1 has been used [4].

$$P\ (Xi|Xi\text{-}1, \ldots\ldots\ldots, X1) = P\ (Xi\ |Parents\ (Xi)) \qquad - \quad \text{Equation 1.}$$

Finally, validate our model through sensitivity analysis.

3. RESULT AND DISCUSSION

To understand the effect of different factors on accident outcomes, Bayesian Belief Network model (BBN) has been used; for modelling of BBN software package, Netica is used, which is based on the Tree Augmented Naïve Bayes Algorithm (TAN algo).

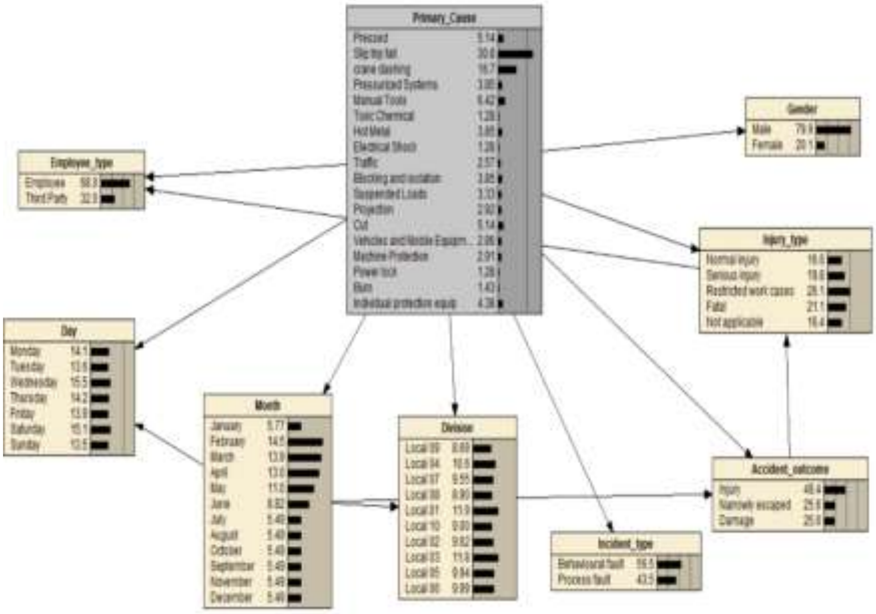

Figure 3. Bayesian Belief Network

Using a tree structure, TAN algo approximates the dependencies between different factors (variables). Based on the experimental analysis of (Jie Cheng Russell Greiner) it has been found that TAN algo is a few times faster than other unrestricted BN-learning methods like BN augmented Naïve Bayes (BANs) and general BNs (GB). Tree Augmented Naive Bayes (TAN) is more suitable for BN than Naive Bayes and selective Naive Bayes [9]. Figure 3. shows the belief network for all the nodes of the Bayesian Network and the probability percentage of each state of each variable. Further, the use of probability percentage of states for sensitivity analysis to analyze how other factors affect accident outcomes has been done.

3.1. Sensitivity analysis

A. Primary Cause

Table 1. Sensitivity analysis of primary cause.

Sensitivity analysis of 'Primary_cause':

Accident_outcome Primary_causes	Injury	Narrowly_escaped	Damage
Initial	24.2	41.8	34
Pressed	28.9	43.1	28
Slip_trip_fall	30.3	41.1	28.6
Crane_dashing	24.5	48.9	26.6
Pressurized_Systems	26.3	55.2	18.6
Manual_Tools	25.5	41.4	33.3
Toxic_Chemical	26.8	43.1	30.1
Hot_Metal	26.4	42.5	30.9
Electrical_Shock	25.8	41.6	32.6
Traffic	21.4	56	22.6
Blocking_and_isolation	23.8	49.3	26.9
Suspended_Loads	27	65.7	7.29
Projection	32.1	35.9	32.1
Cut	28.4	43.1	28.4
Vehicles and Mobile-Equipments	38.9	32	29.1

As shown in Table 1. The initial probability percentage of Injury narrowly escaped and damage (property damage) is 24.2, 41.8, and 34, respectively. As a result, the injury cases are highest for state Vehicles, and Mobile Equipment's probability percentage is 38.9. The instances of narrowly escaped are highest for Suspended state loads, and its probability percentage is 65.7. The property damage cases are significantly low for state Suspended loads and highest for state Manual tools.

B. Employee Type

Table 2. Sensitivity analysis of employee type

Sensitivity analysis of 'Employee_type':

Accident_outcome Employee_type	Injury	Narrowly_escaped	Damage
Initial	49.4	25.6	25
Employee	21.6	29.9	48.5
Thirdparty	31.2	34.8	34.2

Table 2. shows an analysis of employee type and explains how it directly affects accident outcomes. The case of narrowly escaping is the least sensitive, with the highest variation of 9.2% from the initial probability percentage for the state of Third-party employees. Here, injury and property damage cases significantly differ from the initial probability.

C. Month

Table 3. Sensitivity analysis of month

Sensitivity analysis of 'Month':

Accident_outcome Month	Injury	Narrowly_escaped	Damage
Initial	49.4	25.6	25
January	23.2	31.7	45.1
February	24.6	37.2	38.1
March	19.7	43.8	36.5
April	42.2	21.8	36
May	19.7	36.2	44.1
June	16.8	27.3	55.9
July	25.2	36.4	38.4
August	37.4	38.2	24.4
September	51.8	19.2	29
October	21.3	37.6	41.1
November	31.2	43.4	25.4
December	34.8	26.2	39

As depicted in Table 3., the initial probability percentage of injury narrowly escaped and damaged the (property damage) is 24. It's 41.8 and 34, respectively. As a result, the harm cases are highest in September, and its probability percentage is 51.8. The instances of narrowly escaped are highest for March. Its probability percentage is 43.8. Here on, September is significantly sensitive for Property damage; There is a variation of 30.9% from the initial probability rate/percent.

4. Conclusion

The above analysis helps us understand the nature of accidental hazards in the steel industry and how accident outcomes depend on different factors (variables). This analysis will help us understand the situation that can lead to accidents and focus on sensitive parameters to prevent accidents. The above result shows that the factors like primary cause, month, and employee type have a powerful influence on accident outcomes; the rest of the elements were less sensitive to accident outcomes, so their analysis is not discussed. This model can be used for safety enhancement in the steel industry by developing a decision support system.

References

[1] H. Nordlöf, B. Wiitavaara, U. Winblad, K. Wijk, and R. Westerling, "Safety culture and reasons for risk-taking at a large steel-manufacturing company: Investigating the worker perspective," Saf. Sci., vol. 73, no. March, pp. 126–135, 2015, doi: 10.1016/j.ssci.2014.11.020.

[2] C. Li, J. Qin, J. Li, and Q. Hou, "The accident early warning system for iron and steel enterprises based on combination weighting and Grey Prediction Model GM (1,1)," Saf. Sci., vol. 89, pp. 19–27, 2016, doi: 10.1016/j.ssci.2016.05.015.

[3] ILO, Protecting Workplace Safety and Health in Difficult Economic Times. 2013.

[4] G. Ministry of Steel, "25 Safety Guidelines for Iron & Steel Sector," Gov. India, pp. 1–17, 2019, [Online]. Available: https://steel.gov.in/sites/default/files/Framework Document for Safety Guidelines.pdf.

[5] S. Mao, B. Wang, Y. Tang, and F. Qian, "Opportunities and Challenges of Artificial Intelligence for Green Manufacturing in the Process Industry," Engineering, vol. 5, no. 6, pp. 995–1002, 2019, doi: 10.1016/j.eng.2019.08.013.

[6] S. Sarkar, A. Patel, S. Madaan, and J. Maiti, "Prediction of occupational accidents using decision tree approach," 2016 IEEE Annu. India Conf. INDICON 2016, no. December, pp. 3–9, 2017, doi: 10.1109/INDICON.2016.7838969.

[7] S. Sarkar, A. Kumar, S. K. Mohanpuria, and J. Maiti, "Application of Bayesian network model in explaining occupational accidents in a steel industry," Proc. - 2017 3rd IEEE Int. Conf. Res. Comput. Intell. Commun. Networks, ICRCICN 2017, vol. 2017-Decem, pp. 337–342, 2017, doi: 10.1109/ICRCICN.2017.8234531.

[8] G. A. Shirali, M. V. Noroozi, and A. S. Malehi, "Predicting the outcome of occupational accidents by CART and CHAID methods at a steel factory in Iran," J. Public health Res., vol. 7, no. 2, 2018, doi: 10.4081/jphr.2018.1361.

[9] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian Network Classifiers," Mach. Learn., vol. 29, no. 2–3, pp. 131–163, 1997, doi: 10.1023/a:1007465528199.

[10] J. Cheng and R. Greiner, "Comparing Bayesian Network Classifiers," 2013, [Online]. Available: http://arxiv.org/abs/1301.6684.

[11] J. E. Martín, T. Rivas, J. M. Matías, J. Taboada, and A. Argüelles, “A Bayesian network analysis of workplace accidents caused by falls from a height,” Saf. Sci., vol. 47, no. 2, pp. 206–214, 2009, doi: 10.1016/j.ssci.2008.03.004.

[12] S. García-Herrero, M. A. Mariscal, J. M. Gutiérrez, and D. O. Ritzel, “Using Bayesian networks to analyze occupational stress caused by work demands: Preventing stress through social support,” Accid. Anal. Prev., vol. 57, pp. 114–123, 2013, doi: 10.1016/j.aap.2013.04.009.

[13] H. Langseth and L. Portinale, “Bayesian networks in reliability,” Reliab. Eng. Syst. Saf., vol. 92, no. 1, pp. 92–108, 2007, doi: 10.1016/j.ress.2005.11.037.

[14] J. Y. Zhu and A. Deshmukh, “Application of Bayesian decision networks to life cycle engineering in Green design and manufacturing,” Eng. Appl. Artif. Intell., vol. 16, no. 2 SPEC., pp. 91–103, 2003, doi: 10.1016/S0952-1976(03)00057-5.

Biographies

Shivam Kumar received the bachelor's degree in Production engineering from Birla Institute of Technology, Mesra in 2022, and currently working as Graduate Engineer Trainee (Production) in Jindal Steel & Power Limited, Angul, Odisha, India. His research areas include Process safety management, Risk-analysis, Artificial intelligence and lean six sigma.

Arunjeet Chakraborty received the bachelor's degree in Production engineering from Birla Institute of Technology, Mesra in 2022, and currently working as a Business analyst in EXL service. His research areas include Machine learning, Statistics, Image processing, Computer vision and Blockchain technology.

Shatrudhan Pandey is pursuing a Doctorate of Philosophy degree in Production and Industrial engineering from Birla Institute of Technology, Mesra, Ranchi, India. His research areas include Occupational safety and health, Process safety, Hazard identification and analysis, and; Workplace accidents, severity & injury.

Dr. Abhishek Kumar Singh is an Assistant Professor in Production and Industrial engineering at Birla Institute of Technology Mesra, Ranchi, Jharkhand, India. He has over 13.9 years of teaching and research experience. He has published various articles in Peer-reviewed International Journals and authored one book. His research area includes Green manufacturing, Sustainability, Circular economy and Process safety.