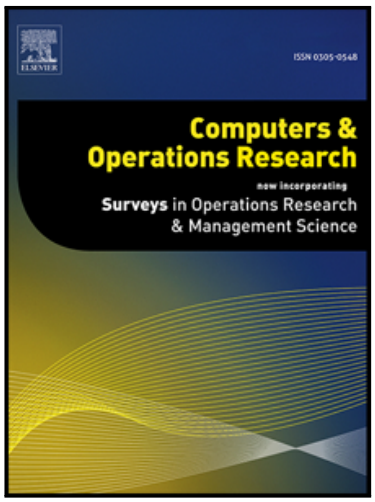

Application of Optimized Machine Learning Techniques for Prediction of Occupational Accidents

Sobhan Sarkar , Sammangi Vinay , Rahul Raj , J. Maiti , Pabitra Mitra

PII: S0305-0548(18)30060-1
DOI: 10.1016/j.cor.2018.02.021
Reference: CAOR 4426

To appear in: *Computers and Operations Research*

Received date: 15 June 2017
Revised date: 11 February 2018
Accepted date: 28 February 2018

Please cite this article as: Sobhan Sarkar , Sammangi Vinay , Rahul Raj , J. Maiti , Pabitra Mitra , Application of Optimized Machine Learning Techniques for Prediction of Occupational Accidents, *Computers and Operations Research* (2018), doi: 10.1016/j.cor.2018.02.021

Highlights

- Both categorical data and unstructured text data have been analysed.
- Incident outcomes are predicted using GA and PSO optimized SVM and ANN approaches.
- PSO-SVM outperforms other algorithms in terms of accuracy (i.e., 90.67%).
- Nine useful rules are extracted using PSO-SVM based C5.0 algorithm.
- Root causes of incidents have been identified.

Title: Applications of Optimized Machine Learning Techniques for Prediction of Occupational Accidents

First author: Sobhan Sarkar
Research Scholar
Department of Industrial & Systems Engineering
Indian Institute of Technology, Kharagpur
Kharagpur: 721302
Email: sobhan.sarkar@gmail.com

Second author: Sammangi Vinay
B. Tech. Student
Department of Mechanical Engineering
Indian Institute of Technology, Kharagpur
Kharagpur: 721302
Email: sammangi.vinay@gmail.com

Third author: Rahul Raj
B. Tech. Student
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur
Kharagpur: 721302
Email: rahul361raj@gmail.com

Fourth author*: Dr. J. Maiti
Professor
Department of Industrial & Systems Engineering
Indian Institute of Technology, Kharagpur
Kharagpur: 721302
Email: jhareswar.maiti@hotmail.com

Fifth author*: Dr. Pabitra Mitra
Professor
Department of Computer Science & Engineering
Indian Institute of Technology, Kharagpur
Kharagpur: 721302
Email: pabitra@cse.iitkgp.ernet.in

* **Corresponding author: Prof. J. Maiti**

Application of Optimized Machine Learning Techniques for Prediction of Occupational Accidents

Abstract

Although, the usefulness of the machine learning (ML) technique in predicting future outcomes has been established in different domains of applications (e.g., heath care), its exploration in predicting accidents in occupational safety domain is almost new. This necessitates the investigation of ML techniques in predicting accidents. But, ML-based algorithms cannot produce best performance until its parameters are properly tuned or optimized. Moreover, only the selection of efficient optimized classifier may not fulfil the overall decision-making purposes as it cannot explain the inter-relationships among the factors behind the occurrence of accidents. Hence, in addition to prediction, decision making rules are required to be extracted from the accident data. Considering the above-mentioned issues, in this research, optimized machine learning algorithms have been applied to predict the accident outcomes such as injury, near miss, and property damage using occupational accident data. Two popular machine learning algorithms, namely support vector machine (SVM) and artificial neural network (ANN) have been used whose parameters are optimized by two powerful optimization algorithms, namely genetic algorithm (GA) and particle swarm optimization (PSO) in order to achieve higher degree of accuracy and robustness. PSO-based SVM outperforms the other algorithms with highest level of accuracy and robustness. Furthermore, rules are extracted by incorporating decision tree C5.0 algorithm with PSO-based SVM model. Finally, a set of nine useful rules extracted to identify the root causes behind the injury, near miss and property damage cases. A case study from a steel plant is presented to reveal the potentiality and validity of the proposed methodology.

Keywords: Occupational accidents, Support Vector Machine, Artificial Neural Network, Genetic Algorithm, Particle Swarm Optimization, Rule extraction.

1. Introduction

According to International Labour Organization (ILO) estimation, globally about 2.3 million workers succumb to death annually due to occupational accidents and diseases which include approximately 3.6 lakh fatal accidents [1]. Overall, nearly 337 million occupational accidents are reported per year. From ILO report, it is revealed that approximately 4% of the annual gross domestic product (GDP), which is equivalent to US $1.25 trillion, is drained off due to occupational accidents [2]. From EUROSTAT, it is reported that each year, 3.2% of workers in the European Union, i.e., EU-27 meet

an accident at their working places [3]. In relation with this, ILO also makes the following comments: ñFatalities are not fated; accident do not just happen; illness is not random; they are causedò [4]. The basic causes of accidents are unsafe conditions or unsafe acts or both. There are multiple factors contributing towards an accident. There are many theories available in literature that explain the causation of accidents. Khanzode et al. [5] explained the various theories in their study behind the accidents such as accident proneness theory [6], Domino theory [7], injury epidemiology [8], system theory [9], sociotechnical system theory [10], and macro-ergonomic theory [11]. An injury event is occurred due to the presence of a chain of events or causal factors. If the causes are known, the outcomes (i.e., accidents) can be predicted. In addition, the predictive models will quantify the contribution of the various causal factors towards an accident to happen.

Predictive models for occupational accidents can be statistical learning based or machine learning (ML) based. Owing to the large amount of data available, ML supersedes traditional statistical counterpart in predicting future events that has been used in various fields such as engineering, medical science, finance, and it renders very useful results [12]. However, a review of literature shows that the ML techniques have been used in occupational accident analysis on a limited basis [13]. So far, studies made on occupational analysis show the use of ML techniques in terms of their predictive power [14] and explanatory capacity [15]. These methods, based on historical data from incident reports, or interview with employees, ensure their advantages over conventional statistics in terms of predictive functions and importance of predictors with a bearing on incident outcomes. The potential benefits of ML can not only be realized from the capability of processing large quantity of data but also from: (i) their capability to deal with large dimensional problems, (ii) their flexibility in reproducing the data generation structure irrespective of complexity, and (iii) their predictive and interpretative potential through the extraction of rules. Due to the capability of ML techniques, it has been used successfully in several domains including occupational accident analyses. However, the ML techniques do not produce good results if their parameters are not tuned. Optimization of parameters can provide better results. Usually, the concept behind the optimization is to search the optimal solution of the key parameter values that helps classifiers perform best on given data set. Several studies in literature are available which show the utility of parameter optimization of ML techniques in different domains using genetic algorithm (GA), particle swarm optimization (PSO) and so on [16].

Therefore, the primary objective of the present study is to develop a prediction model using machine learning techniques, namely SVM, and ANN for the prediction of occupational incident outcomes. In order to achieve the better accuracy, optimization techniques i.e., GA and PSO have been employed on the classifiers. In addition, rule extraction for the occurrence of injuries has been performed by the PSO-SVM-based classifiers combined with decision tree (C5.0). The secondary objective includes the identification of the relevant variables attributable to incident outcomes using chi-square feature

selection technique. The results of the analysis show the utility of the SVM classifier in terms of both prediction as well as rule extraction purposes.

2. Review of literature

In the domain of occupational accident prediction, there are many ML algorithms used such as support vector machine (SVM), artificial neural network (ANN), extreme learning machine (ELM), and decision tree. In the application of DT in accident analysis, some algorithms like C4.5, C5.0, classification and regression tree analysis (CART), Chi-square Automatic Interaction Detector (CHAID) etc. are usually used for prediction of occupational accident. The main aim to use DT is to predict and interpret qualitative and quantitative patterns lying in data which leads to exploration of hidden information. Due to its relaxation on assumptions on distribution of attributes or independence of attributes, DTs have been successfully used in different fields like medicine [17], social sciences [18], business management [19], construction engineering and management [20], process industry [13].

Other than DTs, algorithms like artificial neural network (ANN), Bayesian classifier, adaptive neuro-fuzzy inference system (ANFIS), Bayesian network (BN), support vector machine (SVM), extreme learning machine (ELM) have been used in different domains like construction industry [21], mining industry [21], ship building industry [22], service industry [14] etc. In 2008, Matias et al. used SVM, ELM (i.e., feed forward neural network), BN techniques for the analysis of causes and types of accidents like floor-level falls [14]. They used 148 records obtained from different companies during 2003 to 2006 in Spain. As results, BN is found to be higher predictive capacity than others. Sánchez et al. carried out one study using SVM to classify those workers suffering work-related accidents for a year [1]. They analysed the data consisting of 11,054 responses of the workers employed in all economic activities in Spain. Their findings show that SVM performs better than back-propagation neural network (BPNN) without over-fitting problems. In 2011, Rivas et al. modelled the accidents and incidents in two companies in construction and mining to identify the most important causes of accidents and developed predictive models using BN, SVM, and other ML techniques [21]. Bayesian network (BN)-based prediction model have also been used by many researchers in different sectors like mining [14], construction [14]. For example, Sanmiquel et al. carried out one study in Spanish mining sector and analysed the 69,869 instances of occupational accidents during 2003 to 2012 using BN [23].

Another important prediction model used in this domain is ANN. Due to its important characteristics like the ability to learn from data, distributed memory, parallel operation and fault tolerance, it has been widely used in diverse field of study along with in occupational accident domain. For instances, He et al. attempted to solve the problem of coal and gas outburst by classification technique using backward algorithm of ANN (BA-ANN) and exponent evaluation method (EEM) [24]. Using BA-

ANN, the weights of factors are calculated towards the response variables (i.e., coal and gas outburst). Yi et al. developed an early warning system for the workers in hot and humid environments using ANN [25]. They have collected 550 data related to work, environment, and individuals which are analysed by ANN to predict rating of perceived exertion (RPE) of the workers in the construction sites. Apart from the application of ANN in occupational accidents, there are plenty of literature available for other accidents of which research on road accidents is found to have attained more focus [26]. More interestingly, artificial intelligence (AI) approaches like ANN are found to have greater performance in terms of prediction than regression analysis. Reviewing the literature on accident analysis domain, techniques like SVM and ANN are found to be popular and useful as they have a robust theoretical grounding that enables the successful learning from the data, capability of handling any level of complexity except computational complexity of the problem and flexibility with non-parametric philosophy.

However, all these machine learning algorithms do not provide optimal results like classification accuracy and understandability if the parameters of them are not properly tuned. To tune the parameters of the classification algorithms, optimization methods are found to be most useful than other techniques like manual tuning or grid search. In occupational accident research, hardly any study or no study has been reported that uses optimization techniques on classifiers (like SVM, or ANN) in order to obtain better classification accuracy. From the research of the other domain, it is observed that in order to get enhanced accuracy in SVM model, penalty factor (c), and kernel parameter (γ) are considered to be optimized [27]. There are many optimization techniques used for this purpose like genetic algorithm (GA), particle swarm optimization (PSO), gradient descent method, etc. [28]. Of them, GA and PSO are found to be the most popular methods used for optimizing the parameters of classifiers (e.g. SVM) to achieve higher accuracy [29]. Therefore, in this paper, GA and PSO have been selected for parameter optimization of SVM. Similarly, for ANN, there exist several parameters which can be optimized like number of layers, input and hidden neurons, type of transfer functions, topology of ANN, weights, thresholds etc. Li et al. used initial parameters, network topology, weights, and thresholds of back-propagation neural network (BPNN) based on memetic algorithm with GA [30]. Xue & Liu used only initial weights and threshold values for BP model for predicting liquefaction susceptibility of soil [31]. Das et al. focused on optimizing the weights, transfer functions, and topology of ANN for channel equalization [32]. Most of the studies showing the importance of optimization techniques on classifiers have been carried out to solve the problems in different domains other than occupational accident. Hence, it is required to implement such useful optimization techniques on classifiers to improve their performance like predictive accuracy.

However, the performance of the classifiers not only depends on optimized parameter values, but also the types of data used. As it is a known fact that numerical attributes hold more information than categorical attributes or free-text attributes, thus dealing with different data types also impacts classifierôs performance. Hence, it is really a challenging task for researchers to extract pattern from different types of data like categorical or more specifically, free-text data. Most of the literature in accident domain used either numerical data or categorical data for the analyses of accident scenarios. However, analysis of free-text data remains under-utilized in most of the cases as it is really hard task to extract pattern from the passage of free-text. Narrative text is one of the key resources for the prediction of accident. It provides the valuable additional information in analysis along with other types of data. To investigate the importance of narratives in prediction of occupational accidents, Jones & Lyons showed increase of home injuries identified by 19%, rugby injuries by 137%, and assaults by 26% [33]. Li & Guo tried to analyse the aviation safety data with the help of topic modelling techniques [34]. Related to this, a noteworthy contribution made by Brown is to analyse rail accident data to explore the main contributors behind the accident using text mining associated with other techniques like Latent Dirichlet Allocation (LDA), and Random Forest [35]. Thus, the main challenge lies in the analysis of the unstructured text. To tackle the issue, Tixier et al. tried to develop a system that could overcome the problem by decoding the unstructured reports from accident database [36]. The system developed by them could use the unstructured injury database with 101 attributes, and produced with 95% classification accuracy. Vallmuur, therefore, mentioned in his study that future research on injury analysis would direct a continued growth and advancement in the application of text mining for utilizing information within text [37]. The primary difficulty in the analysis of free unstructured text is the sparsity and high-dimensionality of document-term matrix. Moreover, text mining-based approach cannot capture the order of words and semantic meaning of them. One recent study by Pavlinek & Podgorelec shows that topic modelling of free-text could help in text classification task and reduction of sparsity [38]. The study by Niraula et al. revealed the superiority of topic modelling in a supervised setting [39]. Consequently, many classification algorithms have been implemented using topic modelling in different domains. In road accident analysis, one study by Pereira et al. used topic modelling of incident reports of traffic to extract information in real time to predict incident duration [40]. Therefore, topic modelling of under-utilized free unstructured text has full potential in extraction of latent information within text field that facilitates the prediction of occurrence of accidents.

Prediction analysis, as a standalone tool, may not serve the entire purpose of the accident analysis until a prescriptive analysis is not made for the interpretation of the accident causation. Rule extraction and its interpretation from the accident data set are often considered to be an effective approach. The rules can generally be obtained by using either decision tree (DT), or by association rule mining (ARM) approach. In several studies of occupational accident, DT has been used for rule

extraction and interpretation more than ARM. DT is found to be useful when target function is discrete valued, when it is describable by attribute-value pairs, or when the data sets are noisy trained. DT works well in rule extraction when the data set used for DT analysis are more informative than others. Otherwise, other data sets having less information might lead to generation of low-quality rules. Therefore, selection of the set of data, which is informative, is required for better rule building and rule interpretation. Some of the previous studies showed that the rule extraction based on support vectors identified by support vector machine is useful as the rules are less in number, and interpretable [41]. DT algorithms with SVM have also been attempted to turn the black box of SVM decisions into transparent and comprehensible rules which can be utilized as secondary opinion for any decision-making task.

Based on the above-mentioned literature, it is found that none of the previous literature in occupational accident domain has reported to use text data and categorical data together for the building of prediction model. Moreover, to the best of authorsô knowledge, optimization techniques on classification algorithms to get optimal solutions have not been also addressed by any researchers previously in this domain. Another important point to be noted is that very less studies have been conducted so far for the prevention of accidents in steel industry, whereas previous research focused more on either construction, or mining industry. Therefore, there remains a strong need of research on prevention of occupational accident using machine learning techniques for steel industry.

2.1. Research issues and contribution of the study

Based on the review of literature presented above, the following issues have been identified in the domain of prediction of occupational accident analysis.

(i) None of the previous studies reported have shown the combined analysis using text and non-text attributes together for incident prediction.
(ii) None of them reported the parameter optimization of the classifiers for better prediction accuracy.
(iii) There are no studies reporting the SVM-based rule extraction for incident occurrences.

Realizing the issues in accident literature, our study, therefore, endeavours to contribute in the following ways:

(i) The study takes care of text and categorical attributes together for predicting the incident categories,
(ii) It uses optimization algorithms for parameter optimization of the classifiers for improved prediction accuracies,

(iii) It includes SVM-based rule extraction method for injury, near miss, and property damage cases

(iv) It identifies the importance of the predictors towards occurrence of incidents, and

(v) The developed methodology is validated with a case study in a steel plant.

The rest of the paper is organized as follows: Section 3 describes the methods used in this study; Section 4 presents the case study with data set and data pre-processing tasks; in Section 5, results and discussion are presented; and finally, conclusions with future scopes of the present study are discussed in Section 6.

3. Methods

In the methodological section, topic modelling, SVM, ANN, GA, PSO, and PSO-SVM combined DT-based rule extraction methods are discussed briefly. The total proposed methodological flowchart is depicted in Fig. 1. There are three important phases shown in the flowchart. They are: (i) ***Data pre-processing phase*:** In data pre-processing, three important tasks, namely feature addition, missing value handling, and evaluation of feature importance are performed on the data set. The initial data set has 1500 incident records and 16 attributes (15 categorical and one text) with a very low percent of missing values in three of them. Four categorical attributes, which are found to be interrelated in nature, are combined into one new attribute. In addition, a new attribute or feature is generated from text data using topic modelling technique which will be discussed in subsequent sections in details. Thereafter, missing value imputation has been done using random forest. Finally, feature importance is calculated. The final data set generated after this phase has 1500 records and 13 attributes (all categorical) without any missing value; (ii) ***Optimization & prediction phase:*** In this phase, optimization techniques, namely genetic algorithm (GA) and particle swarm optimization (PSO) have been implemented on two classifiers, namely SVM and ANN using 10-fold cross validation. Then, classifier with the highest accuracy is considered as the best one; and finally (iii) ***Rule extraction phase:*** In this phase, useful rules from the best classifier i.e., PSO-SVM combined with C5.0 decision tree are extracted. All the processes are illustrated in following sections.

<Insert Fig. 1>

3.1 Topic modelling

In machine learning and natural language processing (NLP), a topic model can be described as a type of statistical model to extract the abstract ñtopicò or ñclassesò. In topic modeling, latent Dirichlet allocation (LDA) is a very popular approach. In order to use LDA, we need to fix the number of topics is required to be fixed. There are several metrics developed by researchers to select optimal number of topics for LDA model. We used four of those metrics in our study. óMetric1ô, developed by

Griffiths & Steyvers [42], shows that the number of topics for which log-likelihood of the data becomes maximum is considered to be optimal. Cao et al. [43] developed ómetric2ô where they have used average cosine distance between every pair of topics to measure the stability of topic structure. It was observed that smaller the average distance, better the stability. Similarly, ómetric3ô has been developed by Arun et al. [44]. The measure is computed in terms of symmetric Kullback-Leibler (KL) divergence of salient distributions that are derived from these matrix factors. It was also observed that the divergence values become lowest for optimal number of topics. Recently, another metric ómetric4ô has been developed by Deveaud et al. [45]. They proposed a simple heuristic that estimates the number of latent concepts of a user query by maximizing the information divergence between all pairs of topics of LDA. So, when put together, in order to find optimal number of topics, Metric2 & Metric3 should be minimized and Metric1 & Metric4 should be maximized. The detailed description of basic principle of topic modeling and its application are presented in [40].

3.2 Support Vector Machine (SVM)

SVM, developed by Vapnik [46], is an emerging machine learning technique in statistical learning theory of multi-dimensional function which is used for classification and regression analysis. It holds an ability of being universal approximators of any multivariate functions to any desired level of accuracy. Initially, it was developed for regression tasks, but later was used as a powerful classifier. According to the previous studies [47], SVM has been used in most engineering fields with good accuracy. Theoretically, it has less overfitting problem, and better generalization ability. However, the main problem encountered in constructing SVM model is to adequately select training parameter values as inappropriate parameter setting leads to poor prediction accuracy. The readers may refer [48] for basic understanding of the working principle of SVM.

3.3 Artificial Neural Network (ANN)

ANN is an artificial model of the human brain which can learn through adapting the present situations. It consists of interconnected network of neurons and synapses. Usually, it has three layers (i.e., input, hidden and output) or more (when more than one hidden layer). Hidden layers are considered the root of all calculations in ANN. A network gets activated when a set of inputs are triggered that consequently produce desired results through output layers. Each input value is multiplied by its corresponding weight layers, then it is summed up and added to a scalar parameter called bias, which in turn generates output through final output layer. Modifying connection weights and biases using appropriate learning algorithm, training process can be accomplished. Many evolutionary algorithms or gradient descent methods have been used in this training process by updating weights and biases. At each iteration, they are modified until prediction error of the network gets minimized. Out of many learning algorithms, back propagation (BP), which is a gradient type of

adjustment for the modification of weights, has been used in the paper of Benjio et al. [49]. Basically, the output of any node is determined by a mathematical operation on the input of the particular node. This operation is called the transfer function which facilitates the transformation of inputs into output either in linear or non-linear manner. There are three types of transfer functions used commonly in literature i.e., sigmoid, hyperbolic tangent, and linear. In this paper, sigmoid transfer function has been used.

3.4 Working principles of GA and PSO on classifiers

In this section, the two optimization algorithms i.e., GA and PSO have been described briefly on how they are used to optimize the parameter values of the classifiers, namely SVM and ANN. For detailed description of GA and PSO, interested readers are requested to go through [50,51]. For GA, initial population is generated at random. Then the data is split into training and test sets. Fitness function is developed by which fitness value i.e., accuracy of the classifier is computed for each chromosome for each iteration. Then, the criterion for termination of the algorithm is checked. Here, we have used the maximum number of iteration (i.e., 500) as termination criterion for both GA and PSO operations. If it is not satisfied, it goes for crossover, and then mutation process which ultimately creates a new population. Recursively, this process continues until the termination criterion gets satisfied. Once it is satisfied, optimal parameter values for both SVM and ANN are achieved which will be ultimately used for model building for prediction of incident outcomes (see in Fig. 2). Similarly, in Fig. 3, the process of optimization of SVM/ ANN parameters using PSO has been depicted.

<Insert Fig. 2>

<Insert Fig. 3>

3.5 C5.0

C5.0 is a decision tree algorithm developed from C4.5. In C4.5, when training the model, all training samples are set as the root of the decision tree. Then, the gain information ratio of every feature is calculated based on the entropy of the feature, and the feature with the highest information gain is selected to split the data into multi-subsets. The algorithm repeats this procedure on each subset until all instances in the subset belong to the same class and a leaf node is created. For detailed understanding, interested readers may refer [52].

3.6 SVM and DT-based rule extraction

SVMs and artificial neural network (ANN) have shown better performance than other machine-learning algorithms in some application areas, such as speech recognition, computer vision, and

Table 1. Top eight terms across each topic and extracting a meaningful event from them

Topic	Top eight terms	Meaningful event
1	0.051*person + 0.045*one + 0.044*hit + 0.019*remove + 0.019* piece + 0.016*injury + 0.014*take + 0.014*roof	hitting by foreign body
2	0.054*operate + 0.043*crane + 0.022*roll + 0.019*coil + 0.019* place + 0.018*lift + 0.016*work + 0.016*move	Crane operation failure
3	0.039*fell + 0.029*work + 0.025*fall + 0.019*ground + 0.015* due + 0.015*level + 0.015*plate + 0.014*floor	Falling from heights
4	0.044*fire + 0.031*cable + 0.024*damage + 0.02*excavate + 0.017*power + 0.016*work + 0.015*site + 0.014*weld	Fire incidents
5	0.041*shift + 0.034*left + 0.031*first + 0.03*aid + 0.024*leg + 0.021*near + 0.02*duty + 0.019*plant	First aid incidents
6	0.046*area + 0.028*material + 0.019*belt + 0.018*end + 0.018* conveyer + 0.016*engage + 0.015*clean + 0.013*around	Incidents during cleaning
7	0.03*job + 0.029*due + 0.029*pipe + 0.024*gas + 0.023*line + 0.023*water + 0.019*open + 0.016*came	Pipe leakage
8	0.044*side + 0.028*load + 0.025*dumper + 0.023*driver + 0.021*road + 0.019*vehicle + 0.018*gate + 0.016*toward	Vehicle hitting/collision
9	0.068*got + 0.067*hand + 0.051*injury + 0.043*right + 0.034* cut + 0.033*finger + 0.027*slip + 0.027*left	Slipping

Table 2. The utilised GA parameters

SL	GA Parameter	Value
1	Population size	12
2	Number of generations	500
3	Crossover probability	0.8
4	Mutation probability	0.1
5	Elitism	0.05

Table 3. Optimal parameter setting of SVM models

Model	Iterations for convergence	Iterations	Cost for best solution	Gamma for best solution	Best Accuracy (%)
GA-SVM	212	500	1.1093	0.2474	90.53
PSO-SVM	142	500	1.3405	0.2257	90.67

Table 4. The utilised PSO parameters

SL	Parameter	Value
3	Number of generations	500
4	Swarm size	12
5	Exponent for calculating number of informants	3
6	Exploitation constant	0.721
7	Local exploration constant	1.193
8	Global exploration constant	1.193

Table 5. Optimal parameter setting of ANN models.

Model	Iterations	Iterations for convergence	Number of hidden layers for best solution	Number of nodes in hidden layers for best solution	Learning rate for best solution	Best Accuracy (%)
GA-ANN	500	202	1	15	0.0335	89.07
PSO-ANN	500	18	1	30	0.0189	89.33

Table 6. Rules generated from optimized SVM and C5.0-based model.

Rule no.	Rules	Class	n or n/m	Lift	Confidence
R1	Day of Incident in {Friday, Sunday, Thursday, Wednesday } + Division in {Div10, Div11, Div12, Div2, Div3, Div4, Div9} + Primary Cause in {DC, EF, EMD, RI}+ Topic in {Topic 5, Topic 6, Topic 9}	Injury	35/1	2.0	0.946
R2	Division in {Div13, Div6, Div7} + Injury Type = ITNA	Near Miss	114/2	2.2	0.974
R3	Injury Type = ITNA + Primary Cause in {D, EI, FE, GL, HM, HP, LTT, MA, MH, OI, PI, R, RO, S, SI, STF, TC, WH}	Near Miss	386/16	2.1	0.956
R4	Day of Incident in {Monday, Saturday, Tuesday} + Division in {Div10, Div11, Div12, Div2, Div4, Div8} + Injury Type = ITNA + Topic in {Topic 5, Topic 6, Topic 9}	Near Miss	23/1	2.0	0.920
R5	Division in {Div10, Div12, Div2, Div9} + Injury Type = ITNA + Topic in {Topic 2, Topic 3, Topic 4, Topic 7, Topic 8}	Near Miss	155/15	2.0	0.898
R6	Injury Type = ITNA + Topic in {Topic 2, Topic 3, Topic 7 , Topic 8}	Near Miss	350/38	2.0	0.889
R7	Division in {Div3, Div8} + Injury Type = ITNA+ Primary Cause in {DC, EF}	Property Damage	5	12.6	0.857
R8	Day of Incident in {Friday, Monday, Tuesday} + Division = Div11 + Injury Type = ITNA + Primary Cause in {DC, EF, EMD}	Property Damage	11/3	10.2	0.692

R9	Division = Div4 + Injury Type = ITNA + Primary Cause in {DC, EF, EMD} + Topic = Topic 4	Property Damage	27/8	10.2	0.690

Graphical abstract: