International Journal of Occupational and Environmental Safety

Machine learning in Occupational Safety and Health – a systematic review

S. Maheronnaghsh[a], H. Zolfagharnasab[b], M. Gorgich[c], J. Duarte[d]
[a]Associated Laboratory for Energy, Transports and Aeronautics – LAETA (PROA), Faculty of Engineering, University of Porto, PT (up201600475@fe.up.pt) ORCID 0000-0002-4966-4292, [b]FARO Technologies, PT (hooshiar.h.z@ieee.org) ORCID 0000-0001-8635-9169, [c]Faculty of Engineering, University of Porto, PT (up201306568@fe.up.pt) ORCID 0000-0002-1454-5752, [d]Associated Laboratory for Energy, Transports and Aeronautics – LAETA (PROA), Faculty of Engineering, University of Porto, PT (jasduarte@fe.up.pt) ORCID 0000-0002-5856-5317

Article History

Received May 09, 2022
Accepted April 26, 2023
Published August 02, 2023

Keywords

Artificial intelligence (AI)
Decision support systems
Hazard prediction
Occupational health and safety
Sensor

DOI: 10.24840/2184-0954_007-001_001586

ISSN: 2184-0954

Type: Review

Open Access
Peer-Reviewed

Abstract

With the development of technology, machine learning (ML), a branch of computer science that aims to turn computers into decision-making agents using the most appropriate algorithms, is also paving its way in the modern world. This systematic review arises from the need to understand the impact and report the best practices for applying ML in occupational safety and health. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines were used to provide the best research results. From the 759 identified papers, only 42 were included in the study after applying both exclusion and inclusion criteria. Application is primarily used in accident and risk assessment, and construction and office work are the leaders in applications. The applied methods mainly consist of classification (injuries, accidents, monitoring data), prediction (of hazards), and regression (to find patterns of accidents to prevent them). In conclusion, decision-makers and workers are taking advantage of various artificial intelligence techniques to find solutions in the occupational safety and health environment when experts have access to correct data, either in real-time or recorded datasets. However, it is necessary that in future investigations, limitations of using ML applications in occupational safety and health area be improved and their full potential is achieved.

1. INTRODUCTION

Occupational safety and health (OSH) at workplaces are routinely assessed to assure worker comfort, facilitate equipment use, protect from hazardous exposures, and prevent health hazards (Choi et al. 2020). Not only OSH experts but also researchers and managers have addressed a variety of components, including physical, chemical, biological, ergonomic, and organizational environments, with traditional methods until the Industry 4.0 era (Oliveira, Lopes, and Bana e Costa 2018). Industry 4.0 refers to the improvement of workplaces by employing the digital revolution, such as Artificial Intelligence (AI), the Internet of Things (IoT), and smart devices (Badri, Boudreau-Trudel, and Souissi 2018). Rapid advances and large-scale integration of information technologies (IT) into workplaces, like sensors, mobile systems, massive data processing, and cloud technologies, affect not only the way workers work but also the content and the quality of work (Suarez and Riesgo, 2005). Therefore, this new industrial era based on generated and stored information can also assist in measuring, assessing, and controlling health and safety issues more efficiently while being more effective (Cho et al., 2018; Jin et al., 2019; Xie and Chang, 2019).

Machine learning can be effectively used to create expert systems that expose intelligent behaviors, provide solutions to complex problems, further assist in processing massive data, and find the most appropriate patterns for big data fitting. ML solutions take over traditional computing methods in industries, but ML solutions can also change occupational health performance by modernizing workplace safety approaches.

Machine learning provides statistical models or algorithms to train, provide, or generate models that can be utilized to perform predictions, classifications, estimations, or similar tasks. ML methodologies can also generate models progressively to improve predictions. The models can benefit from past predictions and attempt to improve the success rate of further predictions (Choi et al., 2020).

Several methodologies are being employed in ML, including supervised (Common Algorithms like Neural networks (NN), K-Nearest Neighbor (KNN), Naive Bayes, Decision Tree (DT), Support Vector Machines (SVM), unsupervised (Common Algorithms: k-means clustering), semi-supervised learning, and reinforcement learning (Q-Learning, Temporal Difference (TD), Deep Adversarial Networks) (Weng 2019). Each methodology can be implemented using several techniques, including NN and convolutional NN (mostly known as deep learning or DL), support vector machines, capsule networks, and DT.

In the form of classification or prediction, ML can enhance human decisions by considering factors that might be ignored in a human decision (correlated or hidden factor) or speed up the decision process in terms of real-time reaction. Thanks to the repeatability of several hazards, trained models over a dataset of events can be transferred to other datasets, which makes the trained model portable when it is demanded to train several models. This portability increases the importance of ML approaches for generating models for events that have not been recorded on a large scale.

Nevertheless, safety hazards can be determined; the effects can be recognized or even predicted, and safety methods and equipment can usually be described. This study reviews previous investigations to assess the ML approaches appropriate to OSH issues. Furthermore, it highlights specific ML methodologies, which have been employed successfully in various fields of OSH, as well as the possibilities and challenges.

The obtained results from this review will help researchers, stakeholders, managers, and experts identify new technologies in OSH and help recognize knowledge gaps in the current literature, directing future research that will benefit and support the growth of OSH in the industry.

2. METHODOLOGY

This study methodology was based on the systematic review protocol of Maheronnaghsh et al. (2021), where the Preferred Reporting Items for Systematic reviews and Meta-Analyses (Moher et al. 2009; Page et al. 2021) guidelines were used to help conduct the research and data treatment.

The selected databases for the research were the IEEE, SCOPUS, PubMed, Science Direct, Inspect, and Web of Science. The keyword combinations were sought in "Title/Abstract/Keywords" on every website, and the research results were recorded in an *Excel* table file. The research expression can be summarized as follows:

("machine learning" OR "expert system" OR "Cognitive system") AND ("occupational safety" OR "occupational health" OR "work environment")

Further information regarding exclusion criteria and search strategies are detailed in the protocol (Maheronnaghsh et al., 2021).

Regarding eligibility criteria, a new criterion was added to the proposed protocol. Since this systematic review intended to evaluate standard machine learning models dealing with mathematical and statistical approaches trained via supervised methodologies, studies using either genetic or fuzzy algorithms were not considered. That way, only the methods using separate training and testing data were included to have a fair comparison.

3. RESULTS

3.1. Selected articles

Following the proceedings from the PRISMA Statement (Moher et al. 2009; Page et al. 2021), 759 records were identified in the first phase.

The exclusion criteria were: 1) date – 269 documents were excluded, 2) type of document – 187 excluded papers, 3) source type – one document excluded, and 4) language – also one document excluded. After screening the title and abstract and checking them against the systematic review aim, 236 papers were excluded. Duplicates, 14 articles, were removed.

In the eligibility phase, 51 papers were full-text screened. In this stage, 22 articles were excluded for several reasons: studies that did not use any concept of ML, such as training, dataset, feature extraction, or selection. Besides, those papers that explained underneath layers of methods used (such as fuzzy method) have been excluded. The fuzzy exclusion method was aligned to have evaluation only on mathematical and statistical ML approaches. As mentioned before, such discrimination ensures that there will be only evaluations on methods that use common training and testing techniques. At the end of this phase, 29 papers were considered eligible and included in the research. After applying the snowballing technique (Wohlin 2014), 13 more papers were added to the research.

Figure 1 demonstrates the research strategy and overall review process, including articles excluded at each stage.

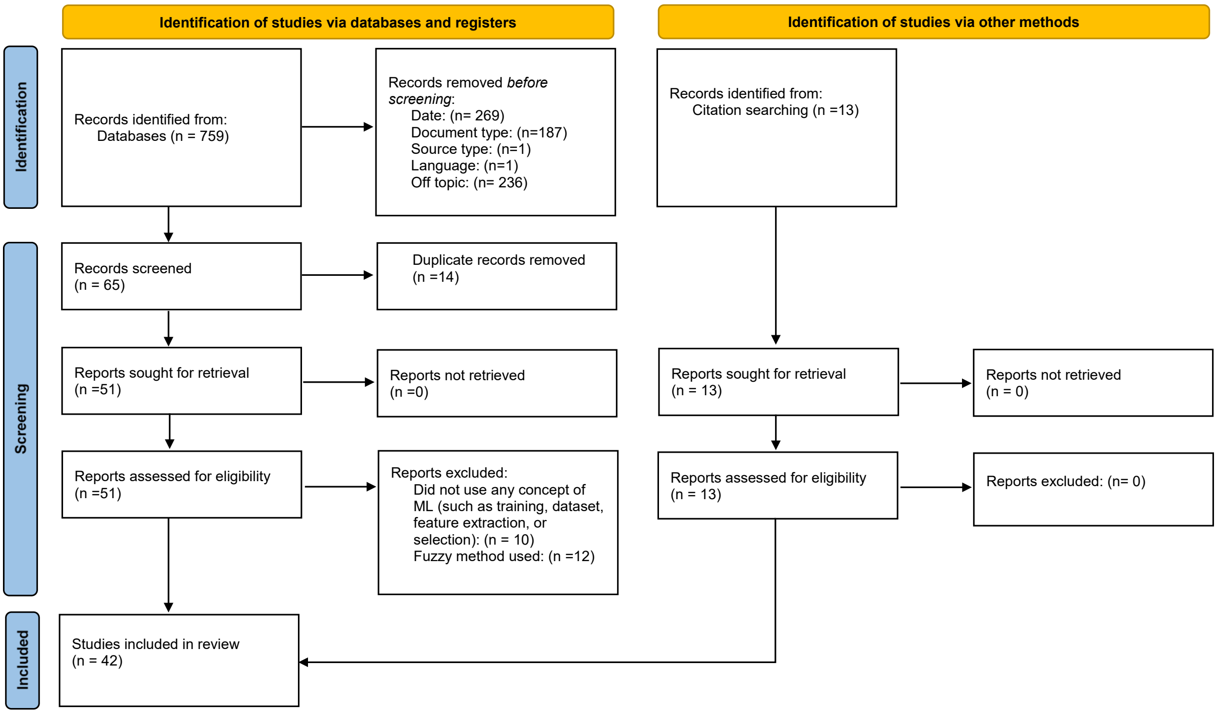

Figure 1. PRISMA Flow Diagram (adapted from Page et al., (2021))

3.2 Presence of Bias Within the Selected Articles

The risk of bias in studies was assessed using an adaptation of the assessment method proposed in the Cochrane Collaboration tool adaptation to evaluate the risk of bias (Higgins et al. 2011). The assessment considered three categories of bias classified as low, high, and "unclear" risk for each topic, where "low risk" is applied to parameters that do not present a significant effect on the results, and "high risk" is used whenever the parameter influences the results, and "unclear risk" applied whenever it is not possible to establish some correlation between cause and effect. The analyzed parameters were an assessment of the proposed classification method, usability of the method, dataset used for training, sensors, reporting quality, and reference quality (Appendix 1).

3.3 Study characteristics

Figure 2 shows that the publication date of the papers ranged between 2013 and 2020, with the number of studies varying each year. According to the same figure, this topic has recently attracted more interest from researchers.

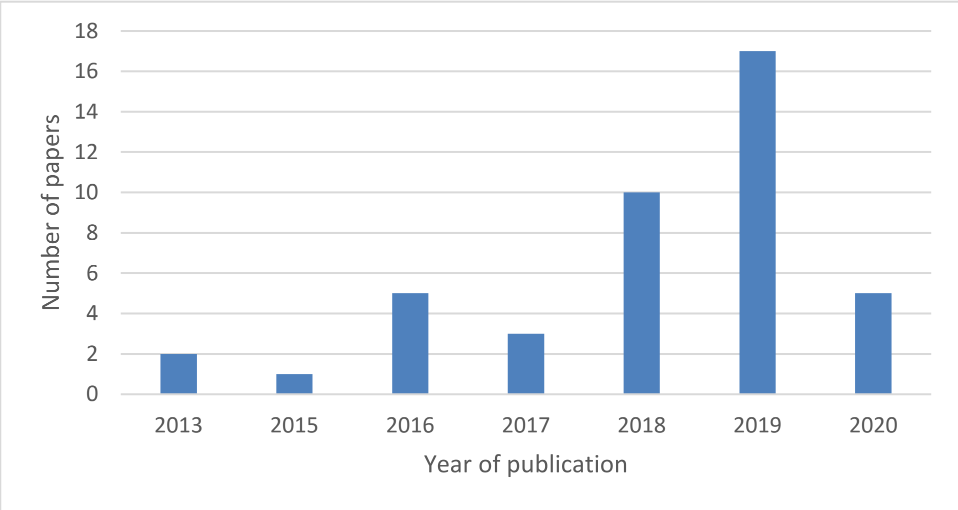

Figure 2. Number of articles by publication year

Studies were published across 11 countries, with the majority of first authors from the United States of America (16 of the 42), China (5), and Korea (5). Figure 3 illustrates the number of articles published across the different countries.

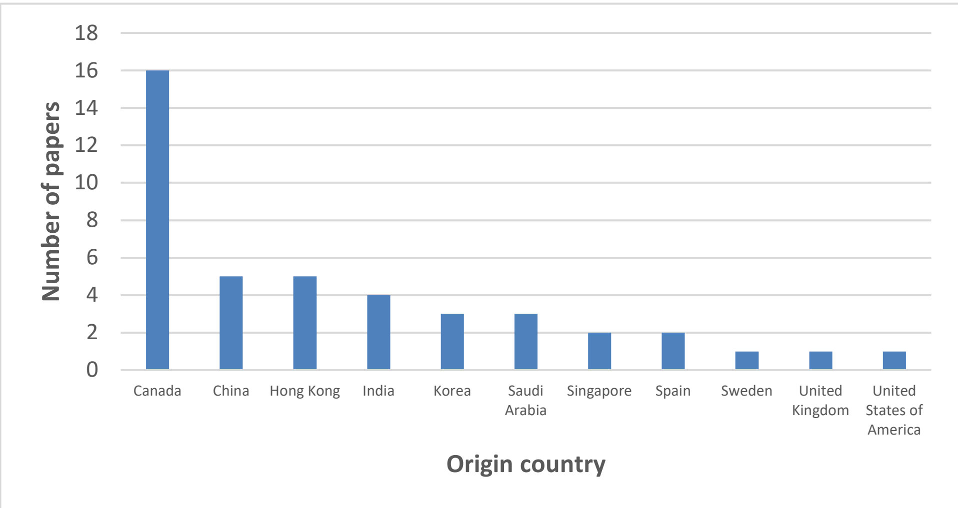

Figure 3. Number of articles by country

3.4 Aim and methodology

The included papers showed a high variability of field (and type of workplace), observed in Figure 4. The construction industry leads the pool with 15 published papers, followed by office work (four papers) and power infrastructure (three papers). Six of the 42 papers did not specify the application field or workplace, and five were in more than one setting.

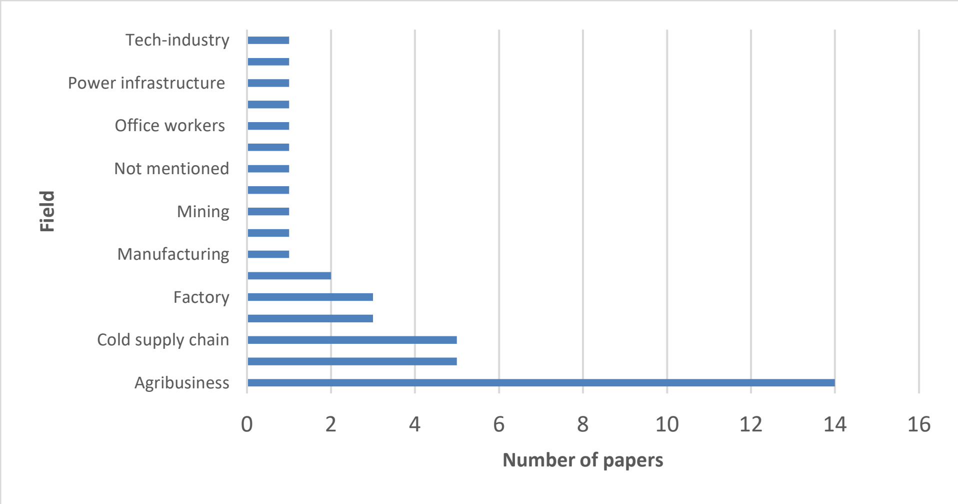

Figure 4. Workplaces identified in the review

4. DISCUSSION

This systematic review provides a comprehensive view of various applications of ML in OSH. To narrow down the ML application, authors have focused on the methods employing supervised machine learning approaches. In general, supervised techniques attempt to find a mathematical association between a set of inputs and one or more outputs. In other words, ML is used to find the function which maps the inputs to provided outputs. By considering the dataset features as inputs and the OSH variable as output (such as risk and stress, among others known as the label), the supervised ML approaches can represent a function that is later used to predict the output of unlabeled inputs. Such analysis can result in finding hidden associations between the causes of an event and its effects.

To the authors' knowledge, this is the first attempt to comprehensively explain the use of ML in OSH. In this regard, several studies that will be mentioned in the discussion part were investigated, focusing on methods of ML from classification (i.e., accident) to prediction (possible hazards) and as well as regression (to find patterns of accidents to prevent them). In the following section, the studies are categorized with a concentration on the application of ML for specific OSH applications. Following each section, the potential and limitations of the evaluated methods are discussed.

4.1 Accidents

To improve the safety and health of the industry's performance, it is mandatory to learn from past accidents effectively. However, while accident reports are typically unstructured or semi-structured textual reports, significant manual annotation before statistical analyses to prepare at an appropriately organized level, these analyses are critical for designing a pattern (or patterns) used directly to predict future accidents (Goh and Ubeynarayana 2017). Handling, keeping, and retrieving accident reports require huge processing efforts. ML is the missing chain between data storing and data processing. Several researchers have studied the occupational accident to classify accident narratives (Choi et al. 2020; Gerassis et al. 2017; Goh and Ubeynarayana 2017; Kang and Ryu 2019; Nanda et al. 2016; Sanmiquel, Rossell, and Vintró 2015; Sarkar et al. 2020; Sarkar, Vinay, et al. 2019a; Zhang 2019). Many ML algorithms, such as SVM, ANN, extreme learning machine (ELM), and DT, are used in occupational accident classification and prediction. Goh et al. (2017) proposed using an SVM model with unigram tokenization in the category of accident narratives. A dataset of 16,323 accidents from 4471 construction industries was classified to provide a training set. The data was manually annotated based on the Workplace Safety and Health Institute instructions. Considering the first occurrence of uncontrolled action as a first event, Goh et al. (2017) made labels including caught in/between objects, the collapse of an object, electrocution, exposure to chemical substances, exposure to extreme temperatures, falls, fires and explosion, struck by a falling object, struck by moving objects, traffic, etc. The evaluation of their trained model showed that employing linear SVM with unigram tokenization resulted in promising classification for automatically code accident narratives (Goh and Ubeynarayana 2017).

In another attempt, Sarkar et al. (2019a) trained two models using SVM and ANN techniques to compare them to predict the incident in construction field workplaces. Sarkar et al. (2019a) reported that SVM surpasses the other methods in predicting accidents correctly. Zhang et al. (2019) used the Word2Vec Skip-Gram model to extract keywords from an OSHA database of work accidents with the same trend. Following the extraction of keywords, they were used to classify the accident reasons via several methods (SVM, KNN, linear regression, decision trees, Naive Bayes, and DNN). The extracted keywords were used as features to identify the difference between each class. They also proved that a deep neural network (DNN) model could find a semantic relation between the reasons for several accidents and classify them into proper clusters (Zhang 2019).

Bayesian network is another method that is being used to build accident prediction or classification models in various industries such as mining (Nanda et al. 2016), construction (Gerassis et al. 2017), and steel plane (Sarkar et al. 2020). They have been evaluated as promising decision support systems to auto-code event information based on attributes. Additionally, it has been applied in many disciplines to quantify the specific causes of different types of accidents (Gerassis et al. 2017) or to find the hidden associations or highly probable interconnected reasons that contributed to a hazard (Nanda et al. 2016), or to predict patterns in accidents (Sarkar et al. 2020). For instance, (Gerassis et al. 2017), authors claimed they used 14 attributes to converge the cause of the accident. Based on their evaluations of the studied database, they reported that overexertion is ranked one (with 13% of total accidents), which has occurred mainly through falling (12%), detachment of tools from height, or losing objects (9%), and stepping unproperly on things.

In another research, Kang et al. (2019) employed Random Forests (RF) to identify the relation between the accidents and hidden reasons followed by a hazard and further classified them into proper clusters. For more details, they tried to find possible associations between causes (called features) that were not directly associated, such as the effects of weather on the falling hazard. They also stated in further investigations that they tend to perform real-time predictions. Since they used to train a specific model per database, one can criticize their work to be offline, in which, for each dataset, new training is required (Kang and Ryu 2019). With the same RF method, Choi et al. (2020) intended to classify the type of accident based on workers' parameters (age, gender, length of service, and kind of construction). They claimed the RF method provided the highest predictions to forecast possible hazards for individual workers. The two last-mentioned works showed that the RF is an independent data approach that can classify textual databases considering the histogram of accident keywords (Choi et al., 2020).

Sanmiquel et al. (2015) utilized a database of almost 70,000 occupational accidents and fatality reports from 2003–2012 in the Spanish mining sector to analyze the main reasons for the reported accidents. The study used supervised methods such as Bayesian classifiers, DTs, and contingency tables, among other data mining techniques. The results revealed that the most important causes of accidents include previous reasons, place, size, physical activity, preventive organization, experience, and age. In addition, the type of accident as a predictor variable and lost workdays as an output variable have served to measure the severity of the accidents. Sanmiquel et al. (2015) suggested that their study's outputs can be used to establish policy improvement to minimize the rate of occupational hazards in the mining sector (Sanmiquel, Rossell, and Vintró, 2015). The same results have been reported by Sarkar et al. (2020), that classified several unstructured accident narratives using the same ML techniques. However, the performance of K Means clustering approaches was highlighted. An overall evaluation of the studied strategies reveals that DT-based methods usually outperformed the other classifiers. From the ML point of view, it is observed that the features are mostly independent of each other. Therefore, a bag of decision trees, each with a separable node, can be constructed effectively (Sarkar et al., 2020). These trees are then used for either classification or prediction. However, the trained models are dependent on the training dataset. It means that most of the time, a model trained over a specific dataset, cannot be employed for the same classification purpose on another dataset. This drawback was directly reported in the work presented by Sarkar et al. (2020), where they could not transfer the learning model to another dataset. In final words, once these ML approaches are intended for classification and prediction, one should pay attention to the changes in the database. Prediction of return to work after sick leave can be considered an issue in which ML approaches are employed to find the correlation between the features indicating the severity of the accident and the number of days required for treatment. The study conducted by Na and Kim et al. (2019) applied a Gradient Boosting Method (GBM) in a DT to classify a text-based dataset of injuries and the number of days required before returning to work. The method seems straightforward in training; however, the trained model is overfitted on the dataset, making it impossible to apply any transfer learning to expand its transportability for other databases (Na and Kim 2019). Kakhki et al. (2019) employed several learning approaches for agribusiness industry workers to overcome such an issue. The result shows SV that SVM-based models represent proper evaluation for predicting injury severity without being overfitting. Although their study could compare several ML approaches in prediction accuracy, they did not present the feature importance, highlighting the ones having a more significant role in the feature selection process (Davoudi Kakhki, Freeman, and Mosher 2019).

4.2 Slip-trip-fall / falling

Falling from height is another leading cause of industry fatalities and injuries. However, ML approaches for predicting STF and analyzing the factors remained an unexplored area of research.

STF has been paid attention to in research by Sarkar et al. (2019a), where they developed a novel methodology using DTs classifiers, regression tree (CART), and RF for prediction. The proposed method was backed up by checking a DT generated over 20 interpretable safety decision rules explaining the factors behind the occurrences of STF (Sarkar, Raj, et al. 2019a). In another work, Chen et al. (2016) introduced an algorithm with the same purpose to provide accurate real-time risk evaluation. The algorithm could generate proactive warnings to alert workers when they are at risk in construction. To train the model, they recorded 1161 falling-related injuries from 2005 to 2015 that were extracted from the OSHA. Several ML approaches were intended to find an appropriate one based on prediction accuracy; K-modes, SVM (with RBF kernel), and DTs. The performance of the proposed models was evaluated using the OSHA injury record data. The result demonstrated that the DT-based falling risk prediction model surpassed other models in predicting true positive hazards. In addition, they provided the feature importance of the best-trained model to indicate the prominent accident features(or simply reasons), including the falling height, the worker's occupation, and the source of the falling (Chen and Luo 2016).

The work presented by Jeong et al. (2019) explains a method for fall detection using images (or videos, as a series of images). Notably, the authors have presented two proposals. As the main proposal, each worker's skeleton structure was extracted in the video to prepare data. Then, the required features were extracted by detecting the position of the legs and arms of each worker in the image. These positions were used as features through the detection of each overall position. As the data were prepared, as the second proposal, a Long short-term memory (LSTM) approach was employed to classify the person's position (Jeong et al., 2019). The main drawback of their proposed system is the number of individuals detected in each image. While Jeong et al. (2019) have not discussed it, LSTM can be trained by default to classify one person per image series.

An overall look at the introduced models shows that the same DTs still predict the SFT hazard better than others because of the features' nature. Nevertheless, STF can be categorized as an accident; hence, those approaches outperformed other ML methods and have shown better performance in STF detection.

4.3 Occupational physical activity and sedentary behavior

Work-related Musculoskeletal Disorders (WMSDs) are considered the third leading reason for disability and early retirement in many occupations. Sedentary behavior (SB) has been identified as a primary reason for WMSDs (Falck et al., 2018). Thus, measuring physical activity and SB in the workplace has been considered in several studies (Maheronnaghsh, Santos, and Vaz, 2018; Neuhaus et al., 2014; Quante et al., 2015). While several studies have characterized workers' postures using traditional measurement methods and questionnaires (Nussbaum et al. 2009; Ravnik, Otáhal, and Fikfak 2008; Yu, Li, Yang, et al. 2019), it is required to consider them together with workers' movement velocity in real-time and real workplaces (Maheronnaghsh et al., 2018). Several studies proposed accelerometer sensors to measure physical activity directly, like Kuster (2018), which used accelerometers on workers' bodies to collect their activities. The individual trait of the features collected from sensors highlights the application of ML techniques such as DTs to consider posture and the activity level of sitting and standing for the classification of safe or hazardous postures (Kuster et al. 2018).

Reviewing the research, the work proposed by Schall et al. (2016) intended to characterize Physical Activity (PA), including full-shift upper arm and trunk postures and movement velocities using inertial measurement units (IMUs) using a custom complementary weighting algorithm developed in MATLAB for nurses. Nurses' occupational, physical activity (PA) was monitored with classified using a waist-worn PA monitor and ML algorithms on raw acceleration data from each IMU. The result suggested that a combination of accelerometers can provide a more representative estimation of physical demands (Schall, Fethke, and Chen 2016). Unfortunately, their work was limited to evaluating the obtained results, and the methods' implementation remained unexplained. Hence based on the reviewed approaches, from the ML point of view, DT surpasses other algorithms.

4.4 Safety risk and risk assessment

Safety-leading indicators are a way to flag sites with higher risks. However, there is a lack of validated leading indicators to classify sites reliably according to safety risk levels. On the other hand, despite the success of ML approaches in other domains, it has not been widely utilized in the construction industry, especially in developing safety-leading indicators.

The research presented by Poh et al. (2018) suggested using five popular ML algorithms to train models to predict accident occurrence and severity. Training data was obtained from a large contractor in construction in the range of 2010 to 2016. About 13 input variables were selected to use a combination of the Boruta feature selection technique and DT. Poh et al. (2018) utilized the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework to find critical data types, including safety inspection records, accident cases, and project-related data. After validating all built models, RF was claimed to provide the best prediction performance. The predictive RF model, an ensemble of individual DTs, can be used as a leading safety indicator of the risk level of a site and can provide a monthly forecast of project safety performance. Besides, it assists in pre-emptive inspections and interventions to be implemented in a more targeted manner (Poh, Ubeynarayana, and Goh 2018).

In another study, Ajayi et al. (2019a) used a text-mining approach to retrieve meaningful terms from data. Six DL models were developed for OSH risk management in construction companies to predict accidents. The DL models include DNN classification (with binary classes: Risk or no risk), DNNreg1 (loss time), DNNreg2 (body injury), DNNreg3 (plant and fleet), DNNreg4 (equipment), and DNNreg5 (environment). An OSH risk database obtained from a leading United Kingdom power infrastructure construction company was used in developing the models using the H_2O framework. The performance of the trained models was assessed and benchmarked with existing models using test data and appropriate performance metrics. The results would guide practitioners to understand OSH challenges better, minimize project costs (such as third-party insurance and equipment repairs), and offer effective strategies to mitigate OSH risks (Ajayi, Oyedele, Davila Delgado, et al. 2019). As a limitation of the work, the approach of DNN demands a large amount of training data. At the same time, it takes a longer time (compared to other ML approaches) for training (Ajayi et al., 2019).

A similar study conducted on power infrastructure (Ajayi et al., 2019) focused on using big data frameworks to manage construction accidents and analyze and predict accidents in power infrastructure. Ajayi et al. (2019) predicted the likelihood of health hazard occurrence using objective data from a UK power infrastructure company. The proposed architecture identified relevant variables and improved preliminary prediction accuracy and explanatory capacities. It has also enabled conclusions to be drawn regarding the causes of health risks. Ajayi et al.(2019) suggested that big data technology could be used to find complex patterns and establishing the statistical cohesion of hidden patterns for optimal future decision-making (Ajayi et al., 2019).

The study conducted by Tixier et al. (2016) proposed injury prediction and employed two tree-based approaches (RF and Stochastic Gradient Tree Boosting). Energy-based data were incorporated into the predictive models to test whether skill affects the level of injury severity. Nevertheless, as a limitation of their work, it can be said they could attribute outcome data extracted from injury reports instead of current predictions based on the occurrence of an accident. (Tixier et al.2016).

4.5 Workers' physical/physiological demand

There are methods to investigate the physical demands of various tasks. Nevertheless, they are mostly limited to detecting only one individual's target (e.g., physiological characteristics) and environmental conditions (e.g., ambient temperature and humidity). Jebelli et al. (2019) developed a procedure to automatically predict physical demand levels based on physiological signals collected with biosensors from workers regularly working in the workplace. They fed the collected data to obtain an Energy-Expenditure Prediction Program (EEPP) as a baseline to classify tasks into low, moderate, and high-intensity activities. In this regard, they benefited from a supervised machine-learning model trained by SVM with a Gaussian kernel. They showed promising results indicating that this method can improve construction workers' productivity, safety, and general well-being by detecting highly physically demanding tasks in the field (Jebelli, Choi, and Lee, 2019).

In another study, Xie et al. (2019) evaluated the life status of workers in a complex environment in real time with a wearable safety assurance system. They tried to detect the risk of safety accidents caused by abnormal physical conditions in the process proactively. They monitored the physiological parameters such as heart rate, body temperature, and blood pressure and classified them with the SVM. Using wearable sensors combined with the worker's location and weather forecasts provides a framework to produce intelligent pre-warning safety hazards that will be possible (Xie and Chang 2019).

Unlike the previous topics where DT was evaluated to outperform other approaches, the studies have shown that SVM was assessed for more accurate predictions. From the ML point of view, when the training features cannot be classified with linear approaches, the dimension mapping technique, common in SVM, can reveal the hidden associations between the inputs.

4.6 Stress prediction

The study carried out by Reddy et al. (2018) focused on analyzing the stress patterns of IT workers with ML methods. Factors obtained from a survey of IT workers in 2017 were evaluated. Multiple features were selected for constructing a DT, and the most prominent factors affecting stress were assessed. DT technique could identify major factors (attributes) influencing the workers' stress (e.g., gender, family history, and availability of health benefits in the workplace). The obtained results can directly assist industries in narrowing their approach to eliminating stress factors and creating a convenient workplace for their employees (Reddy, Thota, and Dharun 2018).

4.7 Noise

Zhao et al. (2019) demonstrated the feasibility of developing machine learning models to predict hearing impairment exposed to complex non-Gaussian industrial noise. For this purpose, a database of several factory noises and the hearing loss of 1,113 subjects was collected. Four potential ML models were evaluated based on SVM, multilayer perceptron, adaptive boosting, and RF. In conclusion, the result showed that SVM outperformed the others in evaluation scores due to the noise source (Zhao et al. 2019).

4.8 Safety and health behavior and human error

Some investigations have been dedicated to improving safety at the workplace by using prediction techniques that can be used to model and formulate each worker's role and the physical characteristics of the job site that may impose incidents. Rashid et al. (2018) studied the feasibility of two trajectory prediction models, using polynomial regression (PR) and the hidden Markov model (HMM) calibrated by factoring in worker's attitudes toward safety, which is a measure of their tendency to or aversion to risky behavior near hazards. Automatic detection of workers' unsafe actions assists in monitoring workers' behavior to improve safety and health at work. It also enables the proactive prevention of an accident by reducing the number of unsafe actions. Rashid et al. (2018) analyzed workers' overall behavior toward risk and aimed at formulating and factoring their attitude instead of separately formulating the effect of each behavioral parameter (e.g., age, gender, and level of experience) to calibrate the output of trajectory prediction. The results demonstrated that HMM could reliably build models to detect unsafe movements and impending collision events (Rashid and Behzadan n.d.). Also, the work carried out by Han et al. (2014) presented a modeling and classification methodology for the recognition of unsafe actions of workers by monitoring their behavior. A one-class SVM was applied with a Gaussian kernel for the classification of action recognition. This method was intended to proactively prevent an accident by reducing the number of unsafe actions.

Studies have shown that head gestures and brain activity reflect human behaviors related to the risk of accidents when using machine tools (Lin et al. 2010; Liu et al. 2009). Instead of a camera, the research carried out by Li et al. (2014) developed a non-invasive Smart Safety Helmet system able to recognize abnormal behaviors of workers to track the head gestures and brain activity, which endanger safety and health. Li et al. (2014) estimated the risk level with electroencephalography(EEG) and the IMU. An AI algorithm was then used to identify the risk level of health issues such as bad posture, accident risk, fatigue, sleepiness, high stress, etc. The risk level was computed based on previous experience and surrounding tools or processes. As in the aerospace industry, risk evaluation using real-time algorithms considers three parameters: the probability of occurrence, the severity of the mishap, and exposure (Li et al. n.d.). Investigations use single sensors, such as cameras, to detect safety hazards. The study by Nath et al. (2020) utilized only images for training a model that detects personal protective equipment (PPE) in workplaces. They benefit from an AI segmentation method based on YOLO (You Only Look Once) network to detect whether an individual wears PPEs. Using three approaches, they evaluated the best model to detect workplace hazards; detect individual PPEs, and decide based on an ML classifier (NN and DT); detect and classify the hazards at once; and detect the whole equipment at once and decide via a ML classifier. They finally reported that the second approach was more accurate for detecting PPE attire with almost 72.3% accuracy. They claimed although the other two approaches were more precise in seeing the equipment since the ML classifier is detached from the detection part, they might be propagation of error that influences the final decision. Their study showed that combined detection and classification of PPE is more efficient in real-time computation. Motion sensors have facilitated workers' real-time monitoring in construction zones(Nath, Behzadan, and Paal 2020). The study intended by Kim et al. (2019) focused on a simulation-based evaluation of multiple motion sensors attached to workers performing typical construction tasks. They employed inertial measurement unit (IMU) sensors (total of 17) to collect motion sensor data from an entire body and get proper readings for generating training data (Kim, Chen, and Cho n.d.). Among the methods they used, RF showed better performance predicting activities that may result in injuries. The final evaluation of their approach shows that if they could have increased the database size, more accurate results could have been obtained. Additionally, clustering of motions in groups has been shown to be more error-prone compared to individual activity prediction. Finally, one can criticize their approach by selecting the feature, as they could have better results if they changed the formation of feature vectors.

4.9 Environmental safety

In the study of Kanazawa et al. (2019), the authors used a Gaussian Mixture Regression (GMR) to predict the worker's trajectory to plan the robot trajectory in the expanded temporal space to decrease collision risk and reduce the latency required for the robot's movement. The proposed system could deal with the worker's regular and irregular motion since it was equipped with an online trajectory generator according to the probabilistic prediction of the worker's motion to justify the robot's trajectory. They evaluated their proposed system experimentally by applying it to an assembly scenario with the two-link planar manipulator. They claimed that the evaluation resulted in confirmation from several workers that the proposed method could successfully enhance work-time efficiency and the worker's safety.

Wang et al. (2019) targeted a mining application. Multi-sensor real-time online monitoring, with pattern recognition and warning systems, was used to provide input data for training a model via NNs. The goaves' three-dimensional stress contour map was drawn, making the stress changes more intuitive and accurate. To facilitate the acquisition of the data sequence of the sensors, a sliding time window was used for sampling; however, to summarize the input data, they proposed to use an efficient, high-precision NN algorithm. More importantly, it can provide short-time warning information about stability and be a scientific basis for predicting geological disasters caused by goaves (Wang, Zheng, and Wang, 2019).

4.10 Ergonomic risks and workload

Considering the diverse and dynamic nature of worker's activities, it isn't easy to unobtrusively collect worker behavior for analysis. To address the issue raised by this impossibility, Yu et al. (2019a) aimed to develop a workload assessment method suitable for the complex nature of construction activities by using biomechanical analysis. They used a model trained with DTs, although the technical parameters and features have not been disclosed, as it was not directly mentioned. The process body shape to have workers pose and decide about joint positions and workload. The method was accomplished with a new data collection approach that can benefit various behavior research fields related to construction safety and help prevent construction workers' ergonomic risks (Yu, Li, Umer, et al., 2019). In a similar study, Meyers et al. (2018) leveraged a state workers' compensation claims database and AI techniques to target injury causation and industry prevention efforts. They coded more than 1.2 million Ohio Bureau of Workers' Compensation for Injury with causation auto-coding methods between 200 industry groups. Workers were ranked according to the soft-tissue musculoskeletal. Finally, they claimed that biomechanical ergonomic (ERGO) or STF interventions could have prevented the injuries. These findings were being used to focus on the prevention resources for specific occupational injury types (Meyers et al. 2018).

Among all the learning approaches employed in work by Sasikumar et al. (2020), they claimed that RF had surpassed others (DT, kNN, ANN, SVM) when it comes to developing a model for predicting the risk of musculoskeletal disorders targeting office workers, especially computer professionals. Their method to construct the training data consisted of a questionnaire and a camera to take random photos while working. However, since the camera cannot represent an accurate 3D model of the body, one can criticize using depth sensors such as Kinect to obtain more reliable readings (Sasikumar and Binoosh 2020).

The other notable attempt has been proposed in the study. Yin et al. (2019) employed a DL method based on a convolutional NN to decompose the EEG signals to understand workers' mental workload. They have used a new transfer dynamical autoencoder (TDAE) to capture EEG signals' dynamical properties and extract valuable features per individual. Then they evaluated and attempted to remove common similarities between the extracted features. While traditional learning methods cannot find the association is the features per individual, the DL approaches have been evaluated as suitable. Yin et al. (2019) showed the potential of DL approaches in decomposing the hidden associations in signals (or features), which traditional methods are unable to detect (Yin et al. 2019).

Not only 1D signals but also 2D features (such as images) can be employed for workload prediction. As an instance, the study followed by the work of Asadi et al. (2020) showed that employing computer vision and machine learning techniques to present an objective and automated approach to predict the force exertions via facial videos and wearable photoplethysmogram (PPG) as input sources can be a trusted method in classifying isometric grip force exertion levels (Asadi et al. 2020). They equipped the workers with 2D and 1D wearable sensors, and similar to the previous study, they employed a DL method for training and production. Their approach of using 2D sensors (images) has been evaluated as not versatile enough, as they stated the facial expression detection model might not work expectedly for older adults. Other studies tried to predict ergonomic risks based on the known attributes, though avoiding giving more details on which features they have employed. An example is a work carried out by Dansie et al. (2013) to train an expert system (AI model) to build a risk assessment tool. They used a Bayesian network to find a correlation between the risk score and the type of job, aiming to improve job conditions; however, they failed to clarify which attributes (features) they had considered through the training (Dansie, Sesek, and Bloswick n.d.).

4.11 Fatigue

Worker fatigue is a highly prevalent phenomenon worldwide. Generally, fatigue happens without adequate rest after either forceful, repetitiveness, prolonged exertions or high stresses on the body (Yu, Li, Yang, et al., 2019).

Fatigue and incomplete recovery can reduce the work capacity, and while it impresses decrease work efficiency negatively, it can increase the risk of injury, accidents, even and death. Baghdadi et al. (2018) tried to provide a practical framework for predicting realistic fatigue levels and proposed a method for classifying non-fatigued compared with fatigued states in manual material handling. The researchers examined using a wearable accelerometer sensor to acquire input data and SVM to classify fatigue-related changes in gait based on a simulated manual material handling task, based on foot acceleration and position trajectories with 90% accuracy. The examined method provided a practical framework for predicting realistic fatigue levels (Baghdadi et al. 2018).

Traditionally, fatigue was monitored with self-reporting or subjective questionnaires, but Yu et al. (2019b) proposed a fatigue assessment model that uses DL algorithms to perform biomechanical analysis to build a physical fatigue model which can classify the physical fatigue level of different construction task conditions (Yu, Li, Yang, et al. 2019b). Considering the wireless sensors, the non-invasively collection of multiple physiological readings in real-time is more available and cheaper. In research led by Umer et al. (2020), they combined cardiorespiratory and thermoregulatory measures to assess real-time physical exertion levels accurately for workers. They analyzed several ML approaches to predict accurate physical exertion modeling from the collected data: KNN, SVM, Discriminant analyses, DTs, and Ensemble classifiers. Evaluations of the constructed dataset showed that the decision tree approaches could obtain accuracy as high as 95%. SVM-based learning was reported to accomplish up to 91% classification accuracy at the next level (Umer et al., 2020).

The authors analyzed multiple physiological measures, including cardiorespiratory and thermoregulatory, to predict physical exertion during manual work. The results highlighted the importance of combining multiple measures to predict exertion levels. Their results claimed the individual measures were almost unable to report excess exertion levels.

4.12 Machinery accidents

Jocelyn et al. (2016) proposed a method benefiting from dynamic risk identification and Logical Analysis of Data (LAD) to evaluate the risk in machinery safety. Their results showed that the second proposal, LAD, is highlighted as an application of AI to extract information from accident reports to analyze machinery-related accidents in the workplace. They also evaluated their proposal, which led to decreased machinery accidents, and kept and updated the accident reports (Jocelyn, Chinniah, and Ouali 2016).

In another work, a practical approach was designed by Marucci-Wellman et al. (2017) to classify injury narratives of Bureau of Labor Statistics Occupational Injury and Illness events for a large workers' compensation database. To type narrative text, they analyzed four machine-learning methods (Naïve Bayes, Single word and Bi-gram models, SVM, and Logistic Regression). The result indicated that human-machine ensemble methods will likely improve performance over manual coding when finding patterns that lead to accidents with injury consequences (Marucci-Wellman, Corns, and Lehto 2017).

5. CONCLUSIONS

This study found several methods of ML can be applied across several fields of occupational health and safety for monitoring, classifying (injuries, accidents, monitoring data), predicting (hazards), and decision-making to improve the safety and health of workers at work. Construction and mining are the leaders in what comes to applications. Decision-makers and workers are taking advantage of various techniques of ML to find solutions for health and safety problems at work, and OSH experts have access to more precise data in real time that can be used instead of traditional methods. The results demonstrated that the successful implementation and adoption of ML-based applications could reduce warning time compared to conventional methods, and OSH experts have faster access to data using real-time and real-work data. Nevertheless, utilizing ML applications may require the availability of OSH incident data and an improved understanding of the output. In addition, in data preprocessing tasks, a lot of manual effort is needed to clean the data suitable for analysis. ML techniques have the potential to be utilized in other fields of OSH. However, it is necessary that in future investigations, the limitations of using ML applications in the OSH area be improved and its full potential be achieved.

CONFLICT OF INTEREST

No conflict of interest is declared.

REFERENCES

Ajayi, Anuoluwapo, Lukumon Oyedele, Juan Manuel Davila Delgado, et al. 2019. "Big Data Platform for Health and Safety Accident Prediction." World Journal of Science, Technology and Sustainable Development 16(1): 2–21. https://doi.org/10.1108/WJSTSD-05-2018-0042

Ajayi, Anuoluwapo, Lukumon Oyedele, Hakeem Owolabi, et al. 2019. "Deep Learning Models for Health and Safety Risk Prediction in Power Infrastructure Projects." Risk Analysis. https://doi.org/10.1111/risa.13425

Asadi, H et al. 2020. "A Computer Vision Approach for Classifying Isometric Grip Force Exertion Levels." Ergonomics 63(8): 1010–26. 10.1080/00140139.2020.1745898

Badri, Adel, Bryan Boudreau-Trudel, and Ahmed Saâdeddine Souissi. 2018. "Occupational Health and Safety in the Industry 4.0 Era: A Cause for Major Concern?" Safety Science 109: 403–11. https://doi.org/10.1016/j.ssci.2018.06.012

Baghdadi, Amir et al. 2018. "A Machine Learning Approach to Detect Changes in Gait Parameters Following a Fatiguing Occupational Task." Ergonomics 0139: 1–14. http://doi.org/10.1080/00140139.2018.1442936

Chen, H, and X Luo. 2016. "Severity Prediction Models of Falling Risk for Workers at Height." In 5th Creative Construction Conference, CCC 2016, eds. Skibniewski M.J. and Hajdu M. Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, Hong Kong: Elsevier Ltd, 439–45. https://doi.org/10.1016/j.proeng.2016.11.642

Cho, C, J Park, K Kim, and S Sakhakarmi. 2018. "Machine Learning for Assessing Real-Time Safety Conditions of Scaffolds." In 35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things, ISARC 2018, Department of Civil and Environmental Engineering and Construction, University of Nevada, Las Vegas, United States: International Association for Automation and Robotics in Construction IAARC). https://doi.org/10.22260/ISARC2018/0008

Choi, J, B Gu, S Chin, and J.-S. Lee. 2020. "Machine Learning Predictive Model Based on National Data for Fatal Accidents of Construction Workers." Automation in Construction 110. https://doi.org/10.1016/j.autcon.2019.102974

Dansie, C, R Sesek, and D S Bloswick. "Ergonomic Risk Assessment Based on a Bayesian-Optimized Expert System." Int. J. Hum. Factors Model. Simul. (Switzerland) 4(1): 23–34. http://dx.doi.org/10.1504/IJHFMS.2013.055782

Davoudi Kakhki, F, S A Freeman, and G A Mosher. 2019. "Evaluating Machine Learning Performance in Predicting Injury Severity in Agribusiness Industries." Safety Science 117: 257–62. https://doi.org/10.1016/j.ssci.2019.04.026

Falck, Ryan S et al. 2018. "Can We Improve Cognitive Function among Adults with Osteoarthritis by Increasing Moderate-to-Vigorous Physical Activity and Reducing Sedentary Behaviour? Secondary Analysis of the MONITOR-OA Study." BMC Musculoskeletal Disorders 19(1): 447. https://doi.org/10.1186/s12891-018-2369-z

Gerassis, S et al. 2017. "Bayesian Decision Tool for the Analysis of Occupational Accidents in the Construction of Embankments." Journal of Construction Engineering and Management 143(2). https://doi.org/10.1061/(ASCE)CO.1943-7862.0001225

Goh, Y M, and C U Ubeynarayana. 2017. "Construction Accident Narrative Classification: An Evaluation of Text Mining Techniques." Accident Analysis and Prevention 108: 122–30. https://doi.org/10.1016/j.aap.2017.08.026

Han, S, S Lee, and F Peña-Mora. 2014. "Comparative Study of Motion Features for Similarity-Based Modeling and Classification of Unsafe Actions in Construction." Journal of Computing in Civil Engineering 28(5). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000339

Higgins, Julian P T et al. 2011. "The Cochrane Collaboration{\textquoteright}s Tool for Assessing Risk of Bias in Randomised Trials." BMJ 343. https://doi.org/10.1136/bmj.d5928

Jebelli, H, B Choi, and S Lee. 2019. "Application of Wearable Biosensors to Construction Sites. II: Assessing Workers' Physical Demand." Journal of Construction Engineering and Management 145(12). https://doi.org/10.1061/(ASCE)CO.1943-7862.0001710

Jeong, S, S Kang, and I Chun. 2019. "Human-Skeleton Based Fall-Detection Method Using LSTM for Manufacturing Industries." In 34th International Technical Conference on Circuits/Systems, Computers and Communications, ITC-CSCC 2019, Department of Information and Communication Technology, University of Science and Technology, South Korea: Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ITC-CSCC.2019.8793342

Jin, Ruoyu et al. 2019. "A Science Mapping Approach Based Review of Construction Safety Research." Safety Science 113: 285–97. https://doi.org/10.1016/j.ssci.2018.12.006

Jocelyn, S, Y Chinniah, and M.-S. Ouali. 2016. "Contribution of Dynamic Experience Feedback to the Quantitative Estimation of Risks for Preventing Accidents: A Proposed Methodology for Machinery Safety." Safety Science 88: 64–75. https://doi.org/10.1016/j.ssci.2016.04.024

Kanazawa, A, J Kinugawa, and K Kosuge. 2019. "Adaptive Motion Planning for a Collaborative Robot Based on Prediction Uncertainty to Enhance Human Safety and Work Efficiency." IEEE Trans. Robot. (USA) 35(4): 817-- 32. https://doi.org/10.1109/TRO.2019.2911800

Kang, K, and H Ryu. 2019. "Predicting Types of Occupational Accidents at Construction Sites in Korea Using Random Forest Model." Safety Science 120: 226–36. https://doi.org/10.1016/j.ssci.2019.06.034

Kim, K, Jingdao Chen, and Y K Cho. "Evaluation of Machine Learning Algorithms for Worker's Motion Recognition Using Motion Sensors." In Computing in Civil Engineering 2019. Data, Sensing, and Analytics, Reston, VA, USA, 51–58. https://doi.org/10.1061/9780784482438.007

Kuster, R P et al. 2018. "Measuring Sedentary Behavior by Means of Muscular Activity and Accelerometry." Sensors (Switzerland) 18(11). https://doi.org/10.3390/s18114010

Li, Ping et al. "A Smart Safety Helmet Using IMU and EEG Sensors for Worker Fatigue Detection." In 2014 IEEE International Symposium on Robotic and Sensors Environments (ROSE). Proceedings, Piscataway, NJ, USA, 55–60. https://doi.org/10.1109/ROSE.2014.6952983

Lin, Chin-Teng et al. 2010. "A Real-Time Wireless Brain-Computer Interface System for Drowsiness Detection." IEEE transactions on biomedical circuits and systems 4(4): 214–22. https://doi.org/10.1109/TBCAS.2010.2046415

Liu, Jiayang, Lin Zhong, Jehan Wickramasuriya, and Venu Vasudevan. 2009. "UWave: Accelerometer-Based Personalized Gesture Recognition and Its Applications." Pervasive and Mobile Computing 5(6): 657–75. https://doi.org/10.1109/PERCOM.2009.4912759

Maheronnaghsh, S., Santos.J., Torres Marques.A., & Vaz.M. 2018. "Methods for Measuring Association between Intervention for Increasing Movement and Productivity: Systematic Review." U.Porto Journal of Engineering 4(1): 46–60. https://doi.org/10.24840/2183-6493_004.001_0003

Maheronnaghsh, S, H Zolfagharnasab, M Gorgich, and J Duarte. 2021. "Machine Learning in Occupational Safety and Health : Protocol for a Systematic Review." 1(Ml): 32–38. https://doi.org/10.24840/2184-0954_005.001_0004

Maheronnaghsh, Sara, Joana Santos, and M. Vaz. 2018. "Methods of Posture Analysis for Computer Workers." International Journal of Occupational and Environmental Safety 2(2): 41–49. https://doi.org/10.24840/2184-0954_002.002_0005

Marucci-Wellman, H R, H L Corns, and M R Lehto. 2017. "Classifying Injury Narratives of Large Administrative Databases for Surveillance—A Practical Approach Combining Machine Learning Ensembles and Human Review." Accident Analysis and Prevention 98: 359–71. https://doi.org/10.1016/j.aap.2016.10.014

Meyers, A R et al. 2018. "Applying Machine Learning to Workers' Compensation Data to Identify Industry-Specific Ergonomic and Safety Prevention Priorities: Ohio, 2001 to 2011." Journal of Occupational and Environmental Medicine 60(1): 55–73. https://doi.org/10.1097/jom.0000000000001162

Moher, David et al. 2009. "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement." Annals of Internal Medicine 151(4): 264–69. https://doi.org/10.1136/bmj.b2535

Na, K.-S., and E Kim. 2019. "A Machine Learning-Based Predictive Model of Return to Work after Sick Leave." Journal of Occupational and Environmental Medicine 61(5): E191–E199. https://doi.org/10.1097/jom.0000000000001567

Nanda, G et al. 2016. "Bayesian Decision Support for Coding Occupational Injury Data." Journal of Safety Research 57: 71–82. https://doi.org/10.1016/j.jsr.2016.03.001

Nath, Nipun D, Amir H Behzadan, and Stephanie G Paal. 2020. "Deep Learning for Site Safety: Real-Time Detection of Personal Protective Equipment." Automation in Construction 112: 103085. https://doi.org/10.1016/j.autcon.2020.103085

Neuhaus, M et al. 2014. "Reducing Occupational Sedentary Time: A Systematic Review and Meta-Analysis of Evidence on Activity-Permissive Workstations." Obes Rev 15. https://doi.org/10.1111/obr.12201

Nussbaum, M A et al. 2009. "Development of a Decision Support System for Residential Construction Using Panellised Walls: Approach and Preliminary Results." Ergonomics 52(1): 87–103. https://doi.org/10.1080/00140130802480869

Oliveira, M D, D F Lopes, and C A Bana e Costa. 2018. "Improving Occupational Health and Safety Risk Evaluation through Decision Analysis." International Transactions in Operational Research 25(1): 375–403. https://doi.org/10.1111/itor.12339

Page, Matthew J et al. 2021. "PRISMA 2020 Explanation and Elaboration: Updated Guidance and Exemplars for Reporting Systematic Reviews." BMJ 372. https://doi.org/10.1136/bmj.n160

Poh, C Q X, C U Ubeynarayana, and Y M Goh. 2018. "Safety Leading Indicators for Construction Sites: A Machine Learning Approach." Automation in Construction 93: 375–86. https://doi.org/10.1016/j.autcon.2018.03.022

Quante, Mirja et al. 2015. "Practical Considerations in Using Accelerometers to Assess Physical Activity, Sedentary Behavior, and Sleep." Sleep Health 1(4): 275–84. https://doi.org/10.1016/j.sleh.2015.09.002

Rashid, K M, and A H Behzadan. "Risk Behavior-Based Trajectory Prediction for Construction Site Safety Monitoring." J. Constr. Eng. Manag. (USA) 144(2): 04017106 (10 pp.)----. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001420

Ravnik, D.a b d, S.a Otáhal, and M.D.c Fikfak. 2008. "Using Different Methods to Assess the Discomfort during Car Driving." Collegium Antropologicum 32(1): 267–76. PMID: 18494213

Reddy, U S, A V Thota, and A Dharun. 2018. "Machine Learning Techniques for Stress Prediction in Working Employees." In 9th IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2018, eds. Krishnan N. and Karthikeyan M. Department of Computer Applications, Machine Learning and Data Analytics Lab, National Institute of Technology, Trichy, India: Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICCIC.2018.8782395

Sanmiquel, L, J M Rossell, and C Vintró. 2015. "Study of Spanish Mining Accidents Using Data Mining Techniques." Safety Science 75: 49–55. https://doi.org/10.1016/j.ssci.2015.01.016

Sarkar, S, S Vinay, et al. 2019a. "Application of Optimized Machine Learning Techniques for Prediction of Occupational Accidents." Computers and Operations Research 106: 210–24. https://doi.org/10.1016/j.cor.2018.02.021

Sarkar, S, R Raj, et al. 2019b. "An Optimization-Based Decision Tree Approach for Predicting Slip-Trip-Fall Accidents at Work." Safety Science 118: 57–69. https://doi.org/10.1016/j.ssci.2019.05.009

Sarkar, S, A Pramanik, J Maiti, and G Reniers. 2020. "Predicting and Analyzing Injury Severity: A Machine Learning-Based Approach Using Class-Imbalanced Proactive and Reactive Data." Safety Science 125. https://doi.org/10.1016/j.ssci.2020.104616

Sasikumar, V, and S.C.A.B. Binoosh. 2020. "A Model for Predicting the Risk of Musculoskeletal Disorders among Computer Professionals." International Journal of Occupational Safety and Ergonomics 26(2): 384–96. https://doi.org/10.1080/10803548.2018.1480583

Schall, Mark C., Nathan B. Fethke, and Howard Chen. 2016. "Working Postures and Physical Activity among Registered Nurses." Applied Ergonomics 54: 243–50. http://dx.doi.org/10.1016/j.apergo.2016.01.008

Shamseer, Larissa et al. 2015. "Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (Prisma-p) 2015: Elaboration and Explanation." BMJ (Online) 349. https://doi.org/10.1136/bmj.g7647

Suarez, A, and P Riesgo. 2005. "Information Technology for Occupational Health and Safety Audit." In 2005 International Conference on Artificial Intelligence, ICAI'05, Department of Business Administration, University of Oviedo - Escuela de Minas de Oviedo, C/Independencia, 13, 33004 Oviedo, Spain, 469–74. https://dblp.org/rec/conf/icai/SuarezR05.bib

Tixier, A.J.-P., M R Hallowell, B Rajagopalan, and D Bowman. "Application of Machine Learning to Construction Injury Prediction." Autom. Constr. (Netherlands) 69: 102–14. https://doi.org/10.1016/j.autcon.2016.05.016

Umer, W et al. 2020. "Physical Exertion Modeling for Construction Tasks Using Combined Cardiorespiratory and Thermoregulatory Measures." Automation in Construction 112. https://doi.org/10.1016/j.autcon.2020.103079

Wang, Yunsen, Guiping Zheng, and Xin Wang. 2019. "Development and Application of a Goaf-Safety Monitoring System Using Multi-Sensor Information Fusion." Tunnelling and Underground Space Technology 94: 103112. https://doi.org/10.1016/j.tust.2019.103112

Weng, Wei-Hung. 2019. "Machine Learning for Clinical Predictive Analytics." Leveraging Data Science for Global Health 199–217. https://doi.org/10.1007/978-3-030-47994-7_12

vWohlin, Claes. 2014. "Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering." In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, EASE '14, New York, NY, USA: Association for Computing Machinery 38: 1-10. https://doi.org/10.1145/2601248.2601268

Xie, X, and Z Chang. 2019. "Intelligent Wearable Occupational Health Safety Assurance System of Power Operation." Journal of Medical Systems 43(1). https://doi.org/10.1007/s10916-018-1122-3

Yin, Z et al. 2019. "Physiological-Signal-Based Mental Workload Estimation via Transfer Dynamical Autoencoders in a Deep Learning Framework." Neurocomputing 347: 212–29. https://doi.org/10.1016/j.neucom.2019.02.061

Yu, Y, H Li, X Yang, et al. 2019. "An Automatic and Non-Invasive Physical Fatigue Assessment Method for Construction Workers." Automation in Construction 103: 1–12. https://doi.org/10.1016/j.autcon.2019.02.020

Yu, Y, H Li, W Umer, et al. 2019. "Automatic Biomechanical Workload Estimation for Construction Workers by Computer Vision and Smart Insoles." Journal of Computing in Civil Engineering 33(3). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000827

Zhang, Fan. 2019. "A Hybrid Structured Deep Neural Network with Word2Vec for Construction Accident Causes Classification." International Journal of Construction Management: 1120-1140 https://doi.org/10.1080/15623599.2019.1683692

Zhao, Yanxia et al. 2019. "Machine Learning Models for the Hearing Impairment Prediction in Workers Exposed to Complex Industrial Noise: A Pilot Study." Ear and hearing 40(3): 690–99. https://doi.org/10.1097/aud.0000000000000649

Appendix 1

Table 1. Risk of bias per study

Paper	properness of the classification method	usability of method	the database used for training	sensors properness	reporting quality	references quality
(Wang, Zheng, and Wang 2019)	+	+	?	+	+	?
(Gerassis et al. 2017)	+	+	?	NA	+	+
(Sanmiquel, Rossell, and Vintró 2015)	+	+	?	NA	+	+
(Nanda et al. 2016)	+	+	?	NA	+	?
(Na and Kim 2019)	+	-	+	NA	+	+
(Yu, Li, Umer, et al. 2019)	+	+	?	+	+	+
(Jebelli, Choi, and Lee 2019)	+	+	+	-	-	+
(Choi et al. 2020)	+	+	+	NA	?	+
(Kang and Ryu 2019)	+	+	+	NA	+	+
(Han, Lee, and Peña-Mora 2014)	+	+	-	-	+	+
(Baghdadi et al. 2018)	+	+	+	-	+	+
(Yu, Li, Yang, et al. 2019)	-	-	-	-	-	+
(Zhang 2019)	+	+	-	?	+	+
(Zhao et al. 2019)	-	+	+	+	+	+
(Nath, Behzadan, and Paal 2020)	+	+	+	+	+	+
(Yin et al. 2019)	?	+	+	+	+	+
(Davoudi Kakhki, Freeman, and Mosher 2019)	+	+	+	?	+	+
(Dansie, Sesek, and Bloswick n.d.)	-	+	-	?	+	+
(Tixier et al. n.d.)	?	+	+	?	-	+
(Kim, Chen, and Cho n.d.)	+	+	-	+	+	+
(Asadi et al. 2020)	+	+	?	?	+	+
(Sasikumar and Binoosh 2020)	+	+	+	-	+	?
(Umer et al. 2020)	+	+	+	+	+	+
(Sarkar et al. 2020)	+	-	-	?	+	+
(Sarkar, Vinay, et al. 2019)	+	+	+	NA	+	+
(Goh and Ubeynarayana 2017)	+	+	+	NA	+	+
(Meyers et al. 2018)	+	+	-	?	-	+
(Chen and Luo 2016)	+	+	+	-	+	+
(Jeong, Kang, and Chun 2019)	+	+	+	+	+	+
(Kuster et al. 2018)	+	+	-	+	?	+
(Schall, Fethke, and Chen 2016)	+	+	?	+	+	+

(Poh, Ubeynarayana, and Goh 2018)	+	+	+	NA	+	+
(Reddy, Thota, and Dharun 2018)	+	+	+	NA	+	+
(36)	+	+	+	NA	+	+
(Xie and Chang 2019)	+	?	+	+	+	+
(Ajayi, Oyedele, Owolabi, et al. 2019)	+	+	-	-	+	+
(Rashid and Behzadan n.d.)	+	?	+	+	+	+
(Kanazawa, Kinugawa, and Kosuge 2019)	+	+	?	+	+	+
(Li et al. 2014)	+	+	+	NA	+	+
(Jocelyn, Chinniah, and Ouali 2016)	+	+	+	NA	+	+
(Sarkar, Raj, et al. 2019b)	-	+	-	NA	+	+
(Marucci-Wellman, Corns, and Lehto 2017)	+	+	+	NA	+	+

\+ Low risk

\- High Risk

? Unclear

NA=Not applicable