

## Machine learning in occupational safety and health: protocol for a systematic review

S. Maheronnaghsh<sup>a</sup>, H. Zolfagharnasab<sup>b</sup>, M. Gorgich<sup>c</sup>, J. Duarte<sup>d</sup>

<sup>a</sup>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, <sup>b</sup>FARO Technologies, PT (hooshiar.h.z@ieee.org), ORCID 0000-0001-8635-9169, <sup>c</sup>Faculty of Engineering, University of Porto, PT (up201306568@fe.up.pt), ORCID 0000-0002-1454-5752, <sup>d</sup>Associated Laboratory for Energy, Transports and Aeronautics – LAETA (PROA), Faculty of Engineering, University of Porto, PT (jadsuarte@fe.up.pt), ORCID 0000-0002-5856-5317

### Article History

Received January 17, 2020  
Accepted April 26, 2021  
Published April 30, 2021

### Keywords

Artificial Intelligence  
Sensor  
Prediction  
Intelligent decision support systems  
Protocol  
Systematic review

### DOI:

10.24840/2184-0954\_005.001.0004

### ISSN:

2184-0954

### Type:

Protocol



Open Access



Peer-Reviewed



CC BY

### Abstract

Industry 4.0 has shaped the way people look at the world and interact with it, especially concerning the work environment with respect to occupational safety and health (OSH). Machine learning (ML), as a branch of Artificial Intelligence (AI), can be effectively used to create expert systems to exhibit intelligent behavior to provide solutions to complicated problems and finally process massive data. Therefore, a study is proposed to provide the best methodological practice in the light of ML. Alongside the review of previous investigations, the following research aims to determine the ML approaches appropriate to OSH issues. In other words, highlighting specific ML methodologies, which have been employed successfully in others areas. Bearing this objective in mind, one can identify an appropriate ML technique to solve a problem in the OSH domain. Accordingly, several questions were designed to conduct the research. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for Protocols and Systematic Reviews were used to draw the research outline. The chosen databases were SCOPUS, PubMed, Science Direct, Inspect, and Web of Science. A set of keywords related to the topic were defined, and both exclusion and inclusion criteria were determined. All of the eligible papers will be analyzed, and the extracted information will be included in an Excel form sheet. The results will be presented in a narrative-based form. Additionally, all tables summarizing the most important findings will be offered.

## 1. INTRODUCTION

### 1.1 Background

Throughout the years, occupational health and safety practitioners, researchers, and managers have addressed the physical, chemical, biological, ergonomic, and organizational environment issues using traditional methods. However, with the rise of the Industry 4.0 era, this picture started to change. Industry 4.0 aims to improve workplaces within the digital revolution, including artificial intelligence (AI), the Internet of Things (IoT), and “smart” devices (Badri et al., 2018). AI is a computer science branch aiming to turn computers into intelligent machines using algorithms (Eicker et al., 1991). This term was first introduced in Dartmouth College in 1956, during a scientific conference (Hsueh and Oliveira, 2019). The main AI applications employ machine learning (ML) mathematical methods (algorithms), which provide statistical (either linear or nonlinear) models to perform predictions, classifications, and regression, on problems to be solved (Nicholson et al., 2019).

Different ML methods have been integrated into everyday life and workplaces (Can et al., 2019; Moore, 2019). The most common applications are those being embedded in physical objects such as sensors (Braun, 2009), robotic devices (Wallach and Marchant, 2019), or intelligent decision support systems (Lee et al., 2019). As an instance for physical sensors, considering the importance of monitoring and controlling particulate matter (PM) 2.5, the research presented in (Loh and Choi, 2019) targeted multiple affordable sensors instead of high-end ones via employing AI methods to control the scattering of light. For robotic devices, Eder et al. (2014) studied the techniques encompassing the presence of robots performing parallel with operators and the dangers and the responsibilities for human safety. They concluded their research by recommending safety considerations into the design of robotic co-workers from a learning algorithms perspective. One study even addressed the potential ML effects on automation and human-machine interactions (Kong et al., 2019), as generated and stored information can also help measure, assess, and control health and safety issues effectively. In the same scope, hazard prediction in complex industrial environments can also be addressed by the studying field (Zhao et al., 2019).

ML methods have also been used for safety training (Cho et al., 2018), personal protective equipment improvement (Ellena et al., 2016), environmental health surveillance (Xie and Chang, 2019), wellness and health promotion in different occupations like construction, petroleum, firefighting, office workers, drivers, among others (Jin et al., 2019). It could improve occupational safety and health (OSH) measures and equipment from one side, and influence the quality of new methods, used to measure, assess, and control hazards from the other side (Moore, 2019).

Considering the explained scope, in a recent study by Akanmu et al. (2020), a cyber-physical postural training environment has been designed and equipped with ML methods to train workers to work with less ergonomic risks. The system involves wearable sensors, VIVE trackers benefiting from ML methods to train workers in a virtual reality environment while tracking their body kinematics. Other similar works are also the study by Amrollahibuki et al. (2018) and Grabowski et al. 2014 (2014) for training mine workers.

While new technologies and new employment patterns have shaped the world of work, it has become more critical than ever to identify the emerging work-related occupational safety and health risks (Sarkar et al., 2017). Recognizing the hazards and risks is crucial and critical to effectively manage them and develop practical preventive actions (Howard, 2019). As an instance, the research done by Albert et al. (2017) can be considered, which include both recognizing micro-level hazard among different types and evaluating their impacts on the studied interventions. Novel ML applications in occupational health and safety raise important issues for understanding the limitation and possibilities and assess user interface effectiveness and usability. A review concerning the impact and best practices on the application of ML related to OSH is still missing from the literature. In light of that, it is proposed research that considers the main applications of ML in OSH and their potential and future challenges. The systematic review aims to answer a list of relevant topics offering a detailed analysis regarding the innovative and automated methods in this field, enabling a clearer view of the stakeholders generating and utilizing ML environments.

## 1.2 Research questions

This protocol aims to define the research methodologies to conduct a Systematic Review within the scope of ML, and more specifically, in the field of OSH. Several research questions were raised to accomplish this primary objective:

1. What are the ML methods implemented in the occupational setting?
2. What is the scope of each ML implementation for a specific OHS issue?
3. In which occupational activities are ML methods being implemented?
4. How effective and usable are ML methods for improving OSH?
5. Are the ML methods improving traditional approaches?
6. What are the limitations of the ML methods?

## 2. METHODOLOGY

The Preferred Reporting Items for Systematic reviews and Meta-Analyses (Page et al., 2021) was used to draw the guidelines for the systematic review protocol, as well as the Preferred Reporting Items for Systematic reviews and Meta-Analysis Protocol (PRISMA-P) checklist (Shamseer et al., 2015).

### 2.1 Eligibility criteria

To be considered eligible, a study has to address at least one of the topics referred to in the list of relevant questions presented in the Objectives (Section 1.2). All papers that consider the effect of ML applied in the public health and clinical field will be excluded. Participants of the study will include workers general population. Literature reviews will be excluded in the first phase. However, after the first set of eligible papers is found, their citations will be screened to identify other (grey) sources. Finally, it is important to mention that only publications written in English will be considered.

### 2.2 Information sources

Not to narrow the research in the first phase, only generalist electronic databases will be considered in this phase: IEEE, SCOPUS, PubMed, Science Direct, Inspect, and Web of Science. Additionally, to allow for the inclusion of as much evidence as possible, other sources will be considered to search for grey literature.

### 2.3 Search strategy

The keywords chosen to carry the primary research were "machine learning" and "expert system" and "Cognitive system" that are going to be sequentially combined with the main topics of interest: "occupational safety", "occupational health" and "work environment". The combinations will be as follows:

"machine learning" AND "occupational health"

"machine learning" AND "occupational safety"

"machine learning" AND "work environment"

"expert system" AND "occupational health"

"expert system" AND "occupational safety"

"expert system" AND "work environment"

"cognitive system" AND "occupational health"

"cognitive system" AND "occupational safety"

"cognitive system" AND "work environment"

In every database, the research will be carried out in the following way: the main combinations will be sought for in the "Title+Abstract+Keywords" field (Scopus) or the "Topic" field (Inspect, Science Direct, Web of Science) and in the "Title" and "Abstract" (PubMed).

After screening the selected manuscripts in the second stage, new keywords will be identified concerning the subject. They will be used in new search combinations with the keywords previously used.

Furthermore, references will also be scanned to find more articles that can provide complementary information. Finally, in the third stage of research, additional sources identified in the analyzed articles will be assessed. This procedure will be repeated in the newly selected papers until no more relevant information is obtained.

Table 1 (Appendix 1) will be used to register the number of articles after the applied criteria that they are going to be: Year (published after 2010), Document type (article, article in press and conference paper), Source type (journals and conference proceedings), and Language (English).

## **2.4 Data management**

An Excel-based file will be used to register, in a table form, the total number of articles obtained throughout the research, as previously mentioned. All of the selected papers will be exported and screened to remove duplicates. All the references will be managed using Mendeley software.

## **2.5 Selection process**

The selection of literature will occur in two screening phases, based on the defined eligibility criteria. In the first phase, titles and abstracts will be screened, and in the second phase, full-text of all potentially relevant literature will be reviewed. If any doubt arises regarding the title and abstract screening, papers will be kept for the next phase as a potentially relevant study to be reviewed in the full-text phase. In the second phase, any disagreement will be resolved by consensus through discussion between two reviewers. If the conflict remains, the paper will be passed to a third reviewer for final resolution.

## **2.6 Data collection process**

Qualitative data will be extracted from papers and included in the systematic review using a pre-structured table built by the authors. The aim will be to collect the data which will accurately answer the research questions and research objective. The table will be filled with the combined results of three reviewers. If any disagreement is observed, reviewers will try to solve it by discussing or using the fourth author's opinion for further disputes.

This will allow reducing bias and data extraction errors. The total number of screened studies, including those assessed for exclusion and inclusion criteria, will be determined. Furthermore, the reason to exclude studies during the selection process will be documented.

## **2.7 Data items**

Tables will be used to record study statistics, including study characteristics (date of publication, country), target group, industry, study design, OSH field (musculoskeletal disorders, ergonomics, machinery, workplace safety, risk assessment), the hardware used, employed ML taxonomy (artificial neural network - ANN, Fuzzy logic), ML aims and scopes (classification, decision making, pattern recognition, prediction, analysis, regression), requirements of these methods, and finally limitation of the study.

## **2.8 Outcomes and prioritization**

This systematic review aims to identify ML methods performed in OSH that involve measurement of various parameters in different workplaces, target groups. The second objective is to analyze the reliability of the results provided by those methods. The multiple methods of ML used in OSH, the methodological studies, and the list of the ML methods that have been examined for OSH with their possibilities and limitations will be reported.

By conducting this systematic review, a summary of the available research investigating ML in OSH in different industries for different phases of OSH, like monitoring, assessing, and controlling, will present. The obtained results from this review will help researchers, stakeholders, managers, and experts identify new technologies in OSH. It will also help to recognize knowledge gaps in the current literature, therefore, directing future research that will benefit and support the growth of OSH of industry.

## **2.9 Risk of bias in individual studies**

The bias assessment in included articles will be performed with three independent reviewers, using the methodology suggested by Higgins et al. (2011) on a two-term scale: "high risk" or "low risk". Whenever it is impossible to discern between those categories, an "unclear risk" classification will be used. Each paper will be analyzed in light of its methodology and results in order to detect possible biases.

## 2.10 Data synthesis

The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (Page et al., 2021) will be used to guide for writing of this review. The systematic review form will be used to write this study, and data synthesis will be carried out by extracting information from the eligible articles collected in the form sheet. In the table, the significant result will be highlighted.

## 2.11 Protocol registration

This protocol is under registration in PROSPERO

All authors read and approved the final version.

## REFERENCES

- Akanmu, A.A., Olayiwola, J., Ogunseiju, O., McFeeters, D., 2020. Cyber-physical postural training system for construction workers. *Autom. Constr.* 117, 103272. <https://doi.org/https://doi.org/10.1016/j.autcon.2020.103272>
- Albert, A., Hallowell, M., Skaggs, M., Kleiner, B., 2017. Empirical measurement and improvement of hazard recognition skill. *Saf. Sci.* 93, 1–8. <https://doi.org/10.1016/j.ssci.2016.11.007>
- Badri, A., Boudreau-Trudel, B., Souissi, A.S., 2018. Occupational health and safety in the industry 4.0 era: A cause for major concern? *Saf. Sci.* 109, 403–411. <https://doi.org/https://doi.org/10.1016/j.ssci.2018.06.012>
- Braun, M., 2009. 1 Introduction. *Word-Formation and Creolisation* 1–2. <https://doi.org/10.1515/9783484970229.1>
- Can, Y.S., Arnrich, B., Ersoy, C., 2019. Stress detection in daily life scenarios using smart phones and wearable sensors: A survey. *J. Biomed. Inform.* 92, 103139. <https://doi.org/https://doi.org/10.1016/j.jbi.2019.103139>
- Cho, C., Kim, K., Park, J., Cho, Y.K., 2018. Data-Driven Monitoring System for Preventing the Collapse of Scaffolding Structures. *J. Constr. Eng. Manag.* 144. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001535](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001535)
- Eder, K., Harper, C., Leonards, U., 2014. Towards the safety of human-in-the-loop robotics: Challenges and opportunities for safety assurance of robotic co-workers', in: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. pp. 660–665. <https://doi.org/10.1109/ROMAN.2014.6926328>
- Eicker, P.J., Miller, D.J., Strip, D.R., 1991. Intelligent Systems and Technologies for Manufacturing. *AT&T Tech. J.* 70, 10–22. <https://doi.org/10.1002/j.1538-7305.1991.tb00133.x>
- Ellena, T., Subic, A., Mustafa, H., Pang, T.Y., 2016. The helmet fit index - an intelligent tool for fit assessment and design customization. *Appl. Ergon.* 55, 194–207. <https://doi.org/10.1016/j.apergo.2016.02.008>
- Grabowski, A., Jankowski, J., 2014. Virtual Reality-based pilot training for underground coal miners, *Safety Science. Saf. Sci., Safety Science* 72, 310–314. <https://doi.org/10.1016/j.ssci.2014.09.017>
- Higgins, J.P.T., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., Savović, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomized trials. *BMJ* 343. <https://doi.org/10.1136/bmj.d5928>
- Howard, J., 2019. Artificial intelligence: Implications for the future of work. *Am. J. Ind. Med.* 62, 917–926. <https://doi.org/10.1002/ajim.23037>
- Hsueh, T., Oliveira, E., 2019. Artificial Intelligence: An Overview. *Cut. Edge Technol. Microcomput. Appl. Dev. Ctries.* 61–65. <https://doi.org/10.4324/9780429042522-6>
- Kong, X.T.R., Luo, H., Huang, G.Q., Yang, X., 2019. Industrial wearable system: the human-centric empowering technology in Industry 4.0. *J. Intell. Manuf.* 30, 2853–2869. <https://doi.org/10.1007/s10845-018-1416-9>



Lee, J., Azamfar, M., Singh, J., 2019. A blockchain enabled Cyber-Physical System architecture for Industry 4.0 manufacturing systems. *Manuf. Lett.* 20, 34–39. <https://doi.org/https://doi.org/10.1016/j.mfglet.2019.05.003>

Loh, B.G., Choi, G.H., 2019. Calibration of Portable Particulate Matter–Monitoring Device using Web Query and Machine Learning. *Saf. Health Work* 10, 452–460. <https://doi.org/https://doi.org/10.1016/j.shaw.2019.08.002>

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & and the PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine*, 151(4), 264–269. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>

Moore, P. V., 2019. OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces, in: Duffy, V.G. (Ed.), *Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Human Body and Motion.* Springer International Publishing, Cham, pp. 292–315. <https://doi.org/10.1016/j.promfg.2020.04.017>

Nicholson, C., 2019. AI Wiki A Beginner’s Guide to Important Topics in AI, Machine Learning, and Deep Learning. <https://wiki.pathmind.com/neural-network>

Page, M.J., Moher, D., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, MM, Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., McKenzie, J.E., 2021. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 372. <https://doi.org/10.1136/bmj.n160>

R. Amrollahibuki, A.H., 2018. Modeling construction equipment in 4D simulation, in: *EWork and EBusiness in Architecture, Engineering and Construction. Proceedings of the 12th European Conference on Product and Process Modelling (ECPM 2018)*, Copenhagen, Denmark. ISBN:9780429506215

Sarkar, S., Kumar, A., Mohanpuria, S.K., Maiti, J., 2017. Application of Bayesian network model in explaining occupational accidents in a steel industry, in: V., P., A., Mukherjee, A., Mondal, A., D., U., M., D., B., H., B., BK, P., T., G., S., B., I., P. (Eds.), *3rd IEEE International Conference on Research in Computational Intelligence and Communication Networks, ICRCICN 2017.* Institute of Electrical and Electronics Engineers Inc., Department of Industrial and Systems Engineering, IIT Kharagpur, India, pp. 337–342. <https://doi.org/10.1109/ICRCICN.2017.8234531>

Shamseer, L., Moher, D., Clarke, M., Gherzi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A., 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 349. <https://doi.org/10.1136/bmj.g7647>

Wallach, W., Marchant, G., 2019. Toward the agile and comprehensive international governance of AI and robotics. *Proc. IEEE* 107, 505–508. <https://doi.org/10.1109/JPROC.2019.2899422>

Xie, X., Chang, Z., 2019. Intelligent Wearable Occupational Health Safety Assurance System of Power Operation. *J. Med. Syst.* 43. <https://doi.org/10.1007/s10916-018-1122-3>

Zhao, Y., Li, J., Zhang, M., Lu, Y., Xie, H., Tian, Y., Qiu, W., 2019. Machine Learning Models for the Hearing Impairment Prediction in Workers Exposed to Complex Industrial Noise: A Pilot Study. *Ear Hear.* 40, 690–699. <https://doi.org/10.1097/AUD.0000000000000649>

*Appendix 1*

**Table 1.** Data records

Summary of the Selected Articles	Records identified through database	Records excluded by screening criteria	Summary of Total Rejected Items by screening criteria						Database	Keyword 1								
			Date	Document type	Source type	Language	Other	Off topic		Number of Selected Articles	Number of Collected Articles	Number of included articles after exclusion criteria						
												Date	Document type	Source type	Language	Other	On topic	