Improving Health Monitoring of Construction Workers Using Physiological Data-Driven Techniques: An Ensemble Learning-Based Framework to Address Distributional Shifts

Amit Ojha, S.M.ASCE[1]; Yizhi Liu, S.M.ASCE[2]; Houtan Jebelli, Ph.D., A.M.ASCE[3]; Hunayu Cheng, Ph.D., A.M.ASCE[4]; and Mehdi Kiani, Ph.D.[5]

[1]Dept. of Civil and Environmental Engineering, Univ. of Illinois Urbana-Champaign, Champaign, IL. Email: amito2@illinois.edu
[2]Dept. of Architectural Engineering, Pennsylvania State Univ., University Park, PA. Email: ykl5505@psu.edu
[3]Dept. of Civil and Environmental Engineering, Univ. of Illinois Urbana-Champaign, Champaign, IL (corresponding author). Email: hjebelli@illinois.edu
[4]Dept. of Architectural Engineering, Pennsylvania State Univ., University Park, PA. Email: huc24@psu.edu
[5]School of Electrical Engineering and Computer Science, Pennsylvania State Univ., University Park, PA. Email: muk42@psu.edu

ABSTRACT

While researchers have used various off-the-shelf physiological sensors and prevalent machine learning (ML) algorithms to objectively assess construction workers' health status, there remain specific challenges for consistent and accurate health monitoring on the jobsite. The existing physiological-based data-driven frameworks for predicting workers' health status in the field are not robust to the distribution shift of physiological signals and face challenges in stability, reliability, and accuracy. To overcome these issues, this paper proposes using an ensemble learning technique implemented on a support vector machine (SVM) with the Adaptive Boosting (AdaBoost) algorithm to develop a resilient predictive performance of the data-driven framework. To examine the performance of the framework, physiological signals were collected from 10 subjects performing material handling tasks with varying levels of physical fatigue. The proposed framework predicted the physical fatigue level with over 88% accuracy, better than single machine learning classifiers. This study has significant implications for improving the accuracy and stability of physiological-sensing-based health monitoring.

INTRODUCTION

Construction workers suffer from extensive work-related mental and physical stress at the job sites, causing a high number of illnesses, injuries, and fatalities (Bureau of Labour 2018). Despite the high rate of work-related injuries and illness among construction workers, there is a lack of an effective, objective, and continuous approach to assessing and evaluating workers' health status at construction sites. The current methods for assessing workers' physical and mental states mostly rely on the use of self-assessment measures (i.e., self-report, questionnaires, and rating scales) (Habibi et al. 2014). Such methods can be challenging to implement in the field due to their inherently subjective and intrusive nature (Rabeiy 2019). In addition, these methods also do not account for workers' physical and physiological characteristics. Previous studies have leveraged non-invasive physiological sensors in conjunction with prevalent machine learning algorithms to identify potential safety hazards and to assess workers' physical and mental states at construction jobsites (Jebelli et al. 2018; Shakerian et al. 2021). However, most of the current physiological-sensing models rely on a single classifier, such as support vector machine (SVM), k-nearest neighbors (kNN), Random Forest (RF), or Naïve Bayes (NB) as decoders in physiologically based data-driven models, which are susceptible to several factors that may compromise the reliability and accuracy of the health assessments (Li et al. 2019; Reddy and Hota 2013). A single classifier may not adequately capture the complexity of workers' health conditions, potentially resulting in reduced accuracy (Wang et al. 2015). Additionally, a single classifier may be susceptible to overfitting when the training data is limited and noisy, leading to poor generalized performance in real-world scenarios (Wang et al. 2015). Likewise, a single classifier is not robust to the distribution shift of the physiological signals (Liu et al. 2021), impeding the capability of accurate health assessments.

Given such challenges, there is a need to develop more accurate and reliable predictive models and algorithms that can securely analyze real-time physiological data for continuous health monitoring of workers. To that end, this study utilizes an ensemble learning-based framework which utilizes a Support Vector Machine (SVM) with the Adaptive Boosting (AdaBoost) algorithm to improve the predictive performance of physiological-sensing-based health monitoring. By using multiple weak SVM classifiers and combining them using the AdaBoost algorithm, the framework is able to capture complementary information from different subsets of the data, resulting in improved classification accuracy and stability. The iterative updating of the weights of each weak classifier further enhances the framework's adaptability to changing data distributions and enables it to focus more on difficult-to-classify samples. To evaluate the performance of the developed framework, electrodermal activity (EDA), photoplethysmography (PPG), skin temperature (ST), and respiratory signals were collected from 10 subjects performing material handling tasks with varying levels of physical fatigue (low and high). The collected physiological signals were decontaminated from noises, resampled into determined timeframes, and informative features were extracted and finally interpreted into distinct states of physical fatigue levels by employing SVM with AdaBoost algorithms.

AI-DRIVEN PHYSIOLOGICAL SENSING-BASED HEALTH MONITORING OF CONSTRUCTION WORKERS

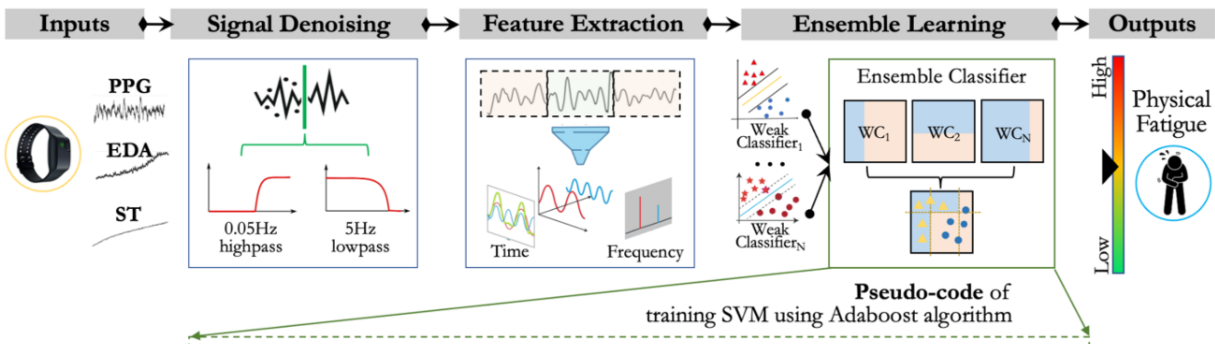
Physiological responses to a workplace stimulus can offer valuable information about the holistic health status of an individual (Awolusi et al. 2018; Jebelli et al. 2019). Previous studies suggested PPG, EDA, and ST signals could capture information about the body responses to an external stressor and be used for assessing the health status of workers (Aryal et al. 2017; Hwang and Lee 2017; Jebelli et al. 2019). The advent of advanced wearable sensing technologies and robust machine learning algorithms has created great opportunities for sustained health monitoring of the field workers (Tixier et al. 2016). Previous studies have used various off-the-shelf physiological sensors, in conjunction with prevalent machine learning algorithms, to identify potential safety hazards and to assess workers' physical and mental states at construction jobsites. To ensure the efficiency and accuracy of physiological sensing-based data-driven framework, a fundamental requirement is to accurately decode physiological signals into meaningful information. Towards that end, several studies (including the authors' previous study) leveraged machine learning (ML) classifiers, such as Support Vector Machine, Artificial Neural Networks, and K-Nearest Neighbor, as decoders in physiologically based data-driven models for objectively discerning different levels of workers' physiological states. In this regard, the classification accuracy of the ML classifier exclusively determines the accuracy of data-driven models. However, such ML classifiers have difficulty in effectively decoding human physiological signals with high accuracy and stability. Firstly, a single classifier may fail to capture the full complexity of workers' health conditions, resulting in reduced accuracy and reliability (Wang et al. 2015). Second, a single classifier may be prone to overfitting, especially when the training data is limited and noisy, leading to poor generalization performance and reduced accuracy in real-world scenarios (Wang et al. 2015). Third, a single classifier may be sensitive to distribution shifts, which can occur when the distribution of the data changes over time or between different populations (Liu et al. 2021). Different workers respond differently to the same stressors; even the same worker can react differently to similar stressors during multiple exposures (Cohen and Hamrick 2003; Matthews et al. 1986). This causes a distribution shift in the data, impeding the capability for accurate assessment of health conditions. In this regard, the single ML classifiers leveraged for physiologically based data-driven models cannot consistently and accurately decode workers' real-time physiological signals into meaningful information for workers' health status.

METHODOLOGY

This study developed an enhanced health monitoring framework for consistently and accurately estimating workers' health status based on their EDA, PPG, and ST signals acquired from wearable biosensors. The overview of the proposed ensemble learning-based approach to improving the predictive performance of physiological-sensing-based health monitoring is shown in Figure 1. The proposed methodology is mainly orchestrated through three consecutive steps:

1) Signal Denoising: To filter out high- and low-frequency noises (electrode, electromagnetic, and thermal noises), which could contaminate EDA, ST, and PPG signals, the authors leveraged a bandpass filter with lowpass and highpass cut off frequencies of 5 filter with the cut-off frequency of 5 Hz, and the 0.05 Hz highpass filter, respectively. Furthermore, the Hampel and moving average filters were applied to reduce outliers from the extracted signals. Further details on how these filtering techniques were applied can be found in our previous studies (Sadat-Mohammadi et al. 2021; Shakerian et al. 2021).
2) Feature Extraction: After artifact removal of the collected signals, all filtered signals were resampled into fixed-sized frames of data points (windows) to extract features. Informative features were extracted from these signals to train a physiologically based data-driven model. The extracted features included measurable metrics in the time domain (e.g., mean value, variance, median value, smallest window element, maximum-to-minimum difference, root-mean-square level, and root-sum-of-squares level) and measurable parameters in the frequency domain (e.g., mean and median frequencies). The details of the feature extraction step can be found in our previous work (Shakerian et al. 2021).
3) Ensemble Learning: The study integrates SVM with the AdaBoost algorithm to develop an improved ensemble learning approach to enhance the predictive performance of physiological-sensing-based health monitoring. SVM with AdaBoost involves using multiple weak SVM classifiers trained on subsets of the data and combining their predictions using the AdaBoost algorithm. The weights of each weak classifier are

iteratively updated based on their performance on the training data, resulting in a final classifier that is a weighted sum of the weak classifiers. The developed ensemble learning-based framework is trained with features extracted from the previous step. More details about the pseudocode of the proposed ensemble learning approach can be seen in Figure 1.



Algorithm 1 – Ensemble learning approach

Input: 1). N training examples, $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, where x_i indicates the extracted features; y_i ($y_i = +1, -1$) is the label for each x_i; D indicates the sample distribution 2) Number of training iterations, T.

Initialize: weight for each sample $\alpha_{t=1} = \frac{1}{N}$

1. **For** $t = 1, 2, \dots, T$
2. Train **weak SVM** (SVM_t) with all training examples $D_i^t \in D^t$ (D^t indicates the distribution of the training samples in the t^{th} iteration; $D_i = (x_i, y_i)$)
3. Compute the error of SVM_t: $\varepsilon_t = \sum_{i=1}^{N} D_i^t \times e_i^t$; $e_i^t = 1$ ($SVM_t(x_i) \neq y_i$) or $e_i^t = 0$ ($SVM_t(x_i) = y_i$)
4. Update weight vector $\alpha_t = \frac{1}{2} \ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$;
5. Update the data distribution D^t based on α_t:
6. For $i = 1, 2, \dots, N$: $D_i^{t+1} = \frac{D_i^t}{Normalization} \times \begin{cases} e^{-\alpha_t} & \text{if } SVM_t(x_i) = y_i \\ e^{\alpha_t} & \text{if } SVM_t(x_i) \neq y_i \end{cases}$

Output: generate the final ensemble classifier E_{svm}:

7. $E_{svm}(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t \times SVM_t(x)\right)$

Figure 1. Details of the proposed enhanced health monitoring framework.

In the ensemble learning approach, multiple weak SVM classifiers are trained on varying distributions of the data, and their predictions are combined using the AdaBoost algorithm, which iteratively updates the weights of each weak classifier based on their performance on the training data. As outlined in Figure 1, the inputs of the approach are the training samples, with the sample distribution: $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, where x is the matrix of input features, and y is the vector of class labels, and the number of iterations T. The developed approach first initializes the weight vector α to be uniform, where each weight is 1/N. The initialization is done to ensure that all the samples are given equal importance in the initial training phase. Line 1 performs T iterations of the AdaBoost. During each iteration, the t^{th} weak SVM classifier (SVM_t) is trained on the training set with distribution D^t, as indicated in Line 2. Once obtained the trained SVM_t, the algorithm computes the classification error (ε_t) of SVM_t using the equation formulated in Line 3. Line 4 computes the weight vector α_t based on the error rate ε_t of the trained weak classifier. The α_t is higher for the weak classifier that performs well and lower for the weak classifier that performs poorly. Then, in lines 5 and 6, the weight of each data point in the distribution (D_i^t) is updated by multiplying by a factor that depends on whether the chosen weak classifier correctly classifies the particular data point or not. In this vein, the correctly classified data point has a factor of $e^{-\alpha_t}$ and the misclassified samples have a factor of e^{α_t}. By doing this, the algorithm will give more importance (e^{α_t}) to the misclassified samples, which can help improve the performance of the weak classifier. In Line 7, the final ensemble classifier E_{svm} is computed as a weighted sum of the T weak SVM classifiers, where each weak classifier is weighted by its weight α_t. The sign($\cdot$) function is used to determine the final class label for each sample. The main advantage of the proposed ensemble learning-based SVM with AdaBoost data-driven framework is its ability to improve the accuracy and stability of SVM-based classification models. By combining multiple weak SVM classifiers into a robust classifier using the AdaBoost algorithm, the framework can adapt to changing data distributions and focus more on data points that are difficult to classify.

CASE STUDY

To examine the performance of the proposed enhanced health monitoring framework, a case study procedure was established to collect physiological signals from subjects in the lab-controlled environment at Penn State University, as illustrated in Figure 2.

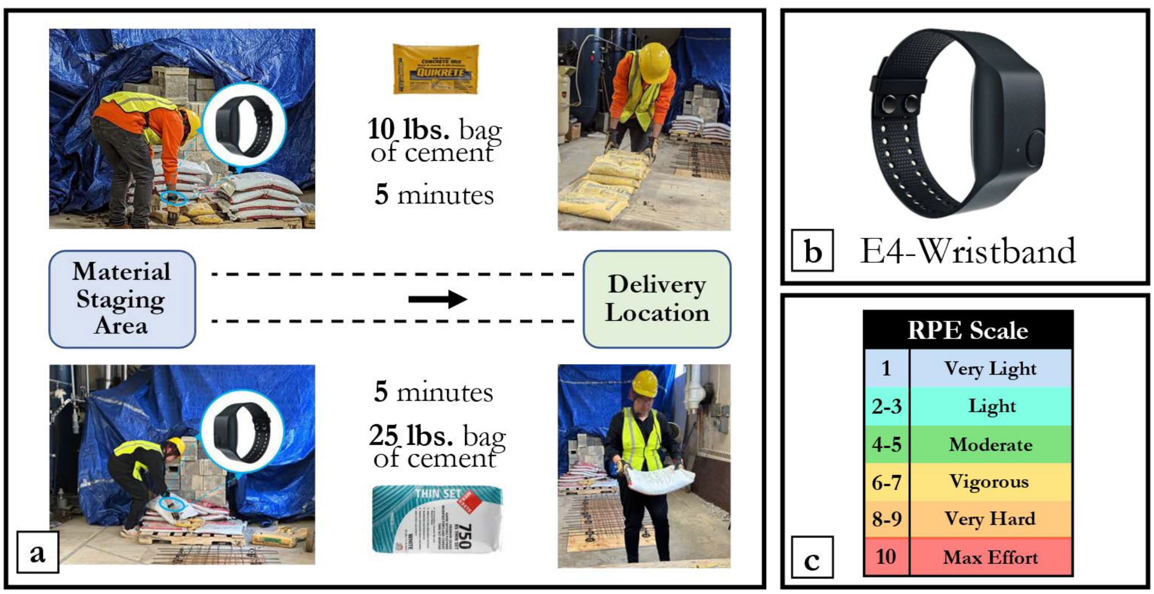


Figure 2. Details of the designed material handling task.

Ten able-bodied subjects were required to complete material handling tasks with varying levels of physical fatigue. Before starting the case study, informed written consent was obtained from all the subjects following the procedure approved by the Institution Review Board (IRB) at Penn State. All the subjects had basic construction engineering knowledge and experience in conducting manual handling tasks on construction sites. Moreover, none of the participants reported any history of mechanical pain or injury. The study used a randomized crossover design where each participant performed the task under two scenarios (Figure 2-a): in one scenario, each subject was required to lift, carry, and lower a 10 lbs. bag of cement for 5 minutes from the material storage area to the delivery location; in the other scenario, they were required to perform the same task with a 25 lbs. bag of cement. During the task, subjects were asked to wear the wristband sensor (Empatica™ E4; Figure 2-b) on their dominant hand. This wearable sensing device included a PPG sensor, an EDA sensor, and an infrared thermopile sensor to collect the PPG, EDA, and ST signals at a sampling rate of 64, 4, and 4 Hz, respectively. After each scenario, subjects were asked to rate their level of physical fatigue on the Borg Rate of Perceived Exertion Scale (Borg RPE Scale; Figure 3-c). Borg RPE Scale was used as a baseline to assess subjects' perceived physical fatigue levels. Based on their perceived physical fatigue levels, each subject’s EDA, PPG, and ST signals were stacked into a dataset and labeled into Low Fatigue Level (RPE scale between 1 and 5) and High Fatigue Level (RPE scale between 6 and 10). Then, the authors used the dataset and their corresponding labels to examine the performance of the proposed enhanced health monitoring framework, the results of which will be reported in the next section.

RESULTS

The proposed framework was applied to the data collected from the subjects after randomly dividing 80% of the data into training datasets and 20% into validation datasets. As stated in the previous section, each captured physiological signal (EDA, PPG, and ST) was stacked into the dataset and employed to examine the performance of the proposed enhanced health monitoring framework. Further, the authors also compared the performance of the proposed framework with traditional ML classifiers, which included kNN, Logistic Regression, SVM, SVM with Gaussian kernel, and Quadratic Discriminant Analysis (QDA). Figure 2-a illustrates the validation accuracy of the ML classifiers in estimating the workers' physical fatigue levels for the captured physiological signals trained using the five-folder cross-validation technique. As demonstrated in Figure 3-a, the proposed E_{svm} had the highest validation accuracy of 88.67%. Likewise, SVM with Gaussian kernel had the second-best performance amongst the ML classifiers, with a validation accuracy of 85.9%. Compared to the SVM with Gaussian kernel, the proposed E_{svm} achieved a performance gain of around 3%.

a) Accuracy (validation) of the ML classifiers

	Validation Accuracy
Ensemble SVM	88.26%[1]
kNN (k=3)	75.9%
Logistic Regression	84.1%
SVM (linear kernel)	84.6%
SVM with Gaussian	85.9%[2]
QDA	83.8%

[1]Ensemble SVM, the proposed method had the best accuracy
[2]SVM with Gaussian had the second based accuracy amongst the other ML classifiers.

b) Performance of the proposed Ensemble SVM and SVM with Gaussian classifier

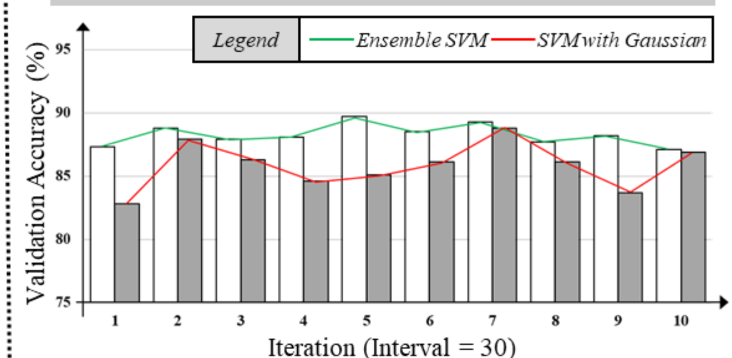


Figure 3. Performance of the proposed enhanced health monitoring framework

To analyze the stability of the performance of E_{svm}, the authors tested and visualized the validation accuracy for every 30 iterations and compared it with the second-best performing classifier, SVM with Gaussian kernel. Fig 3-b shows the comparison of the performance, where the green line represents the validation accuracy of the proposed Ensemble SVM framework, and the red line represents the validation accuracy of the SVM with Gaussian kernel. The validation accuracy of the E_{svm} is stable across the iterations, whereas the validation accuracy of the second-best performing classifier SVM with Gaussian classifier fluctuates. Such stability in the performance of E_{svm} suggests that the proposed framework has a robust generalization performance to the new data and is more likely to be robust to distribution shift.

CONCLUSION

This study aimed to develop an enhanced health monitoring framework for consistently and accurately estimating workers' health status based on their physiological signals acquired from wearable biosensors. For this purpose, the authors proposed an ensemble learning-based framework that utilizes a Support Vector Machine with the Adaptive Boosting algorithm to improve the predictive performance of physiological-sensing-based health monitoring. By applying the developed approach to a material handling task with varying levels of physical fatigue, the results demonstrate that the proposed framework can promptly estimate the likelihood of physical fatigue with an accuracy of 88.26%. The proposed ensemble learning-based framework offers a promising approach to improve the accuracy and stability of physiological-sensing-based health monitoring. This study has significant implications for improving the health and safety of construction workers, as continuous health monitoring can facilitate timely interventions and prevent illness, injury, and fatalities in the workplace. But further research is necessary to validate the proposed framework in larger-scale studies and to explore its potential for real-world applications. Future studies could also investigate the impact of the proposed framework on the overall health outcomes of construction workers and assess the economic feasibility of implementing such a framework in the construction industry.

ACKNOWLEDGEMENT

This material is based upon work supported by the National Science Foundation under Grant No. ECCS-2222654. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Aryal, A., A. Ghahramani, and B. Becerik-Gerber. 2017. "Monitoring fatigue in construction workers using physiological measurements." *Autom Constr*, 82: 154–165. Elsevier B.V.

Awolusi, I., E. Marks, and M. Hallowell. 2018. "Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices." *Autom Constr*, 85: 96–106. Elsevier B.V.

Bureau of Labour. 2018. *Employer-Reported Workplace Injuries and Illnesses*.

Cohen, S., and N. Hamrick. 2003. "Stable individual differences in physiological response to stressors: implications for stress-elicited changes in immune related health." *Brain Behav Immun*, 17 (6): 407–414.

Habibi, P., H. Dehghan, S. Rezaei, and K. Maghsoudi. 2014. Physiological and perceptual heat strain responses in Iranian Veiled women under laboratory thermal Conditions. *Iranian Journal of Health, Safety & Environment*.

Hwang, S., and S. H. Lee. 2017. "Wristband-type wearable health devices to measure construction workers' physical demands." *Autom Constr*, 83: 330–340. Elsevier B.V.

Jebelli, H., B. Choi, and S. Lee. 2019. "Application of Wearable Biosensors to Construction Sites. II: Assessing Workers' Physical Demand." *J Constr Eng Manag*, 145 (12).

Jebelli, H., S. Hwang, and S. Lee. 2018. "EEG-based workers' stress recognition at construction sites." *Autom Constr*, 93: 315–324.

Li, R., L. Zhou, S. Zhang, H. Liu, X. Huang, and Z. Sun. 2019. "Software Defect Prediction Based on Ensemble Learning." *ACM International Conference Proceeding Series*, 6: 1–6. Association for Computing Machinery.

Liu, Y., M. Habibnezhad, and H. Jebelli. 2021. "Brain-computer interface for hands-free teleoperation of construction robots." *Autom Constr*, 123: 103523. Elsevier.

Matthews, K. A., S. B. Manuck, and P. G. Saab. 1986. "Cardiovascular Responses of Adolescents During a Naturally Occurring Stressor and Their Behavioral and Psychophysiological Predictors." *Psychophysiology*, 23 (2): 198–209.

Rabeiy, R. E. 2019. "Evaluation of indoor heat stress on workers of bakeries at Assiut City, Egypt." *International Journal of Environmental Science and Technology*, 16 (6): 2637–2642. Center for Environmental and Energy Research and Studies.

Reddy, J. M., and C. Hota. 2013. "P2P traffic classification using ensemble learning." *ACM International Conference Proceeding Series*. Association for Computing Machinery.

Sadat-Mohammadi, M., S. Shakerian, Y. Liu, S. Asadi, and H. Jebelli. 2021. "Non-invasive physical demand assessment using wearable respiration sensor and random forest classifier." *Journal of Building Engineering*, 44: 103279. Elsevier BV.

Shakerian, S., M. Habibnezhad, A. Ojha, G. Lee, Y. Liu, H. Jebelli, and S. H. Lee. 2021. "Assessing occupational risk of heat stress at construction: A worker-centric wearable sensor-based approach." *Saf Sci*, 142: 105395. Elsevier.

Tixier, A. J.-P., M. R. Hallowell, B. Rajagopalan, and D. Bowman. 2016. "Application of machine learning to construction injury prediction." *Autom Constr*, 69: 102–114.

Wang, C., L. Hu, M. Guo, X. Liu, and Q. Zou. 2015. "imDC: an ensemble learning method for imbalanced classification with miRNA data." *Genetics and Molecular Research*, 14 (1): 123–133.