VOLUME xx, NUMBER xx

2024

Muhammad Romdon Dahlan
Teknik Industri, Institut Teknologi Nasional Bandung

Caecilia Sri Wahyuning
Teknik Industri, Institut Teknologi Nasional Bandung

ATMAE
PMB 219
8865 Norwin Ave STE 27
North Huntingdon, PA 15642

www.atmae.org

JTMAE
The Journal of Technology, Management, and Applied Engineering

Machine Learning Approach for Modeling Prediction of Human Errors due to Stress Based on Work Environment and Job Demands

Keywords:
Stress; Work Environment; Job Demand; Human Error; Machine Learning

APPLIED RESEARCH

Corresponding author: Muhammad Romdon Dahlan, romdonofficial@gmail.com

Dahlan, M.R. & C.S. Wahyuning. Machine Learning Approach for Modeling Prediction of Human Errors due to Stress Based on Work Environment and Job Demands. The Journal of Technology, Management, and Applied Engineering, xx(xx), 2024, 1–13.
https://doi.org/10.31274/jtmae.16882
Submitted: July 5, 2023; Accepted: May 16, 2024

Machine Learning Approach for Modeling Prediction of Human Errors due to Stress Based on Work Environment and Job Demands

ABSTRACT

Previous research has elucidated the prediction of human errors through the integration of eye tracking and machine learning methods. However, work-related stress can directly influence human errors due to the environment and job demands. In addition, the methods applied in previous studies were limited to computer-based tasks only. This research utilizes direct measurements of the human body and the environment, thereby not being limited to specific types of work only. The aim of this study is to develop a predictive model for human error occurrence due to stress based on the work environment and job demands, using a classification algorithm approach in machine learning. Machine learning algorithms such as random forests and decision trees are applied to classify the occurrence of human errors. The research process begins with collecting a dataset, data preprocessing, modeling, and evaluation. The results show that both algorithms achieve a model accuracy and recall of >90%. Both algorithms can be used to predict human errors.

Introduction

Product quality is a primary concern in manual manufacturing processes, where the produced products need to be widely accepted, meeting usability and requirement specifications (Dhafr et al., 2006). Generally, product quality is determined through inspection processes based on acceptance or rejection standards. Product quality implies that the produced items are free from any kind of product defects (Saptari et al., 2014).

In manual manufacturing processes, defects can be influenced by human actions as performance shaping factors, where deviations occurring during the process are considered human errors (Bubb, 2005; Saptari et al., 2015; Torres et al., 2021). These deviations are caused by the inability or failure to perform specific tasks. Failures that occur define the form of human reliability (Dhillon, 2013; Rasmussen & Laumann, 2018; Wahyuning & Atiko, 2022).

Human reliability indicates the likelihood of success or failure of humans in operating a system within a specified minimum time frame (Dhillon, 2013). Failure to perform specific tasks affects work performance and results in operational disruptions (Wahyuning & Atiko, 2022). Decreased performance is one manifestation of fatigue due to the decline in musculoskeletal and cognitive functions, which often occur simultaneously (Wahyuning & Atiko, 2022).

Heavy physical workload is one of the factors that contribute to mental fatigue (Xing et al., 2020; Sui et al., 2022). Workload refers to the magnitude of job demands and the increased muscular requirements during work activities (Wahyuning & Atiko, 2022). This fatigue arises after going through the strain phase. The strain phase is part of the general adaptation syndrome, where the body attempts to cope with the most perceived stressors. Fatigue occurs as a prolonged effect of sustaining and overcoming strain from the most perceived stressors (Landy & Conte, 2013; Wahyuning, 2017).

During stress, the adrenal glands in the brain secrete hormones such as epinephrine and norepinephrine to protect the body in stressful situations (Taelman et al., 2008; Scanlon & Sanders, 2014). The impact of this secretion results in increased blood pressure, which affects changes in heart rate (HR) and heart rate variability (HRV) (Wahyuning, 2017; Wahyuning et al., 2017). Therefore, an individual's stress condition can be indicated by changes in HR and HRV.

Work stress is influenced by various factors such as workload, responsibility, job opportunities, workplace issues, job requirements, health conditions, control span, job demands, and physical environment (National Institute for Occupational Safety and Health, 2017; Sumardiyono et al., 2020; Wahyuning, 2017). Sources of stress can be psychological/psychosocial demands as well as physical stressors. Physical stressors include noise, air temperature, lighting, and vibration, as well as job tasks/demands (work speed, mental workload, and working hours), including production targets that create time pressure (Ifrikar et al., 2006; Landy & Conte, 2013; Saptari et al., 2014; Wahyuning et al., 2017). These factors are the root causes of decreased performance that ultimately lead to human errors (Wahyuning & Atiko, 2022).

Studies on predicting human errors have been conducted by Damacharla et al. (2019), Saboundji & Robert (2020), and Wahyuning and Atiko (2022). Human reliability using the mean time to human error (MTTHE) approach is employed to predict the occurrence time of human errors (Wahyuning & Atiko, 2022). This research models when human errors occur based on stress resulting from physical workload (Wahyuning & Atiko, 2022). However, the predictive model does not directly take into account stress and physical workload but relies on historical data and time intervals between failures.

Another approach to predicting human errors is through machine learning. In the first study, gaze movement and mouse cursor movement features were used to classify task success and failure (Saboundji & Robert, 2020). Several machine learning techniques were applied, including decision tree (DT), logistic regression (LR), support vector machines (SVM), random forests (RF), and long short-term memory (LSTM). LSTM yielded the highest accuracy of 86%. In the second study, the eye tracking metric method was employed, using the SVM machine learning technique, achieving an accuracy of nearly 98% (Damacharla et al., 2019). However, previous studies (Damacharla et al., 2019; Saboundji & Robert, 2020; Wahyuning & Atiko, 2022) did not consider performance decline triggers as the main variables in predicting task success or failure.

In this research, machine learning is used to predict task success or failure/human errors, indicating human performance in manual-based manufacturing processes, using a data-driven approach. Failures/errors in manual-based manufacturing can be influenced by the environment, job demands, and stress.

The utilization of machine learning as a computational tool plays a crucial role in the effectiveness of prediction models (Damacharla et al., 2019; Saboundji & Robert, 2020), data mining, pattern recognition, and information extraction (Ge et al., 2017; Ismail et al., 2021). Machine learning can provide fast results in predicting human errors, enabling the early handling of potential product defects. This research will employ machine learning techniques such as DT and RF to classify the occurrence of errors. These techniques are chosen for their good performance on small and relatively simple datasets and their ability to be understood without relying on specialized statistical knowledge for interpretation (Patel & Prajapati, 2018; Bansal et al., 2022). The aim of this study is to develop a predictive model for human error occurrence due to stress based on the work environment and job demands, using a classification algorithm approach in machine learning.

Methodology

This prediction model is constructed based on quantitative data collected from the garment industry. The study utilizes secondary data obtained from case studies conducted at sewing workstations involving six operators (Lucky, 2022). Measurements were taken on the sewing machine operators.

The prediction model of human error using machine learning involves the use of predictor variables and a target variable (Table 1). The predictor variables are used to predict the target variable. Stress measurement is conducted using the HRV data approach. HRV has parameters such as MeanRR, SDNN, RMSSD, Mode, MxDMn, pNN50, AMo50, CV, Total Power, HF, LF, VLF, LF/HF, HF/LF/VLF, and HR (Tarvainen et al., 2014; Rotenberg & McGrath, 2016; Kim et al., 2018; Järvelin et al., 2019; Sarafinjuk et al., 2020; Attar et al., 2021; Ananda & Wahyuning, 2022; Lucky, 2022; Welltory Team, 2023):

- Mean RR (R-R interval) is the average time between consecutive heartbeats in normal-normal intervals.

Table 1.
Research Variables

Predictor Variables		Target Variable
Stress (HRV)	MeanRR, SDNN, RMSSD, Mode, MxDMn, pNN50, AMo50, CV, Total Power, HF, LF, VLF, LF/HF, HF/LF/VLF, and Heart Rate	Occurrence of Error or Not (Binary)
Work Environment	Air Temperature (°C), Globe Temperature (°C), Wet Bulb Temperature (°C), Humidity (Rh), and Lighting (Lux)	
Job Demands	Production Target	

- MxDMn (difference between maximum and minimum value) indicates a more active parasympathetic nervous system activity, and this is associated with a better body recovery capability.
- MO (Mode) refers to the value or values that occur most frequently in the distribution of RR intervals. In HRV analysis, the mode is often used to indicate the range of RR intervals that are dominant or most commonly occurring during a specific period of time.
- AMo50 (mode amplitude) evaluates sympathetic activity, and a higher value indicates a more active sympathetic system.
- SDNN (standard deviation of normal-to-normal), RMSSD (root mean square successive difference), pNN50 (proportion of NN50), and HF (high frequency) indicate an increase in stress when their values decrease.
- VLF (very low frequency), LF (low frequency)—the higher their values, the higher the stress level.
- Total power indicates how much power the body has and how well it adapts to stress.
- LF/HF indicates which system works harder between the parasympathetic or sympathetic.
- HF/LF/VLF (wave balance) indicates the strength of the system regulating heart activity.
- The coefficient of variation (CV) is the ratio of the standard deviation (SD) of RR intervals to the mean RR interval, typically expressed as a percentage. This parameter is used to measure the relative level of variability of RR intervals in HRV.
- HR can provide an indication of a person's stress level. The HR response to stress is influenced by the autonomic nervous system, which is divided into two main branches: the sympathetic nervous system and the parasympathetic nervous system.

Work environment and job demands are other variables used in this research. In this study, the work environment parameters used are air temperature (°C), globe temperature (°C), wet bulb temperature (°C), humidity (RH), and lighting (Lux). These work environment parameters serve as predictors of human errors. In coping with heat exposure, the human body increases skin temperature and maintains its core temperature close to 37°C, while HR and cardiac output increase and redirect blood circulation to the skin (Wahyuning & Laksemi, 2021). Job demand is a predictor variable that consists of the production targets that operators need to achieve within a batch.

The target variable is the prediction of human errors. Inaccuracy or mistakes in decision-making are often referred to as human errors, which occur in a significant portion of tasks that require cognition. Human performance reliability represents the probability of failure, indicating the likelihood that humans will fulfill all specified human functions under those conditions (Wahyuning & Atiko, 2022). In this study, the target variable is the occurrence of human errors or not (binary).

In this study, machine learning is used to enhance the effectiveness of the prediction model through the modeling process. The modeling process involves several stages, starting from data collection, preprocessing, modeling, and evaluation. The modeling process utilizes classification machine learning algorithms.

DATA COLLECTION

The dataset used in this study consists of measurements of HRV, working environment, job demands, and occurrence of human errors (Lucky, 2022). The collected dataset consists of 270 instances and includes 21 predictor variables and 1 target variable, which is human error. In the target variable, the occurrence of human errors is indicated by the number 1, while the absence of errors is indicated by the number 0 (binary). The assessment of human errors is based on the presence of defects in a production lot. On the other hand, the absence of errors indicates that a production lot is defect-free and in a resting condition.

PREPOCESSING DATA

Data preprocessing is an important stage in the successful implementation of modeling algorithms (García et al., 2015; Cheng et al., 2018; Fernández et al., 2018a). The data preprocessing begins with cleaning the dirty data, which includes handling missing data, correcting erroneous data, and standardizing representations of the same data (García et al., 2015). Additionally, this stage also involves eliminating data collection errors and outliers (Ismail et al., 2021). Outliers are removed from the modeling dataset as they can degrade the performance of machine learning models (Liu et al., 2004; Kao et al., 2017). Control chart techniques are used to detect outliers using 3-sigma rules (Bakar et al., 2006; Kee et al., 2022). As a result of removing outliers, missing values are replaced with the mean value of the attribute or using Bayesian inference (Ismail et al., 2021).

The next step is to balance the data between positive and negative instances of the target variable. The Synthetic Minority Oversampling Technique (SMOTE) is used to balance the data (Fernández et al., 2018b; Asniar et al., 2021). This technique is used for oversampling by generating new synthetic data points for the minority class, which differ from the original instances. This approach helps reduce the impact of overfitting on the minority class (Kovács et al., 2020). SMOTE works by randomly selecting minority data points and creating synthetic data points based on the neighboring data. This technique aids the model in better understanding the minority class and mitigating the bias resulting from imbalanced data. A balanced dataset exhibits distinct value ranges across variables.

Data normalization is used to balance the scale of each feature (attribute/column). The min-max technique is employed to normalize the data based on Equation 1 (García et al., 2015). This technique maps the value range between 0 and 1, making it easily interpretable. This step is crucial for enhancing data quality and the performance of machine learning algorithms (Singh & Singh, 2019).

$$\overline{X} = \frac{X - Xmin}{Xmax - Xmin} \tag{1}$$

Explanation:

$\overline{X}$: Normalized result

Xmax : Maximum value of the variable

X : Original value of x

Xmin : Minimum value of the variable

Xmax : Maximum value of the variable

MODELING AND EVALUATION

The prediction model is built using machine learning algorithms based on Equation 2. This study classifies the occurrence or absence of human errors as an indication of human performance using machine learning techniques, namely DT and RF. The DT algorithm makes decisions like a human,

making it easy to understand in data searching, text extraction, certified medical fields, and search engines (Patel & Prajapati, 2018). The ensemble learning method RF is used for classification. The RF algorithm works by constructing multiple DTs from the training data and combining the results of each DT to make the final decision (Prajwala, 2015). Model validation is performed by dividing the data into 70% for training and 30% for testing (Kavzoglu et al., 2020; Nguyen et al., 2021).

$$Y = \sum_{i=1}^{n} Xi \tag{2}$$

Explanation:

Y : Target variable (human error)

X : Predictor variables of human error,

where i = 1, 2, 3, . . . , n

The prediction model is evaluated using accuracy and recall as evaluation parameters. Accuracy is the ratio of correct predictions (both positive and negative) to the total data (Equation 3) (Vujović, 2021). The higher the accuracy value, the fewer human errors are missed by the model.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{3}$$

Recall is also considered in the evaluation, indicating how many true positive observations are correctly predicted by the algorithm (Vakili et al., 2020). Recall represents the ratio of true positives to the total of true positives and false negatives (Equation 4). The higher the recall value, the more human errors are successfully identified by the model.

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

Explanation:

TP : True Positive

FP : False Positive

FN : False Negative

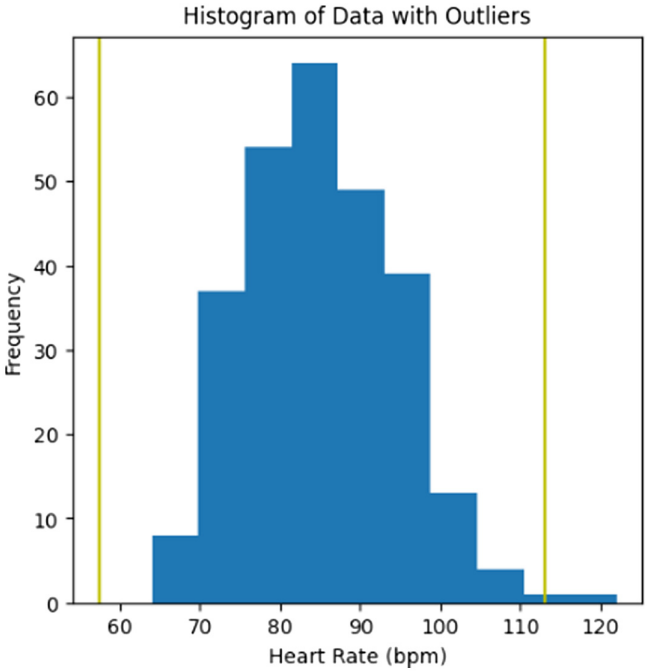
TN : True Negative

Results and Discussion

In the initial stage of data preprocessing, no missing values were found. However, outliers were detected in the dataset, specifically in the HR variable as shown in Figure 1. The outlier data points are removed from the dataset. After removing the outlier data, it was found that there were 55 missing values in 14 out of 21 predictor variables, as indicated in Table 2. The missing values will be replaced with the mean value of the respective variables.

The imbalance data ratio in the target variable is 70%:30%, as shown in Figure 2. The 70%:30% ratio indicates that the majority class accounts for 70% of the total data, while the minority class only accounts for 30% of the total data. Imbalanced data in the target variable can lead to bias in the resulting model, as the model tends to predict the majority class better than the minority class. Therefore, it is necessary to perform up-sampling on the minority class to address the data imbalance issue.

Figure 1.
Histogram of data outlier (heart rate).



SMOTE is used to address the imbalance data issue, resulting in a balanced ratio between the two classes (50:50), as shown in Figure 3. This can enhance the accuracy and performance of the generated model, especially in predicting the minority class. The number of instances also increases to 387. However, there is a significant disparity in the scale of values across the variables.

Data normalization is performed to bring the scale of the variables closer together, based on Equation 2. As a result, all values of variable x are within the range of 0–1. This range of values helps improve the performance of machine learning algorithms and prediction models as it allows the algorithms to work more effectively and efficiently with normalized data (Singh & Singh, 2019).

The application of the DT and RF algorithms yielded excellent accuracy levels, with a high degree of agreement between predicted and actual values of over 90% (Table 3). This indicates that the models are reliable in predicting human errors. The RF algorithm has a higher accuracy and recall rate than the DT algorithm, but both are still suitable for predicting human errors based on stress, working environment, and job demands.

This study successfully predicted human errors using machine learning algorithms. Stress, working environment, and job demands were used as predictor variables, which yielded high accuracy and recall values. The accuracy achieved in this study was lower than a previous study (Damacharla et al., 2019). With an accuracy rate above 90%, the predictive model built in this study remains effective. Additionally, this research utilized predictor variables that directly contribute to the increase in human errors through direct measurements on the human body and the working environment. Previous studies were challenging to apply to manual manufacturing cases as they required computer-based tasks. The predictive model developed in this study is not limited to manual manufacturing industry but can be applied to any job involving human factors.

The prediction results can be influenced by other factors such as the quality and completeness of the data used. Therefore, proper data preprocessing is crucial to generate accurate and reliable results. Factors such as data input errors, incomplete data, or sampling biases can affect the accuracy and

Table 2.
Missing Value Record

Predictor	Total Missing Value Record
Total Power (ms^2)	8
LF/HF	7
Amo50 (%)	6
VLF (ms^2)	6
HF (ms^2)	5
SDNN (ms)	4
CV	4
HF/LF/VLF	4
LF (ms^2)	3
Pencahayaan (Lux)	3
MxDMn (s)	2
RMSSD (ms)	1
Mode (ms)	1
Heart Rate (bpm)	1

Figure 2.
Data imbalance.

Figure 3.
Data balance.

Table 3.
Comparison of Model Evaluation for Algorithm Modeling

Algorithm	Accuracy	Recall
Decision tree	91%	92.86%
Random forest	91.17%	98.21%

validity of the model. Hence, careful assessment of the model's limitations and proper interpretation of the obtained results are necessary.

Conclusion

The utilization of machine learning classification algorithms, namely DT and RF, have demonstrated high accuracy in predicting human error performance. Additionally, this study takes into account the influence of stress, working environment, and job demand as a predictor variable for human error, enhancing the validity and relevance of the predictive model. This research has practical implications in preventing product defects caused by human error. The prediction of human errors provides an initial estimate for taking steps to reduce product defects. Production managers can determine break times for workers based on predictions of the probability of human errors. The study requires expanding the application of the model to other occupations by increasing the dataset, predictor variables, and different types of jobs to improve the validity and generalizability of the predictive model. Therefore, further research can be developed by examining other factors that influence human error in the workplace.

References

Ananda, D., & Wahyuning, C. S. (2022). Strategi pengoptimalan performansi kerja berdasarkan human reliability akibat laju beban kerja mental pada operator mesin jahit ("*Strategy for optimizing work performance based on human reliability due to mental workload rate on sewing machine operators*") (Unpublished thesis). Institut Teknologi Nasional, Bandung -Indonesia.

Asniar, Maulidevi, N. U., & Surendro, K. (2021). SMOTE-LOF for noise identification in imbalanced data classification. *Journal of King Saud University - Computer and Information Sciences*, *34*(6), 3413–3423. https://doi.org/10.1016/j.jksuci.2021.01.014

Attar, E. T., Balasubramanian, V., Subasi, E., & Kaya, M. (2021). Stress analysis based on simultaneous heart rate variability and EEG monitoring. *IEEE Journal of Translational Engineering in Health and Medicine*, *9*(January), 1–7. https://doi.org/10.1109/JTEHM.2021.3106803

Bakar, Z. A., Mohemad, R., Ahmad, A., & Deris, M. M. (2006). *A comparative study for outlier detection techniques in data mining* [conference proceedings]. 2006 IEEE Conference on Cybernetics and Intelligent Systems, Bangkok, Thailand. https://doi.org/10.1109/ICCIS.2006.252287

Bansal, M., Goyal, A., & Choudhary, A. (2022). A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. *Decision Analytics Journal*, *3*(May), 100071. https://doi.org/10.1016/j.dajour.2022.100071

Bubb, H. (2005). Human reliability: A key to improved quality in manufacturing. *Human Factors and Ergonomics in Manufacturing*, *15*(4), 353–368. https://doi.org/10.1002/hfm.20032

Cheng, Y., Chen, K., Sun, H., Zhang, Y., & Tao, F. (2018). Data and knowledge mining with big data towards smart production. *Journal of Industrial Information Integration*, *9*, 1–13. https://doi.org/10.1016/j.jii.2017.08.001

Damacharla, P., Javaid, A. Y., & Devabhaktuni, V. K. (2019). Human Error Prediction Using Eye Tracking to Improvise Team Cohesion in Human-Machine Teams. In: Boring, R. (eds) *Advances in Human Error, Reliability, Resilience, and Performance*. AHFE 2018. Advances in Intelligent Systems and Computing, vol *778*. Springer, Cham. https://doi.org/10.1007/978-3-319-94391-6_5

Dhafr, N., Ahmad, M., Burgess, B., & Canagassababady, S. (2006). Improvement of quality performance in manufacturing organizations by minimization of production defects. *Robotics and Computer-Integrated Manufacturing*, *22*(5–6), 536–542. https://doi.org/10.1016/j.rcim.2005.11.009

Dhillon, B. S. (2013). *Safety and human error in engineering systems*. CRC Press, Taylor & Francis Group.

Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018a). Data level preprocessing methods. In Learning from imbalanced data sets (pp. 79–121). Springer. https://doi.org/10.1007/978-3-319-98074-4_5

Fernández, A., García, S., & Herrera, F. (2018b). SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research*, *61*, 863–905. https://doi.org/10.1613/jair.1.11192

García, S., Luengo, J., & Herrera, F. (2015). Instance Selection. In: *Data Preprocessing in Data Mining*. Intelligent Systems Reference Library, vol *72*. Springer, Cham. https://doi.org/10.1007/978-3-319-10247-4_8

Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: The role of machine learning. *IEEE Access*, *5*, 20590–20616. https://doi.org/10.1109/ACCESS.2017.2756872

Ifrikar, Z. S., Anggawisastra, R., & Tjakraatmadja, J. H. (2006). Teknik Perancangan Sistem Kerja ("*Techniques of work system design*") (2nd ed.). ITB.

Ismail, M., Mostafa, N., & El-assal, A. (2021, March). *Utilization of machine learning techniques for quality monitoring and prediction* [conference paper]. 11th Annual International Conference on Industrial Engineering and Operations Management, Singapore. 4830–4839. https://doi.org/10.46254/AN11.20210838

Järvelin, P. S., Sinikallio, S., & Tarvainen, M. P. (2019). Heart rate variability and occupational stress—Systematic review. *Industrial Health*, *56*(6), 500–511. https://doi.org/10.2486/indhealth.2017-0190

Kao, H. A., Hsieh, Y. S., Chen, C. H., & Lee, J. (2017). Quality prediction modeling for multistage manufacturing based on classification and association rule mining. *MATEC Web of Conferences, 123*. https://doi.org/10.1051/matecconf/201712300029

Kavzoglu, T., Bilucan, F., & Teke, A. (2020, Nov). *Comparison of support vector machines, random forest and decision tree methods for classification of sentinel - 2A image using different band combinations* [conference proceedings]. ACRS 2020: 41st Asian Conference on Remote Sensing, Deqing, China.

Kee, K. K., Yew, S. L. B., Lim, Y. S., Ting, Y. P., & Rashidi, R. (2022). Universal cyber physical system, a prototype for predictive maintenance. *Bulletin of Electrical Engineering and Informatics*, *11*(1), 42–49. https://doi.org/10.11591/eei.v11i1.3216

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018). Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investigation*, *15*(3), 235–245. https://doi.org/10.30773/pi.2017.08.17

Kovács, B., Tinya, F., Németh, C., & Ódor, P. (2020). Unfolding the effects of different forestry treatments on microclimate in oak forests: Results of a 4-yr experiment. *Ecological Applications*, *30*(2), 321–357. https://doi.org/10.1002/eap.2043

Landy, F. J., & Conte, J. M. (2013). *Work in the 21st century: An introduction*. John Wiley and Sons, Inc.

Liu, H., Shah, S., & Jiang, W. (2004). On-line outlier detection and data cleaning. *Computers and Chemical Engineering*, *28*(9), 1635–1647. https://doi.org/10.1016/j.compchemeng.2004.01.009

Lucky, K. (2022). Pengaruh lingkungan kerja fisik terhadap human reliability akibat laju beban kerja pada operator mesin jahit ("*The influence of the physical work environment on human reliability due to workload rate in sewing machine operators*"). (Unpublished thesis). Institut Teknologi Nasional, Bandung -Indonesia

National Institute for Occupational Safety and Health. (2017). *NIOSH generic job stress questionnaire*. Centers for Disease Control and Prevention. Retrieved December 21, 2022, from https://www.cdc.gov/niosh/topics/workorg/detail088.html.

Nguyen, Q. H., Ly, H. B., Ho, L. S., Al-Ansari, N., Van Le, H., Tran, V. Q., Prakash, I., & Pham, B. T. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Mathematical Problems in Engineering, 2021*. https://doi.org/10.1155/2021/4832864

Patel, H. H., & Prajapati, P. (2018). Study and analysis of decision tree based classification algorithms. *International Journal of Computer Sciences and Engineering*, *6*(10), 74–78. https://doi.org/10.26438/ijcse/v6i10.7478

Prajwala, T. R. (2015). A Comparative study on decision tree and random forest using R tool. *International Journal of Advanced Research in Computer and Communication Engineering*, *4*(1). https://doi.org/10.17148/IJARCCE.2015.4142

Rasmussen, M., & Laumann, K. (2018). The evaluation of fatigue as a performance shaping factor in the Petro-HRA method. *Reliability Engineering and System Safety*, *194*, 106187. https://doi.org/10.1016/j.ress.2018.06.015

Rotenberg, S., & McGrath, J. J. (2016). Inter-relation between autonomic and HPA axis activity in children and adolescents. *Biological Psychology*, *117*, 16–25. https://doi.org/10.1016/j.biopsycho.2016.01.015

Saboundji, R. R., & Robert, A. R. (2020, July). *Predicting human errors from gaze and cursor movements* [conference proceedings]. Proceedings of the International Joint Conference on Neural Networks, Glasgow, UK. https://doi.org/10.1109/IJCNN48605.2020.9207189

Saptari, A., Leau, J. X., & Mohamad, N. A. (2015). The effect of time pressure, working position, component bin position and gender on human error in manual assembly line [conference proceedings]. In *2015 International Conference on Industrial Engineering and Operations Management (IEOM)*, Dubai, UAE. (pp. 1–6). IEEE. https://doi.org/10.1109/IEOM.2015.7093793

Saptari, A., Leau, J. X., Ng, P. K., & Mohamad, E. (2014). *2309 Human error and production rate correlation in assembly process of electronics goods* [conference proceedings]. The Proceedings of Design & Systems Conference. The Japan Society of Mechanical Engineers. *24*(0), 2309-1–2309-6. https://doi.org/10.1299/jsmedsd.2014.24._2309-1_

Sarafinjuk, L. A., Syvak, A. V., Radyoha, R. V., Sarafinjuk, P. V., & Matseyko, I. I. (2020). Peculiarities of heart rate variability indicators in track-and-field athletics with mesomorphic somatotype. *World of Medicine and Biology*, *16*(73), 101. https://doi.org/10.26724/2079-8334-2020-3-73-101-106

Scanlon, V. C., & Sanders, T. (2014). *Essentials of anatomy and physiology*. (6th Ed.). F. A. Davis.

Singh, D., & Singh, B. (2019). Investigating the impact of data normalization on classification performance. *Applied Soft Computing, 97*. https://doi.org/10.1016/j.asoc.2019.105524

Sui, S. X., Balanta-Melo, J., Pasco, J. A., & Plotkin, L. I. (2022). Musculoskeletal deficits and cognitive impairment: Epidemiological evidence and biological mechanisms. *Current Osteoporosis Reports*, *20*, 260–272. https://doi.org/10.1007/s11914-022-00736-9

Sumardiyono, S., Wijayanti, R., Hartono, H., & Sri Budiastuti, M. T. (2020). Pengaruh kebisingan terhadap tekanan darah, dengan stres kerja sebagai variabel mediator. *Jurnal Kesehatan Vokasional*, *5*(2), 124. https://doi.org/10.22146/jkesvo.54088

Taelman, J., Vandeput, S., Spaepen, A., & Van Huffel, S. (2008). Influence of mental stress on heart rate and heart rate variability. *IFMBE Proceedings*, *22*, 1366–1369. https://doi.org/10.1007/978-3-540-89208-3_324

Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV: Heart rate variability analysis software. *Computer Methods and Programs in Biomedicine*, *113*(1), 210–220. https://doi.org/10.1016/j.cmpb.2013.07.024

Torres, Y., Nadeau, S., & Landau, K. (2021). Classification and quantification of human error in manufacturing: A case study in complex manual assembly. *Applied Sciences (Switzerland)*, *11*(2), 1–23. https://doi.org/10.3390/app11020749

Vakili, M., Ghamsari, M., & Rezaei, M. (2020). Performance analysis and comparison of machine and deep learning algorithms for iot data classification. arXiv. http://arxiv.org/abs/2001.09636

Vujović, Ž. (2021). Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, *12*(6), 599–606. https://doi.org/10.14569/IJACSA.2021.0120670

Wahyuning, C. S. (2017). Karakteristik kerja, stres, dan beban kerja mental masinis kereta api Indonesia ("*Work Characteristics, Stress, and Mental Workload of Indonesian Train Engineers*"). Institut Teknologi Bandung.

Wahyuning, C. S., & Atiko, S. (2022). Predicting human reliability based on individual's resting period: Effect of physical workload rate. *Jurnal Optimasi Sistem Industri*, *21*(2), 66–74. https://doi.org/10.25077/josi.v21.n2.p66-74.2022

Wahyuning, C. S., & Laksemi, D. B. (2021). Kajian Pengaruh Heat Stress terhadap Beban Kerja Fisik Berat pada Kegiatan Lapangan ("*Study on the Influence of Heat Stress on Heavy Physical Workload*

in Field Activities"). Seminar Nasional Teknik Dan Manajemen Industri, *1*(1), 167–174. https://doi.org/10.28932/sentekmi2021.v1i1.71

Wahyuning, C. S., Rachmatiah, I., & Sutalaksana, I. Z. (2017). Pengaruh lingkungan kerja kabin lokomotif terhadap tingkat stres dan beban kerja mental masinis secara fisiologis. *Jurnal Sistem Teknik Industri*, *19*(1), 20–29. https://doi.org/10.32734/jsti.v19i1.362

Welltory Team. (2023). *Heart rate variability guide by Welltory*. https://welltory.com/heart-rate-variability/

Xing, X., Zhong, B., Luo, H., Rose, T., Li, J., & Antwi-Afari, M. F. (2020). Effects of physical fatigue on the induction of mental fatigue of construction workers: A pilot study based on a neurophysiological approach. *Automation in Construction*, *120*, 103381. https://doi.org/10.1016/j.autcon.2020.103381