PAPER NAME

Springer formatwesad.pdf

AUTHOR

wesad feb2023

WORD COUNT

3271 Words

CHARACTER COUNT

17715 Characters

PAGE COUNT

12 Pages

FILE SIZE

330.0KB

SUBMISSION DATE

Feb 21, 2023 11:18 AM GMT+5:30

REPORT DATE

Feb 21, 2023 11:18 AM GMT+5:30

9% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 6% Internet database
- 7% Publications database
- Crossref database
- Crossref Posted Content database
- 3% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less then 9 words)
- Manually excluded text blocks

ML-Based Stress Classification and Detection Using Multimodal Affect Dataset

Anonymized Authors

Delhi Technological University, Delhi, India

Abstract. Mental stress can pose health risks and affect our daily lives. This study aimed to investigate the use of wearable biosensors for real-time stress detection. The WESAD dataset, which includes bio-signals such as acceleration, respiration, electrodermal activity, ECG, temperature, electromyogram, and blood volume pulse, was analyzed using multiple machine learning models. The results showed that the XGBoost algorithm with hyperparameter tuning produced the best results. The XGBoost algorithm provided the highest accuracy (98.8%) when using a combination of 5 bio-signals (acceleration, EDA, ECG, temperature, and respiration). EDA was found to be the most significant bio-signal for stress detection. The findings highlight the potential of wearable biosensors in detecting stress in real-time environments and could help prevent numerous health problems related to stress by providing early detection and valuable insights.

Keywords: Stress detection, Wearable and Stress Affect Detection (WESAD), 3-axis Acceleration (ACC), Electrodermal activity (EDA), Electrocardiogram (ECG), Electromyogram (EMG), XGBoost

1 Introduction

Stress is the any psycho-physiological or emotional strain in response to a stressor. While ‘acute’ stress or ‘eustress’ is beneficial, positive, manageable and not harmful, ‘chronic’ stress or ‘distress’ can have damaging effects on mental, physical, and emotional health. Chronic stress has become a significant concern worldwide, leading to negative impacts on both individual physical and mental health, as well as on economies and society. The recent COVID-19 pandemic has only exacerbated the problem, with increased stress levels taking a toll on people's well-being. Stress activates the body's flight-or-fight response, releasing a surge of hormones to increase alertness, tense muscles, and heighten blood pressure, to help respond quickly to dangerous situations [1]. However, staying in this heightened sense of arousal for long periods can lead to overexposure to stress hormones, which contributes to health problems like coronary artery disease, cardiac arrest, high blood pressure, memory and concentration impairment, mental health disorders like anxiety, depression, and other physical disorders such as irritable bowel syndrome (IBS), back pain, and gastroesophageal reflux disease (GERD) [2].

Early stress detection can help prevent it from becoming chronic and causing irreparable damage. Traditional methods of stress assessment, such as questionnaires and self-reports, are limited by their subjectivity and potential for bias. Physiological signals, which cannot be controlled voluntarily, are a more reliable measure of stress. The use of mobile health (mHealth) applications and wearables have become increasingly popular for stress detection and management. Wearables can collect and measure physiological signals, such as electrodermal activity, electrocardiogram, and electromyogram, non-invasively and in real-time. This data can be used to measure stress and provide biofeedback to the user.

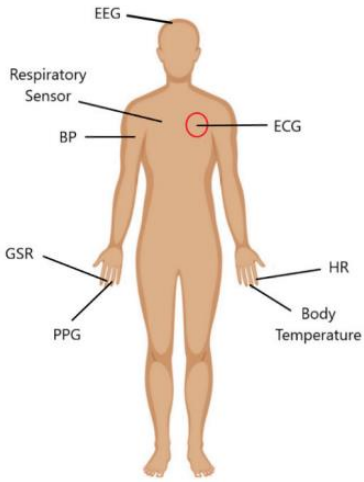

Fig. 1. Schematic diagram showing the different physiological signals collected through wearables

Despite advances in machine learning for stress assessment using bio-signals, challenges remain in obtaining accurate feature points and correctly classifying stress due to fluctuations in signal size and the presence of noise. Some studies have attempted to categorize stress signals using SVM and KNN with varying levels of accuracy. Prasanthi et al. [20] used EMG, GSR, and respiratory signals with an accuracy of 93.65%, but had trouble obtaining correct feature points. Md Fahim Rizwan et al. [21] used ECG, attaining an accuracy of 98.6% by combining RR interval, QT interval, and EDR data with SVM. Still, this result could be better because it depends solely on one signal, the ECG, and ignores other vital bio-signals that are as critical for stress induction.

This research project endeavors to present a novel solution for automatic stress detection by developing simple and efficient Machine Learning algorithms for stress level classification. The algorithms are trained using the WESAD dataset, a comprehensive and multimodal dataset containing data from 15 subjects and various physiological signals collected from wearables. The results of this study also provide insight into the optimal combination of physiological signals for maximum accuracy in stress level classification. The proposed solution is less resource-intensive and quicker to train compared to deep neural networks, making it a more accessible solution for widespread use.

2 Literature Review

In recent years, multiple studies have been conducted on stress detection using machine learning techniques. Physiological signals, such as heart rate variability, cortisol levels, and electrodermal activity, have been integrated to improve the accuracy of stress detection. Different machine learning techniques, such as decision trees, support vector machines, and deep neural networks have been used to model stress levels in various populations. The use of wearable devices, such as wristbands and smartwatches, have been studied for continuous monitoring of stress levels, while self-reported measures like questionnaires and diaries have been found to complement physiological data. However, there is a need for larger, more diverse datasets to better understand stress detection in different populations and contexts. The WESAD dataset, which contains physiological signals from 15 subjects, has become a widely used benchmark in these studies.

Studies on stress detection using machine learning and the WESAD dataset have compared various techniques for accuracy. Garg et al. [4] compared KNN, LDA, Random Forest, AdaBoost, and SVM for binary and three-class classification, with Random Forest performing best with F1-scores of 83.34 and 65.73, respectively. Cosoli et al. [5] evaluated the WESAD dataset with Linear Regression and SVM algorithms using acoustic signals from chest and wrist-worn medical devices, achieving accuracy rates of 75% and 72.62%. Kumar et al. [6] proposed a deep hierarchical model (CNN) for data clustering and emergency alarm triggering with an average accuracy of 87.7% and the highest subject-level accuracy of 96.98%. Bhanushali et al. [7] presented a real-time circuit for chest ECG signal prediction with a Random Forest classifier and 5 low-power time domain features, with an accuracy of 96% and an estimated power consumption of 1.16mW. Simons et al. [8] trained SVM algorithms using data from RespiBAN and Empatica E4 sensors, observing 100% and 99% accuracy for each, respectively. Hsieh et al. [9] developed a feature selection framework for XGBoost-based stress detection using chest and wrist-based EDA signals, with F1 scores of 92.38% and 89.92% and 9 dominant features.

These studies demonstrate the use of different machine learning algorithms and signal processing techniques for stress detection using the WESAD dataset. However, more research is needed to establish the most effective and efficient methods for stress detection.

3 Methodology

3.1 Data Collection

We utilized the data from the publicly available dataset WESAD (Wearable Stress and Affect Detection). WESAD is a publicly available collection of physiological and behavioral data collected from 15 participants while they perform various activities and tasks. The dataset includes data from multiple biometric signals such as:

- Electrocardiogram (ECG): to measure heart rate and rhythm
- Electrodermal Activity (EDA): to measure skin conductance levels
- Respiration: to measure respiration rate and depth
- Motion data: to measure physical activity levels
- Photoplethysmogram (PPG): to measure blood oxygenation levels
- Body Temperature: to measure skin temperature
- Accelerometer and Gyroscope: to measure body movement and posture

These signals are collected using wearable sensors such as ECG and EDA sensors, accelerometers, and temperature sensors. The participants in the dataset performed activities such as resting, stress-induction tasks, and watching videos, allowing researchers to analyze physiological changes in response to different stimuli. The dataset bridges the gap between previous lab studies on stress and emotions by covering three distinct affective states (neutral, stress, and amusement) [3]. Additionally, the dataset includes self-reports from the individuals, obtained through standard questionnaires. The WESAD dataset is widely used in the field of affective computing and stress detection for developing and evaluating algorithms for stress and affect recognition.

Table 1. Dataset Description

Dataset Characteristics	Multivariate, Time Series
Number of Instances	63000000
Attribute Characteristics	Real
Number of Attributes	12
Missing Values?	N/A
Number of Subjects	15

3.2 Data Preprocessing

The WESAD dataset is characterized by its imbalanced nature. As seen in Figure 2, each output state in the dataset has a different number of samples for each label, leading to a greatly unbalanced distribution. Additionally, when comparing the data from the wrist-worn device to that of the chest-worn device, it becomes apparent that the latter provides significantly more samples (4255300 per sensor) compared to the wrist-worn device, which has fewer than 500,000 samples per signal. The data from the chest-worn device is 21 times greater than that of the wrist-worn device. Furthermore, based on the conclusions of the original authors of the WESAD dataset, the physiological data from the chest-worn device alone provides better results than the combination of both devices. Therefore, for our analysis, we utilized only the physiological signals from the chest-worn device.

Fig. 2. Imbalance in WESAD dataset in class and samples

3.3 Feature Extraction

After dropping the wrist data due to its imbalanced distribution, we selected the remaining features from the chest data, including 3-axis acceleration (ACC), respiration (RESP), body temperature (TEMP), electrodermal activity (EDA), electromyography (EMG), and electrocardiogram (ECG).

To extract the most important and relevant information for the model, we performed feature extraction. We used both the Filter method and Wrapper method to identify the significant features. The results showed that the seven features, ACC_Y, ACC_Z, ECG, EMG, EDA, TEMP, and RESP were highly uncorrelated and deemed important by the wrapper method. Therefore, we concluded that all seven features must be used for accurate predictions by the ML model. The feature extraction step played a crucial role in ensuring that the ML model was trained on the most relevant and important information.

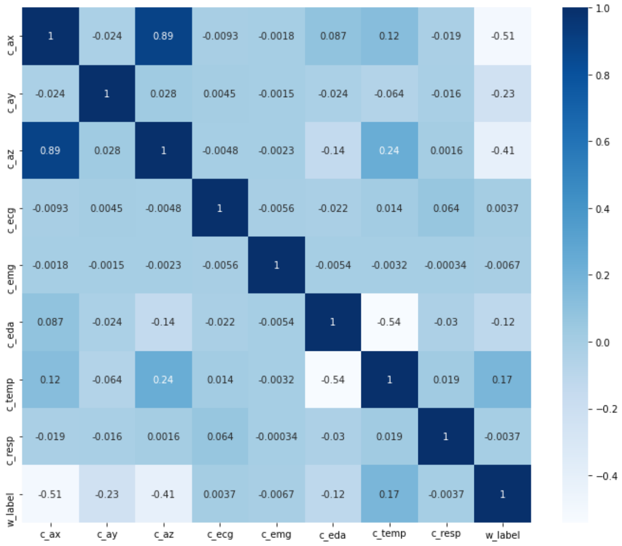

Fig. 3. Correlation Heatmap shows data is highly uncorrelated

3.4 Evaluation Metric

The metrics shown in Table 2 have been used to evaluate our models.

Table 2. Metrics for Model Evaluation

Metrics	Mathematical formula
Accuracy	$\frac{(TP+TN)}{(TP+FP+TN+FN)}$
Precision	$\frac{TP}{(TP+FP)}$
Recall = TPR	$\frac{TP}{(TP+FN)}$
F1 measure	$\frac{2*(Recall*Precision)}{(Recall+Precision)}$

Since the WESAD dataset is highly unbalanced and has multiple labels, the accuracy score has no bearing here. Hence, we have used the F1-scores as the main evaluation metric.

3.5 Classification

In our work, we evaluated the performance of seven machine learning algorithms (Logistic Regression, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Decision Tree, Support Vector Machine, K-Nearest Neighbors, XGBoost) for classifying stress into five levels: neutral, baseline, amusement, stress, and meditation. To achieve optimal results, hyperparameters were tuned for each algorithm and 30 different models were tested on the S2 subject with an 80-20 training-testing data split.

After evaluating the results, we found that the XGBoost model with optimized hyperparameters (learning rate = 0.1, depth = 10, number of estimators = 300, early stopping rounds = 20, gradient based = 0.1, alpha = 10, and gamma = 10) gave the best performance and achieved an impressive AUC score of 0.991. Based on these results, we decided to use this XGBoost model for the remaining subjects. Table 3 provides a comparison of the best results achieved by each classification algorithm.

Table 3. Comparison of the seven classification models on S2 data

Algorithm	**Precision**	**Recall**	**Accuracy**	**F-1 Score**
Logistic Regression	0.82	0.8	0.82	0.81
Linear Discriminant Analysis	0.8	0.74	0.79	0.75
Quadratic Discriminant Analysis	0.88	0.92	0.9	0.9
Decision Tree	0.87	0.86	0.86	0.86
K-Nearest Neighbors	0.88	0.92	0.9	0.9
Support Vector Machine	0.86	0.77	0.85	0.8
Logistic Regression	0.82	0.8	0.82	0.81

3.6 Feature Importance

With the results from XGBoost algorithm on S2, we plotted the F1-score vs. Features graph to identify the significant features. Based on the results from this plot, we ran some more test models to determine how well some individual features or combination of features would perform using the XGBoost algorithm.

Fig. 4. Feature Importance

Table 4. Determining best individual and combination of features on S2

Feature(s)	**Accuracy**	**F1-Score**
ACC, EDA, EMG, ECG, Temp, Resp	**0.99**	**0.99**
ACC, EDA, ECG, Temp, Resp	**0.99**	**0.99**
EMG, EDA, ECG, Temp, Resp	0.90	0.90
ACC, Temp, EDA	0.98	0.98
ACC	0.89	0.89
EDA	0.48	0.50
Temp	0.64	0.68
Resp	0.20	0.44

Based on the results, we found that the combination of all the chest signals (ACC, EDA, EMG, ECG, Temp, Resp) and the combination of 3-axis acceleration, EDA, ECG, temperature and respiration, resulted in excellent performance.

4 Results and Discussion

The results of our test runs on the S2 subject showed that the XGBoost algorithm with hyperparameter tuning (learning rate=0.1, depth=10, number of estimators=300, early stopping rounds=20, gradient based=0.1, alpha=10, and gamma=10) produced the best results. Additionally, we found the combination of signals that yielded the best results. To validate the performance of our proposed XGBoost solution, we split the data of each subject into an 80-20 ratio for training and testing, respectively. The results of these tests are summarized in Table 5, which displays the average accuracy and F1-score of the model.

Our XGBoost model using all chest signals (ACC, EDA, EMG, ECG, Temp, Resp) resulted in an average F1-score of 0.98 and an average accuracy of 98.2%. The highest F1-score of 0.994 was achieved for Subject 3, and the accuracy of the model ranged from 98% to 99.5%.

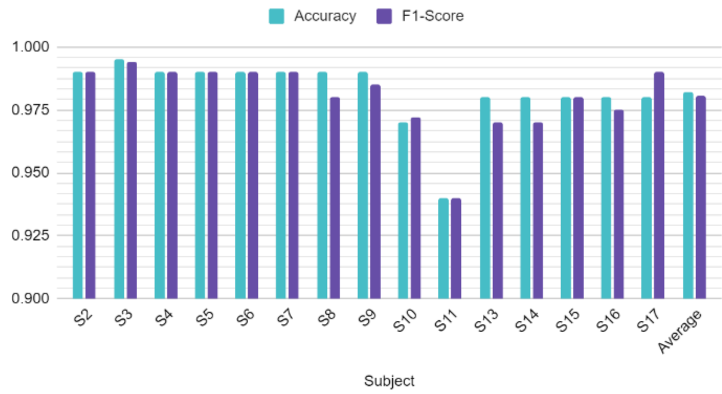

Fig. 5. Performance of proposed model per subject using all signals

When using the combination of 3-axis acceleration, EDA, ECG, temperature and respiration, the XGBoost model produced an average F1-score of 0.987 and an average accuracy of 98.8%. The accuracy of the model ranged from 98% to 99.6%. Again, the best F1-score of 0.996 was achieved for Subject 3.

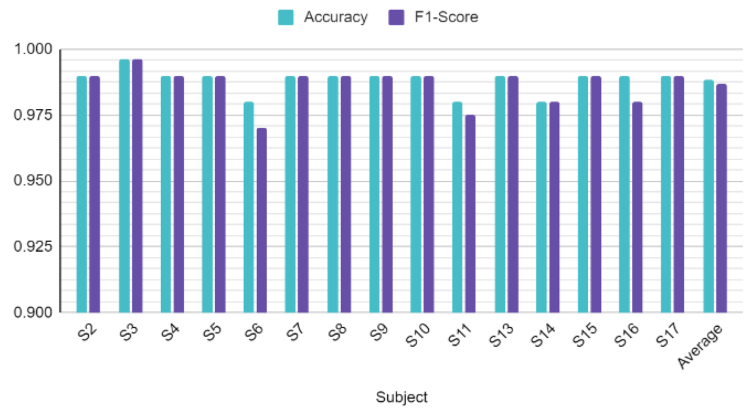

Fig. 6. Performance of proposed model per subject using the combination of 3-axis acceleration, EDA, ECG, temperature and respiration physiological signals

Table 5. Performance of proposed model

Subject	All Signals		ACC, EDA, ECG, Temp, Resp Signals	
	Accuracy	**F1-Score**	**Accuracy**	**F1-Score**
S2	0.99	0.99	0.99	0.99
S3	**0.995**	**0.994**	**0.996**	**0.996**
S4	0.99	0.99	0.99	0.99
S5	0.99	0.99	0.99	0.99
S6	0.99	0.99	0.98	0.97
S7	0.99	0.99	0.99	0.99
S8	0.99	0.98	0.99	0.99
S9	0.99	0.99	0.99	0.99
S10	0.97	0.97	0.99	0.99
S11	0.94	0.94	0.98	0.98
S13	0.98	0.97	0.99	0.99
S14	0.98	0.97	0.98	0.98
S15	0.98	0.98	0.99	0.99
S16	0.98	0.98	0.99	0.98
S17	0.98	0.99	0.99	0.99
Average	**0.982**	**0.98**	**0.988**	**0.987**

In the previous work on the WESAD dataset, Bhanushali et al. [8] achieved an accuracy of 96% with ECG, Hosseini et al. [14] achieved a good accuracy of 97.03%, and Nigam et al. [10] observed that ACC, TEMP and EDA were the most important features and achieved 98% accuracy using this combination of signals with their model as seen in Table 6. The proposed framework shows promising results and effectiveness compared to state-of-the-art methods and outperforms the existing methods in terms of accuracy.

Table 6. Comparison of results of proposed model with existing work on WESAD dataset

	Signals	**Device**	**Accuracy**
Bhanushali *et al.* [8] 2020	ECG	RespiBAN	96
Kumar *et al.* [6] 2021	ACC, RESP, ECG, EMG, EDA, TEMP, BVP, IBI, HR	Empatica E4 RespiBAN	87.7
Nigam *et al.* [13] 2021	ACC, TEMP, EDA	RespiBAN	98
Hosseini *et al.* [14] 2022	EDA	Empatica E4	97.03
Proposed Model	ACC, EDA, ECG, TEMP, RESP	RespiBAN	**98.8**

5 Conclusion

In conclusion, this research project aimed to explore the potential of wearable biosensors in detecting affective states, specifically stress, in real-time environments. The study utilized the WESAD dataset, which contained data from various bio-signals such as acceleration, respiration, electrodermal activity, electrocardiogram, body temperature, electromyogram, and blood volume pulse. The WESAD dataset was analyzed and multiple machine learning classification models were applied to determine the best approach. Results showed that XGBoost algorithm provided the best results with an average F1-score of 0.987 and an average accuracy of 98.8% when using a combination of 3-axis acceleration, EDA, ECG, temperature, and respiration, indicating a promising improvement in comparison with state-of-the-art research work.

This study serves as a foundation for future research in the field of affective state recognition through wearable biosensors. To enhance the results further, future work can involve incorporating self-reports, facial cues, audio/video recordings, and other modalities into the model. The development of a user-friendly platform that leverages this data could provide valuable insights into stress and mental health for individuals and healthcare professionals alike.

References

1. Team, S. C. (2022, April 26). *Stress statistics: How many people are affected in the U.S.?* The Checkup. Retrieved December 8, 2022, from https://www.singlecare.com/blog/news/stress-statistics/
2. Anum Asif, Muhammad Majid, Syed Muhammad Anwar, Human stress classification using EEG signals in response to music tracks, Computers in Biology and Medicine, Volume 107, 2019, Pages 182-196, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2019.02.015
3. Philip Schmidt, Attila Reiss, Robert Duerichen, Claus Marberger, and Kristof Van Laerhoven. 2018. Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection. In Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI '18). Association for Computing Machinery, New York, NY, USA, 400–408. https://doi.org/10.1145/3242969.3242985
4. Prerna Garg, Jayasankar Santhosh, Andreas Dengel, and Shoya Ishimaru. 2021. Stress Detection by Machine Learning and Wearable Sensors. In 26th International Conference on Intelligent User Interfaces - Companion (IUI '21 Companion). Association for Computing Machinery, New York, NY, USA, 43–45. https://doi.org/10.1145/3397482.3450732
5. Cosoli, G., Poli, A., Scalise, L., & Spinsante, S. (2021). Measurement of multimodal physiological signals for stimulation detection by wearable devices. *Measurement*, *184*, 109966. https://doi.org/10.1016/j.measurement.2021.109966
6. Akshi Kumar, Kapil Sharma, Aditi Sharma, Genetically optimized Fuzzy C-means data clustering of IoMT-based biomarkers for fast affective state recognition in intelligent edge analytics, Applied Soft Computing, Volume 109, 2021, 107525, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2021.107525.

7. S. Prashant Bhanushali, S. Sadasivuni, I. Banerjee and A. Sanyal, "Digital Machine Learning Circuit for Real-Time Stress Detection from Wearable ECG Sensor," 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS), 2020, pp. 978-981, doi: 10.1109/MWSCAS48704.2020.9184466.
8. A. Simons, T. Doyle, D. Musson and J. Reilly, "Impact of Physiological Sensor Variance on Machine Learning Algorithms," 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020, pp. 241-247, doi: 10.1109/SMC42975.2020.9282912
9. C. -P. Hsieh, Y. -T. Chen, W. -K. Beh and A. -Y. A. Wu, "Feature Selection Framework for XGBoost Based on Electrodermal Activity in Stress Detection," 2019 IEEE International Workshop on Signal Processing Systems (SiPS), 2019, pp. 330-335, doi: 10.1109/SiPS47522.2019.9020321.
10. Kushagra Nigam, Kirti Godani, Deepshi Sharma, Shikha Jain, “An Improved Approach for Stress Detection Using Physiological Signals,” 2021, SIS, EAI, doi: 10.4108/eai.14-5-2021.169919
11. E. Hosseini et al., "A Low Cost EDA-based Stress Detection Using Machine Learning," 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 2022, pp. 2619-2623, doi: 10.1109/BIBM55620.2022.999509

9% Overall Similarity

Top sources found in the following databases:

- 6% Internet database
- 7% Publications database
- Crossref database
- Crossref Posted Content database
- 3% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	pe.org.pl Internet	2%
2	Elahe Hosseini, Ruijie Fang, Ruoyu Zhang, Anna Parenteau et al. "A Lo... Crossref	<1%
3	ijsret.com Internet	<1%
4	dfki.de Internet	<1%
5	Gilles Bélanger, John R. Walsh, John E. Richards, Paul H. Milburn, Nour... Crossref	<1%
6	Syracuse University on 2019-12-09 Submitted works	<1%
7	"Intelligent Computing and Innovation on Data Science", Springer Scien... Crossref	<1%
8	igi-global.com Internet	<1%

9	github.com Internet	<1%
10	Moudy Sharaf Alshareef, Badraddin Alturki, Mona Jaber. "A transforme... Crossref	<1%
11	National College of Ireland on 2022-05-09 Submitted works	<1%
12	Pramod Bobade, Vani M.. "Stress Detection with Machine Learning and... Crossref	<1%
13	Taeho Yoon, Jaewook Lee, Woojin Lee. "Joint Transfer of Model Knowl... Crossref	<1%
14	Dhananjai Bajpai, Lili He. "Evaluating KNN Performance on WESAD Dat... Crossref	<1%
15	Grazia Iadarola, Angelica Poli, Susanna Spinsante. "Reconstruction of ... Crossref	<1%
16	Liberty University on 2022-04-11 Submitted works	<1%
17	core.ac.uk Internet	<1%
18	deepai.org Internet	<1%
19	researchmgt.monash.edu Internet	<1%

Excluded from Similarity Report

- Bibliographic material
- Cited material
- Manually excluded text blocks
- Quoted material
- Small Matches (Less then 9 words)

EXCLUDED TEXT BLOCKS

an average F1-score of 0

www.scilit.net

axis acceleration (ACC), respiration(RESP

scholar.smu.edu

ACC_Y,ACC_Z, ECG, EMG, EDA, TEMP

openportal.isti.cnr.it