Privacy-Preserving Stress Detection Using Smartwatch Health Data

Lucas Lange[1], Borislav Degenkolb[2] and Erhard Rahm[3]

Abstract: We present the first privacy-preserving approach for stress detection from wrist-worn wearables based on the Time-Series Classification Transformer (TSCT) architecture and incorporating Differential Privacy (DP) to ensure provable privacy guarantees. The non-private baseline results prove the TSCT to be an effective model for the given task. Our DP experiments then show that the private models suffer from reduced utility but can still be used for reliable stress detection depending on the application. Our proposed approach has potential applications in smart health, where it can be used to monitor smartwatch users' stress levels without compromising their privacy and provide timely interventions or suggestions to prevent adverse health outcomes. Another primary contribution is our evaluation, which studies and shows negative effects of DP regarding model training. The results of this work provide perspectives for future research and applications whenever the fields of stress detection and data privacy intervene.

Keywords: Stress Detection, Stress Recognition, Smartwatch, Time-Series Classification Transformer, Differential Privacy, Privacy-Preserving Machine Learning

1 Introduction

Stress is a prevalent issue that can have negative effects on both physical and mental health, leading to various health problems, such as cardiovascular disease, depression, and anxiety [Ya17]. Therefore, early detection is crucial for effective stress management and reduction, which can improve the quality of life for affected individuals.

In recent years, wearable devices such as smartwatches have emerged as a promising tool for monitoring stress levels, as they can collect various physiological signals such as heart rate, skin conductance, and accelerometer data, among others [Gi22a]. Machine learning techniques are widely used to analyze this data and detect patterns associated with stress [Sc18a]. However, the use of sensitive health data raises concerns about user privacy [Ja22; PZ18]. Through providing a provable privacy guarantee, users can benefit from health reports without stressing over privacy concerns.

In this work, we present the first approach to stress detection using smartwatch health data that preserves user privacy through Differential Privacy (DP) [Dw06]. We employ

[1] Leipzig University, Database Group, Augustusplatz 10, 04109 Leipzig, lange@informatik.uni-leipzig.de
[2] Leipzig University, Augustusplatz 10, 04109 Leipzig, bd79howo@studserv.uni-leipzig.de
[3] Leipzig University, Database Group, Augustusplatz 10, 04109 Leipzig, rahm@informatik.uni-leipzig.de

a transformer-based architecture that achieves high accuracy in detecting stress levels from collected time-series data. The used WESAD dataset by Schmidt et al. [Sc18b] is a common standard dataset for this task (e.g. in related work presented in Sect. 3). Our privacy-preserving approach ensures that sensitive health information is protected while still allowing accurate predictions. A primary focus and contribution of our experimental evaluation are the provided insights into numerous drawbacks when training with DP. We thereby contribute to the growing body of research on privacy-preserving machine learning and its potential applications in healthcare.

Sect. 2 provides background information on relevant concepts. We then give an overview of related work in the field of stress detection in Sect. 3. Sect. 4 details our proposed approach and describes the conducted experiments. Their evaluation and results are then presented in Sect. 5. Finally, Sect. 6 summarizes the main contributions, discusses potentials and limitations, and suggests areas for future work.

2 Background

The following section goes over some of the fundamental concepts used in this work.

2.1 Stress detection

Stress detection, also known as stress recognition, is the process of detecting and identifying patterns in physiological signals that are associated with stress. Physiological signals such as heart rate variability, skin conductance, and respiration can provide valuable information about a person's stress level [Gi22a].

Machine learning techniques can be used to analyze these physiological signals and detect patterns associated with stress [Sc18a]. Wearable devices such as smartwatches have become increasingly popular for monitoring physiological signals related to stress and can collect data continuously throughout the day, providing a rich source of data for detection algorithms.

2.2 Time-Series Classification Transformers

Transformers as introduced by Vaswani et al. [Va17], are a type of neural network architecture that has shown excellent performance in natural language processing tasks. They use a self-attention mechanism to weigh the importance of different parts of the input sequence when making predictions.

Time series transformers extend the transformer architecture to handle time-series data [We23]. They incorporate temporal information into the self-attention mechanism by

operating on a sliding window of fixed length that moves along the time axis. This allows them to better handle long-term dependencies and more efficiently train on longer sequences [We23]. Time-Series Classification Transformers (TSCT) are a specific type of time series transformer architecture designed for classification problems. It has no decoder part and is better suited for our classification-based task of stress detection in physiological signals.

2.3 Differential Privacy

The concept of DP [Dw06] is a mathematical definition of privacy that ensures that the inclusion or exclusion of any individual's data in a dataset does not significantly change the output of statistical queries on that dataset. The framework is based on the idea of adding noise to query answers (including sensitive person-related information) to protect individual privacy while preserving the utility of the dataset for analysis.

Formally, an algorithm A training on a set S is called (ε,δ)-differentially-private, if for all datasets D and D' that differ by exactly one record:

$$Pr[A(D) \in S] \leq e^{\varepsilon} Pr[A(D') \in S] + \delta \tag{1}$$

The ε-parameter measures the level of privacy protection provided by the mechanism. It thus determines the amount of random noise that is added to the output of queries on the dataset to protect individual privacy. The smaller the value of ε, the stronger the privacy protection, but also the greater the amount of noise.

2.4 Differentially Private Stochastic Gradient Descent

Differentially Private Stochastic Gradient Descent (DP-SGD) [Ab16] is a variant of the stochastic gradient descent optimization algorithm that provides differential privacy guarantees. DP-SGD works by adding controlled noise to the gradients computed on each mini-batch of data during the training process. The amount of noise added is controlled by the DP privacy parameter ε, which determines the strength of privacy protection. Tuning the privacy parameter can be challenging, and careful selection and calibration of the noise level are required to achieve the desired privacy-utility trade-off.

3 Related Work

Out of the numerous publications on stress detection (e.g. as reviewed in [Sc18a]), we want to evaluate approaches from works also using the WESAD dataset by Schmidt et al. [Sc18b]. Further, we focus on solely using sensors available to wearable wrist devices like

Related work	Signals	Method	Accuracy	F1-score	ε
Schmidt et al. [Sc18b]	chest+wrist	LDA	92.28	90.74	∞
	wrist	Random Forest	87.12 ± 0.24	84.11 ± 0.31	∞
Siirtola [Si19]	wrist	LDA	87.40 ± 10.4	N/A	∞
Gil-Martin et al. [Gi22b]	chest+wrist	CNN	96.62 ± 0.11	96.63 ± 0.11	∞
	wrist	CNN	92.70 ± 0.16	92.55 ± 0.16	∞
Lange et al. (ours)	wrist	Transformer	91.89 ± 0.15	91.61 ± 0.34	∞
	wrist	DP-Transformer	78.88 ± 4.32	76.14 ± 4.27	10
	wrist	DP-Transformer	78.16 ± 8.40	71.26 ± 4.42	1
	wrist	DP-Transformer	71.15 ± 10.5	68.71 ± 5.70	0.1

Tab. 1: Classifier evaluation on the WESAD dataset for modalities collected from either chest or wrist devices regarding the binary (stress vs. non-stress) classification task. We compare accuracy (%) and F1-score (%) and include our achieved models on their ε-settings from DP.

smartwatches as per our goal, and compare performance for the binary task of classifying stress vs. non-stress.

In Tab. 1 we present the collected related works and also include our results in the overview. Approaches conforming to our criteria use the signals designated under the collective term *wrist*, while *chest+wrist* additionally uses the chest wearable data also available in the WESAD dataset. Currently, the best performing approach from related work for both signal collections is a CNN by Gil-Martin et al. [Gi22b]. It also reigns over the other models when omitting chest signals, where we see about 4% reduction in both its accuracy and its F1-score to 92.70% and 92.55%, respectively. Our results are included for comparison but evaluated later in Sect. 5.

We also find transformer models in related work, which however perform worse than the CNN approach and do not conform to our goals: transformer-based model on ECG [Be21] without wrist data, transformer-based model (Husformer) on wrist and chest signals [Wa23] without binary task. In summary, there are no prior transformer stress detection models solely based on wrist device data for our task, which is a gap we want to close.

We further find only one private approach for stress detection by Can; Ersoy [CE21]. However, they employ federated learning without additional privacy implementations, which provides only limited privacy while processing data [Bo23] and also leaves the resulting classification model open to attacks [Sh17]. With DP, we can guarantee provable privacy for our classification model, thus making this work a valuable contribution [DR+14].

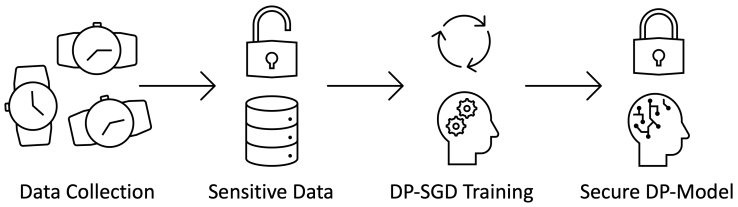

Fig. 1: Illustration of our DP-ensuring machine learning workflow. Instead of privatizing the data itself, we utilize DP-SGD training to directly create a private classification model ready for publishing.

4 Experimental Setup

In this section, we outline our study and give background on the experiments. Reference code is available from our GitHub repository: `https://github.com/BDegenkolb/Privacy-Preserving-Stress-Transformer`.

4.1 Methodology

The general scenario for this work is presented as a basic workflow illustration in Fig. 1. To provide smart health services, a trusted institution or data owner collects the sensitive health data from their users' smartwatches to form a unified dataset for analysis. This data is then funneled into training machine learning models, like in our case for stress detection. To provide good results to all users, especially new device owners, it is important to achieve a generalizing model and this one-for-all approach is therefore favorable compared to creating only a personal model for each user. The final models are later delivered back to the devices for operation, leading to privacy concerns regarding the sensitive training data. A provable protection through DP is achieved by applying the DP-SGD training algorithm. Incorporating this workflow, the final private model is then safe for publishing and can provide health insights at a decreased privacy risk.

Our main objective is to investigate the feasibility of using transformer-based architectures for stress detection using smartwatch health data while preserving user privacy through DP. To evaluate the addition of DP, we first establish a baseline classification model for comparison and as a starting point for our private model. In our baseline performance we want to compete with the current best model by Gil-Martin et al. [Gi22b] (for more related work see Sect. 3). We choose a transformer model instead of CNN because of its capability of capturing long distance dependencies in time-series data [Li22]. On the basis of this baseline, we then evaluate the consequences of using DP regarding model performance.

Our private approach takes the existing baseline model and applies DP-SGD to train for DP. To better quantify the performance loss from DP, we train multiple models for different privacy levels of $\varepsilon = \{10, 1, 0.1\}$. As a non-private model, the baseline is designated with

the privacy level $\varepsilon = \infty$. Values of $\varepsilon \leq 1$ are a common choice for good privacy [Ca19b; Na21] and through also including ε-values a magnitude larger and smaller than $\varepsilon = 1$, we can better analyze the utility-privacy trade-off regarding our task [La23]. Our privacy accounting needs specific considerations due to the time series data and subject-oriented evaluation, which are given in Sect. 4.6.

4.2 Environment

This section details the software and hardware environment in which all experiments were implemented and run. Our description ensures reproducible results, which is also achieved by useing a random seed set to a fixed number of 42 to get comparable training results.

Software-wise, we employ the Python programming language with machine learning workflows implemented using the Tensorflow[4] and Keras[5] libraries. For DP-SGD and other privacy-related implementations, we use Tensorflow Privacy[6]. For running the experiments we relied on Google Colab[7], which enabled us to use a NVIDIA V100 Tensor GPU and 25 GB of system memory for our Jupyter notebooks.

4.3 Dataset

We study the performance of our proposed approach on the publicly available multimodal WESAD dataset [Sc18b], which is the standard dataset for our task. This dataset contains 15 healthy subjects (12 male, 3 female) with about 30 minutes of recorded health data each, which was acquired during a lab study. The physiological and motion data is sampled continuously and simultaneously from a wrist- and a chest-worn device, with the Empatica E4 wrist device offering the following modalities in different sampling frequencies: blood volume pulse (BVP), electrodermal activity (EDA), body temperature (TEMP), and three-axis acceleration (ACC). The three affective states recorded are stress, neutral, and amusement.

We focus on binary classification as stress vs. non-stress and only consider signals available from a wrist-worn wearable like smartwatches—i.e. the Empatica E4 device modalities from the WESAD dataset.

4.4 Preprocessing

An important prerequisite for working with continuous time-series data is signal preprocessing. Our methods are based on prior work by [Gi22b], who presented a throughout process.

[4] Available at `https://github.com/tensorflow/tensorflow`
[5] Available at `https://github.com/keras-team/keras`
[6] Available at `https://github.com/tensorflow/privacy`
[7] Available at `https://colab.research.google.com`

Signal	Sampling frequency	Frequency range	Subwindow length	Model inputs
ACC	64 Hz	0–30 Hz	7 seconds	210
BVP	64 Hz	0–7 Hz	30 seconds	210
EDA	64 Hz	0–7 Hz	30 seconds	210
TEMP	64 Hz	0–6 Hz	35 seconds	210

Tab. 2: Processing details for each signal sampled by the Empatica E4, based on [Gi22b].

Block	Layer	Settings	Data shape
Input	Input layer	N/A	(6, 210)
Encoder (x8)	Layer normalization	epsilon 1e-6	(6, 210)
	Multi head attention	256 head size, 4 heads	(6, 210)
	Dropout	0.25	(6, 210)
	Add	residual connection	(6, 210)
	Layer normalization	epsilon 1e-6	(6, 210)
	Convolution 1D	4 filters of 1x1, ReLU	(6, 4)
	Dropout	0.25	(6, 4)
	Convolution 1D	6 filters of 1x1, linear	(6, 210)
	Add	residual connection	(6, 210)
Output	Avg-Pooling 1D	reduce feature space	(6)
	Fully connected	128 units, ReLU	(128)
	Dropout	0.25	(128)
	Fully connected	2 units, sigmoid	(2)

Tab. 3: Detailed view of our TSCT architecture based on the basic model introduced by Keras [Nt]. Data shape is given as (features, sequence length), which is a ordering we found to be desirable.

In general, our preprocessing workflow is divided into three steps: 1. upsampling (Fourier method), 2. segmentation (windows), and 3. Fast Fourier Transformation (FFT).

The Empatica E4 wristband collects its modalities in differing heterogeneous sampling rates, which for consistency are upsampled to 64 Hz for all signal types using the Fourier method. This adjustment is important to have the signals available at equal time intervals when inputting in our transformer architecture. Aferwards, the data streams are segmented into 60-second windows and subdivided into different length P-second subwindows with a shift of 0.25 seconds. The subwindow length depends on the signal's frequency range and is used to average the 60-second window's spectrum (Fourier transform module) along the subwindows using FFT. This signal-specific spectrum transformation allows to gain a consistent 210 frequency points per 60-second window. These 210 frequency points represent the input for our machine learning model, which then classifies the data at the 60-second window level. We give an overview of preprocessing settings for each signal in Tab. 2. The overall preprocessing result for each subject in the dataset are time series inputs with 6 signals and a sequence length of 210 frequency points per signal.

4.5 None-Private Baseline

The basic structure of our TSCT architecture is shown in Tab. 3 and based on the Keras time series transformer [Nt], where the projection layers are implemented as 1D Convolutions. The model underwent a rigorous hyperparamter tuning process delivering the presented model architecture and settings. We found an optimum at 8 stacked encoder blocks inside the model. We train for 110 epochs with a batch size of 50 and a learning rate of 1e-4. Our input shape takes the form of (features, sequence length), with features being the 6 available signals and 210 points of measurement each. While this is an unusual ordering compared to (sequence length, features), we found this setup to produce better classification results.

For effectively measuring performance in a small dataset of different patients, the evaluation is subject-based using the Leave One Subject Out (LOSO) method. We cycle through the 15 subjects and in each step take one subject as test and the other 14 as training sets for our machine learning model. The overall metrics are then averaged over all 15 single test results. The LOSO method is also used in related work which helps our comparability, especially to the CNN model by [Gi22b].

4.6 Privacy-Preserving Approach

We have to define multiple relevant variables when transferring a non-private baseline model to DP. Total epochs, training set size, batch size, and δ all influence the needed noise scale to achieve our desired ε-guarantees. While the batch size is fixed at 50, there are further calculations needed for the other values. Due to our evaluation strategy, we specifically consider the DP for our averaged LOSO model over all subjects, instead of a single subject.

For the relationship between epochs and DP, we take into account how often a single sample from the training set is seen during the whole training, since each appearance increases privacy risk [DR+14]. For us, this single sample is one data point in the time series. One epoch including the data point translates to one appearance and we therefore assume the maximum epochs of one subject. This is due to the implementation of our LOSO method, which means each of our 15 subject taking part in 14 training sets for the other subjects. Thus the subject's data points are involved in 14 training processes with 110 epochs each, leading to a total epoch count of $e_{total} = 14 * 110 = 1,540$.

Regarding our training set size, preprocessing results in multiple windows with 210 data points and 6 measurements each for our 15 subjects. We want privacy on the data point level of our time series, which is called event-level DP [Ca19a; Mi23]. In our setting however, in addition to hiding just single data points, we additionally want to hide the correlations between windows. We therefore only consider one window per subject to be unique, while the others are treated as copies, which is the worst-case regarding their sampling rate [DR+14]. Again, increasing the sampling rate of an item leads to higher noise needs, because the influence of the single sample on the model is rising proportionally.

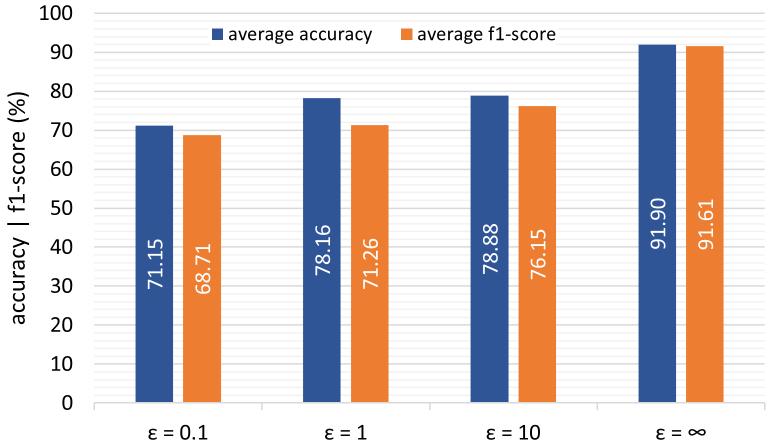

Fig. 2: Average performance over multiple runs each for our non-private and private models (DP).

ε	Accuracy	F1-score	Precision	Recall	Poor models
∞	91.89 ± 0.15	91.61 ± 0.34	91.90 ± 0.42	91.35 ± 0.49	None
10	78.88 ± 4.32	76.14 ± 4.27	77.58 ± 10.1	78.78 ± 16.0	≈ 33%
1	78.16 ± 8.40	71.26 ± 4.42	68.05 ± 10.2	79.09 ± 14.6	≈ 50%
0,1	71.15 ± 10.5	68.71 ± 5.70	60.77 ± 9.30	85.39 ± 18.2	≈ 75%

Tab. 4: Comparison of average model results for non-private and private DP-models. Accuracy, F1-score, precision, and recall are given as %. We also consider their standard deviation. The ratio of poor models tells us how many of the trained model runs failed to reach more than 30% performance.

Meaning, a smaller dataset leads to higher DP noise. Our training set size for determining DP noise is thus given as $n = 210 * 6 * 15 = 18,900$. From that we can also determine $\delta = 1e - 5$ according to $\delta \ll \frac{1}{n}$ [Dw06].

5 Results

To stabilize the results and get a more representative view, we run reach LOSO experiment multiple times and retrieve the best and averaged performance over these runs. On our machine learning task some models either classify all given samples as only stress or only non-stress, thus failing to learn a useful representation of the underlying data. For ease of use we denote these non-representative models as *poor* models in the following. Poor models, in our case, hover around low performance levels of under 30%. When presenting averaged metrics in our evaluation we therefore only consider models achieving more than 30% to not disproportionally decrease the average.

Fig. 2 shows the averaged results for each evaluated privacy level over their tracked experiments. For our non-private performance, we did not manage to outperform the related work by Gil-Martin et al. [Gi22b] but came close with under 1% difference, see Tab. 1 for a full comparison to related work. Our averaged model results show a general trend of decreasing utility when training for better privacy guarantees. The actual reduction in performance from non-private to $\varepsilon = 10$ (max. 15.45%), however, is more significant than inbetween the private models (max. 7.43%). In our case, F1-scores generally takes a bigger hit than accuracy when introducing DP.

In Tab. 4, we present more details in addition to the averages from Fig. 2. We include the recall and precision for each model, which shows that the lowering F1-scores at $\varepsilon = \{1, 0.1\}$ are mostly due to a reduction in precision, while recall keeps higher values. Additionally, the now given standard deviation for our averages is significantly higher when training with DP and for the most part increases further when tightening the ε-setting. We also provide information on the ratio of poor models, i.e. models with less than 30% performance, and see a stern increase when lowering ε for better DP-guarantees. While in non-private training all models were able to achieve feasible results, we recorded between 33–75% failures on $\varepsilon = \{10, 1, 0.1\}$. These results mark the importance of running several rounds of experiments when working with random noise from DP to get an idea of actual performance. We find the induced noise from DP significantly influences model convergence on all levels.

6 Conclusion

We presented the first differentially-private approach to smartwatch stress detection and considering a transformer architecture. While not being able to surpass them, we can however confirm to be close to the related best CNN results from Gil-Martin et al. [Gi22b] regarding our averaged non-private baseline results with under 1% less performance (91.89% accuracy and 91.61% F1-score). The utility-privacy loss recorded for our DP-enabled models lies at over 15%, which is a significant drop to consider. An accuracy of 78.16% and an F1-score of 71.26% at the strong privacy guarantee of $\varepsilon = 1$ is a fair result but the model's feasibility depends on the actual usage scenario. In critical healthcare settings it could be intolerable to lose over 15% utility, while simple lifestyle recommendations from smart health applications on smartwatches might still be reasonable. Regarding the F1-score, we mainly saw a rapid lowering in precision when tightening the DP-guarantees. A low precision score can contribute to increased false positives, which could in the end could annoy users to a point of not using the stress detection feature. We can thus not recommend using very strong privacy guarantees in precision-dependent applications.

A central limitation is the unclear generalizability because of the WESAD dataset characteristics regarding its small size and distribution of gender, age, etc. [Sc18b]. Thus, a highly relevant future venture would be the transition from research to real world studies, where the larger amounts of user data available could also result in better performance for our transformer. In the same vein, generating more training data through synthetization methods

could have a positive impact. Another promising opportunity for future work would be the inclusion of other mobile data sources that could be relevant for stress detection, e.g. mobile phone usage [SP13].

Acknowledgments. We first thank the reviewers for their helpful feedback. We thank Nils Wenzlitschke for his contributions to our preprocessing implementation. The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany and by the Sächsische Staatsministerium für Wissenschaft Kultur und Tourismus in the program Center of Excellence for AI-research "Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig", project identification: ScaDS.AI.

Bibliography

[Ab16] Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H. B.; Mironov, I.; Talwar, K.; Zhang, L.: Deep learning with differential privacy. In: ACM SIGSAC. 2016.

[Be21] Behinaein, B.; Bhatti, A.; Rodenburg, D.; Hungler, P.; Etemad, A.: A Transformer Architecture for Stress Detection from ECG. In: ISWC '21. 2021.

[Bo23] Boenisch, F.; Dziedzic, A.; Schuster, R.; Shamsabadi, A. S.; Shumailov, I.; Papernot, N.: When the Curious Abandon Honesty: Federated Learning Is Not Private, 2023, arXiv: 2112.02918 [cs.LG].

[Ca19a] Cao, Y.; Yoshikawa, M.; Xiao, Y.; Xiong, L.: Quantifying Differential Privacy in Continuous Data Release Under Temporal Correlations. IEEE Transactions on Knowledge and Data Engineering 31/7, 2019.

[Ca19b] Carlini, N.; Liu, C.; Erlingsson, Ú.; Kos, J.; Song, D.: The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks. In: USENIX Security Symposium. Vol. 267, 2019.

[CE21] Can, Y. S.; Ersoy, C.: Privacy-preserving federated deep learning for wearable IoT-based biomedical monitoring. ACM TOIT 21/1, 2021.

[DR+14] Dwork, C.; Roth, A., et al.: The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science 9/3–4, 2014.

[Dw06] Dwork, C.: Differential Privacy. In: Automata, Languages and Programming. Springer Berlin Heidelberg, 2006.

[Gi22a] Giannakakis, G.; Grigoriadis, D.; Giannakaki, K.; Simantiraki, O.; Roniotis, A.; Tsiknakis, M.: Review on Psychological Stress Detection Using Biosignals. IEEE Transactions on Affective Computing 13/01, 2022.

[Gi22b] Gil-Martin, M.; San-Segundo, R.; Mateos, A.; Ferreiros-Lopez, J.: Human Stress Detection With Wearable Sensors Using Convolutional Neural Networks. IEEE Aerospace and Electronic Systems Magazine 37/1, pp. 60–70, 2022.

[Ja22] Jafarlou, S.; Rahmani, A. M.; Dutt, N.; Mousavi, S. R.: ECG Biosignal Deidentification Using Conditional Generative Adversarial Networks. In: Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2022.

[La23] Lange, L.; Schneider, M.; Christen, P.; Rahm, E.: Privacy in Practice: Private COVID-19 Detection in X-Ray Images, 2023, arXiv: 2211.11434 [cs.LG].

[Li22] Li, C.; Huang, X.; Song, R.; Qian, R.; Liu, X.; Chen, X.: EEG-based seizure prediction via Transformer guided CNN. Measurement 203/, 2022.

[Mi23] Miranda-Pascual, À.; Guerra-Balboa, P.; Parra-Arnau, J.; Forné, J.; Strufe, T.: SoK: Differentially Private Publication of Trajectory Data. Proceedings on Privacy Enhancing Technologies 2/, pp. 496–516, 2023.

[Na21] Nasr, M.; Songi, S.; Thakurta, A.; Papernot, N.; Carlin, N.: Adversary instantiation: Lower bounds for differentially private machine learning. In: 2021 IEEE Symposium on security and privacy (SP). IEEE, pp. 866–882, 2021.

[Nt] Ntakouris, T.: Timeseries classification with a Transformer model, https://keras.io/examples/timeseries/timeseries_transformer_classification/, accessed: 05/03/2023.

[PZ18] Perez, A. J.; Zeadally, S.: Privacy Issues and Solutions for Consumer Wearables. IT Professional 20/4, 2018.

[Sc18a] Schmidt, P.; Reiss, A.; Duerichen, R.; Laerhoven, K. V.: Wearable affect and stress recognition: A review, 2018, arXiv: 1811.08854 [cs.HC].

[Sc18b] Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; Van Laerhoven, K.: Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection. In: ACM international conference on multimodal interaction. 2018.

[Sh17] Shokri, R.; Stronati, M.; Song, C.; Shmatikov, V.: Membership Inference Attacks against Machine Learning Models, 2017, arXiv: 1610.05820 [cs.CR].

[Si19] Siirtola, P.: Continuous Stress Detection Using the Sensors of Commercial Smartwatch. In: UbiComp/ISWC '19 Adjunct. 2019.

[SP13] Sano, A.; Picard, R. W.: Stress Recognition Using Wearable Sensors and Mobile Phones. In: Humaine Association Conference on ACII. 2013.

[Va17] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I.: Attention is all you need. NeurIPS 30/, 2017.

[Wa23] Wang, R.; Jo, W.; Zhao, D.; Wang, W.; Yang, B.; Chen, G.; Min, B.-C.: Husformer: A Multi-Modal Transformer for Multi-Modal Human State Recognition, 2023, arXiv: 2209.15182 [cs.HC].

[We23] Wen, Q.; Zhou, T.; Zhang, C.; Chen, W.; Ma, Z.; Yan, J.; Sun, L.: Transformers in Time Series: A Survey, 2023, arXiv: 2202.07125 [cs.LG].

[Ya17] Yaribeygi, H.; Panahi, Y.; Sahraei, H.; Johnston, T. P.; Sahebkar, A.: The impact of stress on body function: A review. EXCLI journal 16/, 2017.