

Unsupervised Detection of Migraine-Related Anomalies in sleep-time Blood Volume Pulse Dynamics Using a Convolutional Autoencoder

Rūta Jankevičiūtė

Department of Electronic Systems
Vilnius Gediminas Technical University
Vilnius, Lithuania
ORCID: 0009-0009-4873-5245

Vytautas Abromavičius

Department of Electronic Systems
Vilnius Gediminas Technical University
Vilnius, Lithuania
ORCID: 0000-0003-1588-6572

Saulius Andruškevičius

Center of Neurology
Vilnius University
Vilnius, Lithuania
ORCID: 0000-0002-4689-4818

Abstract—This study investigates whether sleep-time blood volume pulse (BVP) dynamics exhibit anomalous patterns during migraine days. Rather than framing migraine detection as a supervised classification problem, we adopt an unsupervised anomaly detection approach to identify deviations from patient-specific baseline sleep physiology. Participants were adults (18+ years) who met the criteria of the International Classification of Headache Disorders (ICHD-3) for episodic migraine and reported between 4 and 14 migraine days per month. For each participant, a longitudinal BVP recording containing multiple nights of sleep data was analyzed in conjunction with a migraine diary. Participants with migraine data ($n = 17$) were compared using aggregated anomaly metrics. For most participants, the p99 ratio was higher with the 30 s / 5 s configuration compared to the 60 s / 10 s setup, indicating greater separation between migraine and non-migraine days at finer temporal resolution. The results indicate that migraine days are generally associated with increased anomaly scores in sleep-time BVP signals. In addition, the majority of the participants exhibited a greater sensitivity to migraine-related changes in BVP dynamics. Overall, these findings support the use of a convolutional autoencoder for detecting migraine-associated deviations in sleep-time BVP signal dynamics, particularly in settings where labeled data are limited and inter-subject variability is high.

Index Terms—Convolutional autoencoder, blood volume pulse, Migraine, Sleep-time BVP.

I. INTRODUCTION

MIGRAINE is a complex neurological disorder characterized by significant inter-subject variability and fluctuating physiological manifestations. Traditional clinical monitoring often relies on subjective patient diaries, which are susceptible to recall bias and lack the granularity required for real-time physiological evaluation. Recent advances in wearable technology have facilitated the continuous collection of sleep-time blood volume pulse (BVP) signals, offering a non-invasive window into the behavior of the autonomic nervous system during sleep [1]. However, utilizing these data for migraine detection remains a challenge due to the subtle and highly individualized nature of physiological changes associated with the condition [2], [3].

Existing research predominantly treats migraine detection as a supervised classification task, requiring extensive labeled

datasets that are difficult to obtain in longitudinal studies [4], [5]. Furthermore, population-level models frequently struggle to generalize across different individuals, as a "normal" physiological baseline for one patient may appear anomalous for another. This variability underscores the limitations of feature-engineered supervised models and highlights the need for personalized modeling strategies that can adapt to the unique baseline physiology of each participant.

To address these challenges, this study explores an unsupervised anomaly detection framework based on a one-dimensional convolutional autoencoder. By training the model on patient-specific sleep-time BVP data from non-migraine periods, we aim to characterize migraine-related deviations through reconstruction error rather than predefined migraine classes. This approach avoids training a supervised migraine classifier and instead relies on individualized non-migraine baseline data to identify deviations during migraine days. The study evaluates the feasibility of patient-specific modeling and examines how different temporal resolutions affect anomaly detection performance.

II. RELATED WORK

A. Migraine and Wearable-Based Physiological Monitoring

Wearable devices have enabled continuous monitoring of physiological signals for migraine research. Previous studies using this dataset have explored migraine-related physiological patterns using supervised learning and feature engineering approaches [6], [7], [8]. More broadly, wearable-based migraine and health monitoring studies typically rely on handcrafted features derived from physiological signals such as heart rate or BVP, combined with classical machine learning models [9]. This paradigm is closely aligned with stress detection research, where feature-based supervised models are commonly used [10].

Recent work has explored migraine prediction using wearable-derived physiological and sleep signals [11]. Although these approaches demonstrate feasibility, performance remains variable between individuals, and personalized models often outperform population-level approaches [11], [12]. This

highlights the importance of individualized modeling strategies in migraine research.

B. Anomaly Detection in Physiological Signals

Anomaly detection provides an alternative framework for analyzing physiological data by modeling normal patterns and identifying deviations. Autoencoders are widely used for this purpose, as they learn compact representations of normal data and produce an elevated reconstruction error for anomalous inputs [13]. In wearable health applications, autoencoder-based models have been successfully applied to physiological signals, including stress detection from BVP data [14]. These approaches enable modeling of baseline physiological dynamics without requiring labeled data and have shown strong performance in detecting deviations from normal states. Furthermore, anomaly detection methods have been shown to exhibit improved robustness in cross-context settings compared to supervised models, which often suffer from reduced generalization between datasets and individuals [10]. This makes anomaly detection particularly suitable for highly personalized physiological signals.

C. Positioning of This Work

Despite these advances, anomaly detection has not been widely tested in migraine research. Most existing approaches are based on supervised learning and population-level modeling. In contrast, this work formulates migraine detection as a patient-specific anomaly detection problem, allowing the identification of individualized deviations in the dynamics of BVP during sleep.

III. METHODS

An overview of the proposed pipeline is shown in Fig. 1. The goal is to detect migraine-associated deviations in sleep-time BVP signals using a patient-specific anomaly detection framework.

A. Participants

Participants were adults (18+ years) who met the criteria of the International Classification of Headache Disorders (ICHD-3) [15] for episodic migraine and reported between 4 and 14 days of migraine per month.

Participants were excluded if they had chronic migraine, frequent tension-type headaches, or other primary or secondary headache disorders. Additional exclusion criteria included the use of opioid-based medications, recent migraine preventive treatments, or medications known to affect autonomic function.

All participants underwent a structured clinical assessment by a headache specialist prior to inclusion in the study. A total of 20 participants were included, with substantial variability in the number of days (Table I). On average, participants contributed 22.4 ± 11.7 days of data, with 4.1 ± 2.8 migraine days. Three participants had no recorded migraine days and were excluded from migraine-specific comparisons.

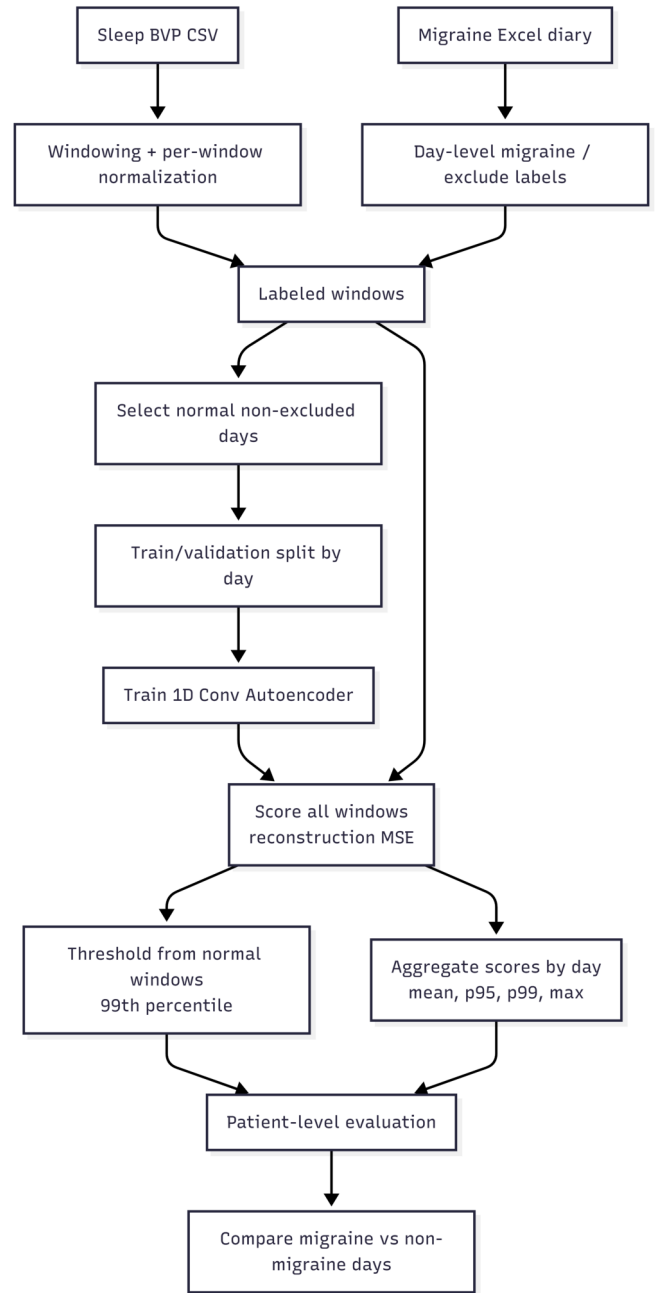


Fig. 1. Overview of the proposed pipeline. Sleep BVP signals are segmented into windows, labeled using migraine diaries, and filtered to retain non-migraine days for training. A patient-specific convolutional autoencoder is trained to model normal physiology. Reconstruction error is aggregated at the day level and compared between migraine and non-migraine days

B. Data and Labeling

For each participant, a longitudinal BVP recording containing multiple nights of sleep data was analyzed in conjunction with a migraine diary. Migraine episodes were converted into daily labels, where any day that contained a migraine episode was labeled as a migraine day. To reduce contamination from postdromal physiological effects, a configurable post-migraine

TABLE I
SUMMARY OF EVALUABLE DATA ACROSS PARTICIPANTS

	Mean $\pm$ Std	Min	Max
Total days	22.4 $\pm$ 11.7	5	46
Migraine days	4.1 $\pm$ 2.8	0	10
Participants with migraine days	17 / 20		

exclusion period of 24 hours was applied. across participants, this corresponded to several thousand windows per individual depending on the recording duration.

C. Preprocessing

sleep-time BVP data was collected using the Empatica EmbracePlus mobile platform. High-frequency raw signals were retrieved in the Avro format. The BVP signal was segmented into overlapping windows (e.g. 30 s window, 5 s stride), and each window was independently normalized using z-score normalization. The windows were timestamped by their start time, enabling alignment with day-level migraine labels. Only windows from non-migraine, non-excluded days were used for model training. Training and validation splits were performed at the day level.

D. Model

A one-dimensional convolutional autoencoder was implemented using the PyTorch deep learning framework. The model was trained separately for each participant using only non-migraine, non-excluded data, with the objective of learning a compact representation of normal sleep-time BVP dynamics. The architecture of the autoencoder is illustrated in Fig. 2.

1) *Input Representation*: Each input sample consisted of a fixed length normalized BVP window L , corresponding to 30 seconds of data sampled at 64 Hz. This results in input tensors of shape $(1, L)$, where the single channel represents the BVP signal. Data preprocessing and window construction were performed using NumPy and pandas, while training and batching were handled using PyTorch Dataset and DataLoader utilities.

2) *Architecture*: The encoder consisted of three one-dimensional convolutional layers with progressively increasing channel depth ($1 \rightarrow 16 \rightarrow 32 \rightarrow 64$). Each layer used a kernel size of 7, stride of 2, and padding of 3. Each convolution was followed by a ReLU activation function. The convolutional output was flattened and passed through two linear layers (with an intermediate dimensionality of 128), forming a low-dimensional latent representation. The decoder mirrored the encoder using transposed convolutional layers ($64 \rightarrow 32 \rightarrow 16 \rightarrow 1$) with matching kernel size, stride, and padding parameters. The decoder progressively reconstructed the signal to its original temporal resolution. The final output had the same shape as the input window.

3) *Training Procedure*: The model was trained to minimize the mean squared error between the input and reconstructed signals. Optimization was performed using the Adam optimizer with a learning rate of 10^{-3} . Training was conducted

for 20 epochs with a mini-batch gradient descent (batch size 256). Training and validation splits were performed at the day level using scikit-learn utilities. Only windows of normal days (migraine = 0, exclude = 0) were used during training.

4) *Anomaly Scoring*: After training, the reconstruction error was computed for each window as the mean squared difference between the input signal and the reconstructed signal. This value served as the anomaly score, with higher values indicating a greater deviation from the learned baseline. Window-level anomaly scores were computed in batches using PyTorch inference and aggregated to the day level using summary statistics (mean, 99th percentile, and maximum). The threshold was not optimized using migraine labels; however, future work should evaluate threshold calibration using strictly held-out baseline data.

E. Evaluation

Migraine and non-migraine days were compared using aggregated anomaly metrics. For each participant, we computed the ratio between migraine-day and non-migraine-day anomaly scores. Ratios greater than 1 indicate higher reconstruction error during migraine days. Tail-based aggregation metrics, such as p99 and maximum, were included because migraine-related physiological deviations during sleep may occur as localized or transient changes rather than uniform shifts throughout the night. The ratio-based metric was chosen to quantify within-subject deviation between migraine and non-migraine states, rather than to optimize classification performance.

Because of the small sample size and the unequal number of migraine days per participant, statistical tests were interpreted as exploratory. We used an exact one-sided binomial test to assess whether the proportion of participants with ratios greater than 1 exceeded chance level. Wilson score 95% confidence intervals were computed for the corresponding participant proportions.

IV. RESULTS

The results were summarized across participants using descriptive statistics, including mean, median, and proportion of participants who exhibited increased anomaly scores during migraine days.

Figure 3 illustrates the day-level anomaly scores for representative participants. Participants 05 and 02 show clear increases in anomaly scores during migraine days, while participants 17 and 08 exhibit weaker or more variable patterns.

Among participants with migraine data ($n = 17$), the majority exhibited higher anomaly scores during migraine days compared to non-migraine days. For the p99 anomaly ratio, 13 out of 17 participants showed higher anomaly scores on migraine days, corresponding to 76.5% of participants (Wilson score 95% CI: 52.7%–90.4%; exact one-sided binomial test, $p = 0.025$). For the maximum anomaly ratio, 11 of 17 participants showed higher migraine-day values, corresponding to 64.7% of the participants (Wilson score 95% CI: 41.3% – 82.7%; $p = 0.166$). Thus, p99 aggregation

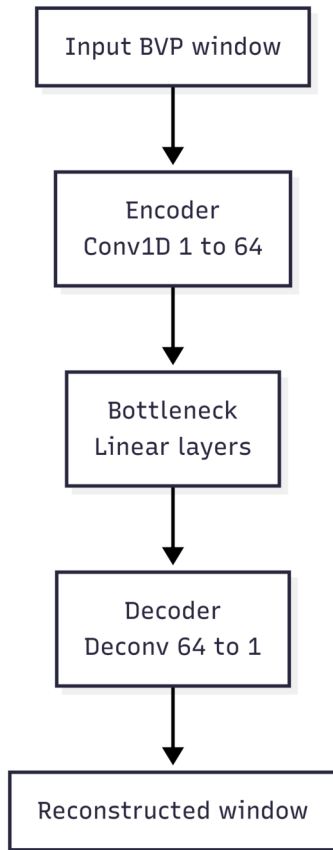


Fig. 2. Architecture of the 1D convolutional autoencoder. The model encodes fixed-length BVP windows into a compressed latent representation and reconstructs the input signal. Reconstruction error is used as an anomaly score

TABLE II
SUMMARY OF ANOMALY SCORE DIFFERENCES BETWEEN MIGRAINE AND NON-MIGRAINE DAYS. RATIOS GREATER THAN 1 INDICATE INCREASED ANOMALY DURING MIGRAINE DAYS

Metric	Mean	Median	% Participants > 1
p99 ratio	1.18	1.13	13 / 17
Max ratio	1.18	1.17	11 / 17

showed stronger participant-level consistency than the maximum anomaly score.

The anomaly scores varied across participants. Some participants showed substantial increases in anomaly ratios (e.g., p99 ratios exceeding 1.5), while others showed minimal differences or ratios below 1. This effect is summarized in Table II.

A. Effect of Segmentation Parameters

The effect of segmentation parameters on anomaly detection performance is summarized in Table III. Increasing the window size and stride from 30 s / 5 s to 60 s / 10 s resulted in higher anomaly ratios across both p99 and maximum aggregation metrics.

The 60 s / 10 s configuration produced the highest mean and median anomaly ratios, indicating stronger separation

TABLE III
EFFECT OF SEGMENTATION PARAMETERS ON ANOMALY RATIOS. RATIOS GREATER THAN 1 INDICATE HIGHER ANOMALY SCORES DURING MIGRAINE DAYS

	30 s / 5 s		60 s / 10 s		120 s / 20 s	
	p99	max	p99	max	p99	max
Mean	1.18	1.18	1.20	1.20	1.06	1.06
Median	1.13	1.17	1.17	1.20	1.05	1.06

between migraine and non-migraine days. In contrast, further increasing window size and stride to 120 s / 20 s resulted in substantially lower anomaly ratios, suggesting reduced sensitivity to migraine-related deviations.

V. DISCUSSION

The results suggest that migraine days are generally associated with increased anomaly scores in sleep-time BVP signals. Because the autoencoder was trained on non-migraine days, an elevated reconstruction error can be interpreted as a deviation from an individual's learned baseline physiological patterns. The observation that most of the participants showed higher p99 anomaly scores on migraine days supports the feasibility of detecting migraine-associated deviations in sleep-time BVP. However, this effect was not uniform across participants and therefore the findings should be interpreted as preliminary evidence of physiological deviation rather than as evidence of diagnostic utility.

Significant variability was observed between participants. Some participants exhibited strong increases in anomaly ratios, while others showed little or no difference. This variability is consistent with the individualized nature of migraine and may reflect differences in migraine phenotype, physiological response, signal quality, or the amount of available baseline data. It also reinforces the relevance of patient-specific modeling approaches.

The statistical evaluation also indicates that the aggregation metric influenced the strength of the observed effect. The p99 anomaly ratio showed a statistically supported participant-level increase, whereas the maximum anomaly ratio did not reach statistical significance. This suggests that p99 may provide a more robust upper-tail summary of reconstruction error, while the maximum score may be more sensitive to isolated artifacts, noise, or unstable signal segments. Therefore, the present results support p99 as the more reliable aggregation metric in this cohort.

The effect of segmentation parameters suggests that migraine-associated deviations may be more clearly reflected at an intermediate temporal resolution in this cohort. The 60 s / 10 s configuration produced the strongest and most consistent anomaly ratios, while smaller windows may have been more sensitive to noise and larger windows may have smoothed localized deviations. These results highlight the importance of temporal resolution when applying anomaly detection to physiological signals, although the optimal segmentation strategy should be validated in larger datasets.

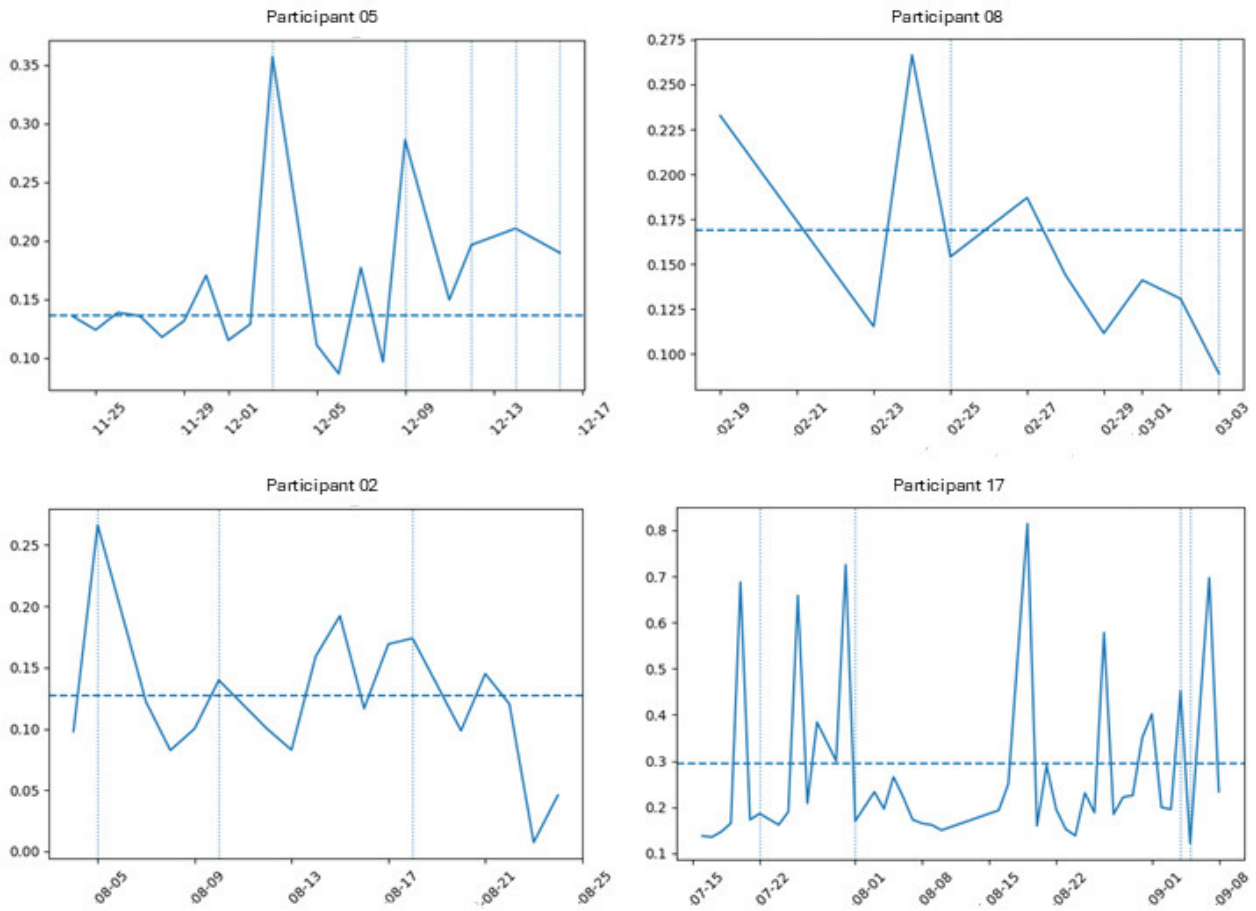


Fig. 3. Day-level anomaly scores (p99 reconstruction error) for four representative participants. The dashed horizontal line indicates the anomaly threshold derived from non-migraine days, while vertical dotted lines mark migraine days. Participants 05 and 02 demonstrate clear anomaly detection patterns; participants 08 and 17 illustrate cases with weaker alignment with the diary labels

The number of usable days varied substantially between the participants, which may have influenced the stability of the anomaly ratios. Participants with fewer non-migraine days provide a less representative baseline for model training, potentially affecting the reliability of reconstruction error estimates. Similarly, a small number of migraine days can lead to less stable comparisons and increased sensitivity to outliers. In addition, this short-paper study did not include comparisons with simpler baselines such as PCA reconstruction, statistical thresholding, or non-convolutional autoencoders. These limitations restrict the generalizability of the results and should be addressed in future work.

In practical terms, the proposed framework should be viewed as a potential support tool for long-term personalized monitoring rather than a standalone clinical method. It could help identify nights with unusual sleep-time BVP dynamics relative to a patient's own baseline, but interpretation would still require diary information, clinical context, and validation in larger and more heterogeneous cohorts.

VI. CONCLUSION

This study demonstrates the feasibility of using patient-specific convolutional autoencoders to identify migraine-associated deviations in sleep-time BVP signals. The results suggest that migraine days are associated with an increase in BVP signal reconstruction error, but the effect is not uniform across participants. The p99 aggregation showed the most consistent participant-level increase, whereas the maximum anomaly score did not provide statistically reliable evidence of separation between migraine and non-migraine days.

These findings support anomaly detection as a promising framework for personalized physiological monitoring in settings with limited labeled data and high inter-subject variability. However, the results remain preliminary due to the limited cohort size, unequal number of migraine and non-migraine days across participants, and absence of simpler baseline comparisons. Future work should evaluate the approach in larger cohorts, compare it with simpler anomaly detection methods, and investigate whether multimodal physiological signals improve sensitivity to migraine-related changes.

VII. ACKNOWLEDGMENTS

During the preparation of the work, the authors used Writeful to refine the use of the English language, correct grammatical errors, and rephrasing sentences for improved clarity. The authors take full responsibility for the content of this publication.

REFERENCES

- [1] S. H. Friend, G. S. Ginsburg, and R. W. Picard, "Wearable digital health technology," pp. 2100–2101, 2023.
- [2] G. Dumkrieger, "The promise of machine learning in predicting migraine attacks," *Cephalalgia*, vol. 45, no. 11, p. 03331024251391207, 2025.
- [3] R. Nosedá and R. Burstein, "Migraine pathophysiology: anatomy of the trigeminovascular pathway and associated neurological symptoms, cortical spreading depression, sensitization, and modulation of pain," *Pain*, vol. 154, pp. S44–S53, 2013.
- [4] K. Mitrović, I. Petrušić, A. Radojčić, M. Daković, and A. Savić, "Migraine with aura detection and subtype classification using machine learning algorithms and morphometric magnetic resonance imaging data," *Frontiers in neurology*, vol. 14, p. 1106612, 2023.
- [5] L. Khan, M. Shahreen, A. Qazi, S. Jamil Ahmed Shah, S. Hussain, and H.-T. Chang, "Migraine headache (mh) classification using machine learning methods with data augmentation," *Scientific Reports*, vol. 14, no. 1, p. 5180, 2024.
- [6] V. Kapustynska, V. Abromavičius, A. Serackis, Š. Paulikas, K. Ryliškienė, and S. Andruškevičius, "Machine Learning and Wearable Technology: Monitoring Changes in Biomedical Signal Patterns during Pre-Migraine Nights," *Healthcare*, vol. 12, no. 17, p. 1701, Jan. 2024, publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: <https://www.mdpi.com/2227-9032/12/17/1701>
- [7] V. Kapustynska, V. Abromavičius, A. Serackis, S. Andruškevičius, K. Ryliškienė, and Š. Paulikas, "Advancing a generalizable model for migraine prediction: Analysis of filtering techniques on physiological signals," *Technology and Health Care*, vol. 33, no. 5, pp. 2184–2193, Sep. 2025, publisher: SAGE Publications. [Online]. Available: <https://doi.org/10.1177/09287329251332415>
- [8] R. Jankevičiūtė, V. Kapustynska, S. Andruškevičius, K. Ryliškienė, and V. Abromavičius, "Heart rate variability as a predictor of migraine: Sleep-time data analysis of pre-migraine nights," *Technology and Health Care: Official Journal of the European Society for Engineering and Medicine*, p. 9287329251412968, Jan. 2026.
- [9] C.-H. Chuang, J.-Y. Li, J.-T. King, W.-T. Chen, S.-P. Chen, Y.-F. Wang, H.-Y. Liu, F.-J. Hsiao, L.-L. H. Pan, S.-J. Wang *et al.*, "Abnormal heart rate variability and its application in predicting treatment efficacy in patients with chronic migraine: An exploratory study," *Cephalalgia*, vol. 43, no. 10, p. 03331024231206781, 2023.
- [10] M. Albaladejo-González, J. A. Ruipérez-Valiente, and F. Gómez Már-mol, "Evaluating different configurations of machine learning models and their transfer learning capabilities for stress detection using heart rate," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 8, pp. 11 011–11 021, Aug. 2023. [Online]. Available: <https://doi.org/10.1007/s12652-022-04365-z>
- [11] L. E. Tomalin, B. R. Kummer, M. C. Campbell, A. N. Erekat, F. Cohen, J. Robinson-Papp, and B. R. Mueller, "Can smartwatches predict mi-graines? Using machine learning (ML) with wearable-derived nocturnal autonomic nervous system (ANS) and sleep metrics for headache prediction," *Research Square*, pp. rs.3.rs–8 195 467, Dec. 2025.
- [12] G. Dumkrieger, "The promise of machine learning in predicting migraine attacks," *Cephalalgia: An International Journal of Headache*, vol. 45, no. 11, p. 3331024251391207, Nov. 2025.
- [13] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
- [14] M. S. Ali, S. S. Bippro, M. A. Motin, S. Kabir, M. Sharma, and M. E. H. Chowdhury, "TEANet: A Transpose-Enhanced Autoencoder Network for Wearable Stress Monitoring," *IEEE Access*, vol. 14, pp. 9560–9571, 2026, arXiv:2503.12657 [eess]. [Online]. Available: <http://arxiv.org/abs/2503.12657>
- [15] Headache Classification Committee of the International Headache Society (IHS), "The international classification of headache disorders, 3rd edition," *Cephalalgia*, vol. 38, no. 1, pp. 1–211, 2018.