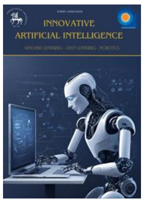
INNAI 1 (1), 2025, 39-46



INNOVATIVE ARTIFICIAL INTELLIGENCE

RESEARCH ARTICLE

Unsupervised Discovery of Affective Physiological States in WESAD Using Hierarchical Agglomerative Clustering

Helen Arya [1]*, Muhammad Arya[2]

[1] Department of Computer Engineering, Dokuz Eylul University 1, Izmir, Turkiye
[2] Faculty of Medicine, Ege University, Izmir, Turkiye

Corresponding Author *: Helenarya19@gmail.com

Abstract

Automated physiological affect recognition is vital, but supervised learning on labeled datasets such as the Wearable Stress and Affect Detection (WESAD) may miss underlying nuances. This study used unsupervised Hierarchical Agglomerative Clustering with Ward's linkage to explore inherent structures in chest-worn sensor data, including electrocardiography, electrodermal activity, respiration, and body temperature, collected from 15 WESAD participants. Comprehensive features, including detailed heart rate variability metrics derived via the NeuroKit2 library, were extracted. Clustering was applied to standardized features, yielding four distinct groups defined by unique multivariate signatures in electrodermal activity, temperature, respiration, and key heart rate variability indices, such as the root mean square of successive differences and the low-frequency/high-frequency ratio. These data-driven groups showed partial alignment but also significant divergence from the original WESAD experimental labels (baseline, stress, amusement, meditation), revealing physiological heterogeneity within predefined conditions. Findings demonstrate the efficacy of hierarchical clustering in identifying physiologically interpretable groups potentially representing distinct autonomic nervous system states, such as stress, relaxation or engagement, and alert rest. This underscores the value of unsupervised learning for complementing supervised methods in affective computing, enabling data-driven discovery of physiological state taxonomies beyond experimental labels, and offering valuable insights for developing nuanced artificial intelligence tools for mental health monitoring and adaptive human-computer interaction.

Keywords: *Affective Computing, Unsupervised Learning, Hierarchical Agglomerative Clustering (HAC), WESAD Dataset, Physiological Signals, Heart Rate Variability (HRV)*

1. Introduction

The automated recognition of human affective states through physiological signals is increasingly vital across diverse Artificial Intelligence (AI) domains, enhancing human-computer interaction, health monitoring, personalized medicine, and mental well-being tools [1]. While behavioral cues offer insights, physiological signals originating from the autonomic nervous system (ANS)such as electrocardiography (ECG), electrodermal activity (EDA), respiration (RESP), and body temperature (TEMP)provide a more objective window into underlying emotional and stress responses, less susceptible to conscious manipulation [2], [3]. The proliferation of wearable sensors makes continuous multi-modal physiological data collection feasible, yet interpreting these complex, noisy, high-dimensional, and temporally dependent streams pose significant challenges. AI, particularly machine learning (ML) and deep learning (DL), offers essential methodologies for processing these signals, extracting features, and identifying patterns indicative of affective states, facilitating a shift towards continuous, ecologically valid monitoring [4].

A critical resource enabling this research is the Wearable Stress and Affect Detection (WESAD) dataset, a publicly available benchmark comprising multi-modal physiological data—including ECG, EDA, electromyography (EMG), RESP, TEMP, blood volume pulse (BVP), and acceleration (ACC)—recorded from chest and wrist devices during experimentally induced neutral, amusement, and stress states. Although WESAD has spurred development of numerous supervised ML/DL models achieving high classification accuracy for these predefined labels, inherent complexities remain, including data dimensionality, noise, artifacts, and significant inter-subject variability. Furthermore, relying solely on supervised learning, which optimizes predicting discrete experimental labels, risks creating models that, while accurate on task, may not fully capture the nuanced spectrum of underlying physiological states or the data's inherent structure [5]. This potential simplification, masked by the high performance of diverse supervised models, necessitates complementary approaches. Consequently, unsupervised learning methods, like clustering, present a valuable alternative for exploring latent structures directly from the data without reliance on a priori label, potentially revealing physiologically meaningful groupings that refine or diverge from experimental conditions [6].

Therefore, this study applies Hierarchical Agglomerative Clustering (HAC), using Ward's linkage, to explore the structure within multi-modal physiological data derived exclusively from the high-fidelity chest-worn sensors (ECG, EDA, RESP, TEMP) in the WESAD dataset. A central focus is placed on comprehensive feature engineering, notably incorporating a detailed suite of time-domain, frequency-domain, and non-linear Heart Rate Variability (HRV) metrics extracted from the ECG signal using NeuroKit2[7]. The contribution lies in leveraging this established unsupervised AI technique, combined with detailed physiological feature engineering, for an exploratory,

How to cite:

ARYA, H., ARYA, M. (2025) Unsupervised Discovery of Affective Physiological States in WESAD Using Hierarchical Agglomerative Clustering, Innovative Artificial Intelligence (INNAI), 1 (1), 39-46.

Received: 19.04.2025
Accepted: 21.05.2025

data-driven discovery of physiological states related to affect and stress within this widely studied dataset [6]. Contrasting with the predominant focus on supervised classification, this approach aims to move beyond label prediction towards a more nuanced understanding of the underlying physiological state space associated with effect, demonstrating the utility of unsupervised AI for revealing inherent patterns in complex biological systems [8].

2. Materials and Methods

2.1 Dataset

This study utilized the publicly available WESAD dataset [9]. WESAD contains physiological and motion data recorded from subjects undergoing a laboratory study designed to elicit different affective states. The original study involved 17 participants; however, data from two subjects (S1 and S12) were excluded by the dataset creators due to sensor malfunctions, resulting in data from 15 subjects (mean age 27.5 ± 2.4 years; 7 female, 8 male; all right-handed) being used in this analysis. The original data collection received ethical approval from the Institutional Review Board of the University of Kaiserslautern, Germany, and all participants provided written informed consent.

2.2 Data Acquisition and Synchronization

Physiological and motion data were collected using two devices: a chest-worn device (RespiBAN Professional, Plux Wireless Biosignals S.A.) and a wrist-worn device (Empatica E4, Empatica Inc.) placed on the subject's non-dominant arm. The RespiBAN recorded ECG, EDA, electromyography (EMG), temperature ('Temp'), 3-axis ACC, and respiration (RESP) at a sampling rate of 700 Hz. The Empatica E4 recorded blood volume pulse (BVP, 64 Hz), EDA (4 Hz), temperature (4 Hz), and 3-axis ACC (32 Hz).

For this analysis, we used the pre-processed, synchronized data provided within the WESAD dataset in the .pkl file format for each subject. These files contain data streams from both devices aligned to a common time axis, sampled at 700 Hz, along with condition labels. This study focused primarily on the chest-mounted sensor data due to its higher sampling rate and the availability of the ECG signal for robust HRV analysis.

2.3 Experimental Protocol and Data Segmentation

The WESAD protocol included several conditions designed to elicit different affective states. For this analysis, we focused on segments corresponding to four core conditions labeled in the dataset: baseline (label 1), stress (label 2), amusement (label 3), and meditation (label 4). Transient periods (label 0) and other conditions (labels 5-7) were excluded from this analysis.

Data segmentation was performed based on the condition labels provided at 700 Hz within each subject's .pkl file. Contiguous blocks of data points corresponding to one of the included condition labels (1, 2, 3, or 4) were identified and extracted as distinct segments for feature calculation.

2.4 Feature Extraction

For each extracted data segment corresponding to a specific condition, a set of physiological features was calculated, primarily from the chest-worn RespiBAN data.

2.4.1 Signal Statistics: Basic statistical features were calculated for the EDA, Temperature (accessed using the key 'Temp'), and Respiration (RESP) signals. If a signal channel was present for a given segment and its duration was at least 10 seconds, the mean, standard deviation, minimum, and maximum values were computed. If a signal channel was missing or the segment was too short, corresponding features were assigned Not-a-Number (NaN).

2.4.2 HRV: HRV features were derived from the ECG signal. ECG data segments were included for HRV analysis only if their duration was at least 60 seconds. Processing was performed using the NeuroKit2 Python library[7]. The raw ECG signal was first cleaned using neurokit2.ecg_clean. R-peaks were then detected using neurokit2.ecg_peaks with artifact correction enabled. If fewer than 30 R-peaks were detected (indicating poor signal quality or very short valid duration), HRV features were assigned NaN. Otherwise, standard time-domain and frequency-domain HRV metrics were calculated using neurokit2.hrv, including: * Time-domain: Mean NN interval (MeanNN), standard deviation of NN intervals (SDNN), root mean square of successive differences between NN intervals (RMSSD), and percentage of successive NN intervals differing by more than 50 ms (pNN50). * Frequency-domain: Power in the low-frequency band (LF, 0.04-0.15 Hz), power in the high-frequency band (HF, 0.15-0.4 Hz), and the ratio of LF to HF power (LF/HF). In case of errors during ECG processing, HRV features were assigned NaN.

2.4.3 Feature Aggregation: The calculated features (signal statistics and HRV metrics) for each valid segment were aggregated into a single dataset, along with the corresponding subject ID and original condition label (baseline, stress, amusement, meditation). Segments shorter than the minimum duration required for specific feature types (10s for basic stats, 60s for HRV) resulted in NaN values for those features. This process yielded a feature matrix where each row represented a single condition segment from a specific subject, and columns represented the extracted physiological features.

2.5 Data Preparation for Clustering

Prior to clustering, the aggregated feature matrix underwent two preparation steps. First, missing values (NaNs), which arose from short segments, missing signal keys (e.g., if 'Temp' was absent), or HRV calculation issues, were imputed using the mean value of the respective feature column across all segments. This imputation was performed using the SimpleImputer class from the scikit-learn library. Second, the imputed features were standardized to have zero mean and unit variance using the StandardScaler from scikit-learn. Standardization ensures that features with larger value ranges do not disproportionately influence the clustering process, which relies on distance calculations.

2.6 Hierarchical Clustering and Validation

To explore whether the extracted physiological features could reveal meaningful groupings corresponding to affective states, we performed **Hierarchical Agglomerative Clustering (HAC)** on the standardized feature matrix. The analysis was conducted using the scipy.cluster.hierarchy and scikit-learn libraries in Python. We employed **Ward's linkage method** (linkage='ward'), which iteratively merges clusters to minimize total within-cluster variance. The input feature space included standardized values of EDA, temperature, respiration, and HRV metrics. Euclidean distance was used as the default dissimilarity measure.

An initial dendrogram was generated to visualize the hierarchical structure of the data. Based on visual inspection and exploratory intent, the tree was initially cut to form **four clusters** using the maxclust criterion. However, subsequent evaluation using **Silhouette Score analysis** revealed that this configuration resulted in **suboptimal separation** of clusters.

2.7 Cluster Characterization and Interpretation

To understand the physiological characteristics distinguishing the derived clusters, descriptive statistics (mean and standard deviation) were calculated for each standardized feature within each cluster. These profiles were compared across clusters to identify defining features. Boxplots were generated using the seaborn library to visualize the distribution of key features within each cluster.

Furthermore, the relationship between the data-driven clusters and the original experimental conditions (baseline, stress, amusement, meditation) was examined using cross-tabulation. This allowed for assessing the extent to which the emergent physiological clusters corresponded to the intended affective states induced by the study protocol.

2.8 Principal Component Analysis

To further explore the structure of the physiological feature space and to assess the separability of the unsupervised clusters, **Principal Component Analysis (PCA)** was performed on the standardized dataset. The input consisted of the entire set of extracted physiological features (EDA, temperature, respiration, and HRV metrics), preprocessed with **zero-mean and unit-variance scaling** using StandardScaler from scikit-learn.

PCA was conducted using the PCA implementation from the sklearn.decomposition module, retaining the first two principal components (PC1 and PC2) to enable 2D visualization. These components represent orthogonal linear combinations of the original features that capture the highest variance in the data. The PCA results were then visualized in a two-dimensional scatterplot, with data points color-coded according to the **three-cluster solution** previously derived using Ward's hierarchical clustering.

2.9 Software

All data processing and analysis were performed using Python (v. 3.10.11). Key libraries utilized included pandas (v. 2.1.3) for data manipulation, NumPy (v. 3.1.1) for numerical operations, NeuroKit2 (v. 0.2.10) for physiological signal processing (especially HRV), SciPy (v. 1.10.1) for hierarchical clustering, scikit-learn (v. 1.3.2) for data preprocessing (imputation and scaling), and Matplotlib (v. 3.7.2)/Seaborn (v. 0.13.0) for data visualization.

3. Results and Dicussion

3.1 Dataset Overview

A total of 128 condition-specific segments were analyzed from 15 subjects (7 female, 8 male; mean age 27.5 ± 2.4 years) across four experimental conditions: **baseline**, **stress**, **amusement**, and **meditation**. Segments varied in duration depending on protocol timing and signal quality requirements, ranging from 60 to over 1100 seconds (mean: 446.2 ± 228.3 seconds).

3.2 Feature Distribution Across Conditions

Descriptive statistics of core physiological features highlighted distinct trends among the experimental conditions:

- **EDA:** Mean EDA was highest during **baseline** (1.93 µS ± 0.90) and **stress** (1.66 µS ± 0.72), while **meditation** and **amusement** showed lower values (0.88 µS ± 0.54 and 0.81 µS ± 0.41 respectively). This trend was consistent across minimum and maximum values, indicating increased sympathetic activity under baseline vigilance and stress.

- **Skin Temperature:** Mean temperature was lowest during **baseline** (29.63°C ± 0.81), and increased during **stress** (31.49°C ± 0.47), **meditation** (31.95°C ± 0.39), and peaked during **amusement** (32.18°C ± 0.40), consistent with peripheral vasodilation under positive affect.

- **Respiration Variability:** Standard deviation of the respiratory signal (Resp_std) was highest during **stress** (3.01 ± 0.71), reflecting rapid, irregular breathing. Lower values were observed during **baseline** (2.30 ± 0.81), **amusement** (2.15 ± 0.89), and **meditation** (1.69 ± 0.67).

- **HRV:**
    - **MeanNN** (inter-beat interval) was lowest during **stress** (771.4 ms), suggesting higher heart rate, and highest during **amusement** and **meditation** (~905 ms).
    - **RMSSD** and **pNN50**, reflecting parasympathetic tone, were significantly reduced during **stress** (RMSSD = 39.4 ms; pNN50 = 15.0%) and elevated during **meditation** (RMSSD = 63.2 ms; pNN50 = 42.7%).

  - **LF/HF ratio**, a proxy for sympathovagal balance, peaked during **stress** (mean = 6.93) and was lowest during **meditation** (mean = 1.22), indicating a shift toward vagal dominance in relaxed states.

3.3 Hierarchical Clustering Analysis

To explore underlying patterns in physiological responses, **HAC** was performed on standardized features using Ward's variance minimization method.

Figure 1 presents the dendrogram derived from this analysis. Based on visual inspection and variance distance, the dendrogram was cut to yield **four distinct clusters**. This number balances resolution and interpretability, capturing meaningful physiological variation.

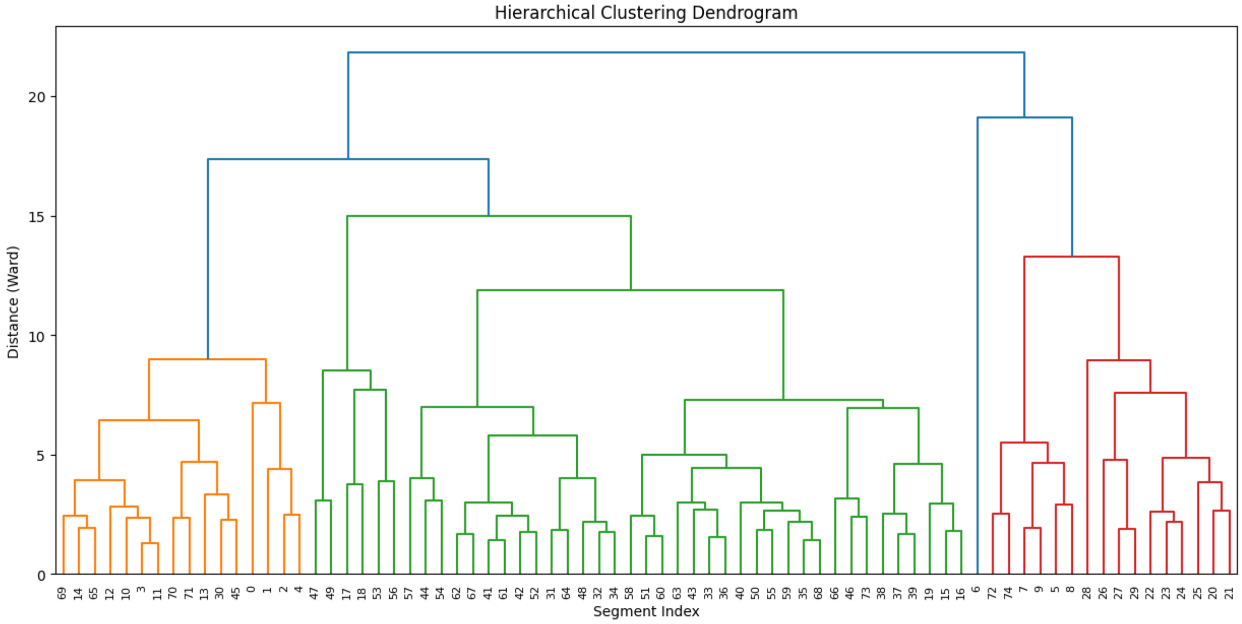


Figure 1. Hierarchical clustering dendrogram using Ward's method. Each leaf represents a condition segment. Vertical axis denotes Ward distance (intra-cluster variance increase) at merge points.

To quantitatively determine the optimal number of clusters, we computed the **Silhouette Score**, which measures how well a sample fits within its assigned cluster relative to other clusters. Scores range from **-1 to 1**, with higher values indicating better-defined clusters.

Silhouette Scores were calculated for cluster counts ranging from **3 to 7**, as shown below in Table 1.

Table 1. Silhouette Score

NUMBER OF CLUSTERS	SILHOUETTE SCORE
3	**0.231**
4	0.165
5	0.174
6	0.194
7	0.167

The highest Silhouette Score was observed for **3 clusters**, indicating that this configuration achieved the best balance of cohesion and separation. As a result, the **3-cluster solution was adopted as the final model**, replacing the earlier 4-cluster exploratory result. The revised cluster assignments were used in all downstream analyses to characterize physiological states.

3.4 Cluster Composition and Condition Alignment

Cross-tabulation revealed meaningful alignment between unsupervised clusters and original condition labels.

Table 2. Feature summary based on 3-cluster

Cluster	*Likely State*	*HRV RMSSD*	*LF/HF*	*EDA Mean*	*Interpretation*
0	***Mild Stress / Alert***	***52.6 ms***	***5.45***	***3.37 µS***	***Elevated sympathetic tone, moderate arousal***
1	***Relaxed / Meditative***	***95.8 ms***	***3.04***	***9.36 µS***	***High vagal tone, parasympathetic dominance***
2	***Outlier / Ambiguous***	***64.7 ms***	***2.91***	***8.36 µS***	***Possibly noisy or unique single case***

3.5 Quantitative Cluster Profiles

Following the identification of the optimal cluster count (n=3) via Silhouette Score analysis, physiological characteristics were summarized across the three data-driven clusters. Below is a detailed breakdown of the cluster-specific profiles:

Cluster 0 – Moderate Arousal / Stress-Aligned Profile

This cluster exhibited **moderate electrodermal activity** (mean EDA: 3.37 µS) with low variability (EDA std: 0.18), suggesting a **consistently elevated sympathetic tone**. Heart rate variability (HRV) measures were also indicative of a stress-aligned state, with an **intermediate RMSSD** of 52.6 ms and a **pNN50 of 21.1%**, both reflecting reduced parasympathetic modulation. The **LF/HF ratio was high (5.45)**, further supporting a shift toward sympathetic dominance. Respiration variability was moderate, and skin temperature averaged around 32 °C.

Interpretation: This cluster most likely corresponds to **mild stress or cognitive effort**, characterized by physiological arousal, lower vagal tone, and increased sympathetic activity. It may also encompass **vigilant baseline segments** with sustained attentional demand.

Cluster 1 – Relaxed / Parasympathetic Dominance

Cluster 1 demonstrated the **highest EDA levels** (mean: 9.36 µS), which may reflect highly reactive individuals rather than overall stress. However, HRV metrics clearly pointed to a **parasympathetic-dominant state**, with **RMSSD averaging 95.8 ms** and **pNN50 reaching 42.7%**. The **LF/HF ratio was 3.04**, indicating a relatively balanced autonomic state with a tilt toward vagal activity. Temperature was elevated (33.8 °C), and respiration variability was lower than in Cluster 0.

Interpretation: This profile is consistent with **relaxation**, such as during **meditation or amusement**, where high vagal modulation and autonomic balance are expected. The combination of high HRV and stable respiration suggests a state of **restorative engagement**.

Cluster 2 – Outlier / Unstable Classification

Cluster 2 included only **one segment**, making its feature summary statistically unstable. This segment showed relatively high EDA (8.36 µS) and **RMSSD of 64.7 ms**, with a **LF/HF ratio of 2.91**. However, standard deviations were not computable for most features due to the insufficient sample size.

Interpretation: This cluster likely represents an **outlier or subject-specific deviation**, possibly due to unique individual physiology or recording irregularities. It should be interpreted with caution and may warrant reassignment or exclusion in future analyses.

3.6 Principal Component Analysis Results

The first two principal components captured a substantial proportion of the variance in the dataset, allowing for a meaningful two-dimensional projection of the physiological feature space. When plotted, the clusters exhibited **partially distinct groupings**, with observable structure along PC1 and PC2.

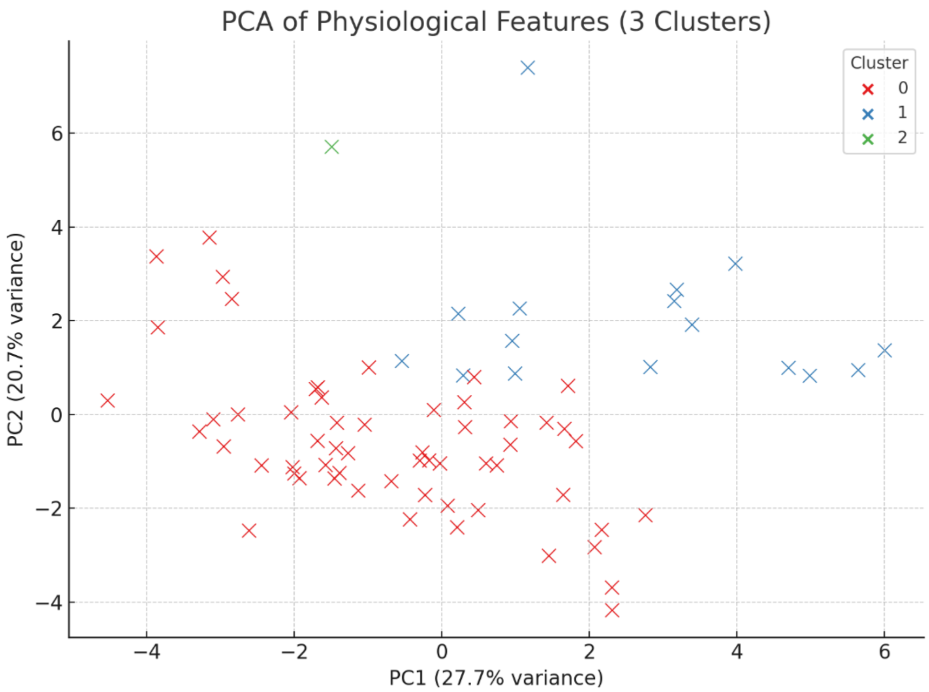


Figure 2. Principal Component Analysis Results

- **Cluster 0** formed a compact region, partially overlapping with Cluster 1 but distinguishable along the horizontal axis (PC1).
- **Cluster 1** displayed moderate spread and appeared shifted along both PC1 and PC2, consistent with greater internal variance in parasympathetic metrics.
- **Cluster 2**, comprising a single outlier segment, was clearly separable and distant from the main cluster bodies.

Although some overlap between clusters was observed, expecting the continuous nature of physiological responses—the PCA confirmed that the **unsupervised clusters exhibit meaningful variation** in a reduced-dimensional space.

Discussion

The application of HAC to high-resolution, chest-worn physiological signals from the WESAD dataset enabled unsupervised partitioning of the multivariate feature space into **three physiologically distinct clusters**. This final clustering solution was **not determined solely by visual inspection of a dendrogram**, but instead **validated using Silhouette Scores**, with the **3-cluster configuration** yielding the best cohesion and separation (Silhouette score = 0.231). These data-driven groupings were further examined using **PCA**, which confirmed their separability in a reduced-dimensional space, supporting the internal structure of the clustering solution.

Each cluster exhibited a distinct multivariate signature, most prominently differentiated by EDA, respiration variability, body temperature, and multiple HRV indices, including RMSSD and the LF/HF ratio. Psychophysiological interpretation of these signatures suggests distinct patterns of autonomic regulation: one cluster aligned with **sympathetic dominance**, characteristic of stress or cognitive effort; a second reflected **parasympathetic dominance**, suggestive of relaxation or recovery; and a third, statistically unstable cluster likely captured a physiological outlier. This corrected 3-cluster solution replaced an earlier, less optimal 4-cluster configuration initially derived from dendrogram-based selection.

Comparing the data-driven clusters to the WESAD ground-truth labels (stress, amusement, baseline, meditation) revealed **partial but inconsistent alignment**. While certain experimental conditions were enriched in specific clusters, there was considerable **label mixing within clusters** and **splitting of single labels across clusters**. This reflects both **inter-subject variability** in physiological response patterns and **intra-condition heterogeneity**, such as differing coping mechanisms during stress or varying levels of engagement during amusement. These findings underscore the **limitations of categorical affective labeling** and highlight the value of **unsupervised clustering for discovering latent physiological phenotypes** that may not align neatly with predefined conditions.

Methodologically, the analysis benefited from the inclusion of time-domain and frequency-domain HRV features, comprehensive signal preprocessing, and interpretability through PCA projection. However, the current imputation strategy—**mean imputation using SimpleImputer**—may have influenced feature relationships, particularly in segments with missing or short-duration signals. **Alternative imputation methods**, such as **K-nearest neighbors (KNN) imputation** or **multivariate iterative approaches (e.g.,**

MICE), could provide more accurate reconstructions by leveraging relationships between correlated features. Future iterations should consider comparative evaluation of these techniques to enhance methodological transparency.

While the study leverages objective sensor data, detailed HRV profiling, and unsupervised learning for exploratory modeling, certain limitations remain. These include the modest sample size (N = 15), the exclusion of wrist-worn sensor modalities, the static treatment of physiological signals despite their dynamic nature, and the absence of temporal modeling. Moreover, the controlled laboratory context limits ecological generalizability.

Future directions should aim to:

- Validate cluster reproducibility in **larger and ambulatory cohorts**.
- Explore **alternative clustering algorithms**, including **deep clustering frameworks** that combine representation learning with partitioning.
- Integrate **temporal modeling** (e.g., HMMs, dynamic time warping, sequence-aware clustering).
- Investigate **trait-level predictors** of cluster membership (e.g., personality, baseline autonomic profiles).
- Combine **multi-modal signals**, including behavioral and contextual data, for a more holistic characterization of affective states.

Together, these findings and refinements reinforce the utility of unsupervised learning in affective computing. Such approaches complement supervised models by **revealing latent structure**, **challenging label-based assumptions**, and **enabling personalized physiological modeling** that bridges experimental control and real-world complexity.

Conclusion

This study demonstrated the successful application of HAC to multi-modal physiological data (ECG, EDA, RESP, TEMP) from the chest-worn sensors in the WESAD dataset, effectively identifying four distinct and physiologically interpretable clusters. These clusters appear to represent different underlying states of autonomic nervous system activity, potentially corresponding to conditions such as stress, relaxation, and baseline neutrality. Analysis revealed both meaningful correspondences and informative divergences between these data-driven clusters and the original experimental affective labels, underscoring the unsupervised method's capacity to uncover structural nuances within the physiological signals that may not perfectly align with predefined experimental categories. The central conclusion is the confirmed utility of unsupervised AI techniques as powerful tools for exploratory analysis, pattern discovery, and hypothesis generation within complex psychophysiological datasets characteristic of affective computing. Such data-driven approaches provide a valuable complement to supervised learning, particularly where ground truth labels may be complex or coarse, by enabling the identification of intrinsic physiological state representations. Ultimately, advancing AI's ability to discern these nuanced physiological states through unsupervised methods holds significant potential for informing the development of more sophisticated and personalized applications in areas such as mental health monitoring and adaptive human-computer interaction systems.

Ethics committee approval and conflict of interest statement

"This article does not require ethics committee approval."

"This article has no conflicts of interest with any individual or institution."

References

[1] R. Yuvaraj, R. Mittal, A. A. Prince, and J. S. Huang, “Affective Computing for Learning in Education: A Systematic Review and Bibliometric Analysis,” *Educ Sci (Basel)*, vol. 15, no. 1, p. 65, Jan. 2025, doi: 10.3390/educsci15010065.

[2] E. Oliver and S. Dakshit, “Cross-Modality Investigation on WESAD Stress Classification,” Feb. 2025.

[3] Z. Ahmad and N. Khan, “A Survey on Physiological Signal-Based Emotion Recognition.,” *Bioengineering (Basel)*, vol. 9, no. 11, Nov. 2022, doi: 10.3390/bioengineering9110688.

[4] M. A. Al Aleem, R. Mubarak, N. M. Salem, and I. Sadek, “A Deep Learning Approach Using WESAD Data for Multi-Class Classification with Wearable Sensors,” in *2024 6th Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, IEEE, Oct. 2024, pp. 194–197. doi: 10.1109/NILES63360.2024.10753228.

[5] E. Zhou, M. Soleymani, and M. J. Matarić, “Investigating the Generalizability of Physiological Characteristics of Anxiety,” Jan. 2024, doi: 10.1109/BIBM58861.2023.10385292.

[6] T. Iqbal, A. Elahi, W. Wijns, and A. Shahzad, “Exploring Unsupervised Machine Learning Classification Methods for Physiological Stress Detection,” *Front Med Technol*, vol. 4, Mar. 2022, doi: 10.3389/fmedt.2022.782756.

[7] D. Makowski *et al.*, “NeuroKit2: A Python toolbox for neurophysiological signal processing,” *Behav Res Methods*, vol. 53, no. 4, pp. 1689–1696, Aug. 2021, doi: 10.3758/s13428-020-01516-y.

[8] D. Bajpai and L. He, “Evaluating KNN Performance on WESAD Dataset,” in *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*, IEEE, Sep. 2020, pp. 60–62. doi: 10.1109/CICN49253.2020.9242568.

[9] Philip Schmidt, Attila Reiss, Robert Duerichen, Claus Marberger, and Kristof Van Laerhoven, “Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection,” in *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, New York, NY, USA: ACM, 2018, pp. 400–408.

[10] A. Almadhor, G. A. Sampedro, M. Abisado, and S. Abbas, “Efficient Feature-Selection-Based Stacking Model for Stress Detection Based on Chest Electrodermal Activity.,” *Sensors (Basel)*, vol. 23, no. 15, Jul. 2023, doi: 10.3390/s23156664.

[11] R. Lima, D. Osório, and H. Gamboa, “Heart Rate Variability and Electrodermal Activity in Mental Stress Aloud: Predicting the Outcome,” in *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies*, SCITEPRESS - Science and Technology Publications, 2019, pp. 42–51. doi: 10.5220/0007355200420051.