



# Electroencephalogram-based emotion recognition using involvement indexes: A machine-learning approach

Alessandro Dell'Orletta<sup>1</sup> , Erica Iammarino<sup>1</sup> , Agnese Sbrollini , Laura Burattini ,  
Ilaria Marcantoni<sup>\*</sup>

Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy

## ARTICLE INFO

### Keywords:

Affective computing  
Machine learning  
Engagement index  
Spectral features  
EEG

## ABSTRACT

Emotional state is often assessed using electroencephalography (EEG), a non-invasive neuroimaging technique that measures brain electrical activity. Machine-learning approaches to EEG-based emotion recognition have been explored, but regression models for predicting emotion intensity remain unexplored. Moreover, few EEG-derived involvement indexes, defined as ratios of the spectral powers of two or more EEG rhythms, have been used for emotion recognition. This study aims to use EEG-derived involvement indexes to develop predictive models capable of identifying and rating specific emotional states, introducing a novel methodological framework that integrates 37 involvement indexes together with the powers of standard EEG rhythms as interpretable input features for predicting emotion intensity. To do so, 64-channel EEG signals recorded from healthy subjects while watching emotionally evocative videos were preprocessed using EEGLAB. Then, 6 EEG rhythms and 37 involvement indexes were extracted and given as inputs to five separate regression models, each corresponding to one target emotion (i.e., Anger, Sadness, Happiness, Disgust, and Fear), implemented using LASSO regression. Model performance was evaluated via mean absolute error (MAE), mean squared error (MSE), root mean square error (RMSE), and coefficient of determination ( $R^2$ ). Results showed MAE values between 0.90 and 1.19, RMSE between 1.23 and 1.56, and  $R^2$  between 0.79 and 0.81, with Anger predicted most accurately and Happiness showing the highest variability. These findings demonstrate that combining the power of EEG rhythms with involvement indexes enables reliable and interpretable prediction of emotion intensity, highlighting the value of LASSO regression in handling high-dimensional EEG data.

## 1. Introduction

Emotions are complex psychological and physiological responses to external stimuli that play a central role in shaping human experience, affecting perception, decision-making, and behavior. These responses are inherently dynamic and rarely isolated; indeed, psychologists have long recognized that individuals often experience multiple emotions simultaneously, reflecting the intricate nature of human emotional processing [1].

Affective computing is an interdisciplinary field that aims to enable machines to recognize, interpret, and simulate human emotions. Positioned at the intersection of computer science, psychology, and cognitive neuroscience, it plays a pivotal role in domains such as human-computer interaction, mental health monitoring, and

development of adaptive systems. By embedding emotional intelligence into technology, affective computing enhances the ability of machines to engage with users in more empathetic and meaningful ways [2]. A core component of this field is emotion recognition, which involves identifying human emotional states based on observable or measurable signals. These signals can stem from facial expressions, speech intonation, body movements, or physiological responses. Among physiological signals, the electroencephalogram (EEG) stands out for its ability to directly capture brain electrical activity, non-invasiveness, and high temporal resolution, making it a valuable tool for studying real-time brain dynamics during emotional processing [2,3].

The EEG measures the electrical activity of the brain via electrodes placed on different regions of the scalp and can provide insights into an individual's psycho-emotional state by analyzing patterns across various

<sup>\*</sup> Corresponding author at: Department of Information Engineering, Università Politecnica delle Marche, Via Brecce Bianche 12, 60131 Ancona, Italy.

E-mail addresses: [S1112278@studenti.univpm.it](mailto:S1112278@studenti.univpm.it) (A. Dell'Orletta), [e.iammarino@pm.univpm.it](mailto:e.iammarino@pm.univpm.it) (E. Iammarino), [a.sbrollini@staff.univpm.it](mailto:a.sbrollini@staff.univpm.it) (A. Sbrollini), [l.burattini@univpm.it](mailto:l.burattini@univpm.it) (L. Burattini), [i.marcantoni@staff.univpm.it](mailto:i.marcantoni@staff.univpm.it) (I. Marcantoni).

<sup>1</sup> Equal contribution.

<https://doi.org/10.1016/j.bspc.2026.110341>

Received 31 July 2025; Received in revised form 17 March 2026; Accepted 15 April 2026

Available online 21 April 2026

1746-8094/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

frequency bands, known as EEG rhythms. These include delta rhythm, associated with deep sleep; theta rhythm, linked to emotional introspection and memory processes; alpha rhythm, typically observed during relaxed wakefulness; sensorimotor rhythm (SMR), associated with motor readiness and attention; beta rhythm, linked to active thinking and problem-solving; and gamma rhythm, associated with sensory integration and higher-order cognition [3–5]. To interpret these EEG patterns in the context of emotions, it is essential to adopt emotion-cognition models, which provide a multidisciplinary framework rooted in psychology and cognitive science for interpreting how emotions are represented, modeled, and classified. To date, two main approaches have emerged: the discrete emotion model and the dimensional space model. The discrete models, exemplified by Ekman's model, conceptualize emotions as distinct categories (e.g., happiness, fear, anger); the dimensional models, instead, describe emotions as continuous coordinates along affective dimensions such as valence, which reflects the psychological polarity of the emotion, ranging from low (negative) to high (positive), and arousal, which measures the intensity of the emotional response, similarly ranging from low (passive) to high (active) [6–8]. Research has shown that changes in the power of EEG rhythms can reflect changes in emotional and cognitive states. For instance, high arousal levels are typically associated with reduced alpha power and increased delta, theta, and gamma power, especially in posterior brain regions [9]. In particular, as emotions shift from states of excitement to calmness, delta power generally increases, whereas alpha power decreases. Increased gamma activity has also been linked to both happiness and sadness, while theta power has shown a negative correlation with emotional arousal levels [10,11]. There is also evidence that these spectral changes correspond to activity in specific cortical regions, which have been linked to specific discrete emotions. The left frontal cortex is often associated with positive emotions like happiness and approach motivation, while the right frontal cortex is linked to negative emotions such as sadness, fear, and withdrawal [4,12,13]. Anger has been associated with activity in both frontal and central regions [4]. The right temporal region is involved in the perception of fear, particularly when processing auditory or social cues, while disgust has been correlated with left frontotemporal regions, potentially reflecting insular activity. The parietal cortex is thought to modulate emotional arousal, occipital sites are involved in processing surprise and visual stimuli, and central regions are frequently linked to general arousal and bodily feedback. Thus, EEG rhythms, with distinct patterns observed across different cortical regions, have been recognized as neural biomarkers for affective states and are increasingly used in EEG-based emotion recognition systems across research, clinical, and real-world applications [12,14]. In addition to analyzing single EEG frequency bands, a growing body of research has investigated the use of more comprehensive biomarkers of mental states, known as involvement indexes. In 2023, Marcantoni et al. [15] systematically identified and reported 37 EEG-derived ratio indexes used in the literature to assess human mental involvement. These indexes are defined as ratios of the spectral power of two or more EEG rhythms. The literature does not provide any definitive physiological interpretation of them that would enable to infer a specific emotion from a particular index. Indeed, so far, a few of these ratio indexes have been used to try to characterize and identify human emotions. For instance, the beta-to-alpha ratio, recognized as index of arousal, has been used to feed emotion recognition models [16,17], or to assess emotional states [18]; while Xu et al. recognized the beta-to-theta ratio as a relevant feature to recognize emotional states [19]. Although these indexes reflect emotional arousal and involvement, to date no published study has directly used them as features for EEG-based emotion recognition models.

In this context, the present study aims to investigate the role of EEG-derived involvement indexes in emotion recognition by developing predictive models capable of identifying and rating specific emotional states. The main contributions of this work are as follows:

- the proposal of a novel methodological framework that integrates EEG-derived involvement indexes as input features for emotion recognition models, possibly extending their application to affective modeling;
- the development of regression models capable of inferring emotions by using interpretable EEG-derived features, using a multi-class (rather than binary) prediction approach;
- the evaluation of five distinct but concurrent emotional states, each assessed according to varying intensity levels, relying on a physiologically realistic coexistence of emotions.

Using an advanced statistical feature-based approach, this study seeks to uncover meaningful patterns in EEG data that are indicative of emotional responses, with the ultimate goal of contributing to the development of emotion recognition systems.

The remainder of this article is organized as follows. Section 2 presents a critical review of the state of the art. Section 3 describes the dataset used to implement and validate the models and provides a detailed description of the models themselves. Section 4 reports the results obtained, while Section 5 comments on these results, compares them with existing literature, and outlines the limitations of the implemented algorithm. Finally, Section 6 concludes with the study take-home messages and the possible future work.

## 2. State of the art

Several studies have explored EEG-based emotion recognition using machine learning techniques. In order to gain a comprehensive summary of the current state of the art, we conducted a literature review using the Scopus database, focusing on publications from the last three years. The search was conducted within the title, abstract, and keyword fields, ensuring that all of the following terms were included: “EEG” or “electroencephalogram”, “emotion recognition”, and “machine learning”.

Only review articles written in English were considered, with exclusion of those focusing on specific applications or pathological conditions. This search strategy yielded 13 studies [6,20–31].

The reviewed literature indicates that EEG-based emotion recognition typically relies on four major categories of features: time-domain, frequency-domain, time-frequency-domain, and nonlinear features. Time-domain features include statistical descriptors such as mean, variance, skewness, kurtosis, peak-to-peak interval, mean square value, and Hjorth parameters [6,20,22,25,29,31]. Frequency-domain features include power and energy metrics quantified across different EEG rhythms, mainly computing the absolute and relative power spectral density [6,20–25,27–29]. Notably, in the review conducted by Huy-Tung et al., also the index beta/(alpha + theta), the second index listed by Marcantoni and co-workers in their systematic review, has been reported as index of arousal [6]. Time-frequency analysis, like wavelet transform, short-time Fourier transform, or wavelet packet decomposition, has also been applied to capture both temporal and spectral information [20–22]. For what concerns nonlinear features, the reviewed studies pointed out entropy-based metrics, such as Shannon entropy [25] and differential entropy [20–24,31], as well as fractal dimension [6,20–22], which describe the complexity and irregularity of the signal. Some studies also examined connectivity and asymmetry features [6,24,25,28,29,31]. Connectivity-related features include correlation, coherence, phase synchronization, phase-locking value, and graph-based measures such as clustering coefficient, degree, efficiency, modularity, and closeness centrality [6,24,31]. Brain asymmetry features were commonly derived from the difference or ratio between channels placed in homotopic regions or through indexes such as the frontal alpha asymmetry index [6,25,29]. Regarding computational approaches, the most popular machine learning techniques resulted to be support vector machines, k-nearest neighbors, and random forests [20–24]. More recently, deep learning models have gained prominence,

including convolutional neural networks, long short-term memory networks, graph neural networks, and hybrid architectures such as convolutional recurrent neural networks [23]. To effectively train machine learning or deep learning systems, a critical step involves defining how emotions are represented, whether as continuous dimensions, binary categories, or discrete emotional states. The majority of studies adopt a dimensional model, mapping EEG signals to continuous values of valence and arousal, as this framework aligns well with the continuous nature of brain activity. Other studies use binary classification, typically distinguishing between high and low valence or arousal, simplifying emotional states into dichotomous categories. Only a small portion of the literature employs discrete emotion models, and these often assume mutually exclusive emotional states. Explicit modeling of coexisting emotions with varying intensities remains rare. Indeed, most systems do not assign multiple simultaneous emotional labels; rather, they infer emotional richness through continuous dimensional outputs. Despite these advances, no prior studies have applied regression models for emotion intensity prediction, and the potential role of EEG-derived ratio indexes as predictive features remains underexplored. Indeed, the reviewed literature points out that these ratios, which express relationships between the spectral powers of different EEG rhythms, have been associated with mental effort and arousal, but have not been systematically employed in emotion recognition models yet.

### 3. Materials and methods

#### 3.1. Description of the Emotional Arousal Pattern dataset

The population analyzed in the present study comes from the Emotional Arousal Pattern (EMAP) database published by Eisenbarth, Oxner, and Shehu on September 7th, 2023. It is an open-access database aimed at studying human emotional responses through behavioral and physiological measures. The database includes EEG and peripheral physiological data acquired from 145 subjects (mean age of  $22.7 \pm 6.7$  years; 93 female, 48 male, 4 other; 8 left-handed, 137 right-handed) recruited from the student body of the Victoria University of Wellington and the surrounding community. Data were recorded while participants watched 24 emotionally evocative videos, each lasting between 13 and 19 s (mean duration of  $15 \pm 1$  s) and presented with no audio [32]. The videos were selected to span a broad spectrum of emotional valence and arousal and were presented to each participant in a unique, random order [32]. Each video was repeated four times in a continuous loop that constitutes a single trial, with no interval between repetitions [32]. During the video presentation, participants used a computer mouse to continuously rate the intensity of emotion they were feeling on a slider scale from 0 (no emotion at all) to 10 (very intense emotion). After each trial, participants provided a single overall rating for the video on nine different scales. Specifically, they rated how positive or negative the video made them feel, with -5 being extremely negative, 5 being extremely positive and 0 being neutral, i.e., neither positive nor negative; the intensity of the emotion they felt while watching the video, with 0 being not intense at all and 10 being extremely intense; how much they liked the video, with 0 being not at all and 10 being very much; how much they wanted to engage with the video, with 0 being they strongly wanted to avoid it and 10 being they strongly wanted to engage with it; how much they felt anger, sadness, happiness, disgust, and fear while watching the video, with 0 being not at all and 10 being very strongly [32].

EEG data were recorded using a BrainVision actiCHamp amplifier and BrainVision Recorder (Brain Products GmbH, <https://www.brainproducts.com>) from 64 active Ag/AgCl electrodes placed according to the international extended 10–20 system. The sampling frequency was 500 Hz, the electrode impedance was kept below 25 k $\Omega$ , and the signals were referenced to electrode FCz. To monitor eye movements

and blinks, an electrooculogram (EOG) was recorded. In addition, a PowerLab amplifier was used to record other peripheral physiological measures (electrocardiogram, respiration, electrodermal activity, and blood volume). In this study, only EEG data were considered and analyzed.

The database was provided in two versions: the Raw Data Set, in which peripheral data were down-sampled to 250 Hz and event/time markers (indicating video onsets and offsets) were co-registered across peripheral, EEG, and behavioral data; and the Clean Data Set, in which EEG data underwent additional basic preprocessing. For the present study, only the Raw Data Set was used.

#### 3.2. Processing and feature extraction

EEG data were analyzed in MATLAB R2021a using EEGLAB [33]. Preprocessing included resampling to 500 Hz, average referencing, band-pass filtering, and noise removal via independent component analysis (ICA). Specifically, EEG data were re-referenced to a robust average reference computed using the PREP Pipeline toolbox with default parameters, except that powerline noise frequencies were adjusted from 60/120 Hz to 50/100 Hz to match the actual noise present in the recordings [34]. Then, filtering was performed using finite impulse response (FIR) filters designed with a Kaiser window, as done by the database developers to obtain the Clean Data Set. EEG data were first high-pass filtered at 0.5 Hz, and then low-pass filtered at 100 Hz [35]. Eventually, noise removal was achieved by performing ICA using the Infomax algorithm and by exploiting the ICLabel plug-in [36], a tool that automatically assigns each independent component (IC) the probability of belonging to seven possible classes, each representing a different signal source. These classes are “Brain”, “Muscle”, “Eye”, “Heart” for biological sources, “Line Noise”, and “Channel Noise” for non-biological sources, and “Other” for unrecognized sources. Based on these probabilities, if a component had the probability of belonging to the “Brain” class lower than 10%, or the probability of belonging to any other class (excluding “Other”) greater than 80%, it was labelled as an artifact and was removed from the EEG data [24]. To further remove powerline noise at 50 Hz, the CleanLine plug-in was applied by setting the bandwidth and taper bandwidth at 1 Hz.

Preprocessed EEG data were then epoched for each subject. Using the information about video onsets and offsets provided by the database developers [32], EEG segments corresponding to single trials were extracted. For each EEG segment, a 30-s sliding window was recursively extracted every second until the end of the signal was reached. To detect any bad channels, each 30-s window was further divided into 2-s non-overlapping consecutive sub-windows, where the standard deviation (SD) of the signal was computed for each channel. If 4 times the SD was greater than 100  $\mu$ V, the 2-s sub-window was considered noisy. In turn, if at least one 2-s sub-window was identified as noisy within the 30-s window, the corresponding channel was rejected from the 30-s window. As a consequence, if all 30-s windows for a given channel were rejected, no data was available for that channel, and analysis could not be performed on it.

For each non-noisy 30-s window, a 6th-order Butterworth filter was implemented to extract six EEG rhythms: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), SMR (12–15 Hz), beta (15–35 Hz), and gamma (35–70 Hz). The power spectral density (PSD) of each rhythm was then estimated using Welch's overlapped segment averaging method and characterized in terms of area under curve [37]. For each accepted channel, the median power of each EEG rhythm was computed across the 30-s windows. Then, the 37 involvement indexes systematically reported by Marcantoni et al. [15] were computed. The specific mathematical formulation of the indexes is reported in Table 1. Thus, a total of 43 features (i.e., 6 rhythms and 37 involvement indexes) were computed for each EEG channel.

**Table 1**

Mathematical formulation of the 37 involvement indexes [12].

|                                       |                                                   |                                                   |                                                            |                                                            |
|---------------------------------------|---------------------------------------------------|---------------------------------------------------|------------------------------------------------------------|------------------------------------------------------------|
| $I_1 = \frac{\beta}{\alpha}$          | $I_9 = \frac{\theta + \alpha}{\alpha + \beta}$    | $I_{16} = \frac{\delta + \theta}{\alpha + \beta}$ | $I_{23} = \frac{\alpha + \beta}{\alpha + \theta}$          | $I_{30} = \frac{\alpha + \beta}{\gamma}$                   |
| $I_2 = \frac{\beta}{\alpha + \theta}$ | $I_{10} = \frac{\theta}{\alpha + \beta}$          | $I_{17} = \frac{\delta}{\alpha}$                  | $I_{24} = \frac{\alpha}{\beta + \gamma}$                   | $I_{31} = \frac{\alpha + \gamma}{\theta + \delta}$         |
| $I_3 = \frac{\beta}{\theta}$          | $I_{11} = \frac{\theta + \alpha}{\gamma}$         | $I_{18} = \frac{\delta}{\beta}$                   | $I_{25} = \frac{\delta + \theta + \alpha}{\beta + \gamma}$ | $I_{32} = \frac{\theta + \alpha}{\delta}$                  |
| $I_4 = \frac{\theta}{\alpha}$         | $I_{12} = \frac{\beta + \theta}{\alpha}$          | $I_{19} = \frac{\theta}{\gamma}$                  | $I_{26} = \frac{\alpha}{\delta + \theta + \alpha}$         | $I_{33} = \frac{\theta + \beta}{\alpha + \gamma}$          |
| $I_5 = \frac{\theta}{\delta}$         | $I_{13} = \frac{\delta + \theta}{\beta}$          | $I_{20} = \frac{\alpha}{\gamma}$                  | $I_{27} = \frac{\alpha}{\theta + \alpha + \beta}$          | $I_{34} = \frac{\beta + \gamma}{\delta + \theta}$          |
| $I_6 = \frac{SMR}{\theta}$            | $I_{14} = \frac{\delta + \theta + \alpha}{\beta}$ | $I_{21} = \frac{SMR + \beta}{\theta}$             | $I_{28} = \frac{\beta}{\theta + \gamma}$                   | $I_{35} = \frac{\delta + \alpha}{\theta + \gamma}$         |
| $I_7 = \frac{SMR}{\beta}$             | $I_{15} = \frac{\delta + \theta}{\alpha}$         | $I_{22} = \frac{\theta + \alpha}{\beta + \gamma}$ | $I_{29} = \frac{\beta + \gamma}{\delta}$                   | $I_{36} = \frac{\theta + \alpha}{\delta + \beta + \gamma}$ |
| $I_8 = \frac{\alpha + \beta}{\delta}$ |                                                   |                                                   |                                                            | $I_{37} = \frac{\alpha + \beta}{\delta + \theta + \gamma}$ |

### 3.3. Statistics

#### 3.3.1. Model implementation

Emotion and feature multicollinearity was assessed by Pearson's correlation analysis: features presenting a correlation coefficient higher than 0.75 were rejected.

Emotion recognition was performed by five regression models, one for each emotion (i.e., anger, sadness, happiness, disgust, and fear). Z-score standardized features were used as predictors, while participants' intensity of emotions was used as response. The models were implemented by using LASSO (Least Absolute Shrinkage and Selection Operator) regression, which is an extension of the ordinary least squares method. LASSO introduces a regularization term based on the sum of the absolute values of the regression coefficients in the loss function [38]. This adjustment prevents overfitting by performing embedded feature selection by shrinking some coefficients exactly to zero and thus excluding these predictors from the model. The LASSO loss function [38] is described by Eq. (1):

$$\text{loss function} = \min_{\alpha, \beta} \left( \sum (Y_i - (\alpha + \beta_1 X_1 + \dots + \beta_N X_N))^2 + \lambda \sum_{j=1}^N |\beta_j| \right) \quad (1)$$

where  $Y_i$  is the gold standard,  $X_1, \dots, X_N$  are the predictors,  $\alpha$  is the intercept, and  $\beta_j$  are the predictor coefficients. The term  $\lambda$  is a regularization parameter that controls the strength of the penalty. Here, the model is represented by Eq. (2):

$$Y = \alpha + \beta_1 X_1 + \dots + \beta_N X_N \quad (2)$$

where  $Y$  is the output of the model,  $X_1, \dots, X_N$  and are the predictors [38].

For each emotion recognition model, the lambda value that minimized the mean squared error (MSE) was determined. Model performance was evaluated by computing the mean absolute error (MAE), the MSE, the root mean squared error (RMSE), and the coefficient of determination ( $R^2$ ) for each regression model using the responses provided by the subjects and the predictions of the models. Moreover, all the levels of intensity (0–10) of each emotion were characterized in terms of mean, variance (expressed as standard deviation), skewness and kurtosis.

#### 3.3.2. Contribution of the predictors

For each emotion, the number of channels used for each feature to develop the model was computed, as well as the number of features (i.e., rhythm powers and involvement indexes) used for each channel to

develop the model. Such number was computed by counting the features for which the regression coefficients were different from zero. Eventually, to assess the consistency of model structure across emotions, we evaluated the similarity in the number of channels as well as in the number of features used using three different approaches: (1) Pearson's correlation coefficients, (2) a permutation test based on RMSE values, and (3) an ANOVA test. Statistical significance was set at 0.05 in all statistical analyses.

In order to evaluate the contribution of each involvement index to the emotional regression models and ensure the interpretability of our findings, a two-step importance analysis was performed: the first relied on the absolute value of LASSO regression coefficients, while the second relied on a resampling-based procedure. In the first step, for each emotion, the absolute values of LASSO regression coefficients were computed; since LASSO performs both regularization and feature selection, these coefficients provide a direct indication of the relative contribution of each feature to the prediction. Then, the median across channels was computed and, based on these aggregated values, the top-ranked predictors were identified both over all features and over involvement indexes only. The second step was specifically intended to assess the stability of feature selection, applying a resampling-based procedure. Specifically, each emotion-related LASSO model was repeatedly estimated with 70 iterations using randomly selected subsets corresponding to 70% of the available observations, where the use of 70% of the observations represents a compromise between preserving statistical power within each subsample and introducing sufficient variability across iterations to evaluate the robustness of feature selection. Moreover, 70 iterations were considered sufficient to obtain stable estimates of feature selection frequency while maintaining a reasonable computational cost. For each feature, both the proportion of iterations in which the feature was selected and the average regression coefficient across iterations were computed. The product of the absolute mean coefficient by the selection frequency was then used as a measure of feature importance. Again, the median across channels was computed, and then, for each emotion, the top-ranked predictors both over all features and over involvement indexes only were identified.

#### 3.3.3. Analysis of model fit

The distribution of residual absolute values was analyzed for each model within a repeated-measures framework to statistically compare model performance, as the same observations (i.e., the same subjects at the same time points) were evaluated across models. Normality was first assessed using the Lilliefors test. When the normality assumption was satisfied, a repeated-measures ANOVA was performed; otherwise, the non-parametric Friedman test was applied. If the test revealed a statistically significant difference, post-hoc pairwise comparisons were performed to identify which pairs of models differed significantly, setting the statistical difference at 0.05. To further investigate the outcomes of these comparisons, the paired  $t$ -test, in case of normal distributions, or otherwise the Wilcoxon signed-rank test for paired data was applied, using Bonferroni correction to control the family-wise error rate associated with multiple comparisons (given 10 comparisons, the corrected significance threshold was set at 0.005, that is 0.05/10). This approach allows for detailed identification of specific pairwise differences, as well as minimizing the risk of Type I errors. In addition, this combination of methods ensures a robust global test, reciprocally strengthening the validity of the results obtained. In the presence of multiple statistically significant differences, the best-performing model was identified as the one with the lowest central tendency of the absolute residuals, considering the mean when normality was satisfied and the median otherwise.

#### 3.3.4. Assessment of model stability

Model stability was assessed through two complementary approaches. In the first approach, the model stability was examined by analyzing the deviation of each absolute residual from the median absolute residual of the corresponding model (linear difference). This

approach provides a robust measure of variability around the median magnitude of the prediction errors and allows for a direct comparison of the consistency of model performance across observations. As in the previous analysis, the normality of these deviation distributions was assessed using the Lilliefors test. Differences between models were evaluated using repeated-measures ANOVA when normality assumptions were satisfied, or otherwise using the Friedman test. When a statistically significant difference was detected, post-hoc pairwise comparisons were performed to identify which pairs of models differed significantly. The paired  $t$ -test, in the case of normal distributions, or otherwise the Wilcoxon signed-rank test with Bonferroni correction, was applied to confirm the pairwise differences between models. In the presence of multiple statistically significant differences, the most stable model was identified as the one with the lowest variability of the absolute residuals, considering the variance when normality was satisfied and the median absolute deviation (MAD) otherwise. In the second approach, the model stability was examined by analyzing the dispersion of the prediction errors. Specifically, a bootstrap resampling procedure was implemented: 5000 bootstrap samples were generated by randomly resampling the original observations with replacement, with each bootstrap sample having the same size as the original dataset, so that some observations could appear multiple times while others might not be selected. Then, variance was used when the assumption of normality was satisfied by the original observation sample, whereas the MAD was used when the distributions deviated from normality. This process generated an empirical bootstrap distribution of the statistics of interest (i.e., variance or MAD) for each model, from which central tendency estimates (i.e., median variance or median MAD for each model) and confidence intervals (CI, ranging from the 2.5th to the 97.5th percentile, representing the 95% confidence interval for the dispersion estimates) were derived.

### 3.3.5. Predictive value assessment of the proposed features and model comparison

The incremental predictive value of the proposed involvement indexes over EEG rhythms alone was evaluated by implementing

additional emotion-related models, using three different feature configurations: the 6 EEG rhythms, the 37 involvement indexes, and the combination of EEG rhythms and involvement indexes. For each feature configuration, both the originally proposed LASSO regression model and an additional Support Vector Regressor (SVR) with a linear kernel were implemented and evaluated. A linear kernel was adopted to avoid introducing unnecessary model complexity, particularly given that the number of observations in our dataset is comparable to the number of available features. Model performance was evaluated using multiple regression metrics, including  $R^2$ , RMSE, MAE, and MSE, in order to provide a comprehensive assessment of predictive performance.

## 4. Results

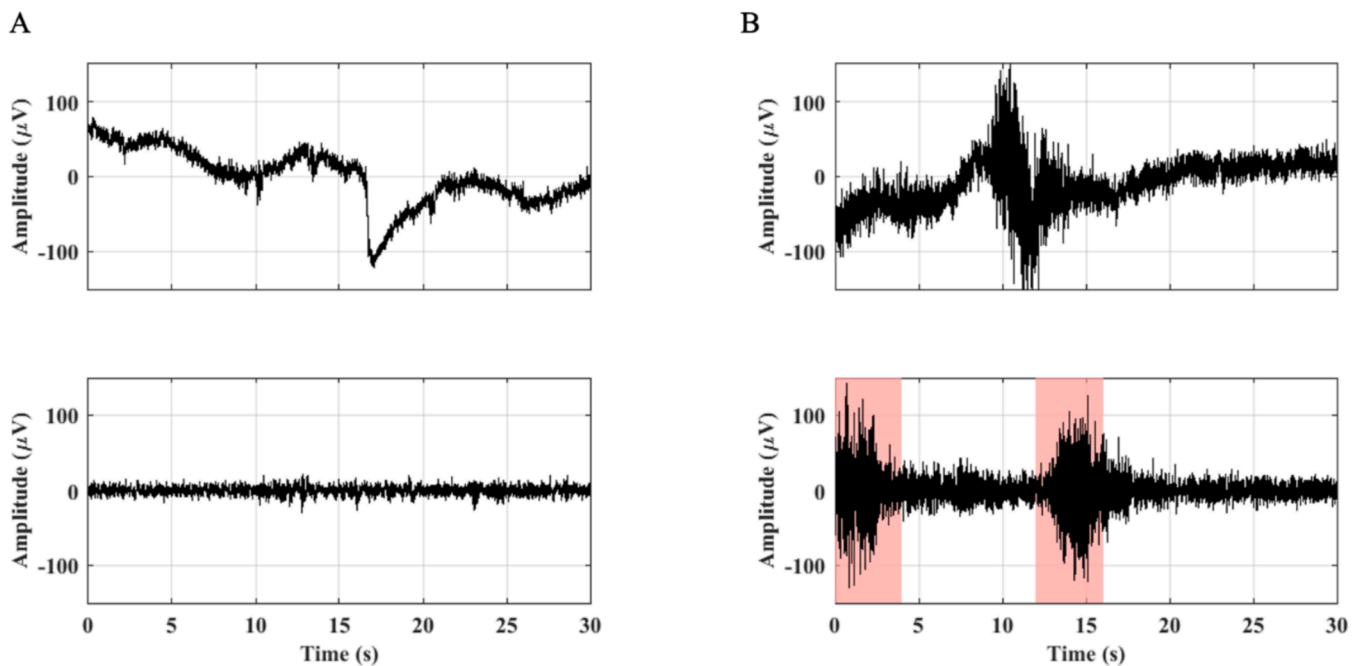
After bad-channel detection, at least 90% of 30-s windows per channel were suitable for the following analysis. Not only was no channel completely excluded from the analysis, but all were adequately represented in the implementation of the model. Indeed:

- up to 1% of windows were rejected for 50 channels,
- between 1% and 2% of windows were rejected for 9 channels,
- between 2% and 3% of windows were rejected for 3 channels,
- between 3% and 10% of windows were rejected for 2 channels.

Fig. 1 shows an example of EEG windows before (top) and after (bottom) noise removal and bad channel detection, with an accepted window shown in the left panels and a rejected window shown in the right panels; in both figures, noisy sub-windows are highlighted through a pink shadow.

All features presented a correlation coefficient lower than 0.75; thus, no feature was excluded when implementing the models. Model performance metrics (MAE, MSE, RMSE, and  $R^2$  coefficients) computed for each emotion recognition model are reported in Table 2, while mean, variance, skewness and kurtosis computed for each emotion intensity are reported in Table A1 of the Appendix A.

Fig. 2 presents a radar plot illustrating the number of relevant EEG



**Fig. 1.** Examples of two EEG windows (referred to channel TP7) before and after noise removal and bad channel detection. Panel A describes an example of accepted window, with the raw EEG window on the top and the corresponding preprocessed EEG on the bottom; panel B describes an example of rejected window, with the raw EEG window on the top and the corresponding preprocessed EEG window on the bottom, with noisy 2-s sub-windows highlighted through the shaded area in pink.

**Table 2**MAE, MSE, RMSE, and  $R^2$  coefficients for each emotion recognition model.

|       | Anger  | Happy  | Sadness | Disgust | Fear   |
|-------|--------|--------|---------|---------|--------|
| MAE   | 0.9020 | 1.1871 | 1.0036  | 1.0183  | 1.1235 |
| MSE   | 1.5194 | 2.4335 | 1.7829  | 1.8690  | 2.1578 |
| RMSE  | 1.2327 | 1.5600 | 1.3353  | 1.3671  | 1.4689 |
| $R^2$ | 0.7873 | 0.8106 | 0.7859  | 0.7908  | 0.8045 |

channels (out of 64) identified for regression, categorized by emotion and feature. Each emotion is represented by a distinct colored-line trend (blue for anger, orange for happiness, green for sadness, cyan for disgust, and violet for fear). The radial axes correspond to individual features, labeled along the outer ring, while the distance from the center indicates the number of EEG channels deemed relevant by the regression model for each feature.

Analogously, Fig. 3 presents a radar plot illustrating the number of relevant features (out of 43) identified for regression, categorized by channel and emotion. The radial axes correspond to individual channels, labeled along the outer ring, while the distance from the center indicates the number of EEG features deemed relevant by the regression model for each channel. When assessing the consistency of model structure across emotions, both Pearson's correlation and permutation test revealed that trends in Fig. 2 were all statistically different ( $p < 10^{-9}$ ), while trends in Fig. 3 resulted all non-statistically different ( $p > 0.05$ ). According to the ANOVA test, for the trends in Fig. 2, the trend of Sadness was statistically different from that of Anger ( $p < 10^{-5}$ ), Happiness ( $p < 0.015$ ), and Fear ( $p < 10^{-3}$ ), while no significant differences in the trends of Fig. 3 were found.

Heatmaps were generated for each emotion to visualize the distribution of feature importance across EEG channels, reported in Fig. A1 and Fig. A2 of Appendix A. In addition, to provide a concise summary of the most relevant predictors, the top features and top involvement indexes, derived from the median importance across channels, were reported for each of the five emotions in Table 3 and Table 4. This dual

representation allows both a comprehensive visualization of all features across the scalp and an attempt at identifying the most informative predictors specific to each emotion.

Considering the absolute residuals, the distributions tested by the Lilliefors test were found to be non-normal. Therefore, the Friedman test was applied, which showed a statistically significant difference between the models ( $p = 5.14 \times 10^{-34}$ ). Given the statistical significance of the difference, a pairwise comparison was conducted, whose results are reported in Table 5, revealing statistically significant differences for all model comparisons, except for Happiness vs. Fear and Sadness vs. Disgust, as reported in the  $p$ -values in the second column of Table 5. Wilcoxon signed-rank test (applied given the non-normality of the distribution) confirmed these outcomes (the pairs of models not showing statistical differences were the same), as appreciated from the  $p$ -values reported in the third column of Table 5.

Considering the first approach for the model-stability assessment, the deviations of the absolute residuals from the median absolute residual of the corresponding model resulted in non-normal distributions. The Friedman test showed a statistically significant difference among the models ( $p = 1.93 \times 10^{-37}$ ), indicating that overall, the models differed in their stability. Then, both Friedman-based pairwise comparison and Bonferroni-corrected Wilcoxon methods yielded consistent results. The only pair that did not show a statistically significant difference was Sadness vs. Disgust. This is shown in Table 6, where both pairwise and Wilcoxon  $p$ -values are reported.

Table 7 shows the median of the absolute residuals and the median of the MAD for each emotion-related model, where the smallest median absolute error value indicates the highest level of predictive accuracy, while the smallest MAD value indicates the most stable model.

Considering the second approach for the model-stability assessment, since the absolute residual distributions were non-normal, MAD was computed on each bootstrap sample. The median and CIs for the MAD values are shown in Table 8, where smaller median and CI width values of the MAD indicate more stable models.

The results of the comparison between LASSO and SVR in the three

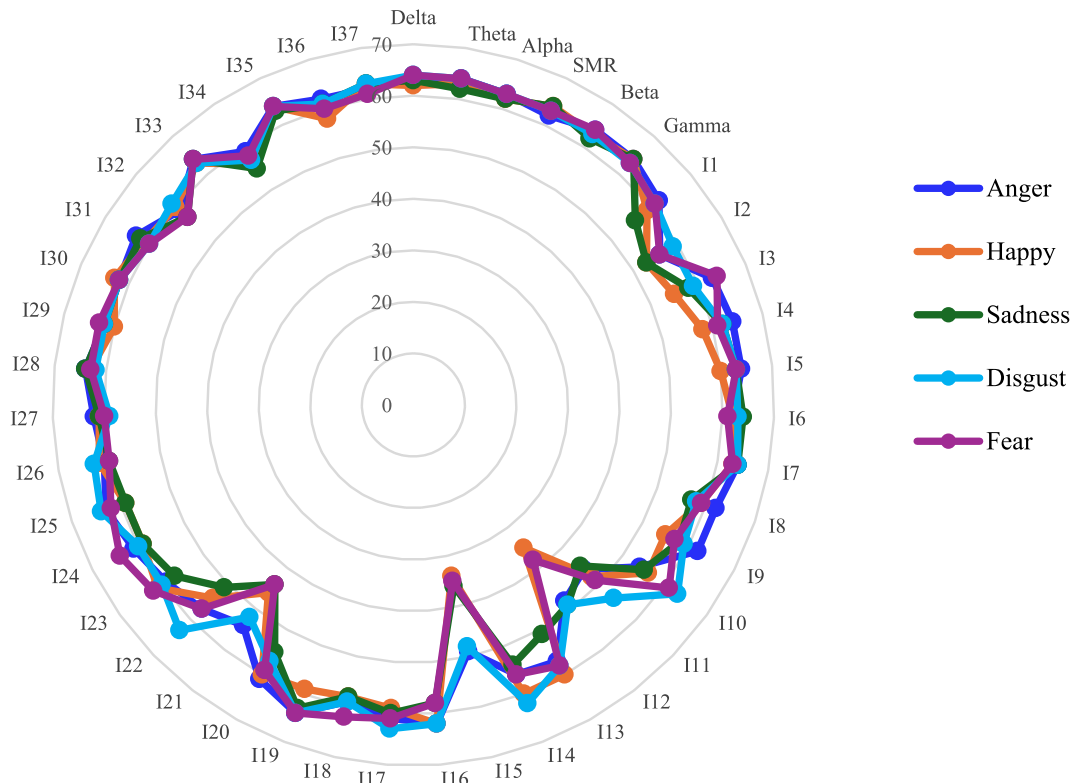


Fig. 2. Number of channels used, for each feature, to develop the regression model of each emotion.

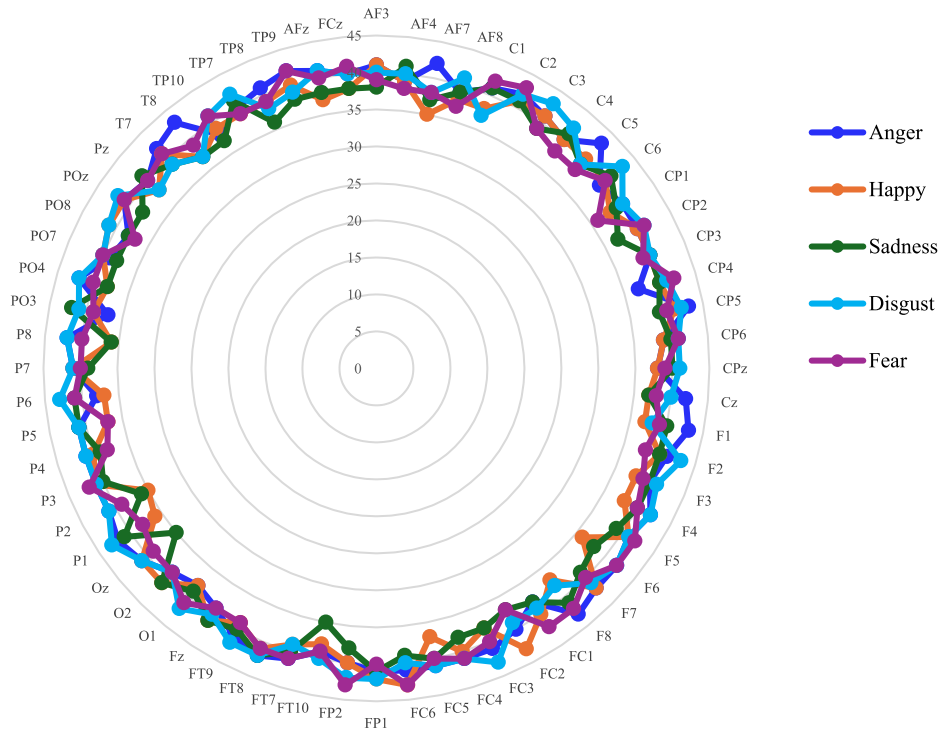


Fig. 3. Number of features used, for each channel, to develop the regression model of each emotion.

Table 3

Top-ranked predictors identified by the LASSO models for each emotion (Anger, Happy, Sadness, Disgust, and Fear). Feature importance values (reported in parentheses) correspond to the median importance across EEG channels, computed as the absolute regression coefficients of the LASSO models. The upper section (Features) reports the ten most relevant predictors considering both EEG rhythms and involvement indexes, while the lower section (Indexes) reports the ten most relevant involvement indexes only.

|          |         |                           |                           |                           |                           |                           |                           |                           |                           |                           |                           |
|----------|---------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Features | Anger   | I <sub>25</sub><br>(1.52) | I <sub>28</sub><br>(1.48) | I <sub>3</sub><br>(1.35)  | Beta<br>(1.34)            | I <sub>8</sub><br>(1.17)  | I <sub>36</sub><br>(1.13) | I <sub>6</sub><br>(1.11)  | Alpha<br>(1.11)           | I <sub>30</sub><br>(1.10) | I <sub>35</sub><br>(1.10) |
|          | Happy   | I <sub>27</sub><br>(2.07) | I <sub>25</sub><br>(2.04) | Alpha<br>(1.97)           | I <sub>37</sub><br>(1.97) | I <sub>26</sub><br>(1.73) | I <sub>30</sub><br>(1.70) | I <sub>11</sub><br>(1.69) | I <sub>16</sub><br>(1.60) | I <sub>23</sub><br>(1.60) | Beta<br>(1.56)            |
|          | Sadness | I <sub>28</sub><br>(1.45) | I <sub>8</sub><br>(1.38)  | I <sub>31</sub><br>(1.35) | Beta<br>(1.35)            | I <sub>26</sub><br>(1.34) | Theta<br>(1.29)           | I <sub>27</sub><br>(1.22) | I <sub>29</sub><br>(1.21) | I <sub>25</sub><br>(1.21) | I <sub>37</sub><br>(1.17) |
|          | Disgust | Alpha<br>(1.87)           | I <sub>27</sub><br>(1.61) | Beta<br>(1.40)            | I <sub>26</sub><br>(1.38) | I <sub>37</sub><br>(1.37) | Theta<br>(1.23)           | I <sub>25</sub><br>(1.23) | I <sub>16</sub><br>(1.22) | I <sub>36</sub><br>(1.14) | I <sub>19</sub><br>(1.13) |
|          | Fear    | I <sub>37</sub><br>(1.52) | I <sub>34</sub><br>(1.51) | Alpha<br>(1.44)           | I <sub>25</sub><br>(1.42) | Beta<br>(1.37)            | I <sub>4</sub><br>(1.31)  | I <sub>24</sub><br>(1.30) | I <sub>36</sub><br>(1.30) | I <sub>23</sub><br>(1.28) | I <sub>16</sub><br>(1.25) |
| Indexes  | Anger   | I <sub>25</sub><br>(1.52) | I <sub>28</sub><br>(1.48) | I <sub>3</sub><br>(1.35)  | I <sub>8</sub><br>(1.17)  | I <sub>36</sub><br>(1.13) | I <sub>6</sub><br>(1.11)  | I <sub>30</sub><br>(1.10) | I <sub>35</sub><br>(1.10) | I <sub>31</sub><br>(1.10) | I <sub>16</sub><br>(1.09) |
|          | Happy   | I <sub>27</sub><br>(2.07) | I <sub>25</sub><br>(2.04) | I <sub>37</sub><br>(1.97) | I <sub>26</sub><br>(1.73) | I <sub>30</sub><br>(1.70) | I <sub>11</sub><br>(1.69) | I <sub>16</sub><br>(1.60) | I <sub>23</sub><br>(1.60) | I <sub>36</sub><br>(1.50) | I <sub>28</sub><br>(1.47) |
|          | Sadness | I <sub>28</sub><br>(1.45) | I <sub>8</sub><br>(1.38)  | I <sub>31</sub><br>(1.35) | I <sub>26</sub><br>(1.34) | I <sub>27</sub><br>(1.22) | I <sub>29</sub><br>(1.21) | I <sub>25</sub><br>(1.21) | I <sub>37</sub><br>(1.17) | I <sub>5</sub><br>(1.11)  | I <sub>14</sub><br>(1.03) |
|          | Disgust | I <sub>27</sub><br>(1.61) | I <sub>26</sub><br>(1.38) | I <sub>37</sub><br>(1.37) | I <sub>25</sub><br>(1.23) | I <sub>16</sub><br>(1.22) | I <sub>36</sub><br>(1.14) | I <sub>19</sub><br>(1.13) | I <sub>28</sub><br>(1.13) | I <sub>33</sub><br>(1.12) | I <sub>31</sub><br>(1.04) |
|          | Fear    | I <sub>37</sub><br>(1.52) | I <sub>34</sub><br>(1.51) | I <sub>25</sub><br>(1.42) | I <sub>4</sub><br>(1.31)  | I <sub>24</sub><br>(1.30) | I <sub>36</sub><br>(1.30) | I <sub>23</sub><br>(1.28) | I <sub>16</sub><br>(1.25) | I <sub>27</sub><br>(1.23) | I <sub>31</sub><br>(1.22) |

feature configurations are reported in Fig. 4. Across all emotions, models based solely on the six EEG rhythms show the lowest predictive performance. In contrast, models trained using the involvement indexes consistently achieve higher predictive performance, suggesting that these indexes contain relevant information for predicting emotional scores. A further improvement is observed when rhythms and involvement indexes are combined.

## 5. Discussion

This study analyzed the EMAP database designed to support research investigating emotional states induced by evocative short videos with no audio. The videos were specifically selected by the database developers

with the aim of including a range of emotional valence and arousal, so able to induce a wide range of emotional states [32]. The choice of the EMAP database was prompted by the relatively large number of subjects, as well as the comprehensive coverage of brain activity recorded using 64 electrodes. Having a large amount of data is very important when developing models for emotion recognition based on physiological signals. A large and diverse subject pool allows the model to learn generalizable patterns across individuals rather than overfitting the specific traits of a few participants. This is especially important for developing subject-independent models. Additionally, a larger sample size increases the statistical power of the study, enabling the detection of subtle and consistent differences between emotional states and reducing the influence of noise or outliers. Comprehensive brain coverage

**Table 4**

Top-ranked predictors identified by the LASSO models for each emotion (Anger, Happy, Sadness, Disgust, and Fear). Feature importance values (reported in parentheses) correspond to the median importance across EEG channels, computed as the product of the absolute mean coefficient and the selection frequency of the LASSO model iterations. The upper section (Features) reports the ten most relevant predictors considering both EEG rhythms and involvement indexes, while the lower section (Indexes) reports the ten most relevant involvement indexes only.

|          |         |                           |                           |                           |                           |                           |                           |                           |                           |                           |                           |
|----------|---------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Features | Anger   | I <sub>26</sub><br>(1.45) | I <sub>25</sub><br>(1.43) | I <sub>36</sub><br>(1.36) | I <sub>3</sub><br>(1.14)  | I <sub>37</sub><br>(1.12) | Alpha<br>(1.09)           | Beta<br>(1.08)            | I <sub>27</sub><br>(1.06) | I <sub>4</sub><br>(1.06)  | I <sub>29</sub><br>(1.05) |
|          | Happy   | I <sub>25</sub><br>(2.00) | Alpha<br>(1.89)           | I <sub>26</sub><br>(1.89) | I <sub>27</sub><br>(1.81) | I <sub>37</sub><br>(1.59) | I <sub>30</sub><br>(1.58) | I <sub>31</sub><br>(1.45) | Beta<br>(1.40)            | I <sub>23</sub><br>(1.39) | I <sub>1</sub><br>(1.38)  |
|          | Sadness | I <sub>26</sub><br>(1.36) | Beta<br>(1.33)            | Alpha<br>(1.32)           | I <sub>31</sub><br>(1.19) | I <sub>27</sub><br>(1.12) | I <sub>28</sub><br>(1.11) | I <sub>25</sub><br>(1.10) | I <sub>8</sub><br>(1.08)  | I <sub>37</sub><br>(1.00) | I <sub>18</sub><br>(0.98) |
|          | Disgust | I <sub>36</sub><br>(1.84) | I <sub>26</sub><br>(1.54) | I <sub>37</sub><br>(1.43) | Alpha<br>(1.40)           | I <sub>16</sub><br>(1.38) | I <sub>25</sub><br>(1.36) | I <sub>24</sub><br>(1.29) | I <sub>27</sub><br>(1.28) | I <sub>31</sub><br>(1.25) | I <sub>4</sub><br>(1.24)  |
|          | Fear    | I <sub>25</sub><br>(1.64) | I <sub>36</sub><br>(1.61) | I <sub>37</sub><br>(1.55) | I <sub>16</sub><br>(1.52) | Alpha<br>(1.49)           | I <sub>23</sub><br>(1.48) | Beta<br>(1.43)            | I <sub>27</sub><br>(1.34) | I <sub>4</sub><br>(1.32)  | I <sub>30</sub><br>(1.31) |
| Indexes  | Anger   | I <sub>26</sub><br>(1.45) | I <sub>25</sub><br>(1.43) | I <sub>36</sub><br>(1.36) | I <sub>3</sub><br>(1.14)  | I <sub>37</sub><br>(1.12) | I <sub>27</sub><br>(1.06) | I <sub>4</sub><br>(1.06)  | I <sub>29</sub><br>(1.05) | I <sub>22</sub><br>(1.01) | I <sub>30</sub><br>(1.01) |
|          | Happy   | I <sub>25</sub><br>(2.00) | I <sub>26</sub><br>(1.89) | I <sub>27</sub><br>(1.81) | I <sub>37</sub><br>(1.59) | I <sub>30</sub><br>(1.58) | I <sub>31</sub><br>(1.45) | I <sub>23</sub><br>(1.39) | I <sub>1</sub><br>(1.38)  | I <sub>4</sub><br>(1.34)  | I <sub>24</sub><br>(1.30) |
|          | Sadness | I <sub>26</sub><br>(1.36) | I <sub>31</sub><br>(1.19) | I <sub>27</sub><br>(1.12) | I <sub>28</sub><br>(1.11) | I <sub>25</sub><br>(1.10) | I <sub>8</sub><br>(1.08)  | I <sub>37</sub><br>(1.00) | I <sub>18</sub><br>(0.98) | I <sub>32</sub><br>(0.96) | I <sub>30</sub><br>(0.96) |
|          | Disgust | I <sub>36</sub><br>(1.84) | I <sub>26</sub><br>(1.54) | I <sub>37</sub><br>(1.43) | I <sub>16</sub><br>(1.38) | I <sub>25</sub><br>(1.36) | I <sub>24</sub><br>(1.29) | I <sub>27</sub><br>(1.28) | I <sub>31</sub><br>(1.25) | I <sub>4</sub><br>(1.24)  | I <sub>28</sub><br>(1.19) |
|          | Fear    | I <sub>25</sub><br>(1.64) | I <sub>36</sub><br>(1.61) | I <sub>37</sub><br>(1.55) | I <sub>16</sub><br>(1.52) | I <sub>23</sub><br>(1.48) | I <sub>27</sub><br>(1.34) | I <sub>4</sub><br>(1.32)  | I <sub>30</sub><br>(1.31) | I <sub>1</sub><br>(1.29)  | I <sub>34</sub><br>(1.26) |

**Table 5**

The  $p$ -values from the pairwise comparisons of model absolute errors between different emotions. The first column indicates the emotion-related model comparison, the second column reports the pairwise  $p$ -values, and the third column presents the Wilcoxon signed-rank test  $p$ -values. The significance level for both tests is set at 0.05, while the corrected significance level after Bonferroni correction is 0.005. Bold cell indicates cases where the  $p$ -values do not reach statistical significance.

| Model 1 vs Model 2        | $p$ -value (pairwise)  | $p$ -value (Wilcoxon)                   |
|---------------------------|------------------------|-----------------------------------------|
| Anger vs Happy            | 0                      | $4.29 \times 10^{-33}$                  |
| Anger vs Sadness          | $2.50 \times 10^{-3}$  | $5.74 \times 10^{-5}$                   |
| Anger vs Disgust          | $1.54 \times 10^{-8}$  | $1.50 \times 10^{-11}$                  |
| Anger vs Fear             | $2.63 \times 10^{-20}$ | $4.72 \times 10^{-25}$                  |
| Happy vs Sadness          | $7.07 \times 10^{-14}$ | $1.66 \times 10^{-19}$                  |
| Happy vs Disgust          | $9.52 \times 10^{-7}$  | $6.45 \times 10^{-12}$                  |
| <b>Happy vs Fear</b>      | <b>0.217</b>           | <b>0.0121</b>                           |
| <b>Sadness vs Disgust</b> | <b>0.116</b>           | <b><math>9.80 \times 10^{-3}</math></b> |
| Sadness vs Fear           | $1.83 \times 10^{-7}$  | $5.77 \times 10^{-12}$                  |
| Disgust vs Fear           | 0.0109                 | $1.94 \times 10^{-6}$                   |

**Table 6**

The  $p$ -values from the pairwise comparisons of model error deviations between different emotions. The first column indicates the emotion-related model comparison, the second column reports the pairwise  $p$ -values, and the third column presents the Wilcoxon signed-rank test  $p$ -values. The significance level for both tests is set at 0.05, while the corrected significance level after Bonferroni correction is 0.005. Bold cell indicates cases where the  $p$ -values do not reach statistical significance.

| Model 1 vs Model 2        | $p$ -value (pairwise)  | $p$ -value (Wilcoxon)  |
|---------------------------|------------------------|------------------------|
| Anger vs Happy            | 0                      | $1.38 \times 10^{-28}$ |
| Anger vs Sadness          | $5.20 \times 10^{-3}$  | $3.32 \times 10^{-5}$  |
| Anger vs Disgust          | $6.19 \times 10^{-6}$  | $4.75 \times 10^{-8}$  |
| Anger vs Fear             | $2.46 \times 10^{-19}$ | $3.54 \times 10^{-20}$ |
| Happy vs Sadness          | $3.54 \times 10^{-17}$ | $3.68 \times 10^{-16}$ |
| Happy vs Disgust          | $1.80 \times 10^{-11}$ | $2.67 \times 10^{-10}$ |
| Happy vs Fear             | 0.0261                 | $1.70 \times 10^{-3}$  |
| <b>Sadness vs Disgust</b> | <b>0.536</b>           | <b>0.0255</b>          |
| Sadness vs Fear           | $1.93 \times 10^{-7}$  | $2.24 \times 10^{-9}$  |
| Disgust vs Fear           | $4.41 \times 10^{-4}$  | $5.74 \times 10^{-5}$  |

through widespread electrode placement is important in EEG-based emotion recognition because emotional processing may involve multiple brain regions, and a broader coverage enables the extraction of more

**Table 7**

The medians and the MADs of model absolute errors.

|        | Anger | Happy | Sadness | Disgust | Fear  |
|--------|-------|-------|---------|---------|-------|
| Median | 0.739 | 1.01  | 0.794   | 0.846   | 0.972 |
| MAD    | 0.483 | 0.645 | 0.518   | 0.552   | 0.618 |

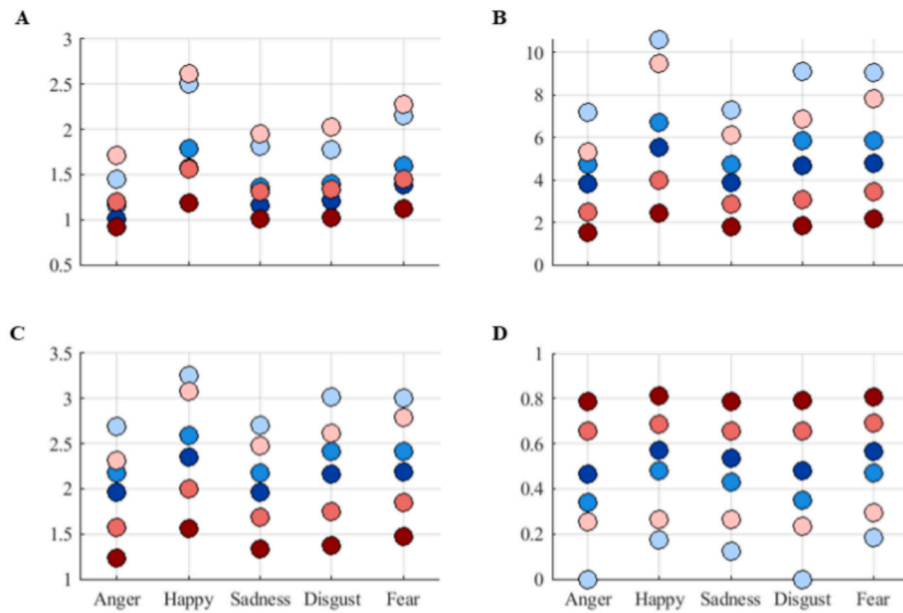
**Table 8**

The table shows the median, as well as the lower bound, the upper bound, and the width of the confidence interval of the bootstrap estimates of the MAD for each model.

| Model   | Median | CI Lower Bound | CI Upper Bound | CI Width |
|---------|--------|----------------|----------------|----------|
| Anger   | 0.482  | 0.454          | 0.506          | 0.0520   |
| Happy   | 0.643  | 0.613          | 0.679          | 0.0660   |
| Sadness | 0.519  | 0.495          | 0.547          | 0.0520   |
| Disgust | 0.550  | 0.523          | 0.579          | 0.0560   |
| Fear    | 0.617  | 0.589          | 0.643          | 0.0540   |

spatially distributed features. Moreover, collecting responses for five different emotions on an intensity scale from 0 to 10 allows to account for the coexistence of different emotions with different intensities, rather than a binary approach (emotion absent/present). This is important because emotional experiences in real life are rarely binary; people often feel multiple emotions at the same time, each with varying intensities. A binary approach, even if it is widely used, oversimplifies the complexity of human affect and can lead to models that fail to capture subtle or mixed emotional states.

Starting from the Raw Data set, we resampled at 500 Hz, as a trade-off to preserve sufficient spectral resolution for emotion-related EEG features, while limiting computational demands, thereby supporting real-time processing and the feasibility of deployment on wearable EEG devices. Next, for the preprocessing, we followed the procedure adopted by the developers of the dataset to go from the Raw Data set to the Clean Data set, to facilitate comparability with previous research based on the same dataset. The only difference is that we used the runica algorithm for ICA decomposition, instead of the binica algorithm, as it ensures full numerical transparency and reproducibility within the MATLAB environment. To further enhance the reliability of the data, an automatic bad-channel detection based on signal standard deviation was performed [39]. This step was crucial, as it ensured more reliable and accurate data for further analysis. Moreover, this approach offers an



**Fig. 4.** Comparison of prediction performance obtained with SVR and LASSO models using different EEG feature sets. The four panels report the evaluation metrics for the five emotions (Anger, Happy, Sadness, Disgust, Fear): (A) MAE, (B) MSE, (C) RMSE, and (D)  $R^2$ . Blue-shaded markers represent SVR models, while red-shaded markers represent LASSO models. Different shades refer to the feature set used: lighter, for rhythms only, intermediate, for indexes only, and darker, for the combination of rhythms and indexes.

objective and reproducible alternative to manual inspection that, although effective, is inherently subjective and time-consuming as it relies on expert annotations [40,41]. Unlike these conventional methods that rely on visual assessment or simple amplitude thresholds, in the literature several quantitative criteria have been proposed to assess EEG signal quality, including amplitude or covariance thresholds, peak-to-peak amplitude, spectral energy, and entropy-based measures [42–45]. Among these, the standard deviation-based criterion here chosen enabled the systematic identification of noisy sub-windows, being mainly sensitive to large-amplitude deviations incompatible with physiological brain activity. In addition, the fact that each 30-s window was further divided into 2-s non-overlapping sub-windows allowed us to localize transient artifacts, thus improving the temporal accuracy of noise detection (Fig. 1).

After the extraction of the EEG rhythms and computation of their powers, we used topographic maps, not reported in the results because of their large number, as a convenient graphical tool to identify detectable differences among different emotional states through visual inspection. When considering the power of the rhythms alone, we observed higher activity in the occipital area for all cases. This may be explained by the occipital lobe's role in visual processing, with distinct cortical areas dedicated to different aspects of visual perception, from basic sensory input to more complex interpretation and recognition. As expected, by visual inspection, it was hard to correctly interpret the large number of features and, for some of them, a clear difference could not be appreciated [46].

A machine learning approach was exploited, since it facilitates the handling of high-dimensional data, providing better interpretability of the results with respect to deep learning approaches. LASSO regression is able to automatically select important features and to remove irrelevant ones, reducing the impact of highly correlated predictors and improving generalization. This makes it especially suitable for our dataset, where many predictors were available and feature selection was critical for robust model performance. As reported in Table 2, good performance was achieved for all regression models, with  $R^2$  values always above 0.75, indicating that the formulated models are reliable in discriminating between emotional states. Among the five emotions, Anger yielded the lowest prediction errors (MAE = 0.902, RMSE =

1.233), suggesting that its intensity was captured most accurately. In contrast, Happiness produced the highest error values (MAE = 1.187, RMSE = 1.560), indicating greater variability or complexity in the corresponding data patterns. The remaining emotions (i.e., Sadness, Disgust, and Fear) exhibited intermediate performance levels, with relatively small differences in error metrics. To further evaluate the performance and reliability of the regression models, skewness and kurtosis were jointly analyzed across each intensity level (0–10) for all five emotions. Skewness analysis revealed a common trend: positive skew at low intensity levels (0–2), indicating a tendency to overpredict emotional responses when weakly expressed, and negative skew at moderate to high levels (5–10), pointing to underprediction when emotional intensity increased. This was particularly evident in the models for Anger and Disgust, where skewness values exceeded  $\pm 0.5$ . Instead, kurtosis values were generally above 3 for most emotions and intensity levels, especially for Anger and Disgust, indicating that while many predictions clustered around the mean, some errors were present towards the extremes. In contrast, models for Fear and Sadness showed more moderate kurtosis (2.5–3.5) and near-zero skewness, reflecting more stable and balanced predictions. The joint analysis of performance metrics and distribution characteristics suggests that all models exhibited reasonable predictive behavior. However, the Anger and Disgust models may benefit from refinements to reduce bias and improve robustness toward extreme intensity values, while the models for Happiness, Sadness, and Fear demonstrated more stable and balanced predictions across the full intensity range.

Referring to the analysis of the model fit to the observations, the pairwise comparisons indicated that the Anger model differs significantly from all the other models. Considering the median absolute error values (Table 7), the Anger model exhibits the lowest median among all models, indicating the highest level of predictive accuracy. Thus, it tends to produce smaller prediction errors than the other emotion-related models. Moreover, when considering the assessment of the model stability, the pairwise comparisons indicated that the Anger, Happy and Fear model differ significantly from each other. By examining the deviation values of the absolute residuals, the Anger model exhibits the lowest MAD among all models (Table 7), indicating that its residuals are the most concentrated around the median value. In contrast, the

Happiness model exhibits the highest MAD, indicating that its residuals are more widely dispersed around the median value. The CIs for the MAD of the Anger model were the narrowest, further confirming that this model has the most stable performance. In contrast, the Happiness model exhibited the widest CIs, reflecting the greatest variability in error prediction and confirming that it is the least stable model among those considered (Table 8). Furthermore, the empirical distribution of the MAD values derived from the bootstrap procedure showed that the Anger model consistently had the lowest MAD across all bootstrap samples, reaffirming its superior stability. Once again, the Happiness model proved to be the least stable, showing the greatest variability in the distribution of the MAD. Therefore, all the results agree that the Anger model is the most accurate and stable, while the Happiness model is the least stable.

Taking into account similarity assessment in the number of channels used for each feature and in the number of features used for each channel, Pearson's correlation coefficients and permutation test gave the same outcomes: the trends of the number of channels per feature were all different (Fig. 2,  $p < 10^{-3}$ ), while there were no differences in the trends of the numbers of features per channel (Fig. 3,  $p > 0.05$ ). The ANOVA test confirmed the outcomes related to these last trends, while for the first ones it pointed out that the trend of Sadness was significantly different from that of Anger ( $p < 10^{-5}$ ), Happiness ( $p < 0.01$ ), and Fear ( $p < 10^{-3}$ ) (Fig. 2). This confirmed that the trends in the number of channels for each feature contributing to the discrimination of Happy and Fear appear to be related, as well as Anger, Sadness, and Fear.

This study underscores the importance of a multiparametric approach in the context of EEG-based emotion recognition, combining a wide range of features and channels to comprehensively capture brain activity and model the complexity of emotional responses. All the channels resulted to provide useful information between 91.16% and 96.74% of total cases, when considering the different emotions and features. When considering the features, they also proved to be useful for over 89% of cases, up to 99.69%, when considering the different emotions and channels. The only exceptions are  $I_{11}$ ,  $I_{21}$ ,  $I_{12}$ , and  $I_{15}$ , that resulted informative respectively in 84.69%, 80%, 79.38 and 65.63% of cases.

The multiparametric approach was confirmed as the best approach by the implementation of both LASSO and SVR in the three feature configurations. Indeed, the joint feature set leads to the best performance across all emotions, suggesting that the proposed involvement indexes provide complementary information. In addition, in the comparison between LASSO and SVR, the former consistently achieved better predictive performance than the latter across all feature configurations and emotions, as reflected by higher  $R^2$  values and lower error metrics (MAE, MSE, and RMSE), proving its effectiveness for the present prediction task (Fig. 4).

Some limitations of the work should be acknowledged. First, the study relies on a gold standard that is inherently subjective, as the targeted emotions are self-assessed by the participants, and the emotional experience may vary from individual to individual, being also influenced by mood, memory, sensitivity, and other internal or external factors. This subjectivity introduces variability that may affect the consistency of model training and evaluation. Second, the study is based on a single dataset recorded under specific experimental conditions. Additional observations collected under comparable settings are crucial for increasing the sample size, improving the reliability of the models, reducing overfitting, and ensuring that the models generalize well to unseen data. Furthermore, another potential limitation may arise from the choice of a linear approach. Even if the emotion-related neurophysiological mechanisms are likely to involve nonlinear dynamics, a

linear approach still allowed us to outline the dominant association between EEG spectral features and emotional intensity. In this context, the assumption of linearity can be seen as a reasonable approximation, at least for preliminary considerations when proposing a new approach, prioritizing model interpretability and robustness during this foundational stage of our research. Indeed, most EEG-based emotion-recognition studies formulate the problem as a classification task, where emotional states are categorized into discrete classes. Since our approach considers the possible coexistence of different emotional states expressed at varying intensities, we believe that our LASSO-based linear regression approach can be used as a benchmark model, that is a performance reference level before evaluating whether more sophisticated approaches actually provide meaningful improvements.

Building on the current results, several promising directions emerge for further research. Future studies could further investigate the role of specific involvement indexes to gain deeper insights and explore which subsets of indexes are most informative for specific emotions, potentially uncovering novel biomarkers of affective states. Analogously, interactions between different brain regions could be further investigated, with the perspective of using low-density EEG systems, which could improve usability and feasibility in real-world scenario applications, allowing emotion recognition approaches to move beyond controlled laboratory environments. Indeed, a high-resolution mapping step is essential before any attempt at scalability, as it enables the determination of informative cortical areas. Without this initial investigation, reducing the number of channels would be arbitrary and could compromise performance. The feature selection performed by LASSO provides encouraging evidence for the use of low-density systems. The fact that the models achieve high performance by relying on a sparse set of stable predictors suggests that a limited number of strategically placed electrodes would likely be sufficient to maintain acceptable accuracy. Another important extension would be the integration of additional physiological modalities, such as electrodermal activity, heart-rate variability, or facial electromyography, within a multimodal analysis approach. These signals can provide complementary information to EEG, significantly enhancing the robustness of emotion recognition models. The present study also highlighted that certain emotions exhibited statistically distinct feature distributions across channels. This observation could motivate research into emotion-specific feature engineering or dynamic modelling approaches that adapt to different affective contexts.

## 6. Conclusions

This study aimed to address the task of emotion recognition, combining EEG rhythm powers with 37 involvement indexes to develop regression models capable of identifying and rating different emotional states. Although acknowledging the limitations of the study linked to the implicit subjectivity of the gold standard and to the single experimental paradigm considered, the obtained results demonstrated that involvement indexes, traditionally applied to assess cognitive states such as attention and workload, can effectively distinguish between emotional states. In addition, this study highlighted the potential of involvement indexes and multi-emotion modeling to advance the development of nuanced EEG-based emotion recognition systems, able to capture emotions as coexisting states expressed at varying intensities.

Thus, our findings contribute to broadening the use of involvement indexes in affective modeling and provide a more realistic and interpretable framework for representing human emotional experience. Future studies will validate the proposed analysis pipeline and regression models with low-density EEG systems, with the ultimate goal of

integrating this monitoring approach into wearable sensors for real-life applications.

### CRedit authorship contribution statement

**Alessandro Dell'Orletta:** Writing – original draft, Visualization, Software, Formal analysis. **Erica Iammarino:** Writing – original draft, Visualization, Methodology, Conceptualization. **Agnese Sbrollini:** Writing – review & editing, Formal analysis. **Laura Burattini:** Writing – review & editing, Project administration, Conceptualization. **Ilaria Marcantoni:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

### Funding

This research has been partially funded by: the project “PR Marche Fondo Sociale Europeo Plus 2021/2027”; Vitality – Project Code

ECS00000041, CUP I33C22001330007 – funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5 – 'Creation and strengthening of innovation ecosystems,' construction of 'territorial leaders in R&D' – Innovation Ecosystems – Project 'Innovation, digitalization and sustainability for the diffused economy in Central Italy – VITALITY' Call for tender No. 3277 of 30/12/2021, and Concession Decree No. 0001057.23–06-2022 of Italian Ministry of University funded by the European Union – NextGenerationEU; Fondazione Cariverona through the project Young Research Group for Future (52563-2022.0313).

### Declaration of competing interest

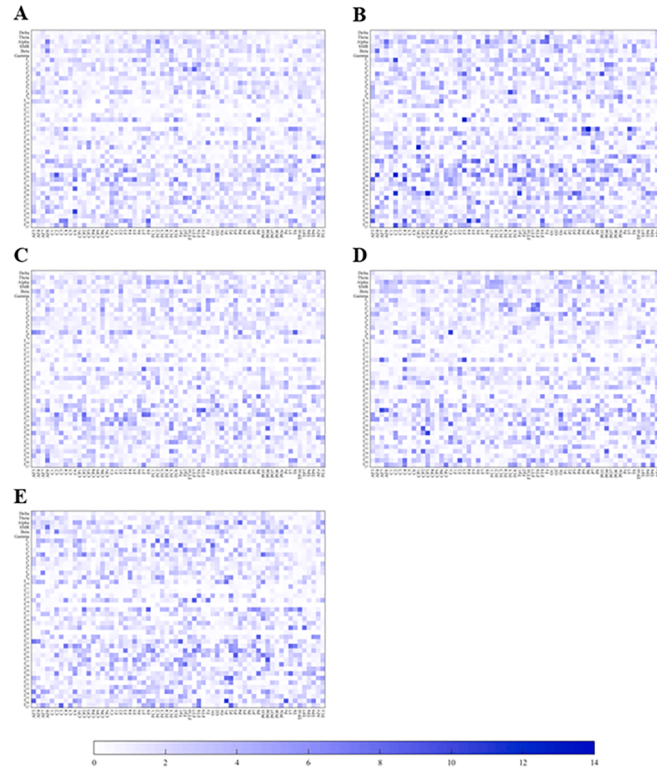
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A

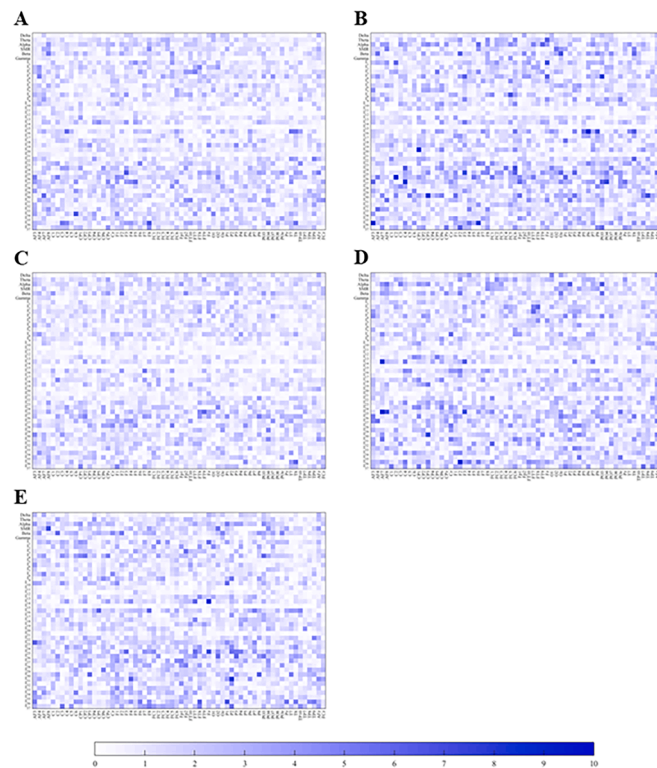
**Table A1**

Values of mean, variance (standard deviation, SD), skewness and kurtosis assessed for each emotion intensity.

| Emotion intensity | Metric   | Anger   | Happy   | Sadness | Disgust | Fear    |
|-------------------|----------|---------|---------|---------|---------|---------|
| 0                 | Mean     | 0.3277  | 0.7323  | 0.4834  | 0.4034  | 0.5449  |
|                   | SD       | 1.0193  | 1.3587  | 1.1049  | 1.1228  | 1.2309  |
|                   | Skewness | 0.3187  | 0.2738  | 0.4142  | 0.3200  | 0.2761  |
|                   | Kurtosis | 3.8635  | 3.2905  | 3.6372  | 3.9742  | 3.5832  |
| 1                 | Mean     | 1.2712  | 1.6295  | 1.3403  | 1.3890  | 1.5199  |
|                   | SD       | 1.0970  | 1.3422  | 1.0940  | 1.0708  | 1.2970  |
|                   | Skewness | 0.4765  | 0.3930  | 0.2270  | 0.6282  | 0.1168  |
|                   | Kurtosis | 3.9527  | 3.3567  | 3.0854  | 3.5899  | 3.3551  |
| 2                 | Mean     | 1.8942  | 2.4782  | 2.0134  | 2.0615  | 2.1630  |
|                   | SD       | 1.0542  | 1.3584  | 1.0224  | 1.0904  | 1.1707  |
|                   | Skewness | 0.4319  | 0.0501  | 0.0173  | 0.1343  | 0.1038  |
|                   | Kurtosis | 3.5303  | 3.6548  | 3.4635  | 3.3880  | 2.9907  |
| 3                 | Mean     | 2.7327  | 3.1550  | 2.8616  | 2.8926  | 3.0107  |
|                   | SD       | 1.0274  | 1.3270  | 1.1504  | 1.3633  | 1.2200  |
|                   | Skewness | −0.3035 | 0.4341  | −0.2679 | −0.5822 | 0.1368  |
|                   | Kurtosis | 3.9435  | 4.1657  | 4.821   | 3.3318  | 3.0538  |
| 4                 | Mean     | 3.4123  | 3.7005  | 3.5862  | 3.3732  | 3.4542  |
|                   | SD       | 0.8792  | 1.2510  | 1.1490  | 1.3418  | 1.2994  |
|                   | Skewness | 0.1837  | −0.0072 | −0.1088 | −0.7322 | −0.2315 |
|                   | Kurtosis | 2.7064  | 3.0812  | 2.6807  | 3.1396  | 2.8501  |
| 5                 | Mean     | 4.3350  | 4.7800  | 4.2726  | 4.0649  | 4.4310  |
|                   | SD       | 0.9422  | 1.0866  | 1.1457  | 1.0411  | 1.3348  |
|                   | Skewness | −0.1604 | −0.0499 | −0.1864 | −0.0670 | 0.0175  |
|                   | Kurtosis | 2.4188  | 2.3651  | 3.2872  | 2.6866  | 2.7435  |
| 6                 | Mean     | 4.7281  | 5.5539  | 5.0557  | 5.0869  | 5.1927  |
|                   | SD       | 1.2075  | 1.3638  | 1.1804  | 1.2487  | 1.1796  |
|                   | Skewness | −0.7425 | −0.4345 | −0.6802 | −0.3048 | −0.5480 |
|                   | Kurtosis | 3.4137  | 3.2489  | 3.3222  | 2.5303  | 3.5526  |
| 7                 | Mean     | 5.8443  | 6.3165  | 5.9876  | 5.7993  | 6.1274  |
|                   | SD       | 1.1263  | 1.1641  | 1.1751  | 1.3961  | 1.3014  |
|                   | Skewness | −0.4713 | −0.2273 | −0.0286 | −0.5603 | −0.0880 |
|                   | Kurtosis | 3.3739  | 4.2802  | 3.1137  | 3.1204  | 2.5745  |
| 8                 | Mean     | 6.4020  | 6.9590  | 6.3015  | 6.6168  | 6.7773  |
|                   | SD       | 1.2786  | 1.4373  | 1.3772  | 1.3272  | 1.3911  |
|                   | Skewness | −0.5814 | −0.6753 | −0.3367 | −0.8134 | −0.4038 |
|                   | Kurtosis | 4.2802  | 3.4857  | 2.8101  | 3.7491  | 2.9308  |
| 9                 | Mean     | 7.0656  | 7.7023  | 7.2638  | 7.1552  | 7.5155  |
|                   | SD       | 1.3126  | 1.4790  | 1.3258  | 1.4779  | 1.4271  |
|                   | Skewness | −0.2667 | −0.4473 | −0.2234 | −0.5723 | −0.2416 |
|                   | Kurtosis | 2.3176  | 2.6596  | 2.5958  | 3.0435  | 2.5303  |
| 10                | Mean     | 8.0739  | 8.6711  | 8.1752  | 8.2486  | 8.3881  |
|                   | SD       | 1.3949  | 1.5041  | 1.5992  | 1.4877  | 1.5494  |
|                   | Skewness | −0.2122 | −0.5438 | −0.1644 | −0.4481 | −0.3945 |
|                   | Kurtosis | 2.2111  | 2.9451  | 2.2160  | 3.2584  | 2.7944  |



**Fig. A1.** Feature–channel importance heatmaps for the five emotions: Anger (A), Happiness (B), Sadness (C), Disgust (D), and Fear (E). The rows represent different features (label top-to-down: Delta, Theta, Alpha, SMR, Beta, Gamma, I<sub>1</sub>, I<sub>2</sub>, I<sub>3</sub>, I<sub>4</sub>, I<sub>5</sub>, I<sub>6</sub>, I<sub>7</sub>, I<sub>8</sub>, I<sub>9</sub>, I<sub>10</sub>, I<sub>11</sub>, I<sub>12</sub>, I<sub>13</sub>, I<sub>14</sub>, I<sub>15</sub>, I<sub>16</sub>, I<sub>17</sub>, I<sub>18</sub>, I<sub>19</sub>, I<sub>20</sub>, I<sub>21</sub>, I<sub>22</sub>, I<sub>23</sub>, I<sub>24</sub>, I<sub>25</sub>, I<sub>26</sub>, I<sub>27</sub>, I<sub>28</sub>, I<sub>29</sub>, I<sub>30</sub>, I<sub>31</sub>, I<sub>32</sub>, I<sub>33</sub>, I<sub>34</sub>, I<sub>35</sub>, I<sub>36</sub>, I<sub>37</sub>), while columns represent different EEG channels (labels left-to-right: AF3, AF4, AF7, AF8, C1, C2, C3, C4, C5, C6, CP1, CP2, CP3, CP4, CP5, CP6, CPz, Cz, F1, F2, F3, F4, F5, F6, F7, F8, FC1, FC2, FC3, FC4, FC5, FC6, Fp1, Fp2, FT10, FT7, FT8, FT9, Fz, O1, O2, Oz, P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO7, PO8, Poz, Pz, T7, T8, TP10, TP7, TP8, TP9, AFz, FCz). In each panel, the heatmap displays the importance values associated with each feature–channel pair, computed as the absolute regression coefficients of the LASSO models.



**Fig. A2.** Feature-channel importance heatmaps for the five emotions: Anger (A), Happiness (B), Sadness (C), Disgust (D), and Fear (E). The rows represent different features (label top-to-down: Delta, Theta, Alpha, SMR, Beta, Gamma, I<sub>1</sub>, I<sub>2</sub>, I<sub>3</sub>, I<sub>4</sub>, I<sub>5</sub>, I<sub>6</sub>, I<sub>7</sub>, I<sub>8</sub>, I<sub>9</sub>, I<sub>10</sub>, I<sub>11</sub>, I<sub>12</sub>, I<sub>13</sub>, I<sub>14</sub>, I<sub>15</sub>, I<sub>16</sub>, I<sub>17</sub>, I<sub>18</sub>, I<sub>19</sub>, I<sub>20</sub>, I<sub>21</sub>, I<sub>22</sub>, I<sub>23</sub>, I<sub>24</sub>, I<sub>25</sub>, I<sub>26</sub>, I<sub>27</sub>, I<sub>28</sub>, I<sub>29</sub>, I<sub>30</sub>, I<sub>31</sub>, I<sub>32</sub>, I<sub>33</sub>, I<sub>34</sub>, I<sub>35</sub>, I<sub>36</sub>, I<sub>37</sub>), while columns represent different EEG channels (labels left-to-right: AF3, AF4, AF7, AF8, C1, C2, C3, C4, C5, C6, CP1, CP2, CP3, CP4, CP5, CP6, CPz, Cz, F1, F2, F3, F4, F5, F6, F7, F8, FC1, FC2, FC3, FC4, FC5, FC6, Fp1, Fp2, FT10, FT7, FT8, FT9, Fz, O1, O2, Oz, P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO7, PO8, Poz, Pz, T7, T8, TP10, TP7, TP8, TP9, Afz, FCz). In each panel, the heatmap displays the importance values associated with each feature-channel pair, computed as the product of the absolute mean coefficient and the selection frequency of the LASSO model iterations.

## Data availability

The record is publicly accessible, but files are restricted to users with access to Zenodo <https://zenodo.org/records/8327283>

## References

- [1] D. Wang, Y. Wang, X. Fu, M. Dou, S. Dong, D. Zhang, Revealing the spatial co-occurrence patterns of multi-emotions from social media data, *Telemat. Inform.* 84 (Oct. 2023) 102025, <https://doi.org/10.1016/j.tele.2023.102025>.
- [2] R.Y. Lim, W.-C.-L. Lew, K.K. Ang, Review of EEG affective recognition with a neuroscience perspective, *Brain Sci.* 14 (4) (Apr. 2024) 364, <https://doi.org/10.3390/brainsci14040364>.
- [3] R.A. Calvo, S. D'Mello, Affect detection: an interdisciplinary review of models, methods, and their applications, *IEEE Trans. Affect. Comput.* 1 (1) (Jan. 2010) 18–37, <https://doi.org/10.1109/T-AFFC.2010.1>.
- [4] M.A. Klados, P. Konstantinidi, R. Dacosta-Aguayo, V.-D. Kostaridou, A. Vinciarelli, M. Zervakis, Automatic recognition of personality profiles using EEG functional connectivity during emotional processing, *Brain Sci.* 10 (2020) 5, <https://doi.org/10.3390/brainsci10050278>.
- [5] Y.-P. Lin, et al., EEG-based emotion recognition in music listening, *I.E.E.E. Trans. Biomed. Eng.* 57 (7) (Jul. 2010) 1798–1806, <https://doi.org/10.1109/TBME.2010.2048568>.
- [6] P. Huy-Tung, I. Eun-Tack, O. Myeong-Seok, G. Gwang-Yong, EEG-based emotion recognition: a review and emerging paths, *IEEE Access* 13 (2025) 165037–165060, <https://doi.org/10.1109/ACCESS.2025.3610918>.
- [7] H.A. Hamzah, K.K. Abdalla, EEG-based emotion recognition systems; comprehensive study, *Heliyon* 10 (10) (May 2024) e31485, <https://doi.org/10.1016/j.heliyon.2024.e31485>.
- [8] Y. Geng, S. Shi, X. Hao, Deep learning-based EEG emotion recognition: a comprehensive review, *Neural Comput. Appl.* 37 (4) (Feb. 2025) 1919–1950, <https://doi.org/10.1007/s00521-024-10821-y>.
- [9] R. Vempati, L.D. Sharma, EEG rhythm based emotion recognition using multivariate decomposition and ensemble machine learning classifier, *J. Neurosci. Methods* 393 (Jun. 2023) 109879, <https://doi.org/10.1016/j.jneumeth.2023.109879>.
- [10] J.W. Li, S. Barma, S.H. Pun, M.I. Vai, P.U. Mak, Emotion recognition based on EEG brain rhythm sequencing technique, *IEEE Trans. Cogn. Dev. Syst.* 15 (1) (Mar. 2023) 163–174, <https://doi.org/10.1109/TCDS.2022.3149953>.
- [11] M. Balconi, C. Lucchiari, EEG correlates (event-related desynchronization) of emotional face elaboration: a temporal analysis, *Neurosci. Lett.* 392 (1–2) (Jan. 2006) 118–123, <https://doi.org/10.1016/j.neulet.2005.09.004>.
- [12] S.K. D'mello, J. Kory, A review and meta-analysis of multimodal affect detection systems, *ACM Comput. Surv.* 47 (3) (2015) 1–36, <https://doi.org/10.1145/2682899>.
- [13] R.J. Davidson, Affective neuroscience and psychophysiology: toward a synthesis, *Psychophysiology* 40 (5) (Sep. 2003) 655–665, <https://doi.org/10.1111/1469-8986.00067>.
- [14] S. Koelstra, et al., DEAP: a database for emotion analysis using physiological signals, *IEEE Trans. Affect. Comput.* 3 (1) (Jan. 2012) 18–31, <https://doi.org/10.1109/T-AFFC.2011.15>.
- [15] I. Marcantoni, et al., Ratio indexes based on spectral electroencephalographic brainwaves for assessment of mental involvement: a systematic review, *Sensors* 23 (2023) 13, <https://doi.org/10.3390/s23135968>.
- [16] Y. Liu and O. Sourina, EEG-based dominance level recognition for emotion-enabled interaction, in *2012 IEEE International Conference on Multimedia and Expo*, Melbourne, Australia: IEEE, Jul. 2012, pp. 1039–1044, <https://doi.org/10.1109/ICME.2012.20>.
- [17] J. Fang, G. Yu, S. Liao, S. Zhang, G. Zhu, F. Yi, Using the  $\beta/\alpha$  ratio to enhance odor-induced EEG emotion recognition, *Appl. Sci.* 15 (9) (Apr. 2025) 4980, <https://doi.org/10.3390/app15094980>.
- [18] S. Lee, T. Lee, T. Yang, E. Seomoon, C. Yoon, and S.-P. Kim, “Neural correlates of anxiety induced by environmental events during driving,” in *TENCON 2018 - 2018 IEEE Region 10 Conference*, Jeju, Korea (South): IEEE, Oct. 2018, pp. 0446–0450, <https://doi.org/10.1109/TENCON.2018.8650325>.
- [19] X. Xu, F. Wei, Z. Zhu, J. Liu, X. Wu, Eeg feature selection using orthogonal regression: application to emotion recognition, in *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain: IEEE, May 2020, pp. 1239–1243, <https://doi.org/10.1109/ICASSP40776.2020.9054457>.
- [20] D. Ganepola, M.W.P. Maduranga, V. Tilwari, I. Karunaratne, A systematic review of electroencephalography-based emotion recognition of confusion using artificial intelligence, *Signals* 5 (2) (2024) 244–263, <https://doi.org/10.3390/signals5020013>.

- [21] J. Chen, Y. Cui, C. Wei, K. Polat, F. Alenezi, Advances in EEG-based emotion recognition: challenges, methodologies, and future directions, *Appl. Soft Comput.* 180 (2025), <https://doi.org/10.1016/j.asoc.2025.113478>.
- [22] N. Gudikandula, R. Janapati, R. Sengupta, S. Chintala, Brain computer interface based emotion recognition with error analysis and challenges: an interdisciplinary review, *Discov. Appl. Sci.* 7 (2025) 8, <https://doi.org/10.1007/s42452-025-06692-0>.
- [23] Y. Geng, S. Shi, X. Hao, Deep learning-based EEG emotion recognition: a comprehensive review, *Neural Comput. Appl.* (2024), <https://doi.org/10.1007/s00521-024-10821-y>.
- [24] M.A. Becerra, C. Duque-Mejia, A. Castro-Ospina, L. Serna-Guarín, C. Mejía, E. Duque, EEG-based biometric identification and emotion recognition: an overview, *Computers* 14 (2025) 8, <https://doi.org/10.3390/computers14080299>.
- [25] H.A. Hamzah, K.K. Abdalla, EEG-based emotion recognition datasets for virtual environments: a survey, *Appl. Comput. Intell. Soft Comput.* 2024 (2024), <https://doi.org/10.1155/2024/6091523>.
- [26] M. Jafari, et al., Emotion recognition in EEG signals using deep learning methods: a review, *Comput. Biol. Med.* 165 (2023), <https://doi.org/10.1016/j.combiomed.2023.107450>.
- [27] S.S. Al-Hadithy, A.S. Abdalkafor, B. Al-Khateeb, Emotion recognition in EEG Signals: deep and machine learning approaches, challenges, and future directions, *Comput. Biol. Med.* 196 (2025), <https://doi.org/10.1016/j.combiomed.2025.110713>.
- [28] E. Gkintoni, A. Aroutzidis, H. Antonopoulou, C. Halkiopoulos, From neural networks to emotional networks: a systematic review of EEG-based emotion recognition in cognitive neuroscience and real-world applications, *Brain Sci.* 15 (2025) 3, <https://doi.org/10.3390/brainsci15030220>.
- [29] M.P. Melo, D.F. Adamatti, M.S. Aguiar, Machine learning techniques to classify emotions from electroencephalogram topographic maps: a systematic review, *Comput. Biol. Med.* 197 (2025), <https://doi.org/10.1016/j.combiomed.2025.111022>.
- [30] S. Patil, R. Patil, S. Goudar, S. Sangani, R.H. Goudar, Review on music emotion analysis using machine learning: technologies, methods, datasets, and challenges, *Discov. Appl. Sci.* 7 (2025) 7, <https://doi.org/10.1007/s42452-025-07178-9>.
- [31] A. Abgeena, S. Garg, Unravelling emotions: exploring deep learning approaches for EEG-based emotion recognition with current challenges and future recommendations, *Cogn. Neurodyn.* 19 (2025) 1, <https://doi.org/10.1007/s11571-025-10328-9>.
- [32] H. Eisenbarth, et al., Emotional arousal pattern (EMAP): a new database for modeling momentary subjective and psychophysiological responding to affective stimuli, *Psychophysiology* 61 (2) (Feb. 2024) e14446, <https://doi.org/10.1111/psyp.14446>.
- [33] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *J. Neurosci. Methods* 134 (1) (Mar. 2004) 9–21, <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- [34] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, K.A. Robbins, The PREP pipeline: standardized preprocessing for large-scale EEG analysis, *Front. Neuroinformatics* 9 (Jun. 2015), <https://doi.org/10.3389/fninf.2015.00016>.
- [35] Kaiser, J.F., Nonrecursive digital filter design using the 10-sinh window function, 1974.
- [36] L. Pion-Tonachini, K. Kreutz-Delgado, S. Makeig, ICLABEL: an automated electroencephalographic independent component classifier, dataset, and website, *Neuroimage* 198 (Sep. 2019) 181–197, <https://doi.org/10.1016/j.neuroimage.2019.05.026>.
- [37] P. Welch, The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, *IEEE Trans. Audio Electroacoustics* 15 (2) (Jun. 1967) 70–73, <https://doi.org/10.1109/TAU.1967.1161901>.
- [38] R. Tibshirani, Regression shrinkage and selection via the Lasso, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 58 (1) (1996) 267–288, <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>.
- [39] A. Sbröllini et al., Evaluation of the low-frequency components in surface electromyography, in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA: IEEE, Aug. 2016, pp. 3622–3625. <https://doi.org/10.1109/EMBC.2016.7591512>.
- [40] S.F. Abbasi, H. Jamil, W. Chen, EEG-based neonatal sleep stage classification using ensemble learning, *Comput. Mater. Contin.* 70 (2022) 3, <https://doi.org/10.32604/cmc.2022.020318>.
- [41] S.F. Abbasi, et al., EEG-based neonatal sleep-wake classification using multilayer perceptron neural network, *IEEE Access* 8 (2020) 183025–183034, <https://doi.org/10.1109/ACCESS.2020.3028182>.
- [42] M. Jas, D.A. Engemann, Y. Bekhti, F. Raimondo, A. Gramfort, Autoreject: automated artifact rejection for MEG and EEG data, *Neuroimage* 159 (Oct. 2017) 417–429, <https://doi.org/10.1016/j.neuroimage.2017.06.030>.
- [43] H. Nolan, R. Whelan, R.B. Reilly, FASTER: fully automated statistical thresholding for EEG artifact rejection, *J. Neurosci. Methods* 192 (1) (Sep. 2010) 152–162, <https://doi.org/10.1016/j.jneumeth.2010.07.015>.
- [44] J. Liu, et al., Effects of signal artefacts on electroencephalography spectral power during sleep: quantifying the effectiveness of automated artefact-rejection algorithms, *J. Sleep Res.* 27 (1) (Feb. 2018) 98–102, <https://doi.org/10.1111/jsr.12576>.
- [45] S. Mariani, A. F. T. Borges, T. Henriques, A. L. Goldberger, M. D. Costa, Use of multiscale entropy to facilitate artifact detection in electroencephalographic signals, in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan: IEEE, Aug. 2015, pp. 7869–7872. <https://doi.org/10.1109/EMBC.2015.7320216>.
- [46] S. Vaid, P. Singh, C. Kaur, EEG signal analysis for BCI interface: a review, in 2015 Fifth International Conference on Advanced Computing & Communication Technologies, Haryana, India: IEEE, Feb. 2015, pp. 143–147. <https://doi.org/10.1109/ACCT.2015.72>.