Chronic Stress Assessment Based on EEG Signals: An Empirical Method for EEG Feature Importance

Rattaphong Laoharungpisit†, Akraradet Sinsamersuk†, Chaklam Silpasuwanchai, Matthew Dailey and Attaphongse Taparugssanagorn

Center for Health and Wellness Technologies, School of Engineering and Technology, Asian Institute of Technology.

Contributing authors: laoharungpisit.r@gmail.com; akraradets@gmail.com; chaklam@ait.ac.th; mdailey@ait.ac.th; attaphongset@ait.ac.th;
†These authors contributed equally to this work.

Abstract

Chronic stress is a prolonged and constant feeling of stress that could lead to depression and anxiety. It would be useful to have a self-diagnosing tool to help notify individuals before chronic stress becomes unmanageable. EEG analysis has proven effective as a means of human emotion prediction, but studies linking EEG analysis with and chronic stress are limited. To fill this gap, we recorded rest-state EEG and perceived stress scale (PSS) scores for 55 participants. Five frequency bands (delta, theta, alpha, beta, and gamma) and lateral asymmetry between electrode pairs calculated for the alpha and beta bands are extracted as features from the rest-state EEG. A feature selection procedure was performed using a t-tests and feature importance scores based on five classifiers. A RBF SVM achieved a 10-cv score over 0.9 using eight different combinations of frequency bands, electrodes, and asymmetries, namely, $\boldsymbol{\beta_f}$, F3$_{\boldsymbol{\delta}}$, F4$_{\boldsymbol{\delta}}$, F3$_{\boldsymbol{\beta}}$, P4$_{\boldsymbol{\delta}}$, F3$_{\boldsymbol{\gamma}}$, P4$_{\boldsymbol{\theta}}$, and C3$_{\boldsymbol{\theta}}$. This paper shows that data-driven feature selection for EEG signals can yield accurate chronic stress classifiers.

Keywords: Long-term stress, Chronic stress, Electroencephalography

1 Introduction

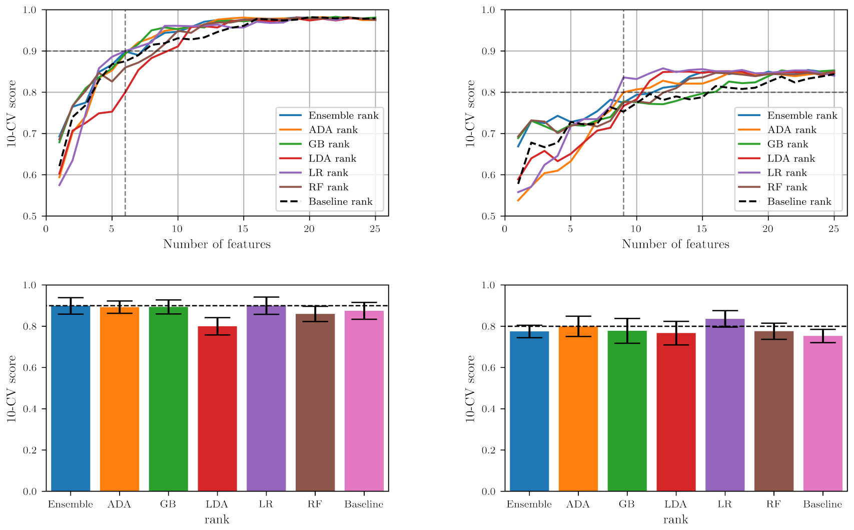
Stress is an unavoidable life phenomenon that can generate temporary discomfort as well as long-term consequences. Stress can have damaging effects on the mental health of individuals and can be categorized into acute (short-term) and chronic (long-term) occurrence [1], where acute stress is usually not a health risk, while the persistence of stress for a longer duration becomes chronic [1], which can lead to a state of depression, anxiety, and other possibly life-threatening issues. Chronic stress also affects the human body at different levels ranging from skin conditions, eating habits, inadequate sleeping to decision-making [2–4] and can be a better predictor of depressive symptoms as compared to acute stress [5]. Early detection of chronic stress can therefore reduce the risk of physical and mental illness if the stress is identified around the individual undertake appropriate stress relief therapies.

Several studies that have developed techniques to assess chronic stress using subjective psychological self-report questionnaires such as the Perceived Stress Scale (PSS) [6]. This approach is practical for capturing signs of stress over a long period, but the result could be biased since participants must reconstruct their feelings in the past, and they may hide their real feelings due to personal circumstances. As a result, genuine stress may not come to light. Recently, objective physiological assessments based on bio-markers such as electroencephalography (EEG) signals have emerged as alternative means to detect mental states including acute stress [7, 8]. However, the relationship between chronic stress and EEG signals has not been widely investigated.

The aim of this study is to identify which EEG features are important contributors to predictive classification accuracy. We performed a study with a total of 55 participants, whose rest-state EEG measurement were recorded for ten minutes with eyes closed. The label of each participant was calculated from the statistics of their PSS score. An initial 133 features including five EEG frequency bands (δ, θ, α, β, γ), two custom bands (slow, β_{low}), one engineered band (γ_{relative}) from 16 electrodes and five asymmetries of Alpha and Beta band are extracted from the recorded EEG. We selected the 25 most important features using t-test. We ranked the 25 features with six methods. The first five are based on five classifiers' feature importance scores. The five classifiers are Logistic Regression (LR), Linear Discriminant Analysis (LDA), AdaBoost (AB), Gradient Boosting (GB), and Random Forest (RF). The sixth ranking was created by combining the first five ranked using ranked voting scheme. Finally, 10-CV score is used to verify the rank and t-Distributed Stochastic Neighbor Embedding (t-SNE) is used for visualizing the similarity of Stress and Non-Stress groups given a set of features.

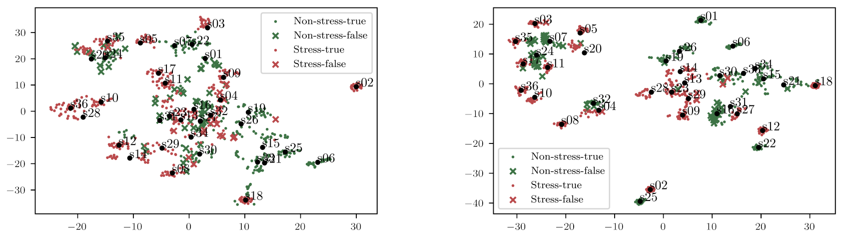
We found that the SVM with Radial Basis Function (RBF) kernel achieves close or up to 0.9 10-CV score when using top six features from Rank_{LR} and $\text{Rank}_{\text{Ensemble}}$ as shown in Figure 1a. Top eight features in $\text{Rank}_{\text{Ensemble}}$ are β_f, $F3_\delta$, $F4_\delta$, $F3_\beta$, $P4_\delta$, $F3_\gamma$, $P4_\theta$, and $C3_\theta$. In addition, t-SNE seems to suggest that top six $\text{Rank}_{\text{Ensemble}}$ are a better representation of a chronic stress when comparing to the $\text{Rank}_{\text{Baseline}}$ as shown in Figure 2.

Fig. 1: 10-CV score of both SVM and LR classifier when using different rank at different number of features. A line plot show the trend of 10-CV when increasing the number of feature based on ranks. The bar plot shows the 10-CV score at the number of feature, 6 for SVM and 9 for LR.



(a) SVM (b) LR

Fig. 2: The top six features from Rank$_{\text{Baseline}}$ and Rank$_{\text{Ensemble}}$ plot in t-SNE space. Color green and red indicate the group Non-stress and Stress respectively and shape dot and x indicates the classification result with RBF-SVM.



(a) Baseline (b) Ensemble

This paper shows the empirical method of feature selection by combining t-test result with coefficient/feature importance score of five different classifiers. Our result confirms that EEG is a viable data for classifying chronic stress and the top eight features agree with the previous work that Delta and Beta are the most important frequency for this task.

2 Related work

Stress is known to cause aberrant reactions in the autonomic nervous system (ANS), which is composed of the antagonistic regulation of the sympathetic (SNS) as well as the parasympathetic (PNS) nervous system [9, 10]. These two systems are associated with stress and relaxation responses, respectively. Changes in the ANS can be detected by EEG signals [11]. In recent years, EEG signals have become an alternative technique to analyze mental states including stress [7, 8]. EEG signals can be decomposed into frequency band[12], where each of the frequency bands can be utilized to differentiate between different brain states [13]. A change of power in β band is associated with alertness, whereas θ band occurred throughout the sleep state, and α band rose with relaxation [14]. Recently, some studies used EEG signals conducted with various signals processing and machine learning methods to detect stress.

Several studies have demonstrated the relationship between stress and EEG signals and show that EEG can be used as an input variable in both short-term and long-term stress analysis. In EEG, the α band is commonly used in stress assessment because it is increasing during the relaxation situation. On the other hand, it is decreasing during stress condition [15]. In contrast, α is not the only band that can indicate mental stress. The β band has been found as a sign of mental stress due to an increase of power in β band associated with an increase in alertness when performing cognitive tasks. The relationship between stress and β band at the anterior temporal lobe has demonstrated when stress-inducing pictures were shown [7, 11, 16]. Meanwhile, theta band was used to assess stress during mental arithmetic tasks [17]. Moreover, the relative gamma band has been proposed as a biomarker for the identification of stress [18].

Previous studies have demonstrated the feasibility of using EEG for chronic stress assessment. Chronic stress is measured without involving stress-induced tasks. Since chronic stress is a greater indicator of depressive symptoms than acute stress [5] as well as the potential of EEG to be employed as a stress indicator in everyday life [19]. Thus, we propose the use of EEG signals to identify chronic stress without inducing stress.

3 Methodology

3.1 Participants

A total of 55 healthy participants (30 male, 25 female) with ages ranging from 21 to 45 years (Mean $= 26, SD = 4.84$) participated in the study. The participants were informed to avoid food and drink with caffeine as well as conditioners, hair creams, sprays, or styling gels on the day of the experiment. Prior to the experiment, all participants gave consent in writing before completing the questionnaire and having their EEG recorded.

3.2 Data Acquisition

The EEG acquisition phase of the experiment was conducted in a room with a controlled environment. An EEG Electrode Cap Kit was used to record EEG signals of the participants at 16 active electrodes, namely FP1, FP2, F3, F4, F7, F8, C3, C4, T3, T4, T5, T6, P3, P4, O1, and O2, with two reference electrodes according to the electrode locations of International 10 - 20 system for EEG (Figure 4) with a sampling rate of 125 Hz. The experimental procedure is shown in Figure 3. All participants were initially instructed regarding the experimental procedure, signed consent form, then took the PSS questionnaire. Each participant was placed comfortably in a chair and asked to close their eyes for 10 minutes while keeping their head motionless to eliminate movement artifacts. The closed-eyes condition was utilized to simplify data processing. Previous research has found that chronic stress can be classified without stress induction [1, 19, 20].

In this study, the PSS-10 questionnaire was used to assessing participants' stress levels and classify them. The questionnaire has 10 questions, each of which asks about the frequency of stress in the last month. Each question is answered on a scale of 0, which indicates that the incident never happened, to 4, which indicates that it occurs frequently. All participants were categorized into three groups based on their PSS scores, as either stressed, non-stressed, or neutral. PSS thresholds were chosen as

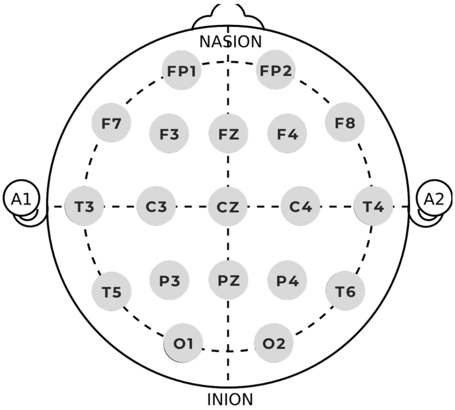
$$T_u = \mu + \frac{\sigma}{2} \tag{1}$$

$$T_l = \mu - \frac{\sigma}{2}, \tag{2}$$

where μ is the mean of PSS scores and σ is the standard deviation of the PSS scores overs all 55 participants, T_u is an upper threshold and T_l is a lower threshold. Any participants with PSS score lower than T_l are classified as non-stressed, and those with PSS scores higher than T_u are classified as stressed. Others are classified as neutral.

Fig. 3: Experimental procedure and data acquisition process

Fig. 4: The electrode locations of International 10-20 system for EEG recording from [21]



3.3 Feature Extraction

The first five minutes of the rest-state EEG signals were segmented into 20 smaller records (15 seconds each). For each record, the five EEG frequency bands were obtained using the `psd_welch` function from in the `mne` library. The function calculates power spectral density (PSD) of multiple smaller windows and average them to get a smoother result. The frequency bands, namely, δ (1-3 Hz), θ (4-7 Hz), α (8-12 Hz), β (13-30 Hz), γ (25-43 Hz), slow (4-13 Hz), and β_{low} (13-17 Hz) were computed from each electrode. We did not perform any other preprocessing. The ratio of gamma to slow bands was used to calculate a γ_{relative}. Moreover, there were five Alpha and Beta asymmetries determined. The following formulae were used to compute the alpha asymmetries:

$$\alpha_f = \frac{\mathrm{F4}_\alpha - \mathrm{F3}_\alpha}{\mathrm{F4}_\alpha + \mathrm{F3}_\alpha} \tag{3}$$

$$\alpha_t = \frac{\mathrm{T4}_\alpha - \mathrm{T3}_\alpha}{\mathrm{T4}_\alpha + \mathrm{T3}_\alpha} \tag{4}$$

$$\alpha_a = \alpha_f + \alpha_t \tag{5}$$

α_f, α_t and α_a indicate frontal, temporal, and α asymmetries, respectively, and channel$_\alpha$ denotes the PSD of α at each electrode. β asymmetry was computed over the frontal and temporal region using the same procedure:

$$\beta_f = \frac{\mathrm{F4}_\beta - \mathrm{F3}_\beta}{\mathrm{F4}_\beta + \mathrm{F3}_\beta} \tag{6}$$

$$\beta_t = \frac{\mathrm{T4}_\beta - \mathrm{T3}_\beta}{\mathrm{T4}_\beta + \mathrm{T3}_\beta} \tag{7}$$

β_f and β_t indicate frontal and temporal asymmetries of the β band, and channel$_\beta$ denotes the PSD of β at each electrode.

In total, the number of extracted features is $(16 \times 8) + 5 = 133$.

3.4 Feature Selection

We segmented data into 15ssecond segments (20 samples from each participant) followed by the feature extraction process. Then, we normalized each of the 133 features across the entire dataset using z-scaling:

$$\mathbf{X}_i' = \frac{\mathbf{X}_i - \mu_i}{\sigma_i}, \tag{8}$$

where μ_i is the mean and σ_i is the standard deviation of the data for the feature i.

Next, we eliminated features unlikely to be useful for classification by compare the means of each features over each class separately using t-test for the null hypothesis of equal means. From the original 133 features, we found that 25 of t-test gave $p < .001$, so we used these 25 features for further consideration.

We ranked the 25 features with the following 7 methods:

1. Rank$_{\text{Baseline}}$: ¡Will investigate t-value¿.
2. Rank$_{\text{LR}}$: ranked according to the logistic regression coefficient.
3. Rank$_{\text{LDA}}$: ranked according to the linear discriminant analysis coefficient.
4. Rank$_{\text{AB}}$: ranked according to the feature importance score of AdaBoost.
5. Rank$_{\text{GB}}$: ranked according to the feature importance score of gradient boosting.
6. Rank$_{\text{RF}}$: ranked according to the feature importance score of random forest.
7. Rank$_{\text{Ensemble}}$: ranked according to a score based on the previous five rankings. The score for feature i is

$$\text{score}_i = \sum_{m \in M} \text{Rank}_m[i], \tag{9}$$

where $M = \{\text{LR}, \text{LDA}, \text{AB}, \text{GB}, \text{RF}\}$ and i is the index of the feature of interest. Rank$_{\text{Ensemble}}$ is obtained by sorting the scores in ascending order. Thus, the feature with the lowest score is the most important one.

3.5 Classification

We utilize two classifiers commonly used in previous studies [13, 19, 22, 23]: logistic regression (LR) and support vector machine (SVM).

3.5.1 Logistic Regression (LR)

In its simplest form, logistic regression is a generalized linear model for the Bernoulli (binary class) distribution that employs a logistic sigmoid in a linear combination of the features to represent a binary dependent variable i.e., p(y|x).

3.5.2 Support Vector Machine (SVM)

A support vector machine performs linear binary maximum margin classification in a feature space. SVMs have been shown to be very effective in EEG-based stress categorization studies [13, 22]. A linear SVM would give a similar result to LR, but with a nonlinear transformation into feature space, can give higher accuracy.

3.6 Performance Metrics

A 10-CV score is used to measure the classification performance. t-SNE is used to visualize the similarity between samples.

4 Results

4.1 Groups Labeling

A total of 55 participants can be divided into 3 groups consisting of Stress, Non-Stress, and Neutral groups based on the PSS questionnaire score. From the Equation 1 and Equation 2, we obtain the μ, σ = (20.75±6.13). The score lower than 17.68 is considered to be Non-Stress group, whereas the score higher than 23.81 is considered to be Stress group. The participants with a score in between the first standard deviation are considered to be in Neutral group and are discarded in the classification. As a result, the Stress group comprises 20 participants, while Non-Stress group comprises 16 participants. In total, the number of subjects in classification is $(20 + 16) = 36$ and $(36 \times 20) = 720$ samples after segmentation.

4.2 t-test result

The independent sample t-test with $p < .001$ was used to select feature in this study. For the t-test, the degree of freedom was 1 and the null hypothesis was tested for EEG feature of Stress and Non-Stress groups. The result using the p-values for different features are shown in Table 1 and Table 2.

There are 25 features with $p < .001$ which are $\mathrm{FP1}_\delta$, $\mathrm{F3}_\delta$, $\mathrm{F4}_\delta$, $\mathrm{C3}_\delta$, $\mathrm{P4}_\delta$, $\mathrm{F3}_\theta$, $\mathrm{C3}_\theta$, $\mathrm{T4}_\theta$, $\mathrm{P4}_\theta$, $\mathrm{F3}_\alpha$, $\mathrm{F4}_\alpha$, $\mathrm{F7}_\alpha$, $\mathrm{T6}_\alpha$, $\mathrm{P4}_\alpha$, $\mathrm{F3}_\beta$, $\mathrm{F3}_\gamma$, $\mathrm{T3}_\gamma$, $\mathrm{F3}_{\mathrm{slow}}$, $\mathrm{F3}_{\mathrm{Low}\beta}$, $\mathrm{T4}_{\mathrm{Low}\beta}$, $\mathrm{C3}_{\mathrm{RG}}$, α_f, α_t, α_a,β_f. We use this as a list for $\mathrm{Rank}_{\mathrm{Baseline}}$.

4.3 Ranking with coefficient and feature importance score

We rank the 25 features from the t-test result using coefficient/feature importance score of a classifier as mentioned in the Section 3.4. To verify any Rank_m where $m \in \{\mathrm{LR}, \mathrm{LDA}, \mathrm{ADA}, \mathrm{GB}, \mathrm{RF}\}$, we (1) iterate over the index of Rank_m; (2) select the features from the Rank_m^1 to Rank_m^i where i is the current iteration; (3) calculate the 10-CV score using rbf-SVM and LR; and (4) plot the t-SNE space to visualize the similarity of samples given the features from the iteration that SVM can achieve over 0.9 10-CV score.

Table 1: Results for the t-test on various frequency bands.

	Frequency bands							
Channel	**Delta (δ)**	**Theta (θ)**	**Alpha (α)**	**Beta (β)**	**Gamma (γ)**	**Slow**	**Low Beta**	**RG**
FP1	**<.001**	0.744	**0.002**	0.899	**0.024**	0.203	0.629	0.221
FP2	**0.016**	**0.010**	**0.041**	**0.008**	0.956	0.682	0.071	0.168
F3	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	0.864
F4	**<.001**	**0.001**	**<.001**	0.488	0.276	**0.020**	0.479	**0.011**
F7	**0.006**	0.550	**<.001**	**0.015**	0.616	**0.026**	0.474	**0.011**
F8	0.115	0.211	0.088	**0.045**	0.251	0.789	0.117	0.677
C3	**<.001**	**<.001**	0.297	0.069	**0.007**	**0.026**	**0.003**	**<.001**
C4	0.551	0.329	0.316	0.895	**0.016**	0.297	0.186	0.146
T3	0.090	0.719	**0.001**	0.650	**<.001**	0.401	0.608	0.078
T4	0.101	**<.001**	0.855	**0.029**	**0.003**	**0.003**	**<.001**	**0.001**
T5	0.095	0.126	0.213	0.217	0.305	0.482	0.687	0.199
T6	0.919	0.840	**<.001**	0.306	**0.003**	**0.010**	**0.014**	0.105
P3	0.639	0.507	0.405	0.114	0.068	0.424	0.077	0.358
P4	**<.001**	**<.001**	**<.001**	0.573	**0.012**	0.630	0.379	0.165
O1	0.057	0.169	0.080	0.271	0.128	0.163	0.778	0.469
O2	**0.001**	0.384	0.388	**0.009**	0.544	0.729	0.252	**0.042**

Table 2: Results for the t-test on asymmetries

Alpha asymmetry (α_a)	**Alpha frontal (α_f)**	**Alpha temporal (α_t)**	**Beta frontal (β_f)**	**Beta temporal (β_f)**
<.001	**<.001**	**<.001**	**<.001**	**0.002**

The 10-CV score threshold of 0.9 for SVM and 0.8 for LR is selected based on the 10-CV score of each model when training the model using 25 features. The SVM exceeds 0.9 10-CV score while LR can reach 0.85 at best. Table that reports every ranks and 10-CV score of all classifiers can be found in Appendix A.

4.3.1 Baseline

When using $\text{Rank}_{\text{Baseline}}$, rbf-SVM can achieve 0.9 10-CV score with top eight features including FP1_δ, F3_δ, F4_δ, C3_δ, P4_δ, F3_θ, C3_θ, and T4_θ as shown in Table 3. Using the same rank with LR, however, the model requires top 16 features in order to reach 0.8 10-CV score.

Fig. 5: The data plot in t-SNE space using top eight features from $\text{Rank}_{\text{Baseline}}$. Although the rbf-SVM achieves 0.915 10-CV score, the data are not separated. This affect the LR performace as shown in Table 3

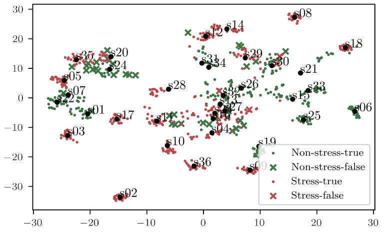


Table 3: The 10-CV score of each classifier when using $\text{Rank}_{\text{Baseline}}$ as a rank. The model is trained on a list of features from Rank^1 to Rank^i where i is the row number.

No.	$\text{Rank}_{\text{Baseline}}$	SVM	LR
1	FP1_δ	0.621±0.046	0.578±0.054
2	F3_δ	0.740±0.061	0.678±0.055
3	F4_δ	0.769±0.043	0.667±0.062
4	C3_δ	0.829±0.044	0.678±0.077
5	P4_δ	0.868±0.051	0.728±0.064
6	F3_θ	0.875±0.041	0.722±0.049
7	C3_θ	0.890±0.020	0.724±0.043
8	T4_θ	**0.915±0.031**	0.764±0.033
9	P4_θ	0.919±0.012	0.753±0.032
10	F3_α	0.931±0.020	0.774±0.041
11	F4_α	0.928±0.024	0.796±0.027
12	F7_α	0.933±0.028	0.781±0.048
13	T6_α	0.946±0.031	0.790±0.036
14	P4_α	0.956±0.016	0.783±0.045
15	F3_β	0.961±0.023	0.788±0.041
16	F3_γ	0.978±0.015	**0.815±0.037**
17	T3_γ	0.976±0.023	0.811±0.044
18	F3_{slow}	0.974±0.015	0.808±0.040
19	$\text{F3}_{\text{Low}\beta}$	0.976±0.011	0.811±0.040
20	$\text{T4}_{\text{Low}\beta}$	0.982±0.014	0.825±0.058
21	C3_{RG}	0.981±0.013	0.838±0.022
22	α_f	0.979±0.007	0.824±0.041
23	α_t	0.981±0.009	0.833±0.037
24	α_a	0.978±0.013	0.838±0.031
25	β_f	0.979±0.014	0.844±0.040

4.3.2 Logistic Regression Coefficient

Table 4 shows the result using the ranked features of Logistic Regression Coefficient. With this rank, the SVM is able to achieve 0.9 10-CV score at top six features. The LR classifier also reach 0.8 10-CV score at top nine features, seven features less than the $\text{Rank}_{\text{Baseline}}$. The top six features of Rank_{LR} are F3_β, P4_δ, β_f, C3_{RG}, F3_δ, F4_δ, and T4_θ.

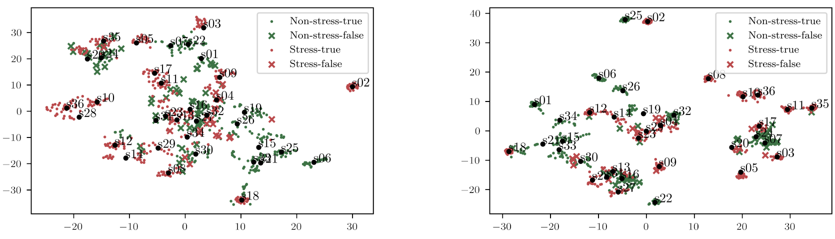
4.3.3 Ensemble

The score for $\text{Rank}_{\text{Ensemble}}$ is calculated follow the Equation 9. Table 5 shows the calculated score and rank them accordingly. In this configuration, SVM, at top six features, achieves 0.899±0.040 10-CV score and, later, achieves over 0.9 10-CV score at iteration eight. The top eight features are β_f, F3_δ, F4_δ, F3_β, P4_δ, F3_γ, P4_θ, and C3_θ.

Table 4: The 10-CV score of each classifier when using Rank_{LR} as a rank. The model is trained on a list of features from Rank^1 to Rank^i where i is the row number.

No.	Rank_{LR}	coeff	SVM	LR
1	F3_β	1.894	0.575±0.051	0.558±0.040
2	P4_δ	1.651	0.635±0.050	0.571±0.048
3	β_f	-1.435	0.749±0.058	0.624±0.051
4	C3_{RG}	-1.346	0.858±0.034	0.646±0.032
5	F3_δ	-1.273	0.886±0.030	0.718±0.047
6	F4_δ	-1.258	**0.900±0.042**	0.735±0.061
7	T4_θ	0.952	0.910±0.029	0.735±0.053
8	$\text{F3}_{\text{Low}\beta}$	-0.841	0.921±0.037	0.765±0.054
9	P4_α	-0.833	0.961±0.014	**0.836±0.040**
10	F7_α	0.767	0.961±0.022	0.832±0.040
11	F4_α	-0.609	0.960±0.018	0.846±0.048
12	F3_α	-0.594	0.961±0.016	0.858±0.045
13	FP1_δ	-0.503	0.964±0.015	0.849±0.052
14	T6_α	-0.465	0.957±0.015	0.854±0.019
15	F3_γ	0.422	0.957±0.031	0.856±0.032
16	P4_θ	0.405	0.971±0.018	0.851±0.049
17	$\text{T4}_{\text{Low}\beta}$	0.396	0.968±0.019	0.850±0.041
18	F3_θ	-0.384	0.969±0.016	0.854±0.037
19	C3_δ	0.350	0.982±0.013	0.846±0.036
20	C3_θ	-0.348	0.979±0.016	0.847±0.020
21	T3_γ	-0.323	0.979±0.011	0.850±0.034
22	α_t	-0.201	0.975±0.019	0.853±0.034
23	α_a	-0.182	0.981±0.023	0.853±0.035
24	α_f	-0.067	0.979±0.017	0.847±0.052
25	F3_{slow}	0.036	0.975±0.015	0.840±0.042

Fig. 6: The top six features from $\text{Rank}_{\text{Baseline}}$ and Rank_{LR} plot in t-SNE space. The two groups are better separated in the case of Rank_{LR}.



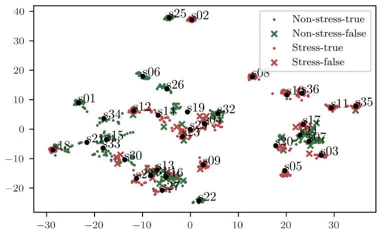
(a) Baseline

(b) LR

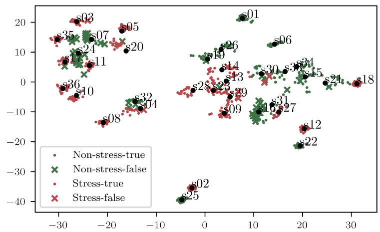
Table 5: The 10-CV score of each classifier when using $\text{Rank}_{\text{ensemble}}$ as a rank. The model is trained on a list of features from Rank^1 to Rank^i where i is the row number.

No.	$\text{Rank}_{\text{ensemble}}$	score	SVM	LR
1	β_f	15	0.693±0.057	0.669±0.046
2	$F3_\delta$	18	0.765±0.029	0.731±0.048
3	$F4_\delta$	34	0.775±0.057	0.725±0.035
4	$F3_\beta$	34	0.849±0.040	0.743±0.047
5	$P4_\delta$	45	0.867±0.044	0.728±0.070
6	$F3_\gamma$	47	0.899±0.040	0.735±0.033
7	$P4_\theta$	48	0.890±0.040	0.754±0.041
8	$C3_\theta$	49	**0.925±0.023**	0.782±0.058
9	$C3_{\text{RG}}$	52	0.944±0.030	0.775±0.030
10	$T4_\theta$	54	0.947±0.017	0.794±0.050
11	$T4_{\text{Low}\beta}$	54	0.958±0.023	0.797±0.058
12	$P4_\alpha$	57	0.971±0.020	**0.811±0.060**
13	$F7_\alpha$	57	0.975±0.025	0.815±0.050
14	$F3_\alpha$	59	0.972±0.025	0.838±0.040
15	$F3_\theta$	61	0.976±0.021	0.849±0.034
16	$C3_\delta$	61	0.978±0.023	0.849±0.058
17	$F4_\alpha$	67	0.978±0.015	0.844±0.056
18	$F3_{\text{Low}\beta}$	69	0.979±0.018	0.849±0.036
19	$T3_\gamma$	71	0.981±0.013	0.840±0.035
20	$T6_\alpha$	73	0.981±0.015	0.850±0.053
21	$FP1_\delta$	79	0.979±0.009	0.844±0.037
22	$F3_{\text{slow}}$	88	0.981±0.013	0.844±0.040
23	α_f	94	0.981±0.011	0.854±0.036
24	α_a	103	0.979±0.013	0.850±0.038
25	α_t	111	0.978±0.015	0.844±0.020

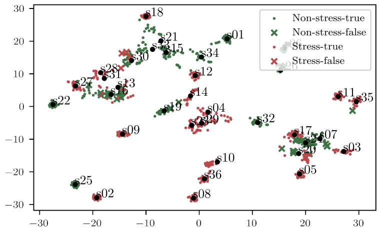
Fig. 7: The top six and top eight features from $\mathrm{Rank}_{\mathrm{LR}}$ and $\mathrm{Rank}_{\mathrm{Ensemble}}$ plot in t-SNE space.



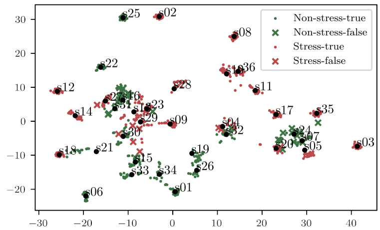
(a) top six LR



(b) top six Ensemble



(c) top eight LR



(d) top eight Ensemble

5 Discussion

At the recording state, we chose to do eye-closed conditions because we want to avoid external factors such as lighting conditions and visual distraction, as well as eye artifacts. However, there are challenges to this approach. First, during closing the eye, the power of the Alpha band in the occipital area is greatly increased. We could argue that this does not affect our result but due to the nature of EEG, the Alpha information in the occipital area could contaminate their neighbor. Maybe this is the reason why any of the Alpha asymmetry features are staying at the bottom of the rank. Second, the participant might fall asleep during the record. Because the record session is five minutes long and the participant is asked to avoid any form of caffeine which many of them have on a daily basis, closing the eyes for five minutes is enough to turn some participants into a relaxing or sleeping state.

All 55 participants can be considered well educated since all of them are, at least, pursuing a degree of bachelor. The nationality range is limited to Asian and the age is between 20 and 40 years old. Our data is somewhat biased by this fact.

The label of the data is based on the questionnaire PSS-10 score, a method subject to error and manipulation. The analysis could be improved if a psychologist expert is involved in the labeling process. Because of our labeling method, our data has a minor imbalance issue which we chose to ignore (16 Non-stress and 20 Stress participants). Our suggestion would be to increase the lower threshold to accommodate more Non-stress samples or we could segment the two groups differently.

In our experiment, RBF-SVM shows to be the best classifier for chronic stress. However, because of the nature of the RBF kernel, we can not rank the features by importance. To be specific, given enough data dimension, RBF-SVM always achieves a 0.9 10-CV score. While we could brute force through every combination of features given enough time, we can not be sure that the model learns chronic stress patterns or participant-specific patterns. Because of this reason, t-SNE is utilized.

We, first, investigated the Logistic Regression Coefficient and found that, in t-SNE space, the two groups are well separated as shown in Figure 6. In addition, the performance of other models is also improved given a smaller set of features (Table A2). Therefore, we extend our idea to other models and, finally, convolute all ranks into the $\text{Rank}_{\text{Ensemble}}$. Every classifier seems to agree that the Beta and Delta band are the most important frequency. The electrode F3 is the most important followed by F4 and P4. Our result agrees with the previous studies [19]

Furthermore, t-SNE plots seem to suggest that our data has an outlier, for instance, s07 and s24 consistently appear in the middle of the stress group. To confirm this assumption, a larger and broader group of participants is needed.

6 Conclusion

In conclusion, the rest-state EEG can be used to classify chronic stress. The classification result shows that SVM with rbf kernel achieves over 0.98 10-CV scores and LR achieves 0.85 using all features with $p < .001$ from the reported t-test result. Extended ranking with various classifiers improve the feature selection and help to narrow down the feature list. Finally, convoluted $\mathrm{Rank}_{\mathrm{Ensemble}}$ shows the top eight Chronic Stress features which are β_f, $\mathrm{F3}_\delta$, $\mathrm{F4}_\delta$, $\mathrm{F3}_\beta$, $\mathrm{P4}_\delta$, $\mathrm{F3}_\gamma$, $\mathrm{P4}_\theta$, and $\mathrm{C3}_\theta$. With these eight features, rbf-SVM, ADA, GB, and RF achieve over 0.9 10-CV score while adding the next four features ($\mathrm{C3}_{\mathrm{RG}}$, $\mathrm{T4}_\theta$, $\mathrm{T4}_{\mathrm{Low}\beta}$, and $\mathrm{P4}_\alpha$), helps LR and LDA to achieve over 0.8 10-CV score (Table A7).

Appendix A 10-CV result of each Ranks

Fig. A1: A bump chart of features in each ranking.

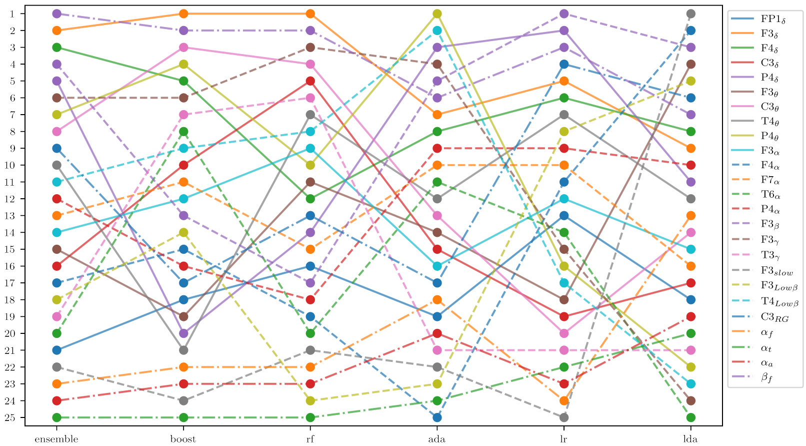


Table A1: The 10-CV score of each classifier when using $\text{Rank}_{\text{baseline}}$ as a rank.

No.	$\text{Rank}_{\text{baseline}}$	SVM	LR	GB	ADA	RF	LDA
1	$FP1_\delta$	0.621±0.046	0.578±0.054	0.606±0.030	0.604±0.036	0.525±0.045	0.585±0.044
2	$F3_\delta$	0.740±0.061	0.678±0.055	0.733±0.047	0.710±0.049	0.724±0.061	0.676±0.045
3	$F4_\delta$	0.769±0.043	0.667±0.062	0.765±0.047	0.714±0.037	0.761±0.028	0.664±0.039
4	$C3_\delta$	0.829±0.044	0.678±0.077	0.832±0.038	0.761±0.036	0.843±0.028	0.675±0.043
5	$P4_\delta$	0.868±0.051	0.728±0.064	0.868±0.030	0.807±0.043	0.885±0.036	0.733±0.036
6	$F3_\theta$	0.875±0.041	0.722±0.049	0.853±0.050	0.821±0.074	0.878±0.058	0.724±0.055
7	$C3_\theta$	0.890±0.020	0.724±0.043	0.875±0.032	0.833±0.046	0.897±0.040	0.722±0.046
8	$T4_\theta$	**0.915±0.031**	0.764±0.033	0.883±0.052	0.843±0.030	**0.907±0.030**	0.751±0.054
9	$P4_\theta$	0.919±0.012	0.753±0.032	**0.901±0.021**	0.840±0.027	0.911±0.045	0.765±0.052
10	$F3_\alpha$	0.931±0.020	0.774±0.041	0.918±0.027	0.878±0.039	0.926±0.021	0.778±0.050
11	$F4_\alpha$	0.928±0.024	0.796±0.027	0.931±0.026	0.876±0.030	0.936±0.026	**0.800±0.035**
12	$F7_\alpha$	0.933±0.028	0.781±0.048	0.928±0.028	0.867±0.034	0.939±0.028	0.793±0.032
13	$T6_\alpha$	0.946±0.031	0.790±0.036	0.940±0.019	0.869±0.038	0.936±0.036	0.789±0.042
14	$P4_\alpha$	0.956±0.016	0.783±0.045	0.933±0.027	0.899±0.037	0.947±0.024	0.788±0.030
15	$F3_\beta$	0.961±0.023	0.788±0.041	0.940±0.028	0.882±0.041	0.957±0.020	0.769±0.033
16	$F3_\gamma$	0.978±0.015	**0.815±0.037**	0.951±0.022	**0.924±0.029**	0.969±0.022	0.814±0.032
17	$T3_\gamma$	0.976±0.023	0.811±0.044	0.956±0.030	0.912±0.027	0.975±0.015	0.814±0.044
18	$F3_{\text{slow}}$	0.974±0.015	0.808±0.040	0.961±0.025	0.914±0.031	0.968±0.024	0.817±0.047
19	$F3_{\text{Low}\beta}$	0.976±0.011	0.811±0.040	0.960±0.020	0.911±0.024	0.971±0.025	0.808±0.042
20	$T4_{\text{Low}\beta}$	0.982±0.014	0.825±0.058	0.958±0.012	0.918±0.027	0.974±0.013	0.812±0.051
21	$C3_{\text{RG}}$	0.981±0.013	0.838±0.022	0.958±0.020	0.904±0.021	0.971±0.018	0.814±0.043
22	α_f	0.979±0.007	0.824±0.041	0.962±0.015	0.912±0.036	0.975±0.020	0.819±0.041
23	α_t	0.981±0.009	0.833±0.037	0.964±0.015	0.924±0.032	0.969±0.021	0.821±0.042
24	α_a	0.978±0.013	0.838±0.031	0.968±0.022	0.914±0.037	0.979±0.016	0.826±0.031
25	β_f	0.979±0.014	0.844±0.040	0.961±0.022	0.917±0.037	0.972±0.014	0.831±0.033

Table A2: The 10-CV score of each classifier when using Rank$_{\mathrm{LR}}$ as a rank.

No.	Rank$_{\mathrm{LR}}$	coeff	SVM	LR	GB	ADA	RF	LDA
1	F3$_{\beta}$	1.894	0.575±0.051	0.558±0.040	0.561±0.072	0.607±0.075	0.564±0.057	0.565±0.046
2	P4$_{\delta}$	1.651	0.635±0.050	0.571±0.048	0.656±0.033	0.610±0.054	0.693±0.036	0.571±0.047
3	β_f	-1.435	0.749±0.058	0.624±0.051	0.753±0.047	0.742±0.039	0.814±0.055	0.624±0.054
4	C3$_{\mathrm{RG}}$	-1.346	0.858±0.034	0.646±0.032	0.856±0.032	0.781±0.034	0.885±0.035	0.647±0.044
5	F3$_{\delta}$	-1.273	0.886±0.030	0.718±0.047	0.875±0.052	0.826±0.044	0.897±0.019	0.715±0.024
6	F4$_{\delta}$	-1.258	**0.900±0.042**	0.735±0.061	0.899±0.028	0.850±0.035	**0.912±0.028**	0.749±0.015
7	T4$_{\theta}$	0.952	0.910±0.029	0.735±0.053	**0.900±0.038**	0.865±0.036	0.922±0.044	0.754±0.041
8	F3$_{\mathrm{Low}\beta}$	-0.841	0.921±0.037	0.765±0.054	0.896±0.052	0.868±0.037	0.931±0.039	0.761±0.044
9	P4$_{\alpha}$	-0.833	0.961±0.014	**0.836±0.040**	0.931±0.026	0.888±0.033	0.940±0.018	**0.836±0.046**
10	F7$_{\alpha}$	0.767	0.961±0.022	0.832±0.040	0.932±0.035	0.893±0.036	0.960±0.028	0.835±0.067
11	F4$_{\alpha}$	-0.609	0.960±0.018	0.846±0.048	0.947±0.024	0.893±0.031	0.946±0.021	0.840±0.049
12	F3$_{\alpha}$	-0.594	0.961±0.016	0.858±0.045	0.942±0.024	**0.900±0.025**	0.949±0.027	0.846±0.040
13	FP1$_{\delta}$	-0.503	0.964±0.015	0.849±0.052	0.938±0.021	0.908±0.031	0.953±0.021	0.839±0.037
14	T6$_{\alpha}$	-0.465	0.957±0.015	0.854±0.019	0.939±0.031	0.917±0.032	0.951±0.024	0.836±0.043
15	F3$_{\gamma}$	0.422	0.957±0.031	0.856±0.032	0.956±0.024	0.912±0.035	0.964±0.020	0.842±0.045
16	P4$_{\theta}$	0.405	0.971±0.018	0.851±0.049	0.951±0.027	0.935±0.026	0.967±0.017	0.846±0.024
17	T4$_{\mathrm{Low}\beta}$	0.396	0.968±0.019	0.850±0.041	0.961±0.010	0.922±0.036	0.968±0.028	0.839±0.048
18	F3$_{\theta}$	-0.384	0.969±0.016	0.854±0.037	0.962±0.018	0.935±0.021	0.968±0.014	0.838±0.064
19	C3$_{\delta}$	0.350	0.982±0.013	0.846±0.036	0.962±0.021	0.938±0.027	0.972±0.021	0.843±0.033
20	C3$_{\theta}$	-0.348	0.979±0.016	0.847±0.020	0.951±0.029	0.936±0.017	0.967±0.021	0.831±0.034
21	T3$_{\gamma}$	-0.323	0.979±0.011	0.850±0.034	0.957±0.025	0.939±0.029	0.974±0.020	0.835±0.023
22	α_t	-0.201	0.975±0.019	0.853±0.034	0.962±0.033	0.925±0.033	0.969±0.018	0.829±0.038
23	α_a	-0.182	0.981±0.023	0.853±0.035	0.954±0.014	0.917±0.039	0.976±0.025	0.829±0.036
24	α_f	-0.067	0.979±0.017	0.847±0.052	0.964±0.029	0.919±0.031	0.976±0.012	0.826±0.040
25	F3$_{\mathrm{slow}}$	0.036	0.975±0.015	0.840±0.042	0.956±0.018	0.925±0.021	0.974±0.016	0.833±0.041

Table A3: The 10-CV score of each classifier when using Rank$_{\mathrm{LDA}}$ as a rank.

No.	Rank$_{\mathrm{LDA}}$	score	SVM	LR	GB	ADA	RF	LDA
1	F3$_{\mathrm{slow}}$	4.658	0.603±0.044	0.589±0.046	0.593±0.052	0.592±0.059	0.610±0.055	0.583±0.043
2	F4$_{\alpha}$	-3.383	0.707±0.040	0.640±0.055	0.703±0.063	0.682±0.050	0.706±0.053	0.643±0.058
3	F3$_{\beta}$	3.258	0.726±0.038	0.658±0.049	0.758±0.051	0.718±0.046	0.783±0.038	0.650±0.064
4	F3$_{\theta}$	-2.143	0.749±0.049	0.633±0.039	0.800±0.024	0.729±0.053	0.818±0.043	0.629±0.041
5	F3$_{\mathrm{Low}\beta}$	-1.710	0.753±0.046	0.651±0.044	0.788±0.032	0.775±0.041	0.804±0.049	0.649±0.066
6	C3$_{\mathrm{RG}}$	-1.595	0.800±0.042	0.679±0.049	0.808±0.044	0.738±0.045	0.847±0.041	0.674±0.044
7	β_f	-1.398	0.854±0.049	0.707±0.041	0.882±0.027	0.824±0.025	0.886±0.038	0.725±0.043
8	F4$_{\delta}$	-1.289	0.883±0.048	0.714±0.053	0.886±0.044	0.856±0.037	**0.917±0.026**	0.725±0.026
9	F3$_{\delta}$	-1.227	0.897±0.031	0.767±0.057	0.893±0.030	0.868±0.023	0.912±0.041	0.782±0.026
10	P4$_{\alpha}$	-1.155	**0.911±0.022**	0.783±0.042	**0.910±0.030**	0.867±0.050	0.918±0.040	0.785±0.058
11	P4$_{\delta}$	1.143	0.958±0.022	**0.828±0.039**	0.925±0.030	0.892±0.036	0.944±0.022	**0.832±0.034**
12	T4$_{\theta}$	1.069	0.960±0.027	0.849±0.037	0.947±0.022	0.899±0.035	0.954±0.022	0.844±0.044
13	α_f	-1.042	0.957±0.027	0.850±0.029	0.942±0.019	**0.906±0.037**	0.953±0.026	0.844±0.046
14	C3$_{\theta}$	-0.639	0.972±0.020	0.850±0.048	0.943±0.027	0.915±0.024	0.967±0.021	0.838±0.044
15	F3$_{\alpha}$	-0.458	0.975±0.021	0.847±0.033	0.944±0.029	0.914±0.044	0.960±0.018	0.839±0.034
16	F7$_{\alpha}$	0.395	0.976±0.014	0.851±0.042	0.946±0.031	0.914±0.030	0.961±0.023	0.835±0.032
17	C3$_{\delta}$	0.364	0.972±0.022	0.851±0.045	0.946±0.025	0.919±0.028	0.968±0.020	0.833±0.024
18	FP1$_{\delta}$	-0.339	0.975±0.016	0.850±0.038	0.950±0.039	0.890±0.031	0.960±0.023	0.832±0.038
19	α_a	-0.279	0.979±0.019	0.839±0.044	0.946±0.023	0.915±0.031	0.961±0.025	0.835±0.036
20	α_t	0.271	0.974±0.025	0.847±0.025	0.953±0.029	0.912±0.028	0.972±0.015	0.836±0.043
21	T3$_{\gamma}$	0.180	0.978±0.019	0.849±0.034	0.953±0.020	0.904±0.029	0.974±0.021	0.835±0.040
22	P4$_{\theta}$	0.077	0.974±0.015	0.849±0.034	0.947±0.039	0.910±0.044	0.968±0.018	0.838±0.044
23	T4$_{\mathrm{Low}\beta}$	0.066	0.979±0.017	0.849±0.024	0.953±0.026	0.911±0.032	0.978±0.015	0.825±0.053
24	F3$_{\gamma}$	-0.042	0.978±0.015	0.842±0.042	0.964±0.029	0.925±0.038	0.971±0.034	0.832±0.033
25	T6$_{\alpha}$	0.022	0.976±0.013	0.847±0.037	0.962±0.019	0.932±0.027	0.975±0.023	0.835±0.051

Table A4: The 10-CV score of each classifier when using $\mathrm{Rank_{ADA}}$ as a rank.

No.	$\mathrm{Rank_{ADA}}$	score	SVM	LR	GB	ADA	RF	LDA
1	$\mathrm{P4}_\theta$	0.10	0.594±0.070	0.538±0.026	0.571±0.040	0.604±0.044	0.517±0.071	0.539±0.032
2	$\mathrm{T4_{Low\beta}}$	0.08	0.703±0.041	0.572±0.045	0.656±0.065	0.633±0.029	0.636±0.054	0.569±0.031
3	$\mathrm{P4}_\delta$	0.08	0.742±0.042	0.604±0.035	0.731±0.032	0.682±0.052	0.750±0.046	0.599±0.052
4	$\mathrm{F3}_\gamma$	0.08	0.835±0.026	0.610±0.051	0.879±0.028	0.774±0.054	0.893±0.027	0.612±0.047
5	$\mathrm{F3}_\beta$	0.08	0.853±0.026	0.633±0.055	0.899±0.038	0.794±0.037	**0.915±0.038**	0.621±0.041
6	β_f	0.08	0.893±0.030	0.678±0.050	**0.924±0.033**	0.846±0.055	0.940±0.029	0.676±0.048
7	$\mathrm{F3}_\delta$	0.06	**0.921±0.033**	0.733±0.044	0.942±0.040	**0.906±0.029**	0.958±0.024	0.739±0.054
8	$\mathrm{F4}_\delta$	0.06	0.932±0.037	0.756±0.035	0.947±0.019	0.901±0.027	0.961±0.014	0.769±0.052
9	$\mathrm{P4}_\alpha$	0.06	0.949±0.028	**0.800±0.049**	0.942±0.038	0.917±0.024	0.964±0.021	**0.804±0.054**
10	$\mathrm{F7}_\alpha$	0.06	0.954±0.021	0.807±0.048	0.950±0.025	0.932±0.023	0.964±0.028	0.803±0.054
11	$\mathrm{T6}_\alpha$	0.04	0.958±0.022	0.811±0.038	0.962±0.022	0.921±0.028	0.968±0.016	0.806±0.039
12	$\mathrm{T4}_\theta$	0.04	0.964±0.023	0.828±0.049	0.965±0.025	0.928±0.030	0.975±0.017	0.810±0.038
13	$\mathrm{C3}_\theta$	0.04	0.976±0.022	0.821±0.031	0.957±0.025	0.926±0.028	0.969±0.014	0.819±0.043
14	$\mathrm{F3}_\theta$	0.04	0.979±0.023	0.821±0.058	0.954±0.028	0.933±0.038	0.972±0.026	0.817±0.047
15	$\mathrm{C3}_\delta$	0.02	0.981±0.017	0.821±0.048	0.964±0.018	0.944±0.022	0.969±0.016	0.825±0.049
16	$\mathrm{F3}_\alpha$	0.02	0.979±0.025	0.832±0.045	0.956±0.008	0.924±0.035	0.975±0.023	0.833±0.037
17	$\mathrm{C3_{RG}}$	0.02	0.978±0.023	0.846±0.040	0.967±0.019	0.935±0.015	0.975±0.018	0.828±0.030
18	α_f	0.02	0.979±0.014	0.847±0.049	0.957±0.020	0.936±0.021	0.976±0.022	0.825±0.029
19	$\mathrm{FP1}_\delta$	0.02	0.979±0.022	0.843±0.044	0.960±0.013	0.929±0.030	0.972±0.023	0.826±0.034
20	α_a	0.00	0.981±0.020	0.846±0.051	0.964±0.021	0.929±0.020	0.975±0.016	0.817±0.043
21	$\mathrm{T3}_\gamma$	0.00	0.982±0.006	0.843±0.034	0.963±0.022	0.928±0.040	0.978±0.017	0.826±0.046
22	$\mathrm{F3_{slow}}$	0.00	0.981±0.017	0.839±0.057	0.967±0.023	0.924±0.027	0.981±0.021	0.829±0.054
23	$\mathrm{F3_{Low\beta}}$	0.00	0.982±0.013	0.843±0.036	0.968±0.015	0.914±0.033	0.979±0.014	0.832±0.037
24	α_t	0.00	0.975±0.021	0.847±0.023	0.958±0.032	0.918±0.018	0.983±0.015	0.832±0.015
25	$\mathrm{F4}_\alpha$	0.00	0.975±0.016	0.847±0.022	0.956±0.026	0.906±0.043	0.972±0.020	0.831±0.037

Table A5: The 10-CV score of each classifier when using $\mathrm{Rank_{GB}}$ as a rank.

No.	$\mathrm{Rank_{GB}}$	score	SVM	LR	GB	ADA	RF	LDA
1	$\mathrm{F3}_\delta$	0.140	0.679±0.049	0.689±0.041	0.668±0.051	0.688±0.047	0.582±0.053	0.688±0.024
2	β_f	0.099	0.767±0.024	0.731±0.050	0.761±0.040	0.749±0.042	0.732±0.049	0.725±0.054
3	$\mathrm{C3}_\theta$	0.082	0.810±0.039	0.718±0.031	0.782±0.050	0.806±0.033	0.803±0.032	0.714±0.042
4	$\mathrm{P4}_\theta$	0.076	0.836±0.025	0.704±0.040	0.822±0.024	0.831±0.037	0.846±0.036	0.707±0.047
5	$\mathrm{F4}_\delta$	0.073	0.858±0.015	0.721±0.054	0.851±0.034	0.836±0.031	0.868±0.025	0.721±0.035
6	$\mathrm{F3}_\gamma$	0.071	0.894±0.034	0.719±0.049	**0.919±0.037**	0.894±0.029	**0.936±0.021**	0.722±0.046
7	$\mathrm{T3}_\gamma$	0.069	**0.917±0.030**	0.732±0.049	0.932±0.042	0.892±0.037	0.953±0.022	0.735±0.039
8	$\mathrm{T6}_\alpha$	0.061	0.950±0.026	0.740±0.063	0.933±0.031	**0.910±0.029**	0.956±0.028	0.740±0.072
9	$\mathrm{T4_{Low\beta}}$	0.050	0.957±0.021	0.778±0.060	0.953±0.027	0.922±0.028	0.956±0.023	0.776±0.035
10	$\mathrm{C3}_\delta$	0.040	0.953±0.015	0.776±0.053	0.947±0.036	0.897±0.020	0.954±0.022	0.776±0.040
11	$\mathrm{F7}_\alpha$	0.036	0.960±0.018	0.772±0.059	0.956±0.023	0.919±0.025	0.971±0.020	0.781±0.034
12	$\mathrm{F3}_\alpha$	0.034	0.957±0.024	0.771±0.063	0.947±0.037	0.915±0.029	0.968±0.015	0.788±0.051
13	$\mathrm{F3}_\beta$	0.026	0.972±0.019	0.779±0.036	0.949±0.029	0.931±0.019	0.976±0.021	**0.801±0.056**
14	$\mathrm{F3_{Low\beta}}$	0.021	0.974±0.021	0.789±0.043	0.953±0.011	0.921±0.027	0.974±0.015	0.817±0.031
15	$\mathrm{F4}_\alpha$	0.016	0.972±0.024	0.797±0.049	0.956±0.022	0.922±0.032	0.972±0.021	0.822±0.025
16	$\mathrm{P4}_\alpha$	0.015	0.975±0.014	**0.801±0.030**	0.961±0.017	0.938±0.034	0.974±0.024	0.804±0.048
17	$\mathrm{C3_{RG}}$	0.014	0.976±0.016	0.826±0.036	0.957±0.024	0.925±0.029	0.971±0.013	0.817±0.028
18	$\mathrm{FP1}_\delta$	0.014	0.972±0.014	0.822±0.051	0.957±0.030	0.929±0.031	0.967±0.014	0.808±0.043
19	$\mathrm{F3}_\theta$	0.013	0.978±0.013	0.824±0.047	0.957±0.021	0.922±0.029	0.972±0.024	0.812±0.043
20	$\mathrm{P4}_\delta$	0.012	0.981±0.013	0.839±0.029	0.960±0.020	0.921±0.032	0.975±0.015	0.826±0.044
21	$\mathrm{T4}_\theta$	0.011	0.979±0.019	0.853±0.040	0.961±0.025	0.932±0.046	0.982±0.015	0.832±0.057
22	α_f	0.010	0.983±0.018	0.849±0.042	0.958±0.016	0.926±0.015	0.975±0.015	0.828±0.037
23	α_a	0.007	0.979±0.017	0.847±0.043	0.954±0.023	0.921±0.033	0.976±0.018	0.835±0.038
24	$\mathrm{F3_{slow}}$	0.006	0.979±0.017	0.851±0.038	0.967±0.017	0.922±0.033	0.978±0.019	0.833±0.061
25	α_t	0.002	0.981±0.013	0.853±0.037	0.964±0.017	0.910±0.026	0.986±0.012	0.828±0.043

Table A6: The 10-CV score of each classifier when using Rank_{RF} as a rank.

No.	Rank_{RF}	score	SVM	LR	GB	ADA	RF	LDA
1	F3_δ	0.103	0.686±0.068	0.693±0.033	0.669±0.048	0.694±0.026	0.568±0.070	0.690±0.039
2	β_f	0.076	0.767±0.046	0.732±0.039	0.751±0.037	0.747±0.056	0.733±0.037	0.725±0.044
3	F3_γ	0.066	0.804±0.033	0.729±0.057	0.868±0.037	0.821±0.039	0.883±0.024	0.725±0.051
4	C3_θ	0.052	0.847±0.038	0.701±0.067	0.879±0.043	0.864±0.050	0.885±0.029	0.711±0.058
5	C3_δ	0.048	0.826±0.038	0.719±0.048	0.897±0.021	0.853±0.032	**0.907±0.027**	0.721±0.050
6	T3_γ	0.045	0.860±0.037	0.724±0.043	**0.904±0.038**	0.865±0.037	0.925±0.024	0.722±0.045
7	T4_θ	0.045	0.872±0.025	0.717±0.031	0.914±0.034	0.868±0.039	0.950±0.023	0.726±0.040
8	$\text{T4}_{\text{Low}\beta}$	0.045	0.890±0.043	0.731±0.057	0.919±0.025	0.894±0.031	0.957±0.017	0.742±0.041
9	F3_α	0.041	**0.917±0.030**	0.776±0.039	0.925±0.023	0.896±0.025	0.951±0.042	0.785±0.034
10	P4_θ	0.038	0.949±0.025	0.778±0.025	0.936±0.023	**0.901±0.031**	0.968±0.015	0.788±0.071
11	F3_θ	0.038	0.944±0.019	0.774±0.033	0.935±0.030	0.901±0.032	0.975±0.016	0.786±0.038
12	F4_δ	0.036	0.965±0.023	0.799±0.037	0.951±0.026	0.915±0.039	0.971±0.021	**0.819±0.040**
13	C3_{RG}	0.036	0.969±0.026	**0.810±0.038**	0.954±0.021	0.921±0.032	0.958±0.014	0.806±0.040
14	P4_δ	0.036	0.969±0.018	0.833±0.059	0.960±0.019	0.919±0.022	0.971±0.024	0.822±0.026
15	F7_α	0.035	0.975±0.018	0.836±0.037	0.957±0.024	0.907±0.022	0.974±0.023	0.828±0.028
16	FP1_δ	0.033	0.978±0.015	0.847±0.038	0.953±0.023	0.912±0.036	0.974±0.013	0.825±0.025
17	F3_β	0.032	0.978±0.011	0.846±0.042	0.958±0.021	0.924±0.027	0.969±0.020	0.828±0.026
18	P4_α	0.032	0.976±0.015	0.843±0.033	0.967±0.019	0.925±0.017	0.972±0.022	0.829±0.032
19	F4_α	0.029	0.981±0.020	0.839±0.035	0.956±0.016	0.915±0.042	0.972±0.015	0.829±0.035
20	T6_α	0.029	0.979±0.019	0.844±0.028	0.957±0.022	0.935±0.016	0.976±0.026	0.831±0.040
21	F3_{slow}	0.027	0.979±0.014	0.844±0.039	0.969±0.025	0.933±0.023	0.975±0.016	0.825±0.040
22	α_f	0.026	0.979±0.024	0.843±0.022	0.958±0.012	0.931±0.024	0.982±0.013	0.825±0.029
23	α_a	0.023	0.981±0.015	0.847±0.036	0.962±0.023	0.911±0.029	0.975±0.016	0.822±0.034
24	$\text{F3}_{\text{Low}\beta}$	0.019	0.979±0.017	0.844±0.045	0.958±0.028	0.917±0.022	0.969±0.017	0.833±0.060
25	α_t	0.012	0.975±0.014	0.851±0.037	0.957±0.037	0.907±0.044	0.978±0.009	0.831±0.040

Table A7: The 10-CV score of each classifier when using $\text{Rank}_{\text{Ensemble}}$ as a rank.

No.	$\text{Rank}_{\text{Ensemble}}$	score	SVM	LR	GB	ADA	RF	LDA
1	β_f	15	0.693±0.057	0.669±0.046	0.646±0.038	0.671±0.035	0.585±0.042	0.667±0.042
2	F3_δ	18	0.765±0.029	0.731±0.048	0.742±0.050	0.736±0.043	0.733±0.036	0.728±0.038
3	F4_δ	34	0.775±0.057	0.725±0.035	0.761±0.051	0.774±0.032	0.769±0.038	0.726±0.045
4	F3_β	34	0.849±0.040	0.743±0.047	0.840±0.028	0.817±0.035	0.860±0.037	0.742±0.064
5	P4_δ	45	0.867±0.044	0.728±0.070	0.860±0.058	0.839±0.032	0.879±0.028	0.728±0.067
6	F3_γ	47	0.899±0.040	0.735±0.033	**0.928±0.025**	0.892±0.034	**0.939±0.037**	0.749±0.052
7	P4_θ	48	0.890±0.040	0.754±0.041	0.940±0.028	**0.908±0.036**	0.947±0.024	0.765±0.036
8	C3_θ	49	**0.925±0.023**	0.782±0.058	0.942±0.029	0.910±0.028	0.956±0.025	0.782±0.035
9	C3_{RG}	52	0.944±0.030	0.775±0.030	0.946±0.024	0.900±0.031	0.953±0.017	0.785±0.037
10	T4_θ	54	0.947±0.017	0.794±0.050	0.960±0.021	0.917±0.036	0.960±0.024	0.796±0.035
11	$\text{T4}_{\text{Low}\beta}$	54	0.958±0.023	0.797±0.058	0.957±0.024	0.914±0.040	0.969±0.010	**0.801±0.052**
12	P4_α	57	0.971±0.020	**0.811±0.060**	0.954±0.019	0.947±0.018	0.975±0.022	0.818±0.038
13	F7_α	57	0.975±0.025	0.815±0.050	0.956±0.020	0.937±0.023	0.971±0.017	0.824±0.042
14	F3_α	59	0.972±0.025	0.838±0.040	0.956±0.025	0.928±0.017	0.974±0.025	0.833±0.026
15	F3_θ	61	0.976±0.021	0.849±0.034	0.958±0.016	0.926±0.030	0.969±0.016	0.828±0.042
16	C3_δ	61	0.978±0.023	0.849±0.058	0.957±0.025	0.932±0.016	0.975±0.019	0.829±0.037
17	F4_α	67	0.978±0.015	0.844±0.056	0.961±0.016	0.925±0.023	0.968±0.015	0.829±0.032
18	$\text{F3}_{\text{Low}\beta}$	69	0.979±0.018	0.849±0.036	0.965±0.024	0.926±0.028	0.974±0.025	0.839±0.043
19	T3_γ	71	0.981±0.013	0.840±0.035	0.961±0.017	0.925±0.028	0.972±0.025	0.836±0.058
20	T6_α	73	0.981±0.015	0.850±0.053	0.961±0.010	0.933±0.036	0.975±0.018	0.836±0.041
21	FP1_δ	79	0.979±0.009	0.844±0.037	0.958±0.027	0.921±0.026	0.974±0.018	0.828±0.049
22	F3_{slow}	88	0.981±0.013	0.844±0.040	0.953±0.039	0.917±0.039	0.974±0.018	0.829±0.036
23	α_f	94	0.981±0.011	0.854±0.036	0.957±0.017	0.925±0.035	0.969±0.017	0.831±0.027
24	α_a	103	0.979±0.013	0.850±0.038	0.960±0.027	0.922±0.038	0.979±0.013	0.831±0.037
25	α_t	111	0.978±0.015	0.844±0.020	0.968±0.019	0.929±0.030	0.978±0.018	0.832±0.039

References

[1] Khosrowabadi, R., Quek, C., Ang, K., Tung, S., Heijnen, M.: A brain-computer interface for classifying eeg correlates of chronic mental stress. Proceedings of the International Joint Conference on Neural Networks, 757–762 (2011)

[2] Thoits, P.A.: Stress and health: Major findings and policy implications. Journal of Health and Social Behavior **51**, 41–53 (2010)

[3] Garg, A., Chren, M.-M., Sands, L.P., Matsui, M.S., Marenus, K.D., Feingold, K.R., Elias, P.M.: Psychological stress perturbs epidermal permeability barrier homeostasis: implications for the pathogenesis of stress-associated skin disorders. Archives of dermatology **137 1**, 53–9 (2001)

[4] Adam, T.C., Epel, E.S.: Stress, eating and the reward system. Physiology & Behavior **91**, 449–458 (2007)

[5] McGonagle, K.A., Kessler, R.C.: Chronic stress, acute stress, and depressive symptoms. American Journal of Community Psychology **18**, 681–706 (1990)

[6] Cohen, S., Kamarck, T., Mermelstein, R.: A global measure of perceived stress. Journal of health and social behavior **24 4**, 385–96 (1983)

[7] Awang, S.A., Paulraj, M.P., Yaacob, S., Yusnita, M.A.: Analysis of eeg signals during relaxation and mental stress condition using ar modeling techniques. 2011 IEEE International Conference on Control System, Computing and Engineering, 477–481 (2011)

[8] Hu, B., Peng, H., Zhao, Q., Hu, B., Majoe, D., Zheng, F., Moore, P.: Signal quality assessment model for wearable eeg sensor on prediction of mental stress. IEEE Transactions on NanoBioscience **14**, 553–561 (2015)

[9] Cohen, H., Benjamin, J., Geva, A.B., Matar, M.A., Kaplan, Z., Kotler, M.: Autonomic dysregulation in panic disorder and in post-traumatic stress disorder: application of power spectrum analysis of heart rate variability at rest and in response to recollection of trauma or panic attacks. Psychiatry Research **96**, 1–13 (2000)

[10] Hughes, J.W., Stoney, C.M.: Depressed mood is related to high-frequency heart rate variability during stressors. Psychosomatic Medicine **62**, 796–803 (2000)

[11] Seo, S.H., Lee, J.-T.: Stress and EEG, (2010)

[12] Kulkarni, N., Phalle, S., Desale, M., Gokhale, N., Kasture, K.: A review on eeg based stress monitoring system using deep learning approach. Mukt Shabd J **9**, 1317–1325 (2020)

[13] Al-Shargie, F., Tang, T.B., Badruddin, N., Kiguchi, M.: Towards multilevel mental stress assessment using svm with ecoc: an eeg approach. Medical & biological engineering & computing **56**(1), 125–136 (2018)

[14] Wang, X.-W., Nie, D., Lu, B.-L.: Emotional state classification from eeg data using machine learning approach. Neurocomputing **129**, 94–106 (2014)

[15] Al-shargie, F., Kiguchi, M., Badruddin, N., Dass, S.C., Hani, A.F.M., Tang, T.B.: Mental stress assessment using simultaneous measurement of eeg and fnirs. Biomedical optics express **7 10**, 3882–3898 (2016)

[16] Hamid, N.H., Sulaiman, N., Mohd Aris, S.A., Murat, Z., Taib, M.N.: Evaluation of human stress using eeg power spectrum. 6th International Colloquium on Signal Processing and Its Applications (CSPA), 1–4 (2010)

[17] Gärtner, M., Grimm, S., Bajbouj, M.: Frontal midline theta oscillations during mental arithmetic: effects of stress. Frontiers in Behavioral Neuroscience **9** (2015)

[18] Arsalan, A., Majid, M., Butt, A.R., Anwar, S.M.: Classification of perceived mental stress using a commercially available eeg headband. IEEE Journal of Biomedical and Health Informatics **23**, 2257–2264 (2019)

[19] Saeed, S.M.U., Anwar, S.M., Khalid, H., Majid, M., Bagci, U.: Eeg based classification of long-term stress using psychological labeling. Sensors (Basel, Switzerland) **20** (2020)

[20] Peng, H., Hu, B., Zheng, F., Fan, D., Zhao, W., Chen, X., Yang, Y., Cai, Q.: A method of identifying chronic stress by eeg. Personal and Ubiquitous Computing **17**, 1341–1347 (2012)

[21] Wikipedia contributors: 10–20 system (EEG). [Online; accessed 10-May-2022] (2022). https://en.wikipedia.org/wiki/10%E2%80%9320_system_(EEG)

[22] Subhani, A.R., Mumtaz, W., Saad, M.N.B.M., Kamel, N.S., Malik, A.S.: Machine learning framework for the detection of mental stress at multiple levels. IEEE Access **5**, 13545–13556 (2017)

[23] Kotsiantis, S.B.: Supervised machine learning: A review of classification techniques. Informatica (Slovenia) **31**, 249–268 (2007)