Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 56-2 (2023) 11627–11632

# EEG-based Affect Classification with Machine Learning Algorithms

**Jianhua Zhang*, Zhong Yin**, and Peng Chen*****

*AI Lab, Department of Computer Science, Oslo Metropolitan University, 0166 Oslo, Norway (Tel: +47-67 23 66 91; e-mail: jianhuaz@oslomet.no)

**School of Optical-Electrical and Computer Eng., University of Shanghai for Science and Technology, Shanghai 200093, P.R. China (e-mail: yinzhong@usst.edu.cn)

***School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, P.R. China (e-mail: chenpeng0538@qq.com)

**Abstract:** In this paper, we aim to study the EEG-based emotion recognition problem. First, we use clustering algorithm to determine the target class of emotions and perform binary classification of emotion along its arousal and valence dimension. Then we compare two different feature extraction methods, i.e., wavelet transform (resulting in wavelet-based features) and nonlinear dynamics analysis (leading to features of approximate entropy and sample entropy). Five feature reduction algorithms are compared in terms of emotion classification accuracy. Furthermore, four types of machine learning classifiers, including k-nearest neighbor (KNN), naive bayes (NB), support vector machine (SVM) and random forest (RF), are also compared. The results on the DEAP physiological data show that the combination of kernel spectral regression (KSR) and random forest leads to the best binary classification of emotions and that the EEG gamma rhythm is closely correlated to variations in emotions.



*Keywords:* Emotion recognition; Affective computing; Electroencephalogram (EEG); Nonlinear dynamics; Wavelets; Kernel Spectral Regression (KSR); Machine learning.

## 1. INTRODUCTION

Affective computing has a wide range of applications. For instance, in human-computer interaction (HCI), if the computer can rapidly and accurately estimate the user's emotional state, the interaction would become more user-friendly and smarter. The application to enhancement of user experience of a product allows manufacturer to monitor in real time the emotional state of its user. In aerospace and defense applications, the risky mental/psychological state of astronauts and soldiers may be detected in real time. In the applications to driving safety, the driver's emotional state can be monitored in real time in order to prevent potential dangers or accidents due to extreme emotional state of the driver during driving.

Emotion recognition is an essential component of affective computing. Human emotions can be identified through the use of facial expressions (video or image), speech (audio), behavior, or physiological signals (Petrushin, 1999; Anderson and McOwan, 2006; Pantic and Rothkrantz, 2000; Zhong et al., 2017; Zhang et al., 2020 ). However, the first three methods may fail when subjects deliberately conceal their true emotions. In contrast, the physiological signals are more reliable and objective (Wang, Nie and Lu, 2014). EEG signals respond to emotion changes more rapidly than other types of peripheral neural signals. It was shown that EEG signals are rich in features of emotional states (Li et al., 2009; Petrantonakis and Hadjileontiadis, 2011). In recent years, there is an increasing need for intelligent HCI. The current studies on emotion recognition focus on:

(i) correlation between physiological signals and emotions;
(ii) different stimulation materials used to evoke various emotion responses;

(iii) different feature extraction algorithms;

(iv) models of emotion; and

(v) emotion recognition by data fusion.

Due to its non-intrusiveness, real-time sensitivity and certain extent of objectivity and robustness, EEG signals are used to recognize human emotional states with several machine learning algorithms in this paper.

## 2. DATASET, AFFECT RECOGNITION FRAMEWORK, AND EEG DATA PREPROCESSING

### 2.1 Dataset and affect recognition framework

In this paper, the DEAP database (Koelstra et al., 2012) is used. Koelstra et al. (2012) selected 40 music videos as emotion stimulation materials based on a 2D emotion model. While 32 subjects (half male, half female, 19-37 years old with an average age of 26.9) watched the 40 music videos, their physiological signals and facial expressions were recorded simultaneously. In the database, for each subject there are 32-channel EEG signals and 8-channel peripheral physiological signals (including galvanic skin response, respiration amplitude, skin temperature, electrocardiogram (ECG), blood volume, electromyography (EMG), and


Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2023.10.486

electrooculography (EOG)). The block diagram of EEG-based emotion recognition proposed in this paper is depicted in Fig. 1.

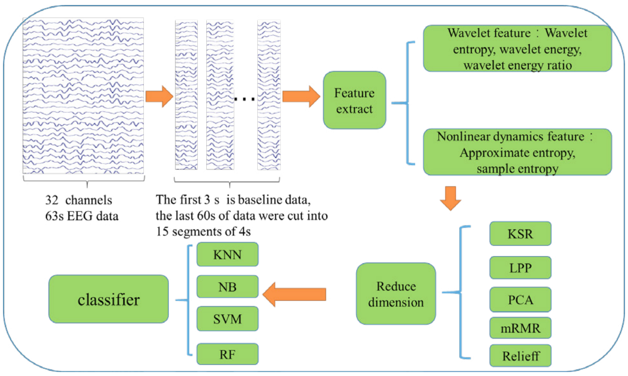


Fig. 1. Block diagram of EEG-based emotion recognition system.

## 2.2 EEG data preprocessing

EEG signals, at a frequency of 512 Hz, were down sampled to 128 Hz during preprocessing. EOG artifacts were removed with a 4.0-45.0 Hz bandpass filter. The preprocessed EEG data contains 60s data while watching the video and 3s baseline data before watching the video. When building the DEAP database , the subject was asked to take a 2-min. break after watching every two videos. In addition, there are marked individual differences in physiological signals across subjects, and even for the same participant and the same stimulation material there would be different emotion evoking patterns at different times and in different environments. In order to minimize the effect of the previous stimulate on the current emotional state and the effect of the cross-subject variability of EEG signals, the EEG features after emotion stimulation were subtracted from those before it. Finally, the (relative) difference features are normalized to the unit interval of [0, 1]. In EEG feature extraction, we obtain the difference features by subtracting the 3s baseline features (before watching music video) from the 60s emotion-related EEG features (when watching it).

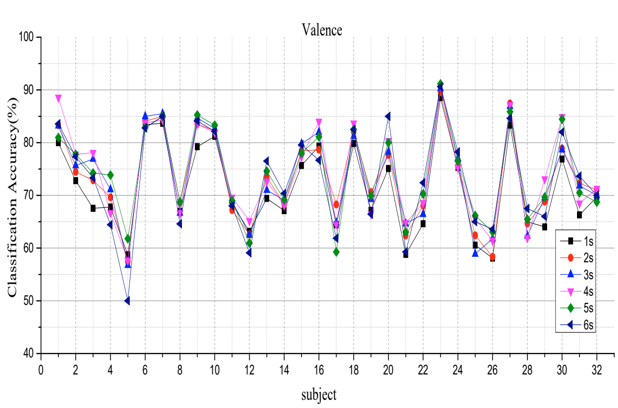


Fig. 1. The average classification accuracy for each of the 32 subjects when using 6 different time-windows on the EEG signals (valence-dimension).

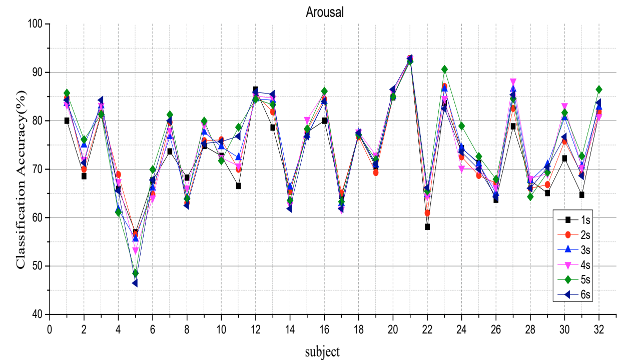


Fig. 2. The average classification accuracy for each of the 32 subjects when using 6 different time-windows on the EEG signals (arousal-dimension).

For each channel the 60s EEG signal is divided into 15 non-overlapping segments, each with the equal length of 4s. Here, the 4s sliding time-window is chosen based on the results shown in Figs. 1 and 2. In order to determine the optimal length of the time-window, we compare the accuracy of binary emotion classification by varying the EEG time-window length from 1-6s. In Figs. 1 and 2, the valence dimension is divided into two classes of emotion: positive vs. negative emotion, while arousal dimension is divided into two classes: high vs. low arousal. The EEG gamma sub-band extracted by wavelet decomposition is used as emotion-relevant features in Figs. 1 and 2. We use PCA to reduce feature dimensionality and random forest as the classifier algorithm.

From Table 1, we can find that regardless of valence or arousal dimension, the correct classification rate (CCR) is the best when the time window is 4s long, so in the subsequent experiments, we choose 4s as the length of time window. After data preprocessing, the number of sample data per subject is 40*15=600. For 32 subjects, a total of 32*600=19200 sample data are available in the dataset.

Table 1. The subject-averaged classification accuracy (%) when using 6 different time-windows on EEG signals.

| | 1s | 2s | 3s | 4s | 5s | 6s |
|---|---|---|---|---|---|---|
| Arousal | 72.99 | 74.18 | 74.80 | 74.83 | 74.57 | 74.41 |
| Valence | 71.31 | 73.37 | 73.55 | 74.18 | 74.01 | 73.42 |

Among the studies on emotion recognition using the DEAP database, many studies focused on the binary classification problem, i.e., two emotion classes are usually considered – either positive and negative emotion (along the valence dimension) or high and low arousal (Yin et al., 2017; Petrantonakis and Hadjileontiadis, 2011; Daimi and Saha, 2014; Yoon and Chung, 2013). The actual (or target) emotion label for each EEG data point in high-dimensional feature space is usually determined by thresholding of the subjective rating data. Unfortunately this threshold method is too simplistic and hard to choose the appropriate threshold. In order to overcome this problem, we also consider the binary classification on the arousal and valence dimension, but instead of hard threshold use the *k*-means clustering algorithm to determine the actual two-class labels. As shown in Fig. 3, the valence dimension is classified into two emotions, viz. negative emotion (Low Valence – LV, Class 1) and positive emotion (High Valence – HV, Class 2). As shown in Fig. 4, the arousal dimension is classified into two emotions: low arousal (LA, Class 1) and high arousal (HA, Class 2). Fig. 5 shows the emotion plane. The cluster centers, boundaries and the size of each emotion class are given in Table 2.

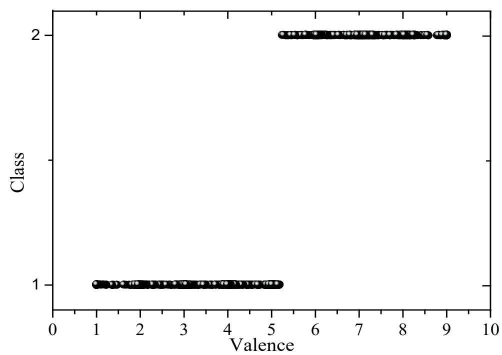


Fig. 3. The clustering results of valence ($k$=2).

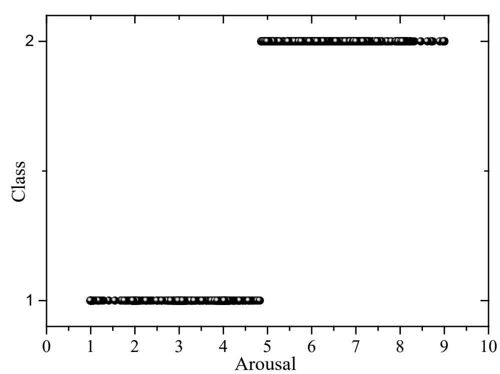


Fig. 4. The clustering results of arousal ($k$=2).

Table 2. Clustering results of each emotion dimension ($k$=2).

| Dimension | Cluster Center | | Class Border | Class Size (Low/High) |
|---|---|---|---|---|
| Valence | 3.33 | 7.10 | 5.21 | 477/803 |
| Arousal | 3.05 | 6.64 | 4.85 | 682/597 |

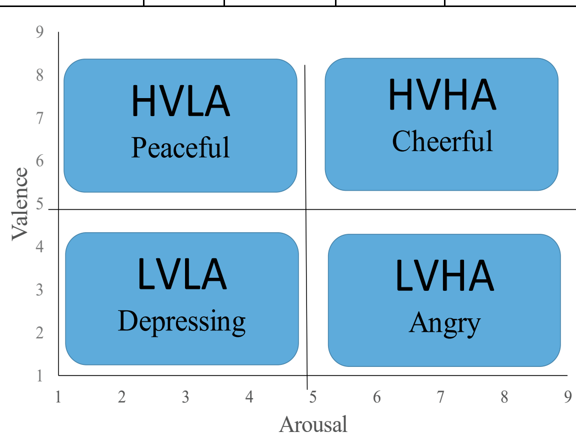


Fig. 5. A 2D model of emotion.

### 2.3 Feature reduction and selection

Dimensionality reduction or feature selection is an important step in EEG-based emotion recognition. An effective feature dimensionality reduction and selection method can not only accelerate model training, but also improve the recognition accuracy of the model.

After each channel of EEG signal is decomposed by 5-level wavelet decomposition, the wavelet coefficients corresponding to the five frequency bands can be obtained. Three features, namely *wavelet energy*, *wavelet energy ratio*, and *wavelet entropy*, are extracted using wavelet coefficients for each frequency band. The feature dimensionality for each frequency band is 32*3=96, while the dimensionality of wavelet feature of all the five frequency bands is 32*5*3=480. If the approximate entropy and sample entropy are used in nonlinear dynamics analysis, the feature dimensionality is 32*2=64. In this paper, three feature reduction methods (KSR, LPP, and PCA) and two feature selection methods (mRMR and Relieff) are employed to process EEG features. PCA is used to provide the baseline performance for comparing those feature reduction and selection algorithms.

Spectral Regression (SR) algorithm is effective when dealing with massive amount of data (Cai, He and Han, 2011; Golub and van Loan, 2012; Franklin, 2005). The SR algorithm includes the following key steps (Zhang et al., 2020):

**Step 1:** regularized least squares

Find ($c$-1) vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{c-1} \in \mathbb{R}^n$ as the solution to the system of equations:

$$(K + \alpha \mathrm{I})\boldsymbol{\alpha}_k = y_k \tag{1}$$

where $K$ is the $m \times m$ Cramer matrix and I is the unity matrix. It is easy to show that the function $f(x) = \sum_{i=1}^{m} \alpha_i^k K(x, x_i)$ is the solution to the minimization problem:

$$\min_{f \in H_k} \sum_{i=1}^{m} (f(x_i) - y_i^k)^2 + \alpha \|f\|_K^2 \tag{2}$$

where $\alpha_i^k$ denotes the i-th component of the vector $\boldsymbol{\alpha}_k$ .

**Step 2:** Spectral regression discriminant analysis

Let $\Theta = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{c-1}]$ be a $m \times (c-1)$ transformation matrix, then the samples can be embedded into ($c-1$)-dim. subspace in the form as follows:

$$x \to z = \Theta^T \mathbf{K}(:, x) \tag{3}$$

where $\mathbf{K}(:, x) = [K(x_1, x), K(x_2, x), \cdots, K(x_m, x)]^T$ .

## 3. AFFECT CLASSIFICATION RESULTS

After wavelet decomposition of the original EEG signals, we obtain the wavelet features corresponding to the five frequency sub-bands. In order to find the frequency band most related to emotion, the EEG features in each frequency sub-band are first used for dimensionality reduction and classification, and then all the five sub-bands are used jointly to check whether or not it can improve the accuracy of emotion recognition. In order to obtain a more reliable recognition accuracy, for each subject, 5-fold cross-validation is used to divide the dataset into training set and testing set. Accordingly, for each subject the size of training and testing set is 480 and 120, respectively. The wavelet feature dimensionality in each subband is 32*3=96, while the concatenated feature in all the five subbands has 96*5=480 dimensions. The feature dimensionality of approximate entropy and sample entropy are 32 for each. We compare four different classifiers, including RF, KNN, NB, and SVM.

The dimensionality reduction methods under comparison include KSR, LPP, mRMR, Relieff and PCA. The parameter settings for each dimension reduction method are as follows:

1) For KSR and LPP, Gaussian kernel function is used;

2) For KSR, L2 norm is used and the regularization parameter is set to 0.01;

3) For LPP, the number of nearest neighbors is set to 5 and the Euclidean distance function is used; For mRMR and Relieff, the number of features is set to 20; and

4) For PCA, the variance contribution is set to 0.98.

NDR in Fig. 6-9 indicates no dimensionality reduction performed.

Binary emotion classification in question involves two separate tasks: (i) classification of high and low emotional arousal; and (ii) classification of positive and negative emotions;. The 32-subject-averaged testing accuracy by using wavelet features is compared in Figs. 6 (classification of high vs. low arousal) and 7 (classification of positive vs. negative emotion).

We can observe from Figs. 6 and 7 that when using wavelet features, the features in the EEG gamma sub-bands lead to the highest higher classification accuracy, for both the valence and arousal dimension, among the five sub-bands. In terms of dimensionality reduction methods, LPP results in the highest arousal classification accuracy of 87.26±3.26% and the highest valence classification accuracy of 87.50±3.10%. By observing Figs. 8 and 9, we see that when using nonlinear dynamic features KSR is more advantageous than other dimensionality reduction methods, with the highest arousal classification accuracy of 94.85±1.77% and the highest valence classification accuracy of 93.90±2.23%.

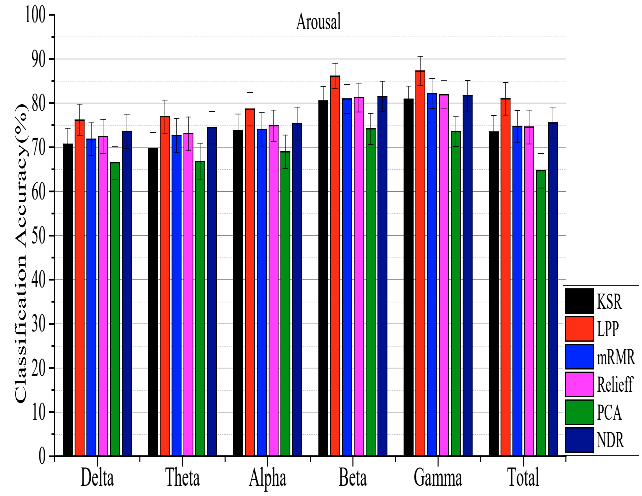


Fig. 6. The binary emotion classification results (wavelet features; Arousal-dimension).

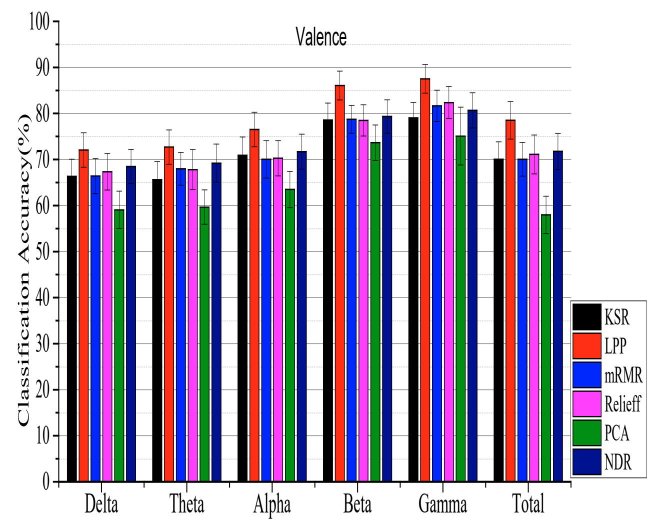


Fig. 7. The binary emotion classification results (wavelet features; Valence-dimension).

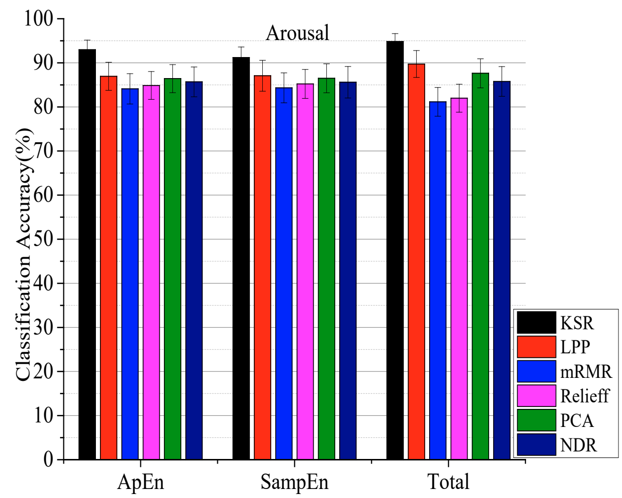


Fig. 8. The binary classification results (ApEn and SampEn features; Arousal-dimension).

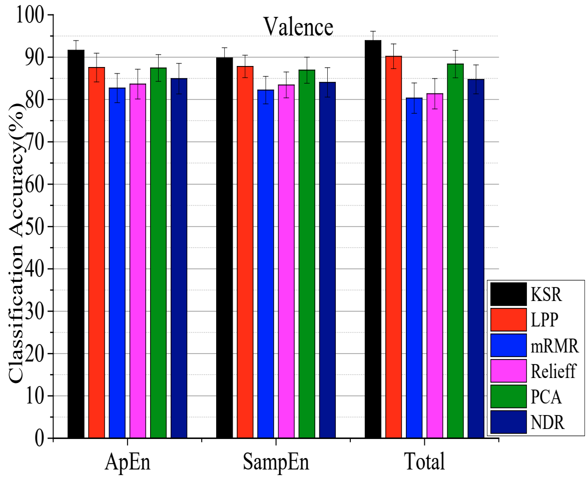


Fig. 9. The binary classification results (ApEn and SampEn features; Valence-dimension).

## 4. SUMMARY AND CONCLUSIONS

In this paper, wavelet features and nonlinear dynamic features of EEG signals are extracted by taking into account the baseline EEG features. The main contributions of the present work are summarized as follows:

(1) Many literature (see for example Daimi and Saha, 2014; Yoon and Chung, 2013; Zhuang, Rozgic and Crystal, 2014; Jirayucharoensak, Pan-Ngum and Israsena, 2014) used threshold method to label physiological data. In this article, the valence and the arousal dimension are classified separately and *c*-means clustering algorithm is employed to determine the ground truth for each EEG data point in the feature space.

(2) In many studies on emotion recognition (for instance, Yoon and Chung, 2013; Daimi and Saha, 2014; Wang, Nie and Lu, 2014; Zhong et al., 2017), researchers used only EEG data during emotion stimulation and evoking while disregarding the baseline data of the subject without emotion stimulus. In this paper, we use the relative difference features between the 60s emotion-related EEG features and the baseline EEG features. The comparative results (omitted here to meet the allotted length of the paper) show that when taking into account the baseline data, the affect classification accuracy can be significantly improved across all the 32 subjects.

(3) Two feature extraction methods, namely wavelet transform and nonlinear dynamics analysis, are investigated and compared. The features in the five EEG frequency sub-bands are extracted by wavelet decomposition, and then the features in each sub-band are separately fed into the classifier. For comparison, the features in all the five sub-bands are also input into the classifier. It is found that the classification accuracy using the EEG gamma sub-band features is the highest, suggesting that the gamma band reflect the change of emotions the best. On the other hand, it is found that improved affect classification accuracy can be achieved when the approximate entropy and sample entropy are used jointly as the emotion features.

From the comparative results presented, we may draw the following conclusions: (i) When using the approximate entropy and sample entropy features and KSR for feature reduction, we can achieve high binary emotion classification accuracy; (ii) When using wavelet features, the EEG gamma sub-band features lead to the highest emotion recognition rate, followed by the beta sub-band; and (iii) KSR is the best dimensionality reduction algorithm, while random forest leads to the most accurate EEG-based emotion recognition.

## REFERENCES

Anderson, K. and McOwan, P.W. (2006). A real-time automated system for the recognition of human facial expressions. *IEEE Trans. on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. **36**(1), pp. 96-105.

Cai, D., He, X., Han, J. (2011). Speed up kernel discriminant analysis. *The Intl. J. on Very Large Data Bases*, vol. **20**(1), pp. 21-33.

Chen, P., Zhang, J., Wen, Z., Xia, J., and Li, J. (2018), Emotion recognition of EEG based on kernel spectral regression and random forest algorithm, *J. of East China University of Science and Technology (Natural Sci. Ed.)*, vol. **44**(5), pp. 744-751.

Daimi, S.N. and Saha, G. (2014), Classification of emotions induced by music videos and correlation with participants' rating, *Expert Systems with Applications*, vol. **41**(13), pp. 6057-6065.

Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, vol. **27**(2), pp. 83-85.

Golub, G.H. and van Loan, C.F. (2012). *Matrix Computations*. JHU Press.

Jirayucharoensak, S., Pan-Ngum, S., and Israsena, P. (2014). EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *The Scientific World J.*, 627892.

Koelstra, S., Muhl, C., Soleymani, M. et al. (2012), DEAP: A database for emotion analysis using physiological signals, *IEEE Trans. on Affective Computing*, vol. **3**(1), pp. 18-31.

Li, X., Hu, B., Zhu, T. et al. (2009). Towards affective learning with an EEG feedback approach. In *Proc. of the 1st ACM Intl. Workshop on Multimedia Technologies for Distance Learning*, pp. 33-38.

Pantic, M. and Rothkrantz, L.J.M. (2000). Automatic analysis of facial expressions: The state of the art. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. **22**(12), pp. 1424-1445.

Petrantonakis, P.C. and Hadjileontiadis, L.J. (2011), A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition, *IEEE Trans. on Information Technology in Biomedicine*, vol. **15**(5), pp. 737-746.

Petrushin, V.A. (1999). Emotion in Speech: Recognition and Application to Call Centers. In *Proc. of the 1999 Conf. on Artificial Neural Networks in Eng.*, pp. 7-10.

Wang, X.W., Nie, D., and Lu, B.L. (2014). Emotional state classification from EEG data using machine learning approach. *Neurocomputing*, vol. **129**(4), pp. 94-106.

Yin, Z., Zhao, M., Wang, Y. et al. (2017), Recognition of emotions using multimodal physiological signals and an ensemble deep learning model, *Computer Methods & Programs in Biomedicine*, vol. **140**(C), pp. 93-110.

Yoon, H.J. and Chung, S.Y. (2013). EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm. *Computers in Biology & Medicine*, vol. **43**(12), pp. 2230-2237.

Zhang, J., Yin, Z., Chen, P., and Nichele S. (2020), Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review, *Information Fusion*, vol. **59**, pp. 103-126.

Zhong, Y., Zhao, M., Wang, Y. et al. (2017). Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. *Computer Methods & Programs in Biomedicine*, vol. **140**(C), pp. 93-110.

Zhuang, X., Rozgic, V., and Crystal, M. (2014). Compact unsupervised EEG response representation for emotion recognition. In *Proc of IEEE-EMBS Intl. Conf. on Biomedical and Health Informatics*, pp. 736-739.