JOURNAL OF SCIENCE
SAKARYA UNIVERSITY

Sakarya University Journal of Science
ISSN 1301-4048 | e-ISSN 2147-835X | Period Bimonthly | Founded: 1997 | Publisher Sakarya University | http://www.saujs.sakarya.edu.tr/

Title: Feature Normalization Effect in Emotion Classification based on EEG Signals

Authors: Orhan Akbulut
Recieved: 2019-09-09 21:04:38

Accepted: 2019-10-07 14:36:40

Article Type: Research Article
Volume: 24
Issue: 1
Month: February
Year: 2020
Pages: 60-66

How to cite
Orhan Akbulut; (2020), Feature Normalization Effect in Emotion Classification based on EEG Signals. Sakarya University Journal of Science, 24(1), 60-66, DOI: 10.16984/saufenbilder.617642
Access link
http://www.saujs.sakarya.edu.tr/tr/issue/49430//617642

Sakarya University Journal of Science 24(1), 60-66, 2020

Feature Normalization Effect in Emotion Classification based on EEG Signals

Orhan Akbulut*[1]

Abstract

Normalization of data in classification-based problem is a fundamental task where binary or multi classifier systems integrate it as a sub-system. Normalization can be thought as a mapping function that makes a transformation from one space to another space. Different types of normalization methods have been proposed depending on the data content. Recently, researches are carried out on whether this process is really necessary. In this paper, the performances of the different normalization methods for Electroencephalogram (EEG) signal based emotion classification are evaluated. Support vector machine based binary classifier is used in emotion classification. Different kernel functions for support vector machine are also considered. Although the experimental findings may not reveal a significant performance difference between different types of normalization, the normalization process increases classification performance of the emotion recognition, in general.

Keywords: normalization, classification, support vector machine, electroencephalogram

1. INTRODUCTION

In many applications, using the raw data without pre-processing can be challenging in terms of obtaining the desired performance. Since the content of the raw data may have different values of range, these data variables can become important. For example, some instances having large values in the data may have larger impact than other instances in the stage of cost function [1]. Hence, making an inference from data, which have different types of range values, can directly affect the performance of the data-dependent system, especially in distance based computations. To handle this issue, one solution is to transform data variables to a different space such that it guarantees to be within similar ranges. This process is known as data normalization or standardization that is widely used in many areas such as image processing, machine learning, pattern recognition and deep learning based applications. Data normalization can also be considered as a pre-processing step of a system.

In general, data to be processed consist of a series of different features or attributes. Thus, in many practical applications, the statement of -feature normalization- is preferred in a more generic way instead of data normalization. Feature normalization step can be critical for a subsystem of the overall system. The necessity of using this stage highly depends on the nature of the features.

* Corresponding Author: orhan.akbulut@kocaeli.edu.tr
[1] Kocaeli University, Computer Engineering Department, Kocaeli, Turkey. ORCID: 0000-0003-0096-0688

In the literature, various types of normalization have been applied to data sets in different application areas. In [2], impact of the normalization of variables in cluster analysis is examined. Data sets generated artificially are exposed different scenarios such as inclusion of outliers, addition of random noise. These data sets are then normalized using several normalization approaches. According to [3], data normalization not only enables the attributes to have equal dynamic range but also provides numerical simplicity during the calculation. In backpropagation based neural network algorithm, normalizing the data can help to speed up the learning process [4]. In [5], the impact of data normalization on the unsupervised classification process is considered as data transformation. With the data transformation, data distributions become more normal and homogeneous. Data transformation, however, results in less class separability because of homogeneity of transformed data. In [6], a general assessment of the normalization effect in terms of different aspects is handled in data mining application. A comparative study on various cancer gene expression data sets for the requirement of normalization is presented in [7]. In [8], the performance of the different normalization methods on the artificial neural network is addressed using diabetes dataset. In [9,10], the performance of the modified K-means clustering method is examined considering pre-processing approaches such as normalization and outlier detection. In [11], specialized feature normalization approaches are addressed for part-based Bag-of-Features models. In [12], a new data normalization approach is proposed for features obtained using triangle coordinate system. The impacts of the feature normalization methods on the optical-synthetic aperture radar remote sensing images based classification are discussed in [13]. In breast tumor based classification, whether the normalization process is important or not, is discussed in [14]. In [15], feature normalization issue is handled in image classification based on convolutional neural networks. Comprehensive study on different feature normalization methods for back-propagation models are applied in [16]. In [17], the performance analysis of the normalization process is carried out in the texture classification. In the texture analysis, textures are obtained using gray-level co-occurrence matrix.

In this paper, the effect of feature normalization is addressed for emotional classification using Electroencephalogram (EEG) signals. The main goal of the study is to reveal whether or not standardization is necessary for the features obtained by EEG signals. Besides, the methodology used in the paper provides insight into how standardization methods effect emotional classification performance. The rest of the paper is organized as follows: Section 2 contains a brief introduction to the emotion analysis and EEG signals. The methodology of the paper including dataset and feature normalization methods is given in Section 3. Experimental results are presented in Section 4. Finally, Section 5 concludes the paper.

2. HIGHLIGHTS ON EEG SIGNALS AND EMOTION ANALYSIS

2.1. EEG Signals

EEG is an electrophysiological method which monitors electrical activity of the brain. Different symptoms such as c/ non-epileptic seizures and emotions cause unusual electrical activities in the brain [18-19]. Thanks to EEG, it can be possible to take prevention early or to take actions at the event time.

In EEG, using a pair of electrodes placed in the brain, voltage differences are measured over the different locations. Each electrode is represented with a letter and a number. Left hemisphere in the brain covers odd numbers while right hemisphere consists of even numbers. Accordingly, two different electrode placements, which are 10/10, and 10/20, are commonly adopted in EEG. The “10/x” terminology corresponds to the positioning of the electrodes between nasion and inion at 10 or 20 percent intervals [20]. The international 10/20 electrode placement system is symbolically illustrated in Figure 1.

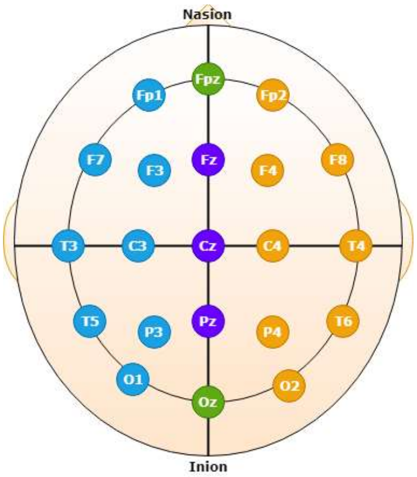

Figure 1. International 10/20 electrode placement system.

Though EEG waveforms consist of time series signal, these waveforms have different frequency bandwidths, as given in Table 1. These frequency ranges are often used in the interpretation of signals emitted by brain.

Table 1. Frequency bandwidths in EEG

Bands	Freq. Range (Hertz)
Delta	1-3
Theta	4-7
Alpha	8-12
Beta	13-30
Gama	31-100

2.2. Emotion Modeling

It is clear that there is a strong relationship between human behaviors and emotions. Understanding and modelling emotions, therefore, can be important to analyze human behaviors. In the literature, there are several emotion models used to categorize the emotions. One of them is the basic emotion model where six major emotions are adopted as in [21]. The other one, [22], is 2-dimensional emotion modelling (2DEM), as shown in Figure 2. In 2DEM, axes are represented by arousal and valence. Arousal corresponds to the energy level of the emotions while valence classifies the emotion with respect to its positive and negative contents.

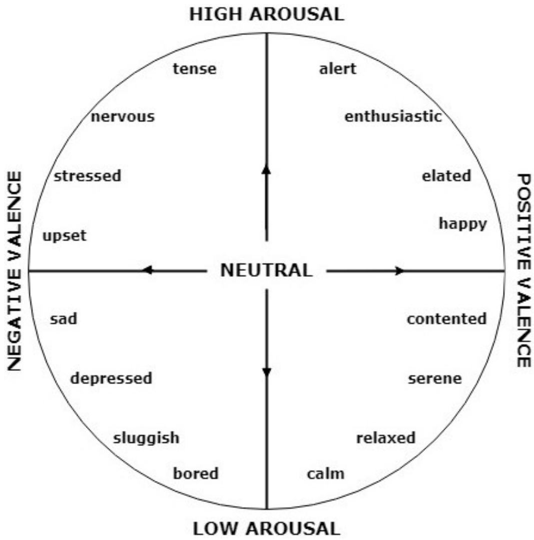

Figure 2. 2-Dimensional emotion model

Emotion classification based many approaches have been proposed using EEG signals [23-26]. EEG based researches on emotion classification generally focus on selection of classification methods, feature extraction, and EEG channel selection. To the best of our knowledge, there is no study about whether feature normalization is required in the EEG based emotion classification.

3. METHODOLOGY

In this study, feature normalization effect in emotion classification is addressed. In this context, emotions are classified using EEG signals. The flowchart of the presented methodology is shown in Figure 3. Each sub-system of the flowchart is handled in the following sub-sections.

3.1. DEAP Dataset

DEAP introduced in [27] is a comprehensive dataset containing EEG signals of 32 participants. Details of the dataset are given in Table 2. The dataset also includes facial information obtained from some participants during the experiment. In this study, EEG signals are provided using DEAP dataset. Note that, EEG signals used in this paper are filtered in default.

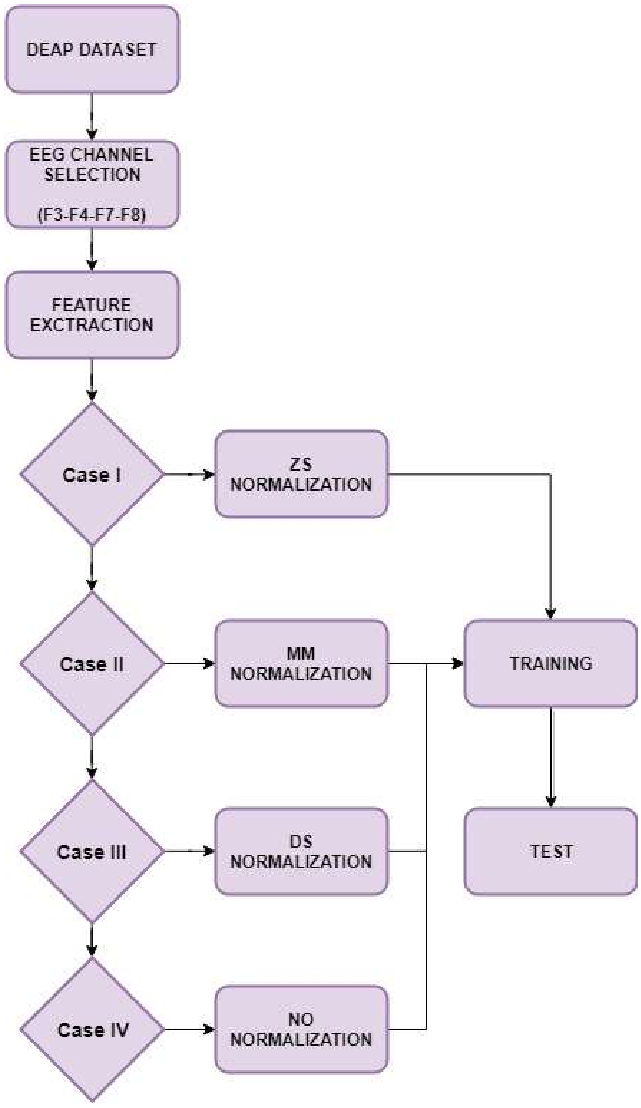

Figure 3. Flowchart of the proposed methodology

Table 2. DEAP dataset

Information of Dataset	Values
Number of participants (P)	32
Number of channels (C)	32
Number of videos (V)	40
Video length (sec)	63
Sampling rate (Hertz)	128
Samples per channel	8064
Valence	Yes
Arousal	Yes
Dominance	Yes
Familiarity	Yes

3.2. Feature Extraction

Since EEG signals consist of time-series signals, extraction of features is a fundamental requirement. Therefore, features are obtained from time-domain and frequency domain of the EEG signals.

Let $\mathbf{f} \in \mathbb{R}^{(P\times C\times V)\times 1}$ be column vector, considering the observations defined in DEAP dataset. Thus, feature matrix can be defined as the following form,

$$F = [\mathbf{f}_1 \quad \dots \quad \mathbf{f}_d]_{(P\times C\times V)\times d} \tag{1}$$

where $P \times C \times V$ corresponds to total observation number. Note that, *P, C,* and *V* refer to number of participants, number of channels and number of videos, respectively. Also, *d* corresponds to the feature number. Seven distinct features (*d*=7) are extracted from the EEG signals, as in [26]. These features are given in Table 3. For the NSI, each feature vector is divided into small segments. Then, the variation of mean of these segments is computed.

Table 3. Extracted features

Domain	Features
Time	Mean of the 1st order derivative
Time	Mean of the 2nd order derivative
Time	Non-stationary index (NSI) [28]
Frequency	Average power of the Delta band
Frequency	Average power of the Theta band
Frequency	Average power of the Alpha band
Frequency	Average power of the Beta band

3.3. Feature Normalization Stage

In this sub-section, we mention some of normalization methods such as min-max normalization, decimal scaling and z-score.

Taking into account the feature vector **f**, it is clear that f_i corresponds to feature value. Let, f_i and f_i^{nor} be original and normalized feature value, respectively. Then, well-known normalization methods can be expressed as follows;

- **Z-score normalization (ZS):** This method is widely used for feature normalization. Normalized features which have zero mean and unit variance, are computed as follows;

$$f_i^{nor\_z} = \frac{f_i - \mu_{\mathbf{f}}}{\sigma_{\mathbf{f}}} \tag{2}$$

where μ_f and σ_f are mean and standard deviation of the related feature vector, respectively.

- **Min-Max normalization (MM):** In this normalization method, feature values are normalized to a given range. Commonly, this range is applied as [0, 1] as follows;

$$f_i^{nor\_mm} = \frac{f_i - min(\mathbf{f})}{max(\mathbf{f}) - min(\mathbf{f})} \tag{3}$$

- **Decimal-scaling normalization (DS):** Unlike the linear normalization methods, depending on the data content nonlinear normalization approach may give better performance. In this method, feature values are normalized to the range (0,1) in nonlinear as follows;

$$f_i^{nor\_ds} = \frac{f_i}{10^{round(\log_{10} max(\mathbf{f}))}} \tag{4}$$

3.4. Emotion Classification

In this study, in order to classify EEG based emotions Support Vector Machine (SVM) algorithm is used as a classifier model. 80% of the total observations are applied for training stage while remaining observations are evaluated for testing phase. Emotions for arousal and valence in DEAP dataset are labeled between 1-9 values. Since a binary classification based SVM algorithm is adopted in this study, Emotions for both arousal and valence are divided into two classes. Positive class is determined under condition $5 \leq positive$ while negative class is assigned under condition $5 > negative$.

4. EXPERIMENTAL RESULTS

In order to assess feature normalization effect in emotion classification, three different normalization methods, ZS, MM, DS, are considered. Emotions are also classified without normalization (W/O). Classification results have been obtained by using different kernel functions such as “linear”, “polynomial” and “radial based functions (RBF)”. Instead of using all channels for emotion recognition, fewer channels have been taken into consideration, as in [26]. In this context, “F3”, “F4”, “F7”, and “F8” channels are used. The performance metrics given in the following Tables are based on classification accuracy.

The classification performances of normalization methods on emotion recognition with respect to valence are given in Table 4. All normalization methods exhibit similar performance in both training and test phases while classification performance deteriorates in case of without normalization. Note that, though training score for the RBF based SVM is very high (W/O), overfitting is occurred. On the other hand, linear SVM algorithm for the W/O has poor generalization capability where underfitting is occurred.

Table 4. Classification performances of normalization methods on emotion recognition with respect to valence

Norm. methods	Valence					
	Linear SVM		Poly. SVM		RBF SVM	
	Training	Test	Training	Test	Training	Test
ZS	56.84	55.58	59.06	55.96	61.91	56.35
MM	56.84	55.57	57.18	55.27	57.08	55.66
DS	56.81	55.56	57.30	55.47	57.47	55.37
W/O	43.17	45.12	56.52	55.37	95.00	55.18

The classification performances of normalization methods on emotion recognition with respect to arousal are given in Table 5. First of all, in general, arousal classification accuracies are higher than valence classification accuracies. Similar to Table 4, all normalization methods has similar classification accuracies in both training and test phases in terms of. As seen in Table 5, in case of without normalization, underfitting and overfitting issues remain.

Table 5. Classification performances of normalization methods on emotion recognition with respect to arousal

Norm. methods	Arousal					
	Linear SVM		Poly. SVM		RBF SVM	
	Training	Test	Training	Test	Training	Test
ZS	58.33	61.52	59.30	61.72	62.55	62.21
MM	58.25	61.32	56.81	61.52	58.59	61.52
DS	58.25	61.32	58.74	61.52	58.57	61.62
W/O	42.53	39.26	51.34	50.88	95.12	61.33

5. CONCLUSION

In this study, feature normalization issue is addressed in EEG signal based emotion classification. In this context, the performance of well-known normalization methods has been handed separately in SVM based emotion classification. Considering the experimental findings, SVM classifier with polynomial kernel presents more robust solution compared to other kernel functions. Moreover, it is greatly possible to obtain better classification performance with the better feature extraction techniques.

6. REFERENCES

[1] S. Theodoridis, and K. Koutroumbas, Pattern Recognition, 4th Edition, Academic Press, 2008.

[2] G. W. Milligan and M. C. Cooper, “A study of standardization of variables in cluster analysis,” Journal of Classification, vol. 5, pp. 181–204, 1988.

[3] C. W. Hsu, C. C. Chang, and C. J. Lin, “A practical guide to support vector classification” Tech. Rep., 2003.

[4] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques. San Mateo, CA, USA: Morgan Kaufmann, 2006.

[5] I. L. Fonseca, “The Impact of data normalization on unsupervised continuous classification of landforms,” International Geoscience and Remote Sensing Symposium, 2003.

[6] L. A. Shalabi, and Z. Shaaban, “Normalization as a preprocessing engine for data mining and the approach of preference matrix,” International Conference on Dependability of Computer Systems, 2006.

[7] M. C. P. Souto, D. S. A. Araujo, I. G. Costa, R. G. F. Soarez, T. B. Ludermir, and A. Schliep, “Comparative study on normalization procedures for cluster analysis of gene expression datasets,” IEEE International Joint Conference on Neural Networks, 2008.

[8] T. Jayalaklashmi, and A. Santhakumaran, “Statistical normalization and back propagation for classification,” International Journal of Computer Theory and Engineering, vol. 3, no. 1, pp. 89-93, 2011.

[9] V. R. Patel, and R. G. Mehta, “Impact of outlier removal and normalization approach in modified k-Means clustering algorithm,” International Journal of Computer Science Issues, vol. 8, no. 5, pp 331-336, 2011.

[10] I. B. Mohamad, and D. Usman, “Standardization and its effects on K-Means clustering algorithm,” Research Journal of Applied Sciences, Engineering and Technology, vol. 6, no. 17, pp. 3299-3303, 2013.

[11] L. Xie, Q. Tian, and B. Zoang, “Feature normalization for part-based image classification,” IEEE International Conference on Image Processing, 2013.

[12] M. S. Azmi, N. A. Arbain, A. K. Muda, Z. A. Abas, and Z. Muslim, “Data normalization for triangle features by adapting triangle nature for better classification,” IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, 2015.

[13] H. Zhang, H. Lin, and Y. Li, "Impacts of feature normalization on optical and SAR data fusion for land use/land cover classification," IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 5, pp. 1061-1065, 2015.

[14] B. K. Singh, K. Verma, and A. S. Thoke, "Investigations on impact of feature normalization techniques on classifier's performance in breast tumor," International Journal of Computer Applications, vol. 116, no. 19, pp. 11-15, 2015.

[15] K. P. Kuntal, and K. S. Sudeep, "Preprocessing for image classification by convolutional neural networks," IEEE International Conference on Recent Trends in Electronics Information Communication Technology, 2016.

[16] A. S. Easa, and W. Arabo, "A Normalization Methods for Backpropagation: A Comparative Study," Science Journal of University of Zakho, vol. 5, no. 4, pp. 319-323, 2017.

[17] M. Kociolek, M. Strlecki, and S. Szymajda, "On the influence of the image normalization scheme on texture classification accuracy," Signal Processing: Algorithms, Architectures, Arrangements, and Applications, 2018.

[18] Harender, and R. K. Sharma, "DWT based epileptic seizure detection from EEG signal using k-NN classifier," International Conference on Trends in Electronics and Informatics, 2017.

[19] J. W. C. Medithe, and U. R. Nelakuditi, "Study of normal and abnormal EEG," International Conference on Advanced Computing and Communication Systems, 2016.

[20] E. K. St Louis, and L. C. Frey (Eds), "Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants", Chicago, IL: American Epilepsy Society, 2016.

[21] P. Ekman, "Basic emotions", Handbook of cognition and emotion, pp. 45-60, 1999.

[22] J. A., Russell, "A circumplex model of affect," Journal of Personality and Social Psychology, vol. 39, no. 6, pp. 1161-1178, 1980.

[23] M. Mikhail, K. El Ayat, J. A. Coan, J. J. Allen, "Using minimal number of electrodes for emotion recognition using brain signals produced from a new elicitation technique," International Journal of Autonomous and Adaptive Comm. System, vol. 6, no. 1, pp.80-97, 2013.

[24] N. Jatupaiboon, S. Pangum, and P. Israsena, "Emotion classification using minimal EEG channels and frequency bands," 10th International Joint Conference on Computer Science and Software Engineering, pp. 21-24, 2013.

[25] J. Zhang, M. Chen, S. Zhao, S. Hu, Z. Shi, and Y. Cao, "ReliefF-based EEG sensor selection methods for emotion recognition," Sensors, vol. 16, no. 10, 2016.

[26] S. Kul, P. O. Durdu, and O. Akbulut, "Performance Comparison of EEG Channels in Emotion Recognition," 27th Signal Processing and Communications Applications Conference (SIU), 2019.

[27] S. Koelstra, et al., "DEAP: A Database for Emotion Analysis ; Using Physiological Signals," IEEE Transactions on Affecting Computing, vol. 3, no. 1, pp.18-31, 2012.

[28] R. Jenke, A. Peer, M. Buss, "Feature extraction and selection for emotion recognition from EEG" IEEE Transactions on Affective Computing, vol. 5, no. 3, pp. 327-339, 2014.