Multichannel Electroencephalography-based Emotion Recognition Using Machine Learning

To cite this article: IN Yulita et al 2019 J. Phys.: Conf. Ser. 1230 012008

View the article online for updates and enhancements.
Multichannel Electroencephalography-based Emotion Recognition Using Machine Learning

IN Yulita¹, RR Julviar¹, A Triwahyun², T Widiastuti²

¹Department of Computer Science, Universitas Padjadjaran, Jl. Raya Bandung – Sumedang Km. 21, Sumedang 45363, Indonesia
²Faculty of Psychology, Universitas Padjadjaran, Jl. Raya Bandung – Sumedang Km. 21, Sumedang 45363, Indonesia

Email: intan.nurma@unpad.ac.id

Abstract. In recent years, research in the field of human-computer interaction (HCI) has focused on strengthening machine functions in recognizing and understanding human emotions. Emotion recognition can be done in several ways, among others, through sounds, facial expressions, or a combination of both. The different sounds and facial expressions from different races and nations cause less accurate in the reading of emotions using these methods. Another method for recognizing emotions can be done by analysing the data from an electroencephalograph (EEG). The EEG signals from the human brain are the result of various activities carried out. One of them is emotion. The EEG signal used in this study came from the DEAP dataset. This dataset consists of 32 files, each of which contains 40 EEG recordings. The emotions from this dataset are classified based on the dimensions of arousal and valence. The signal was then decomposed into three different frequency groups (alpha, beta, and gamma) through band-pass filtering. After that, the principal component analysis (PCA) and resampling were carried out. The classification processes used a number of methods of machine learning. The result was known that the performance of K-star was the highest while naïve Bayes was the lowest. The accuracies of K-star in arousal and valence classification were 81.2, 82.6, respectively. The naïve Bayes got 51.2 for the arousal, and 52.5 for the valence.

1. Introduction

Emotion is a feeling that arises in a person as a result of stimulation, both from within himself and from outside [1]. It is closely related to the psychological condition and mood of a person expressed in certain forms of behaviour [2]. The feeling can be positive emotion (good emotion), and can be negative emotion (bad emotion). Many people interpret this word as a form of anger, but actually the emotion represents various forms of human feelings. Some forms of emotions or human feelings include: sad, happy, and angry. The ability to recognize self-emotion is an important ability for someone [3]. With this ability, he can recognize his own feelings when those feelings or emotions emerge. This is often said to be the basis of emotional intelligence. A person with this emotional intelligence will have a good sensitivity. It will affect the process of making decisions firmly [4].

The existence of emotions is also important related to the interaction between humans and machines [5]. Without the ability to process emotions, the machine or computer will have difficulty communicating with humans naturally. The existence of a machine now can help humans so that work becomes easier. But it will not be achieved when the machine cannot understand language and human intentions, especially human emotions. In recent years, the research in the field of human computer interaction (HCI) is focusing on strengthening machine functions in recognizing and understanding
human emotions [6]. The recognition itself can be done in several ways, among others, through sounds [7], facial expressions [8], or a combination of both. The differences in the sounds or facial expressions from different races and nations cause less accurate reading of emotions using this method [9]. Another method for recognizing emotions can be done by analysing electroencephalography (EEG) [10].

EEG is one of the tests conducted to measure the voltage fluctuations generated by ion currents in brain neurons [11]. In the clinical context, EEG refers to recording spontaneous electrical activity of the brain for a certain period, usually 30-40 minutes. It is recorded from many electrodes mounted on the scalp. The EEG will record the electrical activity of the brain, which is represented in the form of a wave line. Researchers believe that the state of the brain changes with the changes that occur in a person's feelings. This situation can be recorded properly with EEG. Through EEG, recording records changes in brain waves that vary according to feelings or emotions. The EEG has several advantages for studying human emotions. Some of the advantages are high-speed, non-invasive and does not cause pain in the subject. It is important to get genuine emotions from the subject. However, it should be noted that there is some noise outside the measurement object that might appear as well as the body fatigue factor or other external factors. EEG can be decomposed into five different frequencies namely delta, theta, alpha, beta, and gamma [12]. The frequency range indicates a difference in physical condition in a person, but the frequency division is not enough to recognize someone's emotions. Because, in an EEG signal recording there are several different frequencies that form a certain pattern.

The EEG research for classification has been carried out. For example research for the stages of sleep [13], and epilepsy [14]. Some researchers have carried out an analysis of the relationship between EEG and emotional conditions in humans. Some of them are research conducted by LSTM recurrent neural networks [15], fuzzy [16], K-nearest neighbors [17], auto-regressive modelling [18], and support vector machine [19]. Hu et al stated that KNN has good performance in EEG based emotion recognition. The KNN classifies based on neighbors with the closest distance. Generally the distance calculation in KNN is based on the Euclidean distance [20]. But this distance value does not represent an important feature value. This deficiency can be overcome by K-star [21]. Therefore this study proposes the emotion classification based on K-star, also compares it with a number of other machine learning methods including KNN.

2. Research method
This study aims to recognize emotions into two categories, namely valence and arousal. Every data is classified whether the label is in the form of "high" or "low" for each of these categories. The classification based on data from EEG. The methodological flow to achieve these objectives is illustrated by Fig. 1.

![Figure 1. The methodological flow of EEG based emotion recognition.](image-url)
2.1. Data collecting
The EEG signal in this research came from the DEAP dataset. This dataset consists of 32 files, each of which contains 40 trial EEG records [22]. The data consists of 32 EEG data channels and 8 other psychological channel recordings. This 32 channel are from the recording of 32 electrodes installed under the 20-10 international system. During the recording process, each respondent was stimulated with several music videos and asked to do a self-assessment of the level of valence and arousal.

But this study only used data derived from EEG signals. The signal was a signal that has undergone the pre-processed data. The noise and artefact on the data has also been removed. Then the signal went through the sample process so that the sampling frequency became 128 Hz. The dataset also contains ground truth which contains values from valence and arousal. Both of these values have ranges from 1 to 9.

2.2. Preprocessing
At this stage, the signal was decomposed into three groups of brain wave frequencies namely alpha, beta and gamma. Delta and theta were not used because both frequencies occur when humans experience unconsciousness, so that it was not in accordance with the process of emotion recognition in this research. The decomposition process was carried out through band pass filtering. After the filtering process, the dataset consists of 96 features originating from three frequencies from 32 channels. The labelling from this research was done by dividing each value from valence and arousal into two classes. If the value was >= 0.5, the class was "high", and if <5, the class was "low".

2.3. Principal component analysis (PCA)
In statistics, PCA is a technique used to simplify data. It transforms linear data to form a new coordinate system with maximum variance [23]. It can be used to reduce the dimensions of a data without significantly reducing the characteristics of the data. This method changes most of the original variables that correlate with each other into a set of new variables that are smaller and mutually independent (no longer correlated) [24]. So the research implemented PCA to reduce data.

2.4. Resampling
The distribution data for each label is un-balanced. The labels of "high" and "low" are 2: 1. If this was ignored, the system tend to classify the data as "high". The resampling in this study was useful for balancing the amount of data between classes [25]. The process was done by randomly selecting data and homogenizing the number of each label so that the ratio between the two labels was 1: 1.

2.5. Machine learning
Machine learning is the retrieval of knowledge from a data. Its application in recent years has developed everywhere in daily life [26]. The method automates the decision making process by generalizing from known examples. Because of its success in many applications, this research applied it to EEG based emotion recognition. The classification used a number of methods of machine learning, namely naive bayes, bayesian networks, decision trees, K-nearest neighbors (KN), K-stars, artificial neural networks (ANN), support vector machines (SVM), and ensemble methods.

Naive Bayes is a simple opportunity classification based on the application of the Bayes theorem with the assumption that the explanatory variables are independent [27]. In this case, it is assumed that the presence or absence of a particular event from a group is not related to the presence or absence of other events. If the variables are dependent, then the graph is according to the Bayesian networks rule.

The decision tree chosen for this research was C45 with a batch size of 100. The decision tree is one of the most popular classification methods, because it is easy to interpret by humans. The model predicts using tree structures or hierarchical structures. The concept of a decision tree is to convert data into trees and their decision rules.

K-nearest neighbors (KNN) and K-star are methods based on instance based classifier. The decision making on KNN is based on the majority value of the number of k-points closest to the data to be classified. In other words, a test data will be classified based on training data a number of points
desired. This method was tested using from one to five neighbors. The difference between KNN and K-star lies in the mechanism of a neighbor's search. KNN is based on distance, while K-star is based on entropy. Generally KNN determines the distance between the closest points that calculated based on Euclidean distance. The accuracy of the KNN algorithm, one of which is influenced by the presence or absence of features that are irrelevant or too influential value during the classification process. The deficiency in KNN can be overcome by K-star [28].

ANN is an information processing technique or approach inspired by the workings of the biological neural system, especially in human brain cells in processing information [29]. The neural networks consist of a large number of information processing elements (neurons) that are interconnected and work together to solve a particular problem. The neural network works can be analogous to humans learning by using examples. It is called supervised learning. A neural network is configured for certain applications, such as pattern recognition or data classification, and then refined through the learning process. The learning process that occurs in biological systems involves adjusting the synaptic connections that exist between neurons. The synaptic connections between neurons are done by adjusting the weight values that exist in each connectivity both from input, neuron, and output. ANN in this study was implemented in two algorithms namely multilayer percepts (MLP), radial basic function (RBF) networks. The training of the RBF method is almost like MLP, but the difference is the use of Gaussian matrix calculations on radial functions in the hidden layer of RBF, while MLP uses sigmoid functions [30].

Ensemble methods are also implemented in this research. This method is presented in three algorithms namely bagging, boosting, logitboost. Bagging is a combination of machine learning algorithms designed to improve the stability and accuracy of machine learning algorithms [31]. Boosting is a learning ensemble meta algorithm method to reduce bias, and also variance. Unlike the case with bagging which gets predictions from the bootstrap process, boosting refers to a collection of algorithms that can convert weak learners to strong learners. Every time the tree is made, the data used remains as before but has a different distribution of weights in each iteration. The use of weights is also carried out during the process of combining the final predictions of many trees produced through classification. Logitboost is a development of logistics. The way the logitboost works in the initial stage is the same as the logistic method. In the logitboost algorithm, the predictive maximization is performed using binomial probability. Unlike logistics which are only capable of handling binary classes, logitboost has been able to handle multiclass.

2.6. Evaluation
The evaluation started from the division of datasets with the 10-fold cross validation method, which is dividing the dataset into two segments. The first segment was used as training data and the second segment was used as testing data to validate the model. Furthermore, the evaluation stages were applied using accuracy, precision, recall, and F-measure to measure classification performance.

3. Results and analysis
The classification carried out in this study focused on two main targets, namely arousal and valence. The performance of two classification targets was explained in Sections 3.1 and 3.2.

3.1. Arousal classification
Table 1 shows naïve Bayes had the lowest performance. It shows that between features had a correlation with each other. If the correlation was still taken into account, the performance would be better. It was shown by Bayesian networks. The table also shows that there were only two methods that have succeeded in achieving accuracy above 80%, namely KNN and K-star. Both are instance based classifiers. It shows the neighbors from the data made the biggest contribution to predict a class from the data. But the addition of neighbors in KNN caused a decrease in performance. The more neighbors caused the more predictive data bias. Moreover, the search for neighbors was only based on the closest distance. If the search for neighbors was not distance, for example by entropy, the performance was getting better. The K-star in Table 1 had a higher performance than KNN.
Table 1. Performance of machine learning methods against arousal classification.

| Methods                  | %     | Accuracy | Precision | Recall | F-measure |
|--------------------------|-------|----------|-----------|--------|-----------|
| Naïve Bayes              | 51.2  | 51.3     | 51.2      | 49.8   |
| Bayesian networks        | 54.8  | 58.7     | 54.8      | 49.0   |
| Decision tree            | 72.9  | 72.9     | 72.9      | 72.9   |
| KNN (K=1)                | 80.6  | 80.7     | 80.6      | 80.6   |
| KNN (K=2)                | 71.1  | 71.3     | 71.1      | 71.0   |
| KNN (K=3)                | 66.6  | 66.6     | 66.6      | 66.6   |
| KNN (K=4)                | 63.4  | 63.6     | 63.4      | 63.2   |
| KNN (K=5)                | 64.1  | 64.1     | 64.1      | 64.1   |
| K-star                   | 81.2  | 81.2     | 81.2      | 81.2   |
| Multi-layer perceptron   | 61.8  | 63.0     | 61.8      | 60.9   |
| Radial basis function networks | 53.4 | 53.5     | 53.4      | 52.9   |
| Support vector machine   | 53.6  | 53.7     | 53.6      | 53.1   |
| Bagging                  | 77.3  | 77.3     | 77.3      | 77.3   |
| Boosting                 | 55.5  | 57.0     | 55.5      | 53.1   |
| Logitboost               | 58.1  | 58.1     | 58.1      | 58.1   |

For all methods of ensemble methods, only bagging had an accuracy above 70%. The boosting and logitboost only had the accuracy below 70%. The bagging is a method that had a performance above 77%. Its performance was highest after KNN and K-star. After bagging, the highest method was the decision tree. On the other hand, decision trees had the higher performance than RBF and SVM. It showed that more complex systems did not always produce higher performance.

3.2. Valence classification

Based on Table 2, the Bayesian networks had the lowest performance. Its performance was lower than Naïve Bayes. It shows that the correlation between features in this classification was low. The decision tree also showed a decrease in performance compared with Table 1. In Table 2, the decision tree's accuracy was only 54.8%. Even though the bagging was also formed from a tree, its performance remained above 70% despite a slight decrease. This method had the highest performance after KNN and K-star. The performance above 80% was still achieved by KNN and K-star. However, similar to Table 1, the KNN had decreased if the number of neighbors was multiplied. This failure could be solved better by K-star. In this table, K-star had the highest performance compared to all methods of machine learning.

Table 2. Performance of machine learning methods against valence classification.

| Methods                  | %     | Accuracy | Precision | Recall | F-measure |
|--------------------------|-------|----------|-----------|--------|-----------|
| Naïve Bayes              | 52.5  | 53.0     | 52.5      | 50.3   |
| Bayesian networks        | 50.0  | 50.0     | 50.0      | 50.0   |
| Decision tree            | 54.8  | 55.4     | 54.8      | 53.7   |
| KNN (K=1)                | 81.6  | 81.9     | 81.6      | 81.6   |
| KNN (K=2)                | 67.9  | 68.0     | 67.9      | 67.9   |
| KNN (K=3)                | 65.8  | 65.8     | 65.8      | 65.7   |
| KNN (K=4)                | 64.1  | 64.1     | 64.1      | 64.0   |
KNN (K=5) & 62.7 & 62.7 & 62.7 & 62.6 \\
K-star & 82.6 & 82.9 & 82.6 & 82.5 \\
Multilayer perceptron & 62.5 & 62.8 & 62.5 & 62.3 \\
Radial basis function networks & 54.1 & 54.3 & 54.1 & 53.5 \\
Support vector machine & 54.1 & 54.1 & 54.1 & 54.1 \\
Bagging & 75.5 & 75.5 & 75.5 & 75.4 \\
Boosting & 51.1 & 51.2 & 51.1 & 50.3 \\
Logitboost & 54.1 & 54.2 & 54.1 & 54.0 \\

4. Conclusion

The process of emotion classification from EEG wave brain signals was done by using the DEAP dataset. Furthermore, the signals were filtered to produce three features. The features of 32 channels were processed using PCA. Through the implementation of machine learning methods, it can be seen that K-star and KNN were the methods that had the highest performance. It shows that instance based classifiers were the right method for classification in this research. In addition, K-stars always outperformed KNN in these two classifications (arousal and valence). It shows that the search for neighbors based on entropy was better than distance.

Both types of classification had different characteristics. In arousal classification, the correlation between features helped the classification process to produce better performance. But the opposite condition occurs in the valence classification. It was indicated by the performance of Bayesian networks in Table 1, better than naïve Bayes. Table 2 shows naïve Bayes better than Bayesian networks. The decline in performance was also seen from the decision tree and bagging which experienced a decrease in Table 2. Although the bagging decreased in Table 2, the performance was still above 75%. It shows that the ensemble method was very instrumental in forming a model for classification in this study.

References

[1] Reeve J (2014) Understanding Motivation and Emotion John Wiley & Sons
[2] Moksnes UK, Espnes GA and Haugan G 2014 Stress, Sense of Coherence And Emotional Symptoms in Adolescents Psychology & Health 291 32-49
[3] Tice DM, Bratslavsky E and Baumeister RF 2018 Emotional Distress Regulation Takes Precedence Over Impulse Control: If You Feel Bad, Do It Self-Regulation and Self-Control, 275-306
[4] Joseph DL 2015 Why Does Self-Reported Emotional Intelligence Predict Job Performance? A Meta-Analytic Investigation of Mixed EI Journal of Applied Psychology 1002 298
[5] Barros P, Weber CW and Wermter S 2015 Emotional Expression Recognition With A Cross-Channel Convolutional Neural Network For Human-Robot Interaction IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)
[6] Esposito A, Esposito AM and Vogel C 2015 Needs and Challenges in Human Computer Interaction For Processing Social Emotional Information Pattern Recognition Letters 61 41-51
[7] Gunawan TS 2018 A Review on Emotion Recognition Algorithms Using Speech Analysis Indonesian Journal of Electrical Engineering and Informatics (IJEI) 61 12-20
[8] Behere RV 2015 Facial Emotion Recognition Deficits: The New Face of Schizophrenia Indian journal of psychiatry 57(3) 229
[9] Laukka PE, Thingujam NS, Rockstuhl T, Iraki FK, Chui W and Althoff J 2016 The Expression And Recognition of Emotions in The Voice Across Five Nations: A Lens Model Analysis Based on Acoustic Features Journal of personality and social psychology 111(5) 686
[10] Atkinson J and Campos D 2016 Improving BCI-Based Emotion Recognition by Combining EEG Feature Selection and Kernel Classifiers Expert Systems with Applications 4 35-41
[11] Faust O, Acharya UR, Adeli H and Adeli A 2015 Wavelet-based EEG Processing For Computer-Aided Seizure Detection and Epilepsy Diagnosis *Seizure* 26 56-64
[12] Zheng WL and Lu BL 2015 Investigating Critical Frequency Bands and Channels For EEG-Based Emotion Recognition With Deep Neural Networks *IEEE Transactions on Autonomous Mental Development* 7(3) 162-175
[13] Yulita IN, Fanany MI and Arymurthy AM 2018 Fast Convolutional Method for Automatic Sleep Stage Classification *Healthcare informatics research* 24(3) 170-178
[14] Supriya S, Siuly S and Zhang Y 2016 Automatic Epilepsy Detection From EEG Introducing A New Edge Weight Method In The Complex Network *Electronics Letters* 52(17) 1430-1432
[15] Alhagry S, Fahmy AA and El-Khoribi RA 2017 Emotion Recognition based on EEG using LSTM Recurrent Neural Network *Emotion* 8(10)
[16] Lee G, Kwon M, Sri SK and Lee M 2017 Emotion Recognition based on 3D fuzzy Visual and EEG Features in Movie Clips *Neurocomputing* 144 560-568
[17] Hu B, Li X, Sun S and Ratcliffe M 2018 Attention Recognition in EEG-Based Affective Learning Research Using CFS+ KNN Algorithm *IEEE/ACM transactions on computational biology and bioinformatics* 15(1) 38-45
[18] Vijayan AE, Sen D and Sudheer AP 2015 EEG-based Emotion Recognition Using Statistical Measures and Auto-Regressive Modeling *IEEE International Conference on Computational Intelligence & Communication Technology (CICT)* 587-591
[19] Zhang YD, Yang ZJ, Lu HM, Zhou XX, Phillips P, Liu QM and Wang SH 2016 Facial Emotion Recognition Based on Biorhogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation *IEEE Access* 4 8375-838
[20] Adeniyi DA, Wei Z and Yongquan Y 2016 Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method *Applied Computing and Informatics* 12(1) 90-108
[21] Kapur B, Ahluwalia N and Sathyaraj R 2017 Comparative Study on Marks Prediction using Data Mining and Classification Algorithms *International Journal of Advanced Research in Computer Science* 8(3)
[22] Koelstra SE 2012 DEAP: A Database For Emotion Analysis; Using Physiological Signals *IEEE Transactions on Affective Computing* 3(1) 18-31
[23] Pyatykh S, Hesser J and Zheng L 2013 Image Noise Level Estimation By Principal Component Analysis *IEEE transactions on image processing* 22(2) 687-699
[24] Shaharudin SM, Ahmad N, Zainuddin NH and Mohamed NS 2018 Identification of Rainfall Patterns on Hydrological Simulation Using Robust Principal Component Analysis *Indonesian Journal of Electrical Engineering and Computer Science* 11(3)
[25] Marqués AI, García V and Sánchez J S 2013 On The Suitability of Resampling Techniques For The Class Imbalance Problem in Credit Scoring *Journal of the Operational Research Society* 64(7) 1060-1070
[26] Sabri N, Ibrahim Z and Isa D 2018 Evaluation of Color Models for Palm Oil Fresh Fruit Bunch Ripeness Classification *Indonesian Journal of Electrical Engineering and Computer Science* 11(2) 549-557
[27] Lawend HO, Muad AM and Hussain A 2018 An Improved Flexible Partial Histogram Bayes Learning Algorithm *Indonesian Journal of Electrical Engineering and Computer Science* 11(3) 975-986
[28] Williams TP and Gong J 2014 Predicting Construction Cost Overruns Using Text Mining, Numerical Data and Ensemble Classifiers *Automation in Construction* 43 23-29
[29] Zhang Z 2018 Artificial Neural network *Multivariate Time Series Analysis in Climate and Environmental Research* 1-35
[30] Yu H, Reiner PD, Xie T, Bartczak T and Wilamowski BM 2014 An Incremental Design of Radial Basis Function Networks *IEEE transactions on neural networks and learning systems* 25(10) 1793-1803
[31] Bui DT, Ho TC, Pradhan B, Pham BT, Nhu VH and Revhaug I 2016 GIS-based Modeling of Rainfall-Induced Landslides Using Data Mining-Based Functional Trees Classifier With
Adaboost, Bagging, And Multiboost Ensemble Frameworks Environmental Earth Sciences 75(14)

**Acknowledgments**

This work supported by Padjadjaran University Internal Grant funded by Directorate of Research and Service Community of Padjadjaran University.