DOI: https://doi.org/10.24297/ijct.v24i.9650

Positive-Unlabeled Learning Method for Positive Emotion Recognition Using EEG technology

Zizhu Li[1], *Chengyuan Shen*[1], *and Jianting Cao*[1,2]

[1]Graduate School of Engineering, Saitama Institute of Technology, Fukaya 369-0217, Japan

[2]RIKEN Center for Advanced Intelligence Project(AIP), 103-0027, Japan

Email addresses {i3003lqa, f3007bua, cao}@sit.ac.jp

Abstract

Emotion is a reaction of the human brain to external events, and the study of emotion recognition has substantial practical applications. Therefore, accurately recognizing and understanding positive emotions across different populations is crucial. Traditional image recognition technology cannot effectively identify emotions in individuals with impaired facial muscle control, such as elderly people in nursing homes with Alzheimer's disease and patients with facial nerve paralysis (Bell's palsy). Consequently, many machine learning methods have been widely applied to emotion recognition based on electroencephalogram (EEG) signals in recent years. In cases where the number of samples is sufficient, powerful deep learning methods can achieve high performance in emotion recognition. However, obtaining a large amount of reliably labeled emotional EEG data is arduous. We introduce a Positive-Unlabeled (PU) learning method for classifying EEG signals into Positive and Non-Positive emotions using a binary classifier developed with minimal labeled data. This approach utilizes a small volume of labeled data containing only positive emotion signals, combined with unlabeled data that includes both classes, effectively reducing the dependency on extensive, reliably labeled EEG data. The best accuracy achieved by this method is 93.95%. Experimental results on the dataset demonstrate the effectiveness of our approach.

Keywords: EEG, PU learning, emotion recognition

Introduction

Emotion represents a complex reaction of the human brain to external events, and understanding this reaction has substantial practical applications across various fields. The study of emotion recognition has garnered considerable interest due to its potential to enhance human-computer interaction, improve mental health assessments, and offer better support for individuals with communication impairments. Positive emotions have a profound impact on human physical and mental health. They can enhance psychological well-being, reduce depression and anxiety, boost immune function, and improve overall physical health. Therefore, accurately recognizing and understanding positive emotions in different populations is crucial(Alexander et al., 2021).

Traditional image recognition technologies are commonly used to detect human emotions. These systems analyze facial muscle movements to infer emotional states. However, they face significant limitations, particularly for individuals with impaired facial muscle control, such as elderly residents in nursing homes, Alzheimer's patients, and individuals with facial nerve paralysis (Bell's palsy)(Tamata & Mohammadnezhad, 2023). These conditions hinder the accurate transmission of emotional signals through facial muscles, rendering facial expression-based systems ineffective for these populations(Kumari et al., 2015).

To overcome these limitations, EEG signals have been increasingly utilized for emotion recognition. EEG provides a direct measure of brain activity and is less susceptible to the limitations faced by facial expression recognition systems. Consequently, many machine learning methods have been developed to analyze EEG signals for emotion recognition, resulting in significant advancements in the field(Pan, Zheng, et al., 2021)(Özerdem & Polat, 2017)(Abdulrahman et al., 2022).

When sufficient samples are available, deep learning methods have demonstrated high performance in emotion recognition tasks by learning complex patterns and features from the data, resulting in accurate emotion classification. However, acquiring a large amount of reliably labeled emotional EEG data is challenging and resource-intensive. Labeling EEG data with emotional states requires expert knowledge and extensive manual effort, limiting the availability of large, high-quality datasets. This challenge can be mitigated by using Positive-Unlabeled (PU) learning. PU learning is a semi-supervised approach that uses a small amount of positively labeled data from one class and a larger set of unlabeled data containing both positive and non-positive emotions. This method allows us to train a binary classifier that can distinguish the types of emotions we want to recognize (such as positive emotions) with significantly fewer labeled samples(Alhalaseh & Alasasfeh, 2020)(Sohaib et al., 2013)(Zhong et al., 2020)(Wang et al., 2018).

In this paper, we propose a novel approach for the classification of Positive/Non-Positive EEG signals. Our method, when applied for classification, has demonstrated superior performance compared to standard Random Forests (RF). Our main contribution is the formulation of the positive emotion recognition problem under the PU learning framework, which has significantly reduced the necessary number of labeled training samples. This method overcomes the ineffectiveness of traditional image recognition technologies. Moreover, we have extensively classified the EEG data of multiple individuals, segmented the EEG signals, and extracted differential entropy (DE) features from each segment. Then, we performed emotion classification through machine learning. PU learning has demonstrated higher accuracy and superior robustness compared to the commonly used RF algorithm. We have also applied the same data collection methods to individual classification and achieved high accuracy in recognizing positive emotions. This indicates that PU learning performs well in recognizing positive emotion within the dataset and achieves similar performance in individual cases.

Methodology

PU Learning for Classification In the PU learning model, only a small portion of the data is labeled, belonging to one specific class, while the other data remain unlabeled. unlike supervised methods, PU learning requires only a small number of labeled samples along with unlabeled samples. Unlike semi-supervised learning, PU learning does not necessitate labeled data for both classes.

More specifically, PU learning is defined as learning from positive and unlabeled data, which can be regarded as a two-class (positive and negative) classification method. At present, according to the strategy how to deal with unlabeled data, PU learning can be divided into two categories. One category is to find reliable Positive data in unlabeled data(Blum & Mitchell, 1998)(Liu et al., 2002)(Li & Liu, 2003)(Zhao et al., 2018), based on which standard supervised learning methods for binary classifications can be applied. Another category is to treat the unlabeled data as Positive data and give a suitable weight for unlabeled data(Liu et al., 2003)(Elkan & Noto, 2008)(Lee & Liu, 2003).

Let x be the input feature vector calculated from one segment of EEG, $y \in \{\pm 1\}$ be the class label, i.e., $+1$ denotes Positive and -1 denotes Non-Positive signal. The class conditional distributions for Positive signals, denoted by $p_p(x)$, and Non-Positive signals, denoted by $p_n(x)$, are defined by

$$\begin{aligned} p_p(x) &= p(x \,|\, y = +1), \\ p_n(x) &= p(x \,|\, y = -1). \end{aligned} \tag{1}$$

The prior probabilities for each class are denoted by $\pi_p = p(y = +1)$ and $\pi_n = p(y = -1)$, respectively. Thus, it is obvious that $\pi_n = 1 - \pi_p$. In PU learning method, π_p is assumed known in advance. By applying the Bayesian rules, the marginal distribution of unlabeled data (i.e., the distribution of both two classes of data), denoted by $p(x)$, can be written as

$$p(x) = \pi_p p_p(x) + \pi_n p_n(x). \tag{2}$$

To solve the PU problem, the most widely used and successful objective function is the empirical unbiased risk estimator that is proposed by (M. C. Du Plessis et al., 2014) (M. Du Plessis et al., 2015) (Kiryo et al., 2017), which is

$$\hat{R}_{pn}(g) = \pi_p \hat{R}_p^+(g) + \pi_n \hat{R}_n^-(g), \tag{3}$$

where $\hat{R}_p^+(g)$ and $\hat{R}_n^-(g)$ denote the empirical risks for Positive and Non-Positive data, respectively. The $g(\cdot)$ denotes the binary classification function, and $l(g(x), \pm 1)$ denotes the loss function. Thus, the empirical risk for positive signal, i.e., $\hat{R}^+(g)$ in (3), can be calculated by

$$\hat{R}_p^+(g) = \mathbb{E}_{x \sim p_p(x)} l(g(x), +1), \tag{4}$$

and the empirical risk for Non-Positive signal, i.e., $\hat{R}_n^-(g)$, can be calculated by

$$\hat{R}_n^-(g) = \mathbb{E}_{x \sim p_p(x)} l(g(x), -1). \tag{5}$$

Since the distribution of Non-Positive signals is unknown due to the labels being only given for Positive signals, $\hat{R}_n^-(g)$ in (5) cannot be computed straightforwardly. As we can see from (2), the distribution of Non-Positive signals can be represented by using

$$\pi_n p_n(x) = p(x) - \pi_p p_p(x). \tag{6}$$

Hence, the empirical risk for Non-Positive samples can be computed by

$$\pi_n \hat{R}_n^-(g) = \hat{R}_u^-(g) - \pi_p \hat{R}_p^-(g), \tag{7}$$

where $\hat{R}_u^-(g)$ and $\hat{R}_p^-(g)$ are the empirical risks under the distribution of unlabeled data and Positive data, respectively, which are defined by

$$\begin{aligned} \hat{R}_u^-(g) &= \mathbb{E}_{x \sim p(x)} l(g(x), -1), \\ \hat{R}_p^-(g) &= \mathbb{E}_{x \sim p_p(x)} l(g(x), -1). \end{aligned} \tag{8}$$

Finally, the risk estimator in (3) can be approximated indirectly by

$$\hat{R}_{pu}(g) = \pi_p \hat{R}_p^+(g) + \hat{R}_u^-(g) - \pi_p \hat{R}_p^-(g). \tag{9}$$

In general, $g(x)$ can be any classifier functions, such as linear discriminative analysis or support vector machines. Given the strong performance of Random forests across various datasets, in this study, we employ Random Forest(RF) as the binary classifier function $g(x)$. Therefore, based on the objective function shown in (9), we utilize the unique training methods of RF to address the PU problem, without the need for the traditional backpropagation (BP) algorithm. The PU learning model is shown in Fig. 1.

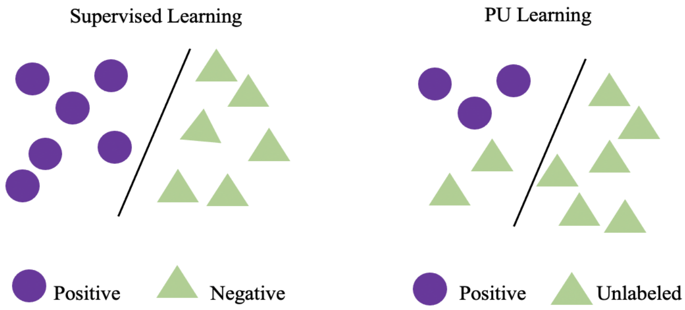

Fig. 1 Illustration of PU learning.

Dataset and Feature extraction In our study, we utilized the SJTU Emotion EEG Dataset (SEED). A detailed description of this dataset is available at the provided link (https://bcmi.sjtu.edu.cn/home/seed/). All signals were filtered using a band-pass filter with a range of 0.05 to 50 Hz. The dataset includes data recordings from 15 subjects, from which we randomly selected 12 for our analysis. The dataset was recorded using 62 channels. Since the lateral temporal area has a relatively high level of activation for positive emotions, we chose the 12 channels in this area for our analysis. Previous studies have consistently concluded that selecting specific channels performs better than using all original channels (Zheng & Lu, 2015). Therefore, we selected these 12 specific channels for our analysis. Each EEG test lasted for 2010 seconds, and the sampling rate for the preprocessed EEG data was 200 Hz. An example of a preprocessed EEG signal is visualized in Fig. 2.

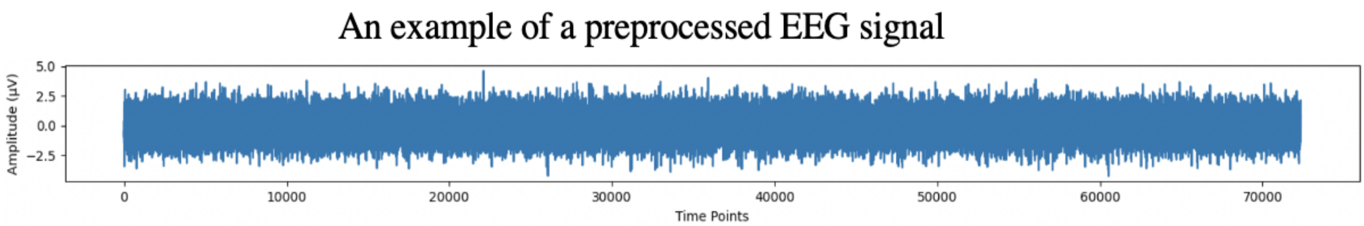

Fig. 2 An example of a preprocessed EEG signal.

Summary and Workflow Each one-second segment of the EEG is treated as a sample, allowing us to collect data samples for training machine learning methods. We use a band-pass filter to process the EEG segments and then compute the DE and other features for each filtered segment, which serve as the feature representation for each EEG segment. We found that differential entropy achieved the highest accuracy through comparison, and we will present the comparative results in the experimental section. Finally, based on partially labeled positive emotion segments, a PU learning classifier is used to classify Positive and Non-positive emotions. Our method workflow is illustrated in Fig. 3.

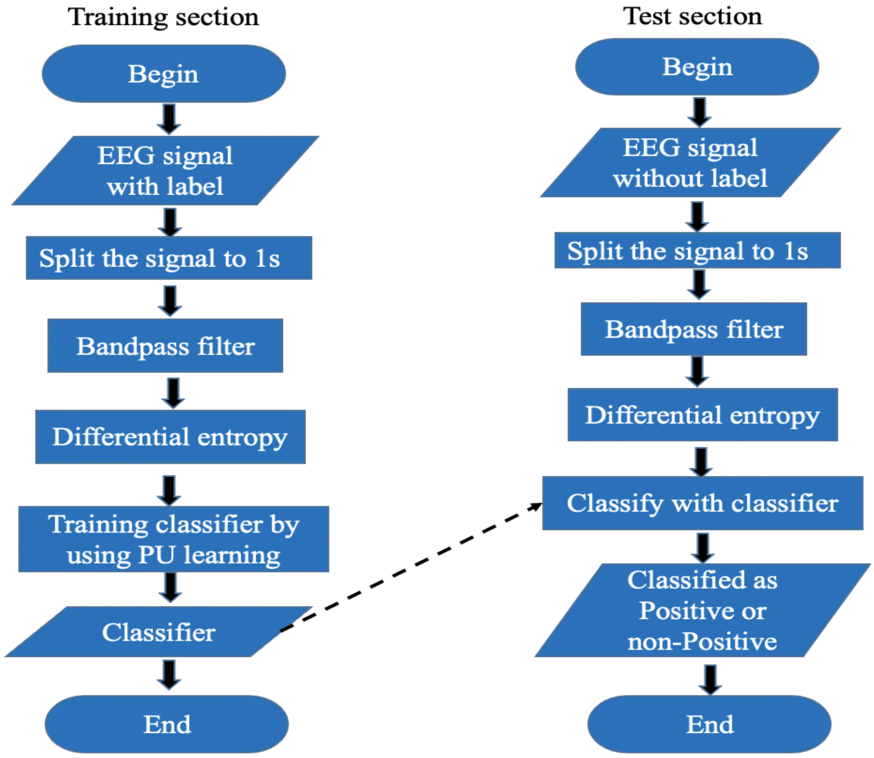

Fig. 3 Flowchart of Positive and non-Positive emotional signal classification.

Experimental Results

In this experiment, we used the SEED dataset and our own recorded data to conduct experiments with PU learning and RF. Our own recorded data refer to the SEED dataset (https://bcmi.sjtu.edu.cn/home/seed/). The rest of the time is recorded in the video, etc.

The dataset contains 72,360 samples. The dataset was processed using a band-pass filter between 0.05 to 50 Hz. In PU learning, 80% of the data samples were randomly chosen as training data and 20% as test data. Positive samples were randomly selected, with labels indicating positive emotions as labeled data, while the rest of the samples were considered unlabeled data. The parameters of the random model were manually chosen. Each experiment began with the conversion of 0 malignant examples to benign, incrementing by 100 malignant examples for each subsequent trial. This process was iteratively conducted across a series of experiments, each consisting of 31 iterations to ensure the robustness and reliability of the outcomes. This methodological approach provided insights into the impact of varying degrees of malignant-to-benign transformations on experimental outcomes. We used PU learning to classify 5 iterations, as shown in Table 1. RF improved model generalization by constructing multiple decision trees and combining their predictions. Its advantages included handling high-dimensional EEG data, reducing overfitting, providing strong noise resistance, capturing nonlinear relationships, and evaluating feature importance. These characteristics made RF highly effective in EEG-based emotion classification. Therefore, we used it to compare with PU learning. When applying the RF, 80% of the data samples were randomly chosen as training data and 20% as test data. We compared the results under the same conditions. The RF used the same set of evaluation criteria.

Table 1.PU learning classify 5 iterations.

iterations	1	2	3	4	5
Accuracy(%)	86.01	87.78	88.92	91.55	90.74

A comparative analysis between PU learning and RF methodologies was conducted across three critical metrics: Precision, Recall, and F1 score. These metrics were evaluated in relation to the number of malignant samples reclassified as benign, demonstrating the dynamic performance of both learning strategies under various degrees of data transformation.

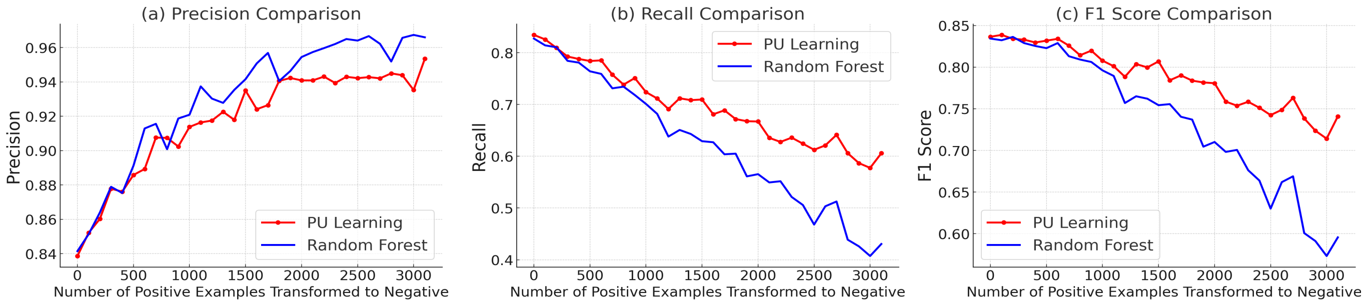

Fig. 4 Comparison of (a)Precision, (b)Recall, and (c)F1 Score between PU Learning and RF for the SEED Dataset.

Regarding Precision in Fig. 4(a), PU learning exhibits slight variations but generally remains slightly inferior to RF, which maintains a higher and more stable precision as the dataset changes. This observation suggests that while PU learning accurately classifies relevant instances, it still lags behind RF in precision as the dataset composition changes. In terms of Recall in Fig. 4(b), PU learning starts with a notably high rate, which, despite a gradual decrease, remains significantly higher than that of RF. The latter shows a more pronounced decline throughout the data transformation process. This differential highlights PU learning's superior capability in consistently identifying all pertinent instances, thereby underscoring its effectiveness across various stages of data modification. A comparative overview of the F1 Score in Fig. 4(c) reveals that PU learning consistently outperforms RF. Notably, as the number of positive examples converted to negative increases, the F1 score of RF exhibits a significant decline, whereas PU learning maintains relative stability. This trend underscores the robustness of PU learning in adapting to changes within the dataset, maintaining its precision in identifying relevant instances despite data transformations. The average accuracy of PU learning is 91.77%, while the average accuracy of RF is 93.06%.

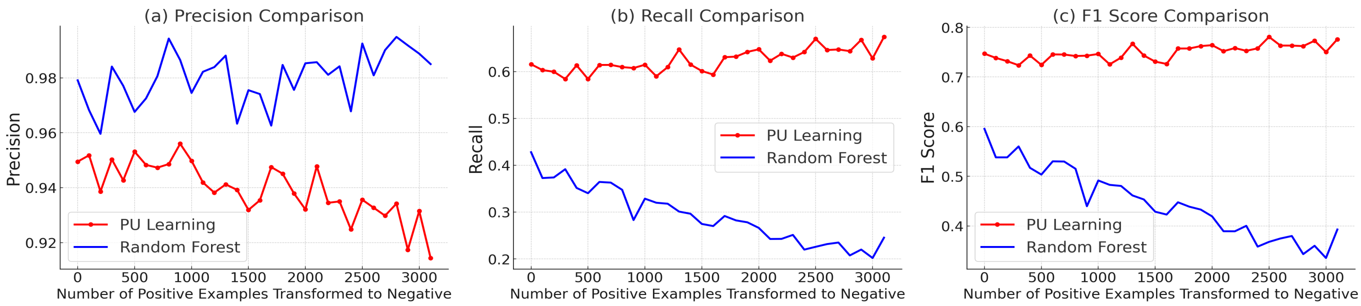

Fig. 5 Comparison of (a)Precision, (b)Recall, and (c)F1 Score between PU Learning and RF for the recorded data.

Regarding Precision in Fig. 5(a), PU learning remains inferior to RF, with RF maintaining higher and more stable precision as the dataset transforms. This observation suggests that, although PU learning can accurately classify relevant instances, its precision still lags behind that of RF amid changes in dataset composition. In terms of Recall in Fig. 5(b), both methods exhibit significant declines. PU learning starts with a notably higher rate, which, despite gradually decreasing, remains substantially higher than that observed in RF. RF, on the other hand, shows a more pronounced decline throughout the data transformation process. This differential highlights PU learning's superior capability in consistently identifying all pertinent instances, underscoring its effectiveness across various stages of data modification. A comparative overview of the F1 Score in Fig. 5(c) reveals that PU learning consistently outperforms RF in terms of performance. Notably, as the number of positive examples converted to negative increases, the F1 score of RF exhibits a significant decline, whereas PU learning maintains relative stability. This trend underscores the resilience of PU learning in adapting to changes within the dataset, maintaining its precision in identifying relevant instances despite data transformations. The average accuracy of PU learning is 93.95%, while the average accuracy of RF is 98.01%. The conclusions drawn from both experiments are consistent.

In summary, PU learning demonstrates superior performance to RF across all evaluated metrics except for accuracy, which is slightly lower. This resilience is particularly noteworthy given the experimental results, where PU learning incurs minimal performance loss. This slight decrease in accuracy is offset by a notable reduction in the labeling workload, showcasing the efficiency and applicability of PU learning in practical settings.

Conclusions and Future Work

In this paper, we have proposed a novel approach for the classification of Positive/Non-Positive EEG signals. Our main contribution lies in formulating the positive emotion recognition problem under the PU learning framework, significantly reducing the required number of labeled training samples. This method overcomes the limitations of traditional facial expression recognition tools, making it more practical than popular supervised learning methods.

We extensively classified the EEG data of multiple individuals, segmented the EEG signals, and extracted DE features from each segment. Then, we performed emotion classification using machine learning. In addition to using the SEED dataset, we also identified positive emotions in individually recorded brainwaves, improved the accuracy. This demonstrated that PU learning performs well in recognizing positive emotions in datasets and achieves high recognition accuracy at the individual level.

Although the accuracy of the PU learning method is slightly lower than that of supervised learning, the reduction in the labeling workload and the ability to handle real-world data without extensive labeling have made it a viable and efficient approach. Future work will focus on improving the accuracy of the PU learning method to approach that of supervised learning as closely as possible, further enhancing its practical applicability and performance.

Acknowledgments

This work was supported by JSPS KAKENHI 20H04249.

Conflicts of Interest

The authors have no conflict of interest about anything in this article.

References

Abdulrahman, A., Baykara, M., & Alakus, T. B. (2022). A novel approach for emotion recognition based on eeg signal using deep learning. *Applied Sciences*, *12*(19), 10028. https://doi.org/https://doi.org/10.3390/app121910028

Alexander, R., Aragón, O. R., Bookwala, J., Cherbuin, N., Gatt, J. M., Kahrilas, I. J., Kästner, N., Lawrence, A., Lowe, L., Morrison, R. G., et al. (2021). The neuroscience of positive emotions and affect: Implications for cultivating happiness and wellbeing. *Neuroscience & Biobehavioral Reviews*, *121*, 220–249. https://doi.org/https://doi.org/10.1016/j.neubiorev.2020.12.002

Alhalaseh, R., & Alasasfeh, S. (2020). Machine-learning-based emotion recognition system using eeg signals. *Computers*, *9*(4), 95. https://doi.org/https://10.3390/computers9040095

Blum, A., & Mitchell, T. (1998). Combining labeled and unlabeled data with co-training. *Proceedings of the eleventh annual conference on Computational learning theory*, 92–100. https://doi.org/https://doi.org/10.1145/279943.279962

Du Plessis, M., Niu, G., & Sugiyama, M. (2015). Convex formulation for learning from positive and unlabeled data. *International conference on machine learning*, 1386–1394.

Du Plessis, M. C., Niu, G., & Sugiyama, M. (2014). Analysis of learning from positive and unlabeled data. *Advances in neural information processing systems*, *27*.

Elkan, C., & Noto, K. (2008). Learning classifiers from only positive and unlabeled data. *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 213–220. https://doi.org/https://doi.org/10.1145/1401890.1401920

Kiryo, R., Niu, G., Du Plessis, M. C., & Sugiyama, M. (2017). Positive-unlabeled learning with non-negative risk estimator. *Advances in neural information processing systems*, *30*.

Kumari, J., Rajesh, R., & Pooja, K. (2015). Facial expression recognition: A survey. *Procedia computer science*, *58*, 486–491. https://doi.org/https://doi.org/10.3390/sym11101189

Lee, W. S., & Liu, B. (2003). Learning with positive and unlabeled examples using weighted logistic regression. *ICML*, *3*(2003), 448–455.

Li, X., & Liu, B. (2003). Learning to classify texts using positive and unlabeled data. *IJCAI*, *3*(2003), 587–592.

Liu, B., Dai, Y., Li, X., Lee, W. S., & Yu, P. S. (2003). Building text classifiers using positive and unlabeled examples. *Third IEEE international conference on data mining*, 179–186. https://doi.org/https://doi.org/10.1109/ICDM.2003.1250918

Liu, B., Lee, W. S., Yu, P. S., & Li, X. (2002). Partially supervised classification of text documents. *ICML*, *2*(485), 387–394.

Özerdem, M. S., & Polat, H. (2017). Emotion recognition based on eeg features in movie clips with channel selection. *Brain informatics*, *4*, 241–252. https://doi.org/https://doi.org/10.1007/s40708-017-0069-3

Pan, B., Zheng, W., et al. (2021). Emotion recognition based on eeg using generative adversarial nets and convolutional neural network. *computational and Mathematical Methods in Medicine*, *2021*. https://doi.org/https://doi.org/10.1155/2021/2520394

Sohaib, A. T., Qureshi, S., Hagelbäck, J., Hilborn, O., & Jerčić, P. (2013). Evaluating classifiers for emotion recognition using eeg. *Foundations of Augmented Cognition: 7th International Conference, AC 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013. Proceedings 7*, 492–501. https://doi.org/https://10.1007/978-3-642-39454-6_53

Tamata, A. T., & Mohammadnezhad, M. (2023). A systematic review study on the factors affecting shortage of nursing workforce in the hospitals. *Nursing open*, *10*(3), 1247–1257.

Wang, F., Zhong, S.-h., Peng, J., Jiang, J., & Liu, Y. (2018). Data augmentation for eeg-based emotion recognition with deep convolutional neural networks. *MultiMedia Modeling: 24th International Conference, MMM 2018, Bangkok, Thailand, February 5-7, 2018, Proceedings, Part II 24*, 82–93. https://doi.org/https://doi.org/10.1007/978-3-319-73600-6_8

Zhao, X., Tanaka, T., Kong, W., Zhao, Q., Cao, J., Sugano, H., & Yoshiday, N. (2018). Epileptic focus localization based on ieeg by using positive unlabeled (pu) learning. *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 493–497. https://doi.org/https://10.23919/APSIPA.2018.8659747

Zheng, W.-L., & Lu, B.-L. (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. https://doi.org/https://10.1109/TAMD.2015.2431497

Zhong, P., Wang, D., & Miao, C. (2020). Eeg-based emotion recognition using regularized graph neural networks. *IEEE Transactions on Affective Computing*, *13*(3), 1290–1301. https://doi.org/https://doi.org/10.1109/TAFFC.2020.2994159