Am J Neurodegener Dis 2024;13(4):23-33
www.AJND.us /ISSN:2165-591X/AJND0160936

Original Article

Comparative analysis of dimensionality reduction techniques for EEG-based emotional state classification

Seyed-Ali Sadegh-Zadeh[1], Nasrin Sadeghzadeh[2], Ommolbanin Soleimani[3], Saeed Shiry Ghidary[1], Sobhan Movahedi[4], Seyed-Yaser Mousavi[5]

[1]*Department of Computing, University of Staffordshire, Stoke-on-Trent, United Kingdom;* [2]*Faculty of Mathematics, University of Qom, Qom, Iran;* [3]*Department of Psychology, University of Shahab Danesh, Qom, Iran;* [4]*Islamic Azad University, Science and Research, Tehran, Iran;* [5]*Department of Psychiatry, Tehran University of Medical Sciences, Tehran, Iran*

Received October 3, 2024; Accepted October 25, 2024; Epub October 25, 2024; Published October 30, 2024

Abstract: Objectives: The aim of this study is to evaluate the impact of various dimensionality reduction methods, including principal component analysis (PCA), Laplacian score, and Chi-square feature selection, on the classification performance of an electroencephalogram (EEG) dataset. Methods: We applied dimensionality reduction techniques, including PCA, Laplacian score, and Chi-square feature selection, and assessed their impact on the classification performance of EEG data using linear regression, K-nearest neighbour (KNN), and Naive Bayes classifiers. The models were evaluated in terms of their classification accuracy and computational efficiency. Results: Our findings suggest that all dimensionality reduction strategies generally improved or maintained classification accuracy while reducing the computational load. Notably, PCA and Autofeat techniques led to increased accuracy for the models. Conclusions: The use of dimensionality reduction techniques can enhance EEG data classification by reducing computational demands without compromising accuracy. These results demonstrate the potential for these techniques to be applied in scenarios where both computational efficiency and high accuracy are desired. The code used in this study is available at https://github.com/movahedso/Emotion-analysis.

Keywords: Dimensionality reduction, electroencephalogram (EEG), feature sets, principal component analysis (PCA), Laplacian score, Chi-square feature selection, linear regression, K-nearest neighbour (KNN), Naive Bayes classifiers

Introduction

Autonomous non-invasive detection of emotional states is potentially useful in multiple domains such as human-robot interaction and mental healthcare [1]. Identifying emotional states provides several benefits, including enhancing personalization in healthcare, enabling more intuitive and engaging human-robot interactions, and offering objective assessments of mental health conditions [2]. In healthcare, this technology can help tailor treatments for emotional disorders and provide continuous monitoring of patients [3]. In human-computer interaction, emotion recognition facilitates the development of empathetic and adaptive systems, leading to improved user experiences [4]. Additionally, emotion recognition has the potential to provide non-verbal cues that are crucial in understanding and supporting individuals in a wide range of social and therapeutic contexts [5].

Identifying emotional states can provide a new level of engagement between the user and the machine and allow for the extraction of concrete information independent of spoken communication [6, 7]. Brain signals data is becoming more accessible to both the consumer industry and research mainly to the expansion of low-cost electroencephalography (EEG) equipment [8, 9]. As a result, it creates the demand for autonomous classification without the necessity for an on-site specialist.

Dimensionality reduction plays a crucial role in EEG-based emotion recognition by reducing the high dimensionality inherent in EEG signals,

https://doi.org/10.62347/ZWRY8401

thereby enhancing computational efficiency [10]. This reduction in dimensions not only mitigates the computational load but also helps in addressing the challenges posed by the 'curse of dimensionality', which can otherwise hinder model performance [11]. By extracting the most relevant features from the EEG data, dimensionality reduction enables faster and more efficient machine learning algorithms, which is essential for real-time applications. The practical application value is particularly significant in scenarios such as healthcare, mental health assessments, and human-robot interaction, where efficient processing of emotional data is vital [12]. For example, it allows the development of more responsive and adaptive systems capable of recognizing and reacting to human emotions, thus fostering more natural and effective interactions.

Real-world data, like brain signals, typically has a high degree of dimension. Its dimensionality should always be decreased in order to appropriately handle such data [13, 14]. The process of transforming highly dimensional data into a useful representation of decreased dimensionality is known as dimensionality reduction [15, 16]. The dimensionality of the reduced representation should ideally match the intrinsic dimensionality of the data [17]. The bare minimum of parameters required to explain the observable qualities of the data is known as the intrinsic dimension of the data [18, 19]. Dimensionality reduction is crucial in many fields because it reduces the negative effects of dimensionality and other high-dimensional spatial characteristics [1, 20]. Therefore, dimensionality reduction makes it easier to classify, visualise, and compress high-dimensional data, among other things [21].

EEG recordings are high-dimensional by nature since the data is spatiotemporally organised. As a result, the multiple comparison problem (MCP) frequently arises when we use statistical analysis to compare brain activities under various settings. Since most experiments have thousands of dependent variables, it is typically challenging to manage the family-wise error rate (FWER) using conventional statistical techniques that work at the level of a single comparison [22, 23]. Furthermore, pre-processing of data to classification with a raw EEG stream is highly challenging due to the high dimensionality, diversity, randomization, and non-stationary features. To find valuable statistics and simplify the model development procedure, dimension reduction can be used. This will allow for time and computing resource savings during the training and classification operations. In order to compare various techniques, this work focuses on dimension reduction of mental emotion EEG data.

In the context of human emotional analysis, dimensionality reduction methods play a crucial role by making it possible to handle high-dimensional EEG data more efficiently while preserving meaningful features that are directly linked to emotional states. The process of reducing the dimensionality helps in highlighting the most relevant signal patterns that can be used to classify different emotional categories, thus enabling the automatic recognition of emotions in real-world scenarios. Understanding which features are most relevant for distinguishing between emotional states can also help us develop more interpretable models, ultimately contributing to advancements in areas like mental health diagnostics and human-robot interaction.

This study focuses on the effectiveness of various dimensionality reduction techniques such as PCA, Laplacian score, and Chi-square feature selection, on the classification performance of an EEG dataset using linear regression, K-nearest neighbour (KNN), and Naive Bayes classifiers. The study aims to evaluate the impact of these techniques on classification performance and their effectiveness in reducing the computational load while maintaining or improving performance. The study concludes that using all dimensionality reduction strategies can generally improve or maintain classification accuracy and that PCA and Autofeat techniques can increase the accuracy of the models. The study's findings have significant implications for various fields, including finance, healthcare, and social sciences.

The five sections of this paper are arranged as follows. There is a literature review in part II. Section III focuses on the conceptual underpinnings of our research and method. In part IV, the results are discussed and assessed. In part V, we discuss and conclude the paper.

Table 1. Table to show Lövheim categories and their encapsulated emotions with a valence label

Emotion Category	Emotion/Valence
A	Shame (Negative) Humiliation (Negative)
B	Contempt (Negative) Disgust (Negative)
C	Fear (Negative) Terror (Negative)
D	Enjoyment (Positive) Joy (Positive)
E	Distress (Negative) Anguish (Negative)
F	Surprise (Negative) (Lack of Dopamine)
G	Anger (Negative) Rage (Negative)
H	Interest (Positive) Excitement (Positive)

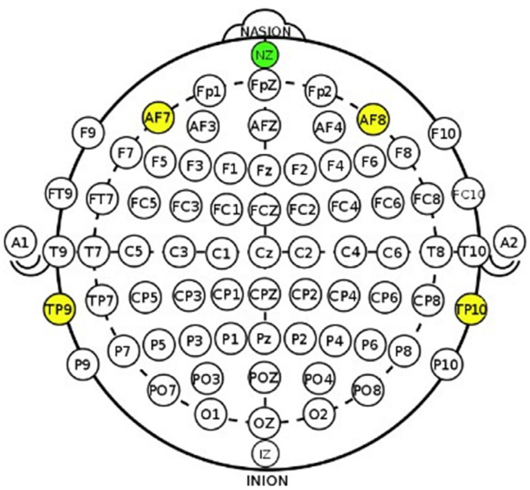

Figure 1. EEG sensors TP9, AF7, AF8 and TP10 of the Muse headband on the international standard EEG placement system [24].

Material and methods

Dataset description

This study employs the EEG Brainwave Dataset: Feeling Emotions, which is publicly available at https://www.kaggle.com/datasets/birdy654/eeg-brainwave-dataset-feeling-emotions. It is a dataset based on EEG brainwave data collected from two subjects, one male and one female, between the ages of 20-22 [24]. The dataset comprises 12 minutes of brain activity data from each subject, recorded during the viewing of six film clips listed in **Table 1**. Additionally, six minutes of neutral brainwave data were collected, resulting in a total of 36 minutes of EEG data per subject. The micro voltage measurements were recorded from TP9, AF7, AF8, and TP10 electrodes, as shown in **Figure 1**. To standardize the variable frequency, the entire dataset was transformed to a total of 324,000 data points at 150 Hz.

The emotional responses of the subjects were classified into eight categories, as outlined in **Table 2**. The categories were differentiated based on valence labels of "positive" or "negative" rather than the specific emotional levels.

In order to minimize the potential for contamination, the collection of neutral data was prioritized prior to the consideration of emotional data. This neutral data was classified as a distinct category to maintain its separation from subsequent emotional data. To further mitigate any potential interference from resting emotional states, data collection was limited to three-minute intervals daily. During the data collection process, participants were instructed to remain motionless and avoid any conscious movements that may produce Electromyographic signals and potentially confound the data. Such precautions have been documented in previous research [25] as necessary to maintain the reliability and validity of physiological data collection in emotional research.

Ethics statement

The EEG Brainwave Dataset: Feeling Emotions used in this study is publicly available at https://www.kaggle.com/datasets/birdy654/eeg-brainwave-dataset-feeling-emotions. The dataset was collected and shared by Bird et al. in accordance with the ethical guidelines of the responsible institution, ensuring participants provided informed consent for the use of their anonymized data for research purposes. Since our study used this publicly available dataset, no additional ethical approval was required.

Methodology

In the realm of real-world data analysis, high dimensionality is a common issue that arises in various fields such as healthcare, genomics, and MRI scans. To address this challenge, dimensionality reduction techniques are applied to reduce the dimensionality of these data categories and improve their manageability [11]. The importance of dimensionality reduction is underscored by its ability to alleviate the curse of dimensionality and other undesirable properties of high-dimensional spaces.

Table 2. Before and after implementing DR techniques

Algorithms	Before	Local Linear	Laplacia	Independent	Isometric	Neighbourhood	Autofeat	PCA	Chi Sq	DT	Autoencoder
	AUC	AUC	AUC	AUC	AUC	AUC	AUC	AUC	AUC	AUC	AUC
LR	50.0	97.1	91.3	95.7	95.2	98.4	99.6	99.5	98.4	97.8	99
KNN	87.7	96.6	90.3	97	97	98.7	99.6	98.1	98.3	97.8	98.7
NB	67.5	96.8	85.5	96.4	95.3	93.1	97.3	85.6	83.1	88.7	83.2
MLP	67.8	97.8	92.5	99	95.8	99	99.4	99.3	99	99.8	98.9
SVM	76.3	98.2	92.8	99.2	97.7	98.6	99.9	99.1	99.1	99.5	98.6

In this study, the use of dimensionality reduction techniques such as PCA, Laplacian score, and Chi-square feature selection is particularly focused on extracting meaningful components from EEG data, which is known to capture complex brain activities underlying human emotions. By using these methods, we can more effectively identify which components of EEG signals are most indicative of different emotional states, allowing us to interpret the data in a more manageable form and to classify emotions accurately with reduced computational overhead.

Numerous dimensionality reduction techniques have been proposed over the years, including those advanced by Matsuda et al. [26], Lespinats et al. [27], Huang et al. [15] and Yan et al. [29]. In this study, we applied ten different dimensionality reduction techniques to the aforementioned dataset, followed by the application of machine learning algorithms to evaluate the results in terms of accuracy performance metrics. The application of dimensionality reduction techniques represents a vital step in managing high-dimensional data in real-world contexts, facilitating its analysis and yielding better insights for decision-making.

For the analysis in this study, we employed several machine learning algorithms, including Linear Regression, KNN, Naive Bayes, MLP, and Support SVM. All analyses were conducted using Python, primarily leveraging the Scikit-learn library for model training, evaluation, and cross-validation. The workflow included splitting the dataset into training and testing subsets, with an 80:20 split. We applied standard scaling to normalize the features and conducted feature selection using different dimensionality reduction techniques such as PCA, Laplacian Score, and Chi-square feature selection. We used k-fold cross-validation (k=5) to validate the model's performance and accuracy, ensuring consistency in results. The evaluation metrics used included accuracy, precision, recall, and F1 score, as appropriate for each algorithm and type of analysis.

Ten dimensionality reduction algorithms, namely Principal Components Analysis (PCA), Laplacian Score, Chi-square feature selection, Decision Tree, Autoencoder, Local Linear, Independent Component Analysis, Isometric Feature Mapping, Neighbourhood Component Analysis, and Autofeat were employed in this study which are discussed below sections.

Inclusion and exclusion criteria

The inclusion criteria for selecting participants were as follows: individuals between the ages of 20-22, with no history of neurological or psychiatric disorders, and available for repeated EEG measurements across different emotional stimuli sessions. Participants were required to have a normal or corrected-to-normal vision to view the emotional stimuli during EEG recording sessions. Individuals who had any form of neurological disorder, or psychiatric conditions, or were under medication that could potentially alter EEG readings were excluded. Moreover, participants were excluded if they had substance use that could affect central nervous system functioning or if they could not maintain the required stillness during EEG recordings. This ensured the reliability and consistency of the EEG data used in the analysis.

Principal components analysis (PCA)

Principal Component Analysis (PCA) is a linear dimensionality reduction technique that projects high-dimensional data onto a lower-dimensional linear subspace. This results in an increase in interpretability while minimizing information loss [30]. PCA is a commonly utilized unsupervised linear technique, particularly in the field of machine learning and data analysis [31].

Laplacian score

The Laplacian score is again an unsupervised linear feature extraction method which evaluates the features according to their locally preserving power. This Unsupervised feature selection method computes the Laplacian score for each variable or feature based on an observation that data from the same class are often close to each other and are probably related to the same topic. The Laplacian score is based on the Laplacian Eigenmaps and Locality Preserving Projection [32].

Chi-square feature selection

The Chi-squared feature selection is a numerical test that measures the deviation from the expected distribution considering the feature event is independent of the class value [33]. It consists of specifying the hypothesis, devising an analysis plan, examining sample data, and deducing results.

Decision tree

The C4.5 is a decision tree algorithm which involves the evaluation of features on an iterative basis. Each node of the decision tree is labelled as a feature and is also known as the non-leaf node indicating that at each iteration a feature needs to be selected towards labelling a non-leaf node of a tree [34]. It is also important to note that this C4.5 algorithm employs entropy or information gain towards heuristic selection of features.

Autoencoder

The Autoencoder, which is an unsupervised Artificial Neural Network, and it encodes the data by compressing it into lower dimensions (bottleneck layer or code) and then decodes them to reconstruct the original input, while this bottleneck layer or the code holds the compressed representation of the input data. It is important to note that the number of output unit must be equal to the number of input data [35].

Local linear

This method which puts forward a method of Non-Linear Dimensionality reduction was proposed by Rowies (2000). This method looks to portray a lower dimensional projection of the data which preserves the distance within local neighbourhoods. The LLE algorithm tries to reduce dimensionality and preserves the geometric features of the original non-linear feature structure.

Independent component analysis

ICA or Independent Component Analysis is technique which is based on higher order moment of input signals. It seeks for new linear transformation that decreases linear dependency between two components and very helpful in separating linearly combined independent sources [36]. For this reason and using such characteristics, Independent Component Analysis has been used to solve problems of blind signal separation such as EEG signals and sound or image separation [37, 38].

Isometric feature mapping

Isomap targets to preserve pair distances for all data points and tries to find the lower dimensions in which any two data points have same distance in high dimension. In this case, Isomap uses the neighbourhood graph connection to find the shortest path between two data as the shortest path between two data points as distance [39].

Neighbourhood component analysis

This is a non-parametric feature selection method to select the features with the goal of maximising prediction accuracy of regression and classification algorithms for metric learning. It learns a linear transformation to improve the classification accuracy of a stochastic nearest neighbours rule in the transformed space [40].

Autofeat

The Autofeat generates non-linear features and different operands while creating the feature space. This results in the growth of feature space and the categorical features are converted into one hot coded feature. Then the Autofeat algorithm removes the highly correlated features and features with low coefficient. This process is repeated until only a few are left, and these features are selected through this iterative process. The mean squared error between the input and the output of the network is trained to minimize [41].

Five machine learning algorithms, namely logistics regression, K-nearest neighbour (KNN), Naïve Bayes, multi-layer perceptron (MLP), and support vector machine (SVM), were employed in this study which are discussed below sections.

Logistics regression

Logistics Regression is normally used to predict the probability in a binary event occuring and commonly used in supervised learning. Here to forecast an output value, input values are mixed linearly with weights or coefficient values (referred to as the Greek capital letter Beta) [42]. In our research we can see the results when Logistics Regression was appiled to the dataset before and after implementing the dimensionality techniques. It is eveident from our research that after the dimensionality techniques is applied on the dataset and Logistics Regression is used the accuracy increases.

K nearest neighbour

When there is little or no prior knowledge about the distribution of the data, K-nearest-neighbour (KNN) classification should be one of the initial options for a classification research. It is one of the most basic and straightforward classification methods. The necessity to perform discriminant analysis when accurate parametric estimates of probability densities are unknown or challenging to calculate led to the development of K-nearest-neighbour classification [43]. KNN was used as an algorithm in our research to see the difference before and after dimensionality techniques is applied on the dataset. It can be seen clearly that there is a significant change in the results.

Naïve Bayes

The Naïve Bayes classifier is another machine Learning model which is used for classification task and based on the Bayes theorem.

$$P(A/B) = \frac{P(B/A)\,P(A)}{P(B)}$$

A is the hypothesis and B is the supporting evidence. When B has already happened, we may use the Bayes theorem to calculate the likelihood that A will also occur. Here, it is assumed that the predictors and features are independent. That is, the presence of one feature does not change the behaviour of another [44]. This algorithm is the third one used in our research to check its effectiveness to assess the before and after position of the accuracy metrics when different dimensionality techniques are used.

Multi-layer perceptron

MLP or Multi-Layer Perceptron classifier is a Neural Network which is used for classification tasks and consists of multiple layers and each layer is fully connected to the following one. With the exception of the nodes in the input layer, the nodes of the layers are neurons with nonlinear activation functions. There may be one or more nonlinear hidden layers between the input and the output layer [45]. In our research, besides machine learning algorithms we have used MLP to understand the difference of results before and after dimensionality reduction techniques. The results show a significance increase in the accuracy metrics after implementing the DR techniques.

Support vector machines

Support vector machines or SVM are a group of supervised learning techniques for classifying data, performing regression analysis, and identifying outliers. Finding a hyperplane in an N-dimensional space (N is the number of features) that categorises the data points clearly is the main goal of the support vector machine algorithm that effectively distinguishes the two classes [46]. We have used this as our fifth and final algorithm to distinguish the difference between the results before and after dimensionality reduction techniques are applied. The results clearly show an increase in the accuracy metrics when the DR technique is applied and then the SVM algorithm is implemented. The accuracy performance metric was employed in this study to evaluate the classification algorithms which is discussed below sections.

Accuracy performance metric

The accuracy is calculated by taking all the correct predictions and then dividing by the total number of datapoints in the datasets. When all the classes are of equal importance, this metric, which describes how the model performs across all classes, is very helpful. The best accuracy is 1.0 and the worst is 0.0. It is generally calculated as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Where TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative.

In our research the Accuracy metric is considered to check our results. It has been observed that after the 10 Dimensionality Reduction techniques were implemented and then 5 Machine Learning algorithms were implemented, the results showed a positive and increased change in the results which can be seen in **Table 1**.

Results

The present study evaluated the effectiveness of dimensionality reduction techniques in combination with machine learning algorithms for analysing a given dataset. The results of the experiment demonstrated promising outcomes, as shown in **Table 2**. Specifically, five machine learning algorithms, namely logistics regression, K-nearest neighbour (KNN), Naïve Bayes, multi-layer perceptron (MLP), and support vector machine (SVM), were employed after the dimensionality reduction technique was applied.

To evaluate the effectiveness of the applied techniques, the pre- and post-analysis results were compared. The results indicate that the dimensionality reduction technique significantly enhanced the performance of the machine learning algorithms. The findings align with previous studies that demonstrated the efficacy of dimensionality reduction techniques in improving the performance of machine learning algorithms [47, 48].

Overall, the study demonstrates the effectiveness of dimensionality reduction techniques in enhancing the performance of machine learning algorithms in analysing complex datasets. The study's findings have significant implications for various fields, including finance, healthcare, and social sciences.

Discussion

The present study evaluated the effectiveness of dimensionality reduction techniques in combination with machine learning algorithms for analysing EEG datasets. To place our findings within the context of existing research, we first summarize the relevant literature. Dimensionality reduction methods have been widely employed to mitigate the complexities associated with high-dimensional EEG data, as noted in numerous studies [15, 49, 50]. Techniques such as PCA, Laplacian score, and Chi-square feature selection have demonstrated significant utility in improving model performance, as they help reduce computational complexity while maintaining high classification accuracy [51-53]. EEG has been extensively used for emotion classification, disease diagnosis, and brain-computer interfaces, with studies highlighting the use of dimensionality reduction for improving computational efficiency and classifier accuracy [54]. Techniques like PCA, ICA, and Laplacian score are particularly prominent for their ability to handle the high dimensionality of EEG signals, thereby reducing noise and enhancing classifier performance [55, 56, 64].

Machine learning algorithms, including K-nearest neighbour (KNN), logistic regression, Naive Bayes, and support vector machines (SVM), have been widely employed for classifying EEG signals, as evidenced by prior research [57-59]. The use of dimensionality reduction techniques such as PCA and Autofeat has been found to boost accuracy across these classifiers, supporting their application in real-world scenarios [60, 61]. Additionally, the use of autoencoder neural networks has shown promise in EEG signal analysis, particularly in reducing noise and extracting essential features [62, 63]. The integration of dimensionality reduction and machine learning models has thus emerged as a promising approach for effectively managing EEG data and improving classification outcomes, which aligns well with the findings of our study.

Table 2 displays the accuracy results attained using various feature extraction methods and machine learning algorithms on the EEG data. These findings show that no single feature extraction technique offers the greatest performance in every situation. This finding is in accordance with the work of García-Laencina et al. [64]. However results show increase in accuracy after applying dimensionality reduction methods, as also reported by Aler [65].

Before applying any dimensionality reduction, the KNN classifier has the highest performance. The curse of dimensionality problem frequently causes the performance of the k-NN algorithm to degrade as the number of features rises. As it is shown, applying dimensionality reduction algorithms increases the performance of KNN method, as well as other algorithms. Choosing the best k-value is crucial when using the KNN algorithm as the choice of a hyper-parameter influences the algorithm's result. The choice of distance metric also affects the model's success [58], even if the metric used to calculate distance is not a hyper-parameter for the k-NN method.

The results demonstrate that dimensionality reduction plays a significant role in capturing the essence of the human emotional mind from EEG data. Techniques like PCA and Autofeat effectively reduce redundant information while retaining essential features that are linked to specific emotional categories. This allows for an improved interpretation of how different brain signal components correlate with emotional experiences, which is particularly valuable for applications in mental healthcare and neurofeedback training. The ability to use dimensionality reduction to focus on key emotional indicators makes the analysis more feasible for real-time applications, thereby enhancing the practical value of emotion recognition systems.

Complex non-linear machine learning models, such as neural networks, perform better when dimensionality reduction techniques are used, but in practise they are difficult to train and even more challenging to explain to non-statisticians who need clear analysis of results as a foundation for crucial decisions. In the other hand linear models often offer lower prediction accuracies, however they are more effective and understandable. Meanwhile, by using appropriate feature reduction techniques, such Autofeat, we can achieve good accuracy for linear models.

Despite the promising results obtained in this study, there are certain limitations that should be noted. Firstly, the limited sample size of the EEG dataset used restricts the generalizability of our findings. Future research should focus on validating the proposed dimensionality reduction techniques with larger and more diverse datasets to ensure robustness across different settings. Furthermore, this study did not delve into advanced deep learning-based dimensionality reduction methods such as convolutional autoencoders, which have the potential to enhance feature extraction beyond what was achieved using traditional techniques. Future studies could investigate these methods to further improve the performance of EEG classification models. Lastly, the applicability of our findings may be constrained by the specific characteristics of the dataset used. Further research should explore these techniques across a broader range of EEG datasets that vary in stimulus type, recording setup, and population demographics to assess the generalizability of our approach.

Autofit [61] uses the initial inputs to automatically generate several tens of thousands of non-linear features, and it then meticulously chooses the most useful ones to add to the linear model's input features. This method produces sufficiently precise predictions on real world data. As can be seen in **Table 1**, it has produced high accuracy for all different classification method applied on our data.

According to the **Table 1**, PCA also has produced very high accurate outcome. Dimension reduction by PCA has often been recommended before ICA decomposition of EEG data, both to reduce the volume of required data and computation time [60, 66].

In addition to processing computational resources, the characteristics of the feature space itself have an impact on the approach selected. The ability of approaches like autoencoder to represent complex non-linear processes allows them to compress data into a low-dimensional latent space more effectively when the features have a non-linear relationship. Therefore they also show increase in accuracy of classification task for our data.

Conclusion

In conclusion, this study demonstrated the effectiveness of various feature extraction methods in enhancing the performance of machine learning algorithms for EEG datasets. Through experimentation with ten different feature extraction techniques and four machine learning algorithms, we have shown that accuracy can be improved by over 20-60% by applying the autofeat dimension reduction method, which removes highly correlated features with a very low coefficient. The highest accuracy was achieved with this method in all algorithms, and our results indicate that it is a promising approach for EEG data analysis. Our models achieved an accuracy of over 90% on the EEG dataset, with LR and SVM achieving accuracy rates of 99.06% and 98.12%, respectively, through the test split method. Additionally, the use of the MLP method with an “adam” optimizer and 400 epochs yielded an accuracy of 98.82%, surpassing the accuracy of previous studies on the same EEG dataset. In conclusion, our study demonstrates the importance of feature extraction in improving machine learning performance and provides valuable insights for future EEG data analysis research.

Disclosure of conflict of interest

None.

Address correspondence to: Seyed-Ali Sadegh-Zadeh, Department of Computing, University of Staffordshire, Leek Road, Stoke-on-Trent ST4 2DE, United Kingdom. E-mail: ali.sadegh-zadeh@staffs.ac.uk

References

[1] Gervasi R, Barravecchia F and Franceschini F. Applications of affective computing in human-robot interaction: State-of-art and challenges for manufacturing. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 2023; 237: 815-832.
[2] Krumhuber EG, Skora LI, Hill HCH and Lander K. The role of facial movements in emotion recognition. Nat Rev Psychol 2023; 2: 283-296.
[3] Sadegh-Zadeh SA, Sakha H, Movahedi S, Fasihi Harandi A, Ghaffari S, Javanshir E, Ali SA, Hooshanginezhad Z and Hajizadeh R. Advancing prognostic precision in pulmonary embolism: a clinical and laboratory-based artificial intelligence approach for enhanced early mortality risk stratification. Comput Biol Med 2023; 167: 107696.
[4] Kaur S and Kulkarni N. Recent trends and challenges in human computer interaction using automatic emotion recognition: a review. Int J Biom 2024; 16: 16-43.
[5] Kalateh S, Estrada-Jimenez LA, Hojjati SN and Barata J. A systematic review on multimodal emotion recognition: building blocks, current state, applications, and challenges. IEEE Access 2024.
[6] El-Nasr MS, Yen J and Ioerger TR. Flame-fuzzy logic adaptive model of emotions. Auton Agent Multi Agent Syst 2000; 3: 219-257.
[7] Jourahmad Z, Habibabadi JM, Moein H, Basiratnia R, Geranqayeh AR, Ghidary SS and Sadegh-Zadeh SA. Machine learning techniques for predicting the short-term outcome of resective surgery in lesional-drug resistance epilepsy. arXiv preprint arXiv: 2302.10901 2023.
[8] Suhaimi NS, Mountstephens J and Teo J. A dataset for emotion recognition using virtual reality and EEG (DER-VREEG): emotional state classification using low-cost wearable VR-EEG headsets. Big Data Cogn Comput 2022; 6: 16.
[9] Nazari MJ, Shalbafan M, Eissazade N, Khalilian E, Vahabi Z, Masjedi N, Ghidary SS, Saadat M and Sadegh-Zadeh SA. A machine learning approach for differentiating bipolar disorder type II and borderline personality disorder using electroencephalography and cognitive abnormalities. PLoS One 2024; 19: e0303699.
[10] Sadegh-Zadeh SA, Soleimani Mamalo A, Kavianpour K, Atashbar H, Heidari E, Hajizadeh R, Roshani AS, Habibzadeh S, Saadat S, Behmanesh M, Saadat M and Gargari SS. Artificial intelligence approaches for tinnitus diagnosis: leveraging high-frequency audiometry data for enhanced clinical predictions. Front Artif Intell 2024; 7: 1381455.
[11] Hu Z, Shukla K, Karniadakis GE and Kawaguchi K. Tackling the curse of dimensionality with physics-informed neural networks. Neural Netw 2024; 176: 106369.
[12] Sadegh-Zadeh SA, Nazari MJ, Aljamaeen M, Yazdani FS, Mousavi SY and Vahabi Z. Predictive models for Alzheimer’s disease diagnosis and MCI identification: the use of cognitive scores and artificial intelligence algorithms. NPG Neurologie-Psychiatrie-Gériatrie 2024.
[13] Van Der Maaten L, Postma E and Van den Herik J. Dimensionality reduction: a comparative. J Mach Learn Res 2009; 10: 13.
[14] Sadegh-Zadeh SA, Bagheri M and Saadat M. Decoding children dental health risks: a machine learning approach to identifying key influencing factors. Front Artif Intell 2024; 7: 1392597.
[15] Huang H, Shi G, He H, Duan Y and Luo F. Dimensionality reduction of hyperspectral imagery based on spatial-spectral manifold learning. IEEE Trans Cybern 2020; 50: 2604-2616.
[16] Sadegh-Zadeh SA and Bagheri M. Harnessing the power of clinical data in dentistry: importance and guidelines for dentists in AI modelling for enhanced patient care. Open J Clin Med Images 2024; 4: 1188.

[17] Bacanin N, Zivkovic M, Bezdan T, Cvetnic D and Gajic L. Dimensionality reduction using hybrid brainstorm optimization algorithm. In: Proceedings of International Conference on Data Science and Applications. Springer; 2022. pp. 679-692.
[18] Fukunaga K. Introduction to statistical pattern recognition. Elsevier; 2013.
[19] Anderberg A, Bailey J, Campello RJ, Houle ME, Marques HO, Radovanović M and Zimek A. Dimensionality-aware outlier detection. In: Proceedings of the 2024 SIAM International Conference on Data Mining (SDM). SIAM; 2024. pp. 652-660.
[20] Wang Z, Zhang G, Xing X, Xu X and Sun T. Comparison of dimensionality reduction techniques for multi-variable spatiotemporal flow fields. Ocean Engineering 2024; 291: 116421.
[21] García-Torres M, Ruiz R and Divina F. Evolutionary feature selection on high dimensional data using a search space reduction approach. Engineering Applications of Artificial Intelligence 2023; 117: 105556.
[22] Westphal M and Zapf A. Statistical inference for diagnostic test accuracy studies with multiple comparisons. Stat Methods Med Res 2024; 33: 669-680.
[23] Groppe DM, Urbach TP and Kutas M. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. Psychophysiology 2011; 48: 1711-1725.
[24] Bird JJ, Manso LJ, Ribeiro EP, Ekart A and Faria DR. A study on mental state classification using EEG-based brain-machine interface. In: 2018 international conference on intelligent systems (IS). IEEE; 2018. pp. 795-800.
[25] Kreibig SD. Autonomic nervous system activity in emotion: a review. Biol Psychol 2010; 84: 394-421.
[26] Matsuda M, Morikuni K and Sakurai T. Spectral feature scaling method for supervised dimensionality reduction. arXiv preprint arXiv: 1805.07006 2018.
[27] Lespinats S, Colange B and Dutykh D. Nonlinear dimensionality reduction techniques: a data structure preservation approach. Springer; 2022.
[28] Rajender N, and Gopalachari MV. An efficient dimensionality reduction based on adaptive-GSM and transformer assisted classification for high dimensional data. International Journal of Information Technology 2024; 16: 403-416.
[29] Yan J, Zhang B, Liu N, Yan S, Cheng Q, Fan W, Yang Q, Xi W and Chen Z. Effective and efficient dimensionality reduction for large-scale and streaming data preprocessing. IEEE Trans Knowl Data Eng 2006; 18: 320-333.
[30] Thakkar A, Kikani N and Geddam R. Fusion of linear and non-linear dimensionality reduction techniques for feature reduction in LSTM-based Intrusion Detection System. Applied Soft Computing 2024; 154: 111378.
[31] Smith LI. A tutorial on principal components analysis. 2002.
[32] Borah K, Das HS, Seth S, Mallick K, Rahaman Z and Mallik S. A review on advancements in feature selection and feature extraction for high-dimensional NGS data analysis. Functional & Integrative Genomics 2024; 24: 139.
[33] Saeed AM, Badawi S and Hassan ADA. Comparison of feature selection methods in Kurdish text classification. Iran Journal of Computer Science 2024; 7: 55-64.
[34] Liu H, Cocea M and Ding W. Decision tree learning based feature evaluation and selection for image classification. In: 2017 International Conference on Machine Learning and Cybernetics (ICMLC). IEEE; 2017. pp. 569-574.
[35] Berahmand K, Daneshfar F, Salehi ES, Li YF and Xu Y. Autoencoders and their applications in machine learning: a survey. Artificial Intelligence Review 2024; 57: 28.
[36] Wisniewski MG, Joyner CN, Zakrzewski AC and Makeig S. Finding tau rhythms in EEG: An independent component analysis approach. Hum Brain Mapp 2024; 45: e26572.
[37] Kotani M, Shirata Y, Maekawa S, Ozawa S and Akazawa K. Application of independent component analysis to feature extraction of speech. In IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339). IEEE 1999; 5: 2981-2984.
[38] Karhunen J, Hyvarinen A, Vigário R, Hurri J and Oja E. Applications of neural blind separation to signal and image processing. In: 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing. IEEE; 1997. pp. 131-134.
[39] Tenenbaum JB, de Silva V and Langford JC. A global geometric framework for nonlinear dimensionality reduction. Science 2000; 290: 2319-2323.
[40] Goldberger J, Hinton GE, Roweis S and Salakhutdinov RR. Neighbourhood components analysis. Adv Neural Inf Process Syst 2004; 17.
[41] Demetgul M, Yildiz K, Taskin S, Tansel IN and Yazicioglu O. Fault diagnosis on material handling system using feature selection and data mining techniques. Measurement 2014; 55: 15-24.
[42] Maalouf M. Logistic regression in data analysis: an overview. International Journal of Data Analysis Techniques and Strategies 2011; 3: 281-299.
[43] Peterson LE. K-nearest neighbor. Scholarpedia 2009; 4: 1883.
[44] Chhogyal K and Nayak A. An empirical study of a simple Naive Bayes classifier based on rank-

ing functions. In: Australasian Joint Conference on Artificial Intelligence. Springer; 2016. pp. 324-331.

[45] Pinkus A. Approximation theory of the MLP model in neural networks. Acta Numerica 1999; 8: 143-195.

[46] Tanveer M, Rajani T, Rastogi R and Shao YH. Comprehensive review on twin support vector machines. Annals of Operations Research 2024; 339: 1223-1268.

[47] Zebari R, Abdulazeez A, Zeebaree D, Zebari D and Saeed J. A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. Journal of Applied Science and Technology Trends 2020; 1: 56-70.

[48] Alhassan AM and Zainon WMNW. Review of feature selection, dimensionality reduction and classification for chronic disease diagnosis. IEEE Access 2021; 9: 87310-87317.

[49] Chhikara P, Jain N, Tekchandani R and Kumar N. Data dimensionality reduction techniques for Industry 4.0: research results, challenges, and future research directions. Softw Pract Exp 2022; 52: 658-688.

[50] Jimenez LO and Landgrebe DA. Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 1998; 28: 39-54.

[51] Jolliffe IT and Cadima J. Principal component analysis: a review and recent developments. Philos Trans A Math Phys Eng Sci 2016; 374: 20150202.

[52] Soni A, Rasool A, Dubey A and Khare N. Data mining based dimensionality reduction techniques. In: 2022 International Conference for Advancement in Technology (ICONAT). IEEE; 2022. pp. 1-8.

[53] Thaseen IS and Kumar CA. Intrusion detection model using fusion of chi-square feature selection and multi class SVM. J King Saud Univ Comput Inf Sci 2017; 29: 462-472.

[54] Subasi A and Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst Appl 2010; 37: 8659-8666.

[55] Comon P. Independent component analysis, a new concept? Signal Processing 1994; 36: 287-314.

[56] Hassan AR and Bhuiyan MI. A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features. J Neurosci Methods 2016; 271: 107-118.

[57] Hosseini MP, Hosseini A and Ahi K. A review on machine learning for EEG signal processing in bioengineering. IEEE Rev Biomed Eng 2021; 14: 204-218.

[58] Isa NEM, Amir A, Ilyas MZ and Razalli MS. The performance analysis of K-nearest neighbors (K-NN) algorithm for motor imagery classification based on EEG signal. In MATEC web of conferences. EDP Sciences 2017; 140: 01024.

[59] Meyer D, Leisch F and Hornik K. The support vector machine under test neurocomputing. Volume 2003; 55: 1-2.

[60] Artoni F, Delorme A and Makeig S. Applying dimension reduction to EEG data by Principal Component Analysis reduces the quality of its subsequent Independent Component decomposition. Neuroimage 2018; 175: 176-187.

[61] Horn F, Pack R and Rieger M. The autofeat python library for automated feature engineering and selection. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer; 2019. pp. 111-120.

[62] Akella A, Singh AK, Leong D, Lal S, Newton P, Clifton-Bligh R, Mclachlan CS, Gustin SM, Maharaj S, Lees T, Cao Z and Lin CT. Classifying multi-level stress responses from brain cortical EEG in nurses and non-health professionals using machine learning auto encoder. IEEE J Transl Eng Health Med 2021; 9: 2200109.

[63] Wang Y, Yao H and Zhao S. Auto-encoder based dimensionality reduction. Neurocomputing 2016; 184: 232-242.

[64] García-Laencina PJ, Rodríguez-Bermudez G and Roca-Dorda J. Exploring dimensionality reduction of EEG features in motor imagery task classification. Expert Syst Appl 2014; 41: 5285-5295.

[65] Aler R and Galván IM. Optimizing the number of electrodes and spatial filters for Brain-Computer Interfaces by means of an evolutionary multi-objective approach. Expert Syst Appl 2015; 42: 6215-6223.

[66] Sadegh-Zadeh SA. Computational methods toward early detection of neuronal deterioration. Diss. University of Hull; 2019.