

Treatment of Imbalance Dataset for Human Emotion Classification

Er. Shrawan Thakur

Department of Science in Computer Engineering, Nepal College of Information Technology, Lalitpur, Nepal
Email: ershrawan014@gmail.com

How to cite this paper: Thakur, Er.S. (2023) Treatment of Imbalance Dataset for Human Emotion Classification. *World Journal of Neuroscience*, 13, 173-191.
<https://doi.org/10.4236/wjns.2023.134012>

Received: August 18, 2023

Accepted: November 1, 2023

Published: November 4, 2023

Copyright © 2023 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Developments in biomedical science, signal processing technologies have led Electroencephalography (EEG) signals to be widely used in the diagnosis of brain disease and in the field of Brain-Computer Interface (BCI). The collected EEG signals are processed using Machine Learning-Random Forest and Naive Bayes- and Deep Learning-Recurrent Neural Network (RNN), Neural Network (NN) and Long Short Term Memory (LSTM)-Algorithms to obtain the recent mood of a person. The Algorithms mentioned above have been imposed on the data set in order to find out what the person is feeling at a particular moment. The following thesis is conducted to find out one of the following moods (happy, surprised, disgust, fear, anger and sadness) of a person at an instant, with an aim to obtain the result with least amount of time delay as the mood differs. It is pretty obvious that the accuracy of the output varies depending upon the algorithm used, time taken to process the data, so that it is easy for us to compare the reliability and dependency of a particular algorithm to another, prior to its practical implementation. The imbalance data sets that were used had an imbalanced class and thus, over fitting occurred. This problem was handled by generating Artificial Data sets with the use of SMOTE Oversampling Technique.

Keywords

Electroencephalography (EEG), Brain Computer Interface (BCI), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), Neural Network (NN), Synthetic Minority Over Sampling Technique (SMOTE)

1. Introduction

Technology is rapidly advancing, bringing enormous benefits into our lives. The emotion recognition concept has the potential to revolutionise human-computer interface technology, identify the feelings of those who are mentally challenged

or autistic who are unable to communicate themselves. A person's psychological condition after a serious injury or sickness is monitored. Emotion recognition could be a field from which business analysts, content creators, and advertising can benefit greatly because consumers are influenced by emotions when making purchasing decisions.

The fascinating field of affective computing, also known as emotion recognition technology, focuses on creating systems and algorithms that can recognize, analyze, and understand human emotions based on a variety of data inputs, including physiological signals, facial expressions, voice tones, body language, and body language. **Data Acquisition:** Emotion recognition technology gathers data from various sources, such as cameras, microphones, sensors, and wearable. These devices capture the user's facial expressions, voice, gestures, or physiological responses like heart rate and skin conductance. **Feature Extraction:** The collected data is processed to extract relevant features. For example, in facial recognition, features like eyebrow movement, lip curvature, and eye openness are analyzed. In voice recognition, pitch, tone, and speech rate are important features. **Machine Learning Algorithms:** Machine learning models, such as deep neural networks, are trained on labelled data sets to recognize patterns and relationships between the extracted features and emotions. These models learn to classify emotions into categories like happiness, sadness, anger, and fear. **Emotion Classification:** Once trained, the algorithm can classify emotions in real-time or based on historical data [1]. A wearable ring created by Mood metric tracks changes in the user's skin conductance to measure emotional health. This technology can be used to monitor mental health and control stress.

We were inspired to do research on emotion recognition, because we believe that if emotion recognition is effectively utilised in the fields indicated, the outcomes will be useful to human civilization. Human understanding, conduct, and communication are all influenced by emotions. It's a brain impulse that propels an organism into action, triggering instinctive reactive behaviour that has evolved as a survival mechanism to meet a survival need [26]. The process of recognizing human emotions is known as emotion recognition. There are two types of emotions: positive and negative. While positive emotions are needed for good health, negative emotions can cause mental health problems, such as depression, stress, and anxiety [2].

Billions of brain cells (neurons), are composed inside the human brain, which are interconnected to each other through synapses to form a neural network (1011 neurons and 104 connections in the human brain). When these cells are activated, electrical activity occurs in the brain. Alpha, beta, delta, gamma, and theta rays' classification of these waves is done using the Brain-Computer Interfaces (BCI) [3].

Electroencephalography (EEG) is a neuropsychological measurement of electrical activity in the brain that is recorded using electrodes. Electrodes are usually placed on the scalp, although in some circumstances they are implanted subdurally and in the cerebral cortex. The EEG measures voltage changes caused by

ionic fluxes within the cerebrum's neurons. Even while EEG has a poor spatial resolution and requires several sensors on the scalp, it has great temporal resolution. Theta, alpha, beta, and gamma subbands make up a pure EEG signal. Individual sub-bands are related with specific physical activities. Theta waves, for example, are associated with REM sleep, profound and raw emotions, and cognitive processes. The Alpha waves are indicated by a drowsy condition. It also serves as a source of relaxation and tranquillity. The conscious state during the cognitive process is represented by the letter beta. When trying to comprehend two different senses at the same time, such as sound and sight, Gamma waves are available [4]. Through electrodes on the scalp, electroencephalography (EEG) monitors brainwave activity. Different emotions are linked to specific patterns of brainwave activity. Consider this: A happier state is associated with more activity in the left prefrontal cortex. Right prefrontal brain activity is more active when someone is sad. Beta and gamma oscillations, together with increased theta activity, are a result of fear. Increased beta power and right prefrontal brain activity are associated with anger. Gamma activity may be more pronounced in surprise. EEG-based emotion identification is a developing field of research since individual diversity and context might affect these patterns [5].

Because of their poor signal-to-noise ratio, classifying these signals is difficult. The electroencephalogram (EEG) is a technique used by scientists to collect information by putting an active electrode on the head and recording electrical activity. Many algorithms may be used to classify EEG data, one of which is the Recurrent Neural Network. It's utilised to tell the difference between seizure and non-seizure events in EEG signals for automated seizure detection.

We discovered core data as well as six emotion labels when working with the DEAP dataset (arousal, valence, dominance and liking). It is possible to determine human emotion from DEAP data utilising only main data and two emotion labels (arousal and valence). Each emotion label is separated into two equal segments, yielding four different emotional states. Also we have minority datasets, so unable to find the desired result. It is, nevertheless, possible to improve on the accuracy that they have gained. Because it only employs six emotional states, this binary classification is unable to detect real-life emotions precisely [6].

Due to imbalance sets, we have introduced SMOTE and PCA in this result. When working with imbalanced datasets, the problem is that most machine learning techniques overlook the minority class, resulting in poor performance, even though it is often the minority class that is most relevant. In this research, we have used given datasets but have not tested with real time EEG datasets [7].

In terms of research, small data sets are limited, thus we may test the models with a larger amount of data to test them to a greater extent. We can use more classifiers to build more models and evaluate them to see if any of them are more accurate than the one we used. The same algorithms can be applied for the brain waves of different people to know if it works the same when the data being fed is varied. Once the algorithm successfully provides a higher degree of output for a

huge number of data sets then it can be further experimented to obtain as recent data as possible. The optimum goal of the research would be to process the live data and provide instant results after the application of algorithms on them. In the future we would like to obtain data in as real time as possible so that the mental state could be known in a profound manner.

There is no real-time data from the human brain. This study only contains a sample data set due to the difficulty of obtaining big data sets. Only a few Machine Learning Algorithms were used in this study. Because this neural network's calculation is slow and training can be challenging, resources are required to train and run the model. In terms of architecture, the models that have been used are quite simple and light.

In comparison to the Support Vector Machine (SVM) and Multiple Layer Perceptron, the K-NN classifier improves classification accuracy and needs less training and testing time (MLP). For time series data Recurrent Neural Network is better than traditional RNN. Furthermore, when compared to Bayes Network, Naive Bayes, SVM, Multi-layer Perceptron, and Random Forest, it is more accurate in identifying EEG signals to detect satisfaction, engagement, frustration, and difficulty. In addition, when compared to Multilayer Perceptron Neural Networks, the k-NN gives superior eye state categorization accuracy [8].

2. Literature Review

Include the contents of section one of literature review here. In this section, we describe the related work on brainwaves and the application of the brain waves in various fields. In the past few decades, many authors and researchers have contributed to brainwaves, their significance, and their applications. Some of the works in the field of brainwaves related to this thesis are mentioned below. Many studies have been conducted to remove these artifacts from the EEG signal so that further processing of the EEG data may be done easily and accurately. In their study, the authors offer a fully automated and online artifacts reduction method for the electroencephalogram (EEG) that may be used in brain-computer interfaces (BCI). A new combination of wavelet decomposition, independent component analysis, and threshold was used to create the approach. During online EEG acquisition, the FORCE was able to operate on a minimal channel set and did not require any extra signals (electrooculogram signals). Blink, electromyogram (EMG), and electrooculogram (EOG) artifacts were among the sorts of artifacts that the approach was able to eradicate. Similarly, authors in paper [9] show how to remove artifacts from EEG signals using Quantum Neural Network-based EEG filtering algorithms. It is a unique neural information processing architecture influenced by quantum physics and includes the well-known Schrodinger wave equation, according to the creators. The researchers suggested an RQNN architecture for representing non-stationary stochastic signals as time-varying wave packets. The RQNN filtering strategy was used to increase signal separability in a two-class motor imagery-based brain-computer interface by fil-

tering electroencephalogram (EEG) data before feature extraction and classification. Biometric authentication systems have also utilised brainwave oscillations. In terms of privacy compliance, resilience against spoofing attacks, capacity to conduct continuous identification, inherent aliveness detection, and universality, brain signals have several qualities not shared by the most regularly used biometrics, such as face, iris, and fingerprints. Because of these characteristics, the use of brain signals is intriguing. The authors in papers [10] [11], and [12] have explained the various methods to use brain wave signals for biometric user authentication. According to these publications, EEG-based authentication systems are made up of four major modules: data collection, pre-processing, feature extraction, and classification. Identification and verification are typically used to evaluate EEG biometric authentication systems. In identification mode, the accuracy of the system is commonly assessed using the average correct recognition rate (CRR) or genuine acceptance rate (GAR). Four main factors impact the performance of EEG biometric authentication systems: EEG acquisition methodology (the procedure used to capture EEG data), pre-processing approach, retrieved EEG signal properties, and classification scheme.

In these papers [13] [14], and [15] authors have explained various feature extraction methods like DWT, FFT, HHT, and classification algorithms like SVM, GA-ANN, Random Forest, etc. for the classification of motor imagery brainwave signals. The authors of the research proposed a method for accurately identifying off-line experimental electroencephalogram (EEG) signals from the BCI Competition 2003. There were three essential steps to the approach. To extract features of EEG for mental tasks, the wavelet coefficient was first recreated using the wavelet transform. At the same time, they employed the AR model power spectral density as the frequency feature in frequency extraction. Second, as the final feature vector, they combine the power spectral density feature and the wavelet coefficient feature. Finally, a linear approach based on iteration was used to classify the feature vector and determine the weight of its components. The categorised result revealed that utilising a feature vector produces a better effect than using a single feature. Similarly, the paper makes an appearance. Finally, a linear approach based on iteration was used to classify the feature vector and determine the weight of its components. The categorised result revealed that utilising a feature vector produces a better effect than using a single feature. The classification of a three-class mental task-based brain-computer interface (BCI) is also presented in the paper, which uses the Hilbert-Huang transform (HHT) for the features extractor and fuzzy particle swarm optimization with cross-mutated-based artificial neural network (FPSOCM-ANN) for the classifier. Electroencephalography (EEG) signals from six channels were used in the trials on five able-bodied participants and five patients with tetraplegia, and several time windows of data were explored to identify the maximum accuracy. The Hilbert-Huang Transform (HHT) for the features extractor and fuzzy particle swarm optimization with cross mutated-based artificial neural network (FPSOCM-ANN) for the

classifier were used to classify a three-class mental task-based brain-computer interface (BCI) [14]. Electroencephalography (EEG) signals from six channels were used in the trials on five able-bodied participants and five patients with tetraplegia, and several time windows of data were explored to identify the maximum accuracy.

Using wavelet transformations, the EEG data are divided into several frequency components and offer coefficients for different scales and places in time. EEG signals are divided into independent components using Independent Component Analysis (ICA). It aids in separating genuine brain activity from artefacts. Thresholding Methodologies to identify and exclude components that are associated with artefacts, apply a threshold to the ICA components. Keep just the parts that match to actual brain impulses. Together, these techniques improve the quality of EEG data by identifying and eliminating undesirable artifacts [16]. Also we can see the table of summary of literature review (Table 1).

3. Model Implementation

We can see the flow diagram of our research where the model implementation in below image Figure 1.

1) Data Collection and Prepossessing:

Because of the sensitive nature of the data, it is extremely difficult to extract; and hence, Open Neuron data set for classification is used [29].

We collect Data from openneuro. The data is normalised using the min-max approach at this point. The signal magnitude is adjusted using normalisation. The range of all waves is mapped between 0 and 1. The following equation is used to calculate min-max normalisation:

Table 1. Major research papers synopsis.

Summary of Literature Review				
Citation	Year	Feature extraction	Modelling technique	Performance in Percentage
Wei Liu <i>et al.</i> [17]	2016	PSD, DE	BDAE	Valence: 85.20 Arousal: 80.50
Wei-Long Zheng <i>et al.</i> [18]	2016	DE	GELM	69.67
Xiang Li <i>et al.</i> [19]	2016	CWT	C-RNN	Valence: 72.06 Arousal: 74.12
Samarth Tri pathi <i>et al.</i> [20]	2017	GD	CNN	Valence: 81.41 Arousal: 73.36
Abeer Al-Naf jan <i>et al.</i> [21]	2017	PSD	DNN	Valence: 82 Arousal: 82
Soheil Ray atdoost <i>et al.</i> [22]	2018	HOC, PSD, DE, HOS	RF	Valence: 60.86 Arousal: 58.08
Guolu Cao <i>et al.</i> [23]	2019	PCA	CNN	Valance: 81.2 ± 3.0 Arousal: 84.3 ± 4.0
Yucel Cimtay <i>et al.</i> [24]	2020	Raw Data	CNN	91.5
Yucel Cimtay <i>et al.</i> [25]	2020	Raw Data	CNN	DEAP-72.81
Yulong Luo <i>et al.</i> [26]	2020	Variance	SNN	Valence: 78 Arousal: 74
Zhongke Gao <i>et al.</i> [27]	2020	DE	CDCN	92.58
Rahul Sharma <i>et al.</i> [3]	2020	DWT	LSTM	82.01
Divya Acharya <i>et al.</i> [28]	2020	FFT	LSTM	89.83

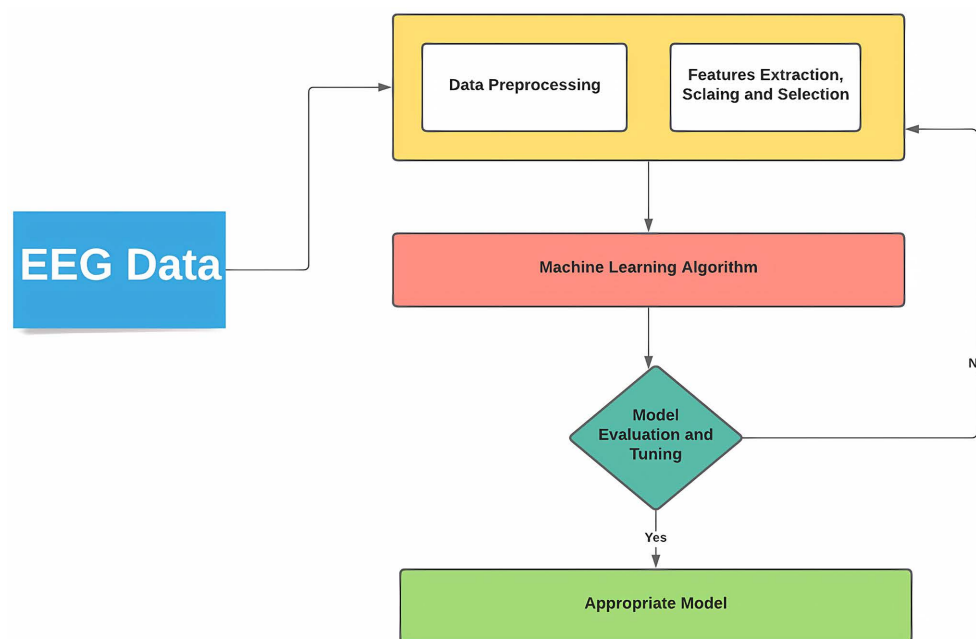


Figure 1. Flow diagram.

$$y = x - \min(x) / \max(x) - \min(x) \quad (1)$$

A genetic algorithm is essentially heuristic inspired by Charles Darwin's natural selection theory. The fittest individuals are picked for reproduction in order to generate offspring for the following generation, and this algorithm is modelled after natural selection [30].

In my research on EEG (Electroencephalography) emotion analysis using Neural Network (NN), LSTM (Long Short Term Memory) and RNN (Recurrent Neural Network) models, Genetic Algorithms (GAs) are integrated into the methodology as follows: Hyperparameter Optimization: GAs are used to optimise hyperparameters for LSTM and RNN models, such as learning rates, batch sizes, and the number of hidden layers. This helps fine-tune the models for better performance in emotion recognition tasks. Architecture Search: GAs explore different model architectures by evolving variations of LSTM and RNN structures. They determine the best network configurations, including the number of neurons in layers, to improve model accuracy.

Feature Selection: GAs assist in selecting the most relevant EEG features for emotion analysis by evolving feature subsets. This reduces noise in the data and enhances the models' ability to capture emotional patterns [31].

Ensemble Creation: Genetic Algorithms can create ensembles of LSTM and RNN models with diverse architectures and hyperparameters. These ensembles improve model robustness and accuracy by combining multiple model predictions.

Objective Function Optimization: GAs optimise an objective function that measures the performance of LSTM and RNN models in emotion recognition tasks. This function typically considers metrics like accuracy, F1-score, or other

relevant criteria.

The goal of this study is to improve the general performance of LSTM NN and RNN models in EEG-based emotion analysis by introducing Genetic Algorithms into the research technique. The search for the best model configurations and hyperparameters is automated using GAs, which enhances the models' capacity to identify and categorise emotional states from EEG data.

2) *Emotion Classification and analysis using Machine Learning:* With change in the mood of a person; changes the their brain signals as certain chemical compositions happen in the neurons and the objective of this research is to tack the most recent mood of a person with highest possible accuracy using various available methods and to find out the best possible way to do so. To achieve the desired goal; altogether five different techniques-Recurrent Neural Network, Random Forest, Naive Bayes, Neural Network and Long Short-term Memory have been used to compare the results and the accuracy against one another.

3) *Model Evaluation and Tuning:* While using the methods mentioned above the undersampling and oversampling of the data has been handled using Synthetic Minority Oversampling Technique (SMOTE) to create a perfect balance between the underfeeding and overfeeding. The result is being presented in the form of graphs and charts for the better and easy understanding of the reader.

In this research five different algorithms-Recurrent Neural Network, Random Forest, Naive Bayes, Neural Network and Long Short-term Memory, have been used to analyze the data set and find out the mood of the person. Algorithms Recurrent Neural Network, Random Forest, Naive Bayes, Neural Network and Long Short-term Memory, has been used to analyze the data set and find out the mood of the person. By partitioning the data sets into numerous partitions, several classification algorithms are employed to develop a model that predicts outcomes. The stages that each of these categorization algorithms has in common are listed below.

- Divide the data into two parts: training and testing.
- Create the model with a machine and deep learning approach utilising a training set.
- As many groups as feasible should be created from the data sets.
- Data Fitting into Algorithms.
- Assess the model's precision.

4. Data Augmentation and Upsampling

Data Augmentation is a technique to increase the size of a training set artificially by changing existing data. Using Data Augmentation is recommended if we want to avoid overfitting, if the original data set is too small to train on or if we want to get more performance out of your model.

It's not simply about avoiding overfitting when it comes to data augmentation. In general, a large data set is required for the performance of both ML and

Deep Learning (DL) models. However, we can improve the model's performance by adding the data we already have. This shows that using Data Augmentation to increase the model's performance is a good idea. In our circumstance, we have about 2100 data sets, which is insufficient to train a machine learning model. As a result, for data augmentation, we have used Oversampling and SMOTE. In most cases, oversampling strategies are better than under sampling methods. Because we tend to eliminate occurrences that may contain essential information when we under sample data, this is the case.

1) *Synthetic Minority Oversampling Technique—SMOTE*:

Synthetic Minority Oversampling Technique (SMOTE) is one of the most extensively utilised oversampling techniques for resolving the imbalance problem. Its purpose is to achieve a more equitable distribution of classes by randomly re-creating examples of minority classes. When one class—the majority class—significantly outnumbers another class in the data set, this is known as a class imbalance [32].

To balance the class distribution, SMOTE creates synthetic samples for the minority class. By interpolating between existing data, SMOTE generates synthetic examples for the minority class. SMOTE chooses the k closest neighbours from the minority class for each sample. Then, it produces synthetic examples by combining the chosen sample and its neighbours in linear fashion. By adding fresh samples to the feature space, this procedure effectively increases the minority class. By giving the classifier a balanced data set, SMOTE makes it easier for it to learn the decision boundary.

SMOTE brings together existing minorities to create new ones. It uses linear interpolation to construct virtual training records for the minority class. For each example in the minority class, these synthetic training records are constructed by selecting one or more of the k -nearest neighbours at random **Figure 2**. The data is rebuilt after the oversampling approach, and several classification models can be used to the updated data. Working Procedure: The overall number of oversampling observations (N) is shown first. One of the most popular

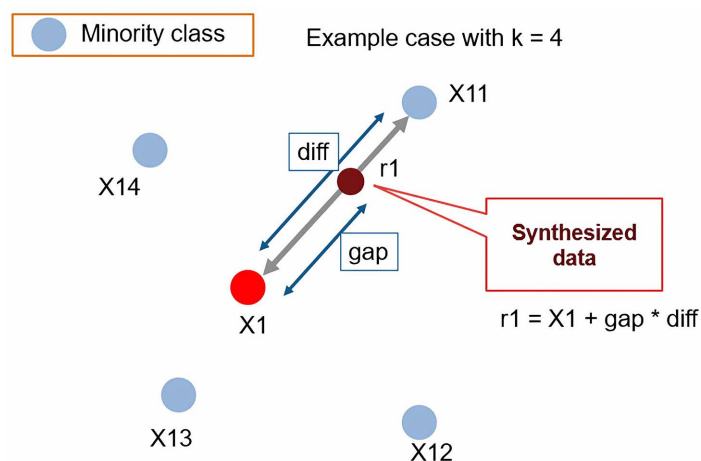


Figure 2. Synthetic minority oversampling technique (SMOTE) [33].

binary class distributions is 1:1. This may, however, be reduced depending on the circumstances. The iteration starts with a choice of a positive class instance at random. The KNNs (by default 5) for that instance are then retrieved.

Finally, N of these K instances are chosen to serve as the foundation for the creation of new synthetic instances. Any distance metric can be used to calculate the distance between the feature vector and its neighbours. This variance is now multiplied by any random value between 0 and 1 and added to the feature vector from before. Though this approach is quite beneficial, it has a few limitations. The synthetic instances are created in the same direction, with an artificial line connecting their diagonal instances. As a result, the decision surface generated by a few classification algorithms grows increasingly complex. SMOTE tends to generate a large number of noisy data points in feature space.

a) *Feature Scaling and Selection:*

The process of limiting the amount of input variables in a predictive model is known as feature selection. The number of input variables should be kept to a minimum to reduce the computational cost of modelling and, in some situations, to increase the model's performance. The link between each input variable and the goal variable is examined using statistics, and the input variables having the strongest association to the target variable are chosen. Despite the fact that the statistical measures utilised are dependent on the data types of the input and output variables, these tactics can be quick and successful.

Using the model's feature importance attribute, we can determine the feature importance of each feature in our dataset **Figure 3**. A score is assigned to each of our data features; the greater the score, the more important or relevant the feature is to our output variable. We'll use Extra Tree Classifier to extract the top 10 features for the dataset because Feature Importance is an intrinsic class that comes with Tree Based Classifiers.

b) *Testing and Training Dataset:*

Separating data into training and testing sets is required for our model evaluation. Only a tiny portion of the data is used for testing, with the rest being used for training. During this step of data preparation, data sets were split into training and testing sets **Figure 4**. To test the model, we construct predictions against the test set after processing the training data. As a result, a pseudo-random number generator was used to generate 80% of the data for the training set and 20% for the testing set.

5. Exploratory Analysis

One of the most typical techniques of dealing with an imbalanced dataset is to resample the data **Figure 5**. Undersampling and oversampling are the two of the most common ways for this. Oversampling techniques are preferred over undersampling techniques in most circumstances. The reason for this is that when we undersample data, we tend to exclude occurrences that may contain crucial information.

Out[15]:

	SN	delta	Alpha	theta	beta	gamm	delta_sqrt	Alpha_sqrt	theta_sqrt	beta_sqrt	...	Min_25	Max_50	Min_50	Ma
0	0	0.368161	0.192478	0.096655	0.106708	0.235999	0.606762	0.438723	0.310893	0.326662	...	50.029000	52.091000	49.918000	52.10
1	2	0.406963	0.163431	0.143101	0.137287	0.149218	0.637936	0.404266	0.378287	0.370523	...	-79.038000	356.210000	-96.599000	406.55
2	3	0.373781	0.151139	0.131902	0.136399	0.206778	0.611376	0.388766	0.363184	0.369323	...	-266.370000	340.240000	-304.110000	370.72
3	4	0.374374	0.152025	0.130324	0.134591	0.208687	0.611861	0.389904	0.361004	0.366867	...	-285.490000	455.320000	-317.060000	492.76
4	5	0.355840	0.147438	0.132670	0.135163	0.228890	0.596523	0.383977	0.364239	0.367645	...	-330.020000	475.220000	-365.540000	512.81
...
25045	1249	0.295102	0.161859	0.105382	0.149832	0.287826	0.543232	0.402317	0.324625	0.387081	...	-339.581759	666.927087	-359.644166	702.82
25046	1249	0.222823	0.136790	0.106329	0.178392	0.355667	0.472041	0.369850	0.326081	0.422365	...	-214.614602	452.613851	-228.004105	476.36
25047	1249	0.297348	0.161449	0.104510	0.149991	0.286702	0.545296	0.401806	0.323280	0.387286	...	-338.300617	664.699768	-358.353816	700.61
25048	1249	0.222686	0.137438	0.106512	0.178370	0.354994	0.471894	0.370716	0.326361	0.422336	...	-214.912729	453.900258	-228.248047	477.66
25049	1249	0.248836	0.147701	0.108542	0.168312	0.326611	0.498834	0.384318	0.329457	0.410258	...	-248.502199	497.983949	-264.825288	526.48

25050 rows x 39 columns

Figure 3. Feature scaling result.

```
Out[3]: Index(['SN', 'delta', 'Alpha', 'theta', 'beta', 'gamm', 'delta_sqrt',
              'Alpha_sqrt', 'theta_sqrt', 'beta_sqrt', 'gamm_sqrt', 'delta_dob',
              'Alpha_dob', 'theta_dob', 'beta_dob', 'gamm_dob', 'PFD', 'DFA',
              'FISHER INFO', 'AR', 'AMV', 'WV', 'SS', 'ZC', 'max_est54', 'min_est54',
              'max_std54', 'min_std54', 'Max_25', 'Min_25', 'Max_50', 'Min_50',
              'Max_75', 'Min_75', 'Max_99', 'Min_99', 'Max_01', 'Min_01',
              'Eye_state'],
             dtype='object')
```

Figure 4. Feature importance.



Figure 5. Sample data sets.

SMOTE is an oversampling approach that generates synthetic samples for the minority group. Overcoming the problem of overfitting produced by random oversampling is easier with this method. It focuses on the feature space to generate new examples by interpolating between positive occurrences that are near in

proximity.

6. Results

1) Recurrent Neural Network Algorithms: Recurrent Neural Networks (RNNs) are great in capturing temporal dependencies in sequential EEG data, making them a good choice for emotion classification applications. RNNs are useful in this situation because they can simulate how emotions change over time.

To function successfully, they need to be carefully tuned because they are vulnerable to the vanishing gradient problem, which can impair their capacity to learn long-term dependencies. Additionally, when working with huge data sets, training RNNs can be computationally expensive. The RNN Classifier revealed that the RNN model was able to predict with a 82% accuracy rate **Figure 6**.

The training and test Accuracy of the RNN classifier are 82.45% and 81.36% respectively **Table 2, Figure 7**.

2) Long Short Term Memory:

Long Short-Term Memory (LSTM) networks are ideal for interpreting EEG signals since they are designed basically for sequential data. They can handle variable-length sequences and are excellent at capturing long-term dependencies. LSTMs work best with enough training data, but they can be sensitive to hyperparameter settings and necessitate careful adjustment.

LSTMs are an excellent option for emotion classification tasks since they are particularly effective at capturing temporal patterns in EEG data despite being computationally demanding in comparison to some other techniques. The training and test Accuracy of the LSTM classifier are 84.05% and 83.09% respectively **Figure 8, Figure 9 (Table 3)**.

Table 2. Result for RNN.

Training Accuracy	Test Accuracy
82.45%	81.36%



Figure 6. RNN model accuracy curve.

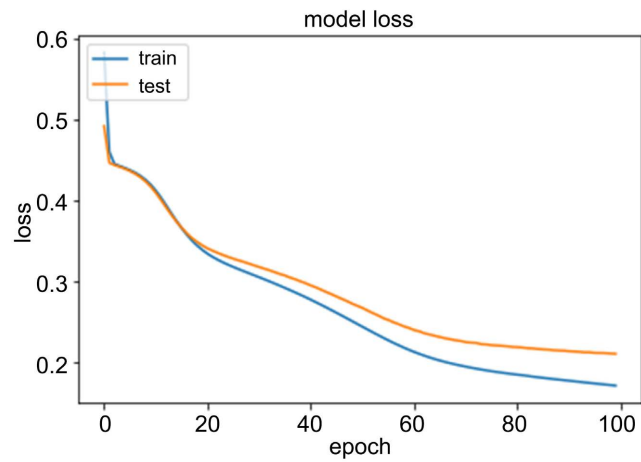


Figure 7. RNN model loss curve.

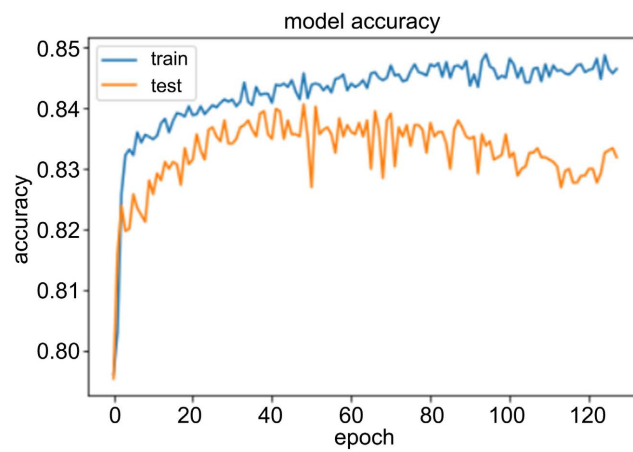


Figure 8. LSTM model accuracy curve.

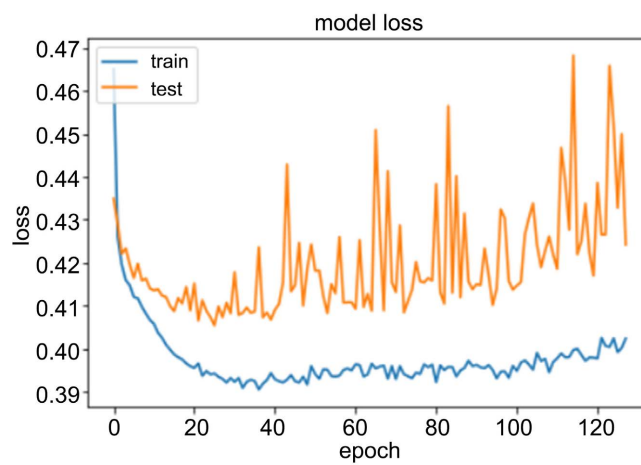


Figure 9. LSTM model loss curve.

Table 3. Result for LSTM.

Training Accuracy	Test Accuracy
84.05%	83.09%

3) *Random Forest Classifier.*

Random Forest is a strong and adaptable method that can handle both numerical and categorical information. Due to its lower propensity for overfitting, it is especially helpful when working with noisy EEG data. Additionally, Random Forest offers feature priority scores that can be used to determine which EEG features are most important for classifying emotions. But because it's not naturally suited to sequential data, like EEG signals, it might have trouble accurately capturing the temporal dynamics of emotions.

The Confusion Matrix of the Random Forest model demonstrated that it was able to predict TP (True Positives) and TN (True Negatives) with a 96% accuracy rate **Figure 10**.

The classification report for the Random Forest model shows that it has a precision, recall, and F1 score **Table 4**.

4) *Naive Bayes Classifier.*

Naive Bayes is a simple but efficient computational technique for classifying emotions, especially for small datasets. It functions well under specific circumstances and can handle high-dimensional feature spaces. For EEG data, where features can be highly correlated, its essential premise of feature independence (the "naive" assumption) may not hold true, which may limit its ability to capture complex emotional patterns and temporal dynamics. The Confusion Matrix of the Naive Bayes model demonstrated that it could predict TP (True Positives) and TN (True Negatives) with a 94% accuracy rate **Table 5**.

Table 4. Result for random forest classifier.

Recall Score	Accuracy	Precision Score	F1-Score
96%	96%	96%	96%

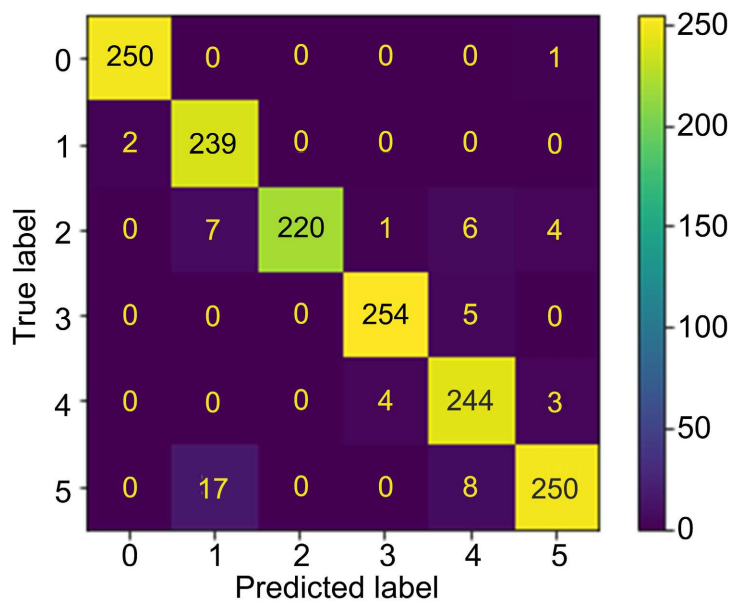


Figure 10. Confusion matrix for random forest classification.

The classification report for the Naive Bayes model shows that it had low accuracy, recall, and F1 scores **Figure 11**.

5) Neural Network:

The versatility of neural networks (NNs) is high, and they can recognize complex patterns in EEG data. They can be used in both shallow and deep architectures, and because they can learn new features, they don't require as much feature engineering. However, NNs are susceptible to overfitting if not adequately regularised and require a significant quantity of data to train deep models effectively. Large neural networks can also be expensive computationally.

The first 38 columns contain our input features. We have the feature that we want to forecast in the last column: by separating features and labels. After that, we divided our data into input features (X) and the label of what we want to predict (Y). We need to determine the optimal numbers for our architecture now that we've defined it. Before we begin training, we must first configure the model by telling it which optimization method to use, which loss function to employ, and which other metrics you want to track in addition to the loss function. According to the Neural Network Model study, it has an accuracy of 81% **Figure 12**.

Because we are fitting the parameters to the data, the function is termed "fit". We must first determine what type of data we will use for training, which is x train and encoded1. Then we figure out how big our mini-batch will be and how long we'll train it (epochs). Finally, we define our validation data so that the model can assess our performance at each level (**Figure 13, Table 6, Table 7**).

Table 5. Naive Bayes classifier result.

Precision Score	Recall Score	Accuracy	F1-Score
94%	94%	94%	94%

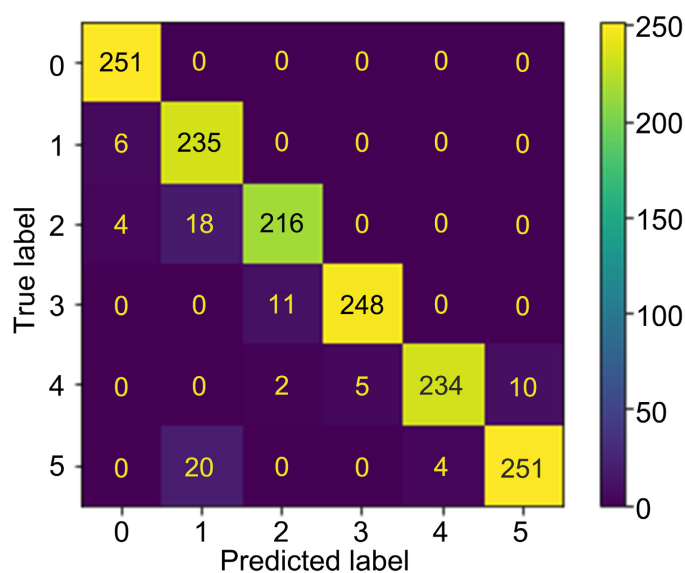


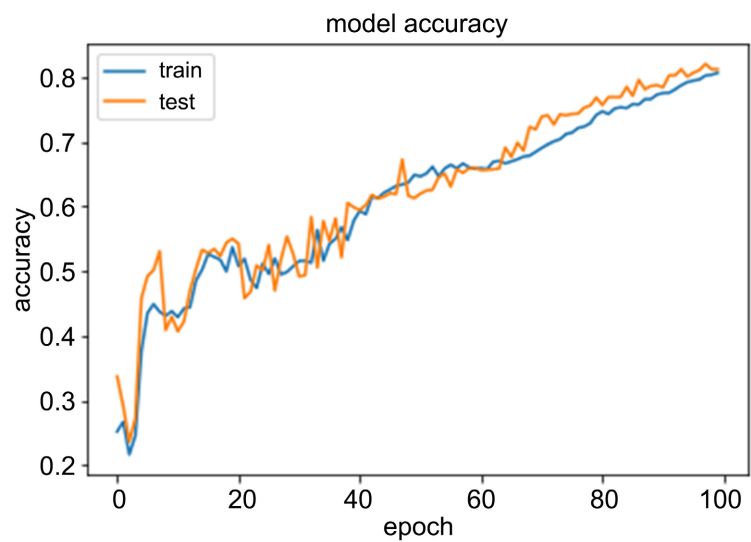
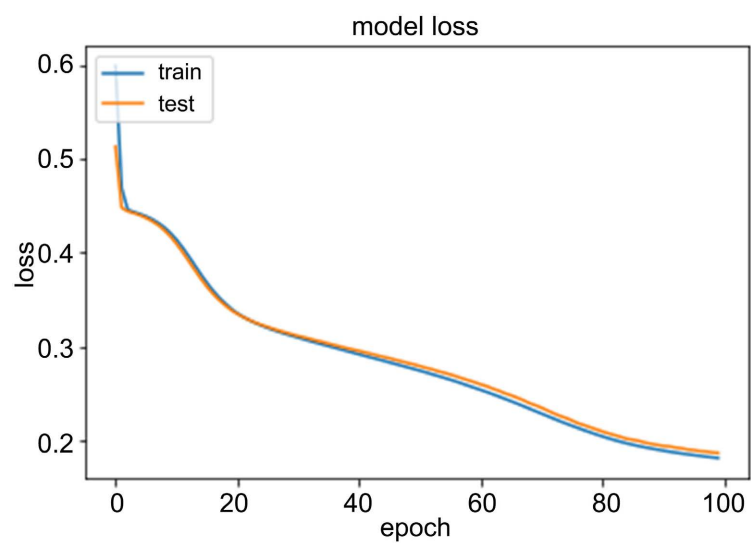
Figure 11. Confusion matrix for Naive Bayes classification.

Table 6. Result for NN.

Training Accuracy	Test Accuracy
80.68%	81.25%

Table 7. Algorithm with their training and test accuracy **Table 7.**

Algorithm	Training Accuracy	Test Accuracy
Random Forest	96.03%	95.4
Naive Bayes	94.72%	93.39
LSTM	84.05%	83.09
NN	80.68%	81.25
RNN	82.45%	81.36

**Figure 12.** Model accuracy curve.**Figure 13.** Model loss curve.

7. Conclusions

While processing the brain waves-alpha, beta, delta, theta and gamma-received from the human brain, minority classes were found in a huge amount in the datasets which led to a subsequent problem of overfitting and to overcome this issue, SMOTE-Oversampling technique was used.

The findings of this study reveal that varying accuracy can be obtained depending upon the Machine Learning Algorithm used. Random Forest, Naive Bayes, RNN, LSTM and NN were used to develop the classification algorithm with a high accuracy.

To improve it, this research can be expanded by decreasing the time unit used to retrieve signals and adjusting the weight value for each type of wave. After incorporating all of the data into our model, we obtained a weighted accuracy of 96.03% while Random Forest was used in clarifying the final simulation.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Hasnul, M.A., Aziz, N.A.A., Alelyani, S., Mohana, M. and Aziz, A.A.J.S. (2021) Electrocardiogram-Based Emotion Recognition Systems and Their Applications in Healthcare—A Review. *Sensors*, **21**, Article No. 5015. <https://doi.org/10.3390/s21155015>
- [2] Acharya, D., Goel, S., Bhardwaj, H., Sakalle, A. and Bhardwaj, A. (2020) A Long Short Term Memory Deep Learning Network for the Classification of Negative Emotions Using EEG Signals. 2020 *International Joint Conference on Neural Networks (IJCNN)*, Glasgow, 19-24 July 2020, 1-8. <https://doi.org/10.1109/IJCNN48605.2020.9207280>
- [3] Sharma, R., Pachori, R.B. and Sircar, P. (2020) Automated Emotion Recognition Based on Higher Order Statistics and Deep Learning Algorithm. *Biomedical Signal Processing and Control*, **58**, Article ID: 101867. <https://doi.org/10.1016/j.bspc.2020.101867>
- [4] Nath, D., Singh, M., Sethia, D., Kalra, D. and Indu, S. (2020) An Efficient Approach to EEG-Based Emotion Recognition Using LSTM Network. 2020 *16th IEEE International Colloquium on Signal Processing Its Applications (CSPA)*, Langkawi, 28-29 February 2020, 88-92.
- [5] Colvonen, P.J., Glassman, L.H., Crocker, L.D., Buttner, M.M., Orff, H., Schiehser, D.M., *et al.* (2017) Pretreatment Biomarkers Predicting PTSD Psychotherapy Outcomes: A Systematic Review. *Neuroscience & Biobehavioral Reviews*, **75**, 140-156. <https://doi.org/10.1016/j.neubiorev.2017.01.027>
- [6] Feidakis, M., Daradoumis, T. and Caballe, S. (2011) Emotion Measurement in Intelligent Tutoring Systems: What, When and How to Measure. *Proceedings of the 2011 3rd International Conference on Intelligent Networking and Collaborative Systems*, Fukuoka, 30 November-2 December 2011, 807-812. <https://doi.org/10.1109/INCoS.2011.82>
- [7] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P. (2002) Smote: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*,

- 16, 321-357. <https://doi.org/10.1613/jair.953>
- [8] Lu, N., Wu, Y., Feng, L. and Song, J. (2018) Deep Learning for Fall Detection: Three-Dimensional CNN Combined with LSTM on Video Kinematic Data. *IEEE Journal of Biomedical and Health Informatics*, **23**, 314-323. <https://doi.org/10.1109/JBHI.2018.2808281>
- [9] Gandhi, V., Prasad, G., Coyle, D., Behera, L. and McGinnity, T.M. (2014) Quantum Neural Network-Based EEG Filtering for a Brain-Computer Interface. *IEEE Transactions on Neural Networks and Learning Systems*, **25**, 278-288. <https://doi.org/10.1109/TNNLS.2013.2274436>
- [10] Campisi, P. and Rocca, D.L. (2014) Brain Waves for Automatic Biometric Based User Recognition. *IEEE Transactions on Information Forensics and Security*, **9**, 782-800. <https://doi.org/10.1109/TIFS.2014.2308640>
- [11] Cecotti, H. and Gräser, A. (2011) Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 433-445. <https://doi.org/10.1109/TPAMI.2010.125>
- [12] Iwasa, H., Horie, T. and Matsuyama, Y. (2016) Verification of Fraudulent PIN Holders by Brain Waves. 2016 *International Joint Conference on Neural Networks (IJCNN)*, Vancouver, 24-29 July 2016, 2068-2075. <https://doi.org/10.1109/IJCNN.2016.7727454>
- [13] Xu, H.Y., Lou, J., Su, R.D. and Zhang, E.P. (2009) Feature Extraction and Classification of EEG for Imaging Left-Right Hands Movement. *International Conference on Computer Science and Information Technology*, Kiev, 25-26 July 2009, 56-59. <https://doi.org/10.1109/ICCSIT.2009.5234611>
- [14] Chai, R., Ling, S.H., Hunter, G.P., Tran, Y. and Nguyen, H.T. (2013) Brain-Computer Interface Classifier for Wheelchair Commands Using Neural Network with Fuzzy Particle Swarm Optimization. *IEEE Journal of Biomedical and Health Informatics*, **18**, 1614-1624. <https://doi.org/10.1109/JBHI.2013.2295006>
- [15] Steyrl, D., Scherer, R., Förstner, O. and Müller-Putz, G. (2014) Motor Imagery Brain-Computer Interfaces: Random Forests vs Regularized LDA-Non-Linear Beats Linear. *6th International Brain-Computer Interface Conference*, Graz, 16-19 September 2014.
- [16] Aeby, A., Santalucia, R., Van Hecke, A., Nebbioso, A., Vermeiren, J., Deconinck, N., De Tiège, X. and Van Bogaert, P. (2021) A Qualitative Awake EEG Score for the Diagnosis of Continuous Spike and Waves during Sleep (CSWS) Syndrome in Self-Limited Focal Epilepsy (SFE): A Case-Control Study. *Seizure*, **84**, 34-39. <https://doi.org/10.1016/j.seizure.2020.11.008>
- [17] Tang, H., Liu, W., Zheng, W.-L. and Lu, B.-L. (2017) Multimodal Emotion Recognition Using Deep Neural Networks. *24th International Conference, ICONIP 2017*, Guangzhou, 14-18 November 2017, 811-819. https://doi.org/10.1007/978-3-319-70093-9_86
- [18] Zheng, W.-L., Zhu, J.-Y. and Lu, B.-L. (2019) Identifying Stable Patterns over Time for Emotion Recognition from EEG. *IEEE Transactions on Affective Computing*, **10**, 417-429. <https://doi.org/10.1109/TAFFC.2017.2712143>
- [19] Li, X., Song, D.W., Zhang, P., Yu, G.L., Hou, Y.X. and Hu, B. (2016) Emotion Recognition from Multi-Channel EEG Data through Convolutional Recurrent Neural Network. 2016 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Shenzhen, 15-18 December 2016, 352-359.
- [20] Tripathi, S., Acharya, S., Sharma, R.D., Mittal, S. and Bhattacharya, S. (2017) Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on

- DEAP Data. *Proceedings of the AAAI Conference on Artificial Intelligence*, **31**, 4746-4752. <https://doi.org/10.1609/aaai.v31i2.19105>
- [21] Al-Nafjan, A., Hosny, M., Al-Wabil, A. and Al-Ohali, Y. (2017) Classification of Human Emotions from Electroencephalogram (EEG) Signal Using Deep Neural Network. *International Journal of Advanced Computer Science and Applications*, **8**, 419-425. <https://doi.org/10.14569/IJACSA.2017.080955>
- [22] Rayatdoost, S. and Soleymani, M. (2018) Cross-Corpus EEG-Based Emotion Recognition. 2018 *IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, Aalborg, 17-20 September 2018, 1-6. <https://doi.org/10.1109/MLSP.2018.8517037>
- [23] Cao, G.L., Ma, Y.L., Meng, X.F., Gao, Y.Y. and Meng, M. (2019) Emotion Recognition Based on CNN. 2019 *Chinese Control Conference (CCC)*, Guangzhou, 27-30 July 2019, 8627-8630. <https://doi.org/10.23919/ChiCC.2019.8866540>
- [24] Cimtay, Y., Ekmekcioglu, E. and Caglar-Ozhan, S. (2020) Cross-Subject Multimodal Emotion Recognition Based on Hybrid Fusion. *IEEE Access*, **8**, 168865-168878. <https://doi.org/10.1109/ACCESS.2020.3023871>
- [25] Cimtay, Y. and Ekmekcioglu, E. (2020) Investigating the Use of Pretrained Convolutional Neural Network on Cross-Subject and Cross-Dataset EEG Emotion Recognition. *Sensors*, **20**, Article No. 2034. <https://doi.org/10.3390/s20072034>
- [26] Luo, Y.L., Fu, Q., Xie, J.T., Qin, Y.B., Wu, G.P., Liu, J.X., Jiang, F., Cao, Y. and Ding, X.M. (2020) EEG-Based Emotion Classification Using Spiking Neural Networks. *IEEE Access*, **8**, 46007-46016. <https://doi.org/10.1109/ACCESS.2020.2978163>
- [27] Gao, Z.K., Wang, X.M., Yang, Y.X., Li, Y.L., Ma, K. and Chen, G.R. (2021) A Channel-Fused Dense Convolutional Network for EEG-Based Emotion Recognition. *IEEE Transactions on Cognitive and Developmental Systems*, **13**, 945-954. <https://doi.org/10.1109/TCDS.2020.2976112>
- [28] Acharya, D., Goel, S., Bhardwaj, H., Sakalle, A. and Bhardwaj, A. (2020) A Long Short Term Memory Deep Learning Network for the Classification of Negative Emotions Using EEG Signals. 2020 *International Joint Conference on Neural Networks (IJCNN)*, Glasgow, 19-24 July 2020, 1-8. <https://doi.org/10.1109/IJCNN48605.2020.9207280>
- [29] Andrzejak, R.G., Lehnertz, K., Mormann, F., Rieke, C., David, P. and Elger, C.E. (2001) Indications of Nonlinear Deterministic and Finite-Dimensional Structures in Time Series of Brain Electrical Activity: Dependence on Recording Region and Brain State. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, **64**, Article ID: 061907. <https://doi.org/10.1103/PhysRevE.64.061907>
- [30] Rayatdoost, S. and Soleymani, M. (2018) Cross-Corpus EEG Based Emotion Recognition. 2018 *IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, Aalborg, 17-20 September 2018, 1-6. <https://doi.org/10.1109/MLSP.2018.8517037>
- [31] Nanthini, B.S. and Santhi, B. (2015) Different Approaches to Analysing EEG Signals for Seizure Detection. *International Journal of Signal and Imaging Systems Engineering*, **8**, 28-38. <https://doi.org/10.1504/IJSISE.2015.067066>
- [32] Chawla, N., Bowyer, K., Hall, L. and Kegelmeyer, P. (2000) SMOTE: Synthetic Minority Over-Sampling Technique. In: *International Conference of Knowledge Based Computer Systems*, Allied Press, Mumbai, 46-57.
- [33] Krawczyk, B. (2016) Learning from Imbalanced Data: Open Challenges and Future Directions. *Progress in Artificial Intelligence*, **5**, 221-232. <https://doi.org/10.1007/s13748-016-0094-0>