Classification of Human Emotion Using DT-SVM Algorithm with Enhanced Feature Selection and Extraction

Adithya Mohanavel
Department of Computer Science & Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu.
E-mail: adithya300800@gmail.com

Dinesh Ram Danaraj
Department of Computer Science & Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu.
E-mail: dineshramdanaraj@gmail.com

Dr.N Sabiyath Fatima
Associate Professor, Department of Computer Science & Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu.
E-mail: sabiyathfathima@crescent.education

Received September 19, 2021; Accepted December 16, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V19I1/WEB19233

Abstract

Emotions are a basic component of human life. It generates different brain waves for emotions such as happiness, sadness, anger, calmness, tension, excitement, etc. The brain waves are electric and their electric impulse can be measured and recorded as a continuous stream of data. These emitted brain waves are recorded using an EEG device. Many existing systems are in use that feeds the recorded data into various Machine learning algorithms to classify the emotions. These systems are huge and complex, thus require a great amount of time for initializing and working. While a lot of algorithms are used and new algorithms are discovered to classify Brain EEG data, most of the time results will be improper and will not be reliable. The proposed system extracts only the data which corresponds to Human-emotions from the continuous stream of EEG data. The system makes use of robust preprocessing algorithms like ANOVA and PCA for feature extraction and selection to identify and extract features associated with Human-emotion. Later, these recording signals are modeled and fed into Dynamic Time wrapping Simple vector machine (DT-SVM) classification algorithm to analyze and predict the emotion of the person during the experiment which produces an improved accuracy of 99.2%compared to existing system.

Keywords

ANOVA, PCA, EEG Classification, Dynamic Time Wrapping SVM, Time Series Algorithm, Emotion Classification.

Introduction

Emotions are psychological states brought about by complex neurological changes. They are so much advanced than our contemporary science that it still doesn't have a scientific consensus definition. The lack of understanding of this subject has not stopped us from using it. The state of our emotion defines our choice of vocabulary, our food choices, and much more. Emotion is not only observed in intelligent species like humans but is also widely found in many other species in the animal kingdom. This prevalent and crucial part of life is still a mystery to our scientific understanding. Among the many scientific advances in the twentieth century is the invention of Electroencephalography (EEG). This is still the widely sought out technique to study our human brain. Its non-invasive approach makes it easy for researchers to experiment with their theory on live human subjects without any side effects. This technique has evolved into being more precise by adapting new algorithms to eliminate stray signals and focus on the signal of interest.

The human brain emits brain waves that are typically measured by a certain methodology called the 10-20 system and it is given in Fig1. This is an internationally recognized testing method for standardizing testing methods for capturing brain waves. The 10-20 refers to the distance between the electrodes on the right-left or front-back of the human head. The points located are Occipital (O), Frontal (F), Parietal (P), Temporal (T), and Central (C). Numbers ranging from one to twenty are also given. Based on the location the electrode is placed in the scalp. Since the human brain is divided into three major units' forebrain, the mid brain, and the hindbrain the readings from the marked electrodes can tell which part of the brain is active on the task.

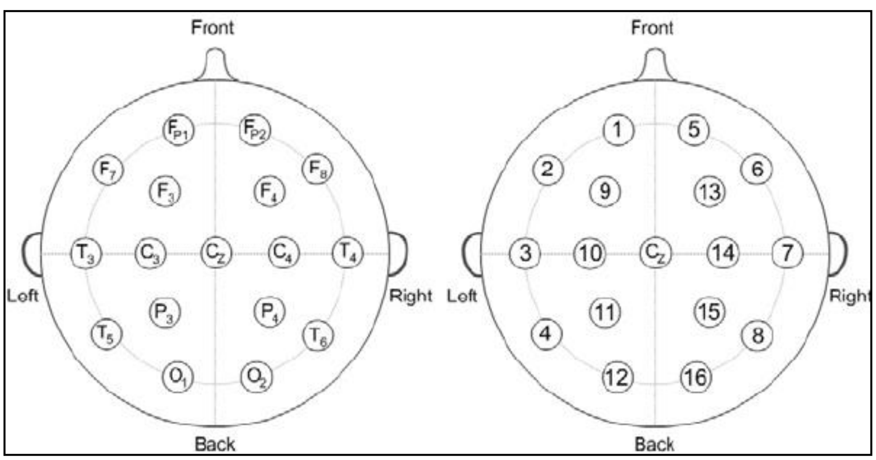

Fig. 1 10-20 EEG system

The wave signal emitted by the human brain is called brainwaves. They are produced by the synchronized electrical pulse from the vastly interconnected neurons when communicating with each other. When our human body is in the process of any task these neurons communicate with each other to bring about the actions related to the task. Analyzing brainwaves, researchers have classified five major types based on their bandwidths. They are Alpha waves, Beta waves, Gamma waves, Delta waves and Theta Waves respectively. All these waves are measured in hertz. Each of the waves is associated with certain human actions or behaviors.

Decoding these EEG signals and assigning them to human actions have improved our understanding of the human brain. The EEG dataset in general is a data mine, it contains huge amounts of data than just to identify a few human characteristics. Our existing instruments are sensitive enough to capture the slightest of signal change in EEG, but what remains is the tedious task of assigning them to all the human emotions and behaviors exhibited by humans. In this experiment, we try to classify human emotions by the use of a well-developed data classification algorithm. We have used IoT devices to sense the Electric impulse generated by the brain and transmit these captured EEG datasets to the system where the algorithm runs to classify the signals of our interest. In order to process and classify the emotion of a person, the algorithm has to identify the particular part of data responsible for human emotion and extract the same from the continuous stream of data which is done by ANOVA Feature selection and PCA Feature extraction. This process of feature selection and extraction is done mainly because our brain is the most complex and advanced computer ever created by nature which processes a lot of things. The recorded readings from the EEG device do not just contain data about mental state and emotion but a lot more than that. Later these processed data will be classified using a Time Series algorithm called Dynamic Time Wrapping to classify the emotion. The overall algorithm proposed in this paper is referred as Dynamic Time wrapping Simple vector Machine (DT-SVM).

With mental health being a serious issue in modern society especially in teenage kids this device can be used to monitor their mental state and help them relieve the depression or anxiety they might be facing. This also can play a crucial role in saving the lives of people who are suicidal after all, a friendly word or statement of encouragement at the right time will make the difference between life or death in their case.

Related Works

Hiram Cavlo et.al used Fast algorithm Dynamic Time Wrap (FDTW) to feature select in sets of EEG records, and found out that LPC, PCA, and ICA to be less accurate than FDTW in terms of classification accuracy.

Steinn Gudmundsson et. al have worked on using Support Vector Machines (SVM) & Dynamic Time Warping (DTW) for the Time Series.

Martin Dinov et. al completed their research on using the Dynamic Time Warping spectrum to predict brain rest state and compared their results with the standard measures of brain dynamics.

H C Huang et.al worked on analyzing EEG waveform with dynamic time warping (DTW) and concluded that (DTW) provided more yields more homogeneous clusters than approaches based on computing peak aligned difference between waveforms and, approaches that were based off feature extracted from the waveforms.

Dinvo Martin released a scientific paper where he has detailed his findings on using Dynamic Time Warping for finding significant changes in FRMI data sets & for quantifying effects in EEG sinusoidal oscillations.

U Rajendra Acharya et. al worked on automating the classification of signals that prove epilepsy in the patient by live monitoring EEG dataset.

T.K Padma Shri et. al paper uses ANOVA graded approximate entropy and classifiers to distinguish EEG signals that can potentially identify people that consume alcohol.

Logan T Trujillo et. al in their paper suggests that Laplacian-transform to be used for computation of EEG Data on scalp level CI(X) and I(X), citing advantages such as high accuracy and reduction of volume conduction in EEG signal quality.

V.M. Coehlo et. al uses time series algorithms to perform on EEG data. These latest time-series models can further be improved by using feature selection like PCA and ANOVA to get much better results.

Huiping Jiang et. al [10] used Long short-term memory as EEG data are in time-series format. But the author’s proposed paper does not use the data pre-processing method to the existing methods or even for the implemented algorithms thus, resulting in a lesser accuracy and processing speed.

Md. Asadur Rahman et. al paper implies that PCA combined with t testing will improve the classification models such as SVM, etc. However, the t-test can only differentiate 2 populations while we are using low alpha, low beta, high alpha, and high beta, in our model. We need an algorithm that is capable of doing such and ANOVA provides such a feature to distinguish between 2 or more populations.

S.S. Lekshmi et. al proposed system which suggests that PCA is clubbed with wavelet transform to predict EEG as per the paper. As wavelet transform is more of an advanced version of Fourier series where it is rather localized in time to record short intervals. This results in a much better outcome however; we can improve it by using statistical testing like ANOVA and then let the pre-processed data run into a time series where it can work well on such short interval cases and then deploy into classification algorithms like SVM.

K. Lugger et. al propose that PCA results are deployed into LDA to process high variance principal components. We can instead use Analysis of Variance (ANOVA) rather than linear discriminant analysis as our system depends on four features to determine the emotions rather than one.

Varsha Harpaleab et. al proposed a system that follows a similar route to the solution by using wavelet transform to improve the classification algorithm limiting its accuracy to 96%. Whereas the proposed system using PCA and ANOVA with time series gives better accuracy than that of the wavelet transform.

System Design

1) System Architecture

The system collects the EEG data using an electrode on EEG headgear that can precisely sense EEG signals from the volunteer. The overall flow of the proposed system as shown in Fig.2 where the collected signals are subsequently sent to computers through Bluetooth, where data will be converted to CSV format and begins preprocessing those data. The results are achieved by removing incomplete time rows and frequencies below 8 HZ and over 30 HZ during the process of data preprocessing. Following preprocessing, the ANOVA algorithm is used to pick features, and the PCA algorithm is used to extract the data. Later these processed data are classified using a Dynamic time wrapping algorithm with SVM classification to determine the emotional state of the person. The findings are then exhibited after evaluating.

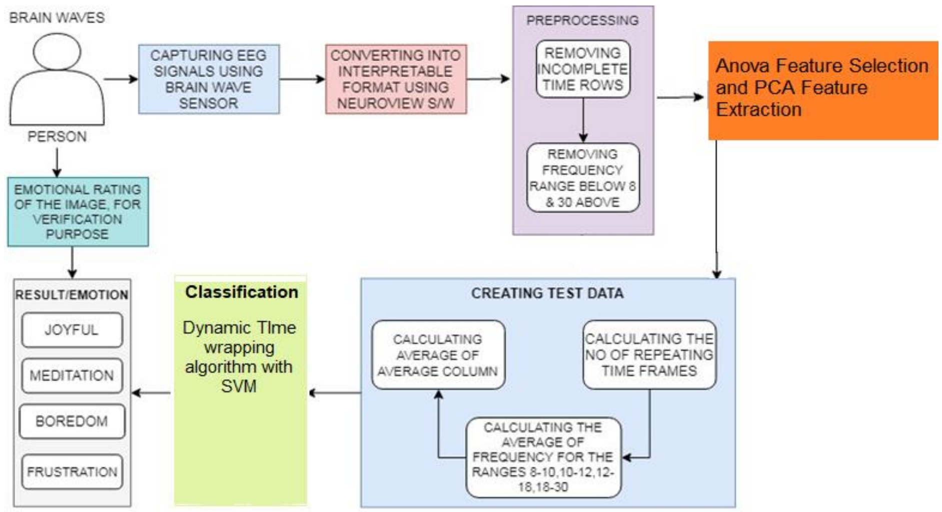

Fig. 2 System Architecture

2) Collection of Data

In order to collect data, we used the Neuro sky headgear with TGAM sensor for EEG signal collection whose flow is shown in Fig.3. We used it not only because of its higher sensitivity but also because of its one-node design, which requires the user to merely stick the node near the left eyebrow and clip one electrode in the left ear. The data is collected by asking volunteers of different age to observe the images that are either calming, revolting images or graphics that depicts a real-world situation that arouses different emotions in the observer. These are standard images used by the University of Stanford which classified these images set on display into several emotions. The collected data of the observer through the sensors are then transmitted to the system in real-time through Bluetooth and saved. Once the EEG data of the observer is transmitted to the system, Neuro sky software converts those EEG data set to a CSV file to allow our algorithms to begin preprocessing the information.

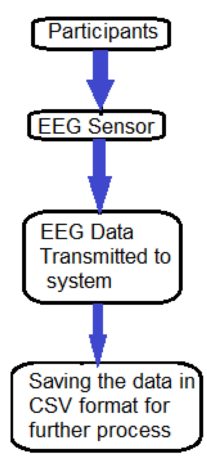

Fig. 3 Flow of collection of data

3) Data Pre-Processing

After the collection of data, it will be stored in CSV file format as shown in Fig 4, the preprocessing is applied to this data to filter it down to the data that is needed by the machine learning algorithms for prediction. The preprocessing is done to find the mean of the Brainwave generated at that second of instance on rows. And also, the mean is applied for the columns to find the mean of different waves emitted during the process. So, at the end of the preprocessing stage, the resultant will be the values of each time frame and the value of each wave listed below. These values obtained are the mean value at that time frame.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Time	0	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5
2	15:20:51	95.5	1.99E+03	490	2.19E+03	1.90E+03	2.41E+03	631	4.07	624	756	3.05
3	15:20:51	233	403	3.41	1.49E+03	2.55E+03	1.99E+03	1.29E+03	218	714	580	223
4	15:20:52	4.43	261	106	1.68E+03	1.39E+03	940	1.88E+03	1.08E+03	509	282	853
5	15:20:52	0.5	670	720	1.57E+03	492	174	1.87E+03	1.79E+03	392	176	1.21E+03
6	15:20:52	270	878	1.18E+03	1.02E+03	199	189	1.34E+03	1.42E+03	510	287	922
7	15:20:52	353	964	661	573	123	226	909	449	632	161	540
8	15:20:52	697	1.34E+03	63.7	367	30.4	107	822	27.3	495	16.9	262
9	15:20:52	909	2.38E+03	18.6	139	275	230	1.45E+03	804	528	23.8	357
10	15:20:52	316	4.16E+03	904	1.32E+03	1.79E+03	1.62E+03	2.42E+03	1.61E+03	956	366	1.09E+03
11	15:20:52	2.60E+03	8.27E+03	6.44E+03	5.79E+03	4.34E+03	3.26E+03	2.59E+03	1.38E+03	1.06E+03	1.03E+03	2.08E+03
12	15:20:53	3.18E+03	1.22E+04	1.67E+04	1.22E+04	4.09E+03	3.62E+03	1.98E+03	559	729	953	3.28E+03
13	15:20:53	3.74E+03	1.66E+04	1.73E+04	1.96E+04	1.70E+03	4.69E+03	1.45E+03	262	413	438	3.96E+03
14	15:20:53	52	1.43E+04	4.59E+03	2.78E+04	7.24E+03	7.53E+03	2.34E+03	63.2	331	1.06E+03	3.62E+03
15	15:20:53	3.44E+03	1.34E+04	4.51E+03	2.88E+04	1.86E+04	1.06E+04	3.37E+03	70.6	421	2.13E+03	2.75E+03
16	15:20:53	8.19E+03	1.61E+04	1.63E+04	2.08E+04	1.98E+04	1.05E+04	3.47E+03	312	443	2.12E+03	1.96E+03
17	15:20:53	7.38E+03	2.29E+04	1.72E+04	9.70E+03	1.06E+04	5.51E+03	3.14E+03	566	440	1.40E+03	984
18	15:20:53	1.36E+03	2.29E+04	5.82E+03	4.77E+03	2.65E+03	2.87E+03	4.11E+03	1.15E+03	653	800	212
19	15:20:53	6.71E+03	3.09E+04	1.18E+04	1.09E+04	1.19E+04	1.42E+04	1.02E+04	2.49E+03	111	1.47E+03	895
20	15:20:54	4.72E+03	3.03E+04	4.04E+04	3.45E+04	4.01E+04	3.49E+04	2.10E+04	5.42E+03	948	3.19E+03	2.79E+03
21	15:20:54	1.46E+03	2.73E+04	6.03E+04	6.15E+04	5.72E+04	4.28E+04	2.67E+04	8.67E+03	2.89E+03	4.82E+03	3.94E+03
22	15:20:54	6.65E+03	3.51E+04	5.50E+04	5.11E+04	4.99E+04	2.45E+04	2.19E+04	9.01E+03	2.81E+03	5.16E+03	2.89E+03
23	15:20:54	3.90E+03	3.17E+04	2.87E+04	1.91E+04	2.65E+04	4.77E+03	1.01E+04	6.72E+03	545	4.75E+03	870
24	15:20:54	676	1.73E+04	9.40E+03	2.92E+03	9.97E+03	68.7	2.37E+03	3.19E+03	645	4.12E+03	485
25	15:20:54	3.54E+03	5.77E+03	4.51E+03	6.69E+03	4.83E+03	595	1.02E+03	542	3.85E+03	2.89E+03	1.24E+03
26	15:20:54	1.14E+03	2.23E+03	5.76E+03	9.42E+03	3.40E+03	93.7	2.08E+03	1.16E+03	5.75E+03	1.77E+03	983
27	15:20:54	1.52E+03	5.03E+03	1.12E+04	6.67E+03	1.46E+03	2.37E+03	4.76E+03	5.35E+03	5.21E+03	1.03E+03	451
28	15:20:55	501	5.60E+03	1.27E+04	4.54E+03	296	5.59E+03	8.31E+03	8.19E+03	3.33E+03	550	792
29	15:20:55	469	6.55E+03	9.33E+03	2.75E+03	1.45E+03	7.00E+03	8.51E+03	6.16E+03	1.90E+03	141	1.09E+03
30	15:20:55	57.6	3.83E+03	4.95E+03	896	2.68E+03	6.83E+03	3.78E+03	2.90E+03	1.10E+03	31.7	625
31	15:20:55	1.55E+03	1.81E+03	1.82E+03	866	3.18E+03	3.89E+03	885	767	676	151	215
32	15:20:55	68.4	37.5	686	2.62E+03	3.92E+03	837	593	583	519	135	72.3
33	15:20:55	2.10E+03	1.72E+03	539	4.45E+03	3.69E+03	326	575	1.11E+03	580	236	201
34	15:20:55	1.56E+03	4.80E+03	2.72E+03	3.40E+03	2.78E+03	737	464	1.17E+03	600	299	347

	M	N	O	P	Q	R	S	T	U	V	W
1	2.75	3	3.25	3.5	3.75	4	4.25	4.5	4.75	5	5.25
2	509	593	568	125	21.2	19	293	515	75.8	22.3	12.8
3	569	585	791	140	44.1	11.5	198	327	133	10.2	34.4
4	548	203	898	261	109	160	96.3	171	223	5.27	50.8
5	655	239	723	289	203	371	117	186	236	38	87.2
6	984	560	407	134	205	447	320	265	147	83.6	101
7	999	584	147	25.5	83.2	374	401	275	40.5	64.9	83.5
8	570	328	30.4	2.06	39.2	272	186	276	0.223	20.5	40.5
9	237	330	80.2	158	253	442	207	383	8.89	19.4	25.2
10	833	880	879	788	1.04E+03	1.01E+03	719	604	96.4	112	229
11	1.89E+03	1.67E+03	2.13E+03	1.79E+03	1.66E+03	1.52E+03	1.09E+03	671	269	239	543
12	2.09E+03	1.34E+03	3.38E+03	1.93E+03	1.27E+03	1.54E+03	657	606	216	332	596
13	1.65E+03	562	3.94E+03	1.11E+03	596	1.04E+03	67.7	514	16.1	336	289
14	1.13E+03	641	3.77E+03	873	293	473	141	529	117	232	38.5
15	564	1.12E+03	3.09E+03	1.61E+03	148	179	570	630	410	138	46.3
16	79.5	810	2.16E+03	1.87E+03	124	137	562	718	373	117	81.9
17	395	351	1.06E+03	1.60E+03	185	103	337	455	209	56	45.9
18	1.35E+03	252	853	1.09E+03	132	38.3	193	80.8	149	28.7	0.154
19	1.49E+03	493	1.17E+03	968	36.2	80.3	239	53.5	339	176	118
20	953	454	1.20E+03	1.22E+03	339	328	289	355	722	686	461
21	538	275	502	1.54E+03	898	329	286	521	1.01E+03	1.13E+03	845
22	764	460	4.61	1.82E+03	1.07E+03	140	147	376	910	918	890
23	1.15E+03	882	455	1.97E+03	1.07E+03	54.7	8.88	206	438	404	524
24	1.16E+03	1.14E+03	728	1.81E+03	1.37E+03	331	294	147	291	122	217
25	1.41E+03	789	483	1.08E+03	1.48E+03	809	875	549	616	252	63.9
26	2.17E+03	226	204	370	1.09E+03	695	1.27E+03	1.32E+03	980	405	0.26
27	2.55E+03	594	145	262	573	139	1.31E+03	1.44E+03	951	424	57.5
28	2.43E+03	1.20E+03	279	369	243	41.9	899	870	453	402	75.8
29	1.94E+03	1.04E+03	400	251	123	140	534	309	86.3	287	24.6
30	1.19E+03	474	243	100	60.5	62.3	269	179	54.5	106	47.6
31	472	144	90	22.5	11.1	0.916	111	143	81.7	17.1	102
32	253	204	44.9	8.79	6.45	24.5	51.9	47	79.2	8.56	112
33	274	265	59	0.439	13.7	39.6	27.6	10.1	58.8	47.3	100
34	245	223	46.1	6.42	26.8	24	25.6	23.5	29.7	62.6	64.5

Fig. 4 Sample EEG data as a CSV format

With reference to the values generated on rows and columns, Table 1 with average values of the waves is obtained. The preprocessed data is further narrowed down to High beta, Low beta, High alpha and low alpha waves. From the values that are obtained, it will be further processed into feature selection and feature extraction process; later Dynamic time wrapping algorithm with SVM classification is applied to it to make for predicting the emotion. These particular waves after feature selection and extraction are fed into the machine learning algorithms.

Table 1 Waves and Bandwidth Range

WAVES	BANDWIDTH
Delta waves	0-4 W/Hz
Theta waves	4-8 W/Hz
Low Alpha waves	8-10 W/Hz
High Alpha waves	10-12 W/Hz
Low Beta waves	12-18 W/Hz
High Beta waves	18-30 W/Hz
Low Gamma waves	30-50 W/Hz
High Gamma waves	50-64 W/Hz

4) Feature Selection and Feature Extraction

The human brain is the most complex and powerful computer that processes a lot of things in the mind. Thus, the EEG data that is being recorded from the brain has more than one feature which is associated with different thoughts and actions. In order to classify human emotion, one has to identify the EEG data that is associated with that feature and then apply a classification algorithm. If not done, there would be redundancy and outliers which could cause digression in our results. We use feature selection to avoid overfitting of the data so that our model does not premeditate any of the EEG data extracted from our brain's electric wave.

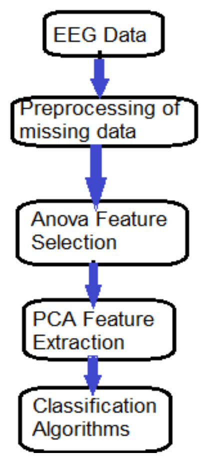

Fig. 5 Flow of EEG data

As the data are extracted from the electric impulse of the waves, the dimensionality would raise a major hurdle for our models. Thus, we extract the data by reducing its dimensionality to process with better and quicker results. Therefore, both feature selection and feature extraction by dimensional reduction help in improving the model. However, even though they engage in similar objectives they are not the same. In the selection process, the ANOVA algorithm removes certain attributes to make it simpler, while with the PCA feature extraction using dimension reduction, it groups a few attributes to reduce the complexity. This is why 2 different algorithms are used namely ANOVA (analysis of variance) and PCA (principal component analysis) respectively as shown in Fig 5.

- **ANOVA Feature Selection**

In order to select the appropriate feature, an ANOVA algorithm is being applied to the different EEG data waves. ANOVA is nothing but the analysis of variance in the wave format EEG data. 8 one-way- ANOVA is applied to the data to analyze the input and select the feature associated with Emotion. The reason for selecting 8-way ANOVA is because the EEG data collected would be of different brain waves, which would be our quantitative component, however, we require our waves to be part of 8 categorical values as mentioned in the data pre-processing. Thus, we use 8 one-way ANOVA for our model. The ANOVA tests the relationship between continuous data of different waves.

As we use 8 one-way ANOVA, it tests the significance between all the 8 groups i.e., the different waves. One-way ANOVA creates n hypothesis where ‘n’ will be the number of groups or categories. In our case, it is 8, so we would have a null hypothesis case to be tested, which states that the mean of all the groups is equal. Where the group's mean is represented as ‘μ’. However, the one-way ANOVA is to find the significant variance so, there is a chance that there could be a significant difference between the means of any 2 categories. This results in, accepting the alternative hypothesis, that the means are not equal to each other. Thus, we eliminate the least significant group or EEG data by using ANOVA.

The major drawback of ANOVA is, that it does not state which group or type of EEG data has been eliminated if there were multiple lesser significant groups found. Thus, certain cases use post-hoc testing to determine the precise groups or types of EEG data. However, in our scenario, we are testing with the reaction or emotion captured at that instance. So, we do not require any irrelevant thoughts generated by people’s brains, so we marginalize the lesser significant thoughts and concentrate only on the sustainable reactions.

- **PCA Feature Extraction**

The selected features from AVONA are further processed with PCA. PCA is applied to the data as the second stage of feature selection and extraction. In PCA, it is being used in order to reduce dimensions of dataset into ‘K’ dimensions. Where k value is less than the number of dimensions in dataset ‘d’ (PCA: $k < d$). PCA uses matrix operations to achieve this. It also ensures better interpretability with less information loss in the data. By doing this, it has been made sure that the training and test data will have features that are only corresponds to Human emotions. PCA (Principal component analysis) in EEG data is used in order to reduce dimensions of large EEG dataset by minimizing variance by using squared deviation keeping the feature that are only corresponds to human emotion.

In our model, we use 8 principal components in order to get the best resulting output. This will lead to simplification of our complex dimensions involved in the data and give out standardized data which can be easily visualized, for us to understand the data. Furthermore, it also improves the processing time and accuracy of our model. PCA not only reduces the dimensions but also helps in improving one of the most powerful classification algorithms i.e., the support vector machine classification. PCA is used here, to obtain the Eigenvectors and Eigenvalues from its covariance matrix. After obtaining these Eigenvalues, it selects the required eigenvectors for the components (in our case:8) and associates them to the largest eigen variable for each dimension respectively in order to sort them from the least to the highest. Later, it transforms these selected eigenvectors into a projection matrix with the new dimension. Thus, it transforms our data into a new reduced form.

5) Classification

Once the feature selection and extraction process are over, the EEG data of the person is then processed further into the classification process where a Time series algorithm is applied to the data. It is this process where the emotion of the person is classified from the data.

- **Dynamic Time Wrapping Classification with SVM**

Dynamic Time wrapping algorithm is a Time series algorithm in which the collection of data points are made sequentially in time. The algorithm usually takes 2 different continuous linear data as shown in Fig. 6 and calculates the optimal match between those two temporal sequence data.

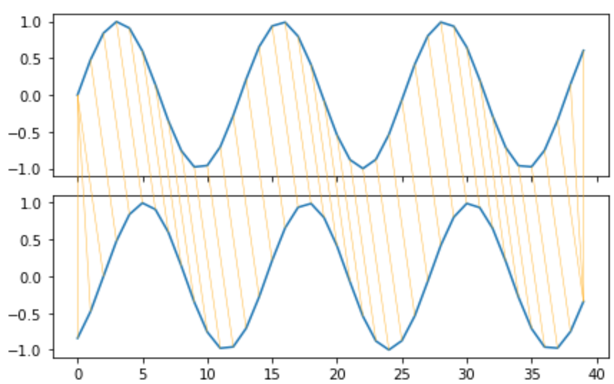

Fig. 6 Sample input data of DTW

Dynamic time warping is one such time series model in which it compares 2-time series data streams. It is considered as one of the most powerful time-series algorithm as it can match different Data streams irrespective of their time constraints or time period. As per our case, the emotions can be prolonged, for instance, happiness can be expressed for a moment or for a longer period. So DTW is used in order to identify the emotion patterns irrespective of their duration. As PCA produces high variance principal components which are countered by implying ANOVA. As it is able to distinguish between multiple populations and works on the variance analysis. It produces a statistical result which may not be effective on short intervals. That is, if there is a continuous sudden change in emotion then the results may not be effective. So, implementing Dynamic Time warping would enhance the result as the algorithm is based on time events or localized with time rather than statistical values.

In this algorithm, the optimal distance between the waveforms of the data in EEG will be used to classify human emotions. The dataset which is processed from the feature selection and extraction contains labeled time series along with different data such as alpha, beta, and gamma waves which is a feature vector with time/frequency domain variables. The DTW algorithm takes the EEG data from ANOVA and PCA as their two continuous linear input data and plots a matrix with the data point that has the optimal match between those two-input data. As the model is now independent of synchronization of time, the results can be more accurate and further clubbed with powerful classification methods like support vector machine, which would predict the right emotion based on the data.

The following Eq.1 formula can is used to process two sequential input data and calculate the optimal sequential data. While calculating the minDist. We have followed the Monotonicity condition, step size condition, and Boundary condition.

$$\mathrm{MinDist}(i, j) = \begin{cases} (A_1 - B_1)^2 & \text{For } i = j = 1 \\ \mathrm{minDist}(i', j') + \mathrm{Min}\begin{cases} (A_i - B_{j+1})^2 \\ (A_{i+1} - B_j)^2 \\ (A_{i+1} - B_{j+1})^2 \end{cases} & \end{cases} \quad (1)$$

Where (i’,j’) is the previous point on warping path.

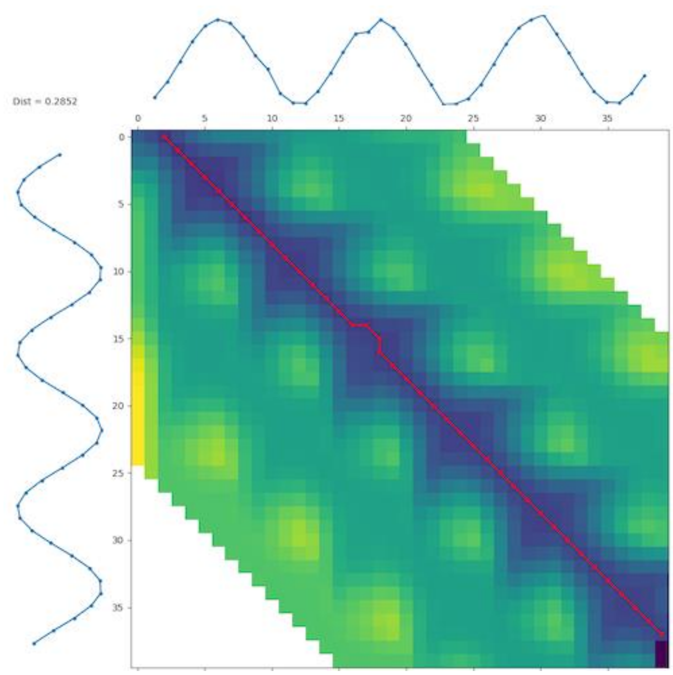

Fig. 7 DTW optimal data warping matrix

The calculated data which produces optimal warping paths as shown in Fig 7 will have same number of values and shape of a full-size matrix. These data are further passed down to SVM for classification of final results.

The Support vector machine algorithm will generate multiple hyperplanes, which will be used as the decision boundary for the algorithm to classify the data into different classes. Among various hyperplanes generated, based on the maximum distance between the data points of the different classes, the SVM will choose the best hyperplane as boundary. Different hyperplanes are generated by polynomial kernel, Radial basis function kernel & Linear kernel. In some cases, the hyperplanes of the SVM algorithm have to be tuned for better performance. The regularization parameter represents the misclassifications or error terms. The regularizations reveal how much error is bearable when the SVM algorithm is optimized. With regularization, it is possible to regulate the compromission between the misclassification terms and the decision boundary. In this system, the data from the optimal warping path of the Dynamic time wrapping algorithm is used to generate the SVM hyperplane data points. After the hyperplane's generation, based on the position of data points the SVM classifies the emotions. The Support Vector Machine algorithms offer better accuracy and perform rapidly when compared to other machine learning algorithms. The algorithms use less memory space for calculations and predictions. The algorithm is sensitive to the kernel that is being used and has poor performance when it has overlapping classes. To avoid the latter, we use PCA to impede the overlapping of classes to get accurate results.

Implementation

The proposed Dynamic Time wrapping Simple vector Machine (DT-SVM) algorithm is used to classify the emotions such as Neutral, Happy, Anxious, Boredom, etc. This proposed system will be implemented and will be evaluated with different metrics. In order to determine the efficiency of proposed algorithm is determined by using metrics such as sensitivity, accuracy, and precision.

i) Accuracy

The algorithms used are evaluated by classification accuracy of the algorithm. The accuracy for classification of an algorithm is given as Eq.2 which is formulated as ratio of the number of precise predictions to the total number of predictions made.

$$Accuracy = \frac{Number\ of\ precise\ predictions}{Total\ number\ of\ predictions\ made} \tag{2}$$

- **Accuracy before Feature Selection and Extraction**

On evaluating the accuracy of the proposed Dynamic Time wrapping Simple vector Machine (DT-SVM) algorithm applying feature selection and extraction, the results were obtained as per Fig 8.

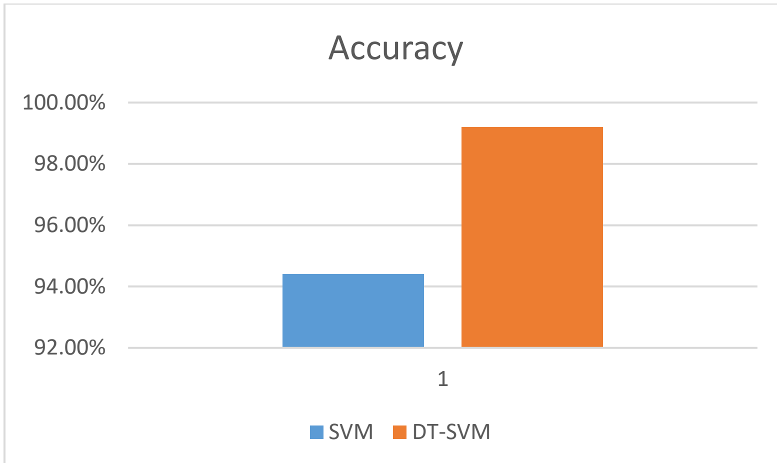

Fig. 8 Accuracy result before and after Feature selection & extraction

From Fig 8, It can be inferred that the accuracy of the existing Support vector machine algorithm is 94.4% while the proposed system of Dynamic Time wrapping Simple vector Machine (DT-SVM) classification with ANOVA and PCA feature selection & extraction has 99.2% accuracy.

ii) Sensitivity

Sensitivity for an algorithm is represented as Eq.3 which is defined as the actual positive that are precisely identified as positive to the sum of false-negative and the true positive values which make up the total number of positive in the model.

$$Sensitivity = \frac{True\ Positive}{(False\ Negative + True\ Positive)} \qquad (3)$$

- **Sensitivity before Feature Selection and Extraction**

On evaluating the sensitivity of the proposed Dynamic Time wrapping Simple vector Machine (DT-SVM) algorithm applying feature selection and extraction, the results were obtained as per Fig 9.

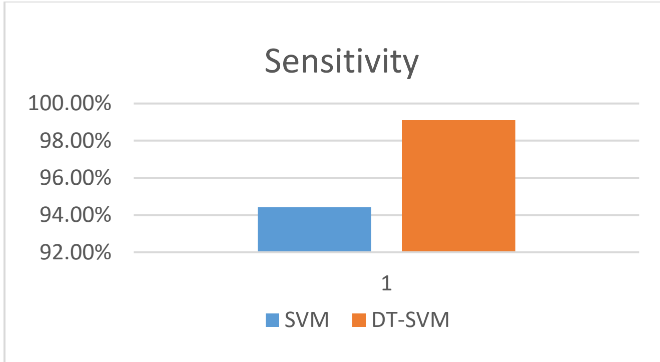

Fig. 9 Sensitivity result before and after Feature selection & extraction

From Fig 9, It can be inferred that the sensitivity of the existing Support vector machine algorithm is 94.4% while the proposed system Dynamic Time wrapping Simple vector Machine (DT-SVM) classification with ANOVA and PCA feature selection & extraction has 99.1% sensitivity.

iii) Precision

The precision will help identify how precise the proposed algorithm has worked on data and model. Precision for an algorithm is defined as the number of actual positive that are precisely identified as positive to the sum of false-positive and the true positive values in the model which is formulated as shown in Eq.4.

$$Precision = \frac{True\ Positive}{(True\ Positive + False\ Positive)} \quad (4)$$

- **Precision before Feature Selection and Extraction**

On evaluating the precision of the proposed Dynamic Time wrapping Simple vector Machine (DT-SVM) algorithm applying feature selection and extraction, the results were obtained as per Fig 10.

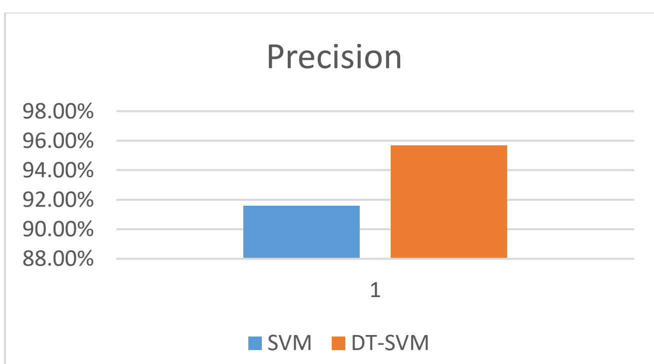

Fig. 10 Precision result before and after Feature selection & extraction

From Fig 10, It can be inferred that the precision of the existing Support vector machine algorithm is 91.6% while the proposed system of Dynamic Time wrapping Simple vector Machine (DT-SVM) classification with ANOVA and PCA feature selection & extraction has 95.7% precision.

Conclusion

It can be clearly observed from the results that the Accuracy, Precision, Sensitivity of the classification has been improved significantly after applying AVONA and PCA for feature selection and extraction. It is inferred that there is an increase in accuracy of around 5% in the algorithm after applying the feature selection and extraction, while the increase in Precision and Sensitivity is around 4% after applying the feature selection and extraction. The Dynamic Time wrapping SVM suits this system and shows high accuracy and efficiency.

In the future, the system can be innovated with better features selection to achieve accuracy close to 100. This system upon innovation can be applied for medical purposes to monitor patients' mental health and this can also be used to study victims while investigating. This system can also be helpful to monitor the astronaut's emotions and ensure they are mentally fit.

References

Calvo, H., Paredes, J.L., & Figueroa-Nazuno, J. (2016). Fast Dynamic Time Warping Feature Extraction for EEG Signal Classification. *In Fifteenth Mexican International Conference on Artificial Intelligence (MICAI)*, 146-151.

Gudmundsson, S., Runarsson, T.P., & Sigurdsson, S. (2008). Support vector machines and dynamic time warping for time series. *In IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence),* 2772-2776. http://doi.org/10.1109/IJCNN.2008.4634188

Dinov, M., Lorenz, R., Scott, G., Sharp, D.J., Fagerholm, E.D., & Leech, R. (2016). Novel modeling of task vs. rest brain state predictability using a dynamic time warping spectrum: comparisons and contrasts with other standard measures of brain dynamics. *Frontiers in computational neuroscience, 10*.

Huang, H.C., & Jansen, B.H. (1985). EEG waveform analysis by means of dynamic time-warping. *International journal of bio-medical computing, 17*(2), 135-144. http://doi.org/10.1016/0020-7101(85)90084-4

Martin, D. (2015). Using dynamic time warping for quantifying effects of sinusoidal oscillation deviations during EEG time series prediction and for finding interesting EEG and fMRI changes. *BMC Neuroscience*, *16*(1), 1-2.

Acharya, U.R., Sree, S.V., Alvin, A.P.C., & Suri, J.S. (2012). Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. *Expert Systems with Applications, 39*(10), 9072-9078. http://doi.org/10.1016/j.eswa.2012.02.040

Shri, T.P., Sriraam, N., & Bhat, V. (2014). Characterization of EEG signals for identification of alcoholics using ANOVA ranked approximate entropy and classifiers. *In International Conference on Circuits, Communication, Control and Computing,* 109-112.

Trujillo, L.T., Stanfield, C.T., & Vela, R.D. (2017). The effect of electroencephalogram (EEG) reference choice on information-theoretic measures of the complexity and integration of EEG signals. *Frontiers in neuroscience, 11*. http://doi.org/10.3389/fnins.2017.00425

Coelho, V.N., Coelho, I.M., Coelho, B.N., Souza, M.J., Guimarães, F.G., Luz, E.D.S., & Oliveira, T.A. (2017). EEG time series learning and classification using a hybrid forecasting model calibrated with GVNS. *Electronic Notes in Discrete Mathematics, 58,* 79-86. http://doi.org/10.1016/j.endm.2017.03.011

Jiang, H., Jiao, R., Wang, Z., Zhang, T., & Wu, L. (2021). Construction and Analysis of Emotion Computing Model Based on LSTM. *Complexity,* 2021, 1-12. http://doi.org/10.1155/2021/8897105

Rahman, M.A., Hossain, M.F., Hossain, M., & Ahmmed, R. (2020). Employing PCA and t-statistical approach for feature extraction and classification of emotion from multichannel EEG signal. *Egyptian Informatics Journal, 21*(1), 23-35. http://doi.org/10.1016/j.eij.2019.10.002

Lekshmi, S.S., Selvam, V., & Rajasekaran, M.P. (2014). EEG signal classification using principal component analysis and wavelet transform with neural network. *In International Conference on Communication and Signal Processing,* 687-690.

Lugger, K., Flotzinger, D., Schlögl, A., Pregenzer, M., & Pfurtscheller, G. (1998). Feature extraction for on-line EEG classification using principal components and linear discriminants. *Medical and Biological Engineering and Computing, 36*(3), 309-314.

Harpale, V., & Bairagi, V. (2021). An adaptive method for feature selection and extraction for classification of epileptic EEG signal in significant states. *Journal of King Saud University-Computer and Information Sciences, 33*(6), 668-676. http://doi.org/10.1016/j.jksuci.2018.04.014