

Stress detection from EEG using power ratio

Taruv Harshita Priya
School of Avionics, IST, JNTUK
Kakinada, India.
taruvharshitha@gmail.com

P. Mahalakshmi
School of Electrical Engineering,
VIT-Vellore, India.
mahalaksh@gmail.com

M. Srinivas
School of Avionics, IST, JNTUK
Kakinada, India.
mallimogla@gmail.com

VPS Naidu
Multi Sensor Data Fusion Lab,
CSIR-NAL, Bangalore, India.
vpsnaidu@gmail.com

Abstract— Stress correlates itself as a mental conscious and emotion within a person that influences mental ability and decision-making skills, which results in an inappropriate work. Studies have recently developed to detect the stress in a person while performing different tasks. One of the methods is through Electroencephalograph (EEG). These are the bioelectrical signals generated in a human body while performing the tasks and thus describes the activity of the brain. Any action taken by a person changes the properties of these signals. This present work focuses on the classification of baseline (relax) and stress detection using EEG sub-band power ratio as features. Support vector machine (SVM) classifier with different kernel function parameters and K-nearest neighbor (KNN) classifier with a different number of neighbors with holdout and 10-fold cross-validation technique were used to classify power ratio features in order to detect stress. To evaluate the classifier performance various performance metrics were used. It is observed that KNN with a number of neighbors as one, with Euclidean distance gives better performance in both validation techniques and also anterior frontal channel Fp1 that is placed at the left side of the brain itself gives a good accuracy of 99.42%. The performance of the proposed method is verified on a publicly available mental arithmetic dataset where stress is induced while performing the mental cognitive workload i.e., mental serial subtraction.

Keywords—Stress analysis, EEG sub-bands, SVM Classifier, KNN Classifier, Power ratio, Machine-learning

I. INTRODUCTION

A stressful situation triggers a cascade of stress hormones that produce well-orchestrated bioelectrical signals to change. These situations can also make the heart pound, muscles tense, sweat appearing and more [1]. It affects both physical and psychological health. People from various professions get affected by stress, which causes a serious problem resulting in life-threatening situations. Also, due to the globalization of the economy and growing diversity in the workplace, the workload is always increasing. A reliable and accurate measurement is needed in order to know about the individual work burden. It is always important to take some steps to relieve the burden. This can be done by various methods one of them is mentioned in literature [2]. Psychological stress is a relationship between the person and the environment. So, when a person does a task it triggers the response in the bioelectrical signals. Electroencephalograph (EEG) is an electrical response of brain activity that is recorded when electrodes and conductive paste is placed on the scalp. This paper deals with the EEG signals as a

bioelectrical signal source to classify the relax condition and stressful condition.

Ref.[3], proposed a method Genetic algorithm(GA), Principle component analysis (PCA) for feature selection of statistical, frequency domain, higher-order crossings, Hjorth parameter & frontal symmetry alpha features and acquired 71.76% of KNN classification accuracy using GA, acquired 65.3% of KNN classification accuracy using PCA for calm and stress analysis. Ref. [4], uses the Teager-Kaiser energy operator(TKEO) with a K-nearest neighbor (KNN), Neural network (NN), classification tree (CT) classifier and acquired 88.3%, 93%, 88.33% respectively for the Fp1 (α band). Ref. [5], performed stress analysis using the power ratio of frequency bands and suggested that alpha to beta and theta to the beta are correlated with stress and indicated that these two power ratio can be used to determined stress. In this present paper, for stress analysis, we are extending the above work by including more power ratio features and classifying using classifiers (SVM, KNN) along with performance evaluation metrics.

II. ELECTROENCEPHALOGRAPH-EEG

Brainwaves are the continuous electrical signal developed by the brain. To evaluate this electrical activity in the brain, an electroencephalogram (EEG) is used. EEG is a fixed range of wave frequencies and amplitudes over a time scale. EEG band has a frequency of about 1-100Hz. There are five types of EEG patterns (brain waves) include Delta waves 0-4Hz, Theta waves 4-8Hz, Alpha waves 8-14Hz, Beta waves 14-30Hz, Gamma waves ≥ 30 Hz. Where the amplitude of the wave implies the intensity of the activity and frequency of the wave implies the speed of activity. In earlier studies, it is known that the alpha rhythm is sensitive to qualitative changes in visual input. The power of alpha increases when the eyes are closed as the amplitude of the EEG increases. Similarly, the power of alpha decreases with the eyes open as the amplitude of the EEG signal decreases. This phenomenon is commonly observed in adults. Theta rhythms are observed during the transition of wakefulness to sleep. Delta rhythms observed when the person is meditating. Studies have also proven about the cause of increasing power is also due to the processing of emotional information and when performing memory loaded tasks. When slower EEG rhythms are dominant we can observe that the person is feeling tired, slow, sluggish or dreamy and higher EEG rhythms gets dominant when a person gets alert. However, gamma rhythms are the

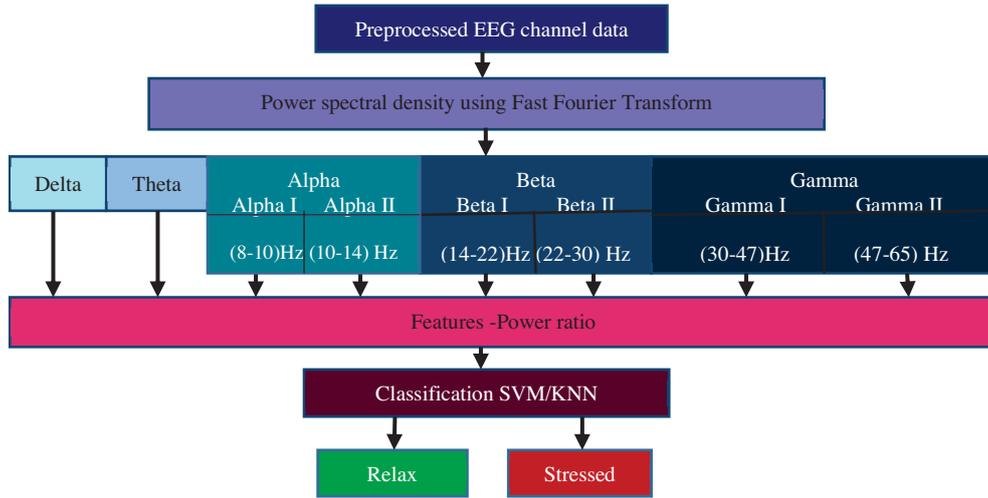


Figure 1: Block diagram for the proposed model

high-frequency waves and it is mainly related to the processing of information from different brain areas [6].

III. STRESS ANALYSIS

A. EEG Dataset

The EEG data used in the present work is taken from the physio bank [7, 8]. This dataset consists of EEG records of 36 subjects using the 10-20 international system. This data was recorded while performing cognitive mental workload; in this case, it is a complex serial subtraction. It contains 19 channels Anterior frontal (Fp1,Fp2), Frontal (F3,F4,F7,F8,Fz), Central (C3,C4,Cz), Parietal (P3,P4,Pz), Temporal (T3,T4,T5,T6) and Occipital (O1, O2) with the sampling rate of 500Hz. Every recording includes artifact-free EEG of 182sec for baseline state and 62sec for task performing state, which is a stressed state.

B. Preprocessing

The aim of this paper is to classify the EEG data whether it is relaxed or stress. It is a binary classification. Figure 1 shows the proposed method block diagram. The baseline data consists of 91000 (182sec X 500Hz) data points and the stress-generated data consists of 31000 (62sec X 500Hz) data points. Note that the data obtained is imbalanced data for both classes. So, only 62sec of the baseline is considered. Now as the data became balanced, it is segmented for every 2^{14} (16384 \approx 32.76sec) data points and with an overlap of about 2^{12} (4096 \approx 8.19sec) data points i.e., a 25% overlap.

C. Feature Extraction

Power spectral density (PSD) is computed from the preprocessed time-domain EEG signal $x(n)$ in order to know the functional characteristics of the signal using a fast Fourier transform (FFT) [9].

$$PSD(f) = 2 \times \frac{|FFT(x(n))|^2}{F_S \times N} \quad (1)$$

Where

N : Length of sequence $x(n)$
 F_S : Sampling frequency (Hz)

In equation 1, the factor '2' is multiplied for the purpose of the removal of negative energies. Here we do not consider this part of power spectral density. The computed PSD is divided into Delta (0-3.9)Hz, Beta (4-7.9)Hz, Alpha I (8-9.9)Hz, Alpha II (10-13.9)Hz, Beta I (14-21.9)Hz, Beta II (22-29.9)Hz and Gamma I (30-46.9)Hz, Gamma II (47-65)Hz bands. Using eq.2 to 9.

$$\text{The power in delta band} \quad P_{\Delta} = \sum_{f=0}^{3.9} PSD(f) \quad (2)$$

$$\text{The power in Theta band} \quad P_{\Theta} = \sum_{f=4}^{7.9} PSD(f) \quad (3)$$

$$\text{The power in Alpha I band} \quad P_{\alpha 1} = \sum_{f=8}^{9.9} PSD(f) \quad (4)$$

$$\text{The power in Alpha II band} \quad P_{\alpha 2} = \sum_{f=10}^{13.9} PSD(f) \quad (5)$$

$$\text{The power in Beta I band} \quad P_{\beta 1} = \sum_{f=14}^{21.9} PSD(f) \quad (6)$$

$$\text{The power in Beta II band} \quad P_{\beta 2} = \sum_{f=22}^{29.9} PSD(f) \quad (7)$$

$$\text{The power in Gamma I band} \quad P_{\gamma 1} = \sum_{f=30}^{46.9} PSD(f) \quad (8)$$

$$\text{The power in Gamma II band} \quad P_{\gamma 2} = \sum_{f=47}^{65} PSD(f) \quad (9)$$

In addition to these features, Fifty-six features are extracted as shown in Table 1.

D. Classifier

The purpose of the classifier is to develop a model using labeled input data (training data + labeled data), which predicts the output of the testing data which is given by the testing data attributes. In the present work, SVM and KNN classifiers are used.

1) Support Vector Machine(SVM)

SVM is a supervised machine-learning algorithm that tries to find suitable hyperplane, which maximizes the margin or the smallest distance. There are several SVM classifiers, which can be controlled by the varying kernel function. Some of the kernel functions are linear, polynomial, radial basis function (or Gaussian), sigmoid, Pearson VII, Fisher kernel, Text intersection, Bayesian Kernels and more. All these kernel functions can be controlled by using their kernel function parameters like Kernel scale (KS) which is a positive scalar value, Polynomial kernel function order (O) which is a positive integer and it is used for polynomial kernel, Box constraint, Kernel offset parameter etc. SVM is capable of performing both linear and non-linear classification [10].

Table 1: Power ratio features of EEG

Sl No.	Feature	Feature Description	Sl No.	Feature	Feature Description
F ₁	P _Δ	Power of delta band	F ₃₃	P _{α2} /P _{β1}	Power ratio of alpha II and beta I
F ₂	P _Θ	Power of theta band	F ₃₄	P _{α2} /P _{β2}	Power ratio of alpha II and beta II
F ₃	P _{α1}	Power of alpha-I band	F ₃₅	P _{α2} /P _{γ1}	Power ratio of alpha II and gamma I
F ₄	P _{α2}	Power of alpha-II band	F ₃₆	P _{α2} /P _{γ2}	Power ratio of alpha II and gamma II
F ₅	P _{β1}	Power of beta-I band	F ₃₇	P _{β1} /P _Δ	Power ratio of beta I and delta
F ₆	P _{β2}	Power of beta-II band	F ₃₈	P _{β1} /P _Θ	Power ratio of beta I and theta
F ₇	P _{γ1}	Power of gamma-I band	F ₃₉	P _{β1} /P _{α1}	Power ratio of beta I and alpha I
F ₈	P _{γ2}	Power of gamma-II band	F ₄₀	P _{β1} /P _{α2}	Power ratio of beta I and alpha II
F ₉	P _Δ /P _Θ	Power ratio of delta and theta	F ₄₁	P _{β1} /P _{β2}	Power ratio of beta I and beta II
F ₁₀	P _Δ /P _{α1}	Power ratio of delta and alpha I	F ₄₂	P _{β1} /P _{γ1}	Power ratio of beta I and gamma I
F ₁₁	P _Δ /P _{α2}	Power ratio of delta and alpha II	F ₄₃	P _{β1} /P _{γ2}	Power ratio of beta I and gamma II
F ₁₂	P _Δ /P _{β1}	Power ratio of delta and beta I	F ₄₄	P _{β2} /P _Δ	Power ratio of beta II and delta
F ₁₃	P _Δ /P _{β2}	Power ratio of delta and beta II	F ₄₅	P _{β2} /P _Θ	Power ratio of beta II and theta
F ₁₄	P _Δ /P _{γ1}	Power ratio of delta and gamma I	F ₄₆	P _{β2} /P _{α1}	Power ratio of beta II and alpha I
F ₁₅	P _Δ /P _{γ2}	Power ratio of delta and gamma II	F ₄₇	P _{β2} /P _{α2}	Power ratio of beta II and alpha II
F ₁₆	P _Θ /P _Δ	Power ratio of theta and delta	F ₄₈	P _{β2} /P _{β1}	Power ratio of beta II and beta I
F ₁₇	P _Θ /P _{α1}	Power ratio of theta and alpha I	F ₄₉	P _{β2} /P _{γ1}	Power ratio of beta II and gamma I
F ₁₈	P _Θ /P _{α2}	Power ratio of theta and alpha II	F ₅₀	P _{β2} /P _{γ2}	Power ratio of beta II and gamma II
F ₁₉	P _Θ /P _{β1}	Power ratio of theta and beta I	F ₅₁	P _{γ1} /P _Δ	Power ratio of gamma I and delta
F ₂₀	P _Θ /P _{β2}	Power ratio of theta and beta II	F ₅₂	P _{γ1} /P _Θ	Power ratio of gamma I and theta
F ₂₁	P _Θ /P _{γ1}	Power ratio of theta and gamma I	F ₅₃	P _{γ1} /P _{α1}	Power ratio of gamma I and alpha I
F ₂₂	P _Θ /P _{γ2}	Power ratio of theta and gamma II	F ₅₄	P _{γ1} /P _{α2}	Power ratio of gamma I and alpha II
F ₂₃	P _{α1} /P _Δ	Power ratio of alpha I and delta	F ₅₅	P _{γ1} /P _{β1}	Power ratio of gamma I and beta I
F ₂₄	P _{α1} /P _Θ	Power ratio of alpha I and theta	F ₅₆	P _{γ1} /P _{β2}	Power ratio of gamma I and beta II
F ₂₅	P _{α1} /P _{α2}	Power ratio of alpha I and alpha II	F ₅₇	P _{γ1} /P _{γ2}	Power ratio of gamma I and gamma II
F ₂₆	P _{α1} /P _{β1}	Power ratio of alpha I and beta I	F ₅₈	P _{γ2} /P _Δ	Power ratio of gamma II and delta
F ₂₇	P _{α1} /P _{β2}	Power ratio of alpha I and beta II	F ₅₉	P _{γ2} /P _Θ	Power ratio of gamma II and theta
F ₂₈	P _{α1} /P _{γ1}	Power ratio of alpha I and gamma I	F ₆₀	P _{γ2} /P _{α1}	Power ratio of gamma II and alpha I
F ₂₉	P _{α1} /P _{γ2}	Power ratio of alpha I and gamma II	F ₆₁	P _{γ2} /P _{α2}	Power ratio of gamma II and alpha II
F ₃₀	P _{α2} /P _Δ	Power ratio of alpha II and delta	F ₆₂	P _{γ2} /P _{β1}	Power ratio of gamma II and beta I
F ₃₁	P _{α2} /P _Θ	Power ratio of alpha II and theta	F ₆₃	P _{γ2} /P _{β2}	Power ratio of gamma II and beta II
F ₃₂	P _{α2} /P _{α1}	Power ratio of alpha II and alpha I	F ₆₄	P _{γ2} /P _{γ1}	Power ratio of gamma II and gamma I

2) K-Nearest Neighbors (KNN)

KNN is a supervised machine-learning algorithm that uses the similarity of features to predict the values of new data points. The selection of optimum algorithms depends on the sparsity of data, a number of neighbors (M) and a number of features. A data is classified by a maximum vote of its neighbors, which is measured by using a distance function. Some of the distance functions are Minkowski, Manhattan, Euclidean, Chebychev, cosine etc. Choosing a distance function and k-value depends upon the dataset and features used.

E. Validation Techniques

The validation technique is used to study the model. The most important significance of validation is to estimate an unbiased generalization performance. Figure 2 shows the classification of EEG power ratio features using the validation technique.

- **K-Fold cross-validation:** In this technique, the total dataset gets splits into K-folds. Here, each K-fold acts as testing data and remaining (K-1) folds acts as training

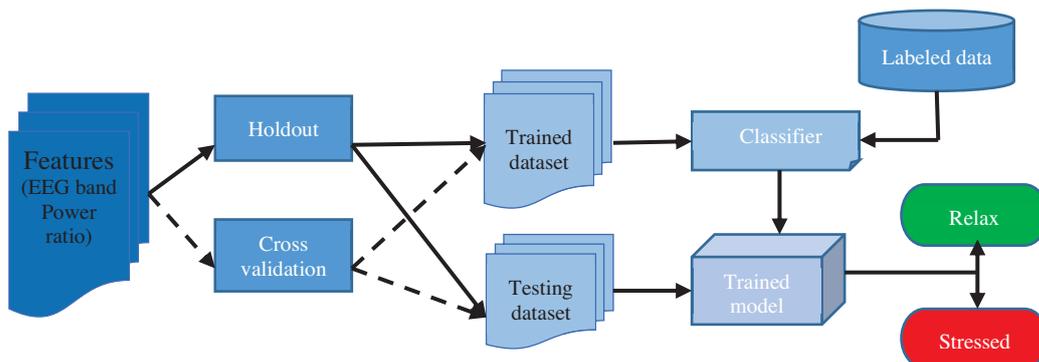


Figure 2: Classification of EEG Power ratio features using validation techniques

Table 2: Classifier Performance evaluation metrics

Accuracy(ACC)	$ACC = \frac{N_{TS} + N_{TR}}{Total}$
Misclassification rate (MCR)	$MCR = \frac{N_{FS} + N_{FR}}{Total}$
Sensitivity(SEN)/ True Positive rate/Recall	$SEN = \frac{N_{TS}}{N_{TS} + N_{FR}}$
False Positive rate (FPR)	$FPR = \frac{N_{FS}}{N_{TR} + N_{FS}}$
True Negative rate /Specificity(SPE)	$SPE = \frac{N_{TR}}{N_{TR} + N_{FS}}$
Positive Predictive value/Precision(PRE)	$PRE = \frac{N_{TS}}{N_{TS} + N_{FS}}$
Negative Predictive Value(NPV)	$NPV = \frac{N_{TR}}{N_{TR} + N_{FR}}$
Cohen's Kappa(CK)	$CK = \frac{P_o - P_e}{1 - P_e}$ <p>where</p> $P_o = \text{accuracy};$ $P_e = \frac{P_{yes} + P_{no}}{total}$ $P_{yes} = \frac{N_{TR} + N_{FS}}{total} * \frac{N_{TR} + N_{FR}}{total}$ $P_{no} = \frac{N_{FR} + N_{TS}}{total} * \frac{N_{FS} + N_{TS}}{total}$
F1 Score(F1)/F measure/ Sorensen- Dice coefficient (DSC)	$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$

data. So, this is done for K times for the given data. The advantage is that all the observations are treated as both training and validation. The value of K is an integer. In present paper, K=10 i.e., 10 Fold cross validation technique is used.

- **Holdout technique:** In this technique, the data is divided into a training set and testing set with the ratio of 60/40,75/25,80/90, etc., Here in 60/40 indicates that 60% of data is considered as a training set while the remaining 20% acts as the testing set. In the present paper, 75/25 holdout is used indicating 75% training data and 25% testing data.

There are also other validation techniques such as Leave one out, nested cross-validation, Time-series cross-validation, Re-substitution, Random sampling, Bootstrapping, Wilcoxon signed-rank test, McNemar's test, 5x2 CV paired t-test, 5x2 CV combined F-test.

F. Performance Evaluation

Performance evaluation is an important and challenging part of signal processing mainly in biomedical applications because it is complicated to accurately discriminate baseline performing subjects and task performing subjects. Here assessment cannot be done because performance cannot be evaluated by observing signals. However, the performance may be evaluated from its ability to correctly discriminate between its tasks. Table 2 describes the performance metrics along with description and formula.

Confusion Matrix: A confusion matrix is not a metric but a table that provides information about the classification performance and used to develop performance evaluation. The data obtained by the confusion matrix are:

- N_{TS} = the number of task performing subjects with a stressed result (True Stressed)
- N_{TR} = the number of baseline performing subjects with relaxing result (True Relax)
- N_{FR} = the number of task performing subjects with a relax result (False Relax) (Type II error)
- N_{FS} = the number of baseline performing subjects with a stressed result (False Stressed) (Type I error)

Total= Number of instances of data.

IV. RESULTS AND DISCUSSIONS

Each channel of EEG data is segmented with a duration of 32.76sec (2¹⁴samples) with 25% overlap. The proposed 64 power ratio features are extracted from the segmented data. Features computed from all channels for all subjects with 100 Monte Carlo using 10-fold cross-validation and holdout technique are shown in Table-3. It is observed that 10-fold cross-validation provides better results compared to the holdout technique. Among the SVM classifiers, Gaussian SVM with KS=2.5 provides better performance. There is no improvement while KS is less than 2.5 and hence KS=2.5 is chosen for the rest of the results. Polynomial SVM with higher-order may provide good classification but it takes

Table 3:Classifier model using 100 Monte Carlo simulations by 10 fold cross-validation technique and Holdout(75/25)technique

Performance Metrics	Classifier															
	10 Fold cross-validation									Holdout						
	Gaussian SVM			Polynomial SVM			KNN			Gaussian SVM			Polynomial SVM		KNN	
	KS= 2.5	KS= 9.9	KS= 90	O-2	O-3	M= 1	M= 10	M= 100	KS- 2.5	KS- 9.9	KS- 90	O-2	O-3	M= 1	M= 10	M= 100
ACC	93.66 ±0.14	78.12 ±0.12	65.96 ±0.12	82.54 ±0.23	87.30 ±0.26	96.96 ±0.10	88.41 ±0.27	74.49 ±0.14	92.78 ±0.65	77.56 ±1.23	65.62 ±1.32	82.18 ±1.10	86.75 ±1.05	96.10 ±0.55	86.81 ±0.78	73.16 ±1.08
MCR	0.063	0.218	0.340	0.174	0.127	0.030	0.115	0.255	0.072	0.224	0.343	0.178	0.132	0.039	0.131	0.268
SEN	0.931	0.748	0.556	0.790	0.848	0.970	0.873	0.700	0.924	0.742	0.558	0.789	0.844	0.962	0.831	0.676
FPR	0.058	0.185	0.237	0.139	0.102	0.031	0.104	0.210	0.068	0.190	0.246	0.145	0.109	0.040	0.095	0.213
SPE	0.942	0.814	0.762	0.860	0.897	0.968	0.895	0.789	0.931	0.809	0.753	0.854	0.890	0.959	0.905	0.786
PRE	0.941	0.801	0.701	0.850	0.892	0.968	0.892	0.768	0.930	0.795	0.694	0.844	0.885	0.959	0.897	0.760
NPV	0.932	0.764	0.632	0.803	0.855	0.970	0.875	0.725	0.924	0.758	0.630	0.802	0.851	0.962	0.842	0.708
CK	0.873	0.563	0.319	0.650	0.746	0.939	0.768	0.489	0.925	0.768	0.645	0.816	0.863	0.959	0.863	0.722
F1	0.936	0.774	0.620	0.819	0.869	0.969	0.882	0.733	0.927	0.767	0.619	0.815	0.864	0.961	0.863	0.716

Table 4: Channel wise accuracy of all 36 subjects using Gaussian SVM and KNN respectively

channel	Gaussian SVM (KS=2.5)	KNN (M=1)
FP1	98.72±0.51	99.42±0.36
FP2	97.90±0.62	98.25±0.61
F3	98.11±0.74	99.07± 0.49
F4	96.17±0.87	97.11± 0.52
F7	97.35±0.76	98.82±0.65
F8	95.74±0.78	97.43±0.57
T3	97.30±0.68	98.05±0.60
T4	96.80±0.73	98.11±0.46
C3	97.66±0.77	99.09±0.55
C4	95.42±0.67	97.16±0.57
T5	96.29±0.71	96.72±0.67
T6	97.17±0.72	97.99±0.57
P3	96.93±0.84	98.26±0.64
P4	95.86±0.59	96.74±0.47
O1	98.25±0.71	99.10±0.44
O2	97.05±0.66	98.22±0.49
FZ	97.63±0.65	98.93±0.45
CZ	95.82±0.80	97.94±0.53
PZ	96.54±0.63	97.78±0.537

more computational time and hence it is discarded. It is also observed from the table that KNN with one neighbor provides better performance while comparing to more neighbors.

Channel-wise accuracy of all subjects using the Gaussian SVM classifier and KNN classifier is performed. Table 4 describes the analysis of all channels over the 36 subjects with 100 Monte Carlo using the 10 Fold cross-validation technique [11]. From table 4, it is observed that the KNN classifier is the best classifier for stress analysis compared to the Gaussian SVM classifier. We can also observe that every channel independently providing almost the same classification accuracy but the Fp1 channel itself provides higher accuracy compare to other channels.

For table 4, KNN channel accuracy is given the ranking i.e. the highest channel accuracy is given as rank 1 and the next highest channe accuracy is given as rank 2. This continues until every channel gets a rank. From the obtained data, different combinations of channels were taken and classification is done. Combinations are taken in such a way that at first, rank 1 channel is selected and given to classifier, then rank 1 and rank 2 channels were given to classifier then rank 1, rank 2, rank 3 channels were given to classifier so on. This is the ranking approach. From figure 3, it is observed that adding the channels results in low accuracy. So, the Fp1

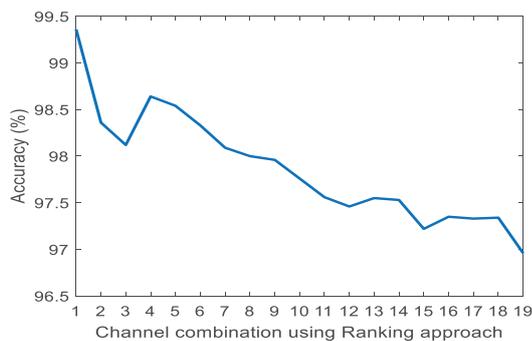


Figure 3: Analysis of EEG channels for KNN (K=1) using ranking approach

Table 5: Brain Lobe accuracy

Lobe	KNN (M=1)	SVM (KS=2.5)
Fp1,Fp2	98.25 ±0.30	98.44±0.40
F3,F4,Fz,F7,F8	97.95±0.29	96.14±0.29
C3,C4,Cz	99.31±0.22	98.11±0.26
P3,P4,Pz	98.30±0.21	97.60±0.22
O1,O2	98.48±0.25	97.98±0.35
T3,T4,T5,T6	96.87±0.32	95.58±0.36
19 Channels	96.96±0.10	93.66±0.14

channel itself provides an accuracy of 99.42% with a deviation of 0.36. Table 5 provides the accuracy of different lobes in the brain for stress classification. Here electrodes present at each lobe are combined together and followed the proposed technique.

Figure 4 shows the mean PSD of Delta, Theta, Alpha, Beta and gamma frequency bands for relaxing and stress detection. The mean PSD is computed for the Fp1 channel by considering all 36 subjects. It is observed that PSD is high during stress than in relaxing conditions for the delta band, theta band and gamma band. PSD of theta band for stress condition is more than relaxing conditions up to 6.4Hz, after that the data is almost overlapped with each other. Hence, it is suggested to split the theta band into two sub-bands for better classification. PSD of the beta band and alpha band for stress condition is overlapped with PSD for relaxing conditions. Gamma band PSD for stress condition is more than relaxing condition and it is also observed that it is almost decreasing linearly till 49Hz. Figure 5 shows the details for frequency >49Hz, it is observed that at 50Hz there is a large spike indicating power line noise and is more during relaxing conditions than at stress condition. But after 51Hz again PSD of gamma-band is more for stress detection than relaxing condition. By observation, changes occur only at 49Hz to 51Hz.

V. CONCLUSION

Stress analysis using EEG data power ratio features has been done and evaluated. EEG data has been preprocessed channel-wise and then computed PSD. EEG band power as features was computed from the PSD and in addition, 56

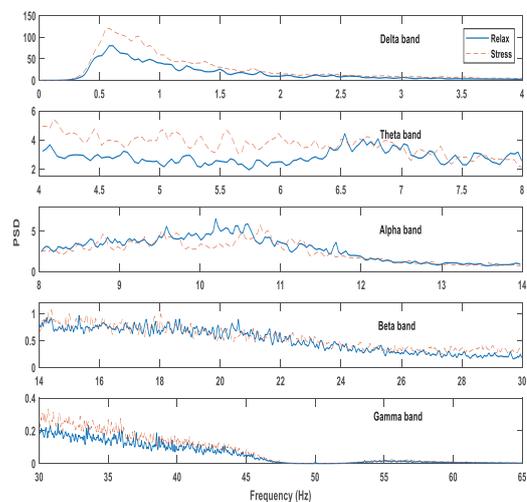


Figure 4: PSD of Fp1 channel bands for stress analysis

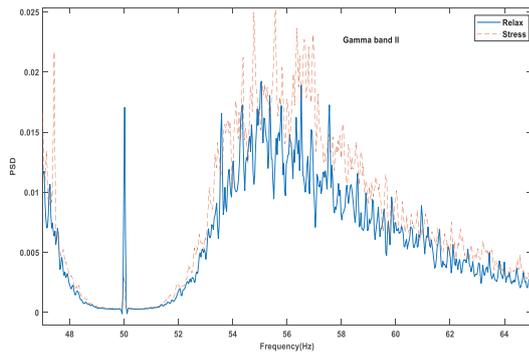


Figure 5: PSD of Fp1 channel gamma II band for stress analysis

features were computed from EEG band powers. SVM with Gaussian & polynomial kernel functions and KNN classifiers were used for stress analysis. These classifiers were evaluated using the 10 Fold cross-validation technique and Holdout technique. The developed algorithms were evaluated using the data obtained from the physio bank. From the results, it is concluded that KNN based classifier performs better. It is also concluded that Fp1 alone provides better classification compared to all channels together or individual channels. Hence, Fp1 channel EEG data processing would provide good stress analysis for real-time applications.

REFERENCES

- [1] Understanding the stress response, Harvard health publishing, Harvard medical school, Downloaded from, <https://www.health.harvard.edu/staying-healthy/understanding-the-stress-response>. Accessed on 30/11/2019.
- [2] Asif A, Majid M, Anwar SM 'Human stress classification using EEG signals in response to the music tracks', *Computers in Biology and Medicine* Vol: 107,182-196, 2019.
- [3] Shonp, Imk, ParkJH,LimDS,JangB,Kim JM,'Emotional stress state detection using genetic algorithm based features selection on EEG signals', *International Journal of environmental research and public health* , Vol.15(11),Nov 2018.
- [4] Prashant Lahane,Mythili thirugnanam,'Human emotion detection and stress analysis using EEG signal', *International journal of innovative technology and exploring engineering* ',Vol-8,Issue 4S2,2019.
- [5] Tee Yiwen,Siti Armiza Mohd Aris,'Electroencephalogram (EEG) Stress analysis on alpha/beta ratio and theta/beta ratio', *Indonesian Journal of electrical engineering and computer science*,Vol 17,No.1,Jan 2020.
- [6] Joni N. Saby, Peter J. Marshall 'The utility of EEG Band power analysis in the study of infancy and early childhood', *Dev Neuropsychol*,Vol. 37(3), pp:253-73 , 2012.
- [7] Zyma I, Tukaev S, Seleznov I, Kiyono K, Popov A, Chernykh M, Shpenkov O,' Electroencephalograms during Mental Arithmetic Task Performance' ,Vol.4(1):14, 2019.
- [8] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. *PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals* (2003). *Circulation*.101 (23):e215-e220. Downloaded from <https://physionet.org/physiobank/database/eeegmat/>, Accessed on 2/11/2019.
- [9] Michael Cerna, Audrey F. Harvey,'The Fundamentals of FFT-Based Signal Analysis and Measurement ', *National instruments, Application note* 041,July 2000.
- [10] Yoshua Bengio,Yves Grandvalet,'No unbiased estimator of the variance of K- Fold Cross-validation', *Journal of machine learning Research*,Vol 5, pp.1089-1105, 2004
- [11] Setti Suresh, VPS Naidu, 'vibration of heterogeneous gearbox faults using EMD features and SVM classifiers',*IOP Conference series; materials science and engineering*, Vol.624, No.1, 2019.