Computer Aided Diagnosis of ASD based on EEG using RELIEFF and Supervised Learning Algorithm

Roopa Rechal. T, Dr.P. Rajesh Kumar

[1]Research Scholar, Dept. of ECE, AUCE (A), Andhra University, Visakhapatnam, India.
[2]Professor and HOD,Dept. of ECE, AUCE (A), Andhra University, Visakhapatnam, India.

Abstract : Autism Spectrum Disorder is diagnosed by physical examination of electroencephalography (EEG) signals that is very responsive to time consuming and bias. Diagnosing autism in existing research experiences low power and unsuitability for processing extensive datasets. An automated diagnosing is an essential assist to medical professionals to eliminate the problems mentioned above. In this article, a novel technique is propounded to diagnose autism from VMD, RELIEFF and supervised learning algorithms. A universal EEG dataset is adopted to explore the proposed method's performance. The technique starts with the extraction of features from EEG signals via VMD, and to recognize the best features RELIEFF is employed. Then, to distinguish typical and autism signals, supervised learning (KNN, SVM, and ANN) methods is employed. The outcome illustrates that the proposed technique attains high accuracy, indicating a powerful way to diagnose and categorize autism.

Keywords: RELIEFF algorithm, Variational Mode Decomposition, Electroencephalogram, Machine Learning Algorithms.

I.   Introduction

Autism spectrum disorder (ASD) is a predominant neurocognitive condition specified by communication and social interaction deficits, along with confined repeated behavioral patterns, preferences and interests [1]. All over the world ~1.5% in developed countries, this state has a particular economic and social effect because of its high prevalence [2]. It has a major impact on daily family life and related to increased morbidity. The Brain-Computer Interfaces (BCI) based on EEG incorporate extensively analyzed communication technics [3]. Nonetheless, concerning the psychological approach in neuro-rehabilitation, the utilization of children with consideration deficiency hyperactivity issue that incorporates the existence of continuous absentminded, incautious, and hyperexcitable practices. A computer-aided detection (CADe) is a set up to support a doctor or a clinician in diagnosing a specific disease or disorder. A CAD system isn't purposive to detect by itself, even so as an aiding tool for the clinician to diagnose, saving them time, increasing accuracy, and providing a second opinion.

A virtual entityP300-based BCI standard with specialized execution is depicted. Intuitive virtual conditions with the consideration associated with P300 brain waveform nature to make a subjective preparing apparatus for ASD. The P300 signal is a renowned neurological sign of observation procedure for diagnosing unique objects in a stimulus sequence [4]. The preparation of joint consideration aptitudes is coupled to the P300 signal because this latter is generally utilized in controlled observational investigations and is identified with a combination of data with context and memory [5].
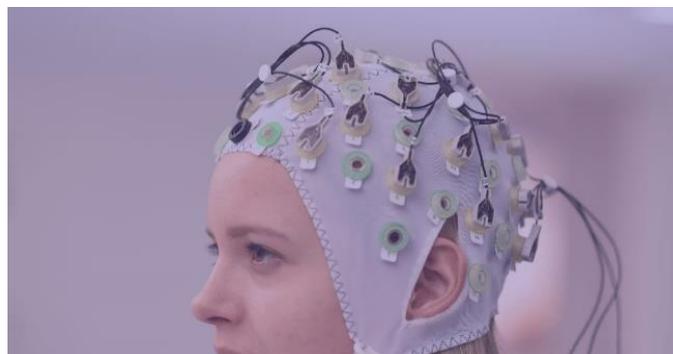
Fig. 1. Placement of Electrodes of an EEG

The electrical signal of an encephalon has a minimum amplitude of $100\mu V$. The frequency is within 0.44 Hz and 80 Hz typically. The frequency ranges and amplitude for each type of wave is shown in Table 1.

Wave	Amplitude	Frequency ranges
Delta band	High	0.5 – 4 Hz
Theta band	Low – Medium	4 - 8 Hz
Alpha band	Low	8 – 15 Hz
Beta band	Very Low	15 – 30 Hz
Gamma band	Smallest	30 – 50 Hz

Table I. Amplitude and Frequency Ranges of wave's

The residue of the article as follows. Existing work review of literature is demonstrated in Section II. Detailed information over the techniques employed (i.e., RELIEF, VMD, KNN, SVM and ANN) is in Section III. The performance parameters of the features extracted from each classifier concerning their confusion plot and the attained experimental results were analyzed in section IV. Section V finally concludes the state of the propounded method.

II.   Related Work

An approach by Kleih et al., [6], The Brain-Computer Interfaces (BCI) based on EEG, incorporate extensively analyzed communication technics.

Mythili & Mohamed Shanavas et al., [7] examined and propounded the best ideal and optimum features to overcome the learning obstacles and to speed up the learning capability. The Principal Component Analysis operates as an evaluator for the Feature Extortion. PSO was deployed for a course of Feature Selection. The PSO appraises the parameter optimization to attain utmost effectual and finest features. The resulting best features were supplied to the SVM classifier. An ultimate outcome evinced that the propounded technique acquires higher accurateness.

Bosl, Tierney, Nelson et al., [8] underlying neurobiological disorders that consistently differentiate autistic and non-autistic brains. The multiclass SVM, NB and KNN algorithms have been implemented to analyze the autistic and typical signals and achieved 80% accuracy.

Ahmadlou et al. [9] examined the fractal dimension (FD) to estimate intricacy and dynamical transitions in the autistic brain. A radial basis function classifier and achieved 90% accuracy.

In work done by Sheikhani et al. [10], A short time Fourier transform (STFT) method employed for extracting features from EEG signals and then given to KNN classifier as input and achieved 82.4% of accuracy. Their next article enhanced the process, employed more extensive data for trial, and achieved up to 96.4% accuracy.

Abdulhay et al. [11] studied frequency 3D mapping and inter-channel stability of EEG signals to analyze the capability for recognizing abnormalities in EEG signals and their relation with ASD. The research originates that, for autism analyzing the order of frequency content and the inter-channel durability of pulsation plot all over the scalp were good measures.

Ridha Djemal et al. [12] proffered computer-aided detection (CADe) of autism grounded upon EEG signal analysis. The propounded method was grounded on entropy (En), artificial NN (ANN) and discrete-wavelet transform (DWT). DWT was implemented to decompose EEG signals to approximate and delineate the coefficients to attain EEG sub-bands. The FV was built by evaluating Shannon En values as of every EEG sub-band. ANN categorizes the equivalent EEG signal to autistic or normal grounded on the extorted features. The experiential outcomes evinced the efficacies of the propounded methodology for aiding the diagnosing of autism. A receiver operating characteristic curve (ROC) metric was deployed to gauge the propounded technique's performance.

III.   Proposed Work

The article aims to develop an algorithm based on EEG signal processing to detect autism. The proposed method employed RELIEFF, VMD and supervised learning algorithms. For used techniques, a concise description and fundamental mathematical formulation given in the following section.

A. Variational Mode Decomposition

Variational mode decomposition (VMD) a novel adaptative signal fragmentation, it fragments each real-time signal to variational modes (uk) or a band limited functions. For the reconstructing an input signal, each method transpired concurrently and exhibited sparsity property. VMD fragments real-time signals into k modes (uk) surrounding its center frequency (ω). Frequency shifting property and Hilbert transform are beneficial variables in formulation and optimization of a problem. The constrained variational problem formulation is [13],

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}$$

$$\sum_k u_k = f \tag{1}$$

The four modes of single fragment features (Coefficient of Variation, Entropy, Inter Quartile Range, Mean, Negentropy, Kurtiosis, Spectral Decrease, Spectral Flatness, Skewness, Spectral Spread, Standard Deviation) of autism and normal EEG shown in table II.

In this work, by making use of four-level variational decomposition method, 200 signals were decomposed into 800 signals, in which 70% (560) of the signals are used for training and remaining 30% (240) of the signals are used for testing. Each EEG signal is fragmented into four sub bands consisting of 1024 samples, from each fragmented part, 11 features (8-statistical, 3-spectral) are extracted. In total, 44 features are extracted from each EEG signal. By observing the extracted feature table, normal and autism signal features have the considerable variations. The highest variation is observed in IQR and standard deviation.

Table II: Extracted feature values of the COV, Entropy, IQR, Mean, Negentropy, Kurtosis, Spectral Decrease, Spectral Flatness, Skewness, Spectral Spread,
Standard Deviation [13].

Mode no	Signal type	COV	Entropy	IQR	Mean	Neg entro	Kurtosis	Spectral decrease	Spectral Flatness	Skewness	Spectral Spread	Standard Deviation
Mode 1	Normal	1.5477	1.0485	30.2210	13.4910	3.4092	458.7429	-0.4880	0.2157	0.1766	216186	20.8806
	Autism	4.8486	1.0337	170.5287	23.3108	5.1127	488.4435	-0.1252	0.2592	0.1673	208105	113.0268
Mode 2	Normal	63.7857	1.2888	25.3614	0.2978	3.0745	500.2189	-0.0091	0.3093	-0.0386	196742	18.9982
	Autism	46.0985	1.0207	188.6512	3.1102	5.3636	499.9619	8.2475E-5	0.3546	0.0352	185487	143.379
Mode 3	Normal	778.5202	1.1972	32.5381	0.0295	3.3569	500.9931	-0.0069	0.1454	-0.0162	178681	22.9919
	Autism	382.2514	1.0207	244.5299	0.5323	5.7139	500.8403	0.00484	0.1517	-0.0045	170118	203.5015
Mode 4	Normal	820.5311	1.3910	12.0869	0.0111	2.2330	500.9958	-0.0044	0.2521	-0.0033	141753	9.07104
	Autism	559.4597	1.0416	226.4214	0.3060	5.5202	500.9086	0.00409	0.2220	0.0019	146350	171.2086

B. K-Nearest Neighbor Classifier (KNN)

Among the non-parametric approaches used for classification of electrophysiological signals, KNN is one. The input comprises K closest training samples (data points), and the output is a class member in the classification problem. A sample will be classified with most of the neighbours and allocated to most common classes among K-nearest neighbours. For classification in KNN class membership is the output. To achieve the classification results testing and training datasets of autistic EEG are employed to K-nearest neighbours. The processing function used is spearman distance [14].

The KNN is simple among all machine learning algorithms and depends on instance-based learning. It is a lazy classifier as its function approximated locally and the calculations are postponed until classification. A drawback occurs with the skewed class distribution, i.e., the most frequent class can dominate the prediction of the new data point. This drawback can be bypassed by limiting the impact of distance from the test data point to each of its k nearest neighbours. Assign a weight to each vote is one way to reduce the effect where weight is a function of the distance between known and unknown data points. If the weight is defined as the inverse squared

distance, votes cast by the nearest data points have much influence on decision process than that of distant data points.

The test data point represented with a circle and the data points of different classes served with squares and triangles. The solid line circle is the case for K=3 where the test data point is at the center and encloses only three data points on the plane. So the test point will be assigned to the class of triangles as there are two triangles and one square. Dashed line circle is another case (K=5) where it assigns the test point to the class of squares. The decision boundary becomes smoother with increasing K value.

Spearman Distance

The distance between the data vectors x_s and y_t are defined as

$$d_{st} = \left(1 - \frac{(r_s - \bar{r}_s)(r_t - \bar{r}_t)}{\sqrt{(r_s - \bar{r}_s)(r_s - \bar{r}_s)'}\sqrt{(r_t - \bar{r}_t)(r_t - \bar{r}_t)'}}\right) \quad (2)$$

C.   Artificial Neural Network (ANN):

 ANNs are motivated with biological neural networks, i.e., animal central nervous systems especially brain. These are used to approximate or estimate functions that can depend on a high number of unknown inputs. Different connections have different numeric weights, which can be turned based on experience, makes the ANNs capable of learning and are more adaptive to inputs. The set input neurons get activated by the input data. The output neuron determines the target class to which the data belongs to. For testing the performance, using ANN for categorization of autism EEG dataset is used [15].

Types of ANNs are differ from those of with a single layer or two layers of one direction logic to multiple many-input layers and multi-directional feedback loops. The networks use some algorithms in programming to control and organize their activation functions. Most of the networks use weights for changing the parameters of throughput and different connections. ANNs can learn through external inputs or can perform self-learning. The abilities of self-learning and decision making make the ANNs suitable for a broad category of problems that may include a large amount of data.

In feed-forward neural network, the connections between the network units cannot form a directed cycle (no loops presented). The information moves only in a forward direction from input nodes to output nodes through some hidden nodes.

The inputs of the network contain some processing functions that convert user data into a form, which is suitable for a network. The outputs also possess corresponding processing functions that are utilized to convert user-provided target vectors. The network outputs are backward processed with similar functions to generate data (output) with comparable characteristics of original targets. The input and output processing functions used in the artificial neural networks. The processing function used is sigmoid function.

Sigmoid Function

Sigmoid function corresponds to the shape of "*S*" (sigmoid curve) and it is a mathematical function which belongs to a special incident of the logistic function. It can be expressed as

$$S(t) = \frac{1}{1 + e^{-t}} \quad (3)$$

A sigmoid function has a positive derivative for all real input values that is defined, it is a bounded differentiable real function.

D.   Support Vector Machines (SVM)

 The classifier is given to above features for Categorization of typical EEG signal. The SVM is developed primarily for two-class classification and can be extended for multi-class optimization problems. The basic approach is to locate a hyperplane which separates the data correctly into two classes. Each object of the training set (set of known objects) contains a feature vector and its corresponding class value. Basis of the training data, the algorithm obtains a decision function for classifying the unknown data. The classifier can reduce the experimental classification error by maximizing geometric margin. Thus, it can be characterized as a maximum margin classifier also. The maximum margin classifier can give better results than other traditional classifiers.

A classifier with a linear decision boundary is called as a linear classifier. The main intention is to attain a decision boundary which can separate the training data. If the separation is not possible with a linear hyperplane, the classifier maps the data into high dimensional feature space using some pre-defined functions (kernel).

The selection of kernel is an important issue in support vector machine classifier. The kernels introduce several nonlinearities into a classification problem by mapping data X, implicitly to Hilbert space through a function φ(X). The SVM classifier requires only inner products of *φ(X)* features, though mapping to Hilbert space or the explicit computation of features may be cumbersome. Kernel functions applied to map data non-linearly to a high dimensional feature space and this mapping is linearly separable. The mapping performed by replacing the inner product (x,y) with Φ(x). Φ(y) and the kernel function is K(x,y)= Φ(x).Φ(y).

Decision function in two class problem is expressed as

$$g\square x\square \square sign[w^T f \square x\square \square b] \quad (4)$$

Optimization problem is given as

Minimize $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad J(w,b,e) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^{N} e_i^2$

(5)

Subject to $\qquad\qquad\qquad\qquad y_i[\, w^T f(x_i) + b \,] = 1 - e_i$, i=1,2….,N $\qquad\qquad$ (6)

Here x_i is N input with i^{th} feature vectors, and y_i is the class label of 1 or -1 for x_i. γ is the parameter of regularization, α_i is a Lagrangian multiplier and b is the bias term. its SVM classifier output derived as

$$g(x) = sing[\sum_{i-1}^{N} \alpha_i y_i K(x,x_i) + b$$

(7)

SVM classifier needs kernel for training. The Gaussian RBF kernel is the efficient one. RBF kernel is expressed

as

$$K(x,x_i) = f^T(x)f(x_i) = e^{(-\|x-x_i\|^2/2\sigma^2)}$$

(8)

Parameter σ is an optimization kernel width [16].

E. RELIEFF

RELIEFF algorithm is widely used to filtrate feature selection in a very effectual method. Distinct and numeric features with binary classification problems are associated with its high-end approach. RELIEFF algorithm use heuristic rule for inductive learning algorithms whereas, Inductive machine learning use greedy search.

The fundamental purpose of RELIEFF is to evaluate attributes to explain how their values differ among those samples which are nearer to each other. For a prescribed instance, RELIEFF seeks for its two nearby neighbors: out of this one from the same class is denominated as the nearest hit (H) and the other from the dissimilar class is denominated as the nearest miss (M). The actual RELIEFF algorithm [17] randomly chooses n training occurrences, where n is the user-defined parameter.

Algorithm is as follows.

Step-1: assign weights W[A] = 0.0;
Step-2: for i=1: n
Step-3: randomly choose an instance R;
Step-4: find nearby hit H and nearby miss M;
Step-5: for A = 1: #all_attributes
Step-6: W[A] = W[A] – diff (A, R, H)/n
Step-7: + diff (A, R, M)/n;
Step-8: end;
Step-9: end;

The estimation of the quality of attributes is denoted by weights W[A]. The basic formulation for upgrading the weights is that a useful feature should have the same value of instances of the same classes (subtracting the difference diff (A; R; H)) and distinct values of instances of the distinct classes (adding the difference diff (A; R; M)).

The difference between the values of an attribute for two instances is calculated by the function diff (Attribute, Instance1, Instance 2). For distinct attributes, the difference is either 1 (different) or 0 (equal) and the actual difference normalized to the interval [0; 1] is the difference of continual attributes. Assigning n guarantees all weights W[A] to be in the interval [-1; 1], however, assigning n is not essential if W[A] is to be used for relative differentiation between the attributes.

The entire distance is just the sum of variations of all characteristics. Actual RELIEFF utilizes the squared difference, where the discrete attributes are equal to diff. In all the observations, there is no vital difference among the results using diff or squared difference. If N is the number of all training instances, then the intricacy of the preceding algorithm is O (n x N x #all_attributes).

IV.   Results and Discussion

In this work, an effective technique comes up with ith RELIEFF, VMD and supervised learning algorithms is evolved and imposed to classify autism. The propounded classification algorithms of autistic signals are enacted and simulation was done in MATLAB, and this section presents the simulated results. The EEG dataset of autism adopted from Kaggle database [18]. The channels distribution in the data patterns is C3, Cz, C4, CPz, P3, Pz, P4, POz. The EEG dataset of typical controls is acquired via Bonn University Hospital of Freiburg [19]. It comprises five separate subsets (A-E) labelled Z, O, N, F, and S. Typical controls were recorded from Set A & B. By using the frequency decimation technique, the frequencies of both datasets are equalized.

Classifier parameters utilized to estimate the performance of the proposed method are sensitivity, accuracy, specificity, G_mean, precision and F_measure.

$$\% \text{ Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} * 100 \tag{9}$$

$$\% \text{ Sensitivity} = \frac{TP}{TP+FN} * 100 \tag{10}$$

$$\%\text{Specificity} = \frac{TN}{TN+FP} * 100 \tag{11}$$

$$\%\text{Precision} = \frac{TP}{TP+FP} * 100 \tag{12}$$

$$\%F_{measure} = \frac{(2*TP)}{(2*TP)+FP+FN} * 100 \tag{13}$$

$$\%G\_mean = \sqrt{\text{Specificity xSensitivity}} * 100 \tag{14}$$

Table 3 illustrates the simulation results of the classification algorithm of autism along with ANN and RELIEFF. In accordance with the attained confusion matrix, classification algorithm based on ANN classifier accomplishes an overall sensitivity 93.33%, overall accuracy 90.41%, overall specificity 87.50%, overall G_mean of 90.36%, overall F_measure 90.68% and overall precision 88.18%.

TABLE III. Classification of autism using ANN confusion matrix with RELIEFF

Signal	Autism (%)	Typical Control (%)
Autism	93.3	12.5
Typical Control	6.7	87.5
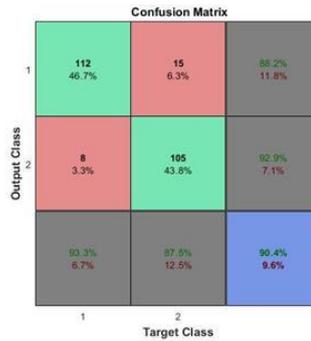
Fig. 2. Autism classification using ANN confusion matrix with RELIEFF

Table 4 illustrates the simulation results of the classification algorithm of autism along with KNN and RELIEFF. In accordance with the attained confusion matrix, classification algorithm based on KNN classifier accomplishes an overall sensitivity 96.67%, overall accuracy 93.33%, overall specificity 90.00%, overall G_mean of 93.27%, overall F_measure 93.54% and overall precision 90.62%.

TABLE IV. Classification of autism using KNN confusion matrix with RELIEFF

Signal	Autism (%)	Typical Control (%)
Autism	96.7	10.0
Typical Control	3.3	90.0
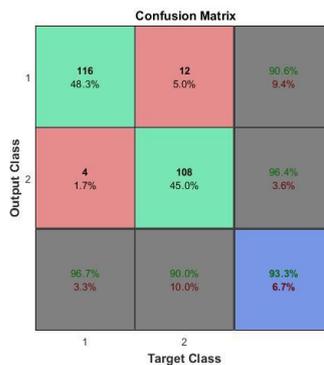
Fig. 3. Autism classification using KNN confusion matrix with RELIEFF

Finally, Table 5 illustrates the simulation results of the classification algorithm of autism along with SVM and RELIEFF. In accordance with the attained confusion matrix, classification algorithm based on SVM classifier accomplishes an overall sensitivity 97.50%, overall accuracy 95.41%, overall specificity 93.33%, overall G_mean of 95.39%, overall F_measure 95.51% and overall precision 93.60%.

TABLE V. Classification of autism using SVM confusion matrix with RELIEFF

Signal	Autism (%)	Typical Control (%)
Autism	97.5	6.7
Typical Control	2.5	93.3
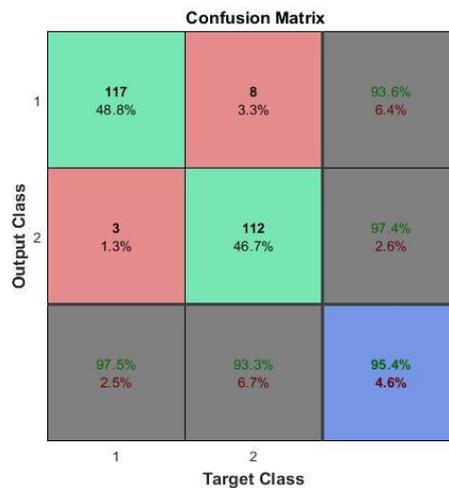
Fig. 4. Autism classification using SVM confusion matrix with RELIEFF
TABLE VI. Autism classification performance summary

S.No	Parameter	RELIEFF + ANN	RELIEFF + KNN	RELIEFF + SVM
1	Sensitivity	93.33	96.67	97.50
2	Accuracy	90.41	93.33	95.41
3	Specificity	87.50	90.00	93.33
4				
5				
6				

5. Comparison of parameters (bar   **Fig. performance graph)**
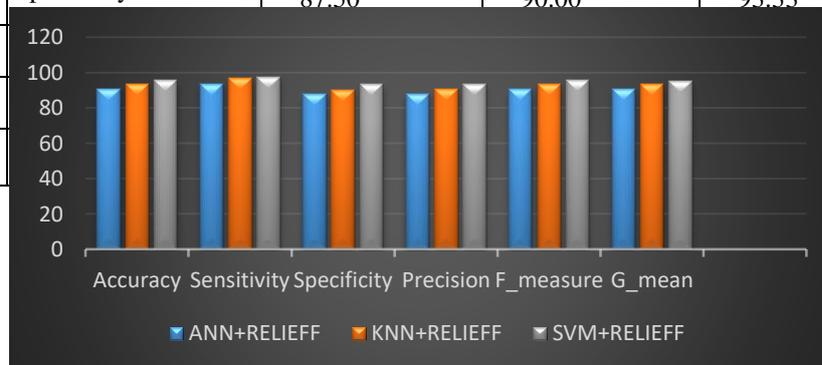
V.   Conclusion

A computer-aided detection (CADe) system has an enormous potentiality to aid therapist in the procedure of diagnosing, increasing the accuracy and to avoid time delay. This paper discussed the methods for diagnosing autism with EEG signals. Firstly, the proposed technique focused on VMD for extracting the features and the acquired

features are fed to RELIEFF for selecting the best features. Secondly, the supervised learning algorithms i.e., KNN, SVM, and ANN are utilized for categorization of autism. Finally, the attained accuracy of ANN with RELIEFF is 90.41% and KNN with RELIEFF is 93.33% while, SVM with RELIEFF is 95.41%. It can be concluded that with a fusion of SVM and RELIEFF highest accuracy achieved from the overall results.

REFERENCES

[1] Astrand, E., Wardak, C., and Ben Hamed, S. (2014) Other Mental Disorders. (n.d.). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition*. doi: 10.1176/appi.books.9780890425596.529303.

[2] Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., and Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. Psychol. Med. 45, 601–613. doi: 10.1017/S003329171400172X.

[3] Farwell, L., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology, 70*(6), 510-523. doi:10.1016/0013-4694(88)90149-6

[4] Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected Studies of the Event-Related Potential. *International Journal of Medical Sciences,* 147-154. doi:10.7150/ijms.2.147

[5] Halgren, E., Baudena, P., Clarke, J. M., Heit, G., Marinkovic, K., Devaux, B., . . . Biraben, A. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. II. Medial, lateral and posterior temporal lobe. *Electroencephalography and Clinical Neurophysiology, 94*(4), 229-250. doi:10.1016/0013-4694(95)98475-n

[6] Kleih, S. C., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., . . . Kübler, A. (2011). Out of the frying pan into the fire—the P300-based BCI faces real-world challenges. *Brain Machine Interfaces: Implications for Science, Clinical Practice and Society Progress in Brain Research,* 27-46. doi:10.1016/b978-0-444-53815-4.00019-4

[7] Mythili, M., & Shanavas, A. M. (2015). An Improved Feature Selection (IFS) Algorithm for Detecting Autistic Children Learning Skills. *Biosciences Biotechnology Research Asia, 12*(1), 499-505. doi:10.13005/bbra/1691

[8] Bosl, W., Tierney, A., Tager-Flusberg, H., & Nelson, C. (2011). EEG complexity as a biomarker for autism spectrum disorder risk. *BMC Medicine, 9*(1). doi:10.1186/1741-7015-9-18

[9] Ahmadlou, M., Adeli, H., & Adeli, A. (2010). Fractality and a Wavelet-Chaos-Neural Network Methodology for EEG-Based Diagnosis of Autistic Spectrum Disorder. *Journal of Clinical Neurophysiology, 27*(5), 328-333. doi: 10.1097/wnp.0b013e3181f40dc8

[10] Sheikhani, A., Behnam, H., Mohammadi, M. R., Noroozian, M., & Golabi, P. (2008). Connectivity Analysis of Quantitative Electroencephalogram Background Activity in Autism Disorders with Short Time Fourier Transform and Coherence Values. *2008 Congress on Image and Signal Processing*. doi:10.1109/cisp.2008.595

[11] Abdulhay, E., Alafeef, M., Hadoush, H., Alomari, N., & Bashayreh, M. (2017). Frequency 3D mapping and inter-channel stability of EEG intrinsic function pulsation: Indicators towards autism spectrum diagnosis. *2017 10th Jordanian International Electrical and Electronics Engineering Conference (JIEEEC)*. doi:10.1109/jieeec.2017.8051416

[12] Djemal, R., Alsharabi, K., Ibrahim, S., & Alsuwailem, A. (2017). EEG-Based Computer Aided Diagnosis of Autism Spectrum Disorder Using Wavelet, Entropy, and ANN. *BioMed Research International, 2017*, 1-9. doi:10.1155/2017/9816591

[13] (2020). Performance Analysis of Supervised Learning Algorithms for Identification of Autism Spectrum Disorder Using EEG Signals. *European Journal of Molecular & Clinical Medicine*, 7(9), 1156-1167.

[14] K Chomboon, P Chujai, and N Kerdprasop, "An Empirical study of distance metrics for K-Nearest Neighbor Algorithm," In Proceedings of 3rd International Conference on Industrial Application Engineering, pp.280-285, 2015.

[15] Sovierzoski, M. A., Azevedo, F. M., & Argoud, F. I. (2008). Performance Evaluation of an ANN FF Classifier of Raw EEG Data using ROC Analysis. *2008 International Conference on BioMedical Engineering and Informatics*. doi:10.1109/bmei.2008.220

[16] Chen, X., Yang, J., & Liang, J. (2011). Optimal Locality Regularized Least Squares Support Vector Machine via Alternating Optimization. *Neural Processing Letters, 33*(3), 301-315. doi:10.1007/s11063-011-9179-8

[17] Kira, K., & Rendell, L. A. (1992). A Practical Approach to Feature Selection. *Machine Learning Proceedings 1992,* 249-256. doi:10.1016/b978-1-55860-247-2.50037-1.

[18] Amaral, C. P., Simões, M. A., Mouga, S., Andrade, J., & Castelo-Branco, M. (2017). A novel Brain Computer Interface for classification of social joint attention in autism and comparison of 3 experimental setups: A feasibility study. *Journal of Neuroscience Methods, 290*, 105-115. doi: 10.1016/j.jneumeth.2017.07.029

[19] http://epileptologiebonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3.