The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.8254867

CLASSIFICATION OF ADHD USING TRADITIONAL MACHINE LEARNING ALGORITHMS USING BEHAVIOURAL DATA OF PATIENTS

SAMYUKTHA S*

Research Scholar, BGSIT, B G Nagara, Affiliated to VTU, India.
*Corresponding Author Email: samyuktha88sk@gmail.com

Dr. M B ANANDARAJU

Professor, Department of ECE, BGSIT, B G Nagara, Affiliated to VTU, India.

Abstract

Accurate diagnosis of attention deficit hyperactivity disorder (ADHD) can be expensive and involves comprehensive assessments like interviews, observations, and evaluations of potential coexisting conditions. With the growing availability of data, there's potential to create machine-learning algorithms that can provide precise diagnostic predictions using cost-effective measures to assist human decision-making. We present the outcomes of employing different classification methods to forecast an ADHD diagnosis that clinicians agree upon. In our proposed study, we evaluate the classification performance of two machine-learning algorithms, Naive Bayes, and K-Nearest Neighbors (KNN), both with and without Principal Component Analysis (PCA). These algorithms are applied to a dataset of 95 samples gathered from open sources. The KNN classifier demonstrates a notable accuracy of 66%, which is significantly higher than the Naïve Bayes accuracy of 55%. Our findings suggest that the KNN classifier performs better in predicting ADHD with improved accuracy.

Keywords: ADHD, clinicians, K-Nearest Neighbors, machine-learning algorithms, Naive Bayes.

INTRODUCTION

ADHD is one of the most common neurodevelopmental disorders of childhood. It is usually first diagnosed in childhood and often lasts into adulthood. Children with ADHD may have trouble paying attention, controlling impulsive behaviors (may act without thinking about what the result will be), or be overly active. It is normal for children to have trouble focusing and behaving at one time or another. However, children with ADHD do not just grow out of these behaviors. The symptoms continue, can be severe, and can cause difficulty at school, at home, or with friends.

Scientists are studying causes and risk factors to find better ways to manage and reduce the chances of a person having ADHD. The causes and risk factors for ADHD are unknown, but current research shows that genetics plays an important role. Recent studies link genetic factors with ADHD [1]. Deciding if a child has ADHD is a process with several steps. There is no single test to diagnose ADHD, and many other problems, like anxiety, depression, sleep problems, and certain types of learning disabilities, can have similar symptoms. One step of the process involves having a medical exam, including hearing and vision tests, to rule out other problems with symptoms like ADHD. Diagnosing ADHD usually includes a checklist for rating ADHD symptoms and taking a history of the child from parents, teachers, and

ISSN 1533-9211

DOI: 10.5281/zenodo.8254867

sometimes, the child. In most cases, ADHD is best treated with a combination of behavior therapy and medication. For preschool-aged children (4-5 years of age) with ADHD, behavior therapy, particularly training for parents, is recommended as the first line of treatment before medication is tried. What works best can depend on the child and family. Good treatment plans will include close monitoring, follow-ups, and making changes, if needed, along the way.

RELATED WORK

Recently, biomedical scientists have suggested several automatic diagnosis approaches for extracting mass features from Functional Magnetic Resonance Imaging (fMRI) data. For example, Waqas Majeed developed a new approach to assess that the reproducible spatiotemporal pattern of BOLD fluctuations is consistent with previous research and may have vital information about brain activity at rest. It indicates that in the resting state, the brain is active. Moreover, several studies have considered the dynamics of brain function connections [2]. Lindquist [3] have proposed a modified exponentially weighted moving average (EWMA) model, which can be applied to FMRI data, and then used it to analyze the change point of time series. Chang et al [5] have investigated the dynamic connectivity of the brain signals by applying the sliding-window methodology [5]. In [7], Ren et al have recommended a dynamic graph metrics strategy to characterize temporal changes of functional brain networks. In [4], Atif Riaz et al have suggested a Hybrid fMRI framework that utilizes affinity propagation clustering and density peak for functional connectivity. In [14], Ahmed et al. presented a classification model that and non-ADHD using ELM with different datasets. We have extracted different FC's using different time atlases and use the FC's for classification from our previous work we conclude that the accuracy of our model increases with the number of extracted FC. During the past few years, researchers have introduced different types of models for the classification of ADHD. As this illness is a medically brain disorders that physicians usually diagnose it by assessing some symptoms through a studies have identified this disorder as a two-category problem, such as ADHD and non-ADHD. Gülay Çıçek as well as created two separate datasets, including gray level cooccurrence matrix and Haralick texture features for classification purposes using the machine learning algorithms [6]. Recently, there have been many advances in this field, in [15] JieWang uses fNIRS signals for functional connectivity and interval features for classification of ADHD and non- ADHD. Shuiqi Lui proposed a novel algorithm for ADHD classification based on (CDAE) convolutional denoising autoencoder and (AdaDT) adaptive boosting decision trees [16]. Yibin Tang et al. [17] have introduced a self-encoding network with non-imaging fusion for ADHD classification, which achieves quite high accuracy. However, it has some limitations such as it does not work well with different datasets and is also not able to extract the required features from the fMRI data.

Miao and Zhang [7] have suggested a relief and VA-relief based feature extraction approach to achieve high precision classification. Later, in 2017 Sudha et al. in [8] have suggested a model to extract the gait signal characteristics of ADHD children from the video signals, which provides the disease diagnosis and strengthen the cognition of sick children [8]. Chang et al. [9] have introduced a feature extraction method based on a texture point of view considering the isotropic local binary patterns on three orthogonal planes (LBP-TOP) that employ the support vector machine (SVM) to classify the identified features. Athena Taymourtash, use sparse based representation method by extracting the feature by cluster ICs and uses KNN classifier to find out the EEG source differences between adults with ADHD and healthy controls [10]. Later in [11], Zhang has introduced the dual diagnosis model, which recognizes the feature space separation by applying sparse representation. Juan L. Lopez Marcano explained that the United States allows using the power ratio (TBPR) as a diagnostic feature of ADHD [20]. F.M. Grisales-Franco uses a Dynamic Sparse Coding (DSC) method based on physiologically motivated Spatiotemporal constraints to construct non-stationary brain activity, they search the difference between ADHD and control groups using statistical results [13].

MATERIALS AND METHODS

Behavioural data considered for the proposed work is shown in the table 1.

Table 1: Behavioural data from neuropsychological data

AQtot	IVA-II : Attentional quotient total
AQaudi	IVA-II : Attentional quotient auditive
AQvis	IVA-II : Attentional quotient visual
RCQtot	IVA-II : Response Control total
RCQaudi	IVA-II : Response Control auditive
RCQvis	IVA-II : Response Control visual

A data set of 95 samples is used in the present experiment. The data set is divided into 80 % train and 20% test set.

Proposed model

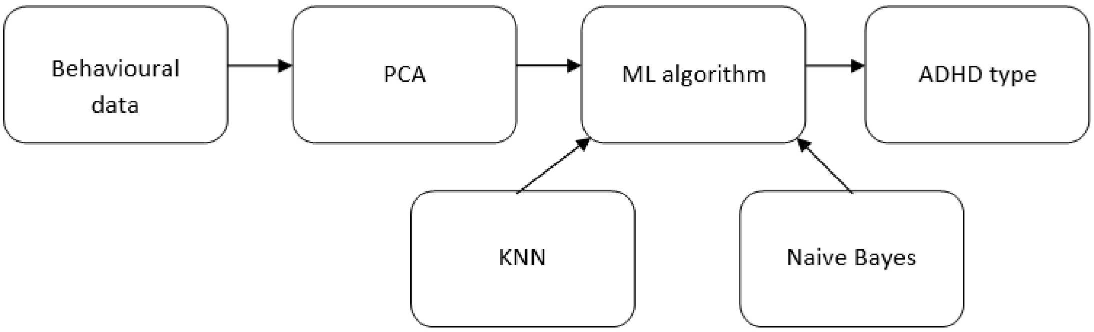


Figure 1: Proposed model

In proposed model shown in fig.1, the behavioural data has been provided as input to the ML algorithm. The ML algorithm is trained to classify using behavioural data. In the present paper, two algorithms have been analysed with respect to ADHD classification. The two ML algorithms used are KNN and Naive Bayes. Before applying KNN and Naive Bayes, the dimension of data has been reduced using PCA. The performance of KNN and Naive Bayes has been compared using PCA and without using PCA.

ISSN 1533-9211

DOI: 10.5281/zenodo.8254867

K Nearest Neighbour

The k-nearest neighbors algorithm, also known as KNN or k-NN, is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point. While it can be used for either regression or classification problems, it is typically used as a classification algorithm, working off the assumption that similar points can be found near one another. Fig. 2 shows working of KNN as a classifier.

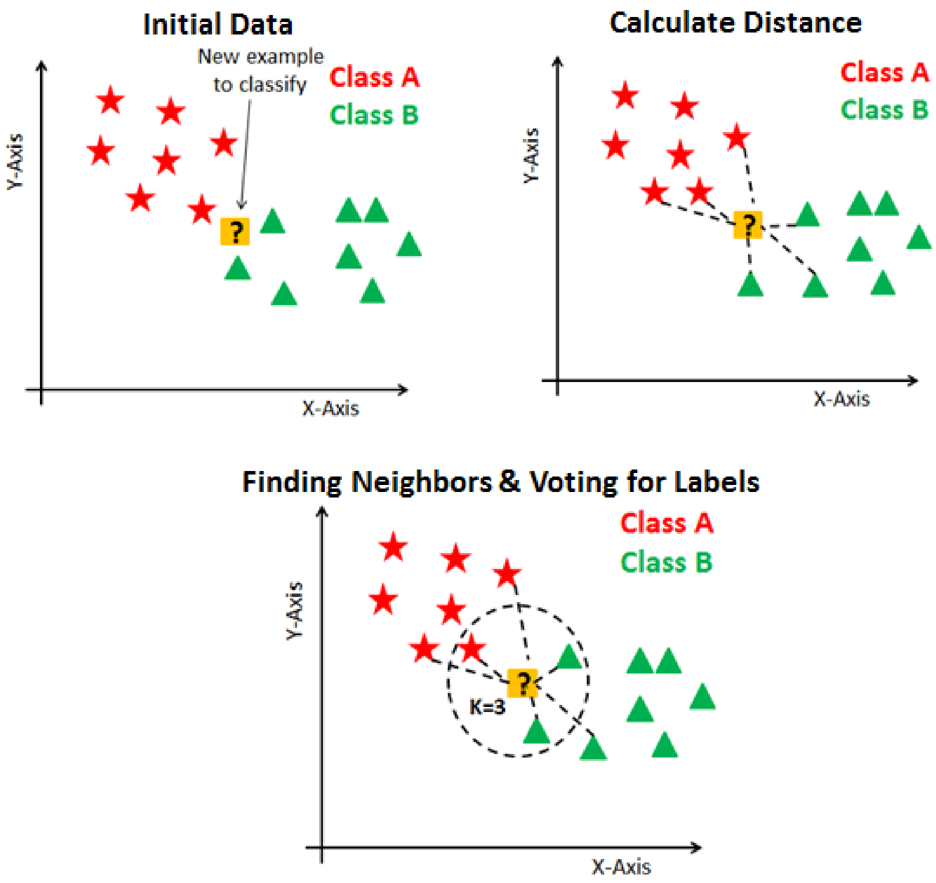


Figure 2: Classification of a query instance

The goal of the k-nearest neighbor algorithm is to identify the nearest neighbors of a given query point, so that we can assign a class label to that point. In order to determine which data points are closest to a given query point, the distance between the query point and the other data points will need to be calculated. These distance metrics help to form decision boundaries, which partitions query points into different regions. Number of distance metrics is existed. But, in the present work, Euclidean distance has been adopted. This is the most commonly used distance measure, and it is limited to real-valued vectors. Using the Eq. (1), a straight line between the query point and the other point is measured.

ISSN 1533-9211
DOI: 10.5281/zenodo.8254867

$$d(x,y)=\sqrt{\sum_{i=1}^{n}(y_i-x_i)^2} \tag{1}$$

Naive Bayes

Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. Bayes' theorem is also known as Bayes' Rule or Bayes' law, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability as given in Eq. (2).

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)} \tag{2}$$

Principal component analysis (PCA)

Principal component analysis, or PCA, is a statistical procedure that allows you to summarize the information content in large data tables by means of a smaller set of "summary indices" that can be more easily visualized and analyzed.

Statistically, PCA finds lines, planes and hyper-planes in the K-dimensional space that approximate the data as well as possible in the least squares sense. A line or plane that is the least squares approximation of a set of data points makes the variance of the coordinates on the line or plane as large as possible as shown in Fig. 3.

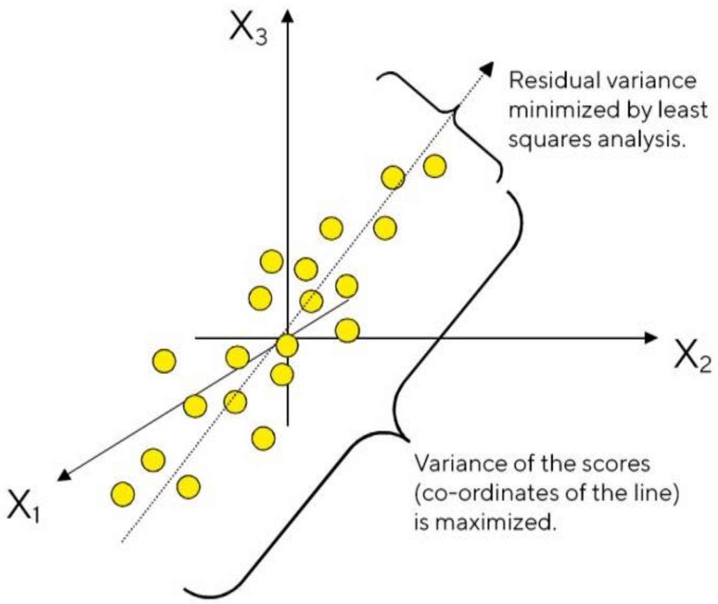


Figure 3: Working principle of PCA

ISSN 1533-9211

DOI: 10.5281/zenodo.8254867

RESULTS AND DISCUSSION

The classification of ADHD type with behavioural data using PCA and KNN

PCA on behavioural data

We computed PCA on the 6 variables from the IVA-II (behavioural data). As we can see from the interactive plot, both components are similar in their distribution. The first component explains 63% of variance, whereas the second component explains 25% of variance.

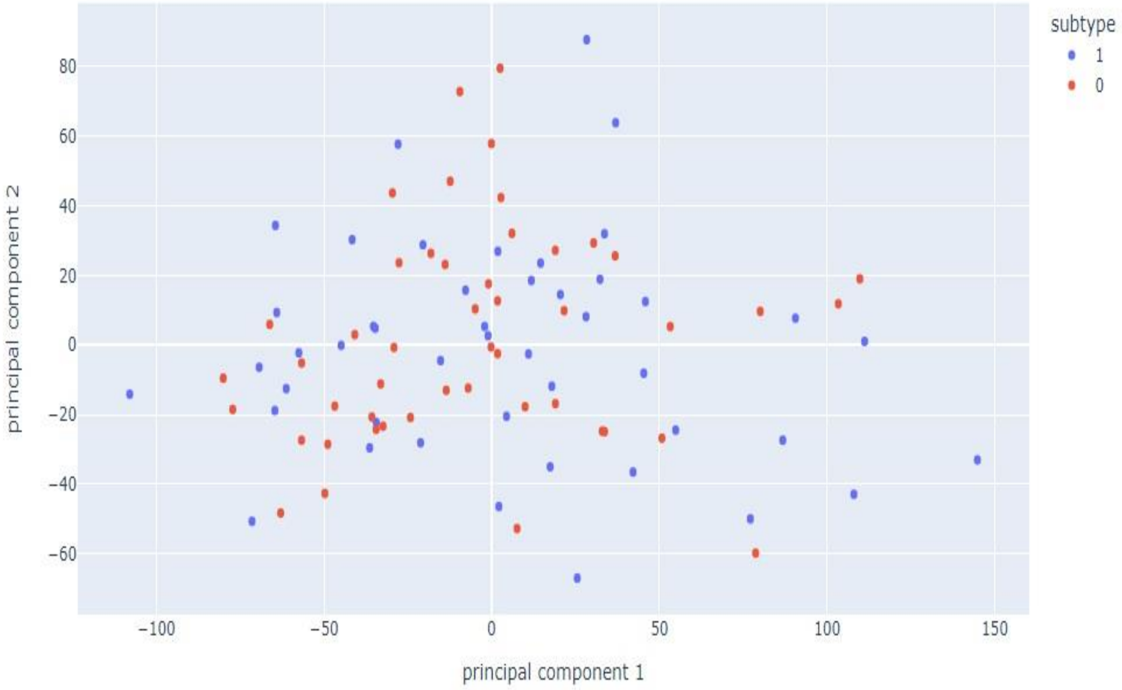


Figure 4: PCA on behavioural data

	principal component 1	principal component 2
0	-34.383521	-22.272995
1	111.214563	1.029841
2	-33.091694	-11.181818
3	-32.373415	-23.333119
4	42.171643	-36.490711
..	...	...
91	1.783274	-2.493788
92	-29.577811	43.674280
93	-1.047448	2.663224
94	144.844632	-33.008058
95	-35.700143	-20.687236

[94 rows x 2 columns]

Figure 5: Principal Components

ISSN 1533-9211

DOI: 10.5281/zenodo.8254867

KNN with these PCAs

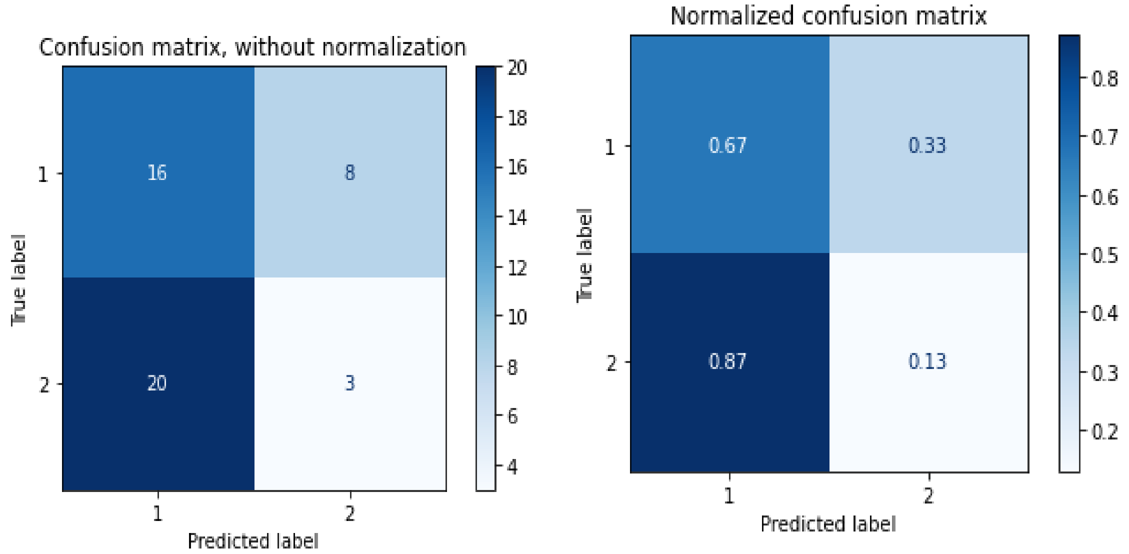


Figure 6: KNN classification

The classification of ADHD type with behavioural data using KNN without using PCA

As with EEG data, classification of ADHD subtype using behavioural data and PCA was low. We decided to test the decoding accuracy without the PCA, using the 6 variables from the IVA-II as features.

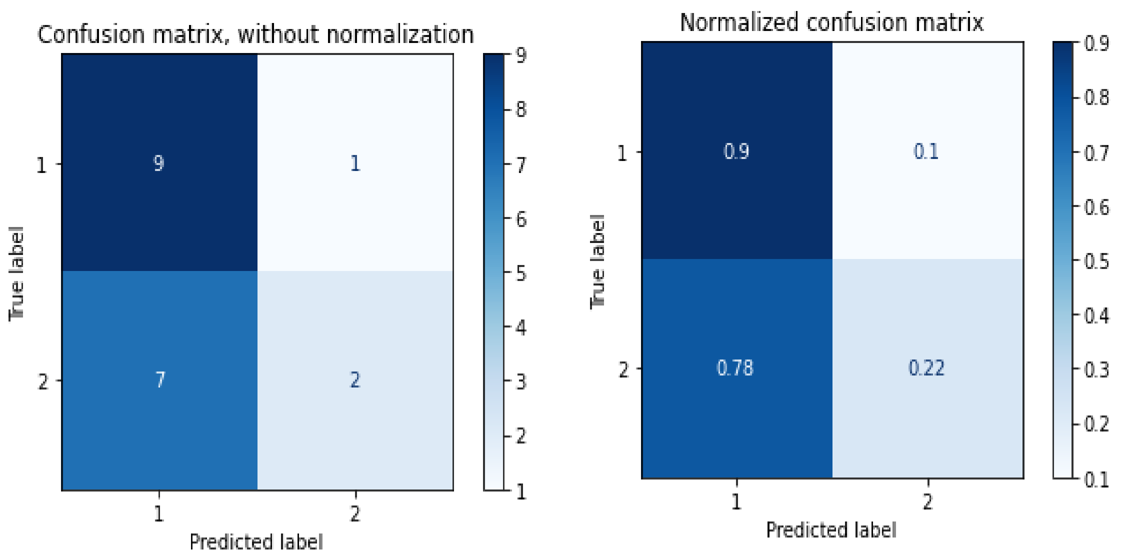
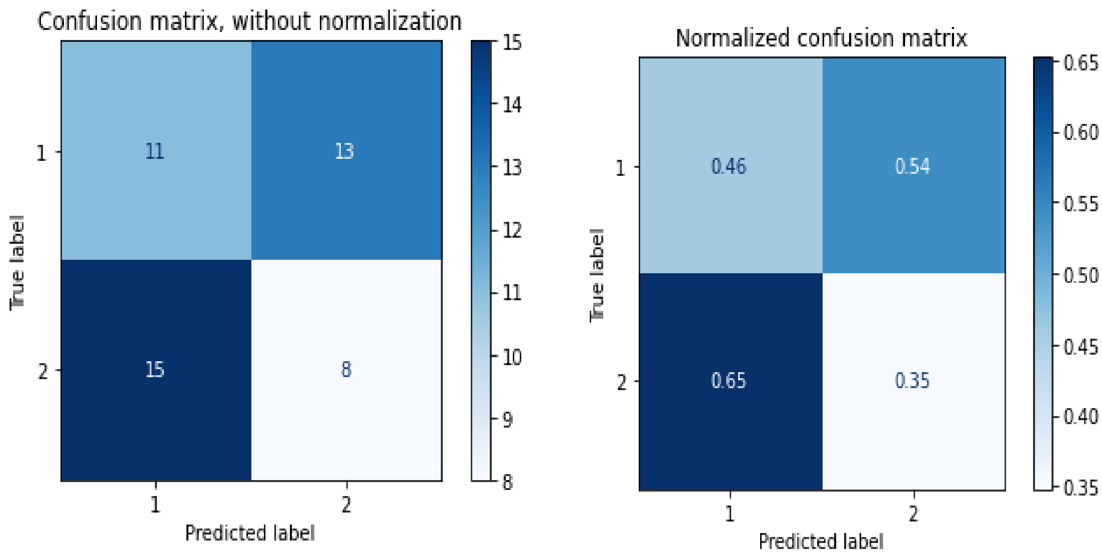


Figure 7: KNN classification without using PCA

The Seybold Report
ISSN 1533-9211
DOI: 10.5281/zenodo.8254867

The classification of ADHD type with behavioural data using PCA and Naive Bayes Classifier

The result of the classifier using PCA has been presented in Fig. 7. The confusion matrix with normalization and without normalization has been shown in the Fig. 8a and Fig. 8b respectively.



(a) **(b)**

Figure 8: Naive Bayes classification using PCA (a) Without normalization (b) With normalization

The classification of ADHD type with behavioural data using Naive Bayes Classifier without using PCA

The classification result of Naive Bayes classifier without using PCA has been presented in Fig. 9. The figure shows the confusion matrix of the classification with and without normalization.

ISSN 1533-9211

DOI: 10.5281/zenodo.8254867

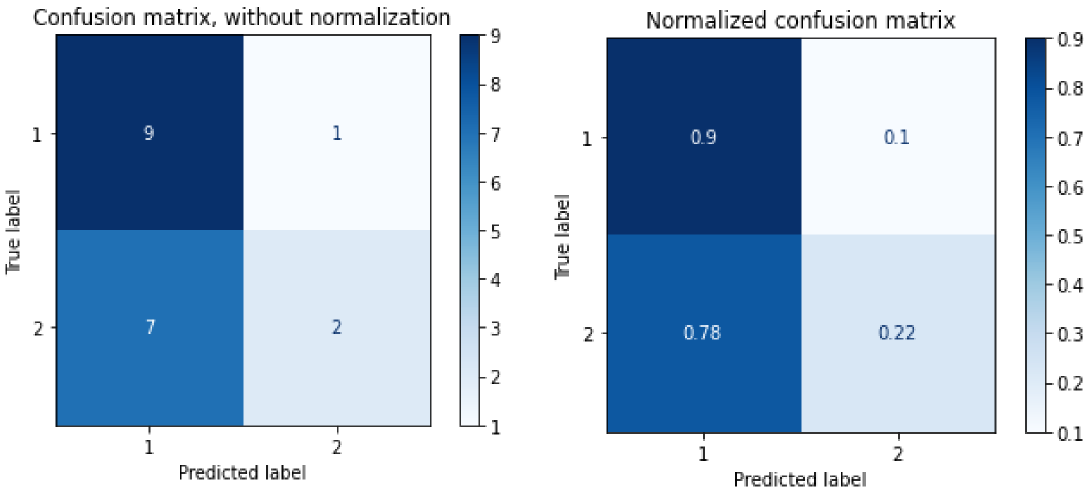


Figure 9: Naive Bayes classification without using PCA (a) Without normalization (b) With normalization

The accuracy and classifier score has been obtained and tabulated as in Table 2.

Table 2: Results of ML models

Using PCA			
Model	**Dataset**	**Accuracy**	**Classification score**
KNN	100	40%	0.57
	200	**55%**	0.51
Naive Bayes	100	40%	0.57
	200	51%	0.58
Without Using PCA			
KNN	100	58%	0.49
	200	**66%**	0.57
Naive Bayes	100	58%	0.49
	200	**65%**	0.57

CONCLUSION

In the present paper, two ML models namely KNN and Naive Bayes have been used for the classification of ADHD. The models have been compared using accuracy and classification score. In conclusion, the study aimed to evaluate the performance of the K-Nearest Neighbors (KNN) algorithm for predicting outcomes using a dataset of 100 and 200 samples. The results indicate that without employing Principal Component Analysis (PCA), the KNN algorithm achieved an accuracy of 66%, which is notably higher compared to the accuracy of 55% when PCA was used. This suggests that the application of PCA might have introduced some degree of feature reduction or noise, impacting the accuracy of the KNN algorithm. These findings underscore the importance of carefully considering the inclusion of PCA and its potential effects on algorithm performance.

ISSN 1533-9211

DOI: 10.5281/zenodo.8254867

References

1. Faraone, S. V., Banaschewski, T., Coghill, D., Zheng, Y., Biederman, J., Bellgrove, M. A., . . . Wang, Y. (2021). The World Federation of ADHD International Consensus Statement: 208 evidence-based conclusions about the disorder. Neuroscience & Biobehavioral Reviews. doi:10.1016/j.neubiorev.2021.01.022
2. W. Majeed, M. Magnuson, et al., "Spatiotemporal dynamics of low frequency BOLD fluctuations in rats and humans," Neuro Image, vol. 54, pp. 1140–1150, 2011.
3. Lindquist, Martin A, C. Waugh, and T. D. Wager. "Modeling state-related fMRI activity using change-point theory." Neuro image 35.3(2007):1125-1141.
4. Lian, Zhi chao, et al. "Exploring functional brain dynamics via a Bayesian connectivity change point model." 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI 2014) IEEE, 2014.
5. Chang, Catie, and G. H. Glover. "Time–frequency dynamics of resting-state brain connectivity measured with fMRI." Neuroimaging 50.1(2010):81-98.
6. Gülay Ç çek, A. Akan, and Z. Orman. "Classification of ADHD by using textural analysis of MR images." Medical Technologies National Congress 2017:1-4.
7. Miao, Bo, and Y. Zhang. "A feature selection method for classification of ADHD." International Conference on Information, Cybernetics and Computational Social Systems 2017:21-25.
8. Sudha, D., and M. P. Rani. "Gait Classification for ADHD Children Using Modified Dual Tree ComplexWavelet Transform."World Congress on Computing and Communication Technologies IEEE Computer Society, 2017:215-218.
9. Che-Wei Chang, Chien-Chang Ho and Jyh-Horng Chen ADHD classification by a texture analysis of anatomical brain MRI data.
10. Athena Taymourtash, Farnaz Ghassemi Independent Component analysis of sparse transformed EEG signal for ADHD/Normal Adult classification 978-1-4799-1972-7/15/> 31.00 c 2015 IEEE
11. Yan Zhang : ADHD classification by feature space separation with sparse representation: DSL 2018: 1-5
12. Marcano, Juan L. Lopez, M. A. Bell, and A. A. L. Beex. "Classification of ADHD and Non-ADHD using theta/beta power ratio features." IEEE Embs International Conference on Biomedical & Health Informatics IEEE, 2017.
13. F.M. Grisales-Franco, J.M. Medina-Salcedo, D.M. Ovalle-Mart nez3, J.D. Mart nez-Vargas1(B), D.G. Garc a-Murillo1, and G. Castellanos-Dominguez1 EEG Source Imaging Based on Dynamic Sparse Coding as ADHD Biomarker IWINAC 2017, Part I, LNCS 10337.
14. S. Ahmed Salman, Z. Lian, M. Saleem and Y. Zhang, "Functional Connectivity Based Classification of ADHD Using Different Atlases," 2020 IEEE International Conference on Progress in Informatics and Computing (PIC), 2020, pp. 62-66, doi: 10.1109/PIC50277.2020.9350749.
15. J. Wang, W. Liao and X. Jin, "Classification of ADHD Using fNIRS Signals Based on Functional Connectivity and Interval Features," 2021 6th International Conference on Computational Intelligence and Applications (ICCIA), 2021, pp. 113-117, doi: 10.1109/ICCIA52886.2021.00029.
16. S. Liu et al., "Deep Spatio-Temporal Representation and Ensemble Classification for Attention Deficit/Hyperactivity Disorder," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 1-10, 2021, doi: 10.1109/TNSRE.2020.3019063
17. Y. Tang, J. Jiang, M. Li, Y. Chen and X. Meng, "ADHD classification via auto-encoding network with non-imaging data fusion," 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2021, pp. 1328-1332.