Research Paper: Discrimination of ADHD Subtypes Using Decision Tree on Behavioral, Neuropsychological, and Neural Markers

Mohammad Rostami, Sajjad Farashi, Reza Khosrowabadi, Hamidreza Pouretemad

1. Institute for Cognitive and Brain Sciences, Shahid Beheshti University, Tehran, Iran.

ABSTRACT

Introduction: Attention-Deficit/Hyperactivity Disorder (ADHD) is a well-known neurodevelopmental disorder. Diagnosis and treatment of ADHD can often lead to a developmental trajectory toward positive results. The present study aimed at implementing the decision tree method to recognize children with and without ADHD, as well as ADHD subtypes.

Methods: In the present study, the subjects included 61 children with ADHD (subdivided into ADHD-I (n=25), ADHD-H (n=14), and ADHD-C (n=22) groups) and 43 typically developing controls matched by IQ and age. The Child Behavior Checklist (CBCL), Integrated Visual And Auditory (IVA) test, and quantitative EEG during eyes-closed resting-state were utilized to evaluate the level of behavioral, neuropsychology, and electrophysiology markers using a decision tree algorithm, respectively.

Results: Based on the results, excellent classification accuracy (100%) was obtained to discriminate children with ADHD from the control group. Also, the ADHD subtypes, including combined, inattention, and hyperactive/impulsive subtypes were recognized from others with an accuracy of 80.41%, 84.17%, and 71.46%, respectively.

Conclusion: Our results showed that children with ADHD can be recognized from the healthy controls based on the neuropsychological data (sensory-motor parameters of IVA). Also, subtypes of ADHD can be distinguished from each other using behavioral, neuropsychiatric and electrophysiological parameters. The findings suggested that the decision tree method may present an efficient and accurate diagnostic tool for the clinicians.
Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is regarded as a neurodevelopmental disorder of childhood. According to the American Psychiatric Association (American Psychiatric Association, 2013), ADHD with the prevalence of 5% can drastically influence children’s achievements in school, their social interactions, and their life quality. The core symptoms of ADHD are inattention, hyperactivity, and impulsivity. Moreover, it is a psychiatric disorder of childhood and continues until adulthood in about 50% of evaluated cases (Cheung, Rijdijk, McLoughlin, Faraone, Asherson, & Kuntsi, 2015; Franke et al., 2018). However, the diagnosis of ADHD is still a challenging issue, which roots in its etiologies. Accordingly, no single test is available to be used as the ‘gold standard’ for its diagnosis. Therefore, researchers, such as Barkley (1997) believed that clinicians should be able to examine multiple sources of evidence and put them together for diagnostics purpose.

One of the most effective methods for understanding ADHD is behavioral rating. In this regard, both parents and teachers are often requested to complete behavioral rating scales, like the Child Behavior Checklist (CBCL) (Achenbach & Rescorla, 2001) to determine the presence or absence of the ADHD symptoms. Zenglein et al. (2016) indicated the availability of four CBCL-derived subgroups related to ADHD. These subgroups are different based on their risk factors, along with the severity of psychopathological symptoms. Although behavioral rating can provide quantitative information on a wide range of behaviors, but parents and teachers are normally failed to make a distinction between the symptoms. On the other hand, there is a poor correlation between assessment results of the parents and the teachers. Moreover, they also do not have a proper understanding of the symptom (Snyder, Regino, Hornig, & Stein, 2015).

The importance of neuropsychological assessment in the evaluation of ADHD is based on the hypothesis that particular features of neuropsychological performance are related to the behavioral symptoms. These features are specified based on the diagnostic and statistical manual of mental disorders (Tinus) criteria (Kofler, Rapport, Bolden, Sarver, & Raiker, 2010). One of the tools that can be used for measuring cognitive performance of the subjects is Continuous Performance Test (CPT).

The CPTs are considered as neuropsychological tests, which measure an individual’s attention and impulsivity and can be implemented in clinical practices as a part of the diagnostic process. However, the utility of CPT as an assessment tool for the diagnosis of ADHD is still a challenging issue. Studies have shown that deficits in the CPT are not observed in all ADHD children (Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005; Trommer, Hoppen, Lorber, & Armstrong, 1988). Therefore, the CPT has not been suggested for individual diagnosis purpose. Nevertheless, the sensitivity and specificity of 94% for the diagnosis of ADHD using the Integrated Visual and Auditory test (IVA) as a CPT test have been reported (Ti-
It is a useful, non-invasive technique, with excellent temporal resolution. EEG-based studies have reported an increase in slow waves, mainly in theta, and a decrease in fast waves, especially in beta in children with ADHD (Barry, Clarke, & Johnston, 2003; Buyck & Wiersema, 2014; Kitsune et al., 2015). Matastra, Lubar, and Linden (2001) implemented theta and beta power to identify the individuals with and without ADHD with 90% sensitivity and 94% specificity. The theta to beta ratio (TBR) was approved by the Food and Drug Administration (FDA) in 2013 as a diagnostic aid marker for the ADHD. Nonetheless, there are still some inconsistent findings about the validity of TBR as a diagnosis marker (Arms, Conners, & Kraemer, 2013; Loo & Makeig, 2012).

The present study aimed at classifying the children based on their behavioral, neuropsychological and electrophysiological data. In this regard, classification algorithms, such as decision trees, discriminant analysis, rule-based methods, logistic regression, and neural networks can be used for categorizing the related features. Studies have shown the capability of the decision trees to accurately handle a multi-class classification problem (Liu & Wiener, 2002). A decision tree is considered as a top-down structure of nodes with directed edges commonly used in data mining studies. The feature data of an input set is divided into different branches based on their information gain. The tree is formed by assigning the feature with the highest gain ratio as the root and the process is continued until either all classes are recognized or the stopping criteria are reached (Han, Pei, & Kamber, 2011). Data gathering, processing, and analysis are described in the following sections.

2. Method

2.1. Sample and procedure

A sample of 104 boys was evaluated using CBCL, IVA, and EEG, respectively. Accordingly, the subjects with an age range of 7-12 years were divided into four groups, including 61 children with combined (n=22), inattentive (n=25), hyperactive/impulsive (n=14) ADHD subtypes, and 43 typically developing controls (TDC). Based on the results, as shown in Table 1, no significant differences were observed in age and IQ scores between the groups.

The children with ADHD were recruited from different Child Psychiatry Clinics in Tehran. The Persian version of the Structured Clinical Interview for DSM-5 (SCID-5) was used as the diagnostic assessment, which was conducted by a board-certified child and adolescent psychiatrist and a senior clinical psychologist. The average SCID scores of the inattentive, hyperactive/impulsive, and combined DSM-5 symptoms in the ADHD group (n=61) were 7.0 (SD=1.3), 8.2 (SD=2.9) and 6.3 (SD=1.3), respectively. The TDC group included the children without any abnormality based on the DSM-5 criteria selected from two elementary schools in Tehran. All subjects had normal intelligence scores using the Raven’s Progressive Matrices (90-124), right-handed based on self-report, with a normal or corrected-to-normal vision/hearing. However, a history of problematic prenatal or neonatal periods, brain damage, central nervous system diseases, convulsive disorders, and sensorimotor deficits were considered as the exclusion criteria.

2.2. Measures

2.2.1. Behavioral rating

All subjects were evaluated by the Child Behavior Checklist (CBCL, 1) after it was answered by their parents. The CBCL/6-18 is a standardized form, which is filled out by parents for explaining their children’s behavioral and emotional symptoms. The questionnaire includes 118 items for rating a child’s behavior during the last six months. The six DSM-oriented scales were used in this study. The validity and reliability of the test could be found at http://www.aseba.org/.

2.2.2. Neuropsychological test

To evaluate inattention and impulsivity in both visual and auditory modalities, the Integrated Visual and Auditory (IVA) test (Sandford & Turner, 2000) is considered as a kind of CPT. Using this test, the commission errors...
(i.e. impulsivity and response inhibition, reflected in IVA “Prudence” scores), omission errors (i.e. inattention, reflected in “Vigilance” scores) and mean reaction time (Speed”) can be evaluated. Furthermore, the IVA can differentiate different types of ADHD, including predominantly inattentive, predominantly hyperactive-impulsive, and combined subtypes (Sanford & Turner, 1995).

2.2.3. Neurophysiology

EEG data were registered using a 19-electrode Mitsar amplifier (www.mitsar-medical.com) with the sampling rate of 250 Hz. Electrodes were placed on the scalp by a standard 10-20 montage (Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, T5, T6, Pz, P3, P4, O1, O2). Besides, the average of ear-channels was utilized as the reference and FPz as the ground electrode, and the electrode impedances were kept below 5 kΩ.

After recording, a self-written program using Matlab (https://www.mathworks.com) and EEGLab (https://sccn.ucsd.edu/eeglab/) software was implemented to process the data. Standard preprocessing consisted of band-pass filtering (1-40 Hz), segmenting the data into epochs of 1-second duration, automatic rejection of disrupted channels conducted using probability, spectrum, and kurtosis criteria. Interpolating the rejected channels was done by averaging their spherical neighbors, eliminating unreliable epochs, and resorting to the mean of all the channels. In the next procedure, Fast Fourier Transform (FFT) was implemented for transforming the preprocessed EEG data into the frequency domain, and the absolute power and relative power of the data was computed in Delta (1-4 Hz), Theta (4-8 Hz), Alpha (8-13 Hz), and Beta (13-30 Hz) frequency bands accordingly. Finally, the power ratio was evaluated for the theta/beta frequency band.

2.2.4. Classification paradigm

In the present study, a decision tree method was used to classify the four above-mentioned groups using their behavioral, neuropsychological and neurophysiological data. To perform the classification using decision tree paradigm, first, a ‘training’ set of input parameters was utilized based on behavioral, neuropsychological, and neurophysiological data of the subjects. Then, a decision tree structure was formed based on the information gain of each input parameter. The extracted model was used as a predictive decision support model for the data of new subjects.

A decision tree has a flowchart-like upside-down tree structure (Han, Pei, & Kamber, 2011), in which a parameter with the highest information gain among all the parameters is considered as the root of the tree (e.g. sensory-motor parameter of IVA). Subsequently, other nodes are created and the tree branches expand. The nodes indicate the best parameter to separate the subjects. The expansion of the tree structure stops when all classes are recognized or the maximum level of expansion is reached. After constructing the model, the decision tree structure schematically demonstrates several features and conditions that are used to categorize all the subjects into their appropriate classes. Interpretation of the tree is based on the rules that can be used for the identification of a class of subjects.

3. Results

The results of the comparison of four classes of the subjects based on their behavioral, cognitive and neural markers using a decision tree algorithm are presented in Figure 1. It should be noted that all features were used as input parameters of the decision tree, and the nodes were selected based on their information gain scores. Table 2 also presents the interpretation of the results presented in Figure 1. Each rule explains how we could discriminate a class of subjects and the most important features and their score criteria are also presented.

### Table 1. The mean and standard deviation of the groups based on age and intelligence

<table>
<thead>
<tr>
<th>Measure</th>
<th>TDC</th>
<th>ADHD-I</th>
<th>ADHD-H</th>
<th>ADHD-C</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>43</td>
<td>25</td>
<td>14</td>
<td>22</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age</td>
<td>9.3±1.4</td>
<td>8.7±1.2</td>
<td>9.4±1.7</td>
<td>9.1±0.99</td>
<td>2.03</td>
<td>0.11</td>
</tr>
<tr>
<td>IQ</td>
<td>102.9±10.9</td>
<td>103.5±10.12</td>
<td>107.5±9.27</td>
<td>103.1±11.02</td>
<td>0.1</td>
<td>0.95</td>
</tr>
</tbody>
</table>

The accuracy of the decision tree classifier for discrimination of the four classes of the subjects was 84.01% based on the test data in a one-leave-out manner, where one sample was left out from each group for evaluating the classifier in each step during the partitioning the data.

A graphical representation of the results is presented in Figures 2 and 3. As presented in Figure 1, the sensory-motor visual as a cognitive score could discriminate typical group from others with 100% accuracy. According to Figure 2, adding mood affective disorder as a behavioral feature was effective to discriminate inattentive subjects.

Table 2. The rule sets for the decision tree to classify four classes of subjects

<table>
<thead>
<tr>
<th>Path</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>If Sensory Motor &lt;54 then class “Typical”</td>
</tr>
<tr>
<td>2</td>
<td>If Sensory Motor ≥54 and Mood Affective Disorder ≥89 then class “Inattentive”</td>
</tr>
<tr>
<td>3</td>
<td>If Sensory Motor ≥54 and Mood Affective Disorder &lt;89 and Visual Prudence ≥82.5 and power ratio (δ/γ)_Fp1 ≤3.78 then class “Hyperactive”</td>
</tr>
<tr>
<td>4</td>
<td>If Sensory Motor ≥54 and Mood Affective Disorder &lt;89 and Visual Prudence ≥82.5 and Absolute power β_Fp1 &lt;4.1 then class “Combined”</td>
</tr>
<tr>
<td>5</td>
<td>If Sensory Motor ≥54 and Mood Affective Disorder &lt;89 and Visual Prudence ≥82.5 and power ratio (δ/γ)_Fp1 &gt;3.78 then class “Inattentive”</td>
</tr>
<tr>
<td>6</td>
<td>If Sensory Motor ≥54 and Mood Affective Disorder &lt;89 and Visual Prudence &lt;82.5 and Absolute power β_Fp1 &lt;4.1 then class “Inattentive”</td>
</tr>
</tbody>
</table>

The accuracy of the decision tree classifier for discrimination of the four classes of the subjects was 84.01% based on the test data in a one-leave-out manner, where one sample was left out from each group for evaluating the classifier in each step during the partitioning the data.

A graphical representation of the results is presented in Figures 2 and 3. As presented in Figure 1, the sensory-motor visual as a cognitive score could discriminate typical group from others with 100% accuracy. According to Figure 2, adding mood affective disorder as a behavioral feature was effective to discriminate inattentive subjects.
Table 3. Performance of the decision tree classifier to identify four different classes of subjects

<table>
<thead>
<tr>
<th>Class of subjects</th>
<th>Classification Accuracy (±SD)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical</td>
<td>100 (±0.00)</td>
<td>0</td>
</tr>
<tr>
<td>Combined</td>
<td>80.41 (±9.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Inattentive</td>
<td>84.17 (±4.56)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hyperactive</td>
<td>71.46 (±8.95)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Figure 2. Decision rules to discriminate typical subjects from Attention Deficit Hyperactivity Disorder (ADHD) groups based on the neuropsychological measures (sensory-motor score in Integrated Visual And Auditory (IVA) test)

Figure 3. Representation of δ/γ ratio at Fp1 location for separating different types of Attention Deficit Hyperactivity Disorder (ADHD)
with high accuracy, whereas it was not effective to differentiate between combined and hyperactive subjects. Figure 3 also shows that adding neuropsychological features (delta/gamma ratio at channel FP1) with visual prudence quotient is also helpful to discriminate a hyperactive subject with a high accuracy rate. These results indicate that behavioral, cognitive, and neuropsychological features have different levels of importance for discrimination of a specific subtype of ADHD.

4. Discussion

A large body of research has focused on the use of data-driven techniques to classify clinical diagnostics data accurately. Our main aim in this study was to evaluate the utility of decision tree algorithm in three levels of behavioral, neuropsychological and electrophysiological data, as well as using it for classification of a population of typically developed, and three subtypes of ADHD children. Based on the results, the decision tree was effective to classify the above-mentioned groups with an accuracy of 84.01%. Specifically, the sensory-motor score in IVA test could perfectly discriminate typical individuals from the ADHD groups (100% accuracy). Besides, the mood affective disorder, as a CBCL parameter, along with the visual prudence in IVA test and beta at Fp1 and delta/gamma ratio at Fp1, as EEG parameters, showed a nearly perfect classified ADHD-C from other groups with 80.41% accuracy. Also, the mood affective disorder plus visual prudence and delta/gamma ratio at Fp1 classified those with the inattentive type of ADHD group with an accuracy of 84.17%. Finally, the mood affective disorder along with visual prudence and delta/gamma ratio at Fp1 distinguished those with hyperactive type of ADHD from other groups with an accuracy of 71.46%. To the best of our knowledge, this is the first study, which represented such a perfect diagnostic classification for the ADHD subtypes.

This classification method is a highly reliable and accurate approach and has been confirmed for clinical diagnosis. Based on the results, the sensory-motor parameter could recognize the children with and without ADHD and consequently can identify ADHD subtypes better when behavioral and electrophysiological and other neuropsychological variables were added. Thus, the result is a highly sensitive and specific when the measures are integrated, compared with the cases, in which each measure is used alone for classifying the ADHD subtypes.

Interestingly, the cutoff score of 54 for the sensory-motor parameter implemented in the decision tree was only at the 56 percentile of standardized samples, which is considered as the average range of the samples. This fact proposes that such cut-off alone does not present a clinical impairment. However, when the sensory-motor score is implied in combination with other measures could enhance the probability of accurate diagnosis. Thus, this method can help the clinicians to design accurate and economically rational diagnostics. In this context, the study has pinpointed the sensory-motor deficits, which are observed in the ADHD individuals, such as concerns with motor sequence learning (Adi-Japha, Fox, & Karni, 2011), motor response to stimuli (Gorman Bozorgpour, Klorman, & Gift, 2013), fine motor skills, and articulation (Iwanaga, Ozawa, Kawasaki, & Tsuchida, 2006).

Previous studies, such as Finch, Davis, and Dean, (2015) have also reported that sensory-motor can identify children with ADHD with an accuracy of 95%. Assessment of sensory and motor deficits is considered as a main element of the neuropsychological assessments because sensory-motor deficits can be associated with dysfunctions in the central nervous system. In general, a sensory-motor deficit represents an underlying neurological etiology that may lead to some behavioral changes in children. Such behavioral changes have been reported in previous studies. For instance, Davis, Pass, Finch, Dean, and Woodcock, (2009) indicated a positive relationship between sensory-motor functioning and academic achievement, and cognitive processing in children with ADHD.

On the other hand, previous studies using machine learning techniques have also demonstrated the effectiveness of the automatic algorithms to categorize children with and without ADHD using behavioral and neuropsychological measures. Various classification algorithms have been used. For instance, Bledsoe et al. (2016) classified healthy controls and ADHD combined type individuals using their behavioral symptoms and neuropsychological performances. Using the decision tree algorithm, they could reach an accuracy of 100%. Also, Cohen (2013) also showed that the decision tree can classify autistic and ADHD subjects from healthy individuals based on their behavioral profiles with an accuracy of 94% and 87%, respectively. Another study by Santos, Bastos, Andrade, Revoredo, and Mattos, (2011) implemented naive Bayes and decision tree algorithms to classify the data obtained for ADHD children while playing a computer game. The naive Bayes algorithm could produce 68% sensitivity and 67% specificity. Nevertheless, the decision tree algorithm was only able to produce 55% of sensitivity and specificity. However, the sensitivity of the algorithm was less or equal to 0.22 to
discriminate ADHD subtypes. Compared with previous studies, the results of the current study showed that EEG features could cover the heterogeneity of ADHD subtypes and make them more distinguishable.

Despite the high accuracy rate of classification of ADHD subtypes and healthy controls, this study had some limitations. For instance, a larger sample size and measurement of other neuropsychological tests must also be considered before generalizing the results to other populations. Moreover, it is proposed to include ADHD comorbid type, such as oppositional defiant and conduct disorder in future studies.

Ethical Considerations

Compliance with ethical guidelines

The subjects were assessed after obtaining the informed consent signed by the parents, and the research protocol was approved by the Ethics Committee of Iran University of Medical Sciences.

Funding

This research was supported by grants from the Cognitive Sciences and Technologies Council of Iran (Grant No.: 1948).

Authors' contributions

All authors contributed in designing, running, and writing all parts of the research.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgments

We would like to thank Dr. Masoud Nosratabadi and Prof. Reza Rostami for data acquisition of the ADHD group.

Reference


