A Rule-Based System for ADHD Identification using Eye Movement Data

Senuri De Silva
Department of Computer Science and Engineering, University of Moratuwa
Sri Lanka
ktsdesilva.15@cse.mrt.ac.lk

Sanuwani Dayarathna
Department of Computer Science and Engineering, University of Moratuwa
Sri Lanka
sanuwaniudara.15@cse.mrt.ac.lk

Gangani Ariyarathne
Department of Computer Science and Engineering, University of Moratuwa
Sri Lanka
gcariyarathne.15@cse.mrt.ac.lk

Dulani Meedeniya
Department of Computer Science and Engineering, University of Moratuwa
Sri Lanka
dulanim@cse.mrt.ac.lk

Sampath Jayarathna
Department of Computer Science, College of Science, Old Dominion University,
Virginia, USA
sampath@cs.odu.edu

Anne M.P. Michalek
Communication Disorders and Special Education, Old Dominion University
Virginia, USA
aperrott@odu.edu

Gavindya Jayawardena
Department of Computer Science, College of Science, Old Dominion University,
Virginia, USA
gavindya@cs.odu.edu

***Abstract*— Attention Deficit Hyperactivity Disorder (ADHD) is one of the common psychiatric disorder in childhood, which can continue to adulthood. The ADHD diagnosed population has been increasing, causing a negative impact on their families and society. This paper addresses the effective identification of ADHD in early stages. We have used a rule-based approach to analyse the accuracies of decision tree classifiers in identifying ADHD subjects. The dataset consists of eye movements and eye positions of different gaze event types. The feature extraction process considers fixations, saccades, gaze positions, and pupil diameters. The decision tree-based algorithms have shown a maximum accuracy of 84% and classification rule algorithms have shown an accuracy of 82% using eye movement measurements. Thus, both algorithms have shown high accuracy with the rule-based component.**

***Keywords*— ADHD, eye movements, rule-based, decision tree, classification rules**

I. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is defined as a consistent pattern of inattention, hyperactivity, and impulsivity, which is at a higher rate compared to other control groups in the level of development [1]. It is known as the most controversial and contentious onset of child disorder with a high rate of prevailing into adulthood. Recent statistics have shown that the ubiquity of the disorder covers 7% of the entire world population while 6.4% children of age between 7 and 14 are diagnosed in the United States [2] and 7.1% of the test participants are identified as ADHD positive in Sri Lanka [3].

Generally, children with ADHD are commonly entangled with learning difficulties. The symptoms of hyperactivity lie in the fields of creativity, curiosity, and enthusiasm for knowledge [4][5]. These behaviours and learning disorders can be developed into comorbid disorders such as obsessive-compulsive disorder, oppositional defiant disorder, conduct disorder, and learning and language disorders, if not treated well [3]. Also, it is found that hyperactivity and aggressive behaviour are prominent behaviours that cause comorbidity [6]. However, the main symptoms of ADHD, including hyperactive, inattentive and impulsive behaviours of children are often ignored by conventional social myths. Thus, early intervention and treatment of ADHD are important to handle their social interactions, attentiveness and control hyperactivity.

Many studies [7], [8], [9] have addressed different data types in classifying ADHD such as fMRI, EEG, clinical data and eye movement data. The integration of these data types would provide a better identification accuracy in identification process than addressing only one data type. Compared to other data types, eye movements have distinguishable characteristics of being less complex and reduced dimensionality of data on its original form for the identification of neurological disorders. However, eye movement-based identification is not widely used in related to ADHD due to the difficulties in the experimental and data collection process. As the novelty of this research, we consider eye movement data to address the existing issues in identifying ADHD.

Studies have shown that there is a relationship between ADHD subjects and their eye movements [10]. For instance, ADHD patients' saccadic movements indicate the activity of frontal cortex and basal ganglia through Oculomotor tasks. Moreover, Mahone et al. [11] have shown that the ADHD children had a significant shortfall in the response preparation (in terms of the response latency and variability) and inhibition of visually guided saccades, through an oculomotor paradigm-based experiment. Eye movement data can be obtained from specially designed experiments to identify ADHD, such as Oculomotor tasks [10], [12]. However, eye movements can be affected by other reasons or disorders, as well. Thus, learning models with ADHD related eye movements data can be used to obtain better results that are distinguished from other disorders.

The general hypothesis of this study is to measure the feasibility of the decision tree and classification rule algorithms in generating the rule-based component. Thus, this study identifies ADHD using eye movement data with a rule-based decision support system. The feature selection and rule derivation are based on optimized data mining algorithms.

The paper is structured as follows. Section II explores the theoretical and practical applications on ADHD classification. Section III presents the design of the proposed solution with the system workflow. Section IV describes the experimental setup with the implementation of the rule-based component. Section V evaluates the results, and Section VI concludes the paper.

978-1-7281-3632-5/19/$31.00 ©2019 IEEE

II. BACKGROUND

A. Overview of ADHD

The symptoms of the neuropsychiatric disorder are often misidentifying as inappropriate hyperactivity. A study [13] shows that 40% of the ADHD subjects tend to impair the disorder into adulthood. Major symptoms of ADHD are issues in socializing, learning difficulties, less productivity than other control groups, relationship issues and anxiety management issues. The severity of this disorder varies with age, based on the treatments [14]. Further, non-psychiatric diseases such as eczema, gastrointestinal problems, respiratory infections, asthma, allergies, and seizures are common among the diagnosed children. Hence, there is a need for early detection and prevention methodologies for this neuropsychiatric disorder.

Many studies have been conducted to diagnose ADHD using clinical data such as EEG and fMRI. Most of them have shown a significant improvement in classification between ADHD and other control subjects [15], [16], [17]. Although these methodologies have performed well, the machine learning approaches [16] still lack the accuracy to be used in real practice. Thus, there is a need for research on efficient approach for early intervention of ADHD to prevent severe future symptoms. One possible approach is to derive an objective measure to classify ADHD with healthy groups [18]. Objective measures provide a robust method to derive information to diagnose ADHD, as sole clinical and other rating methods are not reliable in detecting ADHD, which depends on physicians and other social aspects.

B. Eye Movement Data for ADHD Classification

Eye movements can be classified as horizontal, vertical and torsional movements. The centre of each axis is at the centre of the eyeball. The eye rotations are achieved by contraction and relaxation of six extraocular muscles. Hence the rotations can be classified as (1) ductions, by considering the monocular eye movements, (2) versions that describe the binocular conjugate movements of both eyes and (3) vergences that consider the disjunctive binocular actions [19]. Eye rotation data acquisition require high-end eye trackers such as Tobii Pro X2-60 with over 1000hz sampling rate. The eye movement data in neurological disorders provide rich data set with simple vasomotor baseline and indicate a complex behavioural process of those disorders. Since it does not need any advanced cognitive skills, eye movements can be easily performed by children as well [20].

The relationship between eye movements and ADHD [12], has revealed that the ADHD subjects take longer reaction time and identical variability compared to control subjects. Adults with ADHD have been found with different eye movement patterns, and children with ADHD are in difficulty to maintain fixations. Generally, the measurements of gaze points are taken for the experiment are in the Active Display Coordinate System Pixels (ADCSpx) format, where the eye movement data is mapped to 2D coordinates [21]. Fig. 1 demonstrates the active display coordinate system with the origin at the top left corner.

This paper focuses on gaze types, including fixations and saccades, where the cognitive inferences can be obtained. The fixations are the series of gaze points when the attention is fixed into one point. It is a processed output of a series of gaze points with fixation duration associated apart from the spatial and temporal components [22]. Saccades are known as the rapid movements of fovea from one point to another point of interest [22]. Since saccades also contain useful information on the cognitive state inference, it is important in this study. The data collection of this paper is based on 60hz Tobii eye tracker.

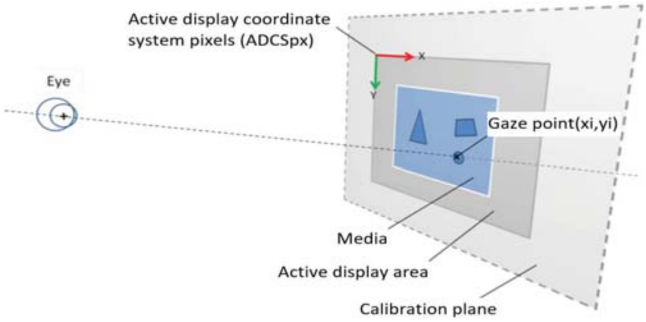

Fig. 1. Gaze Point data collection in ADCSpx [21]

C. Data Mining Techniques

Data mining techniques are used in various applications to classify, analyse and predict data. Table I summarizes some of the classification techniques with their advantages and limitations. A model is trained to make predictions on different class labels in training data set and use to predict distinct classes for a new sample of data. The results from different classification algorithms can be used to make decisions on generalized data set. For the classification of ADHD and non–ADHD classes, decision tree techniques can be used with a series of rules [24].

The decision tree structure consists of both internal and external nodes that are connected via branches. The internal nodes make decisions on determining the next visiting node and external nodes related to class labels. The large decision trees are first generated, and the size is reduced by pruning the tree [24]. Apart from the decision tree algorithms, classification ruling algorithms have used to generate the rules. These algorithms generate disjunctive rules which can be easily implemented in the system. Different decision tree algorithms such as J48, Random Tree, Random Forest, Linear Model Tree (LMT), REPTree, Hoeffding Tree, Decision Stump and classification ruling algorithms such as PART and JRip have discussed in Table I.

D. Rule-Based Decision Support System

A rule-based system is a set of if-then rules for data classification. Each rule specifies an inference step and input data match with each condition in the rules. If the data matches with the left-side, it will execute the decision on the right-side of the rule. One input may satisfy multiple rules, as some conditions are declared under more than one rule [24].

The decision tree is a commonly used classifier due to its accuracy and simplicity compared to other methods. However, there can be interpretation issues when the increase of the scope [26]. Thus, defining a rule-based classifier by extracting If-then rules from decision tree enables easy understanding. One rule is defined for every path from the root node to a leaf node to extract rules from a decision tree. Each splitting principle on different paths is connected using logical AND operations to form the 'IF' part of the rule and the 'THEN' part of the rule is created using the class predictions hold by each leaf node [28].

TABLE I. CLASSIFICATION TECHNIQUES

Technique	Description	Benefits	Limitations
Logistic Model Tree [23]	A decision tree with logistic regression where an attribute is tested in each inner node as in an ordinary tree. This method uses linear regression in each leaf.	High accuracy.	Less efficient. More classification times.
Decision stumps [24]	A decision tree with one-level, with one internal node that is connected to the terminal nodes. In boosting, weights are assigned to models. Missing values are considered separately and expanded to another branch from the stump.	Less time to classify data. Use as a boosting method for a single input model.	Less accurate.
Reduced Error Pruning Tree [24]	Pruning is used to reduce the tree size by removing parts of the tree.	Fast decision learner. Build the tree by reducing the variance, Linear computational complexity.	Lead over pruning if the test set is smaller than the training set.
Hoeffding Tree [25]	A decision tree with Hoeffding bound. An incremental approach to comply with new data in parallel processing. Provides a confidence level for the best attribute to split the decision tree in order to decide the number of instances.	Good for streaming data. High accuracy. Scales diverse attributes with less memory and better sampling.	Classification speed is less. Risk of overfitting.
Random Tree [26]	Uses supervised learning to create a decision tree with random data. Uses the best split in the attributes to split the tree on each node. Considers the random subset of attributes and selects the best split for the defined subset without computing for each node.	Efficient classification, High accuracy with random tree combinations. Use for classification and regression.	Low performance on imbalanced data.
J48 [26]	Builds the decision tree based on labelled data. Analysis using attribute changes that split the data set into subsets. Algorithm iterates on the subsets until all the elements in the subset fits the same class. A leaf node is created to show the class selection.	Simple implementation with high efficiency and accuracy, Handle both nominal and numeric values.	Moderate classification time.
Random Forest [27]	Based on random decision forests that support classification and regression. Build using many decision trees.	Recognizes outliers and anomalies in labelled data. Estimates useful features and support high accuracy. Measures pairwise proximity between samples.	Some classifications are hard to interpret. overfits the data with noisy classification.
PART[28]	A decision list algorithm which uses separate-and -conquer method to generate a practical decision tree making the best leaf to a rule. Then for a rule generation, the leaf with the largest coverage is selected, and the whole pruned tree is discarded.	Provides high accurate and compact rule set. The simplicity of the rule definition and generation. Flexible and efficient. No over-pruning effects.	Global optimization is avoided.
JRip/ Ripper [28]	A propositional rule learner that creates rules using Repeated Incremental Pruning to Produce Error Reduction (RIPPER).	Uses a post-processing phase for optimization.	More sophisticated than other algorithms.

Rule-based Decision Support System (DSS) determines the involvement of decision behaviour and the problem [29]. It allows interactive and efficient exchange of data between the system and the user. DSS supports a problem-oriented approach in problem-solving to determine a justifiable result. DSS is expected to address all the requirements of the decision maker

III. SYSTEM DESIGN

The system is designed using a rule-based component to develop a decision support system to diagnose ADHD with eye movement data. Fig. 2 shows the high-level design process. The pre-processing data module consists of a combination of dimensional reduction methods, including the removal of missing, noisy and inconsistent data [29]. Then, we extract the features from the dataset; fixations and saccades. They are used as related gaze parameters to extract features. The features are extracted from eye movement data by computing the mean, standard deviation and duration of fixations and saccades by measuring the pupil diameters of left and right eyes. Next, we have applied the feature selection methods to identify the relevant features using a rule set based on different techniques such as filtering, wrapping and embedded methods [30].

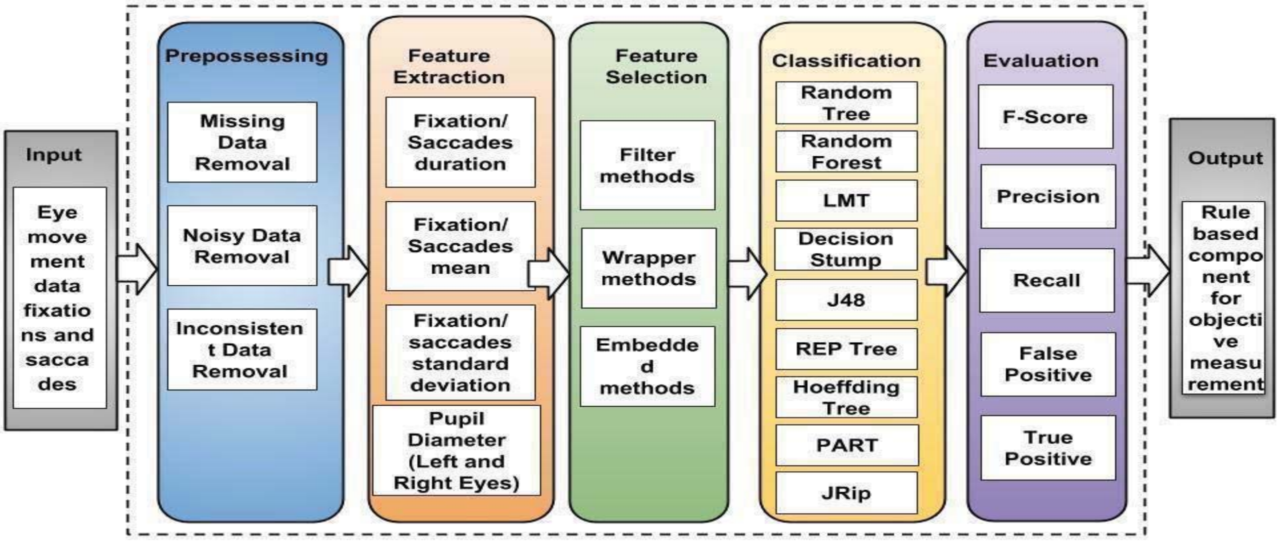

Fig. 2. System overview diagram

Filtering techniques only examine the feature relevance by checking the essentiality of the data. In wrapper methods, a possible feature set is defined, and different feature sets are generated to evaluate. Embedded methods are used to search an optimal set of features in ADHD classifier construction.

The rule-based component consists of a set of classification algorithms, which are iteratively evaluated to produce the optimal result. The best rule set is opted out by considering Random Tree, Random Forest, J48, Linear Model Tree (LMT), REP Tree, Decision Stump algorithms. With random data, the random tree is used to construct the decision tree, and Random forest is generated by many decision trees that apply for regression and classification. The classification model LMT acts as a combination of decision tree and logistic regression. REP Tree develops a decision tree using a gain of information and Hoeffding tree has better attribute comparison capability and less memory consumption compared to other methods [25]. PART and JRip are used as a ruling classification algorithms. Fig. 3 shows the corresponding workflow of the DSS.

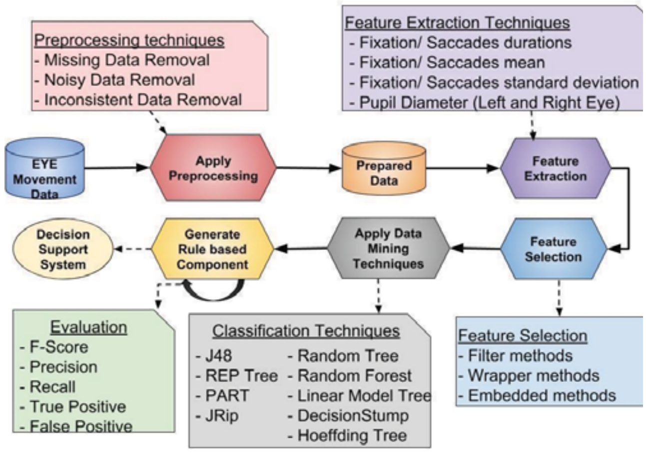

Fig. 3. The workflow of the system

Further, standard information retrieval evaluation metrics are used to evaluate the accuracy of the outcome. Accuracy is the most fundamental criteria of evaluating, and we also used precision, recall, and F1 for the evaluation of the rule-based component [30]. These evaluations are based on the confusion matrix of the evaluation measure to get the ratio of total observations and observations of correctly predicted. Precision is an evaluation of positive observation between total observations to the correctly predicted items. The ratios of positive observations that are correctly predicted to all the observations are determined by the evaluation of Recall. F1- Score which takes both false negative and false positive, is a weighted average of Recall and Precision.

IV. Implementation

A. Experimental Setup

The dataset was obtained from an experiment conducted at Old Dominion University, Virginia, United States. A Tobii pro-X2-60 eye tracker was used to measure the eye movements of participants with Tobii Studio eye tracking software to process and analyse the data gathered by the eye tracker [21]. The Tobii eye tracker was mounted with a computer screen to record the gaze events at a sampling rate of 60Hz. The eye tracker also has an accuracy of 0.4° and precision of 0.32^0 (30Hz version) and 0.34^0 (60Hz version) [22]. A similar experimental setup for eye movement data collection was held for Autism Spectrum Disorder]. This provides details on the gaze positions of the participants and other parameters related to eye movements such as pupil diameter. The eye movements and eye positions were measured based on corneal reflection technology, which directs infrared light towards the eyes and tracks the reflections between the cornea and the pupil by an infrared camera.

A total of 14 participants between the ages of 18-65 years were engaged in this experiment. Among 14 participants, seven adult participants were diagnosed with ADHD, including six females and one male with an average age of 22.85 and standard deviation of age with 3.01. These 7 participants were medically proven with diagnosed with ADHD and confirmed with other formalisms. They were asked to remain medication free for 12 hours before the experiment and informed the associated risks. The participants were native English and reported with normal vision, no cognitive impairments or psychotic symptoms.

In order to gather eye movement data, 14 participants were seated in front of the monitor with the Tobii eye tracker to detect the corneal reflections of the subjects while performing a given task. After describing the experiment for the participants, their position and distance were modified to maintain the same viewing angle in the given task. Then they were asked to read sentences that they see on the computer screen along with the letter at the end of each sentence, followed by a question mark. These sentences were grouped in various sets which have 2-5 sentences. They were also asked to state if the given sentence makes any sense or not, as they understood.

After gathering gaze position details, fixations and saccades were derived as two major gaze related parameters to do further analysis. Saccades are rapid movements of eyes which suddenly change from one point to another interesting point. Fixations are the moments that the gaze position is at one position on the screen. Thus, eye fixations and saccades related features were generated to analyse eye movements. These data include participants' gender, the number of fixations and saccades, the normal, average and standard deviation of the duration of the fixation and saccade in milliseconds, pupil diameters of both left and right eyes and the class label of the participant, which state if the subject is ADHD diagnosed or not.

B. Rule-Based Component

The rule-based component is an expert system which uses the rules generated from the rule and tree algorithms considered for the classification. The main goal of this module is to implement the derived rules to classify ADHD. The rules generated are derived from both decision tree and classification rules. The decision tree algorithms use divide and conquer method to create the rules from the given dataset. Algorithm 1 shows the selection of a suitable classification algorithm.

This study addresses the feasibility of decision trees and classification algorithms in producing rule-based components. There are two types of rule generation in a decision tree approach: (1) compares a single node with a constant, (2) consider two attributes or a function that combines multiple attributes to compare each other and derive value. The leaf node of a decision tree gives the class where a new dataset or a point can traverse through the tree and reach a corresponding class. The decision tree algorithms considered in this paper are J48, Random Forest, Random Tree, REP Tree, LMT, Hoeffding Tree, and Decision Stump.

We used classification rules to generate the rules as another method. The rules are combined with AND operation and derived a set of rules that can implement. Predominantly, the rules generated from classification ruling algorithms are not combined with AND operations, but they are the logical expressions derived after passing a set of tests. The focus is to generate simple rules compared to decision trees, which are not pruned to skip redundancy tests.

Another goal is to take advantage of distinction, which cannot be achieved by decision trees. We used PART, and JRip classification ruling algorithms and the number of rules generated from each algorithm are 43 and 11 rules, respectively.

```
Algorithm 1: Rule-based Module
Dataset: Eye movement data
Result: Best rules with the corresponding accuracy
function generate_rules:

dataset=[]
preprocessed_dataset=preprocess(dataset)
decision_tree_algorithms=[]
decision_tree_rules=[]
decision_tree_accuracy=[]
classification_rule_algorithms=[]
classification_rules=[]
classification_rule_accuracy=[]

for i in decision_tree_algorithms:
  results= decision_tree_algorithms[i].
           apply(preprocessed_dataset)
  rules= generateRules(results.tree)
  if results.accuracy > accuracy :
        accuracy = results.accuracy
        decision_tree_rules= rules

for i in classification_rule_algorithms:
  results= classification_rule_algorithms[i].
           apply(preprocessed_dataset)
  rules=results.rules
  rule_accuracy=0
  rules.sort()
  rules=rules[rules.length/2:]
  best_rules=[]
  for r in rules:
      if rule_accuracy< r.accuracy:
        best_rules=r
        rule_accuracy=r.accuracy
  if classification_rule_accuracy< rule_accuracy:
      classification_rules= best_rules
      classification_rule_accuracy= rule_accuracy

highest_accuracy=max(classification_accuracy,
                     decision_tree_accuracy)
if highest_accuracy ==  classification_accuracy:
   rules= classification_rules
else:
   rules= decision_tree_rules
return rules,highest_accuracy
```

V. Evaluation

The accuracy of each decision tree classifiers is measured in terms of precision, recall, and F1-measures to identify the best classifier for ADHD under a given set of features. According to the results shown in Fig. 4, all the considered algorithms show a similar flow of values in terms of the accuracy measures. Since the rule-based component needs generating rules or decision trees, where the Random Forest and LMT fail to produce, it can be deduced that J48 and Random Tree algorithms would give the best results out there.

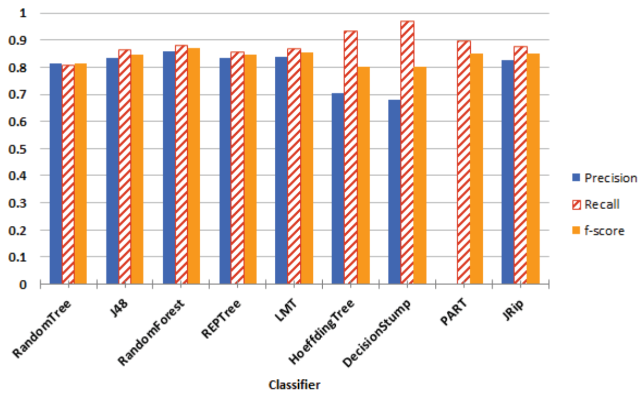

Fig. 4. Accuracy measures decision tree and classification ruling classifiers

J48 is widely known for its efficiency, and Random Tree is a highly accurate algorithm with an attribution selection filter applied. Also, when the classification ruling algorithms are considered, the PART algorithm takes the lead in the considered accuracy measures. The generated Receiver Operating Characteristic (ROC) diagram for the considered decision tree algorithms is shown in Fig. 5 and depicts a close flow of the 7 algorithms.

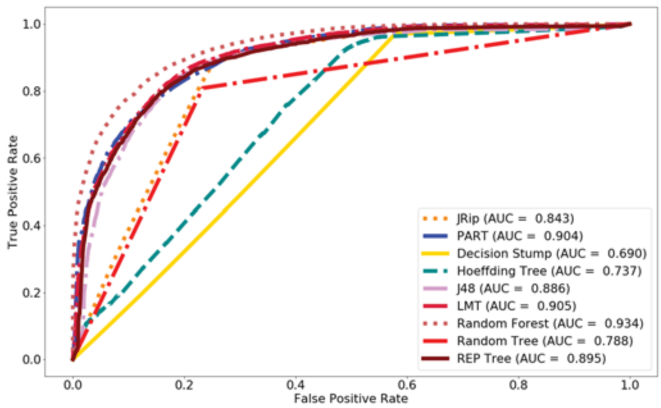

Fig. 5. ROC graph of the decision tree and classifier ruling algorithms

This graph shows the trade-off between the false positive (FP) rates and the true positive (TP) rates of the ADHD as the class variable. Also, J48 and Random Tree algorithms show similar capabilities, and in the rule generation, they could be declared as the best algorithms available. The algorithms are tested using the creation of rules, 10-fold method, and 66 percentage splitting methods. Since the dataset is limited to obtain the classification rules and to test the rule, it is fair that every data should be included in the testing for the model.

Based on the results shown in Table II, the 10-fold evaluation gives high accuracies. In order to validate the 10-folds method is usable in this context, we have also split the dataset into 2:1 as 66% data is used for the training and 33% for data for the testing. Similar approach [9] using eye movement data of the traces of recorded from 15 minutes of videos, has obtained an accuracy of 77.3% in classifying ADHD. Thus, the proposed method shows high accuracy compared to the related studies on eye movement data.

TABLE II. ACCURACY OF DECISION TREE AND CLASSIFICATION RULING

Classifier	Accuracy (10-folds)	Accuracy (66% split)
Random Tree	79.01	78.84
J48	82.72	81.74
Random Forest	85.31	84.48
REPTree	82.59	82.68
LMT	83.46	82.66
Hoeffding Tree	74.19	73.34
Decision Stump	72.82	72.68
PART	82.10	82.11
JRip	82.57	82.57

VI. CONCLUSION

Attention Deficit Hyperactivity Disorder is an impactful disorder that carries a genetic cause and has a higher probability of continuing into adulthood. Hence early detection reduces the effect of the disorder. This paper has addressed the issue of early detection using eye movement data via a rule-based system. The selected algorithms of rule generation have shown between 0.81 and 0.83 precision in terms of the performance measure considered. Also, the rules generated via the classification ruling algorithms have shown the precision of 0.89, which is a better set of rules for the given dataset. Hence, it can be concluded that both decision tree and classification ruling algorithms can be used to generate the rule-based system.

Further, this work can be extended by considering more eye movement features like saccade speed, saccade amplitudes. Moreover, a combination of currently studied ADHD classification data such as EEG, fMRI data also can be used along with the eye movement data to obtain a composite measure or a mathematical score.

ACKNOWLEDGEMENT

This work was supported by Old Dominion University, Virginia. Authors would like to thank for providing the eye movement dataset in the experiment. Also, we acknowledge the support received from the Senate Research Committee Grant SRC/LT/2019/18, University of Moratuwa, Sri Lanka.

REFERENCES

[1] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders: DSM-5*, 5th ed. Washington (D.C.): American Psychiatric Association, 2013.

[2] ADD Resource Center, "ADHD Numbers: Facts, Statistics, and You", *ADD Resource Center*, 2019. [Online]. Available: https://www.addrc.org/adhd-numbers-facts-statistics-and-you. [Jan. 24, 2019].

[3] S. H. Kariyawasam, A. Koralagama, P. Jayawardane, B. Karunathilake, V. Perera and H. Perera. "A descriptive study of attention deficit hyperactivity disorder (ADHD) at Lady Ridgeway," *Sri Lanka Journal of Child Health*, vol. 31, pp. 109-114, 2002.

[4] W. Dodson, "Attention Deficit-Hyperactivity Disorder: The Basics and the Controversies," Underst. Our Gift, vol.14, pp. 17–21, 2002.

[5] L. Honos-Webb, *The gift of ADHD: How to transform your child's Problems into strengths*.Oakland, CA: New Harbinger,2005,pp.89-106

[6] A. Hamed, A. Kauer and H. Stevens, "Why the Diagnosis of Attention Deficit Hyperactivity Disorder Matters", *Frontiers in Psychiatry*, vol. 6, pp. 168-177, Nov. 2015.

[7] J. W. Kim, V. Sharma, and N. D. Ryan, "Predicting methylphenidate response in ADHD using machine learning approaches," *International Journal of Neuropsychopharmacology*, vol. 18, no. 11, pp. 1–7, 2015.

[8] D. Kundra and B. Pandey, "Classification of EEG based Diseases using Data Mining", *International Journal of Computer Applications*, vol. 90, no. 18, pp. 11-15, 2014.

[9] P. H. Tseng, I. G. M. Cameron, G. Pari, J. N. Reynolds, D. P. Munoz, and L. Itti, "High-throughput classification of clinical populations from natural viewing eye movements," *Journal of Neurology*., vol. 260, no. 1, pp. 275–284, 2012.

[10] D. P. Munoz, I. T. Armstrong, K. A. Hampton, and K. D. Moore, "Altered Control of Visual Fixation and Saccadic Eye Movements in Attention-Deficit Hyperactivity Disorder", *Journal of Neurophysiology,* vol. 90, no. 1, pp. 503–514, 2006.

[11] E. Mahone, S. Mostofsky, A. Lasker, D. Zee and M. Denckla, "Oculomotor Anomalies in Attention-Deficit/Hyperactivity Disorder: Evidence for Deficits in Response Preparation and Inhibition", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 48, no. 7, pp. 749-756, 2009.

[12] P. Deans, L. O'Laughlin, B. Brubaker, N. Gay and D. Krug, "Use of Eye Movement Tracking in the Differential Diagnosis of Attention Deficit Hyperactivity Disorder (ADHD) and Reading Disability", *Psychology*, vol. 01, no. 04, pp. 238-246, 2010.

[13] Wiley, "ADHD symptom persistence into adulthood estimated", *ScienceDaily*, 2019. [Online]. Available: https://www.sciencedaily.com/releases/2016/10/161020132300.htm. [Feb. 24, 2019].

[14] National Center for Health Statistics, *Health, United States, 2012*. Hyattsville, MD: U.S. Department of Health and Human Services, 2013, pp. 147-151

[15] A. Hao, B. He and C. Yin, "Discrimination of ADHD children based on Deep Bayesian Network", in Proc. *2015 IET International Conference on Biomedical Image and Signal Processing*, Beijing, China, 2015.

[16] A. Tenev, S. Markovska-Simoska, L. Kocarev, J. Pop-Jordanov, A. Müller and G. Candrian, "Machine learning approach for classification of ADHD adults", *International Journal of Psychophysiology*, vol. 93, no. 1, pp. 162-166, 2014.

[17] B. Solmaz, S. Dey, A. Rao and M. Shah, "ADHD classification using bag of words approach on network features", in Proc. *Medical Imaging: Image Processing*, San Diego, California, USA, 2012.

[18] N. Rommelse, S. Van der Stigchel and J. Sergeant, "A review on eye movement studies in childhood and adolescent psychiatry", *Brain and Cognition*, vol. 68, no. 3, pp. 391-414, 2008.

[19] A. Wong, *Eye movement disorders*. Oxford: Oxford University Press, 2008, pp. 4-5.

[20] Armstrong and D. Munoz, "Inhibitory control of eye movements during oculomotor countermanding in adults with attention-deficit hyperactivity disorder", *Experimental Brain Research*, vol. 152, no. 4, pp. 444-452, 2003.

[21] Tobii, *Tobii Studio User's Manual Version 3.4.5*. Tobii AB, 2016.

[22] Tobiipro, "Types of eye movements", *Tobiipro*, 2019. [Online]. Available: https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/types-of-eye-movements/. [Feb. 20, 2019].

[23] N. Landwehr, M. Hall and E. Frank, "Logistic Model Trees", *Machine Learning*, vol. 59, no. 1-2, pp. 161-205, 2005.

[24] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*, 2nd ed. San Francisco: Morgan Kaufmann, 2005.

[25] A. Kumar, P. Kaur, and P. Sharma, "A Survey on Hoeffding Tree Stream Data Classification Algorithms." CPUH-Research Journal, vol. 1, no. 2, pp. 28–32, 2015.

[26] M. Kantardzic. *Data Mining Concepts, Models, Methods, and Algorithms*, John Wiley & Sons, Inc., Hoboken, 2011.

[27] G. Louppe, "Understanding random forests: From theory to practice", *arXiv preprint arXiv:1407.7502*, 2014.

[28] V. Parsania, N. Jani, N. Bhalodiya, "Applying Naïve Bayes, BayesNet, PART, JRip and OneR Algorithms on Hypothyroid Database for Comparative Analysis", *Journal of Darshan Institute on Engineering Research and Emerging Technologies*, vol. 3(1), pp. 60-64, 2014.

[29] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*, Waltham: Morgan Kaufmann Publishers, 2012.

[30] S.Wibowo and H. Deng, "A rule-based decision support system for evaluating and selecting IS projects", in *International MultiConference of Engineers and Computer Scientists*, Hong Kong, 2008.

[31] S. Thapaliya, S. Jayarathna and M. Jaime, "Evaluating the EEG and Eye Movements for Autism Spectrum Disorder", in *2018 IEEE International Conference on Big Data (Big Data)*, Seattle, WA, 2018, pp. 2328-2336