A Machine Learning Study of Comorbidity of Dyslexia and Attention Deficiency Hyperactivity Disorder

Junaita Davakumar, Arul Siromoney

A Machine Learning Study of Comorbidity of Dyslexia and Attention Deficiency Hyperactivity Disorder

Junaita Davakumar[0000-0002-5947-2446] and Arul Siromoney[0000-0002-4587-7303]

College of Engineering Guindy, Anna University, Chennai, India
junaitadavakumar@gmail.com, arul.siromoney@gmail.com

Abstract. Neurodevelopmental disorders in children like dyslexia and ADHD must be diagnosed at earlier stages as the children need to be provided with necessary aid. Comorbidity of dyslexia and ADHD is very high. Children with comorbidity of dyslexia and ADHD face comparatively more difficulty than children with just one of the disorders. Since all the three, dyslexia, ADHD and comorbid cases share many similar characteristics, it is hard to distinguish between cases which have only dyslexia or ADHD and those which have both. Manual analysis to differentiate based on standard scores of the psycho analysis tests provided inconsistent results. In this paper, we have applied standard machine learning techniques Random Forest, Support Vector Machine and Multilayer Perceptron to the diagnosis test results to classify between ADHD and comorbid cases, and dyslexia and comorbid cases. Analysis using the different individual psycho analysis tests is also done. Application of machine learning techniques provides better classification than the manual analysis.

Keywords: Attention Deficiency Hyperactivity Disorder, Dyslexia, Comorbidity, Machine Learning, Classification.

1 Introduction

Dyslexia and Attention Deficiency Hyperactivity Disorder (ADHD) are both neurodevelopmental disorders which have a high rate of comorbidity or combined occurrence [1]. These neurodevelopmental disorders are chronic and should be diagnosed at childhood, to provide support [2]. These neurodevelopmental disorders make the children depressed as they are unable to perform well academically like their peers with no ailments. This is mainly because of the lack of guidance and proper diagnosis of the disorders during childhood. Dyslexia is the inability to read or spell though the person is highly capable and has adequate intelligence [3]. ADHD is the lack of attention, presence of high impulsivity in individuals [4]. Children with comorbid issues face more difficulties than the children with just one of the disorders [5, 6]. Though the rates of comorbidity or combined occurrences of ADHD and dyslexia are high, they are individual diseases and do not cause each other. Although they are different diseases, individuals with dyslexia and individuals with ADHD share certain similar characteristic traits. It has been found through various studies of families and twins that both ADHD and dyslexia are inherited [7, 8, 9, 10]. ADHD and Dyslexia have been routinely diagnosed using various psycho analysis tests which do not provide precise outcomes. Diagnosis of comorbid cases of dyslexia and ADHD is even more strenuous. When a case is not correctly diagnosed as being comorbid, it means that necessary treatment is not provided for one of the two illnesses.

In this paper we have used standard machine learning techniques Random Forest (RF), Support Vector Machine (SVM) and Multilayer Perceptron (MLP) for classifying and analyzing dyslexia and comorbid cases, ADHD and comorbid cases to achieve better outcomes. The impact of various psycho analysis tests on the classification is also analyzed. To our knowledge, there is no prior work on the analysis of comorbidity of dyslexia and ADHD, and its classification, using machine learning techniques.

2 Related Work

Although there are many genetic studies of dyslexia and ADHD, there are not many conclusive results that link the two disorders. Work has been done to identify genetic, cognitive and neural overlap between the two disorders. Sanchez et al. have studied the association of genes with dyslexia, ADHD and comorbid samples as well [11]. Marino et al. [12] have investigated a strategy to clarify which genes are important for dyslexia. Eva et al. [1] have analyzed the link between dyslexia and ADHD from epidemiological, genetic, neurofunctional, neuropsychological and therapeutic perspectives.

Structural and functional magnetic resonance imaging combined with comprehensive behavioural testing has been used to characterize the behavior of comorbid dyslexia and reading disability (RD) [13]. Lauren et al. have done an analysis of voxel-based morphometry studies to find whether there is any overlap in the gray matter correlates of dyslexia and ADHD [14]. Comorbidity between dyslexia and ADHD has also been clarified by investigating cognitive endophenotypes [15].

3 Materials and Methods

3.1 Methods Used

SVM [16] is a machine learning algorithm that uses supervised learning for classification. Using the training samples, a hyperplane is constructed to separate the samples into two classes. This hyperplane is used to classify the new samples. Random Forest [17] is also a machine learning algorithm that is used for classification and regression. Multiple decision trees are grown using the training samples. To classify a new sample, each tree gives a class and that class is assumed to be given a vote. The class that gets the maximum number of votes is decided as the class for the new sample. MLP [18], a type of feedforward artificial neural networks with at least three layers of nodes and nonlinear activations is used for classification. MLP uses a supervised learning technique called backpropagation for training.

3.2 Dataset Description

In this study a public domain dataset [19, 20] has been used. The dataset includes the results of psycho analysis tests on 26 children with only dyslexia, 27 children with only ADHD and 27 children with comorbid ADHD and dyslexia. The dataset attributes are in four sections namely, demographics, dyslexia tests, ADHD tests and motor skills tests. All the children took up the tests along with their legal guardian and have given a written informed consent [20].

Demographics. This section of the dataset attributes includes age, sex, TONI 4 test and handedness. TONI 4 (Test of non-verbal intelligence, Fourth Edition) tests the non-linguistic and non-motor skills of the examinee, limiting to the analysis of the general intelligence. Abstract reasoning is tested and cognitive abilities such as reading, writing, speaking and listening are avoided. Handedness questionnaire includes eighteen questions based on certain actions testing the hand preference of the examinee.

Dyslexia tests. This section of the dataset attributes includes pseudo-word decoding and spelling tests (Wechsler Individual Achievement Test). Wechsler Individual Achievement Test, Second Edition, is a standard academic based test that is used for the identification of learning disabilities. In Pseudo-Word Decoding subtest, the phonetic knowledge is tested by making the examinee read out meaningless words aloud. In the spelling subtest, words, letters and combination of words are dictated, for which the examinees are expected to write spellings.

ADHD tests. Conners 3-Parent test is a behavioral questionnaire which is a commonly used test for the diagnosis of ADHD is included in this section.

Motor skills tests. The motor skills tests include the attributes Grooved pegboard (GPB) and the Leonard Tapping Task (LTT). Grooved pegboard test is for the analysis of fine motor skills using dexterity and Leonard Tapping Test is for the analysis of simple and complex motor skills.

3.3 Experiments

We have classified between dyslexia and comorbid cases, and ADHD and comorbid cases using the machine learning techniques RF, SVM and MLP. Machine learning classifiers are applied on the results of ADHD tests, dyslexia tests, motor skills tests and demographics, of children with just ADHD, just dyslexia and combined occurrence of dyslexia and ADHD. For the classification using the entire dataset, overall accuracy, receiver operation characteristic and f-measure are found. For the analysis using the individual and combination of psycho analysis tests, overall accuracy alone is found. 3-fold cross validation is done for each classifier. The public domain tool WEKA is used.

4 Results and Analysis

The standard score for the dyslexia tests is 80 and children scoring below 80 are diagnosed with dyslexia. Using the standard score, only 3 out of 26 (11%) children were classified correctly. For the ADHD tests, the standard score is 60 and children scoring above 60 are diagnosed with ADHD. Using this, 19 out of 27 (70%) children were classified correctly. For a child to be diagnosed with comorbidity of dyslexia and ADHD it should pass the criteria of both dyslexia and ADHD tests. Using this, only 6 out of 27 (25%) children were classified correctly. True condition in all the subtests was considered for the classification.

4.1 Comparison of the results of dyslexic and comorbid cases

Table 1: Overall dataset analysis using RF, MLP and SVM (dyslexia vs. comorbid)

CLASSIFIERS	OVERALL PREDICTION ACCURACY	RECEIVER OPERATION CHARACTERISTIC	F-MEASURE
Random Forest	71.6%	0.761	0.717
Support Vector Machine	64.1%	0.642	0.642
Multilayer Perceptron	66.0%	0.660	0.748

In the overall analysis of the results of the psycho analysis tests of dyslexic and comorbid cases using machine learning techniques, it is seen that random forest classifier performs better when compared to support vector machine and multilayer perceptron. The entire dataset has been used for the classification (Table 1).

Table 2: Prediction accuracy using different individual tests (dyslexia vs. comorbid)

CLASSIFIERS TESTS	RF	SVM	MLP
Demographics	62.2%	50.9%	58.4%
ADHD tests	73.5%	81.1%	84.9%
Dyslexia tests	49.0%	50.9%	43.3%
Motor skills test	60.3%	56.6%	50.9%

In the results shown in Table 2, it is seen that ADHD tests provide better classification than any other psycho analysis test. It is seen that dyslexia tests provide bad results. This is due to the fact that, the difference between the comorbid cases and the dyslexic cases would be the lack of ADHD characteristics in the dyslexic cases.

4.2 Comparison of the results of ADHD and comorbid cases

Table 3: Overall dataset analysis using RF, MLP and SVM (ADHD vs. comorbid)

CLASSIFIERS	OVERALL PREDICTION ACCURACY	RECEIVER OPERATION CHARACTERISTIC	F-MEASURE
Random Forest	53.7%	0.533	0.597
Support Vector Machine	66.6%	0.667	0.665
Multilayer Perceptron	62.9%	0.656	0.625

In the overall analysis of the results of the tests of ADHD and comorbid cases using machine learning techniques, it is seen that support vector machine classifier performs better when compared to random forest and multilayer perceptron. The entire dataset has been used for the classification (Table 3).

Table 4: Prediction accuracy using different individual tests (ADHD vs. comorbid)

CLASSIFIERS TESTS	RF	SVM	MLP
Demographics	44.4%	42.5%	40.7%
ADHD tests	37.0%	46.2%	44.4%
Dyslexia tests	66.6%	70.3%	66.6%
Motor skills test	55.5%	57.4%	50%

In the results shown in Table 4, it is seen that dyslexia tests provide better classification than any other psycho analysis test. It is seen that ADHD tests provide bad results. This is due to the fact that, the difference between the comorbid cases and the ADHD cases would be the lack of dyslexia characteristics in the ADHD cases.

5 Conclusion

Differentiating between ADHD and comorbid cases as well as dyslexia and comorbid cases is a difficult task as they share common characteristics. From the above analysis it is seen that, applying machine learning techniques to the results of the psycho analysis tests rather than using standard cut off scores for the diagnosis has proved to classify better. Doctors can thus use machine learning techniques for the preliminary diagnosis of comorbid cases of dyslexia and ADHD.

References

1. Eva, G., Antonella, G.M.D., Paolo, C.: Comorbidity of ADHD and Dyslexia. Developmental Neuropsychology, 35(5), 475-493 (2010). https://doi.org/10.1080/87565641.2010.494748.
2. Tager-Flusberg, H.: Neurodevelopmental disorders, The MIT Press, Cambridge, Massachusetts, 1999.
3. Vellutino, F.R.: Dyslexia: Theory and research, MIT Press, Cambridge, MA, 1979.
4. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Diseases (DSM- V), 5th edn. American Psychiatric Publishing, Washington, DC (2013).
5. Willcutt, E. G., Pennington, B. F., Boada, R., Ogline, J. S., Tunick, R. A., Chhabildas, N. A., et al.: A comparison of the cognitive deficits in reading disability and attention deficit /hyperactivity disorder. Journal of Abnormal Psychology, 110, 157–172 (2001).
6. Mental health problems in neurodevelopmental disorders. http://www.kcl.ac.uk/ioppn/depts/cap/research/MentalHealthProblemsinNeurodevelopmentalDisorders (accessed 05 February 2020).

7. Del'homme, M., Kim, T. S., Loo, S. K., Yang, M. H., Smalley, S. L.: Familiar association and frequency of learning disabilities in ADHD sibling pair families. Journal of Abnormal Child Psychology, 35, 55–62 (2007).
8. Friedman, M. C., Chhabildas, N., Budhiraja, N., Willcutt, E. G., Pennington, B. F.: Etiology of the comorbidity between RD and ADHD: Exploration of the non-random mating hypothesis. American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics, 120, 109–115 (2003).
9. Gayan, J., Olson, R. K.: Genetic and environmental influences on orthographic and phonological skills in children with reading disabilities. Developmental Neuropsychology, 20, 483–507 (2001).
10. Willcutt, E. G., Pennington, B. F., DeFries, J. C.: A twin study of comorbidity between attention-deficit/hyper-activity disorder and reading disability. American Journal of Medical Genetics (Neuropsychiatrics Genetics), 96, 293–301 (2000).
11. Sanchez-Moran, M., Hernandez, J.A., Dunabeitia, J.A., Estevez, A., Barcena, L., Gonzalez-Lahera, A., et al.: Correction: Genetic association study of dyslexia and ADHD candidate genes in a Spanish cohort: Implications of comorbid samples. PLoS ONE, 13(12): e0209718, (2018). https://doi.org/10.1371/journal.pone.0209718.
12. Marino, C., Giorda, R., Vanzin, L., Molteni, M., Lorusso, M. L., Nobile, M., et al.: No evidence for association and linkage disequilibrium between dyslexia and markers of four dopamine-related genes. European Child & Adolescent Psychiatry, 12, 198-202 (2003).
13. Langer, N., Benjamin, C., Becker, B.K.C., Gaab, N.: Comorbidity of reading disabilities and ADHD: Structural and functional brain characteristics. Human Brain Mapping, 40(9), 2677-2698 (2019). doi: 10.1002/hbm.24552
14. Lauren M. McGrath, Catherine J. Stoodley: Are there shared neural correlates between dyslexia and ADHD? A meta-analysis of voxel-based morphometry studies. Journal of Neurodevelopmental Disorders, 11(31) (2019).
15. Debbie, G., Margaret, S., Charles, H,: Time perception, phonological skills and executive function in children with dyslexia and/or ADHD symptoms. Journal of Child Psychology and Psychiatry, 52(2), 195–203 (2011).
16. Cortes, C., Vapnik, V.: Support-vector networks. Machine Learning, 20(3), 273-297 (1995). doi:10.1007/BF00994018.
17. Breiman, L.: Random Forests. Machine Learning, 45(5) (2001). doi: 10.1023/A:1010933404324.
18. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by error propagation, Rumelhart, D.E., McClelland, J.L. (eds.), Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1 MIT Press, (1986).
19. Marchand-Krynski, M-E., Morin-Moncet, O., Belanger, A-M., Beauchamp, M.H., Leonard, G.: Shared and differentiated motor skill impairments in children with dyslexia and/or attention deficit disorder: From simple to complex sequential coordination, PLoS ONE, 12(5) (2017). https://doi.org/10.1371/journal.pone.0177490.
20. [dataset] https://doi.org/10.1371/journal.pone.0177490.s001