Machine Learning Approach Applied to the Prevalence Analysis of ADHD Symptoms in Young Adults of Barranquilla, Colombia

Alexandra Leon-Jacobus[1], Paola Patricia Ariza-Colpas[2(✉)],
Ernesto Barcelo-Martínez[2,3], Marlon Alberto Piñeres-Melo[4],
Roberto Cesar Morales-Ortega[2],
and David Alfredo Ovallos-Gazabon[5]

[1] Universidad Metropolitana, Barranquilla, Colombia
aleonj@unimetro.edu.co
[2] Universidad de la Costa, CUC, Barranquilla, Colombia
{parizal,ebarcelo,rmoralesl}@cuc.edu.co
[3] Instituto Colombiano de Neuropedagogia, Barranquilla, Colombia
[4] Universidad del Norte, Barranquilla, Colombia
pineresm@uninorte.edu.co
[5] Universidad Simón Bolivar, Barranquilla, Colombia
david.ovallos@unisimonbolivar.edu.co

Abstract. Disorder Attention Deficit/Hyperactivity Disorder, or ADHD, is recognized as one of the pathologies of high prevalence in children and adolescents from the global environment population; this disorder generates visible symptoms usually diminish with the passage of time in adulthood, however they remain concealed by demonstrations damnifican personal stability and human development apt. This article shows the results of the research aimed at determining the prevalence of symptoms of attention deficit hyperactivity disorder in Young Adults University of Barranquilla and its Metropolitan Area. The sample of 1600 young adults between 18 and 25 years, which has been estimated at 95% confidence level and a margin of error of 2.44%. The information was acquired through the application of exploratory instruments symptoms of attention deficit hyperactivity disorder. With the application of the algorithm different machine learning algorithms such as: Bagging, MultiBoostAB, DecisionStump, LogitBoost, FT, J48Graft, a high performance in the Bagging algorithm could be identified with the following results in quality metrics: Accuracy 91.67%, Precision 94.12%, Recall 88.89% and F-measure 91.43%.

Keywords: ADHD disorder · Prevalence of symptoms · Pathology · Hyperactivity · Impulsivity · Classification techniques

1 Introduction

Attention Deficit Hyperactivity Disorder, known by its Spanish acronym as ADHD, is a topic of interest for professionals in psychology, psychiatry and neurosciences in general, due to the negative impact it has on social functioning, personal, work,

academic and family of those who suffer from it. Recently, the American Psychiatric Association defines Attention Deficit Hyperactivity Disorder in DSM V as a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development, which is characterized by the presence of symptoms of inattention and/or hyperactivity and impulsivity that are present before the age of 12, and affect the different contexts in which an individual develops.

Attention Deficit Hyperactivity Disorder has been considered one of the most prevalent diseases in children and adolescents worldwide [1]; and although their externally visible symptoms tend to decrease over the years in adulthood, they tend to remain hidden, behind manifestations that affect the personal stability and proper development of the human being [2]. Studies show; that of the patients diagnosed with ADHD in childhood, from 30% to 70% continue to present symptoms that generate difficulties during adolescence and adulthood, in addition; at the age of 19, 38% still fully meet the diagnostic criteria of the pathology (without remission); and 72% have at least one third of the symptoms required for diagnosis (with partial remission) [3, 4].

On the other hand, according to the Diagnostic and Statistical Manual of DSMV IV Mental Disorders, subjects diagnosed with ADHD can reach academic levels lower than those presented by their peers, and their functionality may also be compromised in the work, family, social environment, low self-esteem, disorder, poor planning capacity, lack of concentration, inadequate time management, among others. However, this research aims to determine the prevalence of symptoms of Attention Deficit Hyperactivity Disorder in Young Adults of the Universities of Barranquilla and its Metropolitan Area. The results of the study will also allow obtaining useful information at the level of higher education, since the symptoms may be having an impact on academics, leading to poor performance or even dropout associated with academic factors [5].

2   Methods

The present investigation is framed in the Empirical-Analytical paradigm because this study handles a theoretical approach of analytical cut. Regarding its approach, this research is considered quantitative; of descriptive scope. Finally, it is revealed that the temporality of the research responds to a Transactional study. The population under study corresponds to young university adults, men and women, residing in the city of Barranquilla and its metropolitan area, aged between 18 and 25 who voluntarily express their interest in participating in this research.

For the sample calculation, the reports of the ANDA National Data Archive, arranged by the National Administrative Department of Statistics DANE5, were taken into account, according to which the projection of the population to 2014 by sex and age group between 18 and 25 years in Barranquilla, Atlántico is the following (Table 1):

The final sample was found from a calculation of sample error for finite populations (taking into account a total population of 164,676). The result of the size is 1600 young adults between the ages of 18 and 25, which has been estimated with 95% confidence and a margin of error of 2.44%. The sampling implemented was non-probabilistic,

intentional and the selection technique is of the Expert type [6] based on the following inclusion and exclusion criteria:

Inclusion Criteria. Schooling: University, age: 18–25 years, place of residence: Barranquilla and Metropolitan area, absence of significant clinical history.

Exclusion Criteria

- Individuals who are not in school or have a lower level of education than the University.
- Individuals with ages that do not meet the age range between 18 and 25 years of age.
- Individuals whose fixed place of residence is not Barranquilla and/or municipalities of the metropolitan area.
- Individuals with significant clinical/neuropsychiatric history, for which the presence of the symptoms to be evaluated could be explained.

Table 1. Population projection by sex and age

Gender	2014	Frequency
Men	83.950	51%
Women	80.726	49%
Total	**164.676**	100%

It is necessary to rescue, that in the present study the instruments were applied to 1,674 subjects, of which 74 were excluded from the analyzes for not meeting the inclusion criteria, or for not having answered the questionnaire in its entirety.

3 Materials and Methods

3.1 Procedure

To carry out the present investigation, a procedure described in 6 phases was established:

- Phase 1: Theoretical review and state of the art in ADHD Adult.
- Phase 2: Selection and delimitation of the sample.
- Phase 3: Administration of the instruments.
- Phase 4: Analysis of results.
- Phase 5: Preparation of the report, conclusions and recommendation.

3.2 Dataset Description

For this study, the dataset selected was generate for the experimentation during the applications of neuropsychological tests to 1674 patients and 184 features, its description in the following.

- Sex: Nominal values Male or Female.
- Age: values of the age of a person in years.
- Utah Rating Scale (WURS): Wender Utah Rating Scale for the Attention Deficit and Hyperactivity Disorder, is a battery created to measure the severity of the symptoms presented by adults with ADHD. This test that applies in 10 to 15 min, measure symptoms in seven categories, see Table 2:

Table 2. WURS scale

Number	Description
1	Attention difficulties
2	Hyperactivity/Agitation
3	Humor
4	Affective lability
5	Emotional hyperreactivity
6	Disorganization
7	Impulsivity

Manages a scale of individual item types: 1 to 5, where 1 corresponds to (not at all), 2 (a little), 3 (moderately), 4 (quite a lot) and 5 (a lot). Its heading refers to the symptoms presented in childhood through sentences as: "As a child I was (or had) (or was) [7]. This test can be used to assess status mood and emotional changes that occur.

- ADHD: It is a self-applied checklist designed for ongoing research, the time of which application ranges from 10 to 15 min. Saying instrument is made up of 18 items, its objective is to measure the prevalence of ADHD symptoms in young adults through the quantification of related symptoms with this, taking into account the different factors that affect the configuration of these symptoms and that compromise the feasibility of the evaluation diagnostic and neuropsychological of this disorder [8]. The Exploratory Inventory of ADHD Symptoms (IES-ADHD), is made up of two subscales, which evaluate two criteria such as inattention and hyperactivity/impulsivity respectively. The first of them covers items 1 to 9 and is aimed at evaluating the symptoms of inattentive type in which questions arise on the presence of errors in activities labor, deficiency in the maintenance of attention in certain activities, the difficulty in following instructions etc. The second subscale includes items from 10 to 18, and evaluates the characteristic symptoms of the hyperactive type/impulsive, denoting questions like tinkling of the hands, feeling restless or constant anxiety, impatience, etc. The list of individual item types comprises 0 (never), 1 (rarely), 2 (sometimes), 3 (with frequency) and 4 (very frequently). The objective of this checklist is to establish with both subscales, the prevalence in the symptomatology of the ADHD in young adults whose childhood they presented the disorder and how these affect the functionality of the individual in the different aspects of your life.
- CIE-10 Test: This checklist is self-applicable and is developed to be answered in a period of 10 to 15 min. This instrument is consisting of 18 items, which are divided

according to the evaluation criteria ranging from inattention to hyperactivity to impulsiveness [9]. The first measures the difficulty itself of inattentive predominance that goes from items 1 to 9, the second of hyperactive predominance that goes from items 10 to 14 and the third of impulsive predominance which ranges from 15 to 18. It comprises a scale of individual element types from 0 to 4, where 0 corresponds to (never), 1 (rarely), 2 (sometimes), 3 (with frequency) and 4 (very frequently). The objective of this instrument is to assess membership and prevalence of ADHD symptoms in adults based on the evaluation system criteria OMS ICD-10 in 1992

- Class: In case 0 is a control patients, 1 is the patients has the ADHD disease.

3.3 Algorithms Description

3.3.1 Bagging

Bootstrap aggregation, also known as Bagging, is really a meta-algorithm designed to get model combinations from an initial family, causing a decrease in variance and avoiding overfitting. Although the most common is to apply it with the methods based on decision trees, it can be used with any family. Bagging has been shown to tend to produce improvements in cases of unstable individual models (such as neural networks or decision trees), but it can produce mediocre results or even worsen results with other methods, such as the K closest neighbors [10].

3.3.2 Algorithms Based on Boosting

Unlike bagging, boosting does not create versions of the training set, but always works with the complete input set, and manipulates the weights of the data to generate different models. The idea is that in each iteration the weight of the objects misclassified by the predictor is increased in that iteration, so in the construction of the next predictor these objects will be more important and will be more likely to classify them well [11].

3.3.3 DecisionStump

DecisionStump is an algorithm implemented in weka, where the tree decision will have three branches: one of them will be in case the attribute is unknown, and the other two will be in the case that the value of the attribute of the test example is equal to a specific value of the attribute or different from said value, in case of symbolic attributes, or that the value of the test example is greater or less than a certain value in the case of numerical attributes [12].

3.3.4 J48

Algorithm C4.5 builds decision trees of a training data system of the same form that the ID3 algorithm, which uses the concept of information entropy. Training data they are a system $S = s_1, s_2, \ldots$ of samples already classified. Each example $s_i = \{x_1, x_2, \ldots\}$ is a vector where $x_1, x_2, \ldots$ represent the attributes or characteristics of the example. Training data they are augmented with a vector $C = \{c_1, c_2, \ldots\}$ where c_1, c_2,… represent the class to which it belongs each sample. C4.5 is an extension of the ID3

algorithm previously developed by Quinlan. The trees decision generator by C4.5 can be used for classification, and for this reason, C4.5 is almost always referred to as a statistical classifier [13].

4   Experimentation

For this experimentation the confusion matrix was used. This is a fundamental tool when evaluating the performance of a classification algorithm, since it will give a better idea of how the algorithm is being classified, based on a count of the successes and errors of each of the classes in the classification. This way you can check if the algorithm is classifying the classes poorly and to what extent.

Table 3. Confusion matrix.

True Values		Classification	
		Negative	Positive
	Negative	a	b
	Positive	c	d

The meaning of the result of confusion matrix are explained next:

- *a,* is the number of correct predictions that a case is negative, called *True Negative*.
- *b,* is the number of incorrect predictions that a case is positive, that is, the prediction is positive when the value should really be negative, called *False Positive*.
- *c,* is the number of incorrect predictions that a case is negative, that is, the prediction is negative when the value should really be positive, called *False Negative*.
- *d,* is the number of correct predictions that a case is positive, called *True Positive*.

From the results obtained in the confusion matrix, other quality metrics can be analyzed, which are explained below:

4.1. *Accuracy (AC):* refers to the dispersion of the set of values obtained from repeated measurements of a magnitude [14]. The smaller the dispersion, the greater the accuracy. It is represented by the ratio between the number of correct predictions (both positive and negative) and the total predictions, and is calculated using the equation.

$$AC = \frac{a+d}{a+b+c+d} \tag{1}$$

4.2. *Precision (P):* It refers to how close the result of a true value measurement is. In statistical terms, accuracy is related to the bias of an estimate [15]. It is also known as True Positive (or "True positive rate"). It is represented by the proportion between the real positives predicted by the algorithm and all positive cases, and is calculated using the equation.

$$P = \frac{d}{b+d} \tag{2}$$

4.3. *Recall (TP):* Sensitivity, is also known as the True Positive Rate (TP). It is the proportion of positive cases that were correctly identified by the algorithm [16]. It is calculated according to the equation:

$$TP = \frac{d}{c+d} \tag{3}$$

4.4. *Specificity (TN):* Also known as the True Negative Rate (TN). These are the negative cases that the algorithm has correctly classified [17]. It is calculated according to the equation:

$$TP = \frac{a}{a+b} \tag{4}$$

4.5. *False Positive Rate (FP):* Is the proportion of negative cases that were mistakenly classified as positive by the algorithm [18]. It is calculated according to the equation:

$$FP = \frac{b}{a+b} \tag{5}$$

4.6. *False Negative Rate (FN):* Is the proportion of positive cases incorrectly classified [118]. It is calculated according to the equation:

$$FN = \frac{c}{c+d} \tag{6}$$

4.7. *F-measure:* With the concepts of precision and recall it is possible to define another type of metric called "F-Measure" [19]. This occurs depending on the two metric already seen and can be interpreted as the harmonic mean of both. In particular measure *F* handles a parameter "α" of as follows:

$$F_{\alpha} = \frac{1+a}{\frac{1}{Precision} + \frac{\alpha}{Recall}} \tag{7}$$

5   Results

After carrying out the process of cleaning and preprocessing the data, different algorithms of automatic learning support were used to identify patients with ADHD problems. These results are shown in Table 3 and Figs. 1, 2, 3 and 4.

Table 4. Algorithms results

Algorithm	Accuracy	Precision	Recall	F-measure
Bagging	91,67%	94,12%	88,89%	91,43%
MultiBoostAB	90,74%	94,00%	87,04%	90,38%
DecisionStump	89,81%	93,88%	85,19%	89,32%
LogitBoost	88,89%	90,38%	87,04%	88,68%
FT	87,96%	90,20%	85,19%	87,62%
J48Graft	87,04%	88,46%	85,19%	86,79%
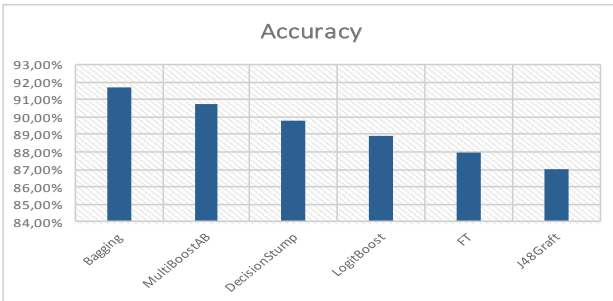
Fig. 1. Algorithms accuracy results. Source: Created by author.

In Fig. 1 the results of the algorithms accuracy can be found: Bagging 91,67%, MultiBoostAB 90,74%, DecisionStump 89,81%, LogitBoost 88,89%, FT 87,96% and J48Graft 87,04% (Table 4).
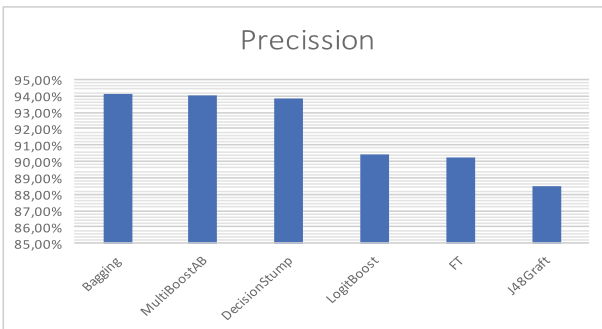

Fig. 2. Algorithms precission results. Source: Created by author.

In Fig. 2 the results of the algorithms precission can be found: Bagging 94,12%, MultiBoostAB 94,00%, DecisionStump 93,88%, LogitBoost 90,38%, FT 90,20% and J48Graft 88,46%.
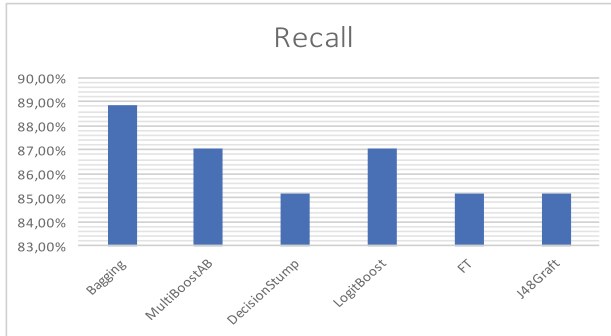
Fig. 3. Algorithms recall results. Source: Created by author.

In Fig. 3 the results of the algorithms recall can be found: Bagging 88,89%, MultiBoostAB 87,04%, DecisionStump 85,19%, LogitBoost 87,04%, FT 85,19% and J48Graft 85,19%.
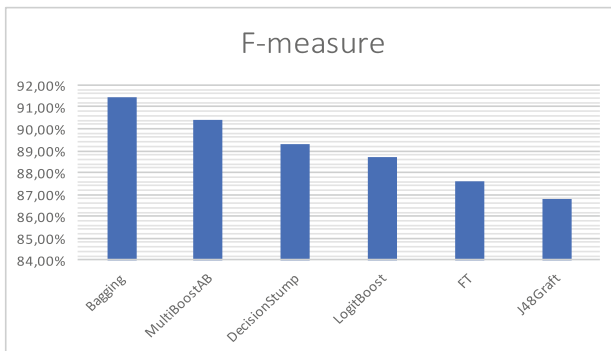

Fig. 4. Algorithms f-measure results. Source: Created by author.

In Fig. 4 the results of the algorithms F-measure can be found: Bagging 91,43%, MultiBoostAB 90,38%, DecisionStump 89,32%, LogitBoost 88,68%, FT 87,62% and J48Graft 86,79%.
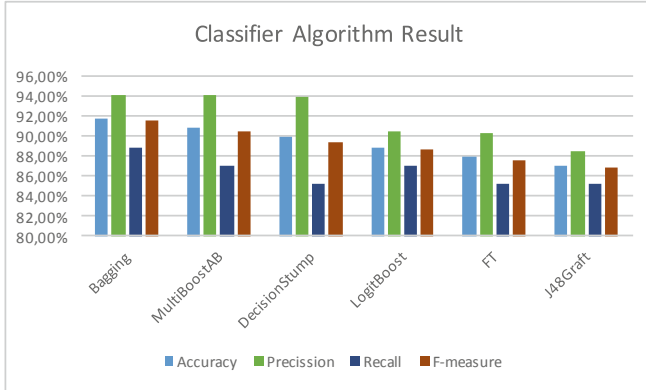
Fig. 5. Classifier algorithm result

6 Conclusions and Discussion

As a result of this experimentation, it can be concluded that machine learning techniques are an effective means or form for the analysis process of neurosychological variables related to attention deficit disorders that currently affect a large population conglomerate not only in Colombia, but around the world. From the results shown in Fig. 5, it can be defined that for this particular case the Bagging technique.

In the specific case of the Bagging classifier, it obtained the following results in the quality metrics. Accuracy 91,67%, Precision 94,12%, Recall 88,89% and F-measure 91,43%

References

1. Velázquez, J., García, M.: Trastorno por déficit de la atención e hiperactividad de la infancia a la vida adulta. Red de Revistas Científicas de América Latina, el Caribe, España y Portugal **9**(4), 176–181 (2007)
2. Ramos-Quiroga, J., Chalita, P., Vidal, R., Bosch, R., Palomar, G., et al.: Diagnóstico y tratamiento del trastorno por déficit de atención/hiperactividad en adultos. Rev. Neurol. **54**(1), 105–115 (2000)
3. Cabanyes, J., García, D.: Trastorno por déficit de atención e hiperactividad en el adulto: perspectivas actuales. Psiquiatría Biol. **13**(3), 86–94 (2006)
4. Faraone, S.V., Biederman, J., Spencer, T., Wilens, T., Seidman, L.J., et al.: Attention-deficit/hyperactivity disorder in adults: an overview. Biol. Psychiatry **48**(1), 9–20 (2000)
5. DANE: Archivo Nacional de Datos ANDA (2014). http://formularios.dane.gov.co/Anda_4_1/index.php/home. Citado 20 Marzo 2016
6. Pimienta-Lastra, R.: Encuestas probabilísticas vs. no probabilísticas. Polít. Cult. **13**, 263–276 (2000)
7. León-Jacobus, A., Valle-Cordoba, S., Florez-Niño, Y.: Diseño y validación piloto del inventario exploratorio de síntomas de TDAH (IES-TDAH) ajustado al DSM-V en jóvenes universitarios (Trabajo de Grado) (2007)

8. Adler, L., Kessler, R., Spencer, T.: Instrucciones para contestar la Escala de Auto-reporte de síntomas de TDAH en Adultos (ASRS-V1.1) (2003). http://www.neuropediatrica.com/descargas/tests/AUTOREPORTE%20TDA%20ADUL.pdf. Citado 15 Feb 2016

9. Barceló-Martínez, E., León-Jacobus, A., Cortes-Peña, O., Valle-Córdoba, S., Flórez-Niño, Y.: Validación del inventario exploratorio de síntomas de TDAH (IES-TDAH) ajustado al DSM-V. Rev. Mex. Neu. **17**(1), 1–113 (2016)

10. Breiman, L.: Bagging predictors. Mach. Learn. **24**(2), 123–140 (1996). https://doi.org/10.1007/BF00058655

11. Friedman, J.H.: Stochastic gradient boosting. Comput. Stat. Data Anal. **38**(4), 367–378 (2002)

12. Pang, J., Huang, Q., Jiang, S.: Multiple instance boost using graph embedding based decision stump for pedestrian detection. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008. LNCS, vol. 5305, pp. 541–552. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-88693-8_40

13. Bhargava, N., Sharma, G., Bhargava, R., Mathuria, M.: Decision tree analysis on J48 algorithm for data mining. Proc. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(6) (2013)

14. Ariza-Colpas, P., et al.: Enkephalon - technological platform to support the diagnosis of Alzheimer's disease through the analysis of resonance images using data mining techniques. In: Tan, Y., Shi, Y., Niu, B. (eds.) ICSI 2019. LNCS, vol. 11656, pp. 211–220. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26354-6_21

15. Davis, J., Goadrich, M.: The relationship between Precision-Recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning, pp. 233–240, June 2006

16. Powers, D.M.: Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation (2011)

17. Ye, K., Anton Feenstra, K., Heringa, J., IJzerman, A.P., Marchiori, E.: Multi-RELIEF: a method to recognize specificity determining residues from multiple sequence alignments using a Machine-Learning approach for feature weighting. Bioinformatics **24**(1), 18–25 (2008)

18. Yih, W.T., Goodman, J., Hulten, G.: Learning at low false positive rates. In: CEAS, July 2006

19. Lane, T., Brodley, C.E.: An application of machine learning to anomaly detection. In: Proceedings of the 20th National Information Systems Security Conference, Baltimore, USA, vol. 377, pp. 366–380, October 1997