Mackenzie

Combining Self-Organizing Maps and Decision Tree to Explain Diagnostic Decision Making in Attention-Deficit/Hyperactivity Disorder

Anderson Martins Silva[1*]
*e-mail: 71953371@mackenzista.com.br

[1]Graduate Program in Electrical and Computer Engineering
Mackenzie Presbyterian University
São Paulo, Brazil

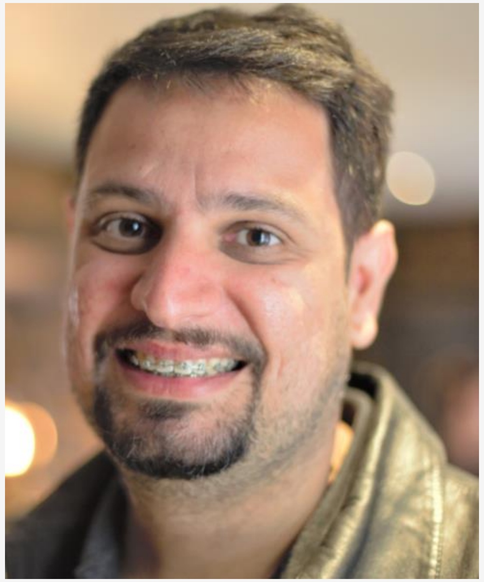
[2]Graduate Program in Developmental Disorders
Mackenzie Presbyterian University
São Paulo, Brazil

Luiz Renato Rodrigues Carreiro[2]
Mayara Miyahara Moraes Silva[2]
Maria Cristina Triguero Veloz Teixeira[2]
Leandro Augusto da Silva[1]

IARIA

BRAININFO 2021
Nice, France

Presenter's Resume



Anderson Martins Silva is a master's student in Electrical and Computer Engineering at Mackenzie Presbyterian University and graduated in 2006 in Statistics at the State University of Rio de Janeiro in both obtained scholarship funded by government program. During his graduation, he was co-author of the article published in the Brazilian Journal of Statistics: Multiple correspondence analysis and cluster analysis in the reduction of dimensionality of life events indicators. He has over 15 years of experience in finance, marketing, and the telecommunications market. He currently works at Oracle as the principal marketing analyst for Latin America.

Mackenzie

Outline

- Introduction
- Proposed Method
- Material and Methods
- Results
- Conclusion and Future Work

Outline

- Introduction
- Proposed Method
- Material and Methods
- Results
- Conclusion and Future Work

Mackenzie

Attention-Deficit/Hyperactivity Disorder (ADHD)

(ASSOCIATION et al., 2014; POLANCCZYK et al., 2007; SIMON et al., 2009)

- It is a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development. The disorder is characterized by inattention involving, for example, difficulty sustaining attention in tasks or play activities.
- The diagnosis of ADHD is clinical, based on the individual's history and expression of symptoms.
- ADHD affects 5.29% of the world's child population. Of this population, 30% up to 70% maintain symptoms into adulthood.
- Computational studies can help professionals in diagnostic assessments, especially using machine learning algorithms.

Mackenzie

Self-Organizing Maps (SOM)

(KOHONEN, 1982; HAYKIN, 2001)

- Created by Kohonen in 1982, the capture of the characteristics of an input data set, with a non-linear distribution, is accomplished by building a complex neural network around a grid of neurons.
- Inspired by neurobiology, the SOM algorithm presents three essential processes for self-organization of the grid: Competition, Cooperation and Adaptation.

Self-Organizing Maps (SOM)

(KOHONEN, 1982; HAYKIN, 2001)

1 **Competition:** The synaptic weight vector is calculated for each *j* neuron of the grid with the same dimension as the input dataset through the inner product between the synaptic weight vector and the input data vector, this function being the basis for choosing the winning neuron. The maximization of this function has mathematical equivalence with the minimization of the Euclidean distance between the synaptic weight and input data vectors.

$$X = [x_1, x_2, ..., x_m]^T$$

→

$$W_j = [w_{j1}, w_{j2}, ..., w_{jm}]^T, \ j = 1, 2, ..., l$$

→

$$i(X) = \arg\ \min_j \left\| X - W_j \right\|, \ j = 1, 2, ..., l$$

X is the input vector of the space m. W_j is the synaptic weight vector of each neuron in the grid. *i(x)* is the index that summarizes the competitive process between neurons.

Mackenzie

Self-Organizing Maps (SOM)

(KOHONEN, 1982; HAYKIN, 2001)

2 **Cooperation:** The basis for cooperation between neighboring neurons is provided by the winner neuron that shows the spatial location of the topological neighborhood of neurons neighboring the winner $h_{j,i(X)}$.

$$h_{j,i(X)} = \exp\left(-\frac{{d^2}_{j,i}}{2\sigma^2}\right)$$

3 **Adaptation:** Neighboring neurons to the winner increase their discriminant function values based on the input dataset and as appropriate adjustments applied to their synaptic weights improve a subsequent input dataset.

$$W_j(n+1) = W_j(n) + \eta(n) h_{j,i(X)}(n)(X - W_j(n))$$

where *n* equals epoch, η(n) is the learning rate, and $h_{i,j(x)}$(n) is the neighborhood function.

Decision Tree

(WITTEN, I. H. et al., 2016; MITCHEL, 1997)

- Decision Trees is an Artificial Intelligence algorithm capable of organizing attributes from a dataset in priority, so that it can generate a path that leads to a decision for a classificatory attribute.
- The information gain represents the expected reduction in entropy when the value of the attribute A is known, since the process calculates the gain for each attribute, choosing the attribute with the highest gain to be tested in the set S. This process creates the division of objects to form the decision tree, giving rise to the node, labeling the attribute, and creating branches for each attribute value.

$$Gain(S, A) = Entropy(S) + \Theta$$

$$\Theta = - \sum_{v \in Values(A)} p(A_v) Entropy(A_v)$$

$$Entropy(S) = \sum_{k=1}^{c} -p_i log_2 p_i$$

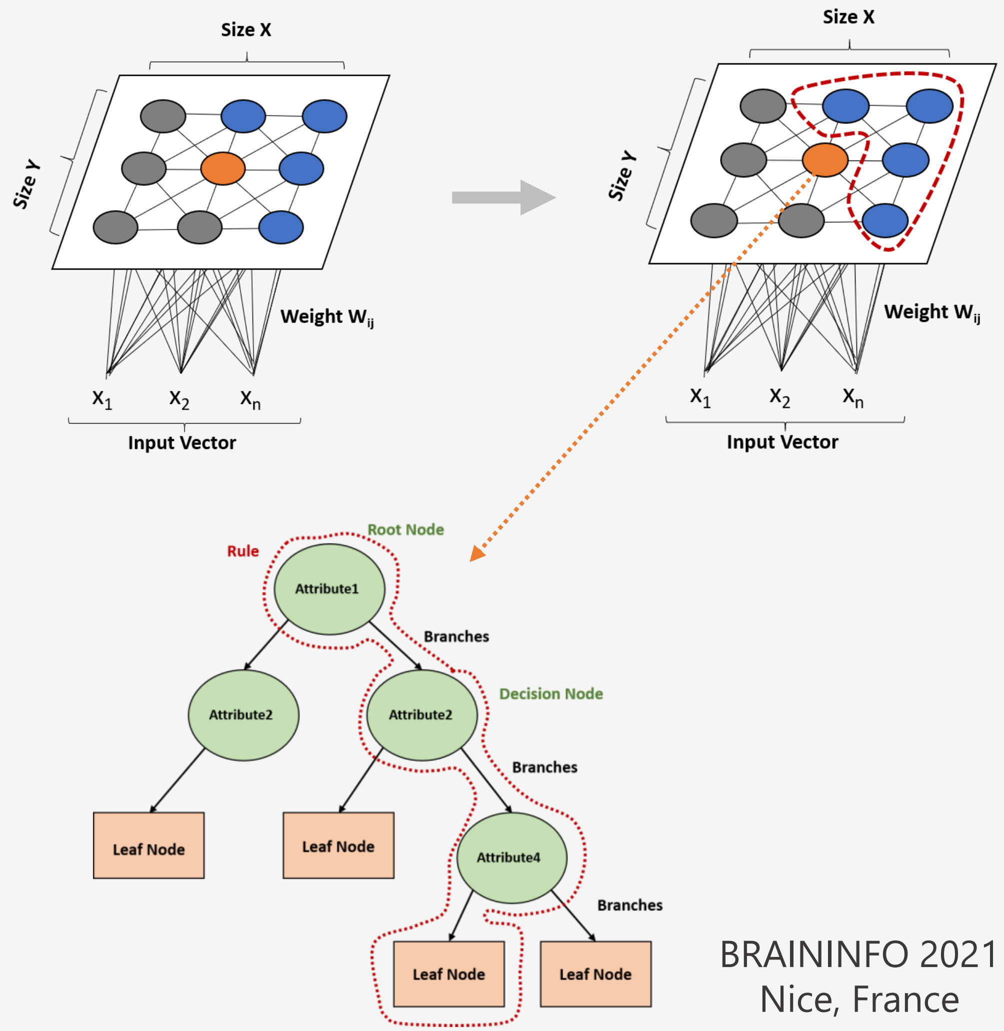
S is the input dataset and *A* are the attributes.

Outline

- Introduction
- Proposed Method
- Material and Methods
- Results
- Conclusion and Future Work

Proposed Method

- It presents a proposal for an unsupervised learning model as a method used in the identification of the neurons of the grid with greater overlapping of diagnoses of the disorder, that is, the class overlap in the neuron shows the case that it is difficult to make the diagnosis for both a machine learning algorithm and an expert.



Mackenzie

Outline

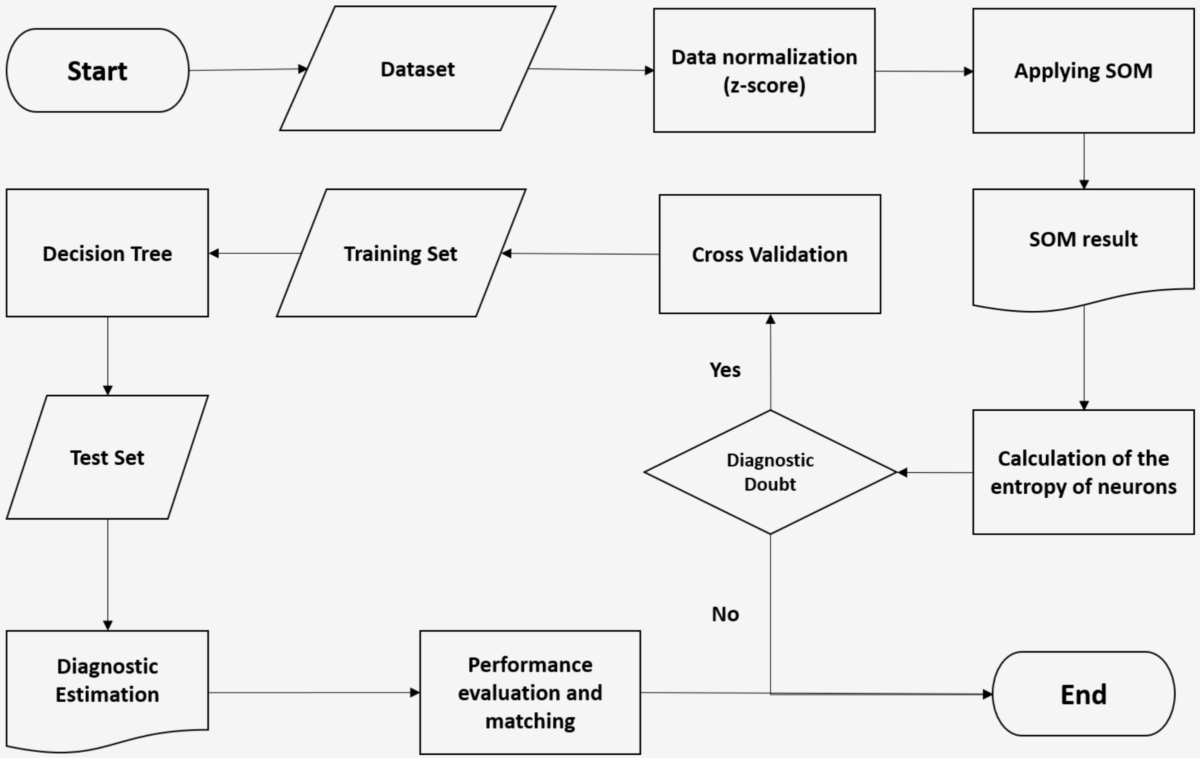
- Introduction
- Proposed Method
- Material and Methods
- Results
- Conclusion and Future Work

Mackenzie

Material and Methods

- The study sample consisted of 127 children and adolescents between 06 and 16 years old, 48 with a clinical diagnosis of ADHD and 79 from the control group, with no diagnosis of ADHD. The attributes that make up the neuropsychological tests and behavioral inventories applied in this study are Attention Cancellation Test (TAC), Trails (TT),Wechsler Intelligence Scale for Children (WISC-III), Wechsler Intelligence Scale for Children (WISC-IV), Wechsler Abbreviated Scale of Intelligence (WASI), Child Behavior Checklist for ages 6-18 (CBCL/6-18) and Teacher's Report Form for ages 6-18 (TRF/6-18).

Dataset Modeling Flowchart



Mackenzie

Outline

- Introduction
- Proposed Method
- Material and Methods
- Results
- Conclusion and Future Work

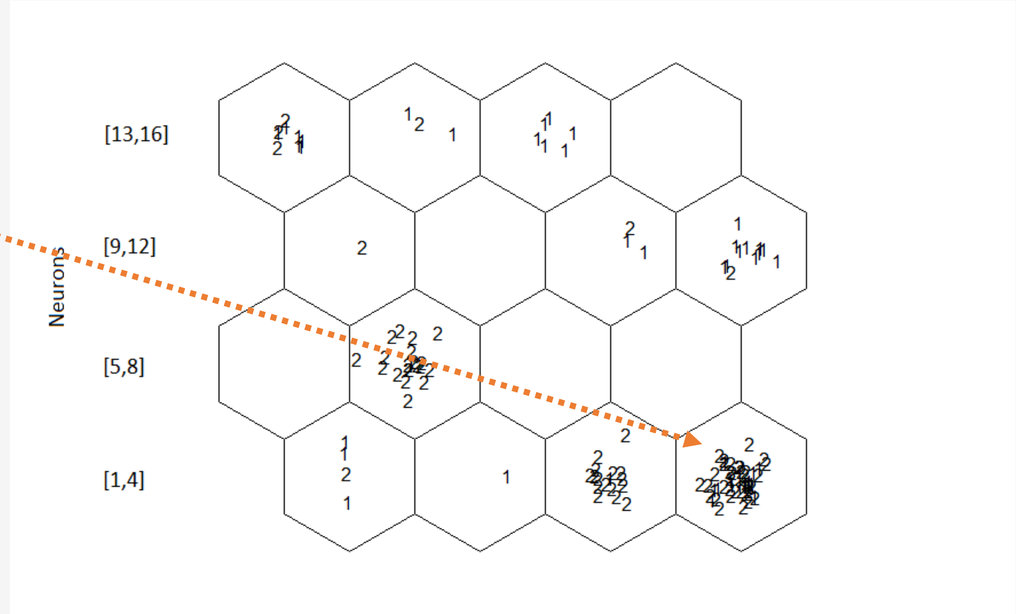
Mackenzie

Results

Comparative Diagnosis by the neuron dimension 4x4

	Diagnostic	1	2	Total
neuron	1	3 (6.2%)	1 (1.3%)	4 (3.1%)
	2	1 (2.1%)	0 (0.0%)	1 (0.8%)
	3	1 (2.1%)	16 (20.3%)	17 (13.4%)
	4	16 (33.3%)	35 (44.3%)	51 (40.2%)
	6	0 (0.0%)	20 (25.3%)	20 (15.7%)
	9	0 (0.0%)	1 (1.3%)	1 (0.8%)
	11	2 (4.2%)	1 (1.3%)	3 (2.4%)
	12	11 (22.9%)	1 (1.3%)	12 (9.4%)
	13	6 (12.5%)	3 (3.8%)	9 (7.1%)
	14	2 (4.2%)	1 (1.3%)	3 (2.4%)
	15	6 (12.5%)	0 (0.0%)	6 (4.7%)
Total		48 (100.0%)	79 (100.0%)	127 (100.0%)

Scattering of objects diagnostic within neurons



1 - group diagnosed with ADHD

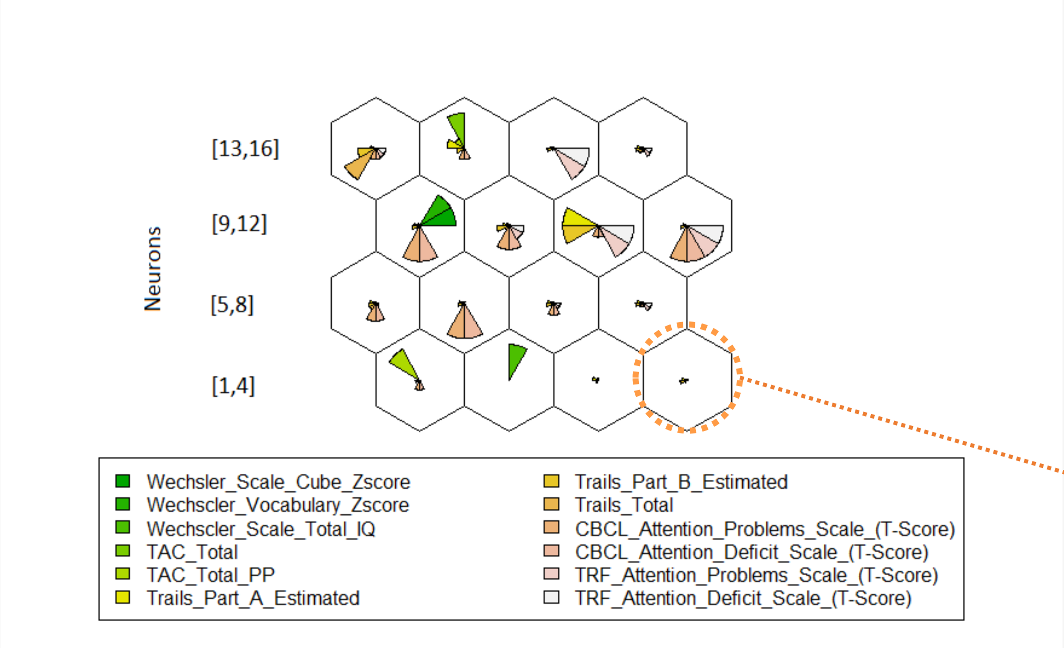
2 – control group diagnosed without ADHD

Neuron 4 – diagnostic doubt

Mackenzie

Results

Contribution of the attributes in the formation of the neuron

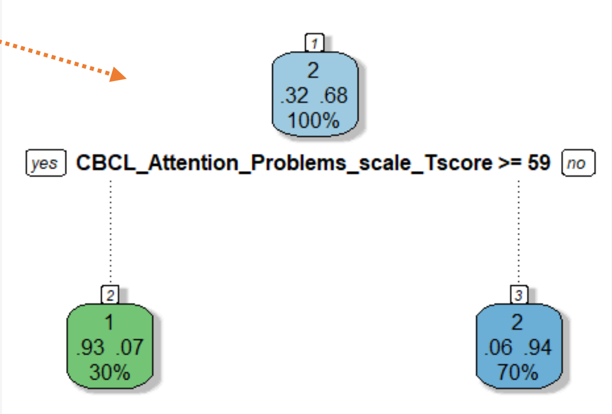


1 - group diagnosed with ADHD

2 – control group diagnosed without ADHD

Neuron 4 – diagnostic doubt

Importance of the attributes in neuron 4 by decision tree



Mackenzie

Outline

- Introduction
- Proposed Method
- Material and Methods
- Results
- Conclusion and Future Work

Mackenzie

Conclusion and Future Work

- The **self-organizing map** contributed especially to the **formation of groups** and the **understanding of clusters** with class overlapping, which is the proposal of this work.
- The application of the **decision tree discriminated six attributes**, being two of cognitive assessment and four of behavioral assessment, both showed relevant discrimination to make the diagnosis, being **the Child Behavior Checklist for ages 6-18 attribute,** the one that showed **the highest discriminative power**. The study presented as a relevant factor the case of overlapping diagnoses of neurons when using the SOM and, in conjunction with the **decision tree**, was able to **separate 88% of the cases**.
- **Future studies** can test the same **decision tree on larger samples** to see if the attributes that showed higher accuracy are maintained. Since this study controlled for no comorbidities in the ADHD group, it is recommended for **future studies to use sample groups with and without ADHD comorbidities** from other psychiatric and neurodevelopmental conditions.

Mackenzie

Thank you for your attention

Acknowledgments

Thank the Academic Excellence Program of the Coordination for the Improvement of Higher Education Personnel (CAPES-PROEX), Process number 1133/2019, (CAPESPROSUC) in mode II, the National Council for Scientific and Technological Development (CNPq, Cases 307730/2017-4 and 307443/2019-1), the Mackenzie Research Fund (MACKPESQUISA) of Mackenzie Presbyterian University, and the Foundation for Research Support of the State of São Paulo (FAPESP, Cases 2018/01063-0 and 2019/20757-1) for financial support.

IARIA