

**UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO**

Anderson Martins Silva

**COMBINAÇÃO DE MAPAS AUTO-ORGANIZÁVEIS E
ÁRVORE DE DECISÃO PARA EXPLICAÇÃO DA TOMADA
DE DECISÃO DIAGNÓSTICA NO TRANSTORNO DE
DÉFICIT DE ATENÇÃO/HIPERATIVIDADE**

Orientador: Prof. Dr. Leandro Augusto da Silva

São Paulo
2021

**UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO**

Anderson Martins Silva

**COMBINAÇÃO DE MAPAS AUTO-ORGANIZÁVEIS E
ÁRVORE DE DECISÃO PARA EXPLICAÇÃO DA TOMADA
DE DECISÃO DIAGNÓSTICA NO TRANSTORNO DE
DÉFICIT DE ATENÇÃO/HIPERATIVIDADE**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação da Universidade Presbiteriana Mackenzie como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica e Computação.

Orientador: Prof. Dr. Leandro Augusto da Silva

São Paulo
2021

S586c Silva, Anderson Martins

Combinação de mapas auto-organizáveis e árvore de decisão para explicação da tomada de decisão diagnóstica no transtorno de déficit de atenção / hiperatividade / Anderson Martins Silva

64 f.: il.; 30 cm

Bibliografia: f. 60-64

Dissertação (Mestrado em Engenharia Elétrica e Computação), São Paulo, 2021.

Orientador: Prof. Dr. Leandro Augusto da Silva

1. Mapas Auto-organizáveis (SOM) 2. Árvore de Decisão 3. Transtorno de Déficit de Atenção e Hiperatividade (TDAH) I. Silva, Leandro Augusto da; Orientador. II. Título.

CDD 616.8589

Bibliotecária Responsável: Maria Gabriela Brandi Teixeira – CRB 8/6339


ANDERSON MARTINS SILVA

COMBINAÇÃO DE MAPAS AUTO-ORGANIZÁVEIS E ÁRVORE DE DECISÃO
PARA EXPLICAÇÃO DA TOMADA DE DECISÃO DIAGNÓSTICA NO
TRANSTORNO DE DÉFICIT DE ATENÇÃO/HIPERATIVIDADE

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação da Universidade Presbiteriana Mackenzie como parte dos requisitos para a obtenção do título de mestre.

Aprovada em 06/08/2021

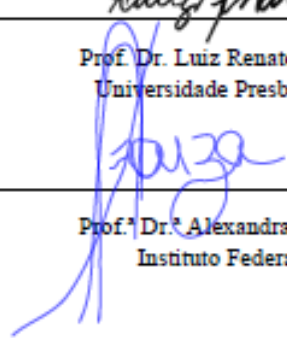
BANCA EXAMINADORA



Prof. Dr. Leandro Augusto da Silva
Universidade Presbiteriana Mackenzie



Prof. Dr. Luiz Renato Rodrigues Carreiro
Universidade Presbiteriana Mackenzie



Prof.ª Dr.ª Alexandra Aparecida de Souza
Instituto Federal de São Paulo

Folha de Identificação da Agência de Financiamento

Autor: Anderson Martins Silva

Programa de Pós-Graduação *Stricto Sensu* em Engenharia Elétrica e Computação

Título do Trabalho: COMBINAÇÃO DE MAPAS AUTO-ORGANIZÁVEIS E ÁRVORE DE DECISÃO PARA EXPLICAÇÃO DA TOMADA DE DECISÃO DIAGNÓSTICA NO TRANSTORNO DE DÉFICIT DE ATENÇÃO/HIPERATIVIDADE

O presente trabalho foi realizado com o apoio de ¹:

- CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
- CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico
- FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo
- Instituto Presbiteriano Mackenzie/Isenção integral de Mensalidades e Taxas
- MACKPESQUISA - Fundo Mackenzie de Pesquisa
- Empresa/Indústria:
- Outro:

¹ **Observação:** caso tenha usufruído mais de um apoio ou benefício, selecione-os.

Apoio:

PROSUC/CAPES Modalidade II

AGRADECIMENTOS

Ao refletir sobre minha vida e passar por uma pandemia de corona vírus, só tenho a agradecer a todos que contribuíram na longa jornada para chegar até aqui. Não foram somente dois anos de estudo, mas um longo período de amadurecimento profissional e acadêmico que me fizeram ter forças para vencer as barreiras criadas por mim e finalizar essa etapa da jornada.

A primeira pessoa que agradeço é o meu orientador Leandro Augusto da Silva, que desde o início do curso conduziu-me de forma a buscar a excelência acadêmica em todos os trabalhos desenvolvidos no curso. Com sua orientação pude aplicar para bolsa de estudo da CAPES e também ingressar no grupo de pesquisa em computação e distúrbios, o qual foi fundamental para elaboração do artigo entregue na Sexta Conferência Internacional sobre Neurociência e Informação Cognitiva do Cérebro (BRAININFO2021) com data de realização entre os dia 18 e 22 de julho de 2021 em Nice, França. Serei eternamente grato ao aprendizado acumulado durante esses dois anos e espero que seja o início da minha caminhada acadêmica.

Ao Prof. Dr. Luiz Renato Rodrigues Carreiro e a Prof^a. Dr^a. Alexandra Aparecida de Souza por participarem da banca de qualificação dando orientações valiosas na composição desse trabalho.

A minha mãe Rosely Ferreira Martins que em todos os momentos disse para não baixar a cabeça e sempre acreditar nos meu sonhos, pois o estudo é a única coisa que ninguém poderia roubar de você. Com base na sua educação consegui ser quem sou hoje e fico orgulhoso de saber que sua missão foi cumprida com sucesso.

Ao meu irmão André Martins Silva que em todos momentos esteve ao meu lado e com palavras de sabedoria fazia-me refletir e buscar a excelência.

A minha esposa Valkiria Scherer de Medeiros por dedicar nesses 15 anos de companheirismo, amor e o ensinamento que a vida pode ser leve e cheia de felicidade. O crescimento vem com o acumulo do aprendizado diário e a Fé em Deus é um pilar importante de sustentação para os momentos difíceis.

A minha amiga e coach Mariana Alves por toda dedicação e orientação durante a jornada de trabalho no Rio de Janeiro e por ter me trazido para São Paulo para fazer parte do seu time e acreditado no meu profissionalismo.

Ao grupo de estudo composto pela Letícia Sakamoto e André Brasileiro por serem fundamentais no aprendizado das matérias e combustível diário para continuar a caminhada, inclusive nesse período de pandemia de covid-19.

Ao amigo Persio de Luca por gastar seu precioso tempo para revisar o artigo do Português para o Inglês e pela troca de experiência enriquecedora que nos faz crescer a cada dia.

Ao gestor Erik Gajardo e a ORACLE por acreditarem que a educação é uma forma brilhante de transformação e empoderamento das pessoas.

RESUMO

O Transtorno de Déficit de Atenção/Hiperatividade (TDAH) apresenta-se em crianças e adolescentes como um padrão persistente de desatenção, hiperatividade e impulsividade que interfere no seu desenvolvimento. Estudos computacionais realizados sobre TDAH focam nas medidas de atividades cerebrais dos participantes e poucos utilizam testes cognitivos padronizados ou inventários comportamentais para avaliação de indicadores objetivos para elaboração do diagnóstico. O trabalho apresenta como proposta computacional a combinação de dois métodos de inteligência artificial para auxiliar a identificação de indicadores diagnósticos para diagnóstico do TDAH. A proposta é combinar uma rede neural de mapas auto-organizáveis para agrupar fatores advindos de testes e inventários padronizados e árvore de decisão para classificar os fatores mais relevantes. O estudo contemplado teve 127 crianças e adolescentes de 6 a 16 anos, sendo 48 com diagnóstico confirmado e 79 sem indicadores diagnósticos para TDAH. O trabalho apresenta como resultado relevante, a forte contribuição de escores do inventário de comportamentos para crianças e adolescentes no diagnóstico do transtorno.

Palavras-chave: *Mapas Auto-Organizáveis (SOM), Árvore de Decisão, Transtorno de Déficit de Atenção e Hiperatividade (TDAH).*

ABSTRACT

Attention-Deficit/Hyperactivity Disorder (ADHD) presents in children and adolescents as a persistent pattern of inattention, hyperactivity, and impulsivity that interferes with their development. Computational studies on ADHD focus on measures of brain activity of the participants and few uses standardized cognitive tests or behavioral inventories to assess objective indicators for diagnosis. The paper presents as a computational proposal the combination of two artificial intelligence methods to aid the identification of diagnostic indicators for ADHD. The proposal is to combine a neural network of self-organizing maps to group factors from standardized tests and inventories, and a decision tree to classify the most relevant factors. The study included 127 children and adolescents from 6 to 16 years old, 48 with ADHD diagnosis and 79 without ADHD (control group). The most relevant result of the study was the strong contribution of the scores of the Inventory of Behaviors for Children and Adolescents in the diagnosis of the disorder.

Keywords: *Self-Organizing Maps (SOM), Decision Tree, Attention-Deficit Hyperactivity Disorder (ADHD).*

Lista de Figuras

1	Tradução da representação de um neurônio	11
2	Tradução da representação dos três estágios do sistema nervoso	12
3	Tradução da arquitetura básica da rede neural de uma camada	13
4	Mapa Topológico - Grade Retangular e Hexagonal	15
5	Visão da Função de vizinhança Gaussiana	17
6	Visualização conceitual do mapa auto-organizável	28
7	Segmentação dos neurônios no mapa auto-organizável	29
8	Estrutura da árvore de decisão	30
9	Fluxograma da modelagem do conjunto de dados	35
10	Visualização das 1.000 iterações aplicadas no SOM - grade 2x2	38
11	Contribuição dos atributos na formação do neurônios	39
12	Dispersão dos objetos dentro dos Neurônios	40
13	Importância dos atributos no neurônio 4 pela árvore de decisão	42
14	Árvore de decisão do neurônio 4	42
15	Visualização das 1.000 iterações aplicadas no SOM	44
16	Contribuição dos atributos na formação do neurônio	45
17	Dispersão dos objetos dentro dos neurônios	46
18	Importância dos atributos no neurônio 4 pela árvore de decisão	48
19	Árvore de decisão do neurônio 4	48
20	Visualização das 1.000 iterações aplicadas no SOM	50
21	Contribuição dos atributos na formação do neurônio	52
22	Dispersão dos objetos dentro dos neurônios	52
23	Importância dos atributos no neurônio 4 pela árvore de decisão	56
24	Árvore de decisão do neurônio 4	56

Lista de Quadros

1	Matriz de Confusão	34
2	Comparativo diagnóstico pelo neurônio dimensão 2x2	38
3	Inércia dos atributos dimensão 2x2 - Parte 1	40
4	Inércia dos atributos dimensão 2x2 - Parte 2	41
5	Inércia dos atributos dimensão 2x2 - Parte 3	41
6	Matriz de Confusão	43
7	Comparativo diagnóstico pelo neurônio dimensão 3x3	44
8	Inércia dos atributos dimensão 3x3 - Parte 1	46
9	Inércia dos atributos dimensão 3x3 - Parte 2	47
10	Inércia dos atributos dimensão 3x3 - Parte 3	47
11	Matriz de Confusão	49
12	Comparativo diagnóstico pelo neurônio dimensão 4x4	51
13	Inércia dos atributos dimensão 4x4 - Parte 1	53
14	Inércia dos atributos dimensão 4x4 - Parte 2	54
15	Inércia dos atributos dimensão 4x4 - Parte 3	55
16	Matriz de Confusão	57

Sumário

1	INTRODUÇÃO	7
1.1	Objetivo	8
1.2	Organização do Trabalho	8
2	REFERENCIAL TEÓRICO	10
2.1	Princípios da Rede Neural Artificial	10
2.1.1	Modelo primário de neurônio artificial	12
2.2	Mapas Auto-Organizáveis (Self-Organizing Maps - SOM)	14
2.2.1	Processo Competitivo	15
2.2.2	Processo Cooperativo	16
2.2.3	Processo Adaptativo	18
2.3	Modelo de Árvore de Decisão	20
2.4	Elementos do Transtorno de Déficit de Atenção e Hiperatividade (TDAH) .	22
3	MÉTODO HÍBRIDO BASEADO EM INTELIGÊNCIA ARTIFICIAL	28
4	METODOLOGIA	31
4.1	Descrição e Experimentos	31
4.2	Avaliação de Desempenho de Classificação	34
5	RESULTADOS	37
5.1	Visualização dos resultados do SOM combinados com a árvore de decisão .	37
5.1.1	Resultados com grade 2x2	37
5.1.2	Resultados com grade 3x3	43
5.1.3	Resultados com grade 4x4	49
6	CONCLUSÃO E TRABALHOS FUTUROS	58
	REFERÊNCIAS BIBLIOGRÁFICAS	64

1 INTRODUÇÃO

O Transtorno de Déficit de Atenção/Hiperatividade (TDAH) é definido de acordo com o Manual Diagnóstico e estatístico das doenças mentais, na sua 5ª edição – DSM-5 (ASSOCIATION et al., 2014), como sendo um transtorno do neurodesenvolvimento definido por níveis prejudiciais de desatenção, desorganização e/ou hiperatividade-impulsividade. Envolvem, por exemplo, a incapacidade de permanecer em uma tarefa, ter comportamento parecido como não escutar quando chamado e perder coisas, expressarem atividade excessiva ou inquietação e incapacidade de aguardar sua vez, sempre de modo excessivos para a idade ou o nível de desenvolvimento. O TDAH tem suas expressões iniciais na infância e costuma persistir na vida adulta, resultando em prejuízos no funcionamento social, acadêmico e profissional. Certos prejuízos são decorrentes de comportamento com reações que prejudicam o ambiente social e se expressa como agressividade, depressão, irritação, nervosismo, impulsividade e impaciência (ACHENBACH, 2019).

O diagnóstico do TDAH é clínico, baseado na história do indivíduo e da expressão dos seus sintomas. Como muitas vezes esse diagnóstico é baseado no relato da severidade dos sintomas e pelo fato desses sintomas também fazerem parte de outras condições clínicas, a dificuldade diagnóstica está presente no dia a dia das equipes interdisciplinares responsáveis pelos processos de avaliação (MAHONE; DENCKLA, 2017; ARETOULI, 2019). Tamanha complexidade demanda essa avaliação diagnóstica que a Associação Americana de Pediatria recomenda o uso de um algoritmo, tanto para avaliação como para o tratamento de crianças e adolescentes com TDAH (WOLRAICH et al., 2019).

Para subsidiar a tomada de decisão clínica, procedimentos de avaliação neuropsicológica, comportamental e de funcionamento adaptativo têm sido frequentemente utilizadas em conjunto a avaliações neurológicas (CARREIRO et al., 2014). Considerando a importância social que implica a emissão adequada de um diagnóstico correto de TDAH tanto em crianças como adolescentes é necessário que sejam propostos estudos que discutam quais os melhores indicadores de avaliação diagnóstica clínica-neurológica, neuropsicológica e comportamental-adaptativa quando crianças e adolescentes apresentam queixas de desatenção e hiperatividade. Além disso, para que avaliações adequadas e intervenções sejam implementadas são necessários critérios diferenciais para caracterizar e identificar corretamente o déficit de atenção/hiperatividade entre crianças e adolescentes. Avaliações

exaustivas nesse sentido permitem um melhor entendimento da complexidade de cada caso para uma adequada orientação, elaboração da intervenção terapêutica e avaliação da necessidade do suporte educacional e emocional para pacientes e familiares (CARREIRO et al., 2014).

Estudos computacionais podem ajudar os profissionais nas avaliações diagnósticas, em especial com uso de algoritmos de aprendizagem de máquina (CHANDANA; VIJAYALAKSHMI, 2018). Kam et al. (KAM et al., 2010) usaram um algoritmo de Inteligência Artificial chamado de árvore de decisão para rastreamento do TDAH através do controle das atividades escolares de 153 crianças com uso de actiógrafo 3-axial e obteve resultados consistentes com estudos anteriores. Por sua vez Lee, Abibullaev & Kang (LEE et al., 2010) analisaram a classificação do TDAH em crianças através medidas de atividades cerebrais. No presente trabalho usaram um algoritmo de rede neural chamado de mapas auto-organizáveis permitindo categorizar características de crianças com e sem indicadores clínicos de TDAH.

1.1 Objetivo

Diferente das propostas anteriores apresentadas na literatura, este trabalho tem como objetivo combinar duas técnicas de inteligência artificial. Em um primeiro momento se faz o agrupamento de resultados de testes padronizados por meio dos atributos e, em um segundo momento, os grupos com alto nível de sobreposição são analisados a partir do uso de algoritmo de árvores de decisão, que ajuda na descoberta de qual atributo é discriminativo no diagnóstico de criança e adolescentes com suspeita de TDAH.

1.2 Organização do Trabalho

Além da introdução ter como finalidade contextualizar o trabalho e apresentar o objetivo, o conteúdo está organizado em mais cinco partes, sendo o referencial teórico, o capítulo que mostra as referências bases para construção do trabalho e o método proposto traz a ideia geral de aplicação do método. Na parte de metodologia descrevem-se os procedimentos para desenvolvimento do estudo com início na obtenção da base de dados e passando por todo o desenvolvimento computacional aplicado pelas duas técnicas

de inteligência artificial. Nos resultados, será descrita a contribuição dos testes cognitivos padronizados ou inventários comportamentais dentro do neurônio de sobreposição do diagnóstico e depois classificá-lo pela árvore de decisão para entendimento das características do diagnóstico do transtorno. Por fim, em discussão e conclusões apresenta-se o resumo do trabalho com destaque para os principais pontos de entendimento abordados no estudo e uma proposta concreta para trabalho futuro.

2 REFERENCIAL TEÓRICO

Conforme descrito por Zumel, Mount e Porzak (2014), existem uma infinidade de ferramentas para exploração e interpretação dos dados utilizando métodos de aprendizado de máquina que são classificados por dois tipos: supervisionado e não supervisionado. A proposta do estudo é usar o método supervisionado que prediz ou estima, com base nos diagnósticos apresentados nos testes aplicados em crianças e adolescentes e com base no método de árvore de decisão, um resultado de classificação da prevalência do TDAH.

Este capítulo apresenta os conceitos elementares sobre rede neural artificial, mapas auto-organizáveis, árvore de decisão e elementos do transtorno de déficit de atenção e hiperatividade (TDAH) que foram envolvidos na realização deste trabalho.

2.1 Princípios da Rede Neural Artificial

A motivação que leva a comunidade científica a efetuar pesquisas em redes neurais é reconhecida pelo fato que o cérebro processa informações mais rápido do que qualquer computador e organiza sua estrutura em neurônios que se comunicam pelas sinapses tornando o sistema complexo, não-linear com processamento paralelo (HAYKIN, 2001).

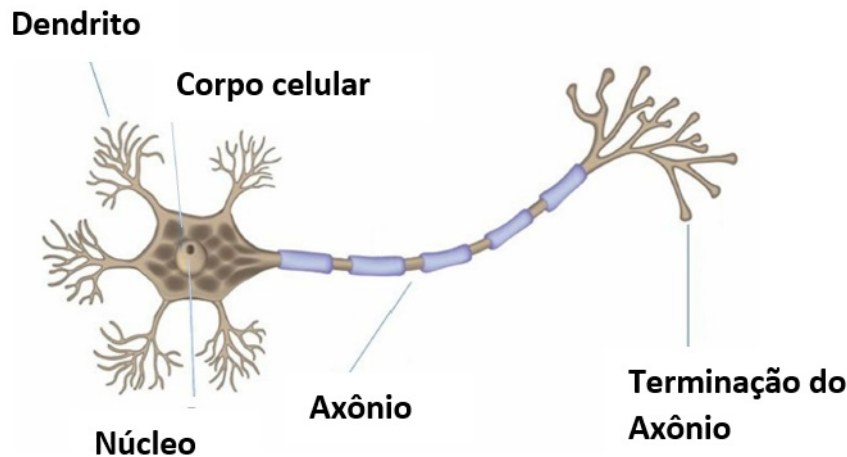
O cérebro pesa aproximadamente 1,5kg com cerca de 86 bilhões de neurônios e é uma unidade de processamento central, sendo o órgão mais importante do corpo humano com uma rede complexa de neurônios interconectados para processar informações (CIA-BURRO; VENKATESWARAN, 2017).

De acordo com Haykin (2001), a rede neural é constituída de processamentos direto e paralelamente distribuído que têm a propensão para guardar dados experimentais e seu uso assemelha-se com o cérebro em dois momentos: o processo de aprendizagem faz o conhecimento ser absorvido pela rede a partir do seu ambiente; os conhecimentos adquiridos são armazenados pelas forças de ligação entre os neurônios que são conhecidas como peso sinápticos.

A Figura 1 mostra os principais componentes de um neurônio sendo o dendrito a parte de entrada que recebe a informação na forma de impulsos elétricos dos outros

neurônios da rede; o corpo celular é o local que gera a inferência e toma as decisões, dado a entrada que vêm dos dendritos; e o terminal axiônio é a parte de saída do impulso elétrico (CIABURRO; VENKATESWARAN, 2017).

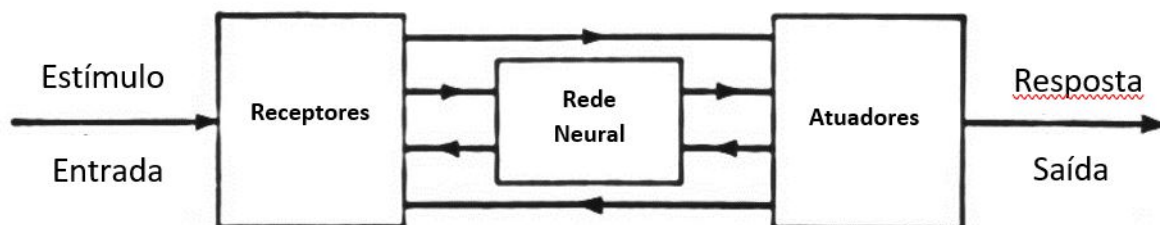
Figura 1: Tradução da representação de um neurônio



Fonte: (CIABURRO; VENKATESWARAN, 2017, p.32)

Já na Figura 2 (ARBIB, 1987) temos um diagrama de blocos para ilustrar o sistema nervoso humano. A rede neural que representa o cérebro recebe as informações de modo recorrente, faz todo o processamento e toma as decisões. As setas apresentadas na Figura 2 tem um significado de transmissão que alimentam (sentido esquerda para direita) e realimentam (sentido da direita para esquerda) o sistema. Os receptores captam do corpo humano e do ambiente externo os estímulos que são convertidos em impulsos elétricos para o cérebro. Já os atuadores pegam os impulsos elétricos enviados do cérebro transformam em resposta que é a saída do sistema (HAYKIN, 2001).

Figura 2: Tradução da representação dos três estágios do sistema nervoso



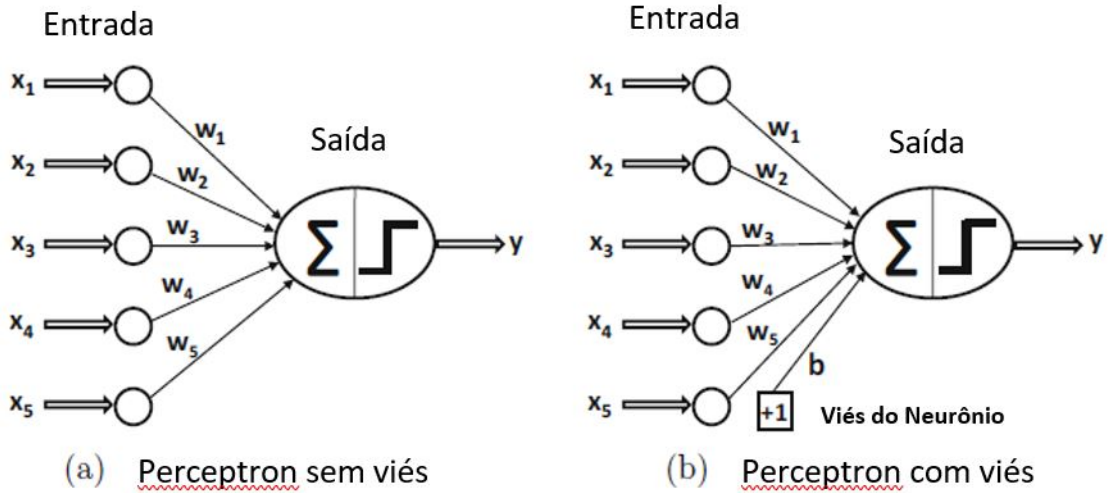
Fonte: (ARBIB, 1987, p.16)

É importante entender a complexidade do cérebro humano e a ciência está longe de recriá-lo em um ambiente digital. Porém, os neurônios artificiais estão no estágio primário de desenvolvimento em relação ao encontrado no cérebro, mas a maneira que o processo evoluiu nas últimas décadas, o que é surpreendente, e tendo a neurobiologia como fonte de inspiração, o aprendizado e a experiência acumulada entre os anos estudados, certamente a cada ano que passa, o entendimento das redes neurais artificiais será mais representativo e esclarecedor (HAYKIN, 2001).

2.1.1 Modelo primário de neurônio artificial

De forma análoga a estrutura biológica do neurônio do cérebro humano, o neurônio artificial é definido como uma unidade de processamento central da rede neural artificial que recebe um conjunto de entrada de dados, faz a operação matemática e gera a saída desse neurônio que é a soma ponderada das entradas mais o bias (CIABURRO; VENKATESWARAN, 2017). Segundo Aggarwal et al. (2018), o conjunto de entrada das redes neurais de camada única está ligado diretamente à saída que utiliza uma generalização da variação de uma função linear. Este processo da rede neural é também conhecido como perceptron. Já em redes neurais multicamadas, os neurônios são organizados em camadas, sendo o conjunto de entrada e saída separadas por um grupo de camadas ocultas. Essa arquitetura da rede é conhecida como “*Feed-Forward*” e está representada na Figura 3, em que (a) representa o perceptron sem viés (*bias*) e o (b) o perceptron com viés (*bias*).

Figura 3: Tradução da arquitetura básica da rede neural de uma camada



Fonte: (AGGARWAL et al., 2018, p.5)

Conforme Haykin (2001) verificam-se três elementos básicos na rede neural que são: a) as sinapses ou elos de conexão que são caracterizados por um peso relativo. Cada entrada da sinapse i recebe um sinal x_i que está conectada a um neurônio k que por sua vez é multiplicado pelo peso sináptico w_{ki} . b) Os pesos sinápticos podem ter valores do seu intervalo tanto negativos quanto positivos; O somatório que é a combinação linear das entradas com suas sinapses (Equação 1);

$$u_k = \sum_{i=1}^N w_{ki}x_i \quad (1)$$

c) a função de ativação ou função restritiva, já que limita o intervalo permitido da amplitude do sinal de saída a um número finito é, geralmente, representada por um intervalo normalizado fechado, devido ao tipo de função identificada como Limiar e a função linear por partes $[0, 1]$ ou $[-1, 1]$ devido a função sigmoide.

Em resumo, a função de saída de cada neurônio é igual ao somatório dos pesos multiplicado pelas entradas e adicionado o bias (Equação 2).

$$y_k = \rho(u_k + b_k) \quad (2)$$

O conjunto organizado de neurônios em um arranjo bidimensional e de maneira não-supervisionada possibilita a disposição dos neurônios em forma de grade e gera visualização de segmentos na rede (SEIFFERT, 2013).

2.2 Mapas Auto-Organizáveis (Self-Organizing Maps - SOM)

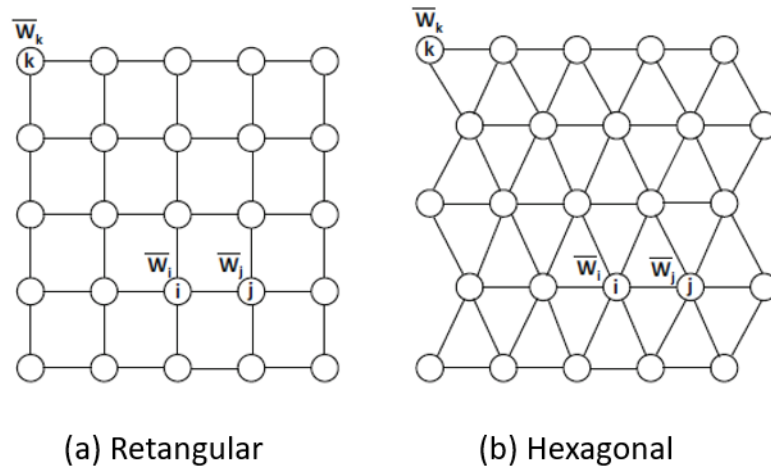
De acordo com Merényi, Mendenhall e O’Driscoll (2016), o modelo SOM é extremamente utilizado por ter uma implementação amigável, com boas propriedades para mineração de dados, além de proporcionar o agrupamento e representação visual dos dados. Porém é um algoritmo complexo que na maioria das vezes é tido como modelo de rede neural com dificuldade de entendimento de suas etapas de forma manual e um estudo recente de Rai (2020) mostra que a inteligência artificial pode sim, ser considerada como um modelo inerentemente interpretável e se tornar uma *caixa de vidro* em que todos com auxílio de técnica de inteligência artificial podem tornar o processo da rede neural explicável.

Com a necessidade de entendimento das características dos atributos combinados e diante de dados com distribuição não-linear, optou-se neste trabalho por um modelo não supervisionado desenvolvido por Kohonen em 1982 denominado de SOM, sendo especialmente adequado para assimilação dos dados porque possui propriedades de visualização como destaque (VESANTO; ALHONIEMI, 2000), mas o conjunto de dados permitia uso de análise de componentes principais ou análise de agrupamento dos dados. O objetivo de métodos não supervisionados é identificar agrupamentos em conjuntos não rotulado de vetores de dados compartilhando semelhanças entre si. Isso ajuda a construir um modelo cognitivo que realiza a inter-relação dos dados (RAUBER; MERKL; DITTENBACH, 2002). Conforme proposto por Kohonen (1982), a captura das características de um conjunto de entrada de dados, com uma distribuição não-linear, é efetivada com a construção de uma rede neural complexa ao redor de uma grade uni ou bidimensional de neurônios. A ordenação dos dados de entrada é estruturada pelos vetores de peso dos neurônios denominados de protótipos, sendo inspirado na neurobiologia, o algoritmo SOM apresenta três processos essenciais para auto-organização da grade iniciada: competição, cooperação e auto amplificação ou adaptação (KOHONEN, 1990; HAYKIN, 2001).

2.2.1 Processo Competitivo

O modelo generalista pode servir para descrever os fenômenos dos dados iniciais de forma a ordenar coletivamente um sistema complexo, que mostra características estatísticas representativas, mesmo com o espaço de entrada totalmente desordenados (HAYKIN, 2001). No entanto, o modelo SOM é associado automaticamente com os nós da grade retangular ou hexagonal, Figura 6, geralmente bidimensional, de forma ordenada e de modo ao qual os modelos mais similares se juntam a nós mais próximos na grade, sendo o contrário para modelos menos semelhantes que ficam equidistantes, assim esse processo possibilita uma visão topológica dos dados (KOHONEN, 2013).

Figura 4: Mapa Topológico - Grade Retangular e Hexagonal



Fonte: (AGGARWAL et al., 2018, p.451)

Para elucidar o funcionamento de um mapa auto-organizável, considere um conjunto de entrada X devidamente normalizado que será usado para o treinamento da rede com tamanho predeterminado a priori e com dimensão do espaço de entrada m , sendo $X = [x_1, x_2, \dots, x_m]^T$. Cada neurônio j tem o seu vetor peso correspondente, sendo $W_j = [w_{j1}, w_{j2}, \dots, w_{jm}]^T$, j indo de 1 até l com escolha inicial aleatória.

Na etapa da competição, segundo Haykin (2001), um vetor de entrada é selecionado aleatoriamente junto com o vetor de peso sináptico do neurônio j , que ao fazer o produto

interno dos dois vetores, sendo $W_j^T X$ para $j = 1, 2, \dots, l$, transpondo o vetor de entrada é possível fazer a seleção do maior produto. A vizinhança topológica dos neurônios excitados é centrada e como a maximização do produto interno dos vetores X e W_j é a minimização da distância euclidiana dos mesmos, ao criar um índice $i(X)$ para identificar o neurônio que melhor se relaciona com o vetor de entrada X , pode-se definir $i(X)$ por:

$$i(X) = \arg \min_j \|X - W_j\|, \quad j = 1, 2, \dots, l \quad (3)$$

Esse processo resume a competição entre os neurônios. A Equação 3 mostra o $i(X)$ é o objetivo desse processo, pois nesta etapa é importante a identidade do neurônio i e o neurônio que satisfaz a condição é denominado neurônio vencedor para o vetor de entrada X .

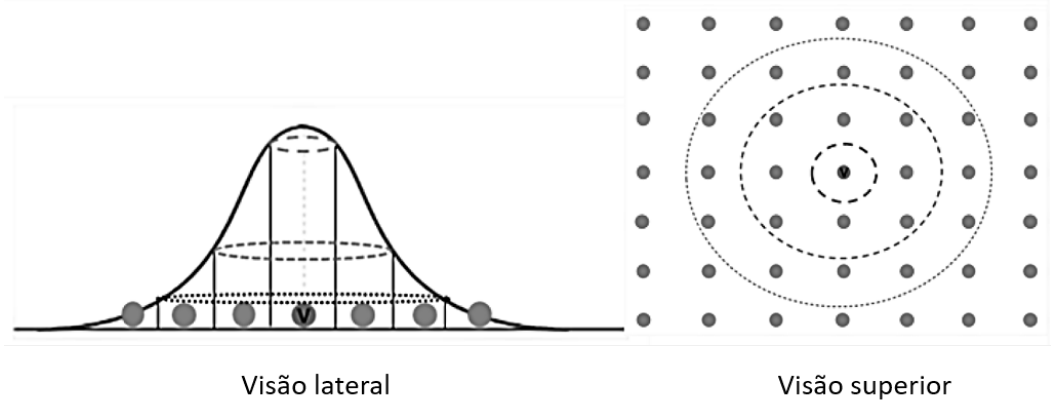
2.2.2 Processo Cooperativo

O processo de cooperação inicia-se quando o neurônio vencedor é atualizado em torno de uma vizinhança topológica dos neurônios mais próximos, sendo semelhantes em torno de um raio r . No entanto, é necessário definir a vizinhança topológica para que apenas os neurônios adjacentes sejam atualizados e, ao mesmo tempo, que se tenha uma forma que a vizinhança decaia de forma suave com a distância lateral (HAYKIN, 2001). Considerando $h_{j,i}$ a vizinhança topológica centrada em torno do neurônio vencedor i que contenha um conjunto de neurônios excitados, sendo um neurônio desse conjunto representado por j e $d_{j,i}$ a distância lateral do neurônio excitado j pelo neurônio vencedor i , então pode-se assumir que a vizinhança topológica Equação 4 é uma função unimodal da distância $d_{j,i}$, desde que satisfaça as condições de simetria em relação ao ponto máximo com $d_{j,i}=0$ e a amplitude decresce com o aumento da distância lateral $d_{j,i}$ (HAYKIN, 2001).

$$h_{j,i(X)} = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right) \quad (4)$$

A Equação 4 é para quando satisfazem as condições descritas anteriormente e pode ser considerada uma função gaussiana como visto na Figura 5.

Figura 5: Visão da Função de vizinhança Gaussiana



Fonte: (SILVA; PERES; BOSCARIOLI, 2017)

Como a vizinhança topológica tem certa dependência com a distância lateral, como visto na Equação 4, então uma grade unidimensional $d_{j,i}$ é um inteiro igual a $|j - i|$. Mas quando for bidimensional é definida pela Equação 5 onde o vetor discreto \mathbf{r}_j é a posição do neurônio excitado j e \mathbf{r}_i é a posição do neurônio vencedor i , sendo ambos medidos no espaço discreto de saída. Uma característica interessante do SOM é que o tamanho da vizinhança topológica diminui com o tempo e isso se faz com a largura σ da Equação 4 diminua com o tempo (RITTER et al., 1992).

$$d_{j,i}^2 = \left\| \mathbf{r}_j - \mathbf{r}_i \right\|^2 \quad (5)$$

$$\sigma(n) = \sigma_0 \exp\left(-\frac{n}{\tau_l}\right) \quad n = 0, 1, 2, \dots, \quad (6)$$

sendo que σ_0 é o valor de σ ao iniciar o SOM e τ_l uma constante de tempo. Desta forma, temos a vizinhança topológica definida como uma variável temporal conforme Equação 7:

$$h_{j,i(X)}(n) = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2(n)}\right) \quad n = 0, 1, 2, \dots, \quad (7)$$

Então pode-se concluir que quando o tempo n aumenta, a vizinhança topológica diminui de forma exponencial, assim como a largura $\sigma(n)$. Essa vizinhança topológica $h_{j,i(X)}(n)$ foi utilizada no estudo e será referência a partir desse ponto.

2.2.3 Processo Adaptativo

Na fase de adaptação, para que a grade seja auto-organizável, o vetor de peso sináptico W_j do neurônio j da grade deve ser alterado em relação ao vetor de entrada \mathbf{X} (HAYKIN, 2001). Entretanto, existe um problema de saturação dos pesos no final do processo, mas pode ser corrigido alterando a hipótese hebbiana com um termo de esquecimento $g(y_i)\mathbf{W}_j$ em que \mathbf{W}_j é o peso sináptico do neurônio j e $g(y_i)$ é a função escalar positiva da resposta y_i , sendo que o termo constante da expansão da série de Taylor de função $g(y_i)$ seja zero (HAYKIN, 2001). Conforme Kohonen (1982), a Equação 8 de atualização é definida por:

$$W_j(n+1) = W_j(n) + \eta(n)h_{j,i(x)}(n)(x - w_j(n)) \quad (8)$$

Em que n é igual a época, $\eta(n)$ é a taxa de aprendizagem e $h_{j,i(x)}(n)$ é a função de vizinhança.

Segundo Kohonen (1982) existem duas fases no processo adaptativo: uma de ordenação, que organiza a topologia dos vetores de peso e outra de convergência, que faz o ajuste do mapa de características, produzindo uma quantização estatística do espaço de entrada. Na fase de ordenação o parâmetro da taxa de aprendizagem $\eta(n)$ inicia-se com valor de 0,1 e vai decrescendo até um valor próximo de 0,01. A função de vizinhança deve conter quase todos os neurônios da grade ao redor do neurônio vencedor i com redução lenta com o tempo e pode requisitar 1.000 ou mais iterações. Na fase de convergência o número de iterações deve ser no mínimo 500 vezes o número de neurônios da grade e para uma boa precisão estatística, o parâmetro da taxa de aprendizagem $\eta(n)$ deve ser próximo de 0,01 e nunca zero. A função de vizinhança tem apenas os vizinhos mais próximos do neurônio vencedor que pode ser um ou zero neurônios vizinhos (HAYKIN, 2001).

Segundo Kaski (1997, p.21), o algoritmo padrão de treinamento pode ser apresentado da seguinte forma:

1. Começando com as definições dos parâmetros
 - Inicialização dos pesos w_{jm} ;
 - Definição dos parâmetros de vizinhança;
 - Definição dos parâmetros de aprendizagem;
2. Enquanto for falso, faça:
 - (a) Para cada j calcule:
 - i. $D(j) = \arg \min_j \|\mathbf{x}_j - \mathbf{w}_j\|$
 - ii. Entrar o índice J tal que $D(J)$ seja mínimo
 - iii. $\forall j \in Nc$ de J , e $\forall i, t$:
 - iv. $\mathbf{w}_j(t + 1) = \mathbf{w}_j(t) + \alpha[\mathbf{x} - \mathbf{w}_j(t)]$
 - (b) Atualização da taxa de aprendizagem
 - (c) Redução do raio de vizinhança

A originalidade do aprendizado do SOM apresentado no estudo, pode contribuir na melhoria do diagnóstico dos pacientes com TDAH ao apresentar um conjunto de protótipos similares que representam a combinação de atributos e posteriormente o utilização com algoritmo de classificação, como por exemplo a árvore de decisão, que separa, de forma direta e rápida, os pacientes propensos ao transtorno. Com isso, o estudo pode ajudar os profissionais da área de saúde e os pesquisadores a desenvolver ferramentas mais precisas e reduzir o custo de aplicação dos testes diagnósticos que atualmente inviabilizam suas aplicações no sistema público de saúde.

2.3 Modelo de Árvore de Decisão

Conforme pesquisas em psicologia cognitiva desenvolvida por Kahneman et al. (1982), a capacidade do ser humano em apreender e manipular conceitualmente vários blocos de conhecimento é restrita pelo processamento físico e cognitivo de uma parte da memória de curto prazo do cérebro. Essa restrição valoriza uso de técnicas de manipulação e apresentação dimensional, como por exemplo a árvore de decisão, capaz de preservar e refletir relacionamentos de alta dimensionalidade de forma prontamente compreensível e faz com que os relacionamentos possam ser consumidos com certa facilidade e aplicados por humanos.

A árvore de decisão é construída com algoritmo que dividem em diversas maneiras um conjunto de dados em segmentos semelhantes a ramos. Estes segmentos criam uma árvore de decisão invertida com início em um nó raiz no topo da árvore até a folha na sua extremidade. Cada objeto do estudo é refletido no nó raiz, sendo uma exibição simples e unidimensional na visão da árvore de decisão. O nome do atributo é exibido junto com a disseminação dos valores que estão contidos no atributo. Sua exibição mostra todos os registros do conjunto de dados, atributos e seus valores são visualizados no objeto da análise. O desenvolvimento da regra de decisão para formação das ramificações sob o nó raiz tem, no método de extração, uma relação entre o objeto da análise e um ou mais atributos que são utilizados como atributos de entrada para criar os ramos ou segmentos, sendo usados para estimar o valor provável do atributo de destino ou resultado ou atributo dependente (VILLE; NEVILLE, 2013).

De acordo com CASTRO e FERRARI (2016) a estrutura da árvore de decisão é composta por nós internos que corresponde a um teste de atributo, cada ramo representa o resultado do teste e as classes ou distribuições de classes são representadas pelos nós folhas. O nó inicial representa o nó raiz e o caminho desse nó até o nó folha é denominado regra de classificação. A construção da árvore de decisão pode ser usada para classificar um objeto de classe desconhecido e a estimativa é feita testando os valores dos atributos na árvore e percorrendo até chegar ao nó folha. Desta maneira, o resultado do algoritmo da árvore de decisão se torna um processo de fácil compreensão e visualização.

O algoritmo de árvore de decisão conforme apresentado por Linoff e Berry (2011) é uma coleção hierárquica de regras que descreve como dividir uma grande coleção de objetos em grupos sucessivamente menores de objetos. A partir de cada divisão sucessiva, os membros dos seguimentos resultantes se tornam mais parecidos entre si.

Segundo Ville e Neville (2013) o processo de geração da árvore de decisão pode ser dividido em seis subprocessos:

1. Pré-processamento dos dados para construção do mecanismo de crescimento da árvore de decisão;
2. Definição das características de modelagem de entrada e destino;
3. Seleção dos parâmetros de crescimento da árvore;
4. Agrupamento e processamento de cada atributo de entrada de formação da ramificação;
5. Seleção das ramificações da árvore de decisão candidata;
6. E por fim, completar a forma e o conteúdo:
 - Pare, cresça, corte ou itere a árvore de decisão;
 - Selecionar a árvore de decisão final.

Para seleção dos atributos a ser escolhido para divisão, pode-se usar o ganho de informação (*Gain*), que é um dos métodos mais conhecidos e tem como base geradora a entropia que é uma medida de pureza (COLIN, 1996). Deste modo, o ganho é a redução esperada da entropia e tem com função principal a divisão dos atributos no conjunto de dados. Já a entropia de Shannon mede a pureza do conjunto de dados (MITCHEL, 1997), sendo uma medida heterogeneidade dos dados de entrada (S) em relação a sua classificação (m). O valor esperado da entropia do atributo $E(A)$ é dado pela Equação 9 e a informação esperada de S é dada pela Equação 10. Assim, o fator chave é o uso de uma função de ganho que permite comparar os atributos (A) para eleger o mais relevante. O atributo escolhido é aquele que maximiza o ganho de informação que é calculado como sendo (COLIN, 1996) (CASTRO; FERRARI, 2016):

Com n_{ij} sendo a quantidade de objetos na classe C_i em um subconjunto S_j , então a informação esperado do atributo pode ser:

$$E(A) = \sum_{j=1}^v \frac{n_{ij} + \dots + n_{mj}}{n} * I(n_{ij}, \dots, n_{mj}) = \sum_{v \in \text{valores}(A)} p(A_v) Entropia(A_v) \quad (9)$$

Seja p_i a proporção de dados em S pertencentes a classe i , sendo estimado como n_i/n e C é o rótulo do atributo de classe que pode assumir m valores diferentes que definem m classes distintas, sendo n_i a quantidade de objetos de S na classe C_i , então a informação esperada para classificação do objeto é:

$$I(S) = I(C_1, C_2, \dots, C_m) = \sum_{k=1}^m -p_i \log_2 p_i \quad (10)$$

$$Ganho(S, A) = I(S) - E(A) \quad (11)$$

Desta forma, o ganho de informação é dado pela Equação 11 e representa a redução esperada na entropia quando conhecido o valor do atributo A , visto que o processo calculo o ganho para cada atributo, escolhendo o atributo com maior ganho para ser testado no conjunto S . Esse processo cria a divisão dos objetos para formação da árvore de decisão, dando origem ao nó, rotulando ao atributo e criando ramos para cada valor do atributo.

Neste trabalho propõe-se usar a base de dados do estudo de (SILVA, 2020) com finalidade de segmentar as crianças e adolescentes com ou sem TDAH, ajudando com Mapas Auto-Organizáveis e classificação dos protótipos, o entendimento e a formação do diagnóstico em caso de prevalência do transtorno de Déficit de Atenção e Hiperatividade (TDAH).

2.4 Elementos do Transtorno de Déficit de Atenção e Hiperatividade (TDAH)

O TDAH faz parte do grupo dos transtornos do neurodesenvolvimento com início na infância, mas que se estende até a vida adulta (KOUMOULA, 2012). Do ponto de

vista cognitivo-comportamental, caracteriza-se por déficits em diversas funções cognitivas como atenção, principalmente a atenção seletiva, a sustentada, a alternada e a dividida, déficits em controle inibitório, velocidade de processamento, organização, capacidade de inibir informações distrativas, déficits na flexibilidade cognitiva, comportamentos de hiperatividade, inquietação e impulsividade (LEAHY, 2017). O TDAH acomete 5,29% da população de crianças mundial (POLANCZYK et al., 2007). Dessa população, 30% a 70% mantêm os sintomas na fase adulta da vida (SIMON et al., 2009).

De acordo com DSM-5 (ASSOCIATION et al., 2014), o TDAH pode ser classificado em função da predominância dos eixos sintomáticos como predominante desatento, predominante hiperativo-impulsivo ou combinado. Padrões comportamentais são importantes no diagnóstico do TDAH. Eis algumas descrições dos pais em relação às crianças: dificuldade de escutar, de obedecer, de seguir regras de rotina, muitas das vezes postergam e esquecem das atividades do cotidiano, dificuldade de seguir instruções diretas, dificuldade em regular sentimento de frustração, exacerbação de atividade motora, impulsividade em mudar de atividade antes de ter uma completada, dificuldade em aguardar sua vez e as crianças podem ter prejuízos nas relações sociais. Esses comportamentos podem contribuir para ambientes familiares ou escolares com alto nível de estresse (BENCZIK; CASELLA, 2015; BARKLEY; MURPHY, 2008).

A desigualdade dos eixos sintomáticos dentro do subgrupo predominante desatento compromete sua validade quando comparado ao subgrupo combinado. Porém, indivíduos considerados apenas desatentos, mas com um nível subclínico de sintomas de hiperatividade-impulsividade (4 ou 5 sintomas), tem sua classificação sem intensidade do subtipo combinado do TDAH. Todavia, ainda são poucas as provas sobre a diferença qualitativa entre os subtipos apresentados, mesmo quando a classificação desatento é estabelecida a indivíduos com três ou menos sintomas de hiperatividade-impulsividade (WILLCUTT et al., 2012). Em diferentes fases da vida o TDAH pode mostrar-se em cerca de 2:1 no caso de crianças e 1,6:1 em adultos. Essa mudança da prevalência de crianças para adultos hipoteticamente se dá porque crianças e adolescentes podem criar métodos que suprem o transtorno ao longo do seu desenvolvimento, tornando-as mais funcionais e ocultando algumas dificuldades e sintomas (POLANCZYK et al., 2007).

De acordo com a nova classificação da Association et al. (2014), os pontos de destaques

são as mudanças nas diversas formas dos sintomas, tentando contextualizar os critérios diagnosticados ao longo da vida do indivíduo; mudança na idade de início do surgimento dos sintomas, saindo de 7 para os 12 anos de idade; alteração do termo “subtipo” por “apresentação atual”; e retirada dos transtornos do espectro autista como fatores excludentes no diagnóstico.

Segundo Association et al. (2014), são cinco os critérios de Diagnósticos, sendo o primeiro critério de diagnóstico caracterizado por um padrão persistente de desatenção e/ou hiperatividade que interfere no funcionamento e desenvolvimento da pessoa. A desatenção apresenta seis ou mais dos seguintes sintomas que duram no mínimo seis meses em um grau que não condiz com o desenvolvimento do indivíduo e impacta diretamente nas atividades sociais e acadêmicas/ profissionais: a) Frequentemente não presta atenção em detalhes ou comete erros por descuido em tarefas escolares, no trabalho ou durante outras atividades (p. ex., negligencia ou deixa passar detalhes, o trabalho é impreciso); b) Frequentemente tem dificuldade de manter a atenção em tarefas ou atividades lúdicas (p.ex., dificuldade de manter o foco durante aulas, conversas ou leituras prolongadas); c) Frequentemente parece não escutar quando alguém lhe dirige a palavra diretamente (p.ex., parece estar com a cabeça longe, mesmo na ausência de qualquer distração óbvia); d) Frequentemente não segue instruções até o fim e não consegue terminar trabalhos escolares, tarefas ou deveres no local de trabalho (p. ex., começa as tarefas, mas rapidamente perde o foco e facilmente perde o rumo); e) Frequentemente tem dificuldade para organizar tarefas e atividades (p. ex., dificuldade em gerenciar tarefas sequenciais; dificuldade em manter materiais e objetos pessoais em ordem; trabalho desorganizado e desleixado; mau gerenciamento do tempo; dificuldade em cumprir prazos); f) Frequentemente evita, não gosta ou reluta em se envolver em tarefas que exijam esforço mental prolongado (p. ex., trabalhos escolares ou lições de casa; para adolescentes mais velhos e adultos, preparo de relatórios, preenchimento de formulários, revisão de trabalhos longos); g) Frequentemente perde coisas necessárias para tarefas ou atividades (p. ex., materiais escolares, lápis, livros, instrumentos, carteiras, chaves, documentos, óculos, celular); h) Com frequência é facilmente distraído por estímulos externos (para adolescentes mais velhos e adultos, pode incluir pensamentos não relacionados); i) Com frequência é esquecido em relação a atividades cotidianas (p. ex., realizar tarefas, obrigações; para adolescentes mais velhos e adultos, retornar ligações, pagar contas, manter horários agendados).

De acordo com Association et al. (2014), a segunda característica do primeiro critério é denominada hiperatividade apresentando seis ou mais dos seguintes sintomas que duram no mínimo seis meses em um grau que não condiz com o desenvolvimento do indivíduo e impacta diretamente nas atividades sociais e acadêmicas/profissionais: a) Frequentemente remexe ou batuca as mãos ou os pés ou se contorce na cadeira; b) Frequentemente levanta da cadeira em situações em que se espera que permaneça sentado (p. ex., sai do seu lugar em sala de aula, no escritório ou em outro local de trabalho ou em outras situações que exijam que se permaneça em um mesmo lugar); c) Frequentemente corre ou sobe nas coisas em situações em que isso é inapropriado. (Nota: Em adolescentes ou adultos, pode se limitar a sensações de inquietude.); d) Com frequência é incapaz de brincar ou se envolver em atividades de lazer calmamente; e) Com frequência “não para”, agindo como se estivesse “com o motor ligado” (p. ex., não consegue ou se sente desconfortável em ficar parado por muito tempo, como em restaurantes, reuniões; outros podem ver o indivíduo como inquieto ou difícil de acompanhar); f) Frequentemente fala demais; g) Frequentemente deixa escapar uma resposta antes que a pergunta tenha sido concluída (p. ex., termina frases dos outros, não consegue aguardar a vez de falar); h) Frequentemente tem dificuldade para esperar a sua vez (p. ex., aguardar em uma fila); i) Frequentemente interrompe ou se intromete (p. ex., mete-se nas conversas, jogos ou atividades; pode começar a usar as coisas de outras pessoas sem pedir ou receber permissão; para adolescentes e adultos, pode intrometer-se em ou assumir o controle sobre o que outros estão fazendo).

Conforme Association et al. (2014), o segundo critério diz respeito a faixa etária da criança que vários sintomas de desatenção ou hiperatividade-impulsividade aparecem antes dos 12 anos de idade e o terceiro critério diz que esses sintomas estão presentes em dois ou mais ambientes como por exemplo em casa, na escola, no trabalho; com amigos ou familiares. O quarto critério comenta da interferência dos sintomas no funcionamento social, acadêmico ou profissional ou na redução de sua qualidade. O quinto e último critério comenta da não ocorrência dos sintomas durante o curso de esquizofrenia ou outro transtorno psicótico e não tem sua explicação em outros transtornos, como por exemplo, transtorno do humor, transtorno de ansiedade, transtorno dissociativo, entre outros.

As crianças que habitualmente estão na fase pré-escolar podem apresentar níveis ele-

vados de atividade motora, déficit de atenção e controle inibitório deficiente que são manifestações comportamentais do TDAH. Entretanto, nos casos clínicos, são mais expressivos e resultam em prejuízo considerável, com elevadas taxas de acidentes e desempenho escolar fraco que pode persistir na idade escolar em 60% a 80% dos casos (CHERKASOVA et al., 2013). Aproximadamente 70% das crianças em idade escolar apresentam pioras nas atividades escolares e prejuízo no convívio familiar e no relacionamento com outras crianças. Geralmente, os sintomas de desatenção em relação aos sintomas de hiperatividade apresentam uma desaceleração com a idade, ou seja, diminui vagarosamente entre a passagem da fase etária de criança, adolescência e adulta com maior persistência (DALSGAARD et al., 2014). Cerca de 33% das crianças com TDAH na fase adulta não apresentam mais os sintomas, ao contrário do restante que continuam a ter o transtorno ou quadros que resultam em perdas. Por um período prolongado, adultos com TDAH apresentam piora no desempenho acadêmico e profissional, assim como aumento de infrações de trânsito, acidentes com veículos automotores e comportamento sexual com elevado risco e aparecimento concomitante de várias doenças psiquiátricas, tais como transtorno de ansiedade, transtornos do humor e ingestão de substâncias (FUENTES et al., 2014).

Os fatores genéticos e condições ambientais contribuem muito para uma etiologia complexa do TDAH com bases neurobiológicas estabelecidas por pesquisas como por exemplo realizadas por (WILLIAMS et al., 2010) que apontam para padrões fora do normal no sistema nervoso central de pessoas com TDAH (PURPER-OUAKIL et al., 2011). Segundo Fuentes et al. (2014), o estudo de endofenótipos neuropsicológicos é motivado pela complexidade etiológica e a inomogeneidade clínica do TDAH. Endofenótipos, por vezes chamados de “fenótipos intermediários”, são traços herdados e aferidos que podem ser encontrados no caminho de ligação de um genótipo a transtornos neuropsiquiátricos complexos. Contudo, os indivíduos com TDAH podem apresentar ou não, diferentes formas de padrões seguindo um perfil cognitivo complexo e múltiplas variações, como por exemplo perda ou baixo foco atencional causando uma falta de flexibilidade, dificuldade em lidar com distratores, dificuldade de auto regulação, impulsividade comportamental, falta de coordenação motora e embaralhamento das informações mentais devido dificuldade de organização. Os sinais apresentados mostram que a criança pode ter um prejuízo no seu desenvolvimento intelectual em diferentes áreas do cérebro e são mostrados por interações lúdicas, por observação e por meio de padronização de instrumentos, tornando-se funda-

mental na mensuração das queixas de desatenção e hiperatividade (CARREIRO et al., 2018).

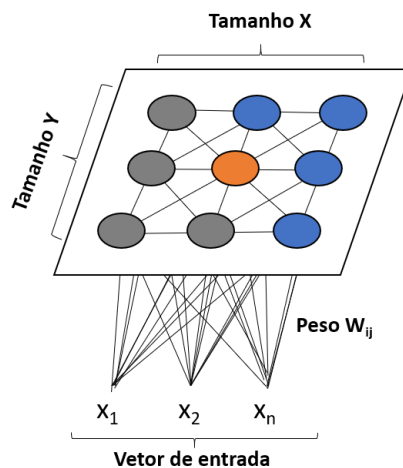
Padrões comportamentais são importantes no diagnóstico do TDAH, os pais relatam que as crianças tem dificuldade de escutar por falta de atenção quando alguém dirige a palavra diretamente, de obedecer e seguir regras e rotinas, muita das vezes postergam e esquecem das atividades do cotidiano, têm dificuldade de seguir instruções diretas, não aceitam frustração, têm exacerbação de atividade motora, podem apresentar alguma forma impulsiva em uma atividade antes de ser completada, não conseguem esperar, relações sociais prejudicadas, ambientes familiares ou escolares com alto nível de estresse (BENCZIK; CASELLA, 2015). A observação de várias pessoas pode aumentar a precisão do diagnóstico, possibilitando nos casos de comorbidades e sintomas inconsistentes, a possibilidade diferenciada do diagnóstico. Com isso, cria-se a necessidade de fazer um relato de forma padronizada, para não ter viés e possibilidade de construção de um perfil de comportamento representativo para avaliação. Na atualidade, a literatura da psicologia, conta com diversos instrumentos que são aplicados em pais e professores, de maneira a extrair o máximo de informações dos dois ambientes das crianças (CARREIRO et al., 2018). Estes instrumentos que são vistos nesse trabalho como atributos, têm um direcionamento com base em protocolo para avaliar queixas de TDAH, sendo os atributos analisados por parâmetros de avaliação cognitivas utilizados, como Teste de Atenção por Cancelemento, Teste de Trilhas, Continuous Performance Test, Teste de Cartas Wisconsin, Escala Wechsler de Inteligência Abreviada e Escala Wechsler de Inteligência para Crianças, além dos inventários de comportamentos para crianças e adolescentes (CBCL/6-18) e para professores (TRF/6-18).

Diante da importância da coleta de várias informações no âmbito cognitivo e da neuropsicologia e da análise do comportamento, o tratamento e análise multivariada dos dados podem nos ajudar a obter informações relevantes no entendimento das queixas do TDAH e as técnicas de inteligência artificial utilizadas se tornam elementos fundamentais na discriminação diagnóstica.

3 MÉTODO HÍBRIDO BASEADO EM INTELIGÊNCIA ARTIFICIAL

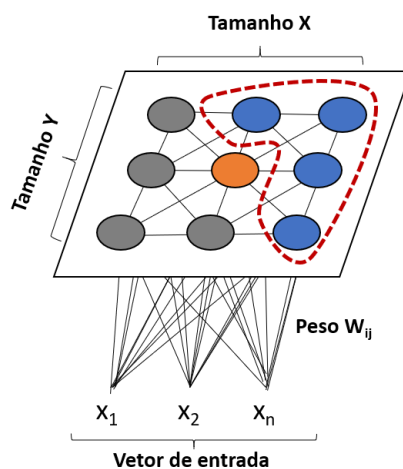
O trabalho apresenta como proposta dentre os modelos de aprendizado não supervisionado o mapa auto-organizáveis como sendo um método para segmentar os neurônios da grade com maior incidência de diagnósticos do transtorno. O processo se inicia pela competição dos neurônios através de uma função de discriminação para ver quem será ativado pelo conjunto de entrada de dados, sendo o neurônio com menor valor de distância o vencedor. O fato interessante é comprovado pelos neurônios vizinhos ao neurônio vencedor que têm suas características semelhantes e ficam mais excitados em relação aos neurônios mais distantes e o processo é conhecido como vizinhança topológica.

Figura 6: Visualização conceitual do mapa auto-organizável



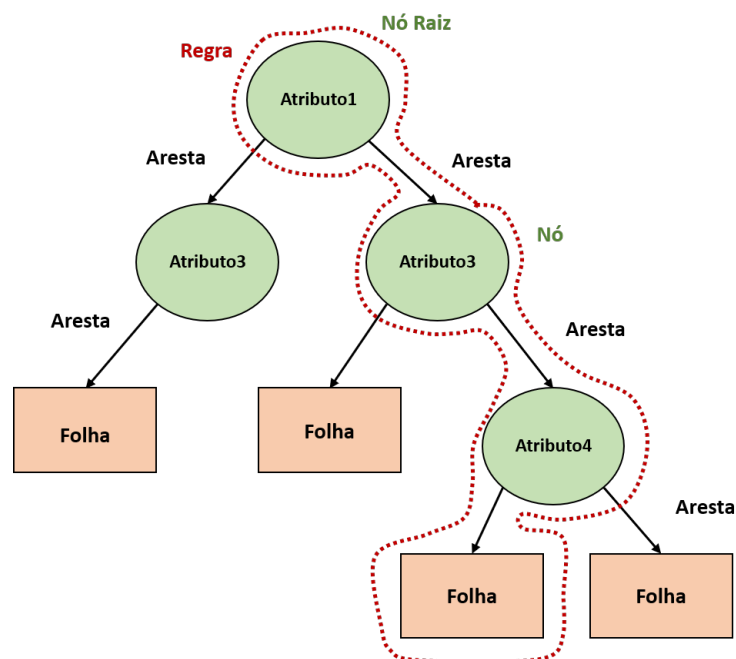
A vizinhança topológica é uma função exponencial que diminui com o tempo e faz a divisão da distância lateral quadrática entre os neurônios com o neurônio vencedor pela quantidade quadrática de vizinhos. Os pesos são calculados considerando a quantidade de iterações e taxa de aprendizagem dos dados no tempo fazendo uma movimentação na vizinhança topológica, ao qual os neurônios mais distantes têm suas atualizações reduzidas. A visualização do mapa topológico do SOM, termina quando não há mais alteração na vizinhança. Desta forma, cada neurônio fica com uma parte das características dos dados e os neurônios com características semelhantes formam um conjunto de atributos que é denominado segmento.

Figura 7: Segmentação dos neurônios no mapa auto-organizável



No estudo, o segmento observado é o diagnóstico do transtorno, sendo um atributo de classe composto por duas classes, uma com TDAH e outra sem TDAH. Entretanto, para alguns neurônios o mapa auto-organizável não consegue fazer a separação, pois há uma confusão interna de classes que denominamos no trabalho como sobreposição. A sobreposição de classes no neurônio é interpretada como sendo uma situação de possível dúvida no diagnóstico por um especialista. A identificação da sobreposição em cada neurônio do mapa topológico se dá pelo cálculo da entropia de Shannon que mede a impureza do neurônio perante as classes, ou seja, quanto mais confuso estiver o diagnóstico dentro do neurônio, maior será o grau de impureza. Feito esse processo em todos os neurônios, opta-se por pegar os neurônios com maior representatividade de dados e com maior grau de impureza. O conhecimento do neurônio com sobreposição nos possibilita utilizar a árvore de decisão para identificar quais atributos discriminam o diagnóstico do transtorno, neste caso, tem a função de ganho de informação como aliada para ordenar os atributos discriminantes e ajudar na construção das regras da árvore de decisão.

Figura 8: Estrutura da árvore de decisão



Diante deste fato, buscou-se a combinação do algoritmo da árvore de decisão que é um modelo supervisionado utilizado para classificação de dados que auxiliam na identificação de um ou mais atributos provenientes de ferramentas padronizadas de avaliação que, neste caso, foram testes cognitivos e inventários de avaliação comportamental. Essas ferramentas são usadas para testar o aprendizado das características do TDAH. No treinamento e teste do algoritmo foi utilizado o método de validação cruzada de omissão "Leave One Out Cross-Validation" que é uma abordagem de validação cruzada considerando cada objeto como um conjunto de teste e o restante como um conjunto de treinamento. O processo é repetido N vezes para cada objeto do conjunto de dados do neurônio com sobreposição. Na finalização tem-se o valor esperado para cada objeto e fazendo a medição com os objetos observados no conjunto de dados. Com isso, o objetivo deste algoritmo da árvore de decisão verifica a acurácia do modelo para confirmação dos casos com diagnóstico de TDAH identificando quais são as ferramentas de avaliação que melhor contribuíram com a confirmação de TDAH. Depois se aplica a estatística de Kappa para medir a concordância e confiabilidade dos atributos de diagnóstico observado com o atributo de diagnóstico esperado que é a medição da concordância do diagnóstico de um especialista comparado ao diagnóstico do algoritmo.

No próximo capítulo será apresentada a metodologia desenvolvida com o conjunto de dados e aplicação dos algoritmos de inteligência artificial para chegar nos resultados do trabalho.

4 METODOLOGIA

O algoritmo de mapas auto-organizáveis (WEHRENS; BUYDENS, 2007) e árvore de decisão (THERNEAU et al., 2015) foram executados utilizando o software R na versão 4.0.3 (CRAN, 2020) que é um programa de computador gratuito que tem sua linguagem voltada a funcionalidades e contempla uma vasta gama de bibliotecas para aprendizagem de máquinas e mineração de dados. O estudo foi desenvolvido utilizando um computador padrão com windows10, processador intel i7 8ª geração, com 16GB de memória RAM e no software R os pacotes: tidyverse, sqldf, kohonen, summarytools, rpart, rpart.plot, ggplot2 e pacman. Este capítulo de metodologia traz uma abordagem da amostra do estudo apresentado por SILVA (2020) e a forma de aplicação dos algoritmos de inteligência artificial no conjunto de dados.

4.1 Descrição e Experimentos

A amostra do estudo foi composta por 127 crianças e adolescentes entre 06 e 16 anos, sendo 48 com diagnóstico clínico de TDAH e 79 do grupo controle, sem diagnóstico de TDAH. Os atributos que compõe os testes neuropsicológicos e inventários comportamentais aplicados neste estudo são: Teste de Atenção por Cancelamento (TAC), Trilhas (TT), Escala Wechsler de Inteligência para Crianças (WISC-III), Escala de Wechsler de Inteligência para Crianças (WISC-IV), Escala Wechsler Abreviada de Inteligência (WASI), Inventário de Comportamento para Crianças e Adolescentes entre 6 e 18 anos/ Child Behavior Checklist for ages 6-18 (CBCL/6-18) e Inventário de Comportamentos para Crianças e Adolescentes entre 6-18 anos - Formulário para o Professor/ Teacher Report Form for ages 6-18 (TRF/6-18) (CARREIRO et al., 2014). O atributo diagnóstico foi

categorizado como 1 - com TDAH e 2 - sem TDAH.

Os atributos que apresentaram dados faltantes foram substituídos pelo número 999 para serem separados de forma natural pela função de distância do algoritmo e todos os atributos foram normalizados pelo método de z-score (CURTIS et al., 2016), no R utiliza o comando *scale* no conjunto de dados, a fim de padronizar as diferentes escalas dos atributos por unidade de desvio padrão (KUHN et al., 2020). Com os dados devidamente normalizados fez-se o treinamento da rede SOM usando o pacote disponível em linguagem R (WEHRENS; BUYDENS, 2007). Na biblioteca de Kohonen são utilizadas as funções *somgrid* com topo igual a hexagonal e *som* com 1.000 iterações e *alpha* igual a 0,05 e 0,01 para parametrizar e treinar o mapa respectivamente. Segundo Rubbo e Silva (2021), o tamanho do mapa pode ser descrito conforme as equações (12) e (13), sendo n o número de objetos e a constante Cm variando de $[-3,3]$ para gerar diferentes tamanhos de mapa:

$$l_{SOM} = \sqrt{n}/2 + Cm \quad (12)$$

$$Tamanho\ do\ Mapa = (l_{SOM})^2 \quad (13)$$

O tamanho da topologia do mapa depende de dois parâmetros x e y que quando multiplicados representam a dimensão da grade (xy) e são representados nos resultados por 2x2, 3x3 e 4x4. Desta forma, a hipótese do estudo foi encontrar neurônios com densidade representativa de objetos e com significativa distribuição de classes. Entender o comportamento dos neurônios depois de executado o algoritmo SOM, traz a possibilidade de entendimento das características dos atributos para formação do diagnóstico do TDAH. No estudo foram plotados dois mapas hexagonais mostrando a contribuição dos pesos dos atributos e a distribuição dos objetos por atributo de classe para construção dos neurônios.

Para selecionar os neurônios que serão submetidos a árvore de decisão, utiliza-se a função *FOR* para percorrer todos os neurônios do mapa hexagonal do SOM com dimensão $x \times y$. Conjuntamente, calcula-se a entropia de Shannon para cada neurônio e seleciona o neurônio com maior entropia e representatividade da quantidade de objetos (MITCHEL, 1997). Ao selecionar o neurônio, faz uso da função *rpat* (THERNEAU et al., 2015) para obter a árvore de decisão com a escolha do método clássico e da função ganho de

informação que tem sua relação inversa a entropia.

O conjunto de dados do neurônio é dividido para validação cruzada pelo método *leave-one-out* que na tradução literal para o português seria deixar um objeto de fora entre todos do conjunto de dados. Desta forma, o conjunto de dados (K) é dividido em conjunto de treinamento e outro de teste, sendo o treinamento um conjunto de dados com menos um objeto ($K-1$) e o teste considerando esse objeto que foi retirado (CAWLEY; TALBOT, 2003). O processo percorre todos os objetos do conjunto de dados e só termina quando o último objeto é testado pelo algoritmo da árvore de decisão.

Na finalização do processo cria-se um atributo de diagnóstico de TDAH estimado e verifica-se o parâmetro de complexidade (cp) para verificar a necessidade de poda da árvore de decisão que no estudo foi levado em consideração um cp de 0,01. Como medida de ajuste da árvore de decisão, utiliza-se a matriz de confusão que mostra em forma de tabela de contingência a classe observada pela predita a fim de medir o *desempenho do algoritmo* ou *acurácia*. A acurácia calcula a probabilidade de explicação do modelo de árvore de decisão em relação ao conjunto de dados. Para avaliar a concordância do diagnóstico observado pelo esperado ou predito, utiliza-se no estudo o coeficiente de Kappa que mede a confiabilidade da concordância entre os dois atributos, ou seja, mede se o diagnóstico obtido com a aplicação dos inventários estão de acordo com o diagnóstico estimado pelo algoritmo de árvore de decisão, sendo o resultado próximo de um, indica uma excelente concordância (KRAEMER, 2014).

4.2 Avaliação de Desempenho de Classificação

No estudo, a classe diagnóstico do TDAH é a classe cujo o valor se deseja prever e também é chamada de classe positiva, o oposto da classe é denominado de classe negativa. O Quadro 1 mostra a matriz de confusão ou matriz de erro que foi usada para analisar o desempenho do algoritmo de classificação da árvore de decisão que relaciona a classe desejada com a classe predita, sendo que a classe positiva e negativa tem suas medidas específicas definidas por:

- VN é o número correto de predições negativas;
- FP é o número falso de predições positivas;
- FN é o número falso de predições negativas;
- VP é o número correto de predições positivas.

Quadro 1: Matriz de Confusão

	Predito Negativo	Predito Positivo
Atual Negativo	VN	FP
Atual Positivo	FN	VP

A partir da matriz de confusão se consegue medir o desempenho do algoritmo com o cálculo da acurácia, que representa o número de classificações corretas dividido pelo número total de classificações, conforme demonstrado na Equação 14 e o erro é o número de classificações erradas dividido pelo número total de classificações como é visto na Equação 15.

$$Acurácia = (VP + VN)/(VP + VN + FP + FN) \quad (14)$$

$$Erro = (FP + FN)/(VP + VN + FP + FN) \quad (15)$$

A medição de concordância ou coeficiente Kappa (K) foi utilizada para medir de forma comparativa o diagnóstico de TDAH das crianças submetidas aos inventários correspon-

dentos ao transtorno pelo diagnóstico de TDAH predito na árvore de decisão. O coeficiente Kappa pode ser calculado com os resultados da matriz de confusão pela Equação 16 (SILVA; PAES, 2012; KRAEMER, 2014):

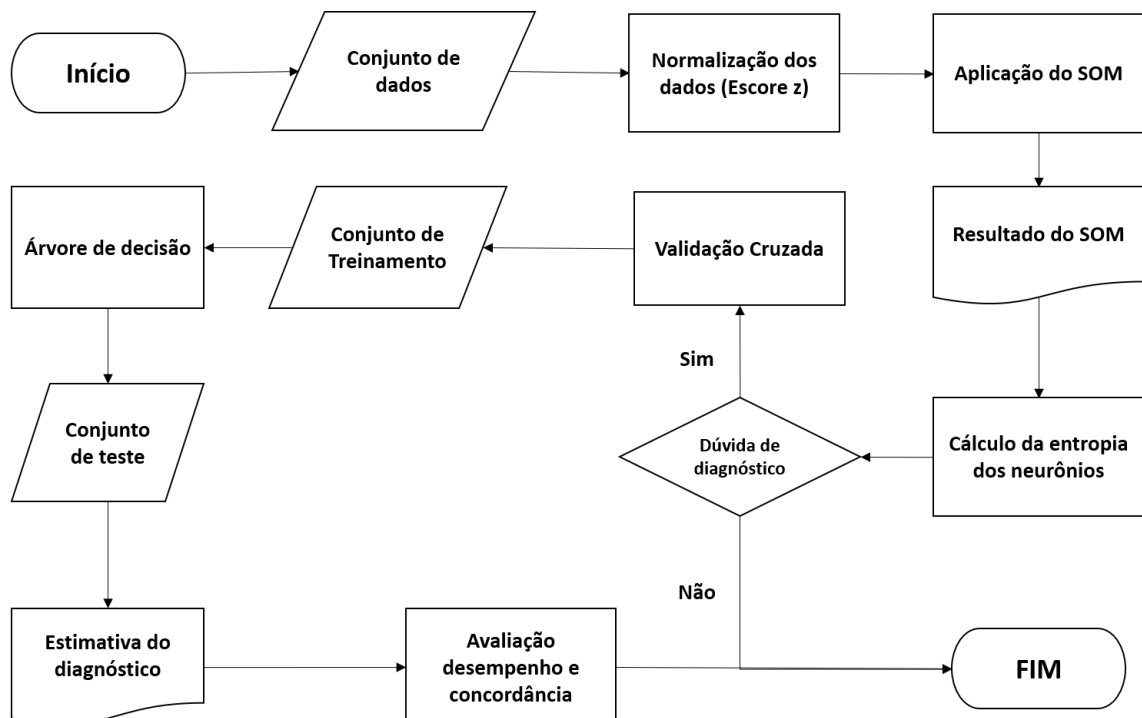
$$kappa = P(O) - P(E)/1 - P(E) \quad (16)$$

Onde $P(O)$ é a probabilidade observada de concordância (soma das respostas concordantes divididas pelo total); $P(E)$ é a probabilidade esperada de concordância (soma dos valores esperados das respostas concordantes dividida pelo total).

Segundo Silva e Paes (2012), o Kappa é uma medida de concordância interobservador que avalia o grau da concordância, assim como se está além do esperado dado ao acaso. É uma medida de valor máximo unitário que corresponde concordância absoluta e valores próximos de zero ou negativos, indicam nenhum ou falta de concordância entre os atributos que estão sendo julgados.

Para simplificar a descrição metodológica desenvolvida no estudo, apresenta um fluxograma que resumidamente descreve o processo através da programação criada para obtenção do resultado e pode ser visto na Figura 9.

Figura 9: Fluxograma da modelagem do conjunto de dados



No próximo capítulo encontram-se os resultados obtidos pelas dimensões 2x2, 3x3 e 4x4 referentes ao SOM e a aplicação da árvore de decisão para classificar e discriminar o diagnóstico do transtorno no neurônio com sobreposição.

5 RESULTADOS

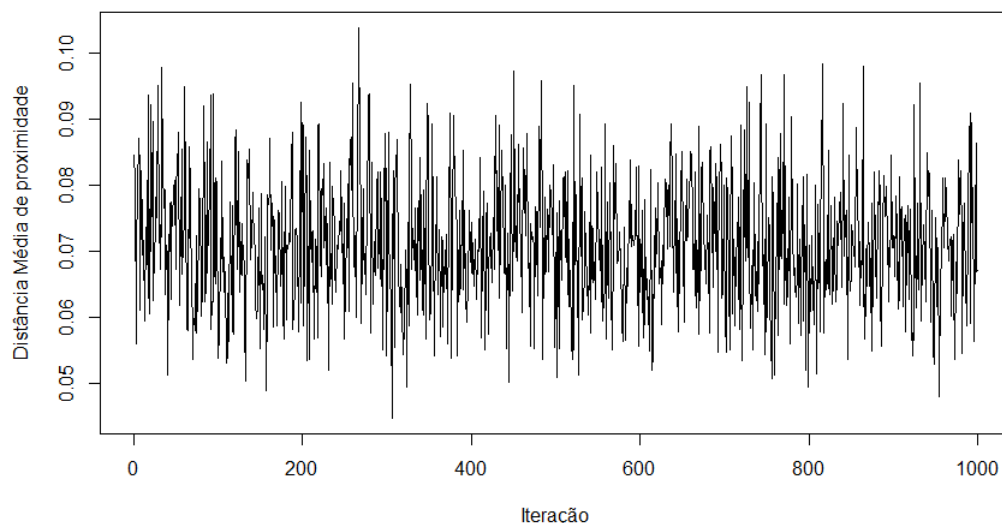
5.1 Visualização dos resultados do SOM combinados com a árvore de decisão

Nesse capítulo, serão mostrados os resultados obtidos ao aplicar o método proposto no conjunto de dados do estudo, conforme descrito na metodologia, assim como, o cálculo das dimensões dos mapas que estão de acordo com a Equação 12 e geram um intervalo entre [2,8], mas no estudo considera-se o intervalo entre [2,4], pois a dimensão 4 apresentou um neurônio com maior sobreposição que segue o objetivo do estudo. Em seguida, serão apresentados os resultados obtidos com a topologia da grade para as dimensões: 2x2, 3x3 e 4x4.

5.1.1 Resultados com grade 2x2

Analisando os resultados esperados da grade 2x2 verifica-se ao iniciar o SOM que a distância média dos neurônios a cada iteração, Figura 10, e ao longo das 1.000 iterações, acaba não sofrendo uma redução das distâncias de proximidade dos neurônios e pode ser demonstrado ao se traçar uma linha no eixo das distância média de proximidade em 0,065 paralela ao eixo das iterações. Isto sugere um aumento da dimensão da grade para identificar uma possível queda na distância de proximidade dos Neurônios que pode melhorar o agrupamento. Entretanto, optou-se por analisar os resultados da dimensão 2x2.

Figura 10: Visualização das 1.000 iterações aplicadas no SOM - grade 2x2



Ao processar o SOM e fazer a comparação do atributo de classe com os neurônios, o Quadro 2 mostra que o neurônio 4 tem a maior participação (66,9%) em relação aos demais e apresenta a maior sobreposição de classe, sendo a classe 1 (37,6%) e a classe 2 (62,4%) e entropia de 0,96.

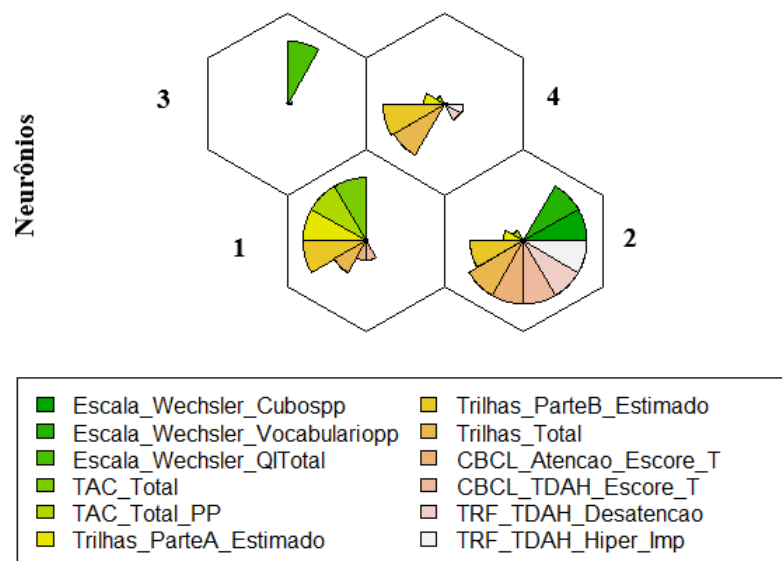
Quadro 2: Comparativo diagnóstico pelo neurônio dimensão 2x2

	Diagnóstico	1	2	Total
Neurônio	1	2 (4,2%)	1 (1,3%)	3 (2,4%)
	2	13 (27,1%)	25 (31,6%)	38 (29,9%)
	3	1 (2,1%)	0 (0,0%)	1 (0,8%)
	4	32 (66,7%)	53 (61,7%)	85 (66,9%)
Total		48 (100,0%)	79 (100,0%)	127 (100,0%)

Ao analisar a Figura 11 tem-se uma explicação de quais atributos são relevantes para formação do neurônio. A contagem dos neurônios começa da esquerda para direita e de baixo para cima, então o neurônio 1 tem como principais atributos o TAC Total, TAC Total PP, Trilhas Part A Estimado e Trilhas Part B Estimado. No neurônio 2, observa-se que o atributo Wechsler Vocabulário Z-Score, Escala Wechsler Cubo Z-Score, Escala TRF

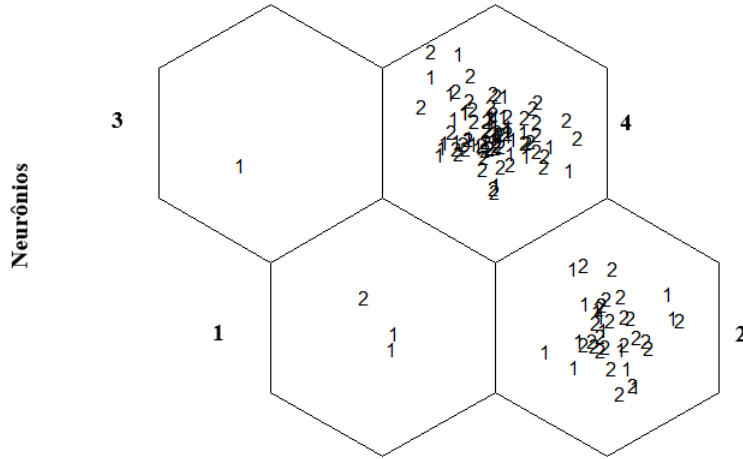
Déficit de Atenção T-Score, Escala TRF Problemas de Atenção T-Score, Escala CBCL Déficit de Atenção T-Score, Escala CBCL Problemas de Atenção T-Score e Trilhas Total têm maiores contribuições. No neurônio 3, a prevalência da Escala Wechsler Vocabulário Z-Score. Por fim, no neurônio 4, nota-se a predominância dos atributos Trilhas Total e Trilhas Parte B Estimado.

Figura 11: Contribuição dos atributos na formação do neurônios



Na Figura 12 observa-se o mapa topológico com a participação dos objetos referente ao atributo de classe. Verifica-se a maior densidade e sobreposição de classe no neurônio 4 com maior entropia entre os neurônios e maior quantidade de objetos (85). No neurônio 4 observa-se elevada sobreposição de objetos comparado ao atributo diagnóstico, mas ao verificar o mesmo neurônio no Quadro 3, Quadro 4 e Quadro 5 identifica-se o atributo CBCL Atenção Escore T e CBCL TDAH Escore T como atributos de maior contribuição na formação do neurônio 4. O CBCL é um atributo de comportamento da criança que é preenchido pelo responsável e poderia ser usado, nos casos duvidosos, como um possível instrumento para melhor entendimento do diagnóstico.

Figura 12: Dispersão dos objetos dentro dos Neurônios



Quadro 3: Inércia dos atributos dimensão 2x2 - Parte 1

Neurônios	Escala Wechsler Cubospp	Escala Wechsler Vocabuláriopp	Escala Wechsler QI Total	TAC Total
Neurônio 1	-0,11	-0,08	-0,25	6,29
Neurônio 2	0,26	0,26	-0,03	-0,22
Neurônio 3	-0,08	-0,08	10,86	-0,15
Neurônio 4	-0,09	-0,09	-0,10	-0,12

Quadro 4: Inércia dos atributos dimensão 2x2 - Parte 2

Neurônios	TAC Total PP	Trilhas Parte A Estimado	Trilhas Parte B Estimado	Trilhas Total
Neurônio 1	1,32	1,57	0,03	-0,19
Neurônio 2	-0,04	-0,05	-0,09	0,10
Neurônio 3	-0,37	-0,84	-0,75	-0,60
Neurônio 4	-0,13	-0,04	0,01	0,05

Quadro 5: Inércia dos atributos dimensão 2x2 - Parte 3

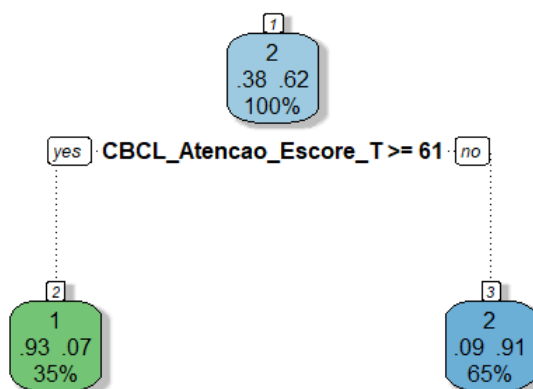
Neurônios	CBCL Atenção Escore T	CBCL TDAH Escore T	TRF TDAH Desatenção	TRF TDAH Hiper Imp
Neurônio 1	0,00	0,00	-0,47	-0,48
Neurônio 2	1,50	1,50	0,33	0,32
Neurônio 3	-0,68	-0,67	-0,47	-0,49
Neurônio 4	-0,66	-0,66	-0,25	-0,25

A escolha do neurônio 4 pela representatividade do conjunto de dados e por ter a maior sobreposição de diagnósticos dos objetos leva ao entendimento dos atributos que são mais discriminantes ao diagnóstico. Desta forma, utilizou-se a árvore de decisão para identificar as regras as quais poderiam ajudar nos casos duvidosos. Na Figura 13 verifica-se que dos 6 atributos presentes, 4 são referentes ao comportamento da criança, sendo o CBCL um inventário respondido pelo responsável e o TRF respondido pelo professor, mas o CBCL apresenta contribuição mais forte para formação de um possível diagnóstico de TDAH que é verificado na árvore de decisão Figura 14 para T-Score abaixo de 61.

Figura 13: Importância dos atributos no neurônio 4 pela árvore de decisão



Figura 14: Árvore de decisão do neurônio 4



A árvore de decisão por ser um algoritmo supervisionado tem como atributo resposta uma estimativa do atributo de classe e o resultado pode ser visto no Quadro 6. Com base na matriz de confusão gerada, observa-se que dos 53 diagnósticos positivos para TDAH coletados com aplicação dos inventários, tem-se 50 diagnósticos positivos estimados pelo algoritmo. Com base nos números apresentados, pode-se calcular a acurácia do algoritmo, assim como utilizar a estatística Kappa para medir a concordância entre os

objetos observados e os objetos estimados (KRAEMER, 2014). O resultados com a grade 2x2 gerou uma acurácia de 91% com uma confiabilidade boa do Kappa de 78%, sendo o p-valor igual a $1,87e^{-13}$.

Quadro 6: Matriz de Confusão

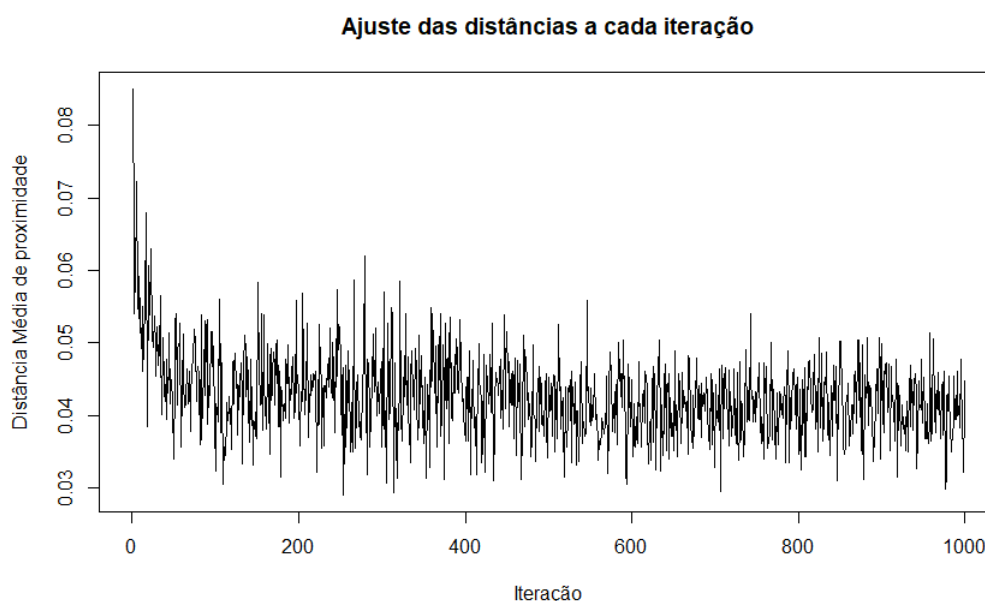
	Predito Negativo	Predito Positivo
Atual Negativo	27	5
Atual Positivo	3	50

A fim de entendimento do diagnóstico do TDAH e dos resultados apresentados do SOM com grade 2x2, optou-se por trazer os resultados do SOM com grade 3x3 para identificar possível divergência de resultado quando comparado o neurônio com maior sobreposição do diagnóstico em relação a discriminação dos atributos pela árvore de decisão.

5.1.2 Resultados com grade 3x3

Ao aumentar a dimensão da grade para 3x3, nota-se que a distância média de proximidade dos neurônios diminui a partir da iteração 100, mas oscila em torno de 0,04 distância média de proximidade, conforme Figura 15. Percebe-se que o aumento da grade reduz a distância média de proximidade, sendo que neste caso ao traçar uma linha imaginária no eixo da distância média de proximidade no ponto 0,045 e paralelo ao eixo da iteração, percebe-se que há um decaimento da curva ao longo das 1.000 iterações que denota um ajuste da vizinhança topológica pela aproximação dos neurônios.

Figura 15: Visualização das 1.000 iterações aplicadas no SOM



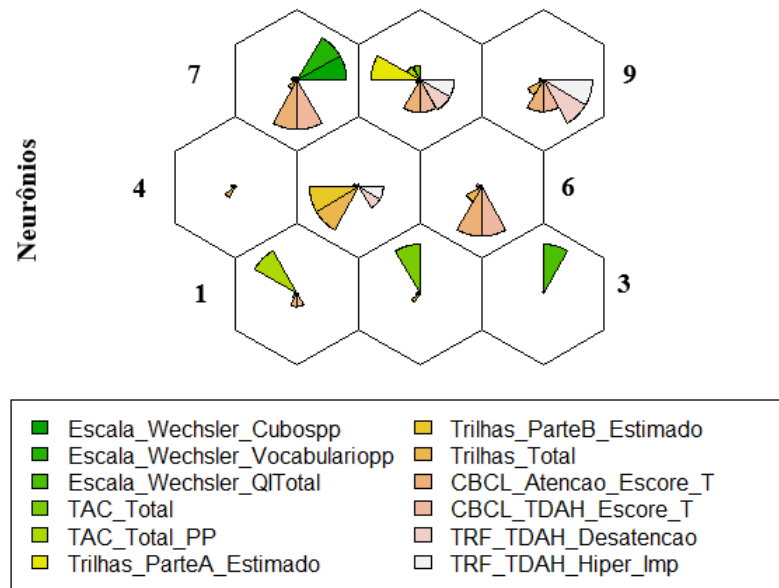
Ao processar o SOM e fazer a comparação do atributo de classe com os neurônios, o Quadro 7 mostra que o neurônio 4 tem a maior participação (55,9%) em relação aos demais e apresenta a maior sobreposição de classe, sendo a classe 1 (64,6%) e a classe 2 (41,7%) e entropia de 0,86.

Quadro 7: Comparativo diagnóstico pelo neurônio dimensão 3x3

	Diagnóstico	1	2	Total
Neurônio	1	3 (6,2%)	1 (1,3%)	4 (3,1%)
	2	2 (4,2%)	0 (0,0%)	2 (1,6%)
	3	1 (2,1%)	0 (0,0%)	1 (0,8%)
	4	20 (41,7%)	51 (64,6%)	71 (55,9%)
	5	0 (0,0%)	2 (2,5%)	2 (1,6%)
	6	0 (0,0%)	22 (27,8%)	22 (17,3%)
	7	0 (0,0%)	1 (1,3%)	1 (0,8%)
	8	2 (4,2%)	1 (1,3%)	3 (2,4%)
	9	20 (41,7%)	1 (1,3%)	21 (16,5%)
Total		48 (100,0%)	79 (100,0%)	127 (100,0%)

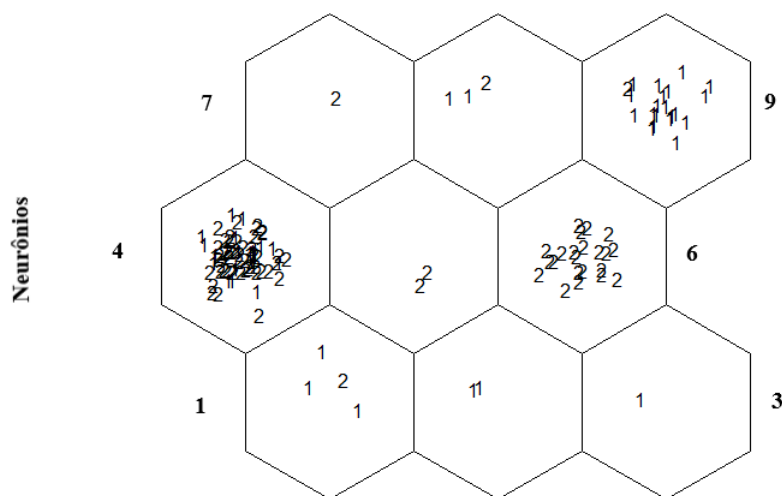
Ao analisar a Figura 16, tem-se o mapa topológico com a contribuição dos atributos para formação dos neurônios, porém no neurônio de maior sobreposição observa-se a contribuição do atributo Trilhas Total que é um instrumento que mede a avaliação das funções executivas e flexibilidade cognitiva.

Figura 16: Contribuição dos atributos na formação do neurônio



Analisando a Figura 17 o neurônio 4 tem a maior concentração de objetos e sobreposição de diagnóstico, mas o atributo que contribui mais para sua formação pode ser visto no Quadro 8, Quadro 9 e Quadro 10 e ao comparar os pesos desses atributos tem-se o atributo CBCL Atenção Escore T e CBCL TDAH Escore T como instrumentos comportamentais descritos pelos responsáveis da criança.

Figura 17: Dispersão dos objetos dentro dos neurônios



Quadro 8: Inércia dos atributos dimensão 3x3 - Parte 1

Neurônios	Escala Wechsler Cubospp	Escala Wechsler Vocabuláriopp	Escala Wechsler QI Total	TAC Total
Neurônio 1	-0,11	-0,12	-0,32	-0,37
Neurônio 2	-0,13	-0,09	-0,33	6,29
Neurônio 3	-0,08	-0,08	10,86	-0,15
Neurônio 4	-0,09	-0,10	-0,09	-0,11
Neurônio 5	-0,09	-0,08	-0,08	-0,02
Neurônio 6	-0,10	-0,11	0,04	-0,13
Neurônio 7	11,17	11,16	0,26	-0,12
Neurônio 8	-0,07	-0,09	-0,24	1,65
Neurônio 9	-0,05	-0,03	-0,16	-0,33

Quadro 9: Inércia dos atributos dimensão 3x3 - Parte 2

Neurônios	TAC Total PP	Trilhas Parte A Estimado	Trilhas Parte B Estimado	Trilhas Total
Neurônio 1	4,87	-0,57	-0,28	-0,35
Neurônio 2	-0,21	-0,43	0,04	-0,14
Neurônio 3	-0,37	-0,84	-0,75	-0,60
Neurônio 4	-0,19	-0,12	-0,15	-0,04
Neurônio 5	-0,24	-0,08	7,64	1,63
Neurônio 6	-0,17	-0,09	-0,08	0,18
Neurônio 7	0,06	-0,01	0,01	-0,14
Neurônio 8	1,28	6,24	-0,02	-0,41
Neurônio 9	-0,30	-0,24	-0,05	0,18

Quadro 10: Inércia dos atributos dimensão 3x3 - Parte 3

Neurônios	CBCL Atenção Escore T	CBCL TDAH Escore T	TRF TDAH Desatenção	TRF TDAH Hiper Imp
Neurônio 1	-0,09	-0,11	-0,48	-0,47
Neurônio 2	-0,65	-0,64	-0,46	-0,48
Neurônio 3	-0,68	-0,67	-0,47	-0,49
Neurônio 4	-0,67	-0,66	-0,48	-0,48
Neurônio 5	-0,67	-0,67	0,79	0,79
Neurônio 6	1,50	1,50	-0,48	-0,48
Neurônio 7	1,50	1,50	-0,49	-0,49
Neurônio 8	0,74	0,74	1,28	1,28
Neurônio 9	0,70	0,70	2,06	2,06

A discriminação do diagnóstico pelos atributos dentro do neurônio 4 pode ser visto na Figura 18 e como analisado na contribuição dos atributos para formação do neurônio tem-se o atributo CBCL com maior poder discriminante em relação ao diagnóstico, o que corrobora com o resultado do SOM. Entretanto, no SOM não se tem um parâmetro para o ponto de corte do diagnóstico de prevalência de TDAH, porém com o algoritmo da árvore de decisão, Figura 19, temos uma estimativa de corte possível para escala CBCL Atenção

maior que 58 Escore T para 27% do conjunto de dados do neurônio 4.

Figura 18: Importância dos atributos no neurônio 4 pela árvore de decisão

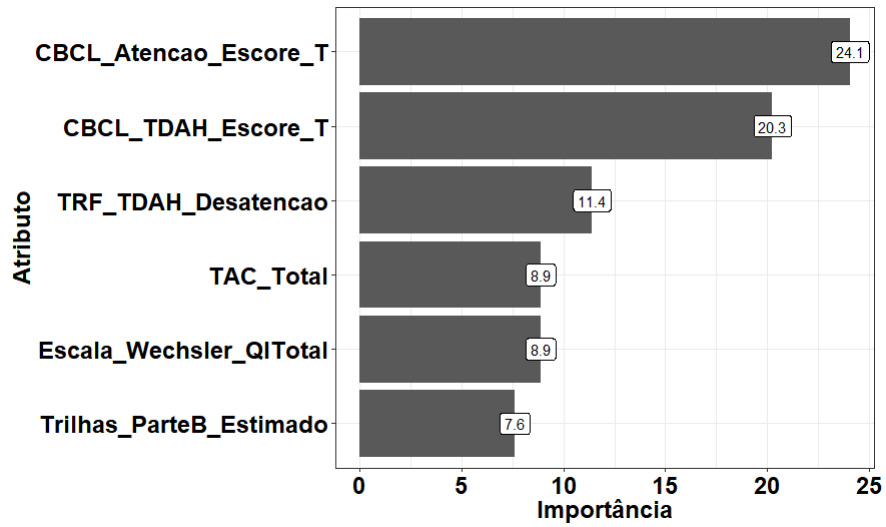
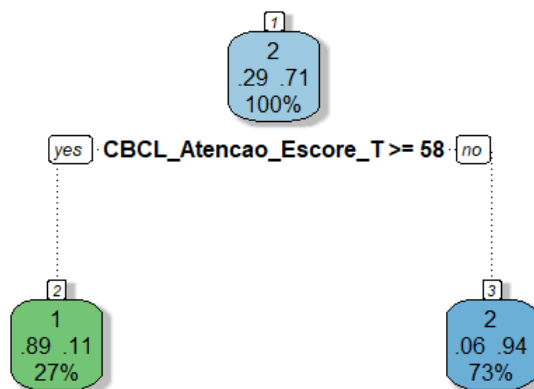


Figura 19: Árvore de decisão do neurônio 4



A árvore de decisão por ser um algoritmo supervisionado tem como atributo resposta uma estimativa do atributo de classe e o resultado pode ser visto no Quadro 11. Com

base na matriz de confusão gerada, observa-se que dos 51 diagnósticos positivos para TDAH coletados com aplicação dos inventários, tem-se 49 diagnósticos positivos estimados pelo algoritmo. Com base nos números apresentados, pode-se calcular a acurácia do algoritmo, assim como utilizar a estatística Kappa para medir a concordância entre os objetos observados e os objetos estimados (KRAEMER, 2014). O resultados com a grade 3x3 gerou uma acurácia de 93% com uma confiabilidade boa do Kappa de 82%, sendo o p-valor igual a $3,88e^{-12}$.

Quadro 11: Matriz de Confusão

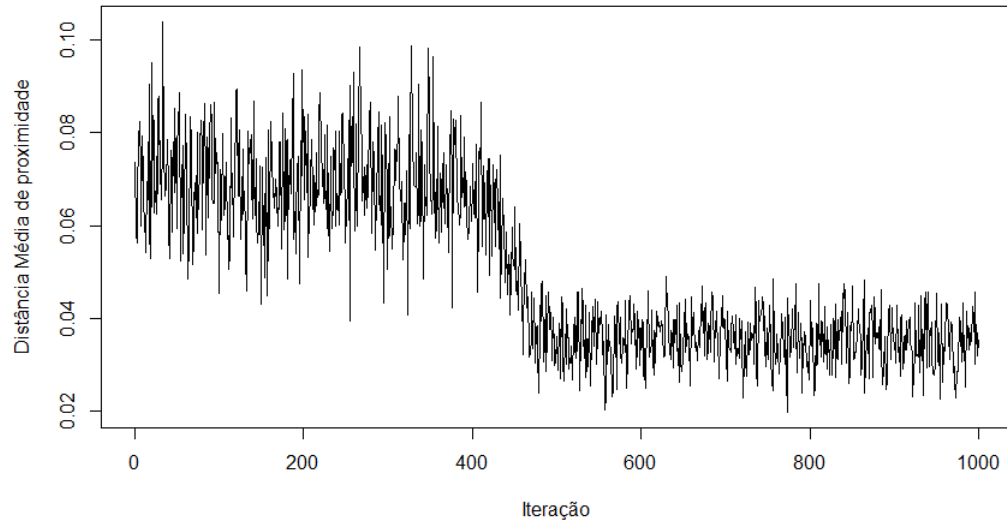
	Predito Negativo	Predito Positivo
Atual Negativo	17	3
Atual Positivo	2	49

O estudo apresentou resultado semelhante ao analisar o mapa do SOM com grade 2x2 e 3x3, sendo que resta analisar o mapa topológico do SOM com grade 4x4 para confirmação dos resultados encontrados.

5.1.3 Resultados com grade 4x4

A visualização da Figura 20 tem-se para 1.000 iterações uma melhora nítida no ajuste a partir da iteração 400 até 500 com redução significativa da distância média de aproximação de 0,07 para aproximadamente 0,03. Desta forma, pode-se dizer que o ajuste é a adequação da vizinhança dos neurônios na grade 4x4.

Figura 20: Visualização das 1.000 iterações aplicadas no SOM



No processamento do SOM tem-se a distribuição de frequência dos objetos em cada neurônio por diagnóstico, porém o Quadro 12 mostra que o neurônio 4 tem a maior participação (40,2%) em relação aos demais e apresenta a maior sobreposição de classe, sendo a classe 1 que representa a prevalência de TDAH (33,3%) e a classe 2 que representa ausência de TDAH (44,3%) e entropia de 0,90.

Quadro 12: Comparativo diagnóstico pelo neurônio dimensão 4x4

	Diagnóstico	1	2	Total
Neurônio	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%)

Devido ao baixo número de objetos em cada neurônio, não se pode tirar grandes conclusões ao visualizar o mapa com grade 4x4, pois o gráfico de setor dentro de cada neurônio acaba refletindo o diagnóstico de um objeto, elevando o peso de contribuição do atributo e acaba influenciando no tamanho do gráfico do neurônio conforme visto na Figura 21. Entretanto, ao verificar o neurônio com sobreposição de diagnóstico na Figura 22 e com maior representatividade no conjunto de dados, fica claro a confusão de diagnóstico dentro do neurônio que poderia ser interpretado como um cenário duvidoso por um especialista.

Figura 21: Contribuição dos atributos na formação do neurônio

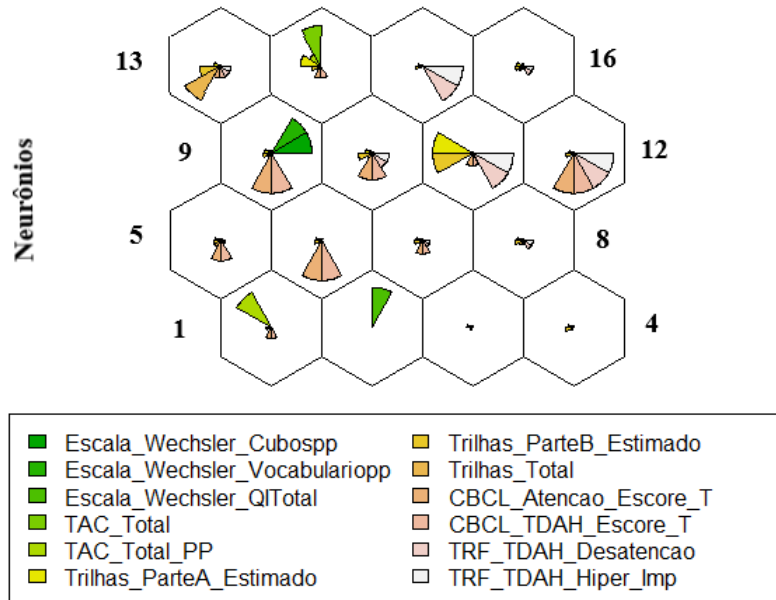
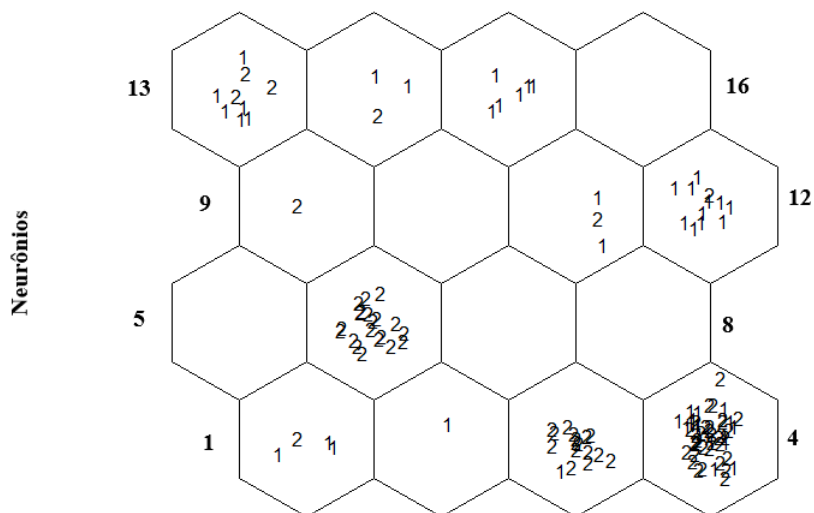


Figura 22: Dispersão dos objetos dentro dos neurônios



No Quadro 13, Quadro 14 e Quadro 15 tem-se os pesos dos atributos para formação do neurônio e para neurônios com um objeto, o pesos elevados é representado por um único atributo. Entretanto, ao verificar o neurônio 4 entre os 16 neurônios da grade, observa-se a presença de confusão de diagnóstico que pode ser confirmado com o resultado da entropia Shannon de 0,9. Esse valor confirma a impureza dos dados e garante a sobreposição de diagnóstico no neurônio. Contudo, tem-se a confirmação de que o atributo CBCL Atenção Escore T e CBCL TDAH Escore T são os atributos com maior contribuição para formação do neurônio 4, mas não se tem o ponto de corte para uma possível definição da prevalência de TDAH. Desta forma, há necessidade de verificar pelo algoritmo de árvore de decisão quais atributos tem maior poder de discriminação do transtorno Figura 23.

Quadro 13: Inércia dos atributos dimensão 4x4 - Parte 1

Neurônios	Escala Wechsler Cubospp	Escala Wechsler Vocabuláriopp	Escala Wechsler QI Total	TAC Total
Neurônio 1	-0,11	-0,12	-0,32	-0,37
Neurônio 2	0,08	-0,08	10,86	-0,15
Neurônio 3	-0,09	-0,11	0,10	-0,05
Neurônio 4	-0,10	-0,09	-0,14	-0,12
Neurônio 5	0,13	0,14	-0,05	-0,11
Neurônio 6	-0,10	-0,11	0,03	-0,10
Neurônio 7	0,07	0,07	-0,03	-0,11
Neurônio 8	-0,09	-0,09	-0,08	-0,11
Neurônio 9	11,17	11,16	0,26	-0,12
Neurônio 10	0,17	0,17	-0,07	-0,12
Neurônio 11	-0,08	-0,10	-0,21	-0,28
Neurônio 12	-0,03	-0,03	-0,23	-0,39
Neurônio 13	-0,09	-0,06	-0,06	-0,28
Neurônio 14	-0,11	-0,08	-0,25	6,29
Neurônio 15	-0,09	-0,03	-0,07	-0,17
Neurônio 16	-0,09	-0,09	-0,08	-0,11

Quadro 14: Inércia dos atributos dimensão 4x4 - Parte 2

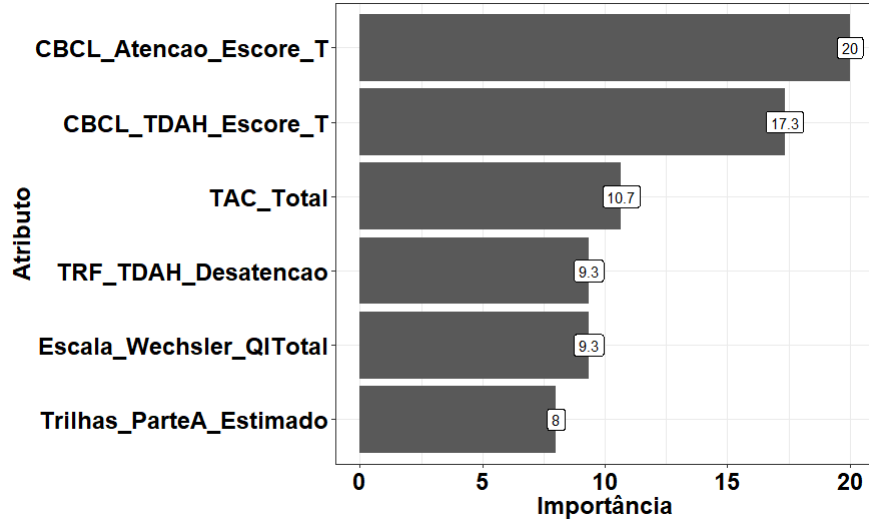
Neurônios	TAC Total PP	Trilhas Parte A Estimado	Trilhas Parte B Estimado	Trilhas Total
Neurônio 1	4,87	-0,56	-0,28	-0,35
Neurônio 2	-0,37	-0,84	-0,75	-0,60
Neurônio 3	-0,20	-0,09	-0,55	-0,25
Neurônio 4	-0,18	-0,12	-0,03	-0,27
Neurônio 5	-0,18	-0,06	-0,07	0,22
Neurônio 6	-0,18	-0,08	-0,08	-0,22
Neurônio 7	-0,18	-0,08	-0,03	-0,27
Neurônio 8	-0,19	-0,08	-0,03	-0,06
Neurônio 9	0,06	-0,01	0,01	-0,14
Neurônio 10	0,08	0,01	0,46	-0,02
Neurônio 11	-0,30	3,96	2,74	-0,41
Neurônio 12	-0,30	-0,26	-0,07	-0,37
Neurônio 13	-0,24	-0,12	1,02	3,58
Neurônio 14	1,30	1,55	0,03	-0,19
Neurônio 15	-0,32	-0,29	-0,10	-0,36
Neurônio 16	-0,19	-0,08	-0,03	-0,02

Quadro 15: Inércia dos atributos dimensão 4x4 - Parte 3

Neurônios	CBCL Atenção Escore T	CBCL TDAH Escore T	TRF TDAH Desatenção	TRF TDAH Hiper Imp
Neurônio 1	-0,08	-0,09	-0,48	-0,47
Neurônio 2	-0,68	-0,67	-0,47	-0,49
Neurônio 3	-0,67	-0,67	-0,48	-0,48
Neurônio 4	-0,67	-0,66	-0,48	-0,48
Neurônio 5	0,43	0,43	-0,24	-0,24
Neurônio 6	1,50	1,50	-0,48	-0,49
Neurônio 7	0,06	0,06	0,00	0,00
Neurônio 8	-0,46	-0,47	0,14	0,14
Neurônio 9	1,50	1,50	-0,49	-0,49
Neurônio 10	0,77	0,77	0,58	0,58
Neurônio 11	0,01	0,01	2,06	2,06
Neurônio 12	1,50	1,50	2,06	2,06
Neurônio 13	-0,02	-0,01	0,24	0,24
Neurônio 14	-0,01	-0,01	-0,47	-0,48
Neurônio 15	-0,64	-0,64	2,06	2,06
Neurônio 16	-0,47	-0,47	0,13	0,13

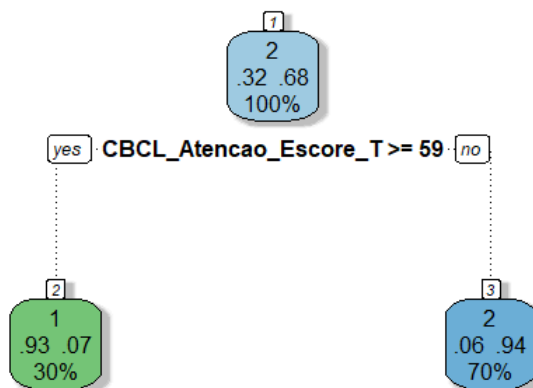
A importância do atributo, Figura 23, é uma medida referente ao ganho de informação da árvore de decisão e quanto maior for o valor do atributo, maior é a discriminação em relação ao atributo de classe que neste caso é o diagnóstico do transtorno. Desta forma, tem-se a concretização do atributo CBCL Atenção Escore T e CBCL TDAH Escore T com sendo atributos com possível utilização nos casos de dúvida de diagnóstico do TDAH.

Figura 23: Importância dos atributos no neurônio 4 pela árvore de decisão



O algoritmo de árvore de decisão mostra-se coerência nas três dimensões de grade apresentada e na grade 4x4 o ponto de corte para prevalência do TDAH são definidos para os objetos que apresentam Escore T superiores a 59 na escala CBCL Atenção.

Figura 24: Árvore de decisão do neurônio 4



Quadro 16: Matriz de Confusão

	Predito Negativo	Predito Positivo
Atual Negativo	12	4
Atual Positivo	2	33

A árvore de decisão por ser um algoritmo supervisionado tem como atributo resposta uma estativa do atributo de classe e o resultado pode ser visto no Quadro 16. Com base na matriz de confusão gerada, observa-se que dos 35 diagnósticos positivos para TDAH coletados com aplicação dos inventários, tem-se 33 diagnósticos positivos estimados pelo algoritmo. Com base nos números apresentados, pode-se calcular a acurácia do algoritmo, assim como utilizar a estatística Kappa para medir a concordância entre os objetos observados e os objetos estimados (KRAEMER, 2014). O resultado com a grade 4x4 gerou uma acurácia de 88% com uma confiabilidade boa do Kappa de 72%, sendo o p-valor igual a $2,68e^{-7}$.

Os resultados apresentados pelo SOM e árvore de decisão corroboraram nas três dimensões analisadas e trazem uma nova perspectiva para casos de dúvidas de diagnóstico de TDAH. Entretanto, no próximo capítulo serão discutidos como o estudo pôde ajudar através da compreensão matemática a interpretação do transtorno do neurodesenvolvimento e a possibilidade de futuros trabalhos utilizarem novas abordagem computacionais supervisionada ou/e não supervisionada para melhorar em grupos com sobreposição de diagnóstico do TDAH.

6 CONCLUSÃO E TRABALHOS FUTUROS

Os dados da área comportamental, por serem baseados no relato de informantes, geralmente podem ser mais susceptíveis ao viés de quem responde ao inventário (CARREIRO et al., 2014). Esse viés é menor quando se utilizam testes cognitivos que são medidas de avaliação aplicadas diretamente na criança. É provável que a compreensão matemática e geração de modelos se torne mais difícil utilizando apenas os inventários comportamentais. Como o TDAH demanda o uso de ambos os tipos de medidas, neste estudo ambas as ferramentas foram utilizadas para a aplicação da árvore de decisão. No estudo conseguiu-se agrupar as crianças com e sem TDAH pelo SOM, o que possibilitou o entendimento sob o olhar de cada agrupamento, que foram mais importantes na sua formação.

O mapa auto-organizável contribuiu especialmente no que se refere a formação dos grupos e no entendimento dos agrupamentos com sobreposição de classe. Neste caso de nebulosidade na definição de um diagnóstico, foi usado a árvore de decisão para classificar os atributos que contribuíram para formação do grupo com TDAH. Com isso, observou-se a predominância de características que ajudaram no entendimento do TDAH em crianças e adolescentes no estudo. A aplicação da árvore de decisão conseguiu fazer a discriminação de seis atributos, sendo dois de avaliação cognitiva e quatro de avaliação comportamental, ambos apresentaram discriminação relevante para fazer o diagnóstico, sendo o Inventário de Comportamentos para Crianças e Adolescentes entre 06 e 18 anos o que mostrou o maior poder discriminativo. Todavia, a incidência de Escores T baixos nas escala de atenção de problema e na escala de déficit de atenção não necessariamente implica que a criança tem TDAH. Os resultados apresentados mostraram a dificuldade e complexidade de encontrar indicadores que definam o TDAH, como já sinalizados por alguns autores (CARREIRO et al., 2014; LEE et al., 2010; CHANDANA; VIJAYALAKSHMI, 2018; CIBRIAN; HAYES; LAKES, 2020).

Importante ressaltar que o diagnóstico do TDAH é diagnóstico clínico e que leva em consideração a aferição de correlatos comportamentais dos déficits atencionais e indicadores de hiperatividade e impulsividade em mais de um ambiente. Sendo o atributo de CBCL/6-18 uma medida relatada pelos pais é provável que se confirme a validade dessas duas escalas para identificação do TDAH. Entretanto, ao desprezar, deve-se considerar

ainda as avaliações feitas com os testes cognitivos que fazem medições cognitivas de forma direta e são essenciais para compor o diagnóstico do TDAH. Neste estudo os testes de maior contribuição no ganho de informação da árvore de decisão foram o CBCL Atenção e o CBCL TDAH.

O Estudo apresentou como fator relevante o caso de sobreposição de diagnósticos dos neurônios ao utilizar o SOM e de forma conjunta com a árvore de decisão conseguiu separar 88% dos casos na grade 4x4 e 93% na grade 3x3. Desta maneira, futuros trabalhos podem colaborar com a técnica abordada no estudo através de outros procedimentos supervisionados e/ou não supervisionados de dados. Essas ferramentas podem auxiliar na realização de comparações entre resultados dos testes padronizados objetivando a redução de possíveis vieses das avaliações comportamentais baseadas em relato de informantes. Os estudos podem testar o algoritmo da árvore de decisão em amostras maiores ou nos conjuntos de neurônios que não foram observados sobreposição para verificar se os atributos mantêm a elevada acurácia e corroboram com os instrumentos de avaliação cognitiva e avaliação comportamental testados no estudo com foco na discriminação do diagnóstico do TDAH. Como este estudo controlou o grupo de TDAH sem a prevalência de comorbidades, recomenda-se para futuros estudos utilizar grupos amostrais com TDAH com e sem comorbidades de outras condições psiquiátricas e de neurodesenvolvimento. Esse tipo de amostra poderá oportunizar a testagem de novos modelos mais complexos devido a sobreposição natural de sinais e sintomas entre TDAH e algumas dessas comorbidades.

REFERÊNCIAS BIBLIOGRÁFICAS

- ACHENBACH, T. M. International findings with the achenbach system of empirically based assessment (aseba): applications to clinical services, research, and training. *Child and adolescent psychiatry and mental health*, Springer, v. 13, n. 1, p. 30, 2019.
- AGGARWAL, C. C. et al. *Neural networks and deep learning*. [S.l.]: Springer, 2018.
- ARBIB, M. A. A historical perspective. In: *Brains, Machines, and Mathematics*. [S.l.]: Springer, 1987. p. 1–14.
- ARETOULI, E. How neuropsychology can inform our understanding of preschool adhd: Clinical and research implications. *Applied Neuropsychology: Child*, Taylor & Francis, v. 8, n. 2, p. 174–181, 2019.
- ASSOCIATION, A. P. et al. *DSM-5: Manual diagnóstico e estatístico de transtornos mentais*. [S.l.]: Artmed Editora, 2014.
- BARKLEY, R. A.; MURPHY, K. R. *Transtorno de déficit de atenção/hiperatividade: exercícios clínicos*. 3. ed. São Paulo: Ed. Porto Alegre: Artmed, 2008.
- BENCZIK, E. B. P.; CASELLA, E. B. Compreendendo o impacto do tdah na dinâmica familiar e as possibilidades de intervenção. *Revista Psicopedagogia*, Associação Brasileira de Psicopedagogia, v. 32, n. 97, p. 93–103, 2015.
- CARREIRO, L. R. R. et al. Avaliação neuropsicológica no tdah: contribuições para identificação de dificuldades cognitivas e orientação escolar. In: AMATO, C. A. d. I. H. et al. (Ed.). *DISTÚRBIOS DO DESENVOLVIMENTO: Estudos Interdisciplinares*. 1. ed. São Paulo: Memnon, 2018. v. 1, cap. 10, p. 142–153. ISBN 978-85-7954-146-9.
- CARREIRO, L. R. R. et al. Protocolo interdisciplinar de avaliação neuropsicológica, comportamental e clínica para crianças e adolescentes com queixas de desatenção e hiperatividade. *Psicologia: teoria e prática*, Universidade Presbiteriana Mackenzie, v. 16, n. 3, p. 155–171, 2014.

- CASTRO, L. N. D.; FERRARI, D. G. *Introdução a mineração de dados*. [S.l.]: Saraiva Educação SA, 2016. 344–360 p.
- CAWLEY, G. C.; TALBOT, N. L. Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. *Pattern Recognition*, Elsevier, v. 36, n. 11, p. 2585–2592, 2003.
- CHANDANA, S.; VIJAYALAKSHMI, K. An approach to measure and improve the cognitive capability of adhd affected children through eeg signals. In: IEEE. *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*. [S.l.], 2018. p. 314–318.
- CHERKASOVA, M. et al. Developmental course of attention deficit hyperactivity disorder and its predictors. *Journal of the Canadian Academy of Child & Adolescent Psychiatry*, v. 22, n. 1, 2013.
- CIABURRO, G.; VENKATESWARAN, B. *Neural Networks with R: Smart models using CNN, RNN, deep learning, and artificial intelligence principles*. [S.l.]: Packt Publishing Ltd, 2017.
- CIBRIAN, F. L.; HAYES, G. R.; LAKES, K. D. Research advances in adhd and technology. *Synthesis Lectures on Assistive, Rehabilitative. and Health-Preserving Technologies*, Morgan & Claypool Publishers, v. 9, n. 3, p. i–156, 2020.
- COLIN, A. *Building decision trees with the ID3 algorithm*. [S.l.]: MILLER FREEMAN, INC 600 HARRISON ST., SAN FRANCISCO, CA 94107 USA, 1996.
- CRAN. *R 4.0.3*. out. 2020. Disponível em: <[https://cran.r-project.org/bin/windows-base/README.R-4.0.3](https://cran.r-project.org/bin/windows/base/README.R-4.0.3)>.
- CURTIS, A. E. et al. The mystery of the z-score. *AORTA Journal*, Thieme Medical Publishers, v. 4, n. 4, p. 124, 2016.
- DALSGAARD, S. et al. Common danish standards in prescribing medication for children and adolescents with adhd. *European child & adolescent psychiatry*, Springer, v. 23, n. 9, p. 841–844, 2014.
- FUENTES, D. et al. *Neuropsicologia-: Teoria e Prática*. [S.l.]: Artmed Editora, 2014.
- HAYKIN, S. *Redes neurais: princípios e prática*. [S.l.]: Bookman Editora, 2001.

- KAHNEMAN, D. et al. *Judgment under uncertainty: Heuristics and biases*. [S.l.]: Cambridge university press, 1982.
- KAM, H. et al. Development of a decision support model for screening attention-deficit hyperactivity disorder with actigraph-based measurements of classroom activity. *Applied clinical informatics*, Thieme Medical Publishers, v. 1, n. 4, p. 377, 2010.
- KASKI, S. Data exploration using self-organizing maps. In: CITeseer. *Acta polytechnica scandinavica: Mathematics, computing and management in engineering series no. 82*. [S.l.], 1997. p. 21.
- KOHONEN, T. Self-organized formation of topologically correct feature maps. *Biological cybernetics*, Springer, v. 43, n. 1, p. 59–69, 1982.
- KOHONEN, T. The self-organizing map. *Proceedings of the IEEE*, IEEE, v. 78, n. 9, p. 1464–1480, 1990.
- KOHONEN, T. Essentials of the self-organizing map. *Neural networks*, Elsevier, v. 37, p. 52–65, 2013.
- KOUMOULA, A. The course of attention deficit hyperactivity disorder (adhd) over the life span. *Psychiatrike= Psychiatriki*, v. 23, p. 49–59, 2012.
- KRAEMER, H. C. Kappa coefficient. *Wiley StatsRef: Statistics Reference Online*, Wiley Online Library, p. 1–4, 2014.
- KUHN, M. et al. Package ‘caret’. *The R Journal*, 2020.
- LEAHY, L. G. Attention-deficit/hyperactivity disorder: A historical review (1775 to present). *Journal of psychosocial nursing and mental health services*, SLACK Incorporated Thorofare, NJ, v. 55, n. 9, p. 10–16, 2017.
- LEE, S. H. et al. Analysis of attention deficit hyperactivity disorder in eeg using wavelet transform and self organizing maps. In: IEEE. *ICCAS 2010*. [S.l.], 2010. p. 2439–2442.
- LINOFF, G. S.; BERRY, M. J. *Data mining techniques: for marketing, sales, and customer relationship management*. [S.l.]: John Wiley & Sons, 2011.

- MAHONE, E. M.; DENCKLA, M. B. Attention-deficit/hyperactivity disorder: a historical neuropsychological perspective. *Journal of the International Neuropsychological Society: JINS*, NIH Public Access, v. 23, n. 9-10, p. 916, 2017.
- MERÉNYI, E.; MENDENHALL, M. J.; O'DRISCOLL, P. Advances in self-organizing maps and learning vector quantization. *Advances in Intelligent Systems & Computing*, Springer, v. 295, 2016.
- MITCHEL, T. Machine learning. mcgraw hill education (ise editions). 1997.
- POLANCZYK, G. et al. The worldwide prevalence of adhd: a systematic review and metaregression analysis. *American journal of psychiatry*, Am Psychiatric Assoc, v. 164, n. 6, p. 942–948, 2007.
- PURPER-OUAKIL, D. et al. Neurobiology of attention deficit/hyperactivity disorder. *Pediatric research*, Nature Publishing Group, v. 69, n. 8, p. 69–76, 2011.
- RAI, A. Explainable ai: From black box to glass box. *Journal of the Academy of Marketing Science*, Springer, v. 48, n. 1, p. 137–141, 2020.
- RAUBER, A.; MERKL, D.; DITTENBACH, M. The growing hierarchical self-organizing map: exploratory analysis of high-dimensional data. *IEEE Transactions on Neural Networks*, IEEE, v. 13, n. 6, p. 1331–1341, 2002.
- RITTER, H. et al. *Neural computation and self-organizing maps: an introduction*. [S.l.]: Addison-Wesley Reading, MA, 1992.
- RUBBO, M.; SILVA, L. A. Filtering-based instance selection method for overlapping problem in imbalanced datasets. *J*, v. 4, n. 3, p. 308–327, 2021. ISSN 2571-8800. Disponível em: <<https://www.mdpi.com/2571-8800/4/3/24>>.
- SEIFFERT, U. *Self-organizing neural networks: recent advances and applications*. [S.l.]: Physica, 2013.
- SILVA, L.; PERES, S.; BOSCARIOLI, C. *Introdução a Mineração de Dados com aplicações em R*. [S.l.: s.n.], 2017. ISBN 9788535284478.
- SILVA, M. M. M. Aplicação do modelo de árvore de decisão no rastreamento de queixas de desatenção e hiperatividade em crianças e adolescentes. São Paulo, 2020.

- SILVA, R.; PAES, Â. Por dentro da estatística: teste de concordância de kappa. *publication/uuid/3E5F4C37-E639-43D6-89C9-96597CA6AB40*, Educ Contin Saúde Einstein, v. 10, n. 4, p. 165–166, 2012.
- SIMON, V. et al. Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis. *The British Journal of Psychiatry*, Cambridge University Press, v. 194, n. 3, p. 204–211, 2009.
- THERNEAU, T. et al. Package ‘rpart’. Available online: *cran.ma.ic.ac.uk/web/packages/rpart/rpart.pdf (accessed on 20 April 2016)*, 2015.
- VESANTO, J.; ALHONIEMI, E. Clustering of the self-organizing map. *IEEE Transactions on neural networks*, Ieee, v. 11, n. 3, p. 586–600, 2000.
- VILLE, B. D.; NEVILLE, P. *Decision Trees for Analytics: Using SAS Enterprise Miner*. [S.l.]: SAS Institute Cary, NC, 2013.
- WEHRENS, R.; BUYDENS, L. kohonen: Supervised and unsupervised self-organising maps. *R package version*, v. 2, n. 2, 2007.
- WILLCUTT, E. G. et al. Validity of dsm-iv attention deficit/hyperactivity disorder symptom dimensions and subtypes. *Journal of abnormal psychology*, American Psychological Association, v. 121, n. 4, p. 991, 2012.
- WILLIAMS, N. M. et al. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. *The Lancet*, Elsevier, v. 376, n. 9750, p. 1401–1408, 2010.
- WOLRAICH, M. L. et al. Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics*, Am Acad Pediatrics, v. 144, n. 4, 2019.
- ZUMEL, N.; MOUNT, J.; PORZAK, J. *Practical data science with R*. [S.l.]: Manning Shelter Island, NY, 2014.