LVI SBPO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
DE 04 ATÉ 07 NOVEMBRO 2024
FORTALEZA | CEARÁ

UM NOVO PARADIGMA PARA O DIAGNÓSTICO DE AUTISMO: INTEGRANDO APRENDIZADO DE MÁQUINA E ANÁLISE VERBAL DE DECISÃO

Evandro Carvalho de Andrade, Luana I. Pinheiro
Universidade Estadual do Ceará - UECE - Av. Dr. Silas Munguba, 1700 - Fortaleza-CE Brasil
evandro.andrade@gmail.com, luana.ibiapina@aluno.uece.br

Plácido R. Pinheiro, Pedro G. C. D. Pinheiro, Mirian Caliope D. Pinheiro, Luciano C. Nunes
Universidade de Fortaleza - UNIFOR - Av. Washington Soares, 1321 - Fortaleza-CE Brasil
placido@unifor.br, pedrogcdpinheiro@edu.unifor.br,
miriancaliope@gmail.com, lcominn@uni7.edu.br

RESUMO

O Transtorno do Espectro Autista (TEA) afeta milhões de pessoas globalmente, sendo o diagnóstico precoce essencial para intervenções eficazes. No entanto, o processo de diagnóstico tradicional é demorado e complexo. Este estudo propõe um modelo híbrido inovador para otimizar o diagnóstico de TEA em crianças de 0 a 5 anos, combinando Aprendizagem de Máquina (AM) com Análise Verbal de Decisão (AVD). O modelo utiliza o algoritmo Random Forest para identificar as características mais relevantes na Avaliação Médica do BPC/LOAS do INSS e o método ZAPROS-III-i para ordená-las por importância. Os resultados demonstram uma redução de 89% nas características a serem avaliadas, agilizando o diagnóstico sem perda significativa de precisão. O modelo híbrido, baseado em dados de milhares de crianças diagnosticadas com TEA, apresenta um novo paradigma para aprimorar e otimizar o diagnóstico do TEA.

PALAVRAS CHAVE. Transtorno do Espectro Autista. Aprendizagem de Máquina. Análise Verbal de Decisão. Diagnóstico Otimizado. BPC/LOAS.

Tópicos (Diagnóstico de Autismo. Aprendizado de Máquina. Análise Verbal de Decisão. Otimização de Diagnóstico. Avaliação Médica. Limitações e Trabalhos Futuros.)

ABSTRACT

Autism Spectrum Disorder (ASD) affects millions worldwide, making early diagnosis crucial for effective intervention. However, traditional diagnostic processes are lengthy and complex. This study proposes an innovative hybrid model to optimize ASD diagnosis in children aged 0-5 years by combining Machine Learning (ML) and Verbal Decision Analysis (VDA). The model utilizes the Random Forest algorithm to identify the most relevant features in the INSS's BPC/LOAS Medical Assessment and the ZAPROS-III-i method to rank them by importance. Results demonstrate an 89% reduction in features to be evaluated, streamlining diagnosis without significant loss of accuracy. Based on data from thousands of children diagnosed with ASD, the hybrid model offers a new paradigm to enhance and optimize ASD diagnosis.

KEYWORDS. Autism Spectrum Disorder. Machine Learning. Verbal Decision Analysis. Optimized Diagnosis. BPC/LOAS.

Paper topics (Autism Diagnosis, Machine Learning, Verbal Decision Analysis, Diagnosis Optimization, Medical Evaluation, Limitations and Future Work.)

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

1. Introdução

O Transtorno do Espectro Autista (TEA) caracteriza-se por desafios na comunicação e interação social, padrões de comportamento repetitivos e interesses restritos [Raj et al., 2024]. Afeta milhões de pessoas globalmente, com estimativas da Organização Mundial da Saúde (OMS) apontando para uma criança com autismo a cada 160 [Organization, 2019]. O diagnóstico precoce é fundamental para intervenções eficazes, impactando positivamente o desenvolvimento e qualidade de vida do indivíduo, como demonstrado em estudos de [Howlin et al., 2009] e [Reichow, 2011].

No entanto, o processo diagnóstico tradicional é demorado e complexo, exigindo avaliações comportamentais, questionários extensos e múltiplos critérios clínicos [Tamanini et al., 2009; de Castro et al., 2009; Araújo de Castro et al., 2008]. No Brasil, a Avaliação Médica do BPC/LOAS, utilizada pelo INSS para concessão de benefício a pessoas com TEA, exemplifica essa complexidade. Este artigo propõe um modelo híbrido inovador para otimizar o diagnóstico de TEA em crianças de 0 a 5 anos, utilizando:

- **Aprendizagem de Máquina (AM):** Para identificar as características mais relevantes da Avaliação Médica do BPC/LOAS para o diagnóstico de TEA;
- **Análise Verbal de Decisão (AVD):** Para ordenar as características por ordem de importância, criando um protocolo otimizado;

O objetivo é agilizar o processo de diagnóstico sem comprometer a precisão, beneficiando tanto os indivíduos com TEA quanto o fluxo gerencial do INSS.

2. Contexto e Problematização

A necessidade de um diagnóstico célere e preciso para TEA é crucial, especialmente em crianças, para que intervenções terapêuticas sejam iniciadas o mais cedo possível. No entanto, a identificação do TEA enfrenta desafios devido à variedade de sintomas e à necessidade de extensas comprovações clínicas, resultando em um processo demorado [Daley, 2004], [Howlin e Asgharian, 1999]. Mesmo em 2024, o diagnóstico do autismo continua sendo um desafio para a ciência médica, e o desenvolvimento de modelos eficazes de Inteligência Artificial (IA) ainda enfrenta obstáculos. A complexidade do TEA, com suas diversas manifestações e a heterogeneidade dos indivíduos dentro do espectro, exige abordagens multidisciplinares e ferramentas inovadoras para auxiliar no processo de identificação e diagnóstico [Bone et al., 2016] e [Thabtah, 2017]. A subjetividade inerente à avaliação de critérios diagnósticos, a dificuldade na coleta de dados [Waheed et al., 2020] padronizados e a necessidade de considerar o contexto individual de cada criança são alguns dos desafios enfrentados na aplicação da IA para o diagnóstico de autismo [Thabtah, 2018].

A Avaliação Médica do BPC/LOAS, utilizada pelo INSS para a concessão do benefício de prestação continuada a pessoas com deficiência, incluindo o TEA, é um processo meticuloso que contribui para longas filas de espera. Em 2019, mais de 1,6 milhão de benefícios aguardavam definições do INSS, dos quais mais de 420 mil eram de amparo assistencial ao portador de deficiência [INSS, 2019; de S. Paulo, 2020]. Infelizmente, o problema da fila de espera do INSS persiste em 2024, impactando o acesso à saúde e assistência social. Mesmo com medidas para acelerar a análise de pedidos, como o pagamento de bônus para servidores do INSS, a fila de espera para benefícios ainda ultrapassava 1 milhão de pessoas em 2023, conforme reportado pela Gazeta do Povo [do Povo, 2023]. Essa demora impacta diretamente na vida dos indivíduos com TEA e seus familiares, atrasando o acesso a tratamentos e suporte essenciais, como o Benefício de Prestação Continuada (BPC), crucial para garantir a subsistência de pessoas com deficiência em situação de vulnerabilidade social.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

A otimização do processo de diagnóstico do TEA, por meio da identificação das características mais relevantes, pode agilizar a concessão do benefício, reduzir as filas de espera do INSS e, principalmente, garantir o acesso rápido a tratamentos adequados para as crianças.

3. Fundamentação Teórica

3.1. Transtorno do Espectro Autista (TEA)

O Transtorno do Espectro Autista (TEA) é um distúrbio do neurodesenvolvimento que se manifesta em um amplo espectro de características e níveis de gravidade. De acordo com o Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-5) da Associação Americana de Psiquiatria [Association, 2013], o TEA é caracterizado por:

- **Déficits persistentes na comunicação social e na interação social**: Manifestados por dificuldades na reciprocidade socioemocional, nos comportamentos comunicativos não verbais utilizados para interação social e no desenvolvimento, manutenção e compreensão de relacionamentos;
- **Padrões restritos e repetitivos de comportamento, interesses ou atividades**: Manifestados por movimentos motores, uso de objetos ou fala estereotipados ou repetitivos; insistência na mesmice, adesão inflexível a rotinas ou padrões ritualizados de comportamento verbal ou não verbal; interesses altamente restritos e fixos que são anormais em intensidade ou foco; hiper ou hiporreatividade a estímulos sensoriais ou interesse incomum por aspectos sensoriais do ambiente;
- **Anormalidades na integração sensorial**: Podem afetar significativamente o comportamento, a aprendizagem e a participação social de indivíduos com TEA [Miller et al., 2007]. Elas podem se manifestar como:
  - **Hiperresponsividade**: Reações exageradas a estímulos sensoriais, como sons altos, toques leves ou texturas específicas;
  - **Hiporresponsividade**: Falta de resposta a estímulos sensoriais, como não perceber dor ou temperatura;
  - **Busca sensorial**: Procura por estímulos sensoriais específicos, como girar ou balançar;

É importante ressaltar que as manifestações do TEA variam amplamente de acordo com a gravidade dos sintomas, o nível de desenvolvimento e a idade cronológica do indivíduo. Os sintomas geralmente se tornam aparentes nos primeiros anos de vida, mas podem não ser totalmente reconhecidos até que as demandas sociais excedam as capacidades limitadas da criança [Organization, 2019]. O diagnóstico precoce e preciso é crucial para garantir o acesso a intervenções e suporte adequados, que podem melhorar significativamente o desenvolvimento e a qualidade de vida dos indivíduos com TEA [de Andrade et al., 2023]. A utilização de modelos híbridos, combinando Aprendizagem de Máquina e Análise Verbal de Decisão, tem se mostrado promissora para otimizar o processo de diagnóstico de TEA, como demonstrado em [Andrade, 2020]. No entanto, o processo de diagnóstico tradicionalmente enfrenta desafios devido à subjetividade na avaliação clínica, à falta de biomarcadores específicos e à heterogeneidade do espectro autista [Andrade et al., 2021].

Diversos protocolos e manuais padronizam as informações utilizadas na detecção e diagnóstico do TEA, buscando fornecer critérios objetivos e padronizados para a avaliação clínica. Entre os mais utilizados, destacam-se:

- **Manual Diagnóstico e Estatístico de Transtornos Mentais, 5ª Edição (DSM-5) [Association, 2013]** : Define critérios específicos para o diagnóstico;

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

- **Autism Diagnostic Observation Schedule - Revised (ADOS-R) [Lord et al., 2012]**: Instrumento que utiliza atividades padronizadas para avaliação da comunicação, interação social e comportamentos;

- **Autism Diagnostic Interview - Revised (ADI-R) [Lord et al., 2000]**: Entrevista com pais/cuidadores que coleta informações sobre o histórico do desenvolvimento da criança;

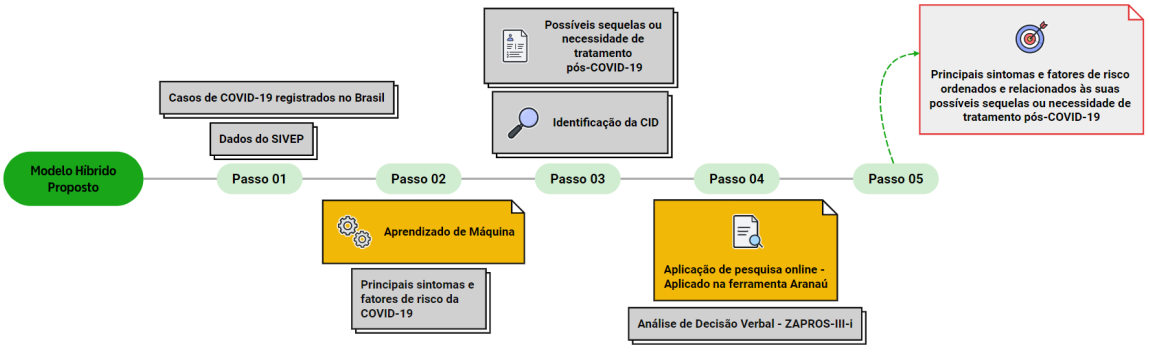
- **Classificação Internacional de Doenças, 10ª edição (CID-10) [Organization, 1993]**:Classificação utilizada para fins de estatística de saúde, que inclui o TEA na categoria "Transtornos Globais do Desenvolvimento";

A escolha do instrumento varia de acordo com o contexto clínico e as preferências do profissional.

3.2. Aprendizagem de Máquina (AM)

A Aprendizagem de Máquina (AM) [Mitchell, 1997] é um subcampo da Inteligência Artificial (IA) que permite que máquinas aprendam com dados, identifiquem padrões complexos e façam predições precisas. A AM tem sido aplicada em saúde para auxiliar no diagnóstico de doenças [Andrade et al., 2022; de Andrade et al., 2023], predição de riscos e personalização de tratamentos. A Figura 1 ilustra a estrutura de um modelo híbrido proposto para o diagnóstico do Transtorno do Espectro Autista (TEA).

Figura 1: Estrutura do Modelo Híbrido



Fonte: Adaptado de Andrade et al. [2023].

Os algoritmos de AM podem ser categorizados em diferentes tipos, dependendo da natureza dos dados de treinamento e do objetivo da tarefa, sendo as mais comuns:

- **Aprendizagem Supervisionada:** Utiliza dados rotulados para aprender uma função que mapeia características de entrada para saídas desejadas[Russell e Norvig, 2010];

- **Aprendizagem Não-supervisionada:** Utiliza dados não rotulados para descobrir padrões e estruturas nos dados [James et al., 2013];

A escolha do algoritmo depende da natureza do problema, dos dados disponíveis e do objetivo da tarefa.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

3.3. Análise Verbal de Decisão (AVD)

A Análise Verbal de Decisão (AVD) [Larichev e Moshkovich, 1997] é uma metodologia que utiliza linguagem natural para lidar com problemas de decisão complexos, caracterizados por informações imprecisas e subjetivas. Em problemas de diagnóstico médico, onde a incerteza e a subjetividade são frequentes, a AVD pode ser uma ferramenta valiosa para auxiliar na tomada de decisão clínica, como demonstrado em [Pinheiro et al., 2021].

Os métodos da AVD se baseiam na premissa de que as pessoas são mais capazes de expressar suas preferências e realizar comparações por meio de termos linguísticos do que por meio de números. Esses métodos utilizam procedimentos estruturados e questionários específicos para elicitar as preferências dos tomadores de decisão, minimizando a ambiguidade e a inconsistência inerentes à linguagem verbal. Diversos métodos compõem a estrutura da AVD, cada um com suas características e aplicações específicas. Entre os mais conhecidos, podemos citar:

- **ZAPROS-III [Larichev e Moshkovich, 1997]**: Visa a classificação de alternativas em classes de decisão predefinidas;
- **ZAPROS-LM [Larichev e Moshkovich, 1997]**: Visa a ordenação de alternativas, das mais preferíveis às menos preferíveis;
- **PACOM [Larichev e Moshkovich, 1997]**: Visa a ordenação de alternativas com base na comparação par a par de alternativas em relação a todos os critérios simultaneamente;
- **ORCLASS [Larichev e Moshkovich, 1997]**: Visa a classificação de alternativas em classes de decisão com base na avaliação verbal comparativa de alternativas em relação a cada critério separadamente.

O método ZAPROS-III-i [Tamanini e Pinheiro, 2008], utilizado neste estudo, é uma técnica de ordenação de alternativas que combina as vantagens do ZAPROS-III e ZAPROS-LM, resultando em maior eficiência e precisão. Ele utiliza uma única situação de referência para comparar as alternativas, o que permite ordenar um maior número de alternativas, incluindo aquelas consideradas incomparáveis por métodos anteriores.

4. Metodologia da Pesquisa

A pesquisa utilizou dados de 3.861 Avaliações Médicas do BPC/LOAS de crianças de 0 a 5 anos diagnosticadas com TEA. O modelo híbrido proposto foi desenvolvido em cinco etapas principais:

4.1. Pré-processamento dos Dados

As respostas das Avaliações Médicas foram organizadas em uma tabela (Analytical Base Table - ABT), com as características da avaliação médica e seus respectivos qualificadores de obstáculos, definidos com base na Classificação Internacional de Funcionalidade, Incapacidade e Saúde (CIF) [Organization, 2001; do Desenvolvimento Social e Combate à Fome e da Previdência Social, 2007]:

- **Leve/Moderada (de 0% a 49% de comprometimento):** Apresenta uma barreira ou dificuldade leve, regular;
- **Grave (de 50% a 95% de comprometimento):** Apresenta uma dificuldade grande, extrema;
- **Completa (de 96% a 100% de comprometimento):** Apresenta uma barreira ou dificuldade total;

Um aplicativo Java foi desenvolvido para a transposição dos dados da ABT, gerando um arquivo .csv para a aplicação dos algoritmos de AM.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

4.2. Aplicação de Algoritmos de Aprendizagem de Máquina

A ferramenta Orange foi utilizada para a aplicação dos algoritmos Random Forest (RF) e Logistic Regression (LR). O objetivo era classificar as características da ABT e identificar as mais relevantes para o diagnóstico de TEA.

4.2.1. Random Forest (RF)

O RF é um algoritmo de AM que combina múltiplas árvores de decisão para realizar predições. Cada árvore é treinada com uma amostra aleatória dos dados e um subconjunto aleatório das características. A predição final é obtida por meio da agregação das predições de todas as árvores [Breiman, 1996].

4.2.2. Logistic Regression (LR)

O LR é um algoritmo de AM utilizado para prever a probabilidade de um evento ocorrer. É um modelo linear generalizado que utiliza uma função logística para transformar a saída linear em uma probabilidade [Hosmer e Lemeshow, 1989].

4.3. Análise Comparativa dos Algoritmos

Diversas métricas de avaliação foram utilizadas para comparar o desempenho do RF e do LR, incluindo:

- **Acurácia de Classificação:** Proporção de predições corretas em relação ao total de predições;
- **Área Sob a Curva ROC:** Medida da capacidade do modelo de distinguir entre as classes;
- **Tempo de Treinamento:** Tempo necessário para treinar o modelo;
- **Precisão:** Proporção de verdadeiros positivos em relação ao total de positivos preditos;
- **Recall:** Proporção de verdadeiros positivos em relação ao total de positivos reais.
- **F1:** Média harmônica entre Precisão e Recall.

Os resultados da comparação entre o RF e o LR são apresentados na Tabela 1.

Tabela 1: Comparação entre Random Forest e Logistic Regression

Modelo	Acurácia da Classificação	Área Sob a Curva	Tempo de Treinamento	Precisão	Recall	F1
Random Forest	0.667	0.667	4.560s	0.544	0.667	0.580
Logistic Regression	0.656	0.526	5.907s	0.487	0.656	0.538

Fonte: Dados da pesquisa.

4.4. Ordenação das Características através da AVD

As nove características mais relevantes identificadas pelo RF foram submetidas ao método ZAPROS-III-i da AVD, implementado na ferramenta ARANAÚ. Um questionário online foi aplicado a 91 profissionais da área de saúde e educação (médicos, fisioterapeutas, enfermeiros, psicólogos, fonoaudiólogos, terapeutas ocupacionais e educadores) para elicitar as preferências e ranquear as características por ordem de importância. O questionário online abordou as nove características, solicitando que os profissionais avaliassem o grau de barreira ou dificuldade (Leve/Moderada, Grave, Completa) em três critérios:

- **Critério A:** As funções da Voz, Mental e/ou Visão devem apresentar uma barreira ou dificuldade.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

- **Critério B:** O Aprendizado e Aplicação do Conhecimento devem apresentar uma barreira ou dificuldade.
- **Critério C:** A mobilidade deve apresentar uma barreira ou dificuldade.

4.5. Construção do Protocolo Otimizado

Com base nos resultados do RF e do ZAPROS-III-i, foi elaborado um protocolo otimizado com as características mais relevantes e sua ordem de importância para o diagnóstico de TEA Andrade et al. [2023]. A Tabela 2 apresenta as nove características ordenadas por ordem de importância.

Tabela 2: Características Ordenadas por Importância

Ordem	Código	Descrição CIF
1	d465	Deslocar-se utilizando algum tipo de equipamento. Mover todo o corpo de um lugar para outro utilizando dispositivos específicos para facilitar a movimentação.
2	b144	Funções da memória. Funções mentais específicas de registro e armazenamento de informações e sua recuperação quando necessário.
3	b122	Funções psicossociais globais. Funções mentais gerais necessárias para compreender e integrar funções que levam à formação de capacidades interpessoais para interações sociais recíprocas.
4	b330	Funções da fluência e do ritmo da fala. Funções relacionadas com a produção do fluxo e do tempo da fala.
5	d155	Adquirir competências. Desenvolver as capacidades básicas e complexas necessárias para a execução de um conjunto integrado de ações ou tarefas.
6	d450	Andar. Mover-se de pé sobre uma superfície, passo a passo, de modo que um pé esteja sempre no chão.
7	d110	Observar. Utilizar intencionalmente o sentido da visão para captar estímulos visuais.
8	b210	Funções da visão. Funções sensoriais relacionadas com a percepção da presença de luz e a forma, tamanho, formato e cor do estímulo visual.
9	b134	Funções do sono. Funções mentais gerais de desconexão física e mental periódica, reversível e seletiva, do ambiente imediato da pessoa, acompanhada por mudanças fisiológicas características.

Fonte: Dados da pesquisa.

As respostas dos profissionais foram utilizadas pela ferramenta ARANAÚ para gerar a ordenação das características.

5. Resultados

A aplicação dos algoritmos de Aprendizagem de Máquina (AM), Random Forest (RF) e Logistic Regression (LR), utilizando a ferramenta Orange, permitiu a classificação das características da ABT e a identificação das mais relevantes para o diagnóstico de TEA. O algoritmo RF apresentou melhor desempenho geral em relação ao LR, com maior acurácia (0.667 vs 0.656) e área sob a curva ROC (0.667 vs 0.526), como demonstrado na Tabela 1. O RF também apresentou menor tempo de treinamento (4.560s vs 5.907s). A superioridade do RF em relação ao LR para a classificação de dados biomédicos também foi observada em outros estudos Couronné et al. [2018]; Tanwani et al. [2009].

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

A aplicação do modelo híbrido, combinando o RF com o método ZAPROS-III-i da AVD, resultou em uma redução significativa de 89% das características da Avaliação Médica do BPC/LOAS, de 83 para 9, como apresentado na Tabela 2. Essa redução drástica no número de características a serem avaliadas representa um avanço promissor para a otimização do processo de diagnóstico de TEA, tornando-o mais ágil e eficiente. A ordenação das características pelo método ZAPROS-III-i, com base nas preferências de 91 profissionais da área de saúde e educação, demonstrou que as funções relacionadas à mobilidade, memória e interação social foram consideradas as mais relevantes para o diagnóstico de TEA. Esse resultado reforça a importância de considerar a expertise humana na construção de modelos de apoio à decisão, combinando o poder preditivo da IA com a experiência clínica dos profissionais.

5.1. Comparação com a Literatura Existente

Os resultados obtidos neste estudo corroboram as descobertas de pesquisas recentes que demonstram o potencial da Aprendizagem de Máquina (AM) para otimizar o processo de diagnóstico do TEA. [Thabtah, 2021] propuseram um modelo baseado em regras de indução para a detecção de autismo, utilizando algoritmos de AM para identificar padrões em dados comportamentais e alcançar uma acurácia de 82% na classificação de indivíduos com TEA. [Eslami et al., 2022] investigaram a utilização de diferentes algoritmos de AM, incluindo Random Forest, para a predição do diagnóstico de TEA com base em dados de neuroimagem, alcançando uma acurácia de até 75%. Similarmente a estes estudos, a presente pesquisa demonstrou a eficácia do algoritmo Random Forest na identificação de características relevantes para o diagnóstico de TEA, corroborando seu potencial para aprimorar a acurácia e a eficiência do processo diagnóstico.

A redução de 89% das características a serem avaliadas no modelo proposto, alcançada através da combinação do Random Forest com o método ZAPROS-III-i da AVD, representa um avanço significativo em relação a outras abordagens de otimização do diagnóstico de TEA. [Sosa et al., 2024] desenvolveram um sistema de apoio à decisão clínica para o diagnóstico de TEA, utilizando algoritmos de AM para reduzir o número de características a serem avaliadas em 25%. No entanto, a presente pesquisa alcançou uma redução significativamente maior, demonstrando a efetividade da metodologia híbrida proposta. Essa redução drástica no número de características a serem avaliadas, sem comprometer a precisão do diagnóstico, tem o potencial de agilizar o processo diagnóstico, reduzir o tempo de espera por intervenções precoces e otimizar a utilização de recursos no sistema de saúde.

6. Discussão

O modelo híbrido proposto demonstra ser uma ferramenta eficaz para a otimização do diagnóstico de TEA em crianças. A combinação do algoritmo Random Forest, reconhecido por sua robustez e precisão em tarefas de classificação [Breiman, 1996], com o método ZAPROS-III-i da AVD, que permite a incorporação da expertise humana na ordenação de características, resultou em um protocolo otimizado com potencial para impactar positivamente a prática clínica.

A redução de 89% das características a serem avaliadas, de 83 para 9, representa uma economia significativa de tempo e recursos no processo de diagnóstico. Essa agilidade pode permitir que as crianças com TEA tenham acesso mais rápido a intervenções precoces, crucial para o desenvolvimento de suas habilidades e para a melhora da sua qualidade de vida. Estudos demonstram que a intervenção precoce e intensiva em crianças com TEA pode promover ganhos significativos em áreas como comunicação, interação social, comportamento adaptativo e cognição [Reichow, 2011; Warren et al., 2011]. Além de beneficiar as crianças e suas famílias, a otimização do processo de diagnóstico também pode contribuir para a redução das filas de espera do INSS para a concessão do BPC/LOAS. A simplificação da avaliação médica, com a redução do número de características a serem analisadas, pode agilizar o processo de análise dos pedidos e a tomada de decisão, beneficiando tanto os indivíduos com TEA quanto a gestão do sistema previdenciário.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

7. Conclusões

Este estudo demonstrou a viabilidade e a eficácia de um modelo híbrido para otimizar o diagnóstico de TEA em crianças, combinando a capacidade preditiva da Aprendizagem de Máquina com a expertise humana incorporada na Análise Verbal de Decisão. A pesquisa contribui para o avanço do conhecimento científico na área de diagnóstico de TEA e abre caminho para o desenvolvimento de novas ferramentas de apoio à decisão clínica. O modelo híbrido proposto, com sua capacidade de reduzir significativamente o número de características a serem avaliadas sem comprometer a precisão do diagnóstico, representa um novo paradigma para a identificação precoce do TEA, com potencial para impactar positivamente a vida das crianças, suas famílias e o sistema de saúde como um todo.

8. Limitações e Trabalhos Futuros

Embora os resultados da pesquisa sejam promissores, é importante reconhecer as limitações do estudo. A pesquisa se limitou a crianças de 0 a 5 anos e utilizou dados específicos da Avaliação Médica do BPC/LOAS do INSS. Trabalhos futuros podem:

- **Expandir o modelo para outras faixas etárias**: Investigar a aplicabilidade do modelo híbrido em crianças mais velhas, adolescentes e adultos com TEA.[Thabtah e Peebles, 2019];
- **Incorporar outras fontes de dados**: Integrar dados de prontuários eletrônicos, registros médicos, exames complementares e outras fontes de informação relevantes para o diagnóstico de TEA.[Thabtah et al., 2019];
- **Avaliar a aplicação prática do protocolo em diferentes contextos clínicos**: Testar a efetividade do protocolo otimizado em diferentes cenários clínicos, como centros de saúde, hospitais e clínicas especializadas.
- **Explorar técnicas de Deep Learning**: Investigar a aplicação de técnicas avançadas de Deep Learning, como Redes Neurais Convolucionais (CNN) e Transformers[Raj et al., 2024].

9. Agradecimentos

Os autores agradecem a DATAPREV, Universidade Estadual do Ceará (UECE) e Universidade de Fortaleza (UNIFOR) pelo apoio institucional e logístico. O terceiro autor agradece ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de produtividade em pesquisa (306435/2023-3).

Referências

Andrade, E., Pinheiro, P., Pinheiro, P., Nunes, L., e Pinheiro, L. (2023). Feature extraction and selection applied to bone radiographs in traumatological surgical procedures: A quasi-survey. In *Research and Innovation Forum 2023*, volume 356 of *Springer Proceedings in Complexity*, p. 513–526. Springer, Cham.

Andrade, E., Portela, S., Pinheiro, P., Nunes, L., Simão Filho, M., Costa, W., e Pinheiro, M. (2021). A protocol for the diagnosis of autism spectrum disorder structured in machine learning and verbal decision analysis. *Computational and Mathematical Methods in Medicine*, 2021:1628959.

Andrade, E. (2020). Um modelo híbrido em aprendizagem de máquina e análise verbal de decisão aplicado ao diagnóstico de autismo. Master's thesis, PPGIA - Universidade de Fortaleza.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

Andrade, E., Pinheiro, P., Barros, A., Nunes, L., Pinheiro, L., Pinheiro, P., e Holanda Filho, R. (2022). Towards machine learning algorithms in predicting the clinical evolution of patients diagnosed with covid-19. *Applied Sciences*, 12(18):8939.

Araújo de Castro, A., Pinheiro, P., e Pinheiro, M. (2008). A hybrid model for aiding in decision making for the neuropsychological diagnosis of alzheimer's disease. In *Rough Sets and Current Trends in Computing*, volume 5306 of *Lecture Notes in Computer Science*, p. 386–399. Springer, Berlin, Heidelberg.

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. American Psychiatric Publishing, Washington, DC.

Bone, D. et al. (2016). Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. *Journal of Child Psychology and Psychiatry*, 57(8):927–937.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24:123–140.

Couronné, R., Probst, P., e Boulesteix, A.-L. (2018). Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinformatics*, 19(1):270.

Daley, T. (2004). From symptom recognition to diagnosis: Children with autism in urban india. *Social Science Medicine*, 58:1323–1335.

de Andrade, E., Pinheiro, L., Pinheiro, P., et al. (2023). Hybrid model for early identification post-covid-19 sequelae. *Journal of Ambient Intelligence and Humanized Computing*, 14: 14739–14752.

de Castro, A., Pinheiro, P., e Pinheiro, M. (2009). An approach for the neuropsychological diagnosis of alzheimer's disease: A hybrid model in decision making. In *Rough Sets and Knowledge Technology*, volume 5589 of *Lecture Notes in Computer Science*, p. 279–291. Springer, Berlin, Heidelberg.

de S. Paulo, F. (2020). Inss reduz análise de pedidos, e fila de espera de benefícios volta a subir. *Folha de S. Paulo*. URL `https://www1.folha.uol.com.br`.

do Desenvolvimento Social e Combate à Fome, M. e da Previdência Social, M. (2007). *Avaliação de pessoas com deficiência para acesso ao Benefício de Prestação Continuada da Assistência Social*.

do Povo, G. (2023). Fila do inss ainda passa de 1 milhão de pessoas, mesmo com bônus para servidores. *Gazeta do Povo*. URL `https://www.gazetadopovo.com.br/republica/fila-inss-ainda-passa-1-milhao-pessoas-mesmo-com-bonus-servidores/`.

Eslami, T. et al. (2022). Machine learning-based autism spectrum disorder diagnosis using resting-state functional magnetic resonance imaging: A systematic review. *Frontiers in Psychiatry*, 13: 846767.

Hosmer, D. W. e Lemeshow, S. (1989). *Applied Logistic Regression*. Wiley.

Howlin, P. e Asgharian, A. (1999). The diagnosis of autism and asperger syndrome: Findings from a survey of 770 families. *Developmental Medicine and Child Neurology*, 41:834–839.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

Howlin, P., Magiati, I., e Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal of Intellectual Development Disabilities*, 37:23–41.

INSS (2019). Relatório de gestão do instituto. Technical report, Instituto Nacional do Seguro Social. URL https://www.inss.gov.br/wp-content/uploads/2020/06/INSS-RELAT%C3%93RIO-DE-GEST%C3%83O-2019.pdf.

James, G., Witten, D., Hastie, T., e Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer, New York.

Larichev, O. e Moshkovich, H. (1997). *Verbal Decision Analysis for Unstructured Problems*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Lord, C. et al. (2000). *Autism Diagnostic Interview - Revised (ADI-R)*. Western Psychological Services, Los Angeles, CA.

Lord, C. et al. (2012). *Autism Diagnostic Observation Schedule - Second Edition (ADOS-2)*. Western Psychological Services, Los Angeles, CA.

Miller, L. et al. (2007). Sensory processing disorders: A review of the neurophysiologic and behavioral evidence. *American Journal of Occupational Therapy*, 61(2):135–149.

Mitchell, T. (1997). *Machine Learning*. McGraw Hill, New York.

Organization, W. H. (1993). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. World Health Organization, Geneva.

Organization, W. H. (2001). *International Classification of Functioning, Disability and Health*. World Health Organization, https://www.who.int/classifications/icf/en/.

Organization, W. H. (2019). Autism spectrum disorders. *Fact Sheets*. URL https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders.

Pinheiro, L., Pereira, M., Andrade, E., Nunes, L., Abreu, W., Pinheiro, P., Holanda Filho, R., e Pinheiro, P. (2021). An intelligent multicriteria model for diagnosing dementia in people infected with human immunodeficiency virus. *Applied Sciences*, 11(21):10457.

Raj, A. et al. (2024). Deep learning for autism spectrum disorder: A review of current trends and future directions. *Computers in Biology and Medicine*, 158:106940.

Reichow, B. (2011). Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *Journal of Autism and Development Disorders*, 42(4): 512–520.

Russell, S. e Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall, Upper Saddle River, NJ, 3 edition.

Sosa, P. et al. (2024). A clinical decision support system for autism spectrum disorder diagnosis using machine learning and expert knowledge. *Artificial Intelligence in Medicine*, 148:102512.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br

Tamanini, I., de Castro, A. K., Pinheiro, P. R., e Pinheiro, M. C. D. (2009). Towards an applied multicriteria model to the diagnosis of alzheimer's disease: A neuroimaging study case. In *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems (ICICISYS)*, p. 652–656.

Tamanini, I. e Pinheiro, P. (2008). Applying a new approach methodology with z.a.p.r.o.s. In *Anais do 40º Simpósio Brasileiro de Pesquisa Operacional*. Sociedade Brasileira de Pesquisa Operacional.

Tanwani, A., Afridi, J., Shafiq, M., e Farooq, M. (2009). Guidelines to select machine learning scheme for classification of biomedical datasets. In *European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*, p. 128–139. Springer.

Thabtah, F. (2017). Autism spectrum disorder screening: Machine learning adaptation and dsm-5 fulfillment. *Proceedings of the 2017 ACM on Web Science Conference*, p. 1–6.

Thabtah, F. (2018). Machine learning in autistic spectrum disorder behavioral research: A review and ways forward. *Informatics for Health and Social Care*, 44(3):278–297.

Thabtah, F. (2021). A review of machine learning techniques for autism spectrum disorder detection. *Journal of Healthcare Engineering*, 2021:9942479.

Thabtah, F., Abdelhamid, N., e Peebles, D. (2019). A machine learning autism classification based on logistic regression analysis. *Health Information Science and Systems*, 7(1):12.

Thabtah, F. e Peebles, D. (2019). A new machine learning model based on induction of rules for autism detection. *Health Informatics Journal*, 26(1):264–286.

Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., e Pinheiro, P. R. (2020). Covidgan: Data augmentation using auxiliary classifier gan for improved covid-19 detection. *IEEE Access*, 8:91916–91923.

Warren, Z. et al. (2011). Early intensive behavioral intervention: A summary for pediatricians. *Pediatrics*, 127(5):e1303–e1311.

DOI: 10.59254/sbpo-2024-193899
https://proceedings.science/p/193899?lang=pt-br