

Auxílio à Decisão no Diagnóstico do Transtorno do Espectro Autista com M-Health

Arthur A. Artoni¹, Bruno C. de Oliveira¹, Cinthyan R. Sachs C. de Barbosa¹

¹Programa de Pós-Graduação em Ciência da Computação – Universidade Estadual de Londrina (UEL) - Cx. Postal 10.011 – 86057-970 - Londrina – PR - Brazil

arthurartoni@uel.br, brunocarazatoo@gmail.com, cinthyan@uel.br

Abstract. *The diagnosis of Autistic Spectrum Disorder (ASD) is a challenge for medical professionals. The ASD is a very common and incurable pathology that affects the behavioral, social and communication part of the individual, but its diagnosis is very difficult and made only by behavioral analysis. In this work, we have developed an m-health tool that uses machine learning to aid in the diagnosis of ASD in children, adolescents and adults. For that, databases were used containing the results of the application of the AQ-10 test and other characteristics that could influence the diagnosis of ASD. Classification algorithms were applied to generate a predictive model, with the use of Linear SVM up to 100% accuracy was obtained.*

Resumo. *O diagnóstico do Transtorno do Espectro Autista (TEA) é um desafio para profissionais da área médica. O TEA, uma patologia muito comum e incurável, afeta a parte comportamental, social e de comunicação do indivíduo, porém seu diagnóstico é muito difícil e feito apenas por análises comportamentais. Neste trabalho foi desenvolvida uma ferramenta m-health que utiliza aprendizado de máquina para auxílio no diagnóstico do TEA em crianças, adolescentes e adultos. Para isso foram utilizadas bases de dados contendo resultados da aplicação do teste AQ-10 e outras características que poderiam influenciar o diagnóstico do TEA. Algoritmos de classificação foram aplicados para gerar um modelo preditivo, com o uso da SVM Linear foram obtidos até 100% de precisão.*

1. Introdução

O Transtorno do Espectro Autista (TEA) consiste em um transtorno global do desenvolvimento, que afeta o comportamento social, o desenvolvimento da função comunicativa e a percepção do indivíduo, manifestando-se na infância, geralmente até os três anos de idade e estará presente durante toda a vida desse [Silva et al. 2012].

Mesmo com os constantes avanços na área médica, os profissionais da saúde ainda enfrentam uma grande dificuldade para realizar o diagnóstico de doenças como o TEA. Isso acontece em decorrência da inexistência de exames laboratoriais que possam confirmar o diagnóstico. Normalmente esse é realizado por um profissional especialista no assunto, como neurologistas e psiquiatras. Para isso, esses profissionais podem utilizar escalas diagnósticas, ou seja, questionários específicos para diagnóstico do TEA desenvolvidos por especialistas, além de análises comportamentais e entrevistas com pais e/ou responsáveis que ajudam o médico a finalizar o diagnóstico [Ferreira 2010].

Vários fatores podem retardar ou mesmo impedir o diagnóstico correto e por consequência, o início do tratamento do TEA. Dentre esses fatores pode-se destacar: a demora na detecção dos primeiros sintomas e grande variação desses, falta de treinamento dos médicos e também dificuldades de acesso aos serviços de saúde [Zanon et al. 2014] [Penteado et al. 2015].

Uma das possíveis alternativas para auxiliar os profissionais da área médica, é empregar técnicas de *machine learning* como um mecanismo de apoio à decisão, a fim de auxiliar na realização o diagnóstico ou pré-diagnóstico do TEA [Bone et al. 2015]. O termo *machine learning* ou em português, aprendizado de máquina, consiste em uma técnica de Inteligência Artificial (IA) aplicada em vários ramos do conhecimento. Nessa técnica, o computador simula ou realiza um estudo do comportamento dos dados buscando um novo conhecimento ou habilidade e organiza, possibilitando a realização de uma melhora progressiva do seu próprio rendimento [Wilson and Obimbo 2011][Ishak and Siraj 2002].

Outra opção interessante é a *M-Health* (saúde móvel), que consiste basicamente no uso de tecnologias móveis aplicadas à saúde, tendo em vista vantagens como: baixo custo de utilização, fácil integração e comunicação com outros sistemas e/ou dispositivos e principalmente a mobilidade. Uma das características dessa tecnologia é a fácil integração com outros sistemas já existentes, como o prontuário eletrônico, uma versão digital do tradicional utilizado em hospitais, que tem como objetivo facilitar o acesso dos médicos ao histórico do paciente [Fuzyi et al. 2015].

É possível notar a necessidade de ferramentas computacionais que possam auxiliar os profissionais da saúde, em especial a um transtorno de difícil diagnóstico como o TEA. Este trabalho está estruturado da seguinte maneira: a seção 2 apresenta a fundamentação teórica na qual o trabalho foi baseado; a seção 3 explica a proposta do aplicativo *M-Health* desenvolvido; a seção 4 aborda os materiais e métodos que possibilitaram a criação do sistema *M-Health*; a seção 5 apresenta alguns resultados; por fim a seção 6 mostra as conclusões deste trabalho, além de apresentar possíveis trabalhos futuros.

2. Fundamentação Teórica

2.1. O Transtorno do Espectro Autista

O TEA atualmente é considerado como uma perturbação do desenvolvimento neurológico, afetando em especial, a maneira como seu portador compreende o mundo ao seu redor [Guedes and Tada 2015]. Existem vários tipos de classificações para o TEA. Para [Didehbani et al. 2016], o autismo pode ser dividido em três níveis, sendo o nível 1 o mais leve, o nível 2 considerado moderado e o nível 3 o mais severo. O primeiro nível tende a ter dificuldades em processar sinais sociais tornando-se oprimido e ansioso em interações sociais, especialmente com pessoas desconhecidas e dificuldades de expor pensamentos e emoções. O segundo nível apresenta graves problemas na sua comunicação, seja verbal ou não, além de uma extrema dificuldade de aceitar mudanças. O terceiro nível apresenta um comprometimento mais grave em interações sociais, precisando de um apoio para sua comunicação ser funcional.

O diagnóstico do TEA leva em conta os critérios estabelecidos pelo Manual Diagnóstico Estatístico de Transtornos Mentais (*DSM - Diagnostic and Statistical Manual*) da Sociedade Norte-Americana de Psiquiatria. Desenvolvido para diagnosticar diversos

tipos de transtornos mentais, não apenas o TEA, sua versão mais recente foi lançada em 2013 e recebe o nome de DSM-V [Association et al. 2014].

Além das escalas de diagnósticos destinadas aos profissionais da área médica, existem as autoaplicáveis, ou seja, escalas criadas para serem utilizadas por pessoas sem nenhuma experiência na área. Um exemplo é o *Autism Quocient* (AQ), que é um teste criado para auxiliar o diagnóstico do TEA proposto por [Baron-Cohen et al. 2001], sendo considerada como a primeira escala de diagnóstico de autismo que poderia ser aplicada sem a necessidade de um especialista. AQ é composta por 50 questões objetivas. Há ainda o AQ-10, que é uma adaptação do AQ, no qual as questões foram reduzidas apenas para as 10 mais relevantes. Além disso, existem três versões voltadas para o diagnóstico do TEA a fim de melhor se adaptar às questões mais impactantes para o diagnóstico do TEA em cada fase da vida [Allison et al. 2012]: crianças, adolescentes e adultos.

De acordo com [Durkin et al. 2015], recentes pesquisas realizadas pelo *WORLD BANK* em 2015, estimam que 86,5% dos diagnosticados com TEA encontram-se na América do Norte, Europa e Japão. Entretanto, essas regiões possuem apenas 10% das crianças que vivem no planeta. Dessa forma, nota-se uma clara dificuldade em diagnosticar o autismo em locais economicamente mais pobres, como é o caso do Brasil e outros países de terceiro mundo. Segundo [Zuckerman et al. 2014], o autismo em crianças latinas é diagnosticado cerca de dois anos e meio anos mais tarde do que em crianças não latinas.

A dificuldade de crianças latinas de encontrar o diagnóstico e o tratamento acaba gerando casos mais severos de autismo, em decorrência dessa demora [Zuckerman et al. 2014]. Uma das principais causas de atraso no diagnóstico do TEA, segundo [Zanon et al. 2014], é a falta de profissionais treinados para reconhecer as primeiras manifestações. Na maior parte das vezes, são os pais e não os profissionais da saúde, que acabam sendo os primeiros a suspeitarem. Quanto mais cedo o TEA for diagnosticado e o tratamento iniciado, melhores poderão ser os resultados, principalmente pela plasticidade do cérebro. Quanto antes o tratamento for iniciado, menores podem ser as expressões sintomática do TEA no indivíduo.

Além dos fatos já mencionados, existem fatores familiares que podem dificultar o diagnóstico, sendo o principal deles a negação. Devido ao TEA ser uma doença crônica, potencialmente grave e que requer intervenção imediata e contínua por toda a vida, muitos pais tendem a entrar em negação com o diagnóstico ou sua possibilidade [Lampreia 2015]. Essa negação também pode atrapalhar o uso de escalas autoaplicáveis, uma vez que o medo da doença pode acabar fazendo com que os pais não sejam totalmente objetivos na realização do teste.

2.2. Inteligência Artificial aplicada à Saúde

Existem algumas maneiras de aplicar a computação na resolução dos problemas do mundo real, como por exemplo, por meio da utilização da IA, a qual consiste na criação de programas que possam apresentar características de inteligência, ou seja, terem capacidade de aprender a executar uma tarefa simples ou resolver problemas complexos [Xue and Zhu 2009].

Diversas técnicas de aprendizado de máquina estão sendo utilizadas para o aprendizado supervisionado e o não supervisionado. O aprendizado não supervisionado consiste em um algoritmo no qual nenhum tipo de resposta esperada é dado para a resolução

do problema, cujo objetivo é descobrir por si mesmo a melhor maneira de reconhecer os padrões de dados. No aprendizado supervisionado antes da aplicação do algoritmo, um rótulo ou classe é aplicado aos dados, a fim de mostrar uma resposta esperada para o problema. Esse aprendizado funciona basicamente com duas fases, na qual a primeira consiste no treinamento, onde uma parte dos dados da base é usada para criar um modelo. Após o treino, outra parte dos dados que não foi previamente utilizada é usada para avaliar o desempenho [Xue and Zhu 2009].

Segundo [Thabtah 2017], tipicamente o diagnóstico do TEA utilizando aprendizado de máquina é encarado como um problema de Classificação, no qual um modelo é construído se baseando em dados previamente classificados de uma base de dados. Essa técnica pode oferecer métodos e soluções bem eficientes para o diagnóstico do TEA, uma vez que possui modelos matemáticos e métodos computacionais capazes de lidar com grandes volumes de dados e regras. O referido trabalho utiliza os dados desses mesmos *datasets* para propor um sistema de apoio à decisão baseado em regras. Isso foi um grande incentivo para a construção do M-Health utilizando de Aprendizado de Máquina, que será mostrado a seguir. Vários trabalhos que também utilizaram de AM, exibidos em [Artoni 2020], serviram de embasamento para a versão inicial desta proposta.

3. Proposta de M-Health

No presente trabalho é apresentado um aplicativo que utiliza M-Health e aprendizado de máquina, visando auxiliar médicos não especialistas na área a realizar um pré-diagnóstico do TEA a partir dos quatro anos de idade.

O objetivo é possibilitar uma rápida análise do paciente por médicos não especialistas em TEA, possibilitando que eles consigam levantar indícios relevantes que possam sugerir a presença do TEA e assim, ajudar a decidir sobre a necessidade de encaminhar o paciente a um especialista. No caso das crianças com suspeitas de TEA, geralmente o primeiro profissional a ter um maior contato com elas é o pediatra e não um neuropediatra ou psiquiatra infantil, que são os mais aptos a oficializar o diagnóstico.

Um diagrama de caso de uso da ferramenta pode ser visto na Figura 1.

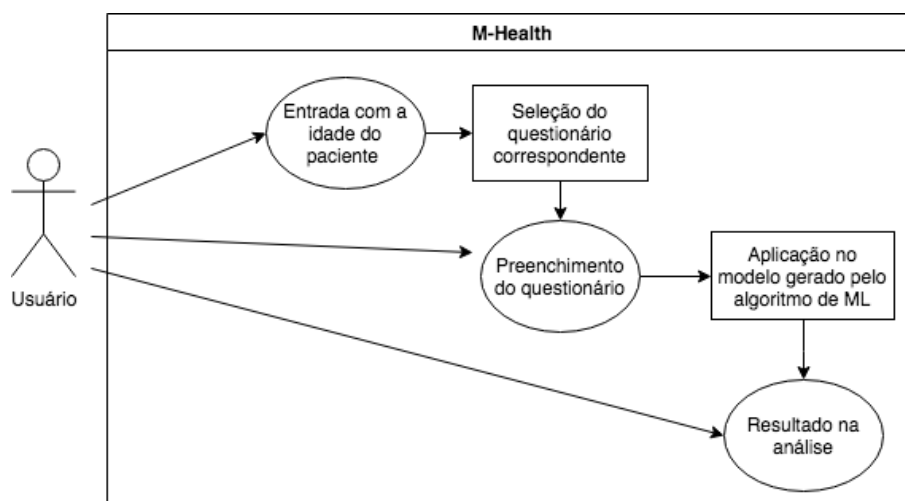


Figura 1. Diagrama da solução M-Health

Na tela inicial do aplicativo, o médico se depara com um campo para entrar com a idade do paciente. Tendo essa disponível o M-Health seleciona o modelo e o questionário mais adequado para o diagnóstico. Por fim, os dados preenchidos são usados como parâmetros de entrada para o modelo gerado pelo aprendizado de máquina, que executa e retorna o resultado da análise. Caso algum indício do TEA for encontrado, o aplicativo sugere ao médico a necessidade de encaminhar o paciente para um especialista, caso contrário, o aplicativo informa que não foi encontrado nenhum indício de TEA, como pode ser visto na Figura 2.

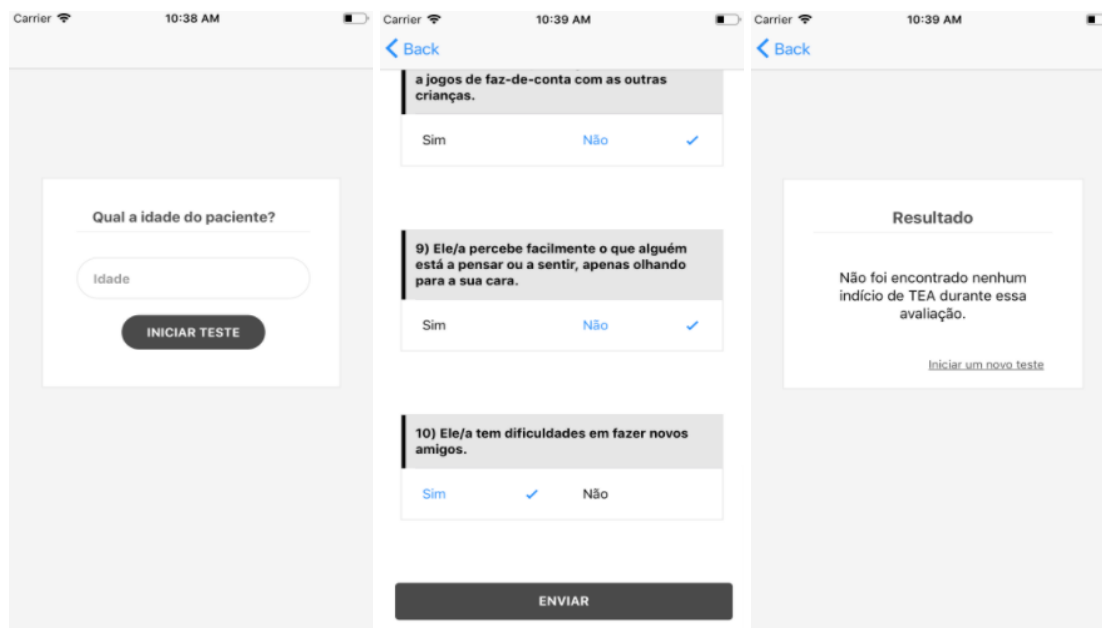


Figura 2. Tela inicial para entrada com a idade, tela do questionário e resultado respectivamente.

4. Materiais e métodos

4.1. Algoritmos de Aprendizado de Máquina

Para realizar a análise das bases de dados e gerar os modelos que compõem este trabalho foram utilizados os seguintes algoritmos:

- **Random Forest (RF)**: esse algoritmo foi proposto por [Breiman 2001] e é baseado no modelo de agregação de ideias. Idealizado a fim de oferecer uma boa resolução genérica de problemas complexos. É um algoritmo do tipo *Ensemble* que utiliza aleatoriedade e uma combinação de múltiplas árvores de decisão do tipo *Classification And Regression Tree (CART)*. Essas árvores são criadas utilizando amostragens geradas da separação da base de dados inicial em diversos subconjuntos por meio de uma técnica chamada de *bootstrap aggregating (bagging)*. O *Bagging* funciona escolhendo randomicamente as amostras que irão compor cada subconjunto, os quais serão utilizados para a construção das árvores de decisão. A RF possui ainda um método de ranqueamento de importância das características do problema. Esse é realizado durante a construção do modelo de aprendizado, onde os atributos mais relevantes tendem a alocar os nós nos níveis mais próximos

à raiz da árvore. Dessa forma, é possível saber quais são os atributos que mais se destacam na base de dados para a resolução do problema [Genuer et al. 2010].

- **Support Vector Machine (SVM)**: funciona traçando hiperplanos com o objetivo de separar as classes da melhor maneira possível. Ela pode ser aplicada aos problemas de classificação e regressão, além de outras funções como a detecção de *outliers*. Inicialmente, o foco da SVM era a solução de problemas binários e por meio dos hiperplanos buscava a separação máxima entre as classes, contudo, esse modelo nem sempre era capaz de oferecer boas soluções para problemas mais complexos. Para solucionar esse problema, um sistema de alteração de *kernels* foi desenvolvido. A SVM pode ser utilizada com quatro tipos de *kernels*: linear, polinomial, radial e sigmoidal. É importante salientar que os problemas de classificação possuem características que podem se adaptar melhor com um dos *kernels* em questão, melhorando assim significativamente a eficácia do modelo gerado [Cortes and Vapnik 1995].

4.2. Obtenção dos dados

Três bases de dados, disponíveis para *download*¹, foram utilizadas para o desenvolvimento deste projeto, contendo dados de resposta do teste AQ-10 e outros 11 atributos que, segundo [Thabtah 2017], poderiam ser relevantes para o auxílio no diagnóstico do TEA se utilizados juntamente com o AQ-10. A Tabela 1 mostra todos os atributos presentes nas três bases de dados.

Tabela 1. Atributos presentes na base de dados

Atributo	Tipo
Idade	Inteiro
Gênero	Texto
Etnia	Texto
Nascido com icterícia	Booleano (Sim ou Não)
Membro da família	Booleano (Sim ou Não)
Distúrbio Difuso de Desenvolvimento	Booleano (Sim ou Não)
Quem está completando o teste?	Texto
País de residência	Texto
Usou o aplicativo de triagem antes	Booleano (Sim ou Não)
Tipo de método de triagem	Inteiro (0,1,2,3)
Questões [1-10]	Binário (0, 1)
Resultado da triagem	Inteiro
Classe	Booleano (Sim ou Não)

Essas bases de dados englobam pacientes a partir dos quatro anos de idade, diagnosticados ou não com TEA. O número total de amostras nas três bases foi de 1100 indivíduos, como pode ser visto na Tabela 2.

Importante salientar que como essas bases de dados são públicas e não identificam os pacientes, não foi necessário submeter este trabalho a um Comitê de Ética, uma vez que não houve nenhuma violação da privacidade das informações dos pacientes.

¹<http://fadifayez.com/autism-datasets/>

Tabela 2. Número total de indivíduos com e sem TEA em todas as bases de dados

	Base Adultos	Base Adolescente	Base Crianças
Indivíduos com TEA	189	63	141
Indivíduos sem TEA	515	41	151
Total	704	104	192

4.3. Desenvolvimento do Modelo

A primeira fase do trabalho foi efetuar o pré-processamento das bases de dados, a fim de retirar dados ruidosos, faltantes e até alguns *outliers* que poderiam atrapalhar o processo de aprendizagem dos algoritmos e por consequência o modelo gerado.

A segunda parte do trabalho foi a de seleção dos atributos mais relevantes presentes nas bases. Para isso, todas as bases de dados foram submetidas individualmente ao processo de ranqueamento de importância dos atributos utilizando a RF. Nas três bases as perguntas relativas ao teste AQ-10 tiveram um índice Gini bem superior aos outros atributos, indicando que elas têm um peso muito maior que os demais e que essas características adicionais poderiam até estar interferindo na criação de um bom modelo.

A terceira parte foi a aplicação dos algoritmos de aprendizado de máquina, a fim de verificar o mais adequado. O que obteve os melhores resultados foi a SVM utilizando o *kernel* linear.

A SVM e todos os outros algoritmos testados anteriormente foram treinados utilizando o *Cross Validation* (Validação Cruzada) com 10 *folds* (dobras). A Validação Cruzada é uma técnica matemática que tem como objetivo evitar que a disposição dos dados na base possa estar arranjada de maneira a beneficiar alguma técnica específica de aprendizado de máquina. Para isso, ela divide a base de dados em K subconjuntos exclusivos e de tamanhos iguais, então um desses é utilizado para teste enquanto os $K-1$ restantes são utilizados para treinar o modelo. Dessa forma, não é possível colocar os dados em uma disposição específica na base de dados [Kohavi et al. 1995].

A Figura 3 mostra o fluxo completo do desenvolvimento para uma das bases de dados. É importante lembrar que esse processo foi repetido três vezes, uma vez para cada base de dados.

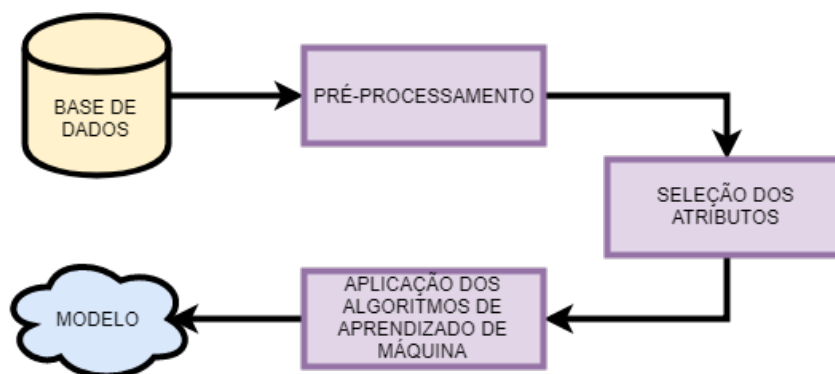


Figura 3. Etapas de desenvolvimento

4.4. Aplicativo

O aplicativo foi desenvolvido utilizando o *framework* de desenvolvimento híbrido Ionic 3. Esse foi escolhido por permite a criação de versões compatíveis com os sistemas operacionais Android e iOS. Dessa forma, grande parte dos médicos não especialistas teria compatibilidade com seus Smartphones ou tablets.

O aplicativo foi testado em simuladores dos seguintes aparelhos:

- iPhone 6, sistema operacional iOS 12.1
- Nexus 4, sistema operacional Android 8.1

Também foi testado em um smartphone LG X Power com sistema operacional Android 6.0.1.

5. Resultados

A Tabela 3 mostra a diferença de desempenho pelos modelos gerados com a SVM-Linear ao fazer uso ou não das características não relacionadas ao teste AQ-10. Ao contrário do esperado, o uso dessas 11 características extras não agregou nenhum tipo de ganho ao modelo nos cenários durante o desenvolvimento deste projeto. Na maioria das vezes, essas características acabaram interferindo de maneira negativa durante a classificação dos indivíduos. Já a *Random Forest* teve um papel de extrema importância na seleção das características mais relevantes de cada base de dados.

Tabela 3. Desempenho da SVM-Linear na classificação em cada base de dados

	Adults	Teenagers	Children
AQ-10 Questions	100%	94,23%	100%
Complete dataset	100%	88,46%	99,65%

Após os experimentos, os modelos escolhidos para fazer parte do aplicativo foram os gerados pela SVM-Linear utilizando apenas as questões referentes ao teste AQ-10.

O aplicativo foi desenvolvido oferecendo uma interface simples, que possibilita a coleta de dados do paciente de maneira rápida e fácil pelos médicos. Após a finalização do questionário, esses dados são usados como parâmetros de entrada para o modelo. Em poucos segundos retorna ao médico um diagnóstico com um bom nível de segurança, uma vez que a taxa de acertos do modelo foi consideravelmente alta alcançando em alguns pontos até 100%.

O modelo com menor grau de precisão gerado foi o de adolescentes. É possível que seja uma consequência de um menor número de amostras na base de dados, o que pode ter limitado o desempenho dos algoritmos de aprendizado de máquina e por consequência o modelo gerado teve um grau de acerto inferior se comparado aos outros dois. Uma possível solução para esse problema poderia ser conseguindo mais dados referentes aos adolescentes para serem incorporados à base de dados e assim verificar se a geração de um novo modelo, seguindo as especificações do primeiro, traria melhores resultados.

6. Conclusões

O TEA é uma patologia de diagnóstico bastante complexo, mesmo com os crescentes avanços nas áreas da medicina e da tecnologia. Ainda hoje é complicado para os profissionais da área médica realizarem diagnóstico de doenças como TEA, uma vez que esse é

baseado em análises comportamentais. Infelizmente nem todos os profissionais da área da saúde têm o conhecimento ou recursos necessários para detectar o TEA, o que torna esse desafio ainda maior. Assim, a aplicação M-Health apresenta-se como uma ferramenta de apoio à decisão para auxiliar os médicos, em especial os que não têm experiência com o diagnóstico de TEA. Além disso, a ferramenta traz maior mobilidade, visto que é um aplicativo para dispositivos móveis que apresenta rapidez e ajuda na prevenção de erros humanos durante a realização do teste AQ-10.

Como contribuição científica também, vale ressaltar a constatação de que em todas as bases de dados, as únicas características que foram realmente relevantes para o modelo de aprendizagem de máquina foram as perguntas referentes ao teste AQ-10. Portanto, foi possível desenvolver um sistema mais simples para os profissionais da saúde com um alto grau de precisão.

Para trabalhos futuros, espera-se aplicar esta ferramenta para verificar aceitação da mesma pelos profissionais da saúde (consequentemente trazendo uma possibilidade maior da evolução do tratamento da doença) e além disso poder expandi-la, oferecendo outros testes diagnósticos.

Pretende-se ainda implantar um banco de dados para armazenar os resultados dos testes, a fim de elicitare um constante incremento de amostras nas bases de dados e fornecer estatísticas relevantes sobre possíveis casos de TEA, categorizando por região, idade, dentre outras características.

Referências

- Allison, C., Auyeung, B., and Baron-Cohen, S. (2012). Toward Brief “Red Flags” for Autism Screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist in 1,000 Cases and 3,000 Controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(2):202–212.
- Artoni, A. A. (2020). Aplicação de Aprendizado de Máquina no Auxílio ao Diagnóstico do Transtorno do Espectro Autista. Dissertação de Mestrado em Ciência da Computação, Departamento de Computação, Universidade Estadual de Londrina.
- Association, A. P. et al. (2014). *DSM-5: Manual Diagnóstico e Estatístico de Transtornos Mentais*. Artmed.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, 31(1):5–17.
- Bone, D., Goodwin, M. S., Black, M. P., Lee, C.-C., Audhkhasi, K., and Narayanan, S. (2015). Applying Machine Learning to Facilitate Autism Diagnostics: Pitfalls and Promises. *Journal of Autism and Developmental Disorders*, 45(5):1121–1136.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

- Didehbani, N., Allen, T., Kandalafi, M., Krawczyk, D., and Chapman, S. (2016). Virtual reality social cognition training for children with high functioning autism. *Computers in Human Behavior*, 62:703–711.
- Durkin, M., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R., Lee, L.-C., Rattazzi, A., Stapel-Wax, J., Stone, W., Tager-Flusberg, H., Thurm, A., Tomlinson, M., and Shih, A. (2015). Autism Screening and Diagnosis in Low Resource Settings: Challenges and Opportunities to Enhance Research and Services Worldwide. *Autism Research*, 8(5):473–476.
- Ferreira, R. S. (2010). Autism Testing: Uma Ferramenta Móvel no Auxílio ao Pré-Diagnóstico do Autismo. In *XXII Conferência Internacional sobre Informática na Educação*, volume 13, pages 178–187, Fortaleza. TISE. Nuevas Ideas en Informática Educativa.
- Fuzyi, E. M., Silva, M. C., and Barbon, S. (2015). M-health Solution on Pre-Diagnosis of Larynx. In *Annual Conference on Brazilian Symposium on Information Systems: Information Systems: A Computer Socio-Technical Perspective*, SBSI 2015, page 501–508, Porto Alegre, BRA. Brazilian Computer Society.
- Genuer, R., Poggi, J.-M., and Tuleau-Malot, C. (2010). Variable selection using Random Forests. *Pattern Recognition Letters*, 31(14):2225–2236.
- Guedes, N. S. P. and Tada, I. N. C. (2015). A Produção Científica brasileira sobre Autismo na Psicologia e na Educação. *Psicologia: Teoria e Pesquisa*, 31(3):303–309.
- Ishak, W. H. W. and Siraj, F. (2002). Artificial Intelligence in Medical Application: an exploration. *Health Informatics Europe Journal*, v. 16.
- Kohavi, R. et al. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *14th International Joint Conferences on Artificial Intelligence*, volume 14, pages 1137–1145, Montreal, Canada.
- Lampreia, A. R. S. (2015). Percepções Parentais sobre a Perturbação do Espectro do Autismo: processo de diagnóstico, interferência e recursos. Dissertação de Mestrado em Psicologia, Faculdade de Psicologia, Universidade de Lisboa.
- Penteado, F. A. O., Salvadeo, D. H. P., Aporta, A. P., Garcia, R. V. B., and Manoni, N. d. V. (2015). Software para Auxílio ao Diagnóstico de Autismo. In *VIII Congresso de extensão universitária da UNESP*, pages 1–4, São Paulo. Universidade Estadual Paulista Júlio de Mesquita Filho.
- Silva, A., Gaiato, M. B., and Reveles, L. T. (2012). *Mundo Singular: entenda o autismo*. Fontana, Rio de Janeiro.
- Thabtah, F. (2017). Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment. In *1st International Conference on Medical and Health Informatics*, Taichung, Taiwan. ACM, 1-6.
- Wilson, R. and Obimbo, C. (2011). Self-Organizing Feature Maps for User-to-Root and Remote-to-Local Network Intrusion Detection on the KDD Cup 1999 Dataset. *2011 World Congress on Internet Security*, pages 42–47.

- Xue, M. and Zhu, C. (2009). A Study and Application on Machine Learning of Artificial Intelligence. In *International Joint Conference on Artificial Intelligence*, pages 272–274, Pasadena, California, USA. IEEE.
- Zanon, R. B., Backes, B., and Bosa, C. A. (2014). Identificação dos Primeiros Sintomas do Autismo pelos Pais. *Psicologia: Teoria e Pesquisa*, 30(1):25–33.
- Zuckerman, K. E., Sinche, B., Mejia, A., Cobian, M., Becker, T., and Nicolaidis, C. (2014). Latino Parents' Perspectives on Barriers to Autism Diagnosis. *Academic pediatrics*, 14(3):301–308.