

UM MODELO COMPUTACIONAL DE RECOMENDAÇÃO DE MÉDICO ESPECIALISTA COM DADOS HISTÓRICOS DE PACIENTES EM UMA CLÍNICA PRIVADA¹

Gabriel U. V. Messina², Rafael Zancan Frantz³

¹ Projeto de pesquisa desenvolvido na Universidade Regional do Noroeste do Estado do Rio Grande do Sul, pelo Programa de Pós-graduação Stricto Sensu em Modelagem Matemática e Computacional.

Computacional da linha de pesquisa Otimização e Integração de Grandes Volumes e Dados

² Bacharel em Ciência da Computação, Estudante do Programa de Pós-graduação em Modelagem Matemática e Computacional (mestrado) da UNIJUI.

³ Doutor em Tecnologia e Engenharia de Software, Professor da Unijui, Orientador

RESUMO

Este estudo foca na aplicação da inteligência artificial (IA) no setor de saúde. Apesar da IA e da análise preditiva serem vistas como instrumentos potenciais para aprimorar a previsibilidade em saúde, o impacto ainda não atingiu o patamar esperado. No entanto, o advento acelerado e o desenvolvimento contínuo do aprendizado de máquina e da análise preditiva estão reformulando quase todos os campos, incluindo a medicina, com a precisão das previsões aumentando proporcionalmente com a crescente quantidade de dados disponíveis.

Ainda que existam avanços significativos na análise preditiva no campo da saúde, muitas dessas aplicações ainda não trouxeram melhorias significativas para a assistência clínica. A meta desta pesquisa é contribuir para o conhecimento de IA, por meio da implementação de um modelo preditivo capaz de direcionar o paciente ao profissional de saúde adequado. O estudo, será feito uso de técnicas de análise exploratória de dados, sistemas de recomendação e conjuntos de dados.

Palavras-chave: Inteligência Artificial. Aprendizado de Máquina. Análise Exploratória de Dados. Sistemas de Recomendação. Modelos Preditivos.

ABSTRACT

This study focuses on the application of artificial intelligence (AI) in the health sector. Although AI and predictive analysis are seen as potential tools to improve predictability in health, the impact has not yet reached the expected level. However, the rapid advent and continuous development of machine learning and predictive analysis are reshaping almost all fields, including medicine, with the accuracy of predictions increasing proportionately with the growing amount of available data.

Even though there have been significant advances in predictive analysis in the field of health, many of these applications have not yet brought significant improvements to clinical care. The goal of this research is to contribute to AI knowledge, through the implementation of a predictive model capable of directing the patient to the appropriate health professional. The study will make use of techniques of exploratory data analysis, recommendation systems, and data sets.

Keywords: Artificial Intelligence. Machine Learning. Exploratory Data Analysis. Recommendation Systems. Predictive Models.

INTRODUÇÃO

A saúde no Brasil apresenta-se como um dos desafios sociais, políticos e econômicos mais complexos do país. Apesar da garantia constitucional de direito à saúde para todos, essa obrigação nem sempre é satisfeita. Uma rota longa e desafiadora ainda precisa ser percorrida para atingir uma capacidade sanitária adequada, o que requer a reestruturação do estado de saúde constitucional e a exploração de novas alternativas de atenção à saúde, (Souza e Botazzo 2013).

A situação precária da saúde pública do país nas décadas de 60 levou algumas das maiores empresas brasileiras a oferecerem serviços médicos aos seus funcionários. Essa necessidade incentivou os médicos a criar empresas de medicina de grupo, que ofereciam vários planos de saúde. Em 1997, essas empresas serviam mais de 17 milhões de pessoas. As cooperativas médicas surgiram como organizações que fornecem serviços médicos e distribuem os lucros entre os médicos associados, (Machado, Baptista e Lima, 2012).

Conforme a pesquisa do Instituto Brasileiro de Geografia e Estatística (IBGE, 2019), o custo dos serviços de saúde privados foi a principal despesa das famílias brasileiras entre 2010 e 2017. Em 2017, o governo desembolsou aproximadamente R\$608,3 bilhões em cuidados e serviços de saúde no Brasil, com mais da metade desse valor sendo gasto em despesas familiares.

A distribuição desigual de médicos no Brasil representa um problema substancial que impacta a qualidade dos serviços de saúde. Apesar de a proporção de médicos no Brasil ser maior do que o ideal recomendado pela OMS, esses profissionais estão majoritariamente concentrados nas capitais e grandes cidades, resultando em superpopulação de médicos em grandes centros e déficits significativos em outras regiões (Machado, 1997).

O tempo de espera para uma primeira consulta pode afetar o desenrolar dos casos e a qualidade de vida dos pacientes, além de refletir a capacidade de resposta do sistema de saúde às necessidades da população (Melo, Araújo, Rodrigues, Jesus e Cardoso, 2021). No Brasil, os longos tempos de espera para procedimentos médicos especializados são uma das principais barreiras para o cuidado integral à saúde. Doenças crônicas prevalentes, falta de

educação e recursos também contribuem para este problema, (Castiglioni, Rundo, Codari, Di Leo, Salvatore, Interlenghi, Gallivanone, Cozzi, D'Amico e Sardanelli, 2021.).

Neste contexto, a necessidade de buscar soluções tecnológicas capazes de acelerar o processo de diagnóstico e encaminhamento de pacientes para especialistas é clara. Uma solução em potencial é o desenvolvimento de um sistema de inteligência artificial para analisar os sintomas do paciente e dados históricos, prevendo a especialidade médica mais adequada para cada caso. Isso poderia reduzir significativamente o tempo de espera para a resolução do problema e diminuir a chance de agravamento dos sintomas.

METODOLOGIA

A Inteligência Artificial (IA) tem grande potencial para analisar enormes volumes de dados médicos. Este capítulo detalha a metodologia de pesquisa adotada neste estudo - a Design Science Research (DSR).

A DSR é uma metodologia que visa criar soluções inovadoras para problemas reais e contribuir para o conhecimento teórico. É adequada em áreas como Ciência da Computação, focada em desenvolver e avaliar sistemas para resolver problemas práticos.

Neste estudo, a DSR é ideal para desenvolver e testar um modelo de IA que identifique a especialidade médica mais apropriada para cada paciente. A pesquisa adota várias etapas da DSR: identificação do problema, design e desenvolvimento do artefato, demonstração da funcionalidade, avaliação do artefato, e comunicação dos resultados (Alpaydin, 2020).

As etapas refletem o processo iterativo da DSR, fornecendo uma estrutura robusta para o desenvolvimento do modelo de IA. Elas permitem uma abordagem organizada para lidar com os complexos desafios do uso da IA na saúde (Harari, 2018).

A DSR também considera possíveis limitações e desafios ao longo do caminho. Assim, a DSR foi escolhida por sua capacidade de lidar com problemas práticos, foco na inovação e contribuição teórica. Seu processo estruturado aumenta a probabilidade de melhorar o acesso dos pacientes à assistência médica especializada.

MATERIAIS E MÉTODOS

Este estudo científico utiliza dados extraídos de várias fontes médicas, incluindo registros eletrônicos, bancos de exames e bases de prescrições. Os dados foram coletados de diferentes hospitais e clínicas e anonimizados para assegurar a privacidade dos pacientes.

O conjunto de dados compreende informações como histórico médico, sintomas e tratamentos anteriores dos pacientes, que são essenciais para treinar o modelo de IA na determinação da especialidade médica adequada.

A coleta de dados seguiu as diretrizes éticas locais e todas as variáveis identificáveis foram eliminadas para preservar a confidencialidade. A qualidade dos dados foi assegurada através de análise e medidas de pré-processamento.

Finalmente, o conjunto de dados foi dividido em grupos de treinamento e teste para o treinamento e a avaliação do modelo de IA.

O modelo de IA foi implementado usando a linguagem Python e a biblioteca Scikit-learn, selecionadas pela sua clareza sintática, diversidade de bibliotecas científicas e facilidade de uso. Eles oferecem variados algoritmos de aprendizado de máquina, técnicas de pré-processamento e métodos de avaliação, fundamentais para um modelo de IA eficaz.

Ferramentas complementares, como as bibliotecas Pandas para manipulação e análise de dados, e Matplotlib e Seaborn para visualização, também foram empregadas. O Jupyter Notebook foi utilizado como plataforma de desenvolvimento devido à sua interface interativa e capacidade de documentação.

O desenvolvimento do modelo de IA seguiu o ciclo iterativo de aprendizado de máquina, englobando pré-processamento de dados, seleção de atributos, treinamento e avaliação do modelo, e ajuste de hiperparâmetros.

No pré-processamento, técnicas de limpeza e transformação de dados foram aplicadas, como a remoção de dados ausentes e outliers e a codificação de variáveis categóricas. Na seleção de atributos, técnicas de análise de correlação e importância de atributos foram usadas para identificar variáveis relevantes.

Diversos algoritmos de aprendizado de máquina foram testados durante o treinamento, e o modelo final foi escolhido com base em sua precisão e robustez. A avaliação utilizou métricas como precisão, recall, F1-score e AUC-ROC.

O ajuste de hiperparâmetros foi realizado para otimizar o desempenho do modelo, empregando técnicas como busca em grade e busca aleatória.

A pesquisa descrita nesta dissertação adotou a metodologia de Design Science Research (DSR), uma abordagem comum em ciência da computação, focada na criação e avaliação de artefatos para resolver problemas. No contexto deste estudo, o artefato é o modelo de IA para direcionamento de pacientes à especialidade médica correta.

A DSR foi escolhida por sua aplicabilidade no desenvolvimento e avaliação de soluções tecnológicas, alinhada com o objetivo da pesquisa. A metodologia inclui seis etapas: identificação do problema, definição dos objetivos da solução, design e desenvolvimento, demonstração, avaliação e comunicação dos resultados.

A pesquisa identificou o problema do direcionamento subjetivo de pacientes, definiu o objetivo de desenvolver um modelo de IA para a tarefa, projetou e desenvolveu a solução, demonstrou sua eficácia com dados reais, avaliou o desempenho usando métricas relevantes e comunicou os resultados através de publicações e apresentações.

A DSR, com sua abordagem sistemática e ênfase na resolução de problemas reais, forneceu um quadro metodológico robusto para o desenvolvimento e avaliação do modelo de IA, alinhando-se com a meta de melhorar o atendimento ao paciente na prática médica.

REFERENCIAL TEÓRICO

A inteligência artificial (IA) busca simular habilidades cognitivas humanas, como aprendizado e raciocínio. Desde o "Teste de Turing" de Alan Turing em 1950, que propôs um critério para avaliar a inteligência de máquinas, a IA tem evoluído consideravelmente. Hoje, ela permeia diversas áreas, da medicina à economia, auxiliando em diagnósticos precisos, otimizações e análise de grandes volumes de dados (Turing, 2007).

Entretanto, essa ascensão traz desafios. Questões éticas sobre privacidade e impacto no mercado de trabalho surgem à medida que máquinas inteligentes se tornam mais integradas à sociedade. O debate não se limita à técnica: reflete sobre a natureza da própria inteligência e se máquinas podem replicar a cognição humana.

A IA promete transformações significativas, capazes de melhorar vidas e revolucionar setores. Contudo, é imperativo abordar seus avanços com uma visão crítica, garantindo que a tecnologia evolua de forma ética e consciente. Em resumo, a IA é simultaneamente uma promessa de inovação e um chamado à reflexão sobre suas implicações no mundo.

O Aprendizado de Máquina (ML) é um subcampo da inteligência artificial (IA) focado em desenvolver sistemas que aprendem com dados. Diferentemente da programação tradicional, onde instruções explícitas são fornecidas, o ML utiliza algoritmos para descobrir padrões e fazer previsões. Estes algoritmos são diversificados, escolhidos de acordo com a natureza do problema e as características dos dados (Mahesh, 2020).

Nos dias de hoje, o crescimento exponencial de dados e o avanço tecnológico no processamento e armazenamento têm propiciado o auge do ML. Isso permitiu aplicações práticas em vários campos, destacando-se a medicina, onde sistemas de ML são usados para diagnosticar doenças, sugerir tratamentos e analisar dados não estruturados (Alpadyn, 2020).

O ML tem raízes na década de 1950, com Arthur Samuel desenvolvendo um programa de aprendizagem. Entretanto, a consolidação como campo autônomo ocorreu nas décadas subsequentes, com ênfase no desenvolvimento de algoritmos robustos e na teoria por trás do aprendizado.

A essência do ML é o aprendizado a partir de dados. À medida que mais dados são fornecidos, o modelo se torna mais preciso. Há interseções com estatística e IA, levando a debates sobre sua natureza e objetivos. Por fim, como toda tecnologia disruptiva, é crucial abordar o ML com responsabilidade, ponderando suas implicações éticas e práticas.

O cérebro humano, composto por mais de 10 bilhões de neurônios interconectados, possui uma capacidade impressionante de processar tarefas complexas, desde reconhecimento facial a cálculos relacionados à fala (Abraham, 2003). Buscando emular essa habilidade, foram desenvolvidas as Redes Neurais Artificiais (RNAs), representações matemáticas do sistema nervoso. Estas redes, com origens na década de 1950, não são programadas para tarefas específicas, mas aprendem a partir de exemplos, assim como o aprendizado humano (Jiang, Jiang, Zhi, Dong, Li, Ma, Wang, Dong, Shen e Wang 2017).

Cada RNA é composta por nós, semelhantes aos neurônios, que se comunicam através de conexões, simulando as sinapses cerebrais. Essas conexões ajustam-se com base nas saídas dos neurônios, permitindo à RNA adaptar-se a tarefas complexas. Esta adaptabilidade fez com que as RNAs ganhassem relevância em áreas variadas, desde pesquisa oncológica até descoberta de medicamentos. No entanto, apesar dos avanços, a compreensão e aplicação desses sistemas ainda são incipientes.

Uma RNA pode lidar com múltiplas tarefas, desde classificação até filtragem e tomada de decisão. Porém, antes de sua implementação, é essencial compreender sua teoria e interação com diferentes algoritmos de aprendizado (Suzuki, 2011). Escolher o algoritmo e modelo adequados é crucial para otimizar o desempenho da rede, ressaltando a importância de conhecer as propriedades e limitações dessas ferramentas ao desenvolver sistemas baseados em RNAs.

MODELO COMPUTACIONAL PROPOSTO

O modelo computacional modularizado discutido neste trabalho tem o intuito de otimizar o atendimento médico ao fornecer informações personalizadas aos pacientes. Constituído por três módulos principais, o primeiro usa inteligência artificial (IA) para identificar palavras-chave em prescrições e exames, criando conjuntos de dados baseados no histórico médico do paciente. O segundo módulo aplica tecnologia de reconhecimento facial para identificar o paciente, recolhendo dados vitais e sintomas. Já o terceiro módulo combina informações dos dois anteriores para determinar a especialidade médica mais adequada ao paciente.

Comparativamente, este modelo supera sistemas médicos tradicionais ao aplicar IA para uma identificação mais precisa de condições médicas, especialmente quando especialidades raras ou específicas são necessárias. Tal abordagem tem o potencial de revolucionar o atendimento ao paciente, garantindo diagnósticos mais precisos, menor tempo de espera e melhor gestão do fluxo hospitalar. Adicionalmente, permite atualizações constantes, adaptando-se aos avanços médicos e servindo como ferramenta de apoio à decisão para médicos.

No entanto, há desafios associados, incluindo questões de privacidade e segurança, treinamento profissional e aceitação do sistema por parte dos envolvidos. Mesmo com tais desafios, os benefícios podem transformar significativamente o atendimento médico, priorizando um tratamento mais individualizado e eficaz.

O sucesso da aplicação de inteligência artificial (IA) na medicina está profundamente atrelado à qualidade dos dados utilizados. No modelo em foco, coletamos informações detalhadas de pacientes e as armazenamos no MongoDB, um banco de dados NoSQL

conhecido por sua eficiência e flexibilidade. Esses dados incluem históricos médicos, sintomas e exames, permitindo que a IA faça recomendações precisas de tratamento.

Implementamos também microsserviços, que são pequenas funções especializadas, como o reconhecimento facial do paciente e o acesso rápido a seus registros médicos. Isso aumenta a eficiência e facilita atualizações.

A atualização constante dos dados é vital. Eles são complementados com novas pesquisas médicas para manter a IA informada. Simultaneamente, garantimos a segurança e privacidade das informações, alinhando-nos a regulamentos rigorosos, como GDPR e LGPD (W. N. Price e I. G. Cohen, 2019).

A representatividade dos dados é crucial. Garantimos que eles abranjam diferentes gêneros, idades e origens, evitando vieses e assegurando eficácia da IA para todos. Em resumo, um armazenamento de dados robusto, combinado com constantes atualizações e foco na privacidade, é o alicerce do sucesso da nossa IA na prática médica (Obermeyer, Powers, Vogeli e Mullainathan, 2019).

A metodologia aplicada para criar um dataset de exames médicos não-imagem passou por quatro etapas principais: seleção de documentos, extração de informações, limpeza e pré-processamento, e armazenamento do dataset (Rajkomar, Dean e Kohane, 2019).

Na fase inicial, documentos médicos foram analisados, selecionando-se aqueles com informações relevantes e representativas da população. Na extração de informações, utilizamos as bibliotecas pdfminer e textblob para identificar palavras-chave e entidades médicas nos documentos. Esse processo também envolveu técnicas de processamento de texto para assegurar a qualidade das informações coletadas (Zhang, Pakhomov, McInnes e Melton, 2011).

Na segunda fase, a extração de informações, as bibliotecas pdfminer e textblob foram utilizadas para identificar e extrair palavras-chave e informações relevantes dos documentos selecionados. A extração de informações foi realizada em várias etapas, incluindo a identificação de entidades médicas, como nomes de exames, doenças, diagnósticos e valores numéricos associados a essas entidades (Heinrich, 2005).

Durante a extração de informações, também foi necessário lidar com a variabilidade dos formatos dos documentos e a qualidade do texto extraído. Técnicas de processamento de

texto, como a remoção de caracteres especiais, correção ortográfica e tokenização, foram utilizadas para garantir a consistência e a qualidade das informações extraídas.

A terceira fase concentrou-se na limpeza dos dados usando a biblioteca pandas, garantindo a correção de erros e a padronização das informações. Finalmente, os dados foram armazenados de forma compatível com ferramentas de análise e aprendizado de máquina (Cohen, Aviram, Elhadad e Elhadad, 2014).

Na quarta e última fase, armazenamento e disponibilização do dataset gerado, os dados foram armazenados em um formato compatível com as ferramentas de análise de dados e aprendizado de máquina. O armazenamento e disponibilização do dataset garantem que ele possa ser facilmente acessado e utilizado em futuras análises e treinamento de modelos de inteligência artificial (Pons, Braun, Hunink e Kors, 2016).

O dataset resultante é vital para treinar modelos de inteligência artificial na medicina. Tais modelos, quando integrados aos sistemas de saúde, têm o potencial de transformar o atendimento médico, auxiliando profissionais na tomada de decisão com informações mais precisas. Contudo, a expertise humana continua sendo essencial no processo de diagnóstico e tratamento. Em suma, este dataset é um avanço significativo para a aplicação de IA na área da saúde.

As fases essenciais envolvidas na elaboração do sistema de recomendação, sublinhando os desafios, as soluções propostas e as perspectivas de aprimoramento.

- Coleta de dados: Este passo inicial é decisivo para a eficácia do sistema de recomendação. Aqui, dados de várias origens, como histórico do usuário, avaliações e dados demográficos, são adquiridos. A integridade destes dados é crucial. Também é fundamental garantir que as práticas de coleta respeitem a privacidade e a ética.
- Preparação dos dados: Após a coleta, os dados são organizados e adaptados para análise. Este processo envolve correção de erros, categorização e transformação dos dados. Esta etapa é crucial, pois a qualidade do pré-processamento influencia diretamente o desempenho do sistema de recomendação.
- Construção do modelo de aprendizado: Com os dados prontos, inicia-se a construção do modelo. O sistema aqui discutido utiliza a filtragem

colaborativa, uma abordagem que se baseia nas avaliações de usuários para proporcionar recomendações personalizadas (Ricci, Rokach e Shapira, 2011).

- **Treinamento e teste:** Aqui, os dados são segmentados em treinamento e teste. O treinamento ajusta os parâmetros do modelo, enquanto o teste avalia sua precisão. Durante esta fase, ajustes são realizados para otimizar o modelo e evitar o sobreajuste.
- **Avaliação da eficácia:** Após o treinamento, é vital verificar o desempenho do sistema. Diversas métricas são usadas para esta análise, considerando tanto a precisão das recomendações quanto a diversidade e satisfação do usuário.
- **Implementação e manutenção:** Uma vez validado, o sistema é colocado em prática. Contudo, monitorar e atualizá-lo é essencial para que continue atendendo às mudanças nas preferências dos usuários.
- **Abordagem híbrida:** Como aperfeiçoamento, pode-se considerar a integração de outras técnicas ao sistema, como filtragem baseada em conteúdo e aprendizado profundo, visando um sistema mais robusto e adaptável.

Em resumo, o sistema de recomendação evidencia o valor das técnicas de aprendizado de máquina, como a filtragem colaborativa, em potencializar a experiência do usuário em variados contextos. Através de uma abordagem metódica, desde a coleta de dados até a avaliação, foi possível conceber um sistema eficaz e acurado. A contínua avaliação e a possibilidade de incorporar abordagens híbridas podem oferecer melhorias significativas, garantindo recomendações sempre pertinentes e úteis.

EXECUÇÃO E COLETA DE DADOS

A pesquisa buscou implementar um modelo de Inteligência Artificial (IA) capaz de redirecionar pacientes à especialidade médica adequada, baseando-se em dados clínicos e prontuários dos pacientes e seus familiares diretos. Os dados, coletados de julho a dezembro de 2022, foram ampliados mensalmente: começando com 89 pacientes (192 exames) e culminando em 712 pacientes (1189 exames). Cada etapa de coleta passou por rigoroso pré-processamento, incluindo organização, verificação de qualidade, eliminação de duplicatas e tratamento de inconsistências.

Uma análise exploratória foi feita para identificar correlações e padrões nos dados. Estes serviram para o treinamento e validação do modelo proposto. Em sua construção, o modelo dividiu os dados em conjuntos de treinamento e validação. Utilizou-se aprendizado supervisionado e técnicas de validação cruzada, otimizando hiperparâmetros e evitando sobreajuste e subajuste. Posteriormente, o modelo foi validado com um conjunto não utilizado no treinamento, analisando seu desempenho e capacidade de generalização.

Indicadores de desempenho, como precisão e revocação, foram estabelecidos para monitorar sua eficácia contínua. O modelo foi mantido sob revisão constante, ajustando-se à evolução da medicina e integrando novas pesquisas e dados. Conclui-se que a pesquisa demonstrou a viabilidade da IA em direcionar eficientemente pacientes à especialidade médica correspondente, com a manutenção e atualização do modelo sendo essenciais para sua aplicação contínua e eficaz.

RESULTADOS

O modelo computacional desenvolvido apresentou um desempenho com 62% de precisão, 46% de revocação e 53% de F1-Score ao identificar especialidades médicas usando dados clínicos e registros médicos, incluindo de familiares. A abordagem reflete situações reais, onde nem sempre é necessário um especialista. Esse avanço em IA aplicada à medicina evidencia a potencialidade de assistência aos profissionais na determinação de encaminhamentos, otimizando tratamentos e melhorando a eficácia em serviços de saúde.

A análise dos erros destacou dificuldades do modelo em casos de sobreposição de sintomas entre especialidades. Fatores como histórico médico familiar, exames específicos e idade do paciente foram cruciais para o desempenho do modelo. A adição de dados familiares melhorou significativamente a precisão.

Durante o desenvolvimento, enfrentou-se desafios como qualidade e disponibilidade de dados. Alguns registros estavam incompletos ou desatualizados e a obtenção de dados familiares foi limitada por questões de privacidade. Outra dificuldade foi a diversidade das condições médicas. Condições com sintomas similares podem resultar em erros de indicação pelo modelo. É vital melhorar o modelo para enfrentar essa complexidade e diversidade, assim como atualizá-lo conforme avanços médicos.

VALIDAÇÃO

O processo de validação do modelo computacional proposto, abordando as técnicas adotadas para assegurar a generalização e confiabilidade dos resultados, bem como as medidas para mitigar ameaças à validade.

O modelo foi validado usando diversas técnicas para garantir a confiabilidade e generalização dos resultados. Estas incluíram:

- Validação cruzada: Durante o treinamento e avaliação, os dados foram divididos em conjuntos de treino e teste, permitindo verificar o desempenho do modelo em diferentes subconjuntos e reduzindo o risco de superajuste (Kohavi, 1995).
- Teste com dados externos: Avaliou-se a validade externa do modelo, testando seu desempenho em dados externos de diferentes fontes e contextos, examinando sua capacidade de generalização (Weiss e Provost, 2003).
- Análise de sensibilidade e especificidade: Esta análise foi realizada para verificar a eficiência do modelo na correta identificação das especialidades médicas e na exclusão das não pertinentes (Powers, 2011).

Para mitigar as ameaças à validade, algumas medidas foram adotadas:

- Dados representativos: Dados de pacientes e familiares foram coletados para serem representativos da população-alvo, assegurando aplicabilidade ampla (Zhou, Menche, Barabási e Sharma, 2014).
- Avaliação de algoritmos e pré-processamento: Vários algoritmos de aprendizado supervisionado foram avaliados, e técnicas de pré-processamento foram utilizadas para reduzir erros (Kotsiantis, Zaharakis e Pintelas, 2007).
- Validade externa: A validação com dados externos e atenção às especificidades locais permitiram avaliar e ampliar a generalização do modelo (Smith e Seaman, 2013).

Os resultados da validação reafirmaram o robusto desempenho do modelo, mostrando métricas consistentes em diferentes dados e contextos. Isso sugere a capacidade do modelo de indicar especialidades médicas adequadas para variados pacientes (Rajkomar, Dean e Kohane, 2019).

O processo de validação ressaltou a capacidade de generalização e confiabilidade do modelo. O uso de técnicas variadas e a mitigação de ameaças ampliaram a robustez e a aplicabilidade. A implementação eficaz em ambientes clínicos reais pode exigir integração com sistemas de apoio à decisão e adaptação às especificidades locais, considerando o feedback dos profissionais de saúde.

Em conclusão, com uma validação rigorosa, o modelo tem potencial para melhorar encaminhamentos médicos, elevando a qualidade do atendimento. Contudo, pesquisas e desenvolvimentos contínuos são essenciais para garantir sua evolução e relevância duradoura.

CONSIDERAÇÕES FINAIS

Ao chegarmos ao fim desta pesquisa, é imprescindível refletirmos sobre os insights e as lições aprendidas, assim como delinear perspectivas para estudos futuros e a aplicabilidade prática dos resultados.

Ao longo deste trabalho, esforçamo-nos para explorar e validar o modelo computacional proposto, enfrentando desafios, superando obstáculos e chegando a resultados significativos que atestam a viabilidade e relevância da abordagem adotada. O processo de validação do modelo, em especial, deu-nos uma visão clara de sua eficácia e das potenciais áreas de melhoria.

Um dos pontos mais relevantes desta pesquisa foi a importância da integração de diferentes técnicas e a necessidade de se ter uma abordagem holística no desenvolvimento e validação de modelos computacionais. A combinação de validação cruzada, testes com dados externos e análises de sensibilidade e especificidade mostrou-se fundamental para garantir um modelo robusto e aplicável em variados contextos clínicos.

Ainda, ficou evidente que, por mais avançados que sejam os modelos, a interação e o feedback dos profissionais de saúde são cruciais para garantir sua relevância e eficácia no mundo real. Tecnologia e expertise humana devem caminhar juntas para maximizar os benefícios dos modelos computacionais na área da saúde.

Além disso, é essencial considerar as implicações éticas e sociais da implementação de tais modelos. Enquanto as vantagens de otimização e eficiência são evidentes, a humanização

do atendimento médico e o respeito à individualidade do paciente não devem ser negligenciados.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAHAM, A.; U. C. S. Department, O. S. Meta-learning evolutionary artificial neural networks. 2003.
- ALPAYDIN, E. Introduction to machine learning. MIT press, 2020.
- MELO, B. A. P. et al. Qualificação profissional e humanização: Um estudo de caso sobre a qualidade no atendimento do setor de saúde no município de Abaetetuba-pa. Colóquio Organizações, Desenvolvimento e Sustentabilidade, v. 11, 2021.
- MAHESH, B. Machine learning algorithms-a review. International Journal of Science and Research (IJSR), v. 9, 2020.
- SOUZA, C. R.; BOTAZZO, C. Construção social da demanda em saúde. Physis: Revista de Saúde Coletiva, jun. 2013.
- MACHADO, C. V.; BAPTISTA, T. W. d. F.; LIMA, L. D. de. Políticas de saúde no Brasil: continuidades e mudanças. SciELO - Editora FIOCRUZ, jan. 2012.
- COHEN, R. et al. Redundancy-aware topic modeling for patient record notes. PLOS ONE, 2014.
- PONS, E. et al. Natural language processing in radiology: A systematic review. Radiology, v. 279, n. 2, maio 2016.
- JIANG, F. et al. Artificial intelligence in healthcare: past, present and future. BMJ, v. 2, jun. 2017.
- RICCI, F.; ROKACH, L.; SHAPIRA, B. Introduction to recommender systems handbook. Springer, Boston, MA, 2011.
- SMITH, G. C.; SEAMAN, S. R. The application of machine learning to problems in medicine. The Lancet Digital Health, v. 1, n. 2, p. e56–e67, 2013.
- HEINRICH, G. Parameter estimation for text analysis. Relatório técnico, Citeseer, 2005.
- WEISS, G. M.; PROVOST, F. Learning when training data are costly: the effect of class distribution on tree induction. Journal of Artificial Intelligence Research, v. 19, 2003.
- SUZUKI, K. Artificial neural networks: Methodological advances and biomedical applications. BoD – Books on Demand, abr. 2011.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Em: Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2, 1995.

MACHADO, M. H. Os médicos no Brasil: um retrato da realidade. Editora Fiocruz, 1997.

PRICE, W. N.; COHEN, I. G. Privacy in the age of medical big data. Nature medicine, v. 25, n. 1, p. 37—43, jan. 2019.

ZHANG, R. et al. Evaluating measures of redundancy in clinical texts. AMIA ... Annual Symposium proceedings. AMIA Symposium, 2011.

RAJKOMAR, A.; DEAN, J.; KOHANE, I. Machine learning in medicine. New England Journal of Medicine, v. 380, n. 14, p. 1347–1358, 2019.

ZHOU, X. et al. Human symptoms-disease network. Nature Communications, v. 5, 2014.

HARARI, Y. N. 21 lições para o século 21. Ago. 2018.