

Desenvolvimento de Modelos Interpretáveis de Aprendizagem
de Máquina Aplicáveis na Predição de Prognóstico de Pacientes
com Doença de Chagas

Daniel Oliveira dos Santos

Março, 2024

Resumo

A doença de Chagas, ou tripanossomíase americana, afeta entre 6 e 7 milhões de pessoas, principalmente na América Latina, incluindo Brasil e México. A infecção é causada pelo protozoário *Trypanosoma cruzi*, transmitido por insetos da subfamília Triatominae. A doença evolui em duas fases: aguda e crônica. Cerca de 30% dos indivíduos na fase crônica desenvolvem complicações cardíacas, como arritmias e insuficiência cardíaca, que podem resultar em morte súbita se não tratadas. A insuficiência cardíaca é uma das principais causas de morte em pacientes com doença de Chagas.

Medicamentos disponíveis podem melhorar a qualidade de vida e reduzir a mortalidade decorrente da insuficiência cardíaca associada à doença de Chagas. Contudo, fora dos centros urbanos, a falta de serviços de saúde especializados dificulta o diagnóstico precoce. O exame eletrocardiográfico, amplamente disponível, é uma ferramenta útil para detectar sinais de insuficiência cardíaca.

Diante desse contexto, este projeto propõe desenvolver algoritmos de aprendizado de máquina para identificar automaticamente sinais de insuficiência cardíaca em pacientes com doença de Chagas, utilizando dados de exames eletrocardiográficos de longa duração (Holter). A abordagem inclui técnicas de processamento digital de sinais, extração automática de parâmetros e aprendizado de máquina, abrangendo tanto modelos clássicos quanto modernos. Os dados para este estudo vêm do Hospital Universitário Clementino Fraga Filho (HUCFF/UFRJ) e incluem registros de eletrocardiogramas e dados clínicos de 219 pacientes com cardiomiopatia chagásica.

O objetivo central do projeto é comparar diferentes métodos computacionais para identificar insuficiência cardíaca nesses pacientes. Essa análise busca aprimorar o diagnóstico e, consequentemente, melhorar a qualidade do tratamento. O projeto também pretende fornecer uma solução adequada para ambientes de saúde com recursos limitados, facilitando diagnósticos precoces e reduzindo os riscos para pacientes com doença de Chagas.

Introdução

A doença de Chagas é uma infecção parasitária que pode levar a sérias complicações cardíacas, incluindo insuficiência cardíaca e morte súbita cardíaca (MSC). O diagnóstico precoce dessas complicações em pacientes com doença de Chagas

é fundamental para evitar desfechos fatais. No entanto, muitas áreas onde a doença é mais prevalente sofrem com a falta de infraestrutura médica adequada, dificultando diagnósticos precisos e oportunos.

Este projeto busca criar um modelo de aprendizado de máquina para analisar dados de eletrocardiogramas e identificar sinais de insuficiência cardíaca e outros problemas cardíacos associados à doença de Chagas. O modelo também pretende detectar padrões relacionados a várias condições clínicas, como a ruptura de cordoalhas cardíacas, uso de dispositivos implantáveis (como marcapassos e CDI), e eventos como morte súbita cardíaca, ablação cardíaca, bem como condições associadas, como psoríase, diverticulite, amiloidose cutânea e acidente vascular cerebral (AVC).

Os objetivos específicos do projeto incluem:

- **Deteção de Insuficiência Cardíaca:** Identificar indicadores de insuficiência cardíaca em pacientes com doença de Chagas utilizando eletrocardiogramas, especialmente dados de longa duração.
- **Identificação de Morte Súbita Cardíaca (MSC):** Descobrir padrões associados a eventos de morte súbita cardíaca, permitindo respostas mais rápidas e melhor prognóstico para pacientes em risco.
- **Análise de Condições Relacionadas:** Explorar a relação entre outras condições médicas, como AVC, psoríase e diverticulite, e complicações cardíacas em pacientes com doença de Chagas.

Para alcançar esses objetivos, o modelo usará algoritmos de aprendizado de máquina aplicados a um conjunto de dados clínicos que incluem registros de eletrocardiogramas, diagnósticos médicos, procedimentos cardíacos e resultados como morte súbita cardíaca, implantações de marcapassos ou CDI, além de causas de morte, como AVC.

Com essa abordagem de aprendizado de máquina, espera-se melhorar tanto a precisão do diagnóstico quanto a qualidade do tratamento, além de oferecer insights valiosos que possam contribuir para a redução de eventos fatais em pacientes com doença de Chagas e outras condições relacionadas.

Justificativa

A doença de Chagas é uma infecção parasitária que impacta milhões de pessoas na América Latina, principalmente no Brasil e no México. É uma das principais causas de cardiomiopatia e insuficiência cardíaca na região, com cerca de 30% a 40% dos pacientes com a forma crônica desenvolvendo complicações cardíacas, como arritmias, insuficiência cardíaca, aneurismas ventriculares e tromboembolismo. A insuficiência cardíaca decorrente da doença de Chagas é uma das principais causas de morte nesses pacientes, com um alto risco de morte súbita cardíaca.

Apesar de medicamentos acessíveis poderem ajudar a controlar a doença, o diagnóstico precoce e o tratamento oportuno são limitados devido à escassez de serviços de saúde especializados, especialmente em regiões fora dos centros urbanos. Esta falta de infraestrutura e recursos contribui para um aumento no número de casos não diagnosticados e, consequentemente, para um maior risco de complicações fatais.

Dado que a detecção precoce é crítica para prevenir a progressão da doença e melhorar os resultados clínicos, há uma necessidade significativa de métodos inovadores para facilitar o diagnóstico em ambientes com recursos limitados. O uso de tecnologias modernas, como o aprendizado de máquina, permite analisar grandes volumes de dados médicos, proporcionando uma oportunidade de aprimorar a identificação de padrões associados à insuficiência cardíaca e outras complicações cardíacas relacionadas à doença de Chagas.

Portanto, o desenvolvimento de um modelo de aprendizado de máquina para detectar sinais precoces de complicações cardíacas em pacientes com doença de Chagas é justificado. Esta abordagem pode oferecer uma solução eficaz para melhorar o diagnóstico precoce, facilitar o tratamento oportuno e, por fim, reduzir as taxas de mortalidade associadas à doença de Chagas e suas complicações cardíacas.

Objetivos

O objetivo geral deste projeto é desenvolver modelos de aprendizado de máquina para detectar insuficiência cardíaca em pacientes com doença de Chagas, utilizando dados clínicos e registros de monitoramento cardíaco de 24 horas. Para atingir esse objetivo, os seguintes objetivos específicos foram identificados:

- Organizar um banco de dados: Reunir e estruturar dados clínicos e registros de monitoramento cardíaco de pacientes com doença de Chagas, fornecidos pelo Hospital Universitário Clementino Fraga Filho da Universidade Federal do Rio de Janeiro (HUCFF-UFRJ).
- Desenvolver modelos de aprendizado de máquina: Aplicar técnicas de aprendizado de máquina para prever insuficiência cardíaca em pacientes com doença de Chagas. Testar diferentes algoritmos, como Redes Neurais, Florestas Aleatórias e Máquinas de Vetores de Suporte, para identificar os modelos mais eficazes.
- Avaliar métricas de desempenho: Analisar a precisão, sensibilidade, especificidade, entre outras métricas, para avaliar a eficácia dos modelos desenvolvidos.
- Analisar cenários de classificação: Avaliar a eficácia dos modelos em diferentes cenários, utilizando registros de monitoramento cardíaco de várias durações para determinar quais configurações oferecem maior precisão na detecção de insuficiência cardíaca.

Metodologia

A metodologia para atingir esses objetivos inclui os seguintes passos:

- Coleta e organização de dados: Criar um banco de dados robusto contendo informações clínicas e registros de monitoramento cardíaco de pacientes com doença de Chagas. A amostra consiste em 308 pacientes, totalizando 385 registros de monitoramento cardíaco, abrangendo um ou mais dias de exames por paciente e desfechos associados (como morte súbita, implante de marcapasso, CDI, entre outros).

- Desenvolvimento e teste de modelos de aprendizado de máquina: Testar uma variedade de algoritmos de aprendizado de máquina para criar modelos preditivos de insuficiência cardíaca. Utilizar métodos como Redes Neurais, Florestas Aleatórias, e Máquinas de Vetores de Suporte para identificar quais são mais eficazes. Os dados serão organizados em janelas de diferentes tamanhos para analisar qual abordagem resulta em melhor precisão na previsão de insuficiência cardíaca.
- Avaliação e validação dos modelos: Aplicar métricas de desempenho para validar a precisão dos modelos e identificar quais são mais eficazes para a detecção de insuficiência cardíaca. As métricas incluem acurácia, precisão, sensibilidade e especificidade.
- Rotulação manual de registros de monitoramento cardíaco: Criar uma interface para rotular manualmente os registros, identificando características como ritmos cardíacos irregulares, bradicardia, taquicardia e outros eventos adversos. Essa rotulação será usada para treinar e validar os modelos de aprendizado de máquina.

Resultados e Impactos Esperados

O projeto tem como objetivo desenvolver um modelo de aprendizado de máquina capaz de identificar sinais de insuficiência cardíaca em pacientes com doença de Chagas. Se bem-sucedido, este modelo pode gerar resultados significativos, entre eles:

- Melhora na detecção precoce de insuficiência cardíaca: Um modelo eficaz pode facilitar a identificação precoce de insuficiência cardíaca em pacientes com doença de Chagas, permitindo uma intervenção mais rápida e eficaz.
- Ferramenta para ambientes de saúde com recursos limitados: Ao criar um modelo que utiliza dados acessíveis e algoritmos eficazes, espera-se fornecer uma ferramenta que possa ser implementada em ambientes de saúde com recursos limitados, onde o acesso a diagnósticos complexos é mais difícil.
- Redução das complicações associadas à doença de Chagas: Ao identificar precocemente a insuficiência cardíaca, o modelo pode ajudar a reduzir o número de complicações graves, como arritmias e tromboembolismo, que frequentemente resultam em desfechos fatais.
- Diminuição da taxa de mortalidade: Ao melhorar a detecção e o tratamento de insuficiência cardíaca em pacientes com doença de Chagas, o projeto pode contribuir para a redução da taxa de mortalidade, resultando em melhores resultados para os pacientes.
- Avanços na pesquisa em aprendizado de máquina para a saúde: O desenvolvimento deste modelo também pode servir como base para futuros trabalhos na aplicação de aprendizado de máquina em outras doenças cardíacas ou problemas de saúde, expandindo o campo de aplicação desta tecnologia.

Esses resultados e impactos esperados apontam para um potencial significativo na melhoria da qualidade de vida dos pacientes com doença de Chagas e oferecem uma oportunidade para aplicar a inovação tecnológica para enfrentar problemas de saúde críticos em regiões com infraestrutura médica limitada.

Plano de Atividades

Etapa 1: Organização do Banco de Dados

Reunir e organizar os dados clínicos de pacientes com doença de Chagas fornecidos pelo Hospital Universitário Clementino Fraga Filho da Universidade Federal do Rio de Janeiro (HUCFF-UFRJ). Definir critérios para classificação dos pacientes e desfechos associados.

Etapa 2: Processamento de Dados

Aplicar algoritmos de pré-processamento para limpar e estruturar os dados. Extrair parâmetros essenciais para análise e caracterização dos registros de monitoramento cardíaco.

Etapa 3: Desenvolvimento de Modelos de Aprendizado de Máquina

Testar diferentes algoritmos de aprendizado de máquina para prever insuficiência cardíaca em pacientes com doença de Chagas. Experimentar diferentes métodos para segmentação de dados, identificando abordagens que melhorem a eficácia do modelo.

Etapa 4: Avaliação e Validação dos Modelos

Calcular métricas de desempenho, como precisão, sensibilidade e especificidade, para avaliar a eficácia dos modelos desenvolvidos. Identificar quais algoritmos ou configurações têm melhor desempenho na detecção de insuficiência cardíaca.

Etapa 5: Implementação do Processo de Rotulação

Desenvolver uma interface para rotular manualmente os dados clínicos, estabelecendo padrões para futuras análises. Usar algoritmos de aprendizado de máquina não supervisionados para auxiliar no processo de rotulação, identificando padrões que podem não ser imediatamente visíveis.

Etapa 6: Finalização do Projeto

Analisar os resultados obtidos durante o desenvolvimento do projeto e compilar a documentação final. Apresentar conclusões e recomendações para futuras pesquisas e possíveis aplicações práticas do modelo desenvolvido.

Assinatura do Candidato

Assinatura do Orientador

Assinatura do Orientador