



Engenharia de Computação

Tema: Análise Formal de Segurança Crítica (*safety*) de Inteligência Artificial em Sistemas de Detecção de Arritmias Cardíacas: Diferenciação entre Batimentos Saudáveis e Não Saudáveis.

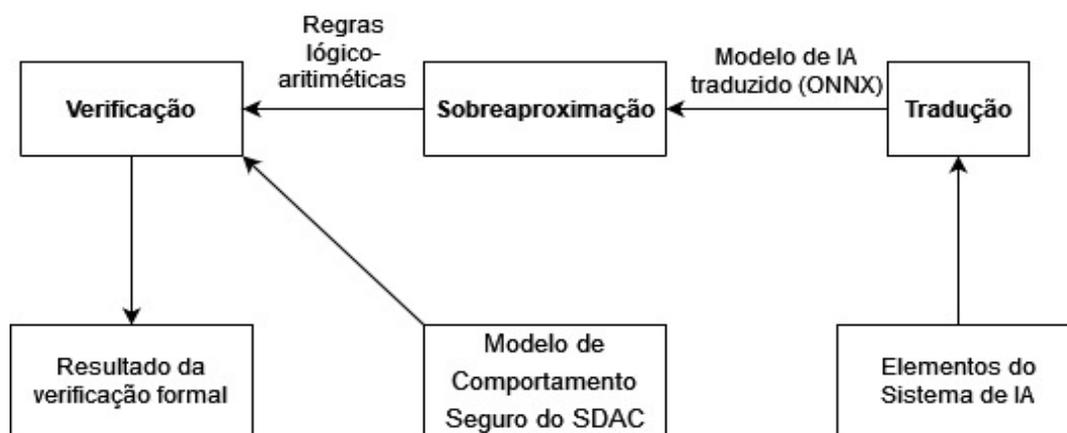
A crescente integração de sistemas baseados em inteligência artificial (IA) em aplicações críticas, como diagnósticos médicos, reforça a importância de garantir a segurança (*safety*) desses sistemas. Aproveitando-se de lacunas identificadas em pesquisas dessa área, o aluno Gabriel Stephano Santos desenvolveu seu trabalho de conclusão de curso visando analisar formalmente o uso de IA em aplicações médicas.

O objetivo desse trabalho foi explorar a avaliação formal de segurança de um sistema de detecção de arritmias cardíacas (SDAC) baseado em redes neurais residuais profundas. As atividades foram conduzidas sob a orientação do Prof. Dr. Paulo Sergio Cugnasca e co-orientação do Prof. Dr. Antonio Vieira da Silva Neto.

A estratégia de análise inicial previa (i.) traduzir a rede neural para um padrão da área, denominado ONNX (*Open Neural Network Exchange*), (ii.) estimar saídas garantidamente seguras por uma técnica chamada de Sobreaproximação e (iii.) verificar se essas saídas de fato atendem às restrições de segurança. Contudo, devido às limitações de ferramentas, o passo (ii.) não pôde ser finalizado.

Como alternativa de análise formal, foi aplicada uma técnica que busca identificar situações sabidamente inseguras. Essa técnica, denominada ataque adversário, permitiu comprovar cenários inseguros; todavia, várias dessas situações são irreais por não representarem batimentos cardíacos típicos de um ser humano – seja ele saudável ou não.

Os próximos passos preveem contornar as limitações das ferramentas usadas, analisar outras arritmias e comparar a IA com normas de segurança e equipamentos comerciais.



Integrante: Gabriel Stephano Santos

Professor Orientador: Prof. Dr. Paulo Sérgio Cugnasca

Co-orientador: Prof. Dr. Antonio Vieira da Silva Neto