



## IMPLEMENTAÇÃO DE MODELOS DE LINGUAGEM PARA EXTRAÇÃO DE DIAGNÓSTICOS EM NOTAS CLÍNICAS

QualiJuntos, 1ª edição, de 30/09/2025 a 01/10/2025

ISBN dos Anais: 978-65-5465-161-5

**REBELO; Marina de Fatima de Sá<sup>1</sup>, GUMIEL; Yohan<sup>2</sup>, CRUZ; Gustavo<sup>3</sup>, MORENO; Ramon Alfredo<sup>4</sup>, BARRA; Cláudia Cabral Moro<sup>5</sup>, KRIEGER; Jose Eduardo<sup>6</sup>, GUTIERREZ; Marco Antonio<sup>7</sup>**

### RESUMO

Introdução A adoção de Registros Eletrônicos de Saúde (RES) transformou a prática médica, ampliando o acesso a grandes volumes de dados clínicos. Uma parte significativa dessas informações está registrada em narrativas livres — como evoluções de ambulatório, laudos e anotações de prontuário. Por serem dados não estruturados, sua utilização direta em pesquisas epidemiológicas ou em sistemas de apoio à decisão é extremamente difícil. Nos últimos anos, os modelos de linguagem de grande porte (Large Language Models – LLMs) vêm alcançando resultados expressivos na análise desse tipo de dado. Eles conseguem reconhecer padrões, organizar informações e resumir textos com eficácia. Na área médica, uma vantagem relevante é a capacidade de lidar com linguagem heterogênea e uso frequente de abreviações, características comuns em narrativas clínicas. Esse diferencial abre novas perspectivas para sua aplicação na prática assistencial e na pesquisa. Ainda assim, permanecem desafios, como interpretações incorretas (“alucinações”) e a necessidade de garantir a privacidade dos pacientes em conformidade com a Lei Geral de Proteção de Dados (LGPD). Objetivo Avaliar a aplicação de LLMs na extração automática de diagnósticos a partir de evoluções ambulatoriais, com foco em cardiologia. O tema é particularmente relevante porque muitos diagnósticos não estão registrados nos campos estruturados dos sistemas de informação, mas apenas em textos livres, que podem conter informações valiosas para uma compreensão mais abrangente da condição clínica dos pacientes. Método utilizado Partimos de modelos de código aberto de última geração e aplicamos técnicas de adaptação ao domínio da saúde, combinando ajuste fino supervisionado com dados anotados por especialistas e pré-treinamento adicional em grandes volumes de textos médicos não anotados. Esse processo resultou em modelos com conhecimento específico, avaliados quanto à sua capacidade de identificar diagnósticos em evoluções ambulatoriais. O desenvolvimento ocorreu inteiramente em infraestrutura de alto desempenho localizada na Instituição, adquirida especialmente para aplicações de inteligência artificial em saúde. Essa decisão assegurou maior aderência às normas de segurança e privacidade e também possibilitou o uso de modelos mais robustos, com maior número de parâmetros. Foram explorados modelos de diferentes famílias, como LLaMA, GPT-

<sup>1</sup> Instituto do Coração - HCFMUSP, marina.rebelo@hc.fm.usp.br

<sup>2</sup> Instituto do Coração - HCFMUSP, yohan.gumiel@hc.fm.usp.br

<sup>3</sup> Instituto do Coração - HCFMUSP, gustavo.cruz@hc.fm.usp.br

<sup>4</sup> Instituto do Coração - HCFMUSP, ramon.moreno@hc.fm.usp.br

<sup>5</sup> Instituto do Coração - HCFMUSP, claudia.moro@hc.fm.usp.br

<sup>6</sup> Instituto do Coração - HCFMUSP, j.krieger@hc.fm.usp.br

<sup>7</sup> Instituto do Coração - HCFMUSP, marco.gutierrez@hc.fm.usp.br

OSS, DeepSeek e Gemma. Para viabilizar os experimentos, construímos um corpus próprio composto por 600 evoluções ambulatoriais devidamente anonimizadas, contendo os respectivos diagnósticos. Para esse fim, a equipe do projeto desenvolveu tanto o algoritmo de anonimização quanto uma ferramenta de pré-annotação baseada em LLM. Essa ferramenta, concebida para otimizar o tempo da equipe clínica, forneceu uma identificação inicial dos diagnósticos, posteriormente revisados e validados por especialistas médicos. O conjunto curado constituiu o padrão ouro utilizado para o treinamento e a avaliação dos modelos. Resultados Os testes foram conduzidos com 89 textos de evolução clínica. Os modelos apresentaram desempenho promissor, mesmo sem treinamento prévio dedicado exclusivamente ao domínio médico. O melhor resultado foi obtido com o modelo GPT-OSS. De forma simplificada, a taxa média de concordância entre os diagnósticos identificados automaticamente pelo modelo e aqueles registrados por especialistas foi de aproximadamente 70%. Medidas complementares de similaridade textual superaram 85%, indicando que, na maior parte dos casos, o modelo reconheceu os diagnósticos de forma precisa e consistente. Discussão Os resultados confirmam o potencial dos LLMs para lidar com dados clínicos não estruturados, mas também expõem limitações relevantes. A variação de desempenho entre diferentes casos sugere a necessidade de estratégias adicionais para reduzir erros e aumentar a confiabilidade. Além disso, persiste o desafio da integração desses modelos a sistemas de saúde já existentes, de forma segura e alinhada às normas de proteção de dados. Ainda assim, a possibilidade de extrair diagnósticos de forma automatizada a partir de notas clínicas representa um avanço importante para pesquisa e assistência. Conclusão Este estudo demonstra que os LLMs podem estabelecer um novo patamar para a extração de informações clínicas a partir de dados não estruturados, oferecendo soluções viáveis para organizar textos médicos em contextos especializados. Os resultados iniciais, obtidos em português e em ambiente seguro, mostram que essa tecnologia pode apoiar tarefas complexas, como a análise automatizada de diagnósticos em cardiologia. Embora persistam desafios, esta aplicação piloto reforça a viabilidade da abordagem e abre caminho para novas frentes, como a sumarização de prontuários e a construção longitudinal da jornada clínica dos pacientes.

**PALAVRAS-CHAVE:** Processamento de Linguagem Natural, Modelos de Linguagem, Diagnósticos Clínicos, Inteligência Artificial, Notas Clínicas

<sup>1</sup> Instituto do Coração - HCFMUSP, marina.rebelo@hc.fm.usp.br

<sup>2</sup> Instituto do Coração - HCFMUSP, yohan.gumiel@hc.fm.usp.br

<sup>3</sup> Instituto do Coração - HCFMUSP, gustavo.cruz@hc.fm.usp.br

<sup>4</sup> Instituto do Coração - HCFMUSP, ramon.moreno@hc.fm.usp.br

<sup>5</sup> Instituto do Coração - HCFMUSP, claudia.moro@hc.fm.usp.br

<sup>6</sup> Instituto do Coração - HCFMUSP, j.krieger@hc.fm.usp.br

<sup>7</sup> Instituto do Coração - HCFMUSP, marco.gutierrez@hc.fm.usp.br