



UNIVERSIDADE FEDERAL DE MINAS GERAIS



Plano de Trabalho do projeto: PediAI: Uma ferramenta para auxiliar estudantes e pediatras processo de raciocínio clínico e diagnóstico utilizando grandes modelos de linguagem

Projeto Tipo B

Coordenador: Eduardo Araujo de Oliveira

Instituição do Coordenador: Universidade Federal de Minas Gerais

1. DADOS DAS PARTÍCIPES

Identificação da Partípice 1 - UFMG

Razão Social: Universidade Federal de Minas Gerais (UFMG)

CNPJ: 17.217.985/0001-04.

Endereço: Av. Antônio Carlos, 6627, Pampulha, CEP 31.270-901, Belo Horizonte, MG.

Representante legal: Sandra Regina Goulart Almeida.

Cargo: Reitora.

Telefone: (31) 3409-4124.

E-mail: chefia@gabinete.ufmg.br.

Identificação da Partípice 2 - Unimed BH

Razão Social: Unimed BH

CNPJ: 16.513.178/0001-76

Endereço: Rua dos Inconfidentes, 44, Funcionários, BH/MG, CEP: 30.140-120

Representante legal: Sr. Frederico José Amédée Péret

Cargo: Diretor Presidente

Telefone: 3229-6021

E-mail: frederico.peret@unimedbh.com.br

Identificação da Partípice 3 – Instituto Federal de Minas Gerais

Razão Social: Instituto Federal de Ciência e Tecnologia de Minas Gerais

CNPJ: 10.626.896/0001-72

Endereço: Avenida Professor Mário Werneck, 2590, Buritis, Belo Horizonte - MG

Representante legal: Rafael Bastos Teixeira

Cargo: Reitor

Telefone: (31) 2513-5179

E-mail: gabinete@ifmg.edu.br

Identificação Fundação de Apoio - FUNDEP

Razão Social: Fundação de Desenvolvimento da Pesquisa - FUNDEP

CNPJ: 18.720.938/0001-41

Endereço: Av. Antônio Carlos, 6627, Pampulha, CEP 31.270-901, Belo Horizonte, MG.

Representante legal: Jaime Arturo Ramirez

Cargo: Presidente

Telefone: (31) 3443-6804

E-mail: fundepatende@fundep.com.br

COORDENAÇÃO DO PROJETO

Nome: Eduardo Araujo de Oliveira

Instituição: Universidade Federal de Minas Gerais

Unidade: Divisão de Nefrologia Pediátrica

Departamento: Pediatria

Nível (Titular/Associado/Adjunto/...): Titular

SIAPE:

CPF: 51543311687

Telefone: 3134099773

E-mail: eduolive812@gmail.com

Nome do coordenador técnico científico responsável pelas atividades por parte da **Partípice 3**: Adriano Lages dos Santos

IDENTIFICAÇÃO DO PROJETO

PediAI: Uma ferramenta para auxiliar estudantes e pediatras processo de raciocínio clínico e diagnóstico utilizando grandes modelos de linguagem

Tipo: Inovação

Área de conhecimento: Inteligência Artificial; Medicina; Pediatria; Clínica Médica.

CARACTERIZAÇÃO DO PROJETO

Contextualização do problema a ser resolvido e análise do estado da arte

O diagnóstico médico de sepsis é um processo importante para o cuidado à saúde, exigindo expertise, conhecimento médico aprofundado e a capacidade de integrar diversas informações do paciente. Apesar da experiência e do conhecimento dos médicos, o processo de diagnóstico desta doença pode ser desafiador e sujeito a erros, assim como em casos de outras doenças raras ou complexas [1, 2]. Muitos médicos já ouviram a seguinte frase: "A clínica é soberana." Isso se deve ao fato de que o quadro clínico do paciente muitas vezes é mais importante que exames em caso de discordância entre os dados clínicos e os diversos exames solicitados pelo médico.

A sepse pediátrica é um problema de saúde pública que causa cerca de 3,3 milhões de mortes anualmente pelo mundo [3]. No entanto, os critérios atuais para diagnosticar sepse pediátrica, publicados em 2005 após a Conferência Internacional de Consenso sobre Sepse Pediátrica está desatualizada, tem baixa especificidade, não permite estratégias de risco fixação em ambientes com recursos mais baixos e mais altos, e pode ser discordante com o diagnóstico baseado no médico [3]. O diagnóstico precoce é importante. Os pacientes devem receber antibioticoterapia adequada o mais rápido possível. Culturas de sangue, bem como outras culturas de locais sob suspeita de infecção, devem ser colhidas na tentativa de detectar o agente causador da doença. A sepse é uma emergência médica e seu tratamento deve ser priorizado [4].

Nesse contexto, a inteligência artificial (IA) surge como uma ferramenta promissora para auxiliar os profissionais de saúde na tomada de decisões clínicas, especialmente no campo do diagnóstico médico. Os grandes modelos de linguagem (LLMs), como Llama, ChatGPT e Gemini, demonstram potencial significativo para processar e analisar grandes volumes de dados médicos, incluindo textos, imagens e registros eletrônicos de saúde, e identificar padrões que podem auxiliar no diagnóstico de doenças [4, 5].

Em estudo publicado em abril de 2024 na revista JAMA Internal Medicine [6], O ChatGPT demonstrou maior acurácia para realizar diagnósticos em consultas em comparação com diagnósticos realizados por clínicos. Foram comparadas as justificativas de diagnósticos feitos por médicos assistentes, estudantes e pelo ChatGPT e todos os diagnósticos realizados pelo LLM obtiveram nota máxima, ou seja, pontuação 10 pela ferramenta r-IDEA [7] enquanto médicos assistentes e residentes obtiveram notas 9 e 8 respectivamente.

Esses estudos demonstram o potencial dos LLMs para auxiliar no diagnóstico médico, contribuindo para a redução de erros diagnósticos. Além disso, são diversos os benefícios de um sistema que utiliza LLMs para realizar diagnóstico médico, principalmente em localidades remotas que contam com equipes de saúde reduzidas. Um sistema desta natureza poderia auxiliar médicos com uma segunda opinião e em caso de encaminhamento para hospitais em grandes centros os pacientes seriam encaminhados já com um relatório diagnóstico maior acurácia. Destaca-se o potencial de sistemas como este como ferramentas valiosas no aprendizado e treinamento de estudantes de medicina de graduação e residentes médicos. Complementares aos laboratórios de simulação do exame físico já existentes nas escolas médicas, há a oportunidade de se desenvolver laboratórios de coleta de história clínica e raciocínio clínico. Estes laboratórios poderão auxiliar no ensino da semiologia para graduandos e no treinamento de especialistas e residentes.

Este projeto de pesquisa busca formular a seguinte pergunta principal:

De que forma um sistema de suporte à decisão (SSD) que utiliza modelos de linguagem grande (LLMs) pode auxiliar estudantes de Medicina e médicos na realização de diagnósticos médicos de sepse com maior acurácia.

Hipótese:

A implementação de um SSD com LLMs pode contribuir para o aprimoramento do processo de diagnóstico médico de sepse, fornecendo sugestões de diagnósticos com maior probabilidade de acerto, permitindo a comparação entre os resultados da ferramenta e os diagnósticos realizados por estudantes e médicos, possibilitando a identificação de erros e o aprimoramento do julgamento clínico, além de contribuir para tratamentos mais rápidos e sistematizados.

Metodologia

3.2.1 População envolvida ou problema

O projeto visa auxiliar estudantes e médicos no processo de diagnóstico médico de sepse, utilizando Grandes Modelos de Linguagem (LLMs) como Llama, ChatGPT e Gemini. O sistema será disponibilizado de forma gratuita e terá como público-alvo:

Estudantes de medicina e médicos residentes de Pediatria: O sistema poderá auxiliar os estudantes na aprendizagem de semiologia pediátrica e diagnósticos médicos, fornecendo sugestões e comparando-as com seus próprios diagnósticos.

Médicos Pediatras: O sistema poderá auxiliar os médicos na tomada de decisões diagnósticas, fornecendo uma segunda opinião e ajudando a identificar pontos de aprimoramento.

O problema que o projeto se propõe a solucionar é a dificuldade de diagnóstico preciso e eficiente para sepse. A sepse é uma resposta séria e generalizada do corpo à bactеремia ou a outra infecção juntamente com o mau funcionamento ou insuficiência de um sistema essencial do organismo. Choque séptico é a pressão arterial perigosamente baixa (choque) e insuficiência orgânica com risco de morte devido à sepse. O diagnóstico incorreto ou tardio pode ter graves consequências para a saúde do paciente, como atrasos no tratamento, prescrição de medicamentos inadequados e até mesmo o agravamento da doença.

Esse projeto apresenta dois desafios, sendo um deles clínico e um computacional. O desafio computacional consiste em realizar o ajuste fino da LLM escolhida para a tarefa do projeto. O ajuste fino de LLMs com dados rotulados, ou seja, o ajuste fino supervisionado consome uma quantidade significativa de recursos computacionais. Entretanto, estudos recentes [8] mostram que o ajuste fino de modelos como o Llama foram suficientes para que as respostas deste modelo fossem preferidas por humanos em relação às respostas do ChatGPT versão 4. O desafio clínico vem na formação dos dados que serão organizados para o treinamento da LLM escolhida. Essa é uma tarefa complexa e que exige uma equipe médica qualificada para analisar os dados ou até mesmo criar exemplos de diagnósticos como exemplos de treinamento.

Requisitos:

Base de dados: O sistema deve ter acesso a uma base de dados de casos clínicos pediátricos, caso não consigamos uma base de dados de casos clínicos, o corpo de médicos desse projeto pode construir casos clínicos para treinamento do LLM.

Reconhecimento de padrões: O sistema deve utilizar técnicas de inteligência artificial para reconhecer padrões nos dados clínicos e identificar características relevantes para o diagnóstico.

Suprimento à tomada de decisão: O sistema deve fornecer aos pediatras suporte à tomada de decisão durante o exame físico, sugerindo possível diagnóstico de sepse e exames complementares a serem realizados.

Simulação de cenários clínicos: O sistema deve permitir que os pediatras pratiquem a semiologia em cenários

clínicos simulados, recebendo feedback personalizado sobre seu desempenho.

Aprendizado personalizado: O sistema deve se adaptar às necessidades individuais de cada pediatra, oferecendo conteúdo e atividades personalizadas de acordo com seu nível de conhecimento e experiência.

3.2.2 Questões éticas

O estudo proposto envolve questões éticas importantes relacionadas à privacidade e segurança dos dados dos pacientes, bem como à responsabilidade do uso de inteligência artificial na prática médica. A garantia de anonimidade de dados dos pacientes deve ser garantida para que não haja violações de princípios éticos individuais.

Além disso, é fundamental abordar as questões éticas relacionadas à IA Responsável. Isso inclui preocupações com o uso de dados não representativos da população, possíveis vieses nos modelos de linguagem que podem surgir de contextos humanos e valores sociais, bem como suposições sobre gênero, raça e geografia que podem influenciar os diagnósticos. Para mitigar esses riscos, serão implementadas técnicas de diversificação de conjuntos de dados e avaliação rigorosa dos modelos quanto a vieses e equidade. É necessário definir claramente quem é responsável por decisões diagnósticas incorretas ou por danos causados pelo sistema. Os dados dos pacientes devem ser armazenados e utilizados de forma segura e confidencial, de acordo com as leis e regulamentações de proteção de dados.

Este projeto estará de acordo com as determinações da Anvisa para softwares como dispositivos médicos - SAMD que diz respeito à privacidade e segurança dos dados e no que diz respeito à infraestrutura de segurança. Essas diretrizes de segurança se encontram em uma resolução da diretoria colegiada da Anvisa - RDC - RDC Nº 657, DE 24 DE MARÇO DE 2022 [8]. Este projeto também estará alinhado às diretrizes propostas no Projeto de Lei 2338 de 2023 [9] que estabelece normas gerais de caráter nacional para o desenvolvimento, implementação e uso responsável de sistemas de inteligência artificial (IA) no Brasil, com o objetivo de proteger os direitos fundamentais e garantir a implementação de sistemas seguros e confiáveis, em benefício da pessoa humana, do regime democrático e do desenvolvimento científico e tecnológico.

3.2.3 Intervenção, exposição ou solução proposta

A solução proposta consiste no desenvolvimento de um protótipo de sistema de suporte à decisão para diagnóstico médico que utilize LLMs. O sistema funcionará da seguinte forma:

- O usuário irá inserir os dados do paciente no sistema, como anamnese, exame físico, história pregressa, estado vacinal, uso de medicamentos, exames e outros dados relevantes, como história familiar e epidemiologia.
- O sistema utilizará os LLMs para analisar os dados do paciente e fornecer sugestões diagnósticas.
- O sistema também apresentará a probabilidade de cada diagnóstico, permitindo que o usuário compare os resultados do LLM com seu próprio diagnóstico.
- A ferramenta IDEA será integrada ao sistema para validação dos diagnósticos fornecidos pelos LLMs.
- O sistema estará disponível de forma gratuita para estudantes e médicos.

É importante destacar que o LLM utilizado no sistema sofrerá ajuste fino (fine tuning). Os dados para o ajuste fino serão diagnósticos médicos corretos feitos por humanos.

3.2.4 Desfechos esperados

- Contribuição para o ensino da semiologia médica, permitindo que professores desenvolvam instrumentos de aprendizagem e avaliação dos graduandos em medicina que iniciam o ciclo profissional da escola médica.
- Redução do tempo de diagnóstico: O sistema poderá auxiliar na redução do tempo necessário para o diagnóstico de sepse, permitindo que os pacientes recebam tratamento mais rapidamente.

- Aumento da precisão do diagnóstico: O sistema poderá auxiliar no aumento da precisão do diagnóstico de sepse, o que pode levar a uma redução do número de erros médicos e à melhora dos resultados para os pacientes.
- Melhora da qualidade de vida dos pacientes: O diagnóstico preciso e oportuno de sepse pode levar à melhora da qualidade de vida dos pacientes, reduzindo o tempo de internação hospitalar, o número de complicações e a mortalidade.
- Democratização do acesso ao diagnóstico médico: O sistema estará disponível de forma gratuita, o que permitirá que estudantes e médicos tenham acesso a uma ferramenta gratuita de diagnóstico. O acesso ao sistema será permitido somente para médicos e estudantes de medicina mediante controle de acesso para evitar acesso de outras pessoas que não atuam na área de saúde.

3.2.5 Métrica para avaliação dos desfechos

- Acurácia do diagnóstico: A acurácia do diagnóstico com e sem o uso do sistema será comparada. A nota será calculada com base na ferramenta IDEA.
- Tempo médio para o diagnóstico: O tempo médio necessário para o diagnóstico de doenças com e sem o uso do sistema será comparado.
- Deverão ser desenvolvidos ensaios clínicos para comparar a performance do uso desta tecnologia no ensino médico e no treinamento de residentes.
- Deverão ser desenvolvidos ensaios clínicos para comparar a performance do sistema no auxílio à tomada de decisões dos médicos.
- Deverão ser desenvolvidos ensaios clínicos para comparar o custo benefício da implantação do sistema como instrumento de ensino e na prática médica.

3.2.6 Desafios e riscos

Desenvolvimento de um protótipo de sistema de IA robusto e confiável: O desenvolvimento de um sistema de IA que seja capaz de reconhecer padrões complexos em dados clínicos e fornecer suporte à tomada de decisão é um desafio técnico significativo.

Garantia da segurança e da privacidade dos dados: Os dados clínicos dos pacientes são sensíveis e precisam ser protegidos de acessos não autorizados. O sistema de IA deve ser desenvolvido com segurança e privacidade em mente.

Integração com o fluxo de trabalho clínico: O sistema de IA precisa ser integrado ao fluxo de trabalho clínico de forma a ser útil e prático para os pediatras.

Avaliação da efetividade do sistema: A efetividade do sistema de IA precisa ser avaliada por meio de estudos rigorosos para garantir que ele realmente melhora a qualidade do diagnóstico e a segurança do paciente.

Garantir a qualidade e representatividade dos dados de treinamento: Os LLMs devem ser treinados em conjuntos de dados de alta qualidade e representativos da população real, a fim de evitar vieses algorítmicos que possam levar a diagnósticos incorretos ou discriminatórios.

Assegurar a transparência e explicabilidade do sistema: O sistema deve ser transparente e explicar aos usuários como chegou a cada diagnóstico, permitindo que os profissionais de saúde compreendam o raciocínio do LLM e possam tomar decisões informadas.

Garantir a segurança e confidencialidade dos dados dos pacientes: Os dados dos pacientes serão armazenados de forma anonimizada e utilizados de forma segura e confidencial, de acordo com as leis e regulamentações de proteção de dados.

Objetivos:

3.3.1 Objetivo Geral

Desenvolver um sistema de suporte à decisão para diagnóstico médico em Pediatria que utilize Grandes Modelos de Linguagem (LLMs) para auxiliar estudantes e médicos no processo de diagnóstico de sepse, com o objetivo de reduzir o tempo de diagnóstico, aumentar a precisão do diagnóstico e melhorar a chance de sobrevida dos pacientes.

3.3.2 Objetivos Específicos

1. Desenvolver um sistema de IA que auxilie estudantes de medicina e pediatras no aprendizado e aperfeiçoamento da semiologia pediátrica, a arte de interpretar sinais e sintomas clínicos para diagnosticar sepse em crianças.
2. Treinar LLMs em conjuntos de dados de alta qualidade e representativos da população real para minimizar o risco de vieses algorítmicos.
3. Desenvolver um sistema transparente e explicável que permita aos profissionais de saúde compreender o raciocínio dos LLMs e tomar decisões informadas.
4. Estabelecer protocolos claros de responsabilidade para garantir a segurança e a confiabilidade do sistema.
5. Implementar medidas robustas de segurança para proteger os dados dos pacientes contra acesso não autorizado, perda ou violação.
6. Validar o sistema em ensaios clínicos para demonstrar sua eficácia e segurança.
7. Tornar o sistema disponível para escolas médicas e unidades de saúde, democratizando o acesso a sistemas médicos que muitas vezes são vendidos por grandes corporações por valores inacessíveis.

Descrição das atividades a serem executadas

Universidade Federal de Minas Gerais / Faculdade de Medicina / Departamento de Pediatria

O Departamento de Pediatria ficará responsável pelas seguintes etapas do projeto:

1. Propiciar o ambiente para o desenvolvimento do projeto
2. Propiciar que o corpo docente da semiologia possam auxiliar no desenvolvimento do sistema

Prof. Cristiane Santos Dias

A partícipe ficará responsável pelas seguintes etapas do projeto:

1. Elaboração de cenários e casos clínicos para o desenvolvimento do sistema
2. Elaboração dos testes para avaliação dos resultados da aplicação do sistema

Prof Eduardo Araújo Oliveira

A partícipe ficará responsável pelas seguintes etapas do projeto:

1. Coordenar as etapas clínicas e pedagógicas do processo
 1. Elaboração de cenários e casos clínicos para o desenvolvimento do sistema
 2. Elaboração dos testes para avaliação dos resultados da aplicação do sistema

Prof Zilma Silveira Nogueira Reis

A partícipe ficará responsável pelas seguintes etapas do projeto:

1. Coordenar as etapas clínicas e pedagógicas do processo
1. Elaboração de cenários e casos clínicos para o desenvolvimento do sistema
2. Elaboração dos testes para avaliação dos resultados da aplicação do sistema

Instituto Federal de Minas Gerais/Departamento de Informática

Prof. Adriano Lages dos Santos

O partícipe ficará responsável pelas seguintes etapas do projeto:

1. Análise de Requisitos.
2. Coleta, limpeza e pré-processamento dos dados.
3. Análise e seleção de características para treinamento dos LLMs.
4. Treinamento e validação dos LLMs.
5. Avaliação de desempenho dos modelos.
6. Desenvolvimento e implantação de um sistema web que será a interface para utilização dos modelos pelos médicos e estudantes.
7. Divulgação científica e escrita de artigos científicos

Unimed-BH

A Unimed-BH, na qualidade de partícipe não executora do presente projeto, autoriza a aplicação dos recursos financeiros já aportados no âmbito do Acordo de Parceria para a Constituição de Aliança Estratégica firmado em 08 de julho de 2022 para execução do presente plano de trabalho. O recurso será gasto de acordo com o previsto no Quadro de Recursos Alocados no Projeto constante no Plano de Trabalho. Não havendo novos aportes no presente Acordo por parte da Unimed.

A Unimed não executa atividades de pesquisa neste Acordo.

Atividades conjuntas das instituições partícipes

Propiciar o espaço para o desenvolvimento do projeto.

Propiciar a compra/fornecimento de equipamentos para desenvolvimento do projeto.

Propiciar ambientes e equipe disponível para testes do sistema.

Realização entre os membros das instituições partícipes para tomada de decisões de desenvolvimento do projeto.

Cronograma de atividades:

Ações – Primeiro Ano	Meses											
	1	2	3	4	5	6	7	8	9	10	11	12
Assinatura do Acordo de Parceria/Plano de Trabalho	x											
Revisão de literatura e estado da arte		x	x									
Definição dos requisitos do sistema			x	x								
Coleta de dados					x	x						
Pré-processamento de dados							x	x				
Treinamento dos LLMs									x	x	x	
Desenvolvimento da interface do sistema											x	x
Testes da interface do sistema												x

Ações – Segundo Ano	Meses											
	13	14	15	16	17	18	19	20	21	22	23	24
Refinamento da interface do sistema	x											
Integração dos LLMs com a interface	x	x										
Refinamento do sistema integrado			x	x								
Validação clínica do sistema					x	x	x					
Documentação do sistema e escrita de artigos científicos							x	x				
Disseminação dos resultados								x	x			
Monitoramento e avaliação do sistema										x	x	x

RESULTADOS

Resultados esperados

4.1.1 Impacto Esperado

Melhoria na qualidade do diagnóstico: O sistema de IA pode auxiliar os pediatras a realizar diagnóstico de sepse de forma mais precisa e rápida, reduzindo o risco de erros médicos e contribuindo para uma assistência sistematizada e integral ao paciente gravemente enfermo.

Aprimoramento da semiologia: O sistema de IA pode ajudar os estudantes e pediatras a aprimorarem suas habilidades de semiologia, permitindo que identifiquem sinais e sintomas sutis com mais facilidade.

Segurança do paciente: O diagnóstico preciso e precoce de sepse é essencial para garantir a segurança do paciente. O sistema de IA pode contribuir para isso, auxiliando os pediatras na tomada de decisões mais assertivas.

Eficiência do atendimento: O sistema de IA pode auxiliar os pediatras a otimizar o tempo de consulta, permitindo que dediquem mais atenção aos pacientes e seus familiares.

Formação médica continuada: O sistema de IA pode ser utilizado como ferramenta de formação médica continuada, permitindo que os pediatras se mantenham atualizados sobre as últimas novidades em semiologia e diagnóstico pediátrico de sepse.

Aumento da precisão do diagnóstico: Os LLMs podem ser treinados em conjuntos de dados muito grandes e

complexos, o que lhes permite identificar padrões que podem ser difíceis de serem detectados por humanos. Isso pode levar a um aumento da precisão do diagnóstico e a uma redução do número de erros médicos.

Redução do tempo de diagnóstico: O sistema poderá auxiliar na redução do tempo necessário para o diagnóstico de doenças, permitindo que os pacientes recebam tratamento mais rapidamente.

Democratização do acesso ao diagnóstico médico: O sistema estará disponível em plataforma, o que permitirá que estudantes e pediatras brasileiros tenham acesso a uma ferramenta auxiliar de diagnóstico. Tem como potencial o uso supervisionado em Telemedicina, principalmente em regiões remotas com escassos recursos de saúde.

Estímulo à pesquisa e desenvolvimento em IA na área da saúde: O projeto poderá estimular a pesquisa e o desenvolvimento de novas aplicações para IA na área da saúde, o que pode levar a avanços no diagnóstico, tratamento e prevenção de doenças.

4.1.2 Perspectivas para Novos Avanços

Desenvolvimento de sistemas de diagnóstico mais precisos e eficientes: A utilização de LLMs em sistemas de diagnóstico médico pode levar ao desenvolvimento de sistemas mais precisos, eficientes e fáceis de usar.

Personalização do diagnóstico e tratamento: Os LLMs podem ser utilizados para personalizar o diagnóstico e o tratamento de doenças, levando em consideração as características individuais de cada paciente.

Desenvolvimento de novos métodos de diagnóstico: A IA pode ser utilizada para desenvolver novos métodos de diagnóstico, como o diagnóstico de doenças a partir de imagens ou exames de sangue.

Melhoria da comunicação entre médicos e pacientes: A IA pode ser utilizada para melhorar a comunicação entre médicos e pacientes, fornecendo informações mais precisas e fáceis de entender.

Integração: O sistema de IA poderia ser integrado a um sistema de prontuário eletrônico, facilitando o acesso às informações dos pacientes.

Ensino: O sistema de IA poderia ser utilizado para gerar relatórios personalizados sobre o desempenho dos alunos e pediatras na semiologia, auxiliando-os a identificar suas áreas de força e fraqueza.

Ensino: O sistema de IA poderia ser utilizado para desenvolver cursos e treinamentos online sobre semiologia pediátrica.

4.1.3 Aspectos Inovadores do Projeto

Utilização de LLMs em um contexto de ensino e treinamento médico: A utilização de LLMs em um contexto médico é uma inovação significativa, com o potencial de revolucionar a forma como os diagnósticos são realizados.

Foco na democratização do acesso ao diagnóstico médico: O projeto visa tornar o sistema de diagnóstico disponível, o que é uma inovação importante para garantir o acesso à saúde de qualidade para todos.

Abordagem multidisciplinar: O projeto reúne profissionais de diferentes áreas, como medicina e ciência da computação, o que permite uma abordagem mais abrangente e inovadora do problema.

4.1.4 Tipo de Inovação

A inovação do projeto de diagnóstico médico com LLMs pode ser classificada como disruptiva, pois tem o potencial de transformar a forma como o ensino da semiologia e os diagnósticos médicos são realizados.

4.1.5 Horizonte de Aplicação

O horizonte de aplicação do projeto é médio a longo prazo (1 a 10 anos).

4.1.6 Nível de Maturidade da Tecnologia (TRL)

O nível de maturidade da tecnologia do projeto de diagnóstico médico com LLMs é TRL 4 - Validação em ambiente de laboratório.

Metas e resultados esperados

Metas a serem atingidas	Prazos previstos para execução (mês de execução)	Indicadores (parâmetros a serem utilizados para a aferição do cumprimento das metas)
Levantamento de requisitos e definição do escopo do projeto	1-2	Documentação completa dos requisitos do sistema. Definição clara do escopo do projeto
Coleta de dados para treinamento dos LLMs	3-4	Quantidade de dados coletados. Qualidade dos dados coletados. Representatividade dos dados coletados.
Treinamento dos LLMs	4 a 6	Precisão dos modelos treinados. Tempo de treinamento dos modelos. Consumo de recursos computacionais.
Desenvolvimento da interface do sistema	7 a 8	Protótipo funcional da interface do sistema. Usabilidade da interface do sistema. Satisfação dos usuários com a interface do sistema Controle de acesso. Banco de dados.
Testes da interface do sistema	9	Identificação de erros e falhas na interface do sistema. Nível de satisfação dos usuários com a interface do sistema. Sugestões de melhorias para a interface do sistema
Integração dos LLMs com a interface	10-11	Funcionamento integrado dos LLMs com a interface do sistema. Precisão dos diagnósticos fornecidos pelos LLMs. Tempo de resposta dos LLMs.

Testes do sistema integrado	12 a 14	Identificação de erros e falhas no sistema integrado. Precisão dos diagnósticos fornecidos pelo sistema. Tempo de resposta do sistema.
Refinamento do sistema integrado	15	Implementação de melhorias no sistema integrado. Redução do número de erros e falhas no sistema integrado. Aumento da precisão dos diagnósticos fornecidos pelo sistema.
Validação clínica do sistema	16-17	Precisão dos diagnósticos fornecidos pelo sistema em comparação com o diagnóstico tradicional. Tempo de diagnóstico com o sistema em comparação com o tempo de diagnóstico tradicional. Satisfação dos médicos com o sistema.
Documentação do sistema	18-19	Documentação completa do sistema, incluindo manual do usuário, guia técnico e documentação de código. Escrita de artigos científicos.
Disseminação dos resultados	20-21	Publicações em revistas científicas. Apresentações em congressos e eventos científicos. Adoção do sistema por instituições de saúde.
Implementação do sistema	22-24	Número de instituições de saúde que implementaram o sistema. Número de usuários do sistema. Impacto do sistema na qualidade do diagnóstico e na saúde dos pacientes. Acompanhamento do sistema em ambiente de produção.

CONTRAPARTIDAS ENVOLVIDAS NO PROJETO

A tabela abaixo apresenta os pesquisadores vinculados às ICTs (listadas como partícipes) que são participantes do projeto, mas para os quais não há previsão de pagamento de bolsa. Os valores desta tabela farão parte da contrapartida econômica do projeto para fins de estimativa de percentual de cotitularidade.

5.1. Participação de PESQUISADORES das partícipes (não prevista remuneração de bolsas):

Nº	Instituição	Nome do Participante	CPF	Vínculo (Docente/Discente/Pesquisador externo/...)	Horas/SEMANA dedicadas ao PROJETO	Remuneração Básica Bruta Mensal (R\$)	Contrapartida mensal considerando dedicação ao PROJETO (R\$)
1	UFMG	Eduardo Araújo Oliveira	515.433.116-87	Docente	4 horas	R\$ 15,500,00	RS 1.000,00
2	UFMG	Cristiane Santos Dias	039.961.496-62	Docente	4 horas	R\$ 25000,00	R\$1.000,00
3	UFMG	Maria Chrsitina Lopes Araujo Oliveira	588.928.776-15	Docente	4 horas	RS 5.000,00	RS 316,00
4	UFMG	Zilma Silveira Nogueira Reis	442.155.336-00	Docente-DE	4 horas	R\$ 22.073,00	R\$ 2.207,00
						Total mensal	R\$ 4.523,00
						Meses de projeto	24
						Total projeto	R\$ 108.552,00

Infraestrutura da UFMG (contrapartida econômica)

	Valor
3 Notebooks	R\$15.000,00
Equipamentos do Centro de Informática em Saúde da Faculdade de Medicina da UFMG, por 24 meses: Utilização de laboratório de informática com computadores de alta performance. Utilização de servidor de dados institucional (até 96 gigas de RAM, representando 20% do total de RAM do servidor e até 8 núcleos de processamento, representando 20% do total de processamento). Compartilhamento de serviços de manutenção de sala de servidor (manutenção do servidor, energia elétrica, no-breaks, ar condicionado, monitoramento de cibersegurança).	R\$82.000,00
Serviços do Centro de Informática em Saúde da Faculdade de Medicina da UFMG: Integração da API-LLM com o sistema de prescrição de instituição parceira, estimado em 66 horas de trabalho técnico de informática.	R\$8.000,00
Custeio de tradução e revisão de textos científicos, estimado em 40 horas	R\$6.000,00
Valor Total da Contrapartida de Infraestrutura	R\$111.000,00

ORÇAMENTO DO PROJETO:

Recursos alocados no projeto:

FAPEMIG

Recursos alocados - FAPEMIG (gestão Fundep)	Valor unitário	Quantidade	Meses	Valor Total
Bolsa FAPEMIG - BDCTI-I (doutor ou graduado + 6 anos) - 50%	R\$ 2.860,00	1	24	R\$ 68.640,00

Bolsa FAPEMIG - BDCTI-VI (graduando)	R\$ 770,00	2	12	R\$ 18.480,00
Despesas com publicações científicas - FAPEMIG	R\$16.000,00	2	1	RS 32.000,00
TOTAL				R\$ 119.120,00

A seleção de um bolsista BDCTI-I, externo à UFMG, justifica-se pela necessidade de incorporar à nossa equipe pesquisador com sólida expertise em Inteligência Artificial (IA). Essa competência é fundamental para o projeto, dada sua complexidade técnica, que demanda conhecimento avançado em Grandes Modelos de Linguagem (LLM) modelagem de dados e validação clínica.

O bolsista selecionado atuará mediante a anuência de sua instituição de origem, assegurando que as atividades serão realizadas sem a utilização de sua infraestrutura física e de pessoal, caracterizando-se como atividades extras e sem prejuízo às suas atribuições regulares.

As atribuições do bolsista compreenderão a estruturação dos modelos, a modelagem dos dados e o auxílio na validação clínica. Tais atividades são inherentemente de pesquisa, desenvolvimento e inovação, não configurando, em nenhuma hipótese, contraprestação de serviços.

Cronograma de Execução do Orçamento da FAPEMIG (Item 6.1.1) pela FUNDEP

SEMESTRE	Bolsas FAPEMIG (R\$)	Diárias e passagens (R\$)	Serviços de Terceiros (R\$)	Material de Consumo (R\$)	Despesas com publicações científicas (R\$)
1	21.780,00				
2	21.780,00				
3	21.780,00				R\$16.000,00
4	21.780,00				R\$16.000,00
Total	R\$ 87.120,00				R\$ 32.000,00

INFRAESTRUTURA DAS INSTITUIÇÕES

Conforme § 3º do art. 35 do Decreto nº 9.283/2018, “As instituições que integram os acordos de parceria para pesquisa, desenvolvimento e inovação poderão permitir a participação de recursos humanos delas integrantes para a realização das atividades conjuntas de pesquisa, desenvolvimento e inovação, inclusive para as atividades de apoio e de suporte, e também ficarão autorizadas a prover capital intelectual, serviços, equipamentos, materiais, propriedade intelectual, laboratórios, infraestrutura e outros meios pertinentes à execução do plano de trabalho.”

A infraestrutura reportada no plano de trabalho serve como referência para cálculo de percentuais de propriedade intelectual (PI), caso venha a ser gerada, e faz parte da contrapartida das ICTs conforme descritas no quadro abaixo. Os percentuais de titularidade de PI serão discutidos em instrumento jurídico próprio, e irão contabilizar o uso da infraestrutura e horas de fato executadas.

VALOR TOTAL DO PROJETO

	Contrapartida para fins de estimativa de percentual de cotitularidade	
--	---	--

Orçamento do projeto	Pesquisadores	Infraestrutura UFMG	Total
R\$ 119.120,00	R\$108.552,00	R\$111.000,00	R\$ 338.672,00

9. PROPRIEDADE INTELECTUAL

Não se aplica ao presente projeto.

CONSIDERAÇÕES FINAIS

Um sistema de IA para auxiliar estudantes no aprendizado de semiologia e pediatras no raciocínio clínico tem potencial de revolucionar a forma como ensino da semiologia e o diagnóstico pediátrico de sepse são realizados. entanto, o desenvolvimento de um sistema robusto, confiável e seguro é um desafio significativo. Com pesquis desenvolvimento contínuos, este tipo de sistema pode se tornar uma ferramenta importante para melhora qualidade do ensino e do cuidado pediátrico.

O projeto de diagnóstico médico de sepse com Grandes Modelos de Linguagem (LLMs) apresenta potencial p revolucionar a forma como os diagnósticos são realizados, com o objetivo de reduzir o tempo de diagnóst aumentar a precisão do diagnóstico e melhorar a abordagem dos pacientes.

A utilização de LLMs em um contexto médico pode superar as limitações dos métodos tradicionais de diagnóstico sistema tem o potencial de ser mais preciso, eficiente e acessível do que os métodos tradicionais, e pode ser utilizada para diagnosticar uma ampla gama de doenças.

O projeto também é inovador por seu foco na democratização do acesso ao diagnóstico médico de sepse. O sistema estará disponível de forma gratuita, o que permitirá que estudantes e médicos de todo o mundo tenham acesso a uma ferramenta poderosa de diagnóstico. Além disso, poderá ser utilizado em locais remotos onde os recursos de saúde são escassos. Imagine como um sistema deste tipo pode ajudar um único médico em situações de trabalho em vilas e locais sem infra-estrutura longe de grandes centros e sem outro médico para poder dialogar sobre suas dúvidas clínicas. O projeto reúne profissionais de diferentes áreas, como medicina, ciência da computação e engenharia, o que permite uma abordagem mais abrangente e inovadora do problema. Essa colaboração multidisciplinar é essencial para o sucesso do projeto.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 - Gamborg ML, Mehlsen M, Paltved C, Vetter SS, Musaeus P. Clinical decision-making and adaptive expertise in residency: a think-aloud study. *BMC Med Educ.* 2023 Jan 12;23(1):22. doi: 10.1186/s12909-022-03990-8. PMID: 36635669; PMCID: PMC9835279.
- 2 - Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *J Ambient Intell Humaniz Comput.* 2023;14(7):8459-8486. doi: 10.1007/s12652-021-03612-z. Epub 2022 Jan 13. PMID: 35039756; PMCID: PMC8754556.
- 3 - Sanchez-Pinto LN, Bennett TD, DeWitt PE, Russell S, Rebull MN, Martin B, Akech S, Albers DJ, Alpern ER, Balamuth F, Bembea M, Chisti MJ, Evans I, Horvat CM, Jaramillo-Bustamante JC, Kissoon N, Menon K, Scott HF, Weiss SL, Wiens MO, Zimmerman JJ, Argent AC, Sorce LR, Schlapbach LJ, Watson RS; Society of Critical Care Medicine Pediatric Sepsis Definition Task Force; Biban P, Carrol E, Chiotos K, Flauzino De Oliveira C, Hall MW,

Inwald D, Ishimine P, Levin M, Lodha R, Nadel S, Nakagawa S, Peters MJ, Randolph AG, Ranjit S, Souza DC, Tissieres P, Wynn JL. Development and Validation of the Phoenix Criteria for Pediatric Sepsis and Septic Shock. JAMA. 2024 Feb 27;331(8):675-686. doi: 10.1001/jama.2024.0196. PMID: 38245897; PMCID: PMC10900964.

4 - Diagnóstico precoce de sepse. Disponível em: <https://www.gov.br/pt-br/noticias/saude-e-vigilancia-sanitaria/2022/09/diagnostico-precoce-e-fundamental-para-tratar-a-sepse-conhecida-como-infeccao-generalizada>

5 - Omiye JA, Gui H, Rezaei SJ, Zou J, Daneshjou R. Large Language Models in Medicine: The Potentials and Pitfalls : A Narrative Review. Ann Intern Med. 2024 Feb;177(2):210-220. doi: 10.7326/M23-2772. Epub 2024 Jan 30. PMID: 38285984.

6 - Thirunavukarasu, A.J., Ting, D.S.J., Elangovan, K. et al. Large language models in medicine. Nat Med 29, 1930–1940 (2023). <https://doi.org/10.1038/s41591-023-02448-8>

7 - Cabral S, Restrepo D, Kanjee Z, et al. Clinical Reasoning of a Generative Artificial Intelligence Model Compared With Physicians. JAMA Intern Med. Published online April 01, 2024. doi:[10.1001/jamainternmed.2024.0295](https://doi.org/10.1001/jamainternmed.2024.0295)

8 - Baker EA, Ledford CH, Fogg L, Way DP, Park YS. The IDEA Assessment Tool: Assessing the Reporting, Diagnostic Reasoning, and Decision-Making Skills Demonstrated in Medical Students' Hospital Admission Notes. Teach Learn Med. 2015;27(2):163-73. doi: 10.1080/10401334.2015.1011654. PMID: 25893938.

9 - Zhou, C., Liu, P., Xu, P., Iyer, S., Sun, J., Mao, Y., Ma, X., Efrat, A., Yu, P., Yu, L., Zhang, S., Ghosh, G., Lewis, M., Zettlemoyer, L., & Levy, O. (2023). LIMA: Less Is More for Alignment. Disponível em: <https://arxiv.org/abs/2305.11206>

10 - RESOLUÇÃO DE DIRETORIA COLEGIADA - RDC Nº 657, DE 24 DE MARÇO DE 2022. Disponível em: <https://in.gov.br/en/web/dou/-/resolucao-de-diretoria-colegiada-rdc-n-657-de-24-de-marco-de-2022-389603457>

11 - Projeto de Lei nº 2338, de 2023. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9347622&ts=1715114415295&disposition=inline>

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Professora Sandra Regina Goulart Almeida

Reitora

Prof. Eduardo Araújo Oliveira

Coordenador do Projeto

UNIMED BH

Sr. Frederico José Amédée Péret

Diretor Presidente

INSTITUTO FEDERAL DE CIÊNCIA E TECNOLOGIA DE MINAS GERAIS

Professor Rafael Bastos Teixeira

FUNDAÇÃO DE DESENVOLVIMENTO DA PESQUISA

Professor Jaime Arturo Ramírez

Presidente



Documento assinado eletronicamente por **Frederico Jose Amedee Peret, Usuário Externo**, em 04/11/2025, às 11:37, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo de Souza Esteves, Usuário Externo**, em 07/11/2025, às 11:08, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Sandra Regina Goulart Almeida, Reitora**, em 07/11/2025, às 18:36, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jaime Arturo Ramírez, Usuário Externo**, em 10/11/2025, às 16:10, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eduardo Araujo de Oliveira, Professor do Magistério Superior**, em 14/11/2025, às 13:21, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rafael Bastos Teixeira, Usuário Externo**, em 11/12/2025, às 12:50, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4705359** e o código CRC **073F578A**.