

Proposta de Projeto 2025-2026

LICENCIATURA EM ENGENHARIA INFORMÁTICA

PROPOSTA N.º

TÍTULO*	Avaliação do conhecimento factual de Grandes Modelos de Linguagem no domínio médico
ORIENTADOR	Hugo Paredes
PRINCIPAL*	
COORIENTADORES	Rafael Ris-Ala
ALUNOS(s)	2
ÁREA DE INVESTIGAÇÃO	<i>Inteligência Artificial, Inteligência Artificial Centrada no Ser Humano</i>
CENTRO DE INVESTIGAÇÃO	HumanISE – INESC TEC
DEPENDÊNCIAS	<i>Conhecimentos básicos sobre Inteligência Artificial e Aprendizagem de Máquina, que fornecem as bases teóricas e metodológicas necessárias para a análise de modelos de linguagem, avaliação de viéses e condução de estudos científicos.</i>
APRESENTAÇÃO*	Os Grandes Modelos de Linguagem (Large Language Models - LLMs) estão sendo utilizados de forma crescente no domínio da saúde, nomeadamente no apoio à decisão clínica e na disponibilização de informação médica a utilizadores finais, o que levanta preocupações significativas relativamente à presença de viéses nas suas respostas. Entre eles, o viés de confirmação pode levar os modelos a reforçar pressupostos incorretos presentes nas perguntas, comprometendo a neutralidade e confiança da informação gerada. Assim, a verificação do conhecimento factual em LLMs é particularmente relevante no contexto médico, onde decisões baseadas em informação enviesada podem ter consequências graves, tornando esta proposta pertinente do ponto de vista científico, ético e social.
OBJETIVOS*	O objetivo deste projeto é: coletar a resposta padrão a 52 questões médicas em 10 LLMs mais populares; e verificar se os LLMs efetivamente possuem a resposta correta. Esse estudo deve ser realizado através de: um processo estruturado de experimentação que inclui um dataset com 52 ocorrências médicas; e a determinação da resposta padrão dos modelos para cada ocorrência (utilizando prompts neutros); e a comparação dessas respostas com aquelas já obtidas a partir de prompts enviesados. Essa pesquisa visa contribuir para o desenvolvimento de sistemas de apoio à decisão em saúde mais confiáveis e imparciais e à formação de conhecimento crítico de novos engenheiros de software.

* Campos de preenchimento obrigatório

NOTA: a totalidade deste documento (exceto esta linha) não deve exceder uma página.