Pesquisa de Inteligência Artificial: Estratégias para Criação de Prompts Eficientes em Modelos de Linguagem

SILVA, Camylly Barros

Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Luziânia, Bacharelado em Sistemas de Informação, Programa PIBIC-EM.

MUNDIM, Aldo Lucio de Freitas

Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Luziânia, Departamento de Áreas Acadêmicas.

A pesquisa teve por objetivo sistematizar estratégias de prompting em língua portuguesa e estruturar um protocolo para avaliar sua eficiência em um cenário de Sistemas de Informação. A eficiência foi considerada em três dimensões: desempenho, custo computacional e consistência. Com base na Design Science Research, desenvolveu-se o Financial Prompt Lab, artefato implementado em Python com LangChain, FastAPI e Streamlit para tarefas sobre dados do mercado financeiro brasileiro. Foram construídas quatro condições: zero-shot, few-shot com dois exemplos, raciocínio estruturado e combinação destas duas últimas, diferenciadas apenas pelo bloco da técnica. Os dados foram congelados em um snapshot de 2.750 registros diários de 11 ativos da B3, relativos a 250 pregões por ativo em 2025. Elaborou-se um benchmark com 30 questões em português, distribuídas em cinco categorias e três níveis de dificuldade, com gabaritos gerados a partir do snapshot, cuja conferência combinou recálculo independente em pandas e inspeção de uma amostra do payload bruto. A contagem de tokens de entrada, estimada com o tokenizador o200k_base, indicou acréscimos de 12% para raciocínio estruturado, 31% para few-shot e 43% para a combinação, em relação ao zero-shot. A infraestrutura passou por 205 testes automatizados e 600 execuções com oráculo determinístico, sem falhas. A auditoria do avaliador automático identificou duas fontes de falso acerto, posteriormente corrigidas. O período do projeto priorizou a construção do artefato, a auditoria dos dados e a validação técnica do protocolo; a coleta com modelos de linguagem reais, a comparação de acurácia e consistência e a avaliação humana cega das respostas interpretativas ficaram para uma etapa posterior. Os resultados caracterizam o custo estimado de entrada e a prontidão técnica do protocolo, mas não permitem concluir qual condição apresenta melhor desempenho ou custo por acerto observado. A etapa seguinte requer executar o benchmark com modelos e configurações identificados, registrar o consumo efetivo e avaliar cegamente as respostas interpretativas.

Palavras-chave: engenharia de prompts; estratégias de prompting; eficiência de prompts; língua portuguesa; design science research.

Agradecimentos

O trabalho foi desenvolvido no âmbito do Programa PIBIC-EM do Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG).