VII Mostra de Extensão, Ciência e Tecnologia
XXXII Seminário de Iniciação Científica
XVII Salão de Ensino e Extensão
VII Mostra da Pós-Graduação Stricto Sensu
VI Seminário de Inovação Tecnológica

Título:	**AVALIAÇÃO DE UM FRAMEWORK MULTIAGENTE ADAPTATIVO PARA RACIOCÍNIO EM MÚLTIPLAS ETAPAS COM MODELOS DE LINGUAGEM DE LARGA ESCALA**		
Autores:	**Leonardo Antonio Patatt** **Rejane Frozza (orientadora)**		
Área	[ ] Humanas [ ] Sociais Aplicadas [ ] Biológicas e da Saúde [X] Exatas, da Terra e Engenharias	Dimensão:	[X] Ensino [ ] Pesquisa [ ] Extensão [ ] Inovação

Introdução: Modelos de linguagem de larga escala (LLMs) têm sido empregados em aplicações que demandam geração, interpretação e transformação de informações. Contudo, a fluência textual das respostas não garante que o raciocínio subjacente seja correto, completo ou eficiente. Em tarefas de múltiplas etapas, observam-se erros de cálculo, perda de informações relevantes, conclusões sem verificação e aumento de custo computacional, quando os modelos são induzidos a gerar cadeias de pensamento excessivamente longas. Neste contexto, o tema desta pesquisa refere-se ao desenvolvimento e avaliação de um protótipo computacional para comparação de estratégias de raciocínio em LLMs. A proposta é contemplar as estratégias *Chain-of-Thought, Chain-of-Draft, Skeleton-of-Thought* em um ambiente com múltiplos agentes de raciocínio **Objetivo:** Desenvolver e avaliar um *framework* multiagente adaptativo, capaz de selecionar e combinar estratégias de geração, decomposição e validação de respostas conforme as características de cada tarefa. **Metodologia:** A pesquisa é de natureza aplicada, exploratória e descritiva, com etapa experimental planejada para comparar, sob as mesmas condições de modelo, exemplos e configuração, sete abordagens: direct prompting, Chain-of-Thought, Self-Consistency, Chain-of-Draft, Skeleton-of-Thought, uma estratégia multiagente e o roteamento adaptativo proposto. O procedimento metodológico compreende levantamento bibliográfico sobre raciocínio em LLMs, *prompting*, verificação, debate entre agentes, eficiência computacional e *benchmarks*; organização de uma seleção inicial de estudos; documentação da arquitetura do *framework*; definição do protocolo de avaliação; execução dos testes; e análise comparativa dos resultados. A avaliação utilizará o GSM8K, voltado a problemas matemáticos, e o HotpotQA, voltado a perguntas que exigem articulação de informações em múltiplas etapas, observando métricas de desempenho e confiabilidade, como acurácia, taxa de abstenção e indicadores de eficiência: latência, quantidade de chamadas, *tokens* de entrada e saída e custo estimado. **Principais resultados:** Um protótipo inicial do sistema foi estruturado, contemplando um fluxo capaz de receber a tarefa, encaminhá-la a estratégias distintas, organizar as respostas geradas e apoiar sua validação, além da definição das condições necessárias à comparação experimental. Essa etapa permitiu delimitar o problema de pesquisa, explicitar os critérios de medição e identificar que a escolha da estratégia deve considerar, simultaneamente, qualidade da resposta, complexidade da tarefa, tempo de execução e consumo de recursos. **Conclusão:** Espera-se que a pesquisa resulte em uma comparação reprodutível entre estratégias de raciocínio, contribuindo para compreender em quais cenários múltiplos agentes e o roteamento adaptativo podem ampliar a confiabilidade das respostas sem impor custo computacional

desnecessário. A reflexão possibilitada por esta etapa indica que aprimorar as respostas de LLMs não depende apenas de gerar mais texto ou de ampliar o número de etapas, mas de selecionar procedimentos de raciocínio e verificação compatíveis com o objetivo, os limites de custo e o nível de incerteza de cada problema.O tema possui aderência aos Objetivos do Desenvolvimento Sustentável, mais especificamente ao ODS 9, relacionada à inovação e infraestrutura, e à ODS 4, pela possibilidade de apoiar ferramentas educacionais e de apoio ao aprendizado quando empregadas com validação adequada.

Link do Vídeo: Apresentação VII Mostra.mp4