

Elton Cardoso do Nascimento - 233840

IA024 - Redes Neurais Profundas para Processamento de Linguagem Natural - 1s2024

Leitura do Artigo "A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications (Sahoo et al.)"

O artigo realiza uma revisão de técnicas para realização de engenharia de prompts. Estas técnicas utilizam prompts ou formas de geração específicas com LLMs para melhorar a performance em tarefas específicas sem precisar realizar um ajuste fino na tarefa.

O trabalho separa as diversas técnicas analisadas em 12 categorias: apresentação básica de novas tarefas, onde pouco é feito para ajudar o modelo a realizá-las; uso de processos passo-a-passo elaborados para gerar respostas; redução de alucinação; geração automática de prompts; uso de ferramentas; aprendizado de exemplos incorretos; prompts afetivos; geração e execução de código; otimização; rephraseamento de prompts; metacognição e interface de anotação.

Embora diverso, o trabalho faz pouco para destacar as diferenças entre as técnicas. Uma outra crítica está na forma de apresentação, onde exemplos concretos ou diagramas de cada técnica poderiam ter sido empregados.