

Classification accuracy of reasoning LLMs

