XQuAD on Gemini with Word Substitution Attack - LLM Judge Score 3.88 3.88 3.71 3.96 3.92 4.13 3.52

