BERT Scores for All Subsets

Random-Samples

Mean score

GPT-4o   GPT-4o-mini   LLAMA3.1 8b   Open Mixtral 8x22b   Mistral-largest-latest   Gemma2

Shortest-Samples

Mean score

GPT-4o   GPT-4o-mini   LLAMA3.1 8b   Open Mixtral 8x22b   Mistral-largest-latest   Gemma2

Longest-Samples

Mean score

GPT-4o   GPT-4o-mini   LLAMA3.1 8b   Open Mixtral 8x22b   Mistral-largest-latest   Gemma2