

Comparison of non_llm_string_similarity across models
Friedman $\chi^2 = 51.9200$, $p = 0.0000$ (significant)

