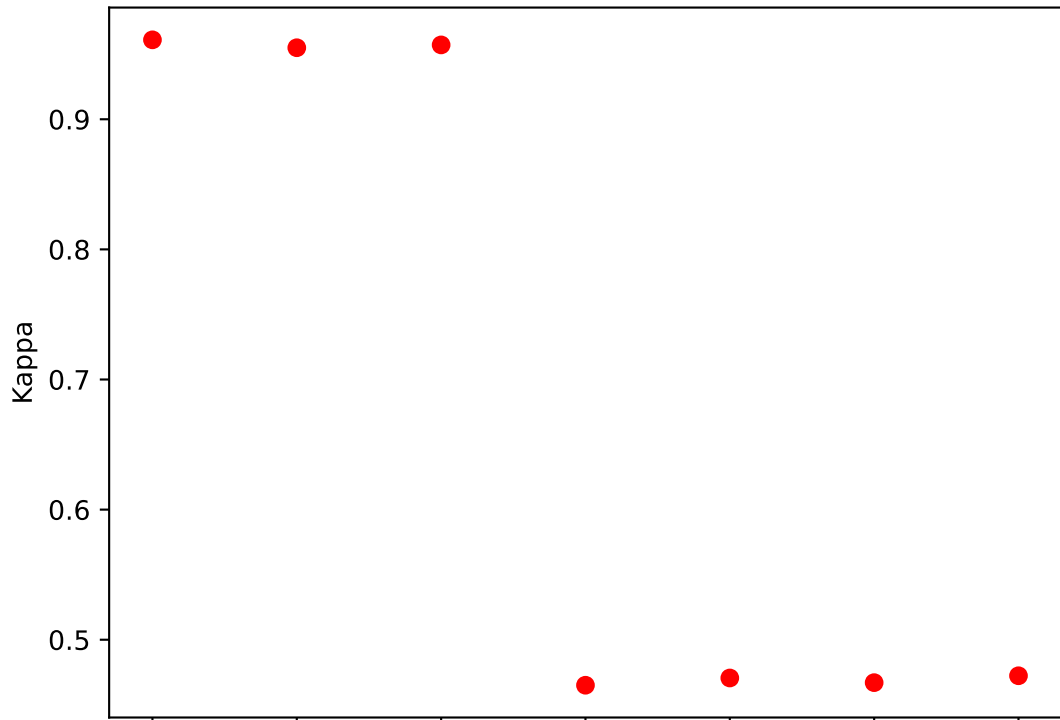


Kappa of ChatGPT gpt4 vs Experts



GPT-Dataset score_1 vs score_2
Expert-Dataset score_1 vs score_2
GPT-Rating-Expert-Dataset score_1 vs score_2
GPT-VS-Expert score_1 vs score_1
GPT-VS-Expert score_1 vs score_2
GPT-VS-Expert score_2 vs score_1
GPT-VS-Expert score_2 vs score_2

Amount of samples the model was trained with

The ChatGPT gpt4-Dataset consists of 8000 samples.
The Expert-Dataset consists of 3592 samples.