BERT Model trained with samples created by gpt4 rating answers created by text-davinci-003 annotated by gpt-3.5-turbo used for testing 0.5 -0.4 Карра - 8.0 0.2 0.1 1000 1500 500 2000 2500 3000 Amount of samples the model was trained with

Each value represents the agreement of exactly 1576 ratings