BERT Model trained with samples created by text-davinci-003 annotated by gpt-3.5-turbo rating answers created by text-davinci-003 annotated by gpt-3.5-turbo used for testing 0.45 -0.40 -0.35 -0.30 -Харра - 52.0 -0.20 -0.15 -0.10 -0.05 -

1500

Amount of samples the model was trained with

2000

500

1000

Each value represents the agreement of exactly 1576 ratings

2500

3000