BERT Model trained with samples created by text-davinci-003 annotated by gpt-3.5-turbo rating answers created by text-davinci-003 used for testing 0.40 -0.35 -0.30 (4 V 20.25 (4 V 20.25 (5 V 20.25 0.20 0.15 0.10 0.05 400 500 100 200 300 600 700 800 Amount of samples the model was trained with Each value represents the agreement of exactly 1577 ratings