XGB Model trained with samples created by human experts rating answers created by text-davinci-003 annotated by gpt-3.5-turbo used for testing 0.04 -0.03 0.02 -Карра 0.01 0.00 -0.01400 500 800 100 200 300 600 700 Amount of samples the model was trained with Each value represents the agreement of exactly 1576 ratings