XGB Model trained with samples created by gpt4 rating answers created by human experts used for testing 0.65 -0.60 0.55 0.50 Карра 0.45 0.40 0.35 0.30 0.25 -2000 500 1000 1500 2500 3000 Amount of samples the model was trained with Each value represents the agreement of exactly 392 ratings