XGB Model trained with samples created by human experts rating answers created by text-davinci-003 used for testing 0.35 0.30 0.25 0.20 0.15 -0.10 0.05 0.00 -0.05400 200 600 800 1000 1200 1400 1600 Amount of samples the model was trained with Each value represents the agreement of exactly 1577 ratings