XGB 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 training 0.7 -0.6 -0.5 -Kappa - 5.0 0.3 -0.2 -700 800 100 200 300 400 500 600 Amount of samples the model was trained with Each value represents the agreement of exactly 6401 ratings