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 0.65 -0.60 -Kappa Kappa Kappa 0.50 -0.45 -0.40 -100 200 300 400 500 600 700 800 Amount of samples the model was trained with Each value represents the agreement of exactly 6401 ratings