XGB Model trained with samples created by gpt4 rating answers created by human experts used for training 0.55 0.50 (Yabba Yabba Yabba 0.40 0.35 0.30 -200 300 400 500 100 600 700 800 Amount of samples the model was trained with Each value represents the agreement of exactly 3200 ratings