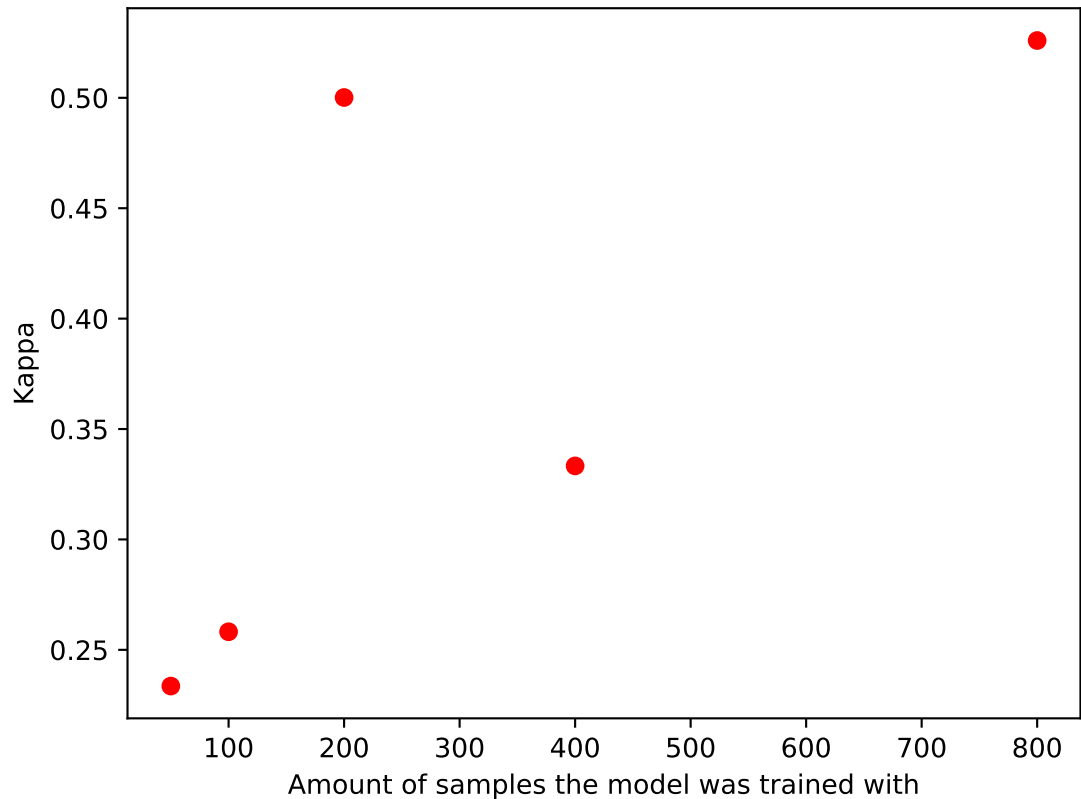QWK of score_1 assigned to samples created by experts-training by two XGB Models
One trained with samples created by gpt4 one trained with samples created by experts.

Kappa

Amount of samples the model was trained with

Each value represents the agreement of exactly 3200 ratings