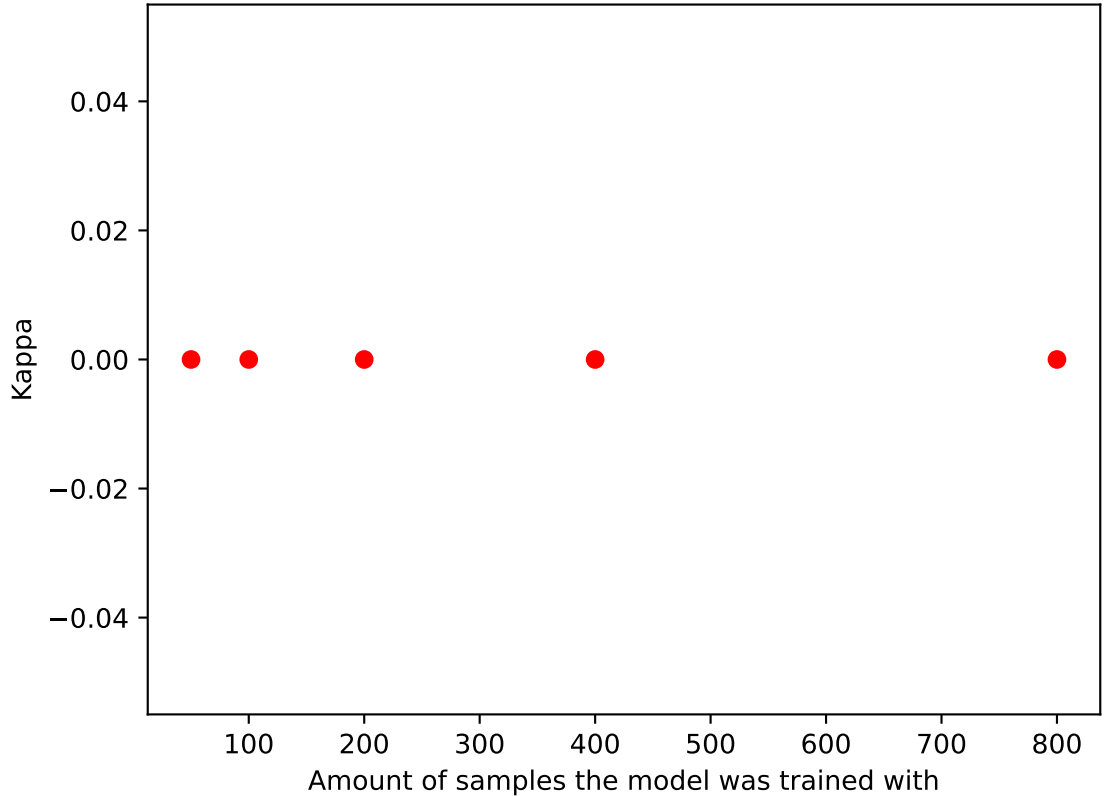


BERT Model trained with samples created by human experts rating answers created by gpt4 used for testing



Each value represents the agreement of exactly 1600 ratings