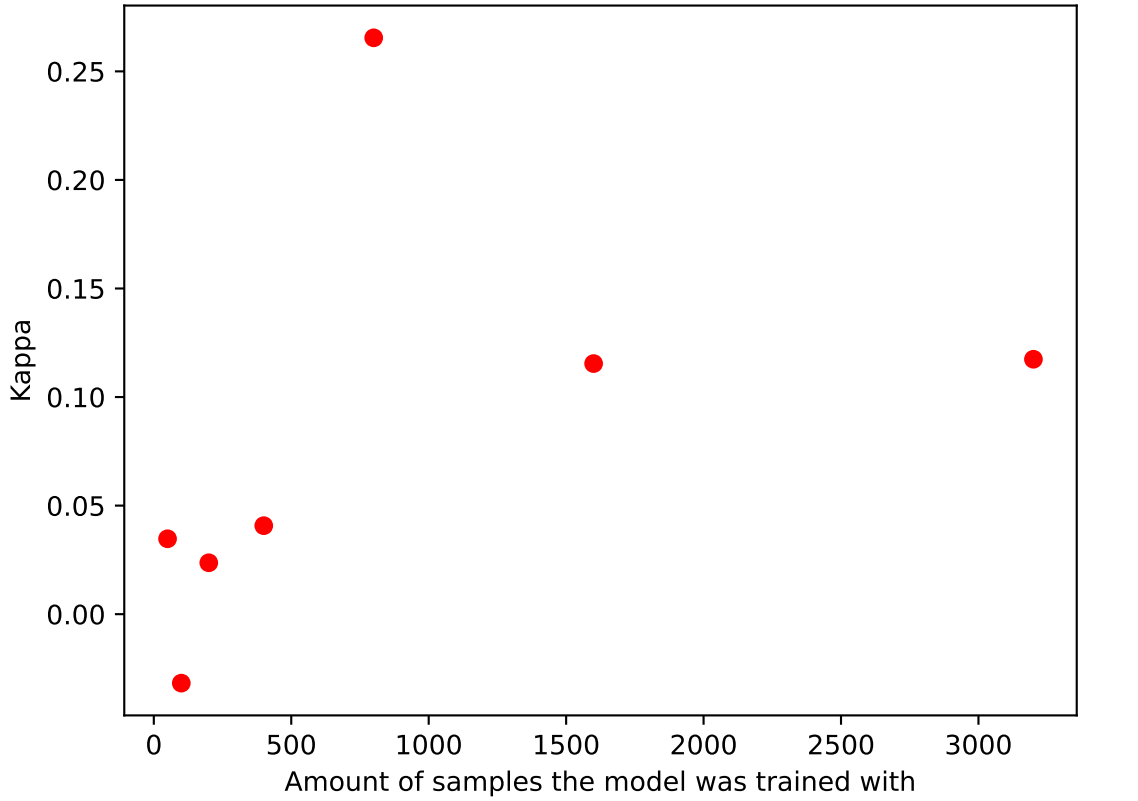
BERT Model trained with samples created by gpt4 rating answers created by text-davinci-003 annotated by gpt-3.5-turbo used for testing



Each value represents the agreement of exactly 1576 ratings