QWK per Prompt after Finetuning