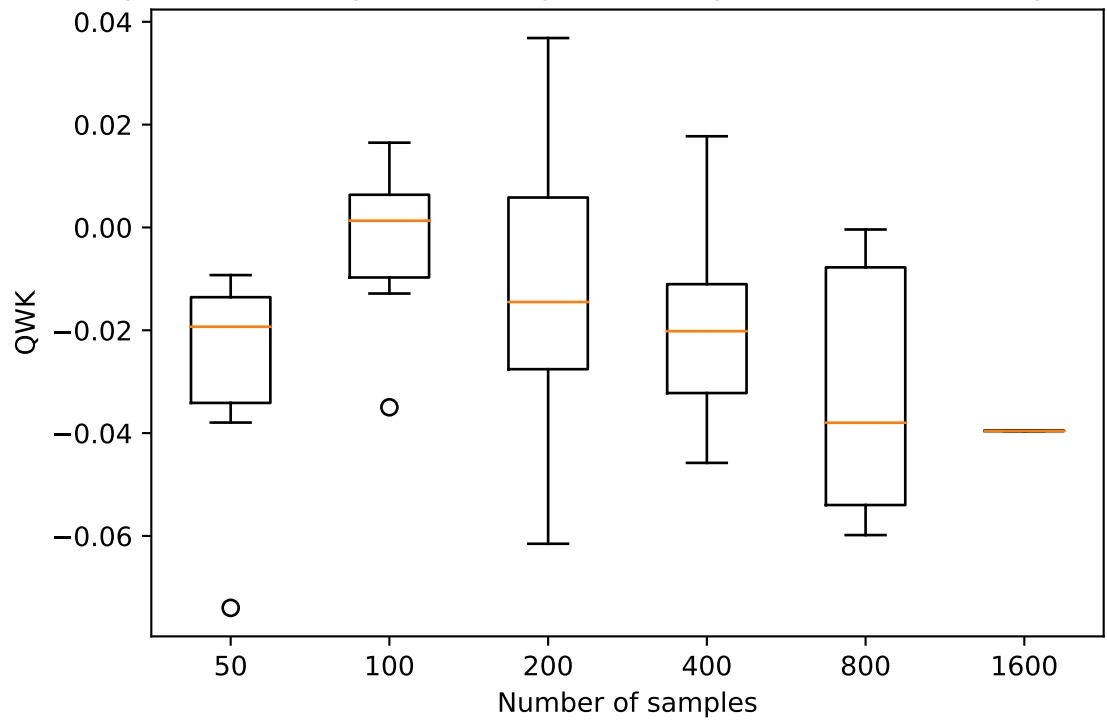XGB Model trained with samples created by human experts rating answers created by text-davinci-003 used for testing

Each box represents exactly 6 kappa values