

KL Differences Relative to GPT2 - \* Smaller Models \*

