

KL Differences Relative to GPT2 - * Earlier Models *

