

KL Differences Relative to GPT2 Large - * Earlier Models *

