

Surprisal (bits)

GPT2M

GPT2L

GPT2XL

Embedding Bias

