We are studying the functionality of each layer during LLM inference. We have selected the most likely output tokens for each layer's hidden state. Please check the activation of the hidden state's output for the following tokens and summarize in a single sentence what the layer's function is. Don't list examples of words. The activation format is token<tab>activation. Activation values range from 0 to 10. A layer finding what it's looking for is represented by a non-zero activation value. The higher the activation value, the stronger the match. We have provided the input prompt for the LLM. You can use this input prompt as a reference to evaluate the functionality of each layer. Layer 0 '\t1 disappearing\t6 earing\t10 aping\t9 the\t4 affecting\t3 <end> Explanation of Layer 0 behavior: the main thing this layer does is to focus on present tense verbs ending in 'ing' Laver 1 arrest\t10 iao\t9 <end> Explanation of Layer 1 behavior: the main thing this layer does is to find words related to physical medical conditions Layer 2 together\t3 ness\t7 town\t1 we\t2 're\t4 all\t3 in\t7 this\t10 together\t5 <end>

Explanation of Layer 2 behavior: the main thing this layer does is to focus on phrases related to community