HDBSCAN attention_logit cluster sizes (randomtext_bert-large-uncased_params-trained_dense-on_layers-24) after layer 1 after layer 4 after layer 2 after layer 3 12.5 10.0 cluster size 7.5 5.0 2.5 0.0 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 6 after layer 5 after layer 7 after layer 8 cluster size k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 9 after layer 10 after layer 11 after layer 12 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 13 after layer 14 after layer 15 after layer 16 150 -60 · cluster size 200 k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 17 after layer 18 after layer 19 after layer 20 30 -k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster after layer 22 after layer 21 after layer 23 after layer 24 cluster size k-th largest cluster k-th largest cluster k-th largest cluster k-th largest cluster