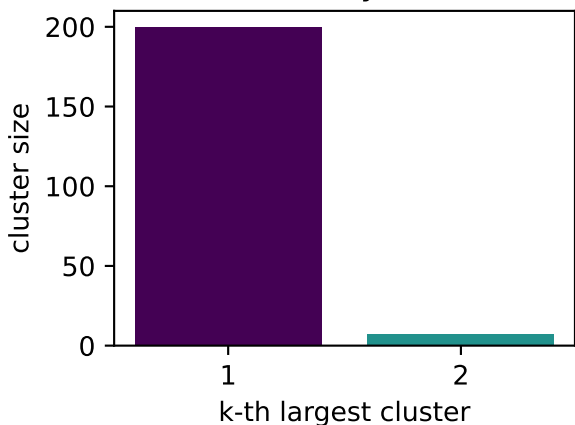
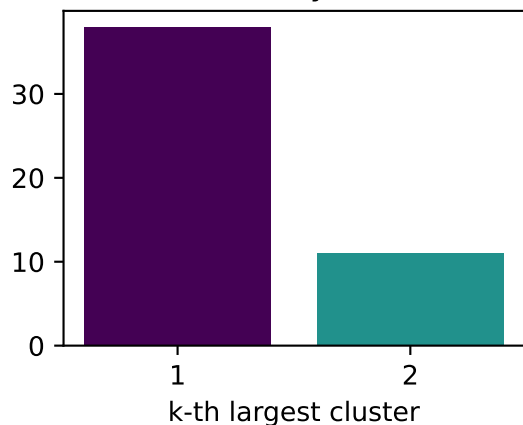


HDBSCAN attention_logit cluster sizes
(wikitext_albert-large-v2_params-trained_dense-on_layers-24)

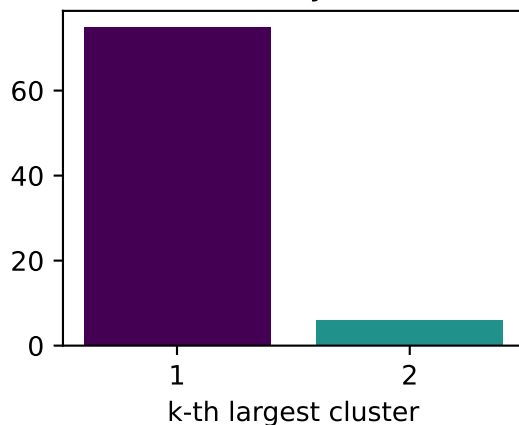
after layer 1



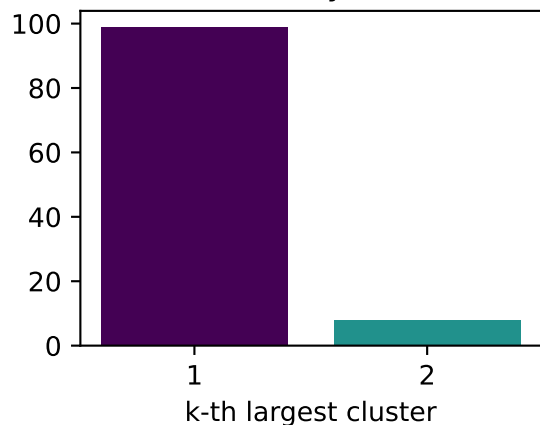
after layer 2



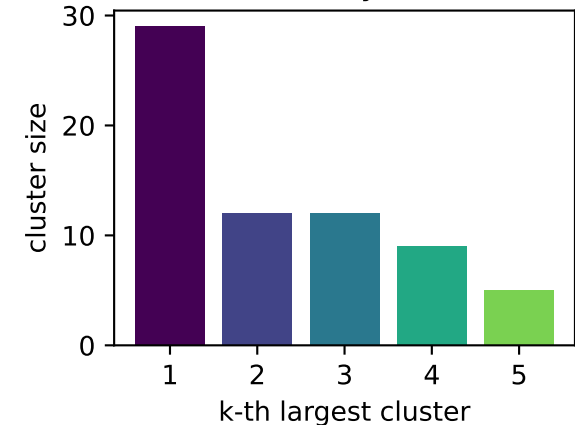
after layer 3



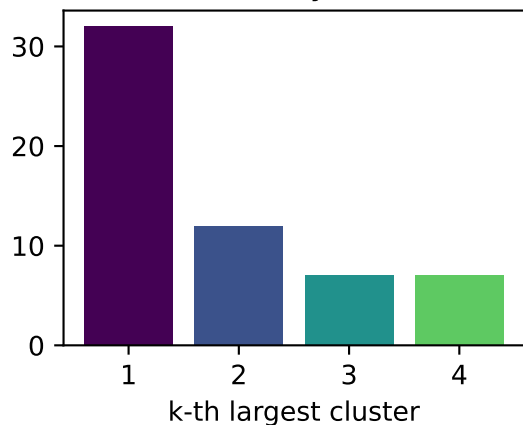
after layer 4



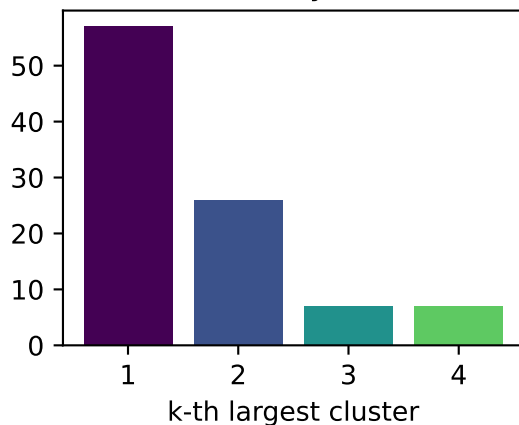
after layer 5



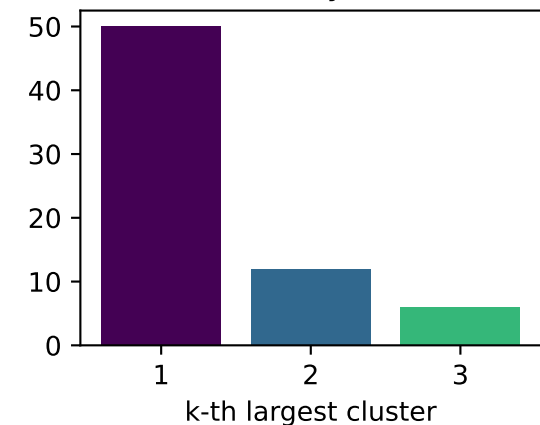
after layer 6



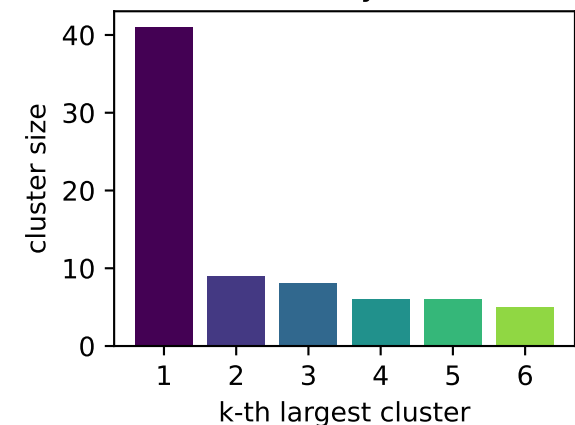
after layer 7



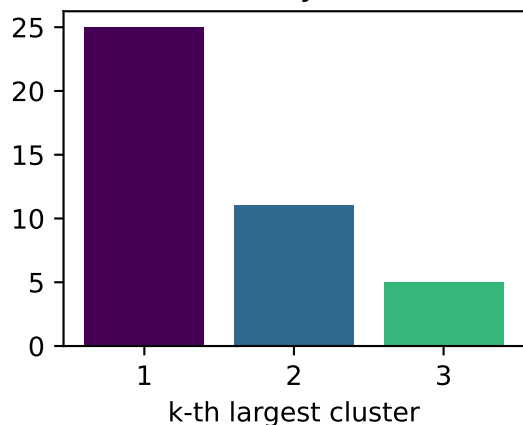
after layer 8



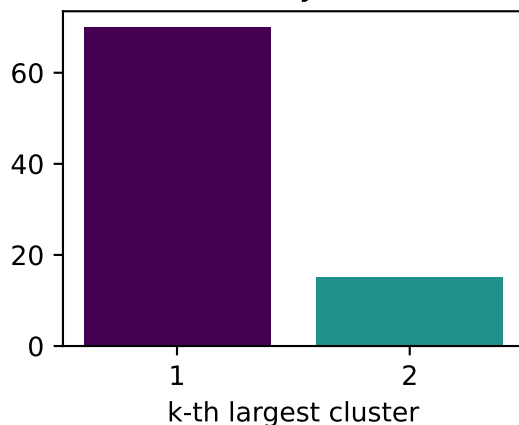
after layer 9



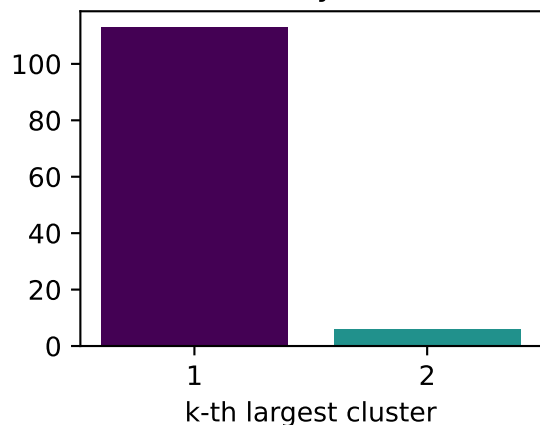
after layer 10



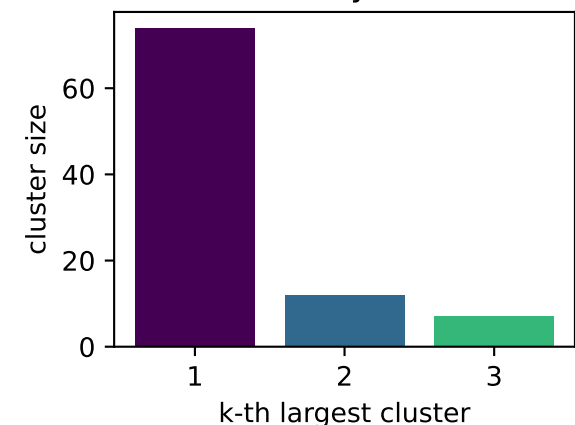
after layer 11



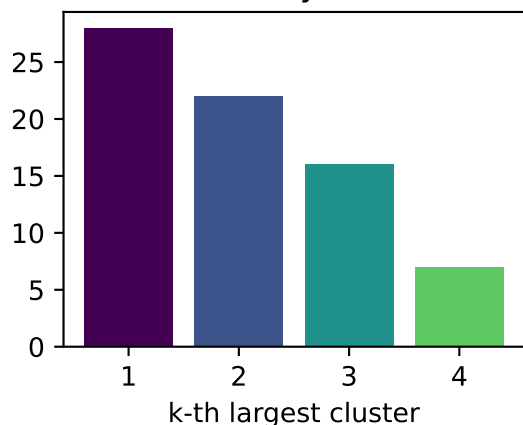
after layer 12



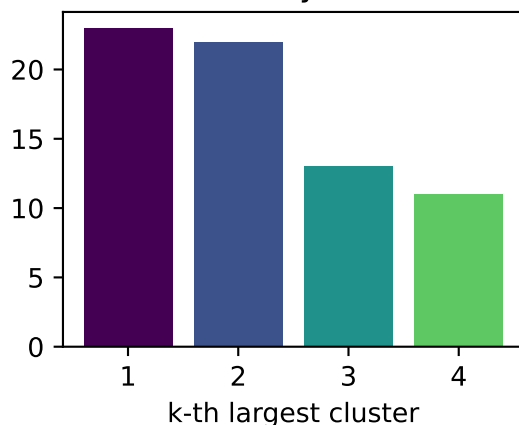
after layer 13



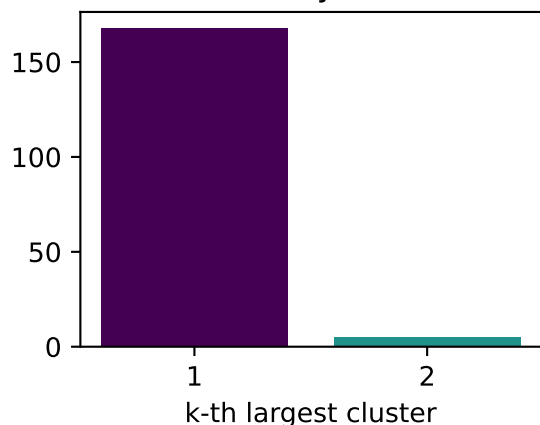
after layer 14



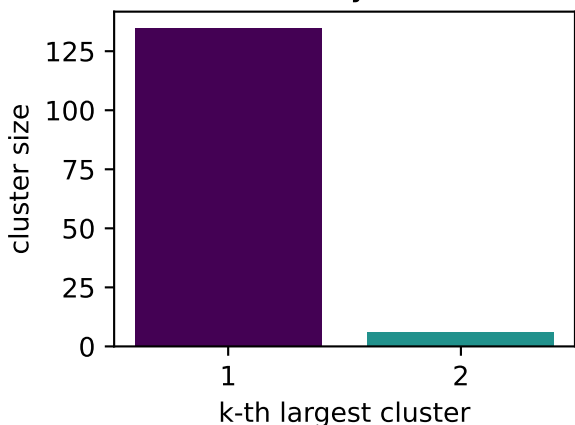
after layer 15



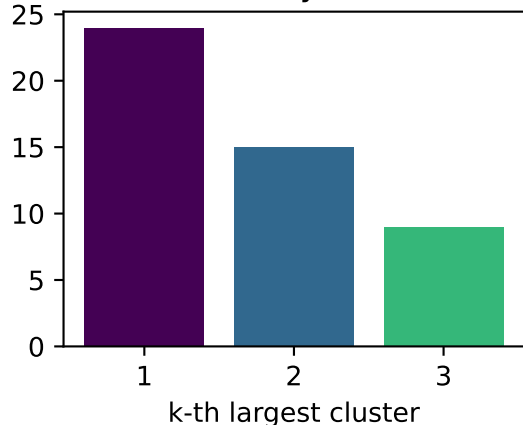
after layer 16



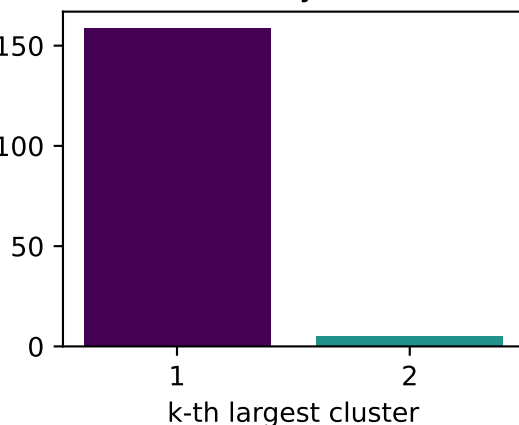
after layer 17



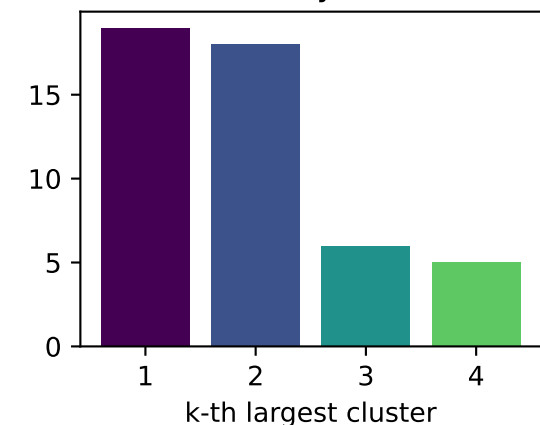
after layer 18



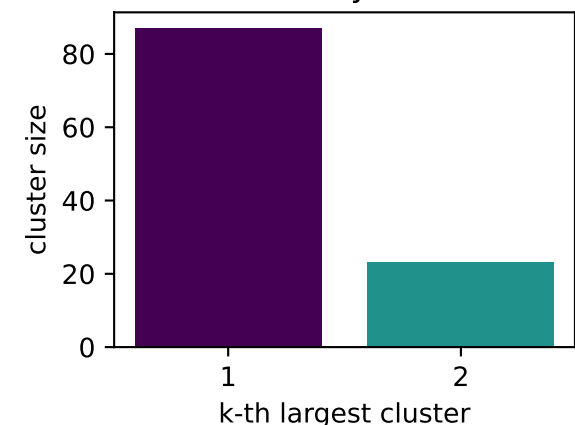
after layer 19



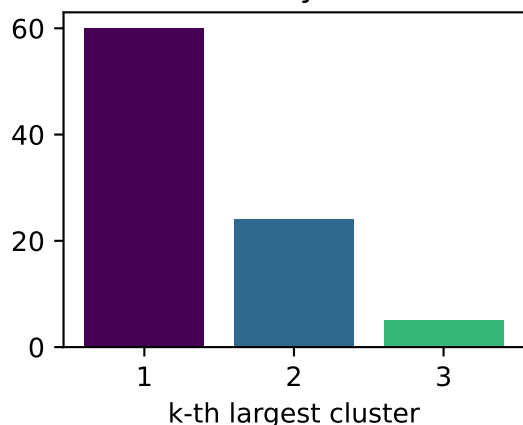
after layer 20



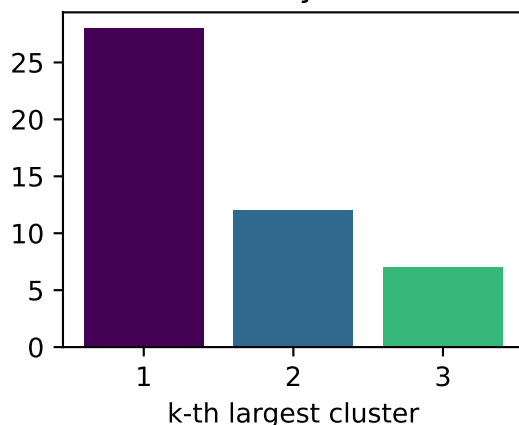
after layer 21



after layer 22



after layer 23



after layer 24

