wikitext\_bert-large-uncased\_params-trained\_dense-off\_layers-24, sample 9 number of clusters  $\begin{array}{c} & & \text{outlier rate} \\ 30 & & & \\ 20 & & & \\$ 

HDBSCAN cluster evaluation of tokens  $x_i$ 

