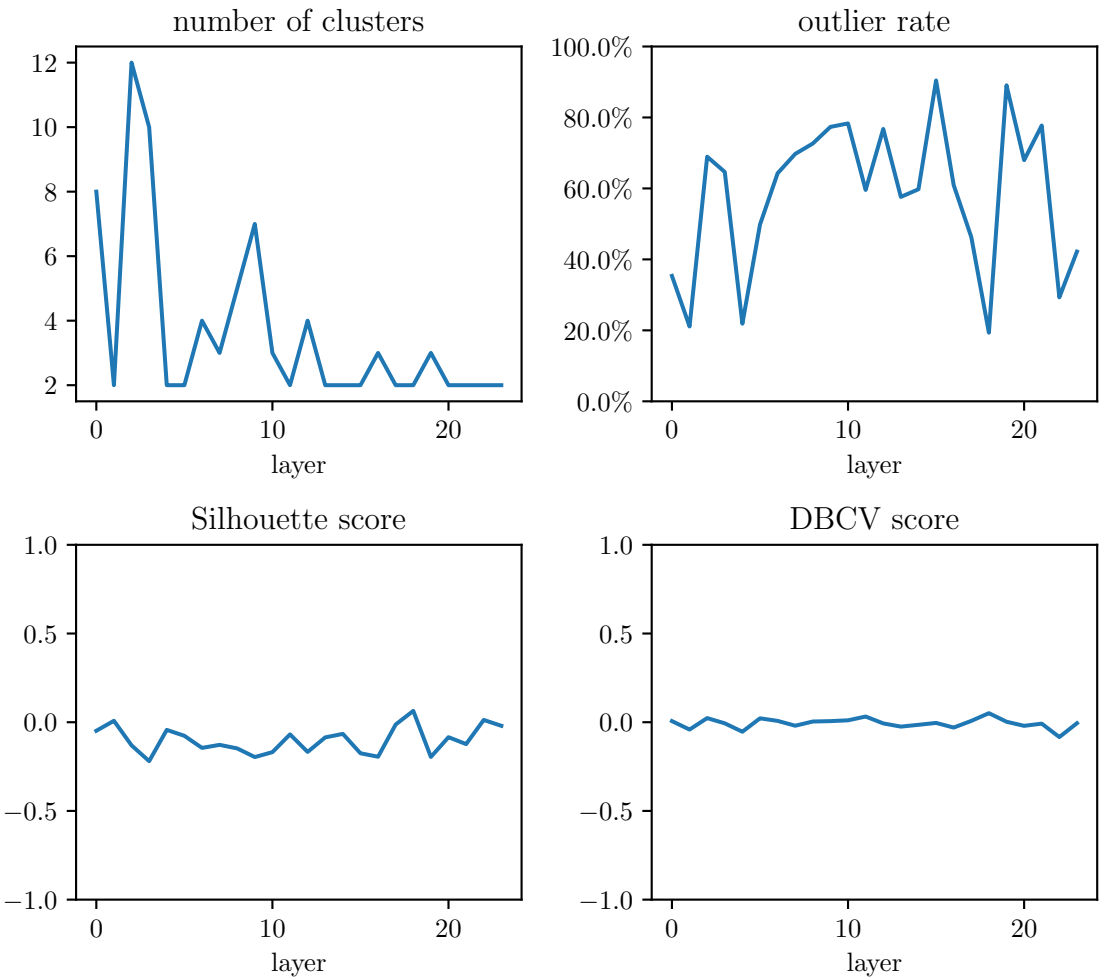
HDBSCAN cluster evaluation of tokens x_i
(imdb_bert-large-uncased_params-random_dense-off_layers-24)
number of clusters
outlier rate
Silhouette score
DBCV score
layer
layer
layer
layer
12
10
8
6
4
2
100.0%
80.0%
60.0%
40.0%
20.0%
0.0%
1.0
0.5
0.0
−0.5
−1.0
0
10
20