token_similarity histograms (imdb_albert-large-v2_params-trained_dense-off_layers-48) after layer 26 after layer 25 after layer 27 after layer 28 5 4 4 4 3 3 3 3 2 2 2 2 1 0.5 -0.50.0 0.5 -0.5 0.0 0.0 0.5 -0.50.0 0.5 1.0 -1.01.0 -0.51.0 1.0 -1.0-1.0after layer 32 after layer 29 after layer 30 after layer 31 5 3 3 3 3 2 2 2 2 1 1 1 1 0 + -1.0 0 + -1.0 -0.50.5 1.0 -0.50.0 0.5 -1.0-0.50.0 0.5 1.0 -0.50.5 -1.01.0 after layer 33 after layer 34 after layer 35 after layer 36 5 5 5 5 3 3 2 2 2 2 1 0 0 -0.5 0.5 0.0 0.5 -1.0-0.50.5 -1.0-0.50.0 1.0 -1.0-0.50.0 1.0 -1.0-0.51.0 0.0 1.0 after layer 39 after layer 37 after layer 38 after layer 40 4 3 3 3 3 2 2 1 1 -0.50.5 -0.5 0.0 0.5 0.0 0.5 -0.50.0 0.5 -1.00.0 -1.0-1.0-0.5-1.01.0 1.0 1.0 1.0 after layer 41 after layer 42 after layer 43 after layer 44 3 3 3 3 2 2 2 1 1 1 0 + -1.0 -1.0 -1.0-0.50.0 0.5 -0.50.0 0.5 -0.50.0 0.5 -0.5 1.0 1.0 -1.01.0 0.0 0.5 1.0 after layer 45 after layer 46 after layer 47 after layer 48 5 -5 5 5 4 4 4 4 3 3 3 3 2 -2 2 -2 1 1 1 1

0

1.0

-1.0

-0.5

0.0

0.5

0 -

-1.0

1.0

-0.5

0.0

0.5

1.0

0 -

-1.0

-0.5

0.0

0.5

-1.0

1.0

-0.5

0.0

0.5