token_similarity histograms (randomtext_albert-large-v2_params-trained_dense-on_layers-48) after layer 26 after layer 27 after layer 25 after layer 28 40 40 40 40 30 30 30 30 20 20 20 20 10 10 10 10 0.0 0.5 0.0 0.5 0.0 0.5 0.0 -0.51.0 -0.51.0 -0.5-0.50.5 -1.0-1.0-1.01.0 -1.01.0 after layer 30 after layer 29 after layer 31 after layer 32 40 40 40 40 30 30 30 30 20 20 20 20 10 10 10 10 0.0 0.0 0.0 0.0 -0.51.0 -1.0-0.51.0 -1.0-0.51.0 -1.0-0.5-1.0after layer 33 after layer 36 after layer 34 after layer 35 40 40 40 40 30 30 30 30 20 20 20 20 10 10 10 10 0 0 0 -1.0-0.50.0 0.5 1.0 -1.0-0.50.0 0.5 1.0 -1.0-0.50.0 0.5 1.0 -1.0-0.50.0 0.5 1.0 after layer 37 after layer 38 after layer 39 after layer 40 40 40 40 40 30 30 30 30 20 20 20 20 10 10 10 10 0.0 0.0 0.0 -0.50.0 -0.50.5 -0.50.5 -1.0-0.50.5 -1.0-1.0-1.00.5 1.0 1.0 1.0 1.0 after layer 41 after layer 42 after layer 43 after layer 44 40 40 40 40 30 30 30 30 20 20 20 20 10 10 10 10 -1.00.0 0.5 -0.50.0 0.5 -0.50.0 -0.5 0.0 -0.51.0 -1.01.0 -1.00.5 1.0 -1.00.5 1.0 after layer 45 after layer 46 after layer 47 after layer 48 40 40 40 40 30 30 30 30 20 20 20 20 10 10 10 10

0

1.0

-1.0

-0.5

0.0

0.5

0

-1.0

-0.5

0.0

0.5

1.0

1.0

0 -

-1.0

-0.5

0.0

0.5

0

-1.0

-0.5

0.0

0.5

1.0