token_similarity histograms (imdb_albert-large-v2_params-trained_dense-on_layers-48) after layer 1 after layer 2 after layer 3 after layer 4 5 5 5 5 4 3 3 3 2 2 1 0.0 0.5 -0.50.0 0.5 -0.50.0 0.5 1.0 -0.50.0 0.5 -0.51.0 -1.01.0 -1.01.0 -1.0-1.0after layer 5 after layer 6 after layer 7 after layer 8 5 5 3 3 2 2 1 1 1 1 -0.50.0 0.5 1.0 -0.50.0 0.5 1.0 -1.0-0.50.0 0.5 1.0 -0.50.0 0.5 -1.0after layer 12 after layer 9 after layer 10 after layer 11 5 5 3 3 3 2 2 2 1 1 0 0 --1.00.5 0.0 0.5 1.0 -0.50.0 0.5 0.0 1.0 -0.50.0 0.5 -1.0-0.51.0 -1.0-0.5-1.0after layer 15 after layer 13 after layer 14 after layer 16 5 5 5 5 3 3 3 2 2 2 2 1 1 1 -0.50.0 0.5 -0.5 0.0 0.5 -1.00.0 0.5 -0.50.0 0.5 1.0 -1.0-0.51.0 -1.0-1.01.0 after layer 17 after layer 18 after layer 19 after layer 20 5 5

