token_similarity heatmaps (imdb_albert-large-v2_params-random_dense-on_layers-48) after layer 25 after layer 26 after layer 27 after layer 28 0 0.5 token_similarity 0.5 token_similarity 0.5 token_similarity 0.5 100 100 100 100 token 500 token :-000 --0.0 200 0.0 0.0 200 0.0 300 300 300 300 400 400 400 400 200 200 200 200 40 0 400 0 40 -1.0-1.0-1.0-1.0token j token j token j token j after layer 29 after layer 30 after layer 31 after layer 32 0 0 0 0 token_similarity 0.5 token_similarity 0.5 token_similarity 0.5 100 100 100 100 token 500 token i 200 0.0 200 0.0 0.0 200 0.0 300 300 300 300 400 400 400 400 200 200 200 200 40 0 40 40 40 -1.0-1.0-1.0-1.0token j token j token j token j 1.0 1.0 1.0 after layer 34 after layer 33 after layer 35 after layer 36 0 0 0 0 0.C co.c co.c similarity 0.5 0.5 token_similarity token_similarity 0.5 0.5 100 100 100 100 token 500 token i 200 token i 200 0.0 0.0 0.0 200 300 300 300 300 -0.5 400 400 -400 400 0 200 0 200 200 400 200 400 400 0 0 -1.0-1.0-1.0-1.0token j token j token j token j 1.0 1.0 1.0 after layer 40 after layer 37 after layer 38 after layer 39 0 0 0.5 token similarity 0.5 0.5 token_similarity 0.5 token_similarity token_similarity 100 100 100 100 token i 200 token i 200 token i token i 200 0.0 200 0.0 0.0 0.0 300 300 300 300 400 400 -400 400 200 200 40(200 40 200 40 0 40 -1.0-1.0-1.0-1.0token j token j token j token j 1.0 1.0 1.0 1.0 after layer 41 after layer 42 after layer 43 after layer 44 0 o c. G. token_similarity o o c token_similarity 0.5 token_similarity 0.5 token_similarity 0.5 0.5 100 100 100 100 token 200 token i token i 0.0 0.0 0.0 200 200 0.0 300 300 300 300 400 -400 400 400 200 200 200 200 0 40 40 40 0 40 -1.0-1.0-1.0-1.0token j token j token j token j after layer 45 after layer 46 after layer 47 after layer 48 0 0 c. 20 5 c. token_similarity o o c c token_similarity 0.5 0.5 100 100 100 100 to 500 token i to ken :token i 0.0 300 300 300 300 400 400 -400 400 200 200 200 200 400 -1.0-1.0-1.0-1.0

token j

token j

token j

token j