cosine T-SNE visualisations (bert-large-uncased_imdb_(x,y)_trained-params_dense-on_24-layers) layer 1 layer 0 layer 2 layer 3 100 100 100 100 50 50 50 50 0 0 -50-50 -50-50 -100 -100-10050 0 50 100 0 50 100 -500 100 -50 0 50 100 -50 -100 -50-100 -100-100 layer 4 layer 5 layer 6 layer 7 100 100 100 100 50 50 50 50 0 -50 -50 -50 -50 -100-100-100 -50 0 50 100 -100 -50 0 50 100 -100-50 50 100 -100-50 50 100 layer 8 layer 9 layer 10 layer 11 100 100 100 100 50 50 50 50 0 0 0 -50-50 -50-100 -100 -100-10050 50 100 50 50 100 -500 100 -100 -500 -100-500 100 -100-50 0 -100layer 12 layer 13 layer 15 layer 14 100 100 100 100 50 50 50 50 -50**-**50 -50-50 -100 -100 -100-100 -50 0 50 -100 -50 0 50 100 -50 0 50 -50 0 50 100 -100 100 -100100 -100layer 16 layer 17 layer 18 layer 19 100 100 100 100 50 50 50 50 -50 -50 -50 -50 --100 | −100 -100 | -100 -100 | -100 -100 | -100 50 0 -50 0 50 100 -50 0 50 100 -50 0 100 -50 50 100 layer 20 layer 22 layer 23 layer 21 100 100 100 100 50 50 50 50 0 0 0 -50-50 -50 -100 --100 --100-100 --100 -50 50 100 -100 -50 50 100 -100-50 50 100 -100 **-**50 50 100