cosine T-SNE visualisation – layer 3 (bert-large-uncased_wikitext_(x,y)_trained-params_dense-on_24-layers) 75 50 25 0 -25-50-75-100 -

50

75

100

25

-50

-25

-100

-75