cosine T-SNE visualisation - layer 8 (bert-large-uncased_wikitext_(x,y)_trained-params_dense-on_24-layers) 75 50 25 0 .3 -25-50-75-10050

0

25

75

100

-100

-75

-50

-25