cosine T-SNE visualisation - layer 6 (bert-large-uncased\_wikitext\_(Qx,Ky)\_trained-params\_dense-off\_24-layers) 75 50 25 0 -25-50-75-100-100-75-50-250 25 50 75 100