cosine T-SNE visualisation – layer 12 (albert-xlarge-v2_uniform-tokens_(Qx,Ky)_trained-params_dense-off_24-layer 75 50 25 0 -25-50-75-100-50 50 -100-75-2525 75 100