

t-SNE embeddings of tokens  $x_i$   
(wikitext\_albert-xlarge-v2\_params-trained\_dense-off\_layers-48, sample 7)

