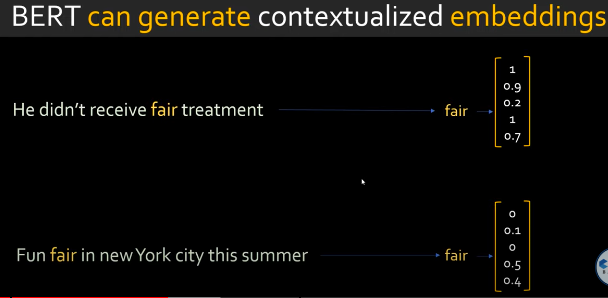
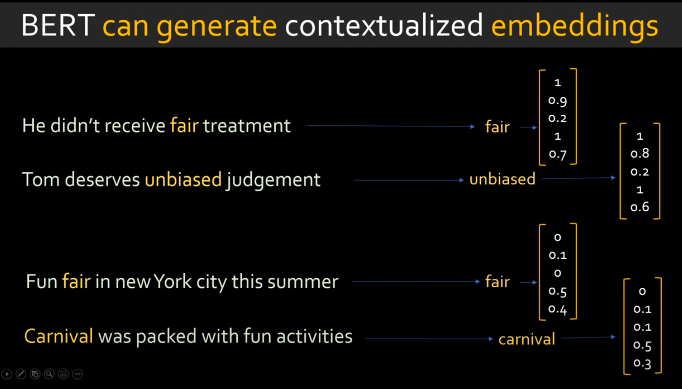
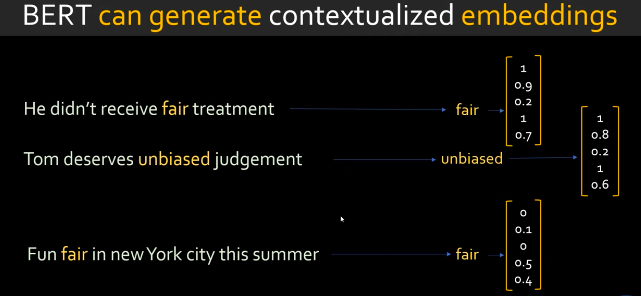
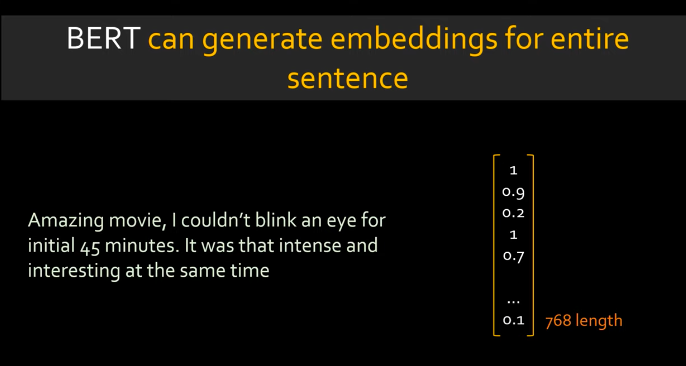
BERT (Bidirectional Encoder Representations from Transformers)

We need a model that will generate contextual meaning of a word.

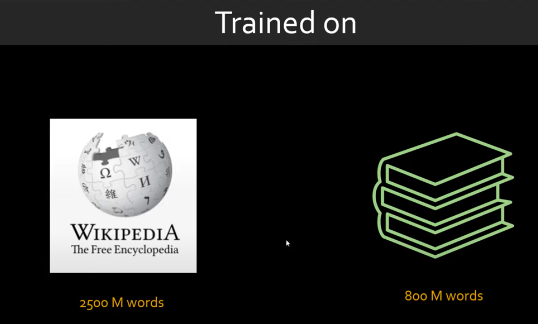
BERT can generate contextualized embeddings



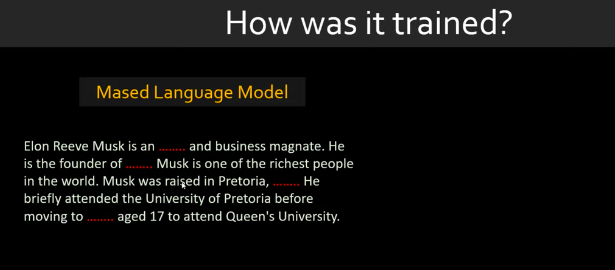
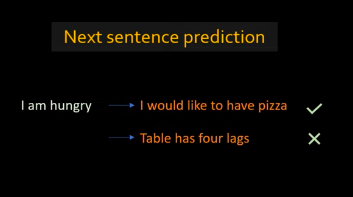


BERT can generate one single vector for the entire sentence

BERT was trained by Google



BERT was trained using 2 approaches

1. BERT was trained by masking the words on wikipedia articles, where as a side effect they got word embeddings
2. The other training approach was Next sentence prediction

Search