NLP with BERT for Sentiment Analysis

Text Classification using Ktrain, which is a wrapper for Tensorflow Keras library

WordEmbeddings (Word2Vec, GloVe, Bert, etc)

Adds two extra tokens to the sentence

At the beginning'[CLS]' At the end'[SEP]'

Len(sen1) = 5

Len(sen2) = 8 => we choose 8 as max. length and add 0's at the end of sen1 (padding)

Now take each sentence and assign a vector for each word in it (imagine every value as one aspect of the word e.g., Masculinity, Age, …)

GloVe => always assigns the same embeddings for a word no matter what the context is

BERT => Gives each word a vector based on the context it was used in

For BERT:

After the tokenization and giving each token a vector value we chose only one token value to represent the sentence and decide the label based on that and that token is [CLS] which stands for classification