

Word2Vec.

Kaleb Azael Avila Padilla.

In the time I was looking for a different way of use for the Word2Vec vectorization that Paty did I encounter some good things about how to use embeddings and vectorization of words and sentences with Word2Vec.

First of with what Paty told me about the project I know that she used the average of the words vectors, something that is of common use in Machine Learning, Deep Learning and text mining in the data science world, but there's another one that is common used and that is the sum of the vectors, I will compare the two of them by researching and conclude if one is the better option for the project, and if I encounter with another use I will compare it too.

In the research I was looking a paper about what are the bests ways to do this was the normalized sum of Word Vectors, actually it it's the best but since it doesn't improve the model they used as an example I thought that in the project maybe it could not work, maybe to try it in an experiment and see if it can work, but if the way in the project works I thinks is better to leave it that way.

References:

<https://www.mdpi.com/2076-3417/8/9/1589> , <https://aclanthology.org/Q15-1016/> ,
<https://datascience.stackexchange.com/questions/110718/sum-vs-mean-of-word-embeddings-for-sentence-similarity> ,