Center for Information and Lang Processing
Prof. Dr. Alexander Fraser

Erweiterungsmodul: Machine Translation **SS 2023**

Exercise Sheet

Monolingual and Bilingual Embeddings

Please submit your writeup in a single PDF file called `yourlastname_yourfirstname_ex5.pdf` (e.g., `fraser_alexander_ex5.pdf`).

Exercise 1. Consider the famous analogy solvable using monolingual word embeddings "King – Man + Woman = Queen"[1].
If Man is labeled a, King is labeled b, Woman is labeled c, and Queen is labeled d, (i.e., we are applying the analogy Man is to King, as Woman is to ?) then we can view the analogy operation as searching for word d, by evaluating the relevant embeddings x for minimal cosine distance (see also Slide 5 in the embeddings lecture):

$$\text{argmin}_d\, cos(x_b - x_a + x_c, x_d) \tag{1}$$

Here we hope that when we set d to "Queen", we will be closest to the left-hand term. If we set d to any of the other words in our vocabulary, d should be farther away from the left-hand term. Suppose you have monolingual word embeddings of dimension 5 which happen to have learned clearly-defined semantic properties along the 5 axes. What are examples of properties that would make sense? Make up 5-dimensional values for these four words as well as "Prince", "Mother" and "Palace". Choose the dimensions and values such that the words are meaningfully embedded, and such that this analogy holds (over these 7 words). Explain how each word is embedded. Show that this analogy does hold.

Exercise 2. Give two examples of linguistic properties (other than part-of-speech and word-sense) that you think cannot be easily captured in a word2vec word-type embedding space, and explain why this is so.

Exercise 3. Suppose we would like to project the English embeddings from Exercise 1 into German, and we have learned the projection reasonably well. We have learned this projection matrix:

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0.1 & 0 & 0 & 0 & 100 \end{matrix}$$

[1] Taken from Linguistic Regularities in Continuous Space Word Representations – Mikolov et al. 2013. NAACL

Give German embeddings for a translation of each of the 7 English words in exercise 1 which will map perfectly (i.e., the projected cosine distance is zero). Explain the semantic differences between the English and German spaces if any. You may find it useful to project a vector where you simply write the semantic properties you defined previously in each dimension. The result of this projection (matrix * vector) will then describe the German space semantically.

Exercise 4. Suppose we have a trained BERT model. Recall that BERT outputs a contextual embedding. Provide 3 example sentences in English containing the word "bank" as follows:

- bank (verb)
- bank (financial, noun)
- bank (river, noun).

Define a semantically meaningful 5-dimensional space (as we did before), that will allow you to disambiguate the three tokens of "bank". Provide sample 5-dimensional embeddings for bank in these sentences.

Exercise 5. Discuss how we can translate the BERT embeddings of bank from the previous exercise to German using bilingual word embeddings. First translate your example sentences to German. Then give sample embeddings for the translation of bank in the second and third examples and discuss the semantics. Finally, discuss how the translation of your first example could be embedded in German (by a German BERT model), and discuss what the correspondence is with the embeddings of the English "bank" token.

Exercise 6. OPTIONAL: carry out Exercise 1 using the freely available package Gensim together with word2vec embeddings (use freely available embeddings or train them yourself). Test whether the analogy holds. Pay particular attention to the relationship of the left-hand term with the embeddings of the words "King", "Man" and "Woman". Discuss your results.

Exercise 7. OPTIONAL: use Gensim (and the same embeddings as before) to show that the two properties you claimed cannot be represented in word2vec embeddings in Exercise 2 really aren't represented well, by presenting the embeddings of relevant example words (with and without these properties), and their distances.