Word Embedding Methods

Classical Methods

- word2vec (Google, 2013)
- *Continuous bag-of-words (CBOW):* the model learns to predict the center word given some context words.
- *Continuous skip-gram / Skip-gram with negative sampling (SGNS)*: the model learns to predict the words surrounding a given input word.
- *Global Vectors (GloVe) (Stanford, 2014)*: factorizes the logarithm of the corpus's word co-occurrence matrix, similar to the count matrix you've used before.
- *fastText (Facebook, 2016):* based on the skip-gram model and takes into account the structure of words by representing words as an n-gram of characters. It supports out-of-vocabulary (OOV) words.

Deep learning, contextual embeddings

In these more advanced models, words have different embeddings depending on their context. You can download pre-trained embeddings for the following models.

- BERT (Google, 2018):
- ELMo (Allen Institute for AI, 2018)
- GPT-2 (OpenAI, 2018)