You will first develop a Turkish Word2Vec model applying the steps in [1] using the latest Wikipedia dump in [2] via Gensim library [3]. There is another Turkish model trained in 2018 [4] that we will also use. Additionally, we need four additional English Word embedding models [5], [6] (fasttext-wiki-news-300, word2vec-google-news-300, glove-wiki-gigaword-300, glove-twitter-200) (Note: You may adjust the dimensions of the vector models based on executability considerations on your own computer or in Google Colab). Then you are able to retrieve the top-1, top-3 and top-10 most similar words to the given words [7].

In the second stage, you will pick two connotation lexicons [8], [9] and [10], [11]. The first lexicon includes the valence, arousal and dominance values whereas the second lexicon has additional features pointing out various emotions. You are going to create a table finding the mean values of these lexical dimensions for top-k words by getting the results for each six Word2Vec models. The test words are 'sport (spor), magazine (magazin), politics (siyaset), economy (ekonomi) '. You should use the related functions in [12] for this operation or you need to parse the lexicon by yourself. If you are unable to access the relevant word from the lexicon, keep the ignorance of that word. For example, print the number of top words as 2, even if you are searching for top 3 words and omitting one word due to its absence in the lexicon.

Note: The test word set will be updated.

[1] https://github.com/akoksal/Turkish-Word2Vec/wiki/
[2] https://dumps.wikimedia.org/trwiki/
[3] https://radimrehurek.com/gensim/
[4] https://drive.google.com/drive/folders/1IBMTAGtZ4DakSCyAoA4j7Ch0Ft1aFoww
[5] https://radimrehurek.com/gensim/models/word2vec.html
[6] https://github.com/piskvorky/gensim-data
[7] https://tedboy.github.io/nlps/generated/generated/gensim.models.Word2Vec.similar_by_word.html
[8] https://saifmohammad.com/WebPages/nrc-vad.html
[9] https://saifmohammad.com/WebDocs/Lexicons/NRC-VAD-Lexicon.zip
[10] https://github.com/JULIELab/MEmoLon
[11] https://zenodo.org/record/3756607/files/MTL_grouped.zip?download=1
[12] https://shifterator.readthedocs.io/en/latest/cookbook/sentiment_analysis.html