

Artificial Intelligence 2021 -2022

Assignment

Word Embeddings

Consider the **pre-trained embeddings** from **Mikolov** et al, 2013 (Google Embeddings).

Use the code provided in the Colab notebook
(https://colab.research.google.com/drive/1S_zbac0r69X8Szh1UGP7MOVCi_pOfKWY?usp=share_link)
to:

1) Find the most similar word to “laptop”, “pillow” in word2vec. 2)
Compute the cosine similarity between “apple”, “banana” and
“book”.

3) Use the most similar function

(https://tedboy.github.io/nlps/generated/generated/gensim.models.Word2Vec.most_similar.html) to find analogies:

a) Find three capital-country analogies that are solved by the

word2vec model (e.g., Rome:Italy=? :France).

- b) Find three verb analogies that are solved by the word2vec model (e.g., swim : ? = eat : eating).
- c) Do the analogies work also for adjectives (like big : bigger?)