

Embedded Words

@author Louise Deeth

@version Java 20.0.1

Description

This is a Java application that performs similarity searches using word embeddings. It uses a dataset of word embeddings to find words with similar meanings based on various metrics.

To Run

From console at .jar file directory and execute the following command:

```
java -cp ./dsa.jar ie.atu.sw.Runner
```

Navigate through the options

Specify the embedding file e.g. ./word-embeddings.txt

This will then load which will be shown by the progress bar

Specify an output file. If none is chosen the default is set as ./out.txt

Enter a word or text.

Once you have entered a word or short sentence, you are given the options of 3 comparison methods.

Choose which one you want to use

You will then be given the choice of how many results to display.

The results will shown together with the average result.

There is also an option configure options which gives you preset options for background and font colours.

Features

- Specify Embedding File: allows users to input a file path for the word embedding to be used in the similarity search.
- Specify Output File: allows users to specify a file path for saving the output of their searches, with a default option available.
- Enter a word or text: users can input a word or short sentence to search for similar words based on the selected similarity metric.
- Configure options: offers customization options for console output, i.e. background and font colours.
- Different ways to compare words:
 - Dot Product
 - Euclidean Distance
 - Cosine Similarity