

# Similarity Search with Word Embeddings

@author - Cian Dicker-Hughes \_ [G00415413@atu.ie](mailto:G00415413@atu.ie)

@version Java 17

## Description

This is a Console-based application that Checks the Similarity of Words by using Word-Embedding. This program will have a HashMap to hold Word-Embedding and the user can pick which word they want to Compare to the rest of the words in the map for Similarity by use Dot Product Comparing Vectors for Similarity. The user can save the results of the Compares to a text Document(.txt)

## To Run

From console at .jar file directory:

```
java -cp .;\dsa.jar ie.atu.sw.Runner
```

When Opened the Application it loads in the in the default File (word-embeddings.txt), but if it can't find (word-embeddings.txt) it will continue without the file and load in the Menu Display which you can then add the new file path to the text file (word-embeddings.txt).

## Features

There six Features in this application

1. Specify Embedding File
  - Input a new path of text file (.txt) to read and load from and put into the HashMap to hold the Word-Embedding.
  - If user Enter no new path or empty, it does nothing
  - If user enters non-existing File Path, it will default to (word-embeddings.txt)
2. Similarity search & Save to File (default: oneSimilarityOut.txt)
  - Compare the Similarity of the Chosen Word, Display the Similarity Words and Save to a text file (oneSimilarityOut.txt is the default file if user don't enter a name).
3. Similarity Search & Save to Map
  - Compare the Similarity of the Chosen Word, Display the Similarity Words and save to a Map so more Compare the Similarity can be added.
4. Save Similarity Search Map to File (default: multipleSimilarityOut.txt)
  - Save The Map that the user Search (3) to a text file (multipleSimilarityOut.txt is the default file if user don't enter a name).
5. Search for Word in Embedding File
  - Search for the word in the Embedding File to see if it exists in the File.
6. Which Embedding File being Used
  - Displays what File is Currently being used, displays how many words are in the file and can Display the full file the user picks yes or no see the full file.
7. When Picking an Option there will be a loading bar.
8. When picking an invalid Option, it will ask the user to enter a valid Option