README

- The aim of the project is to implement a Cross Lingual Document Translator using Statistical Machine Translation Model.
- The use of the translator is to translate a document from English to Dutch and vice-versa.
- Jaccard's coefficient and cosine similarity between the translated document and the actual translation can also be calculated.

REQUIREMENTS TO RUN THE PROGRAM:

- The English and the Dutch corpus should be included in the program directory.
- All the text files should be included in the program directory.

The text files used:
1.input.txt – File to be translated.
2.output.text – translated file.
3.actual.txt – actual translation.

- nltk.stopwords should be installed in the python environment.

RUNNING THE PROGRAM:

- The program can be started by running the Main.py file in command prompt.

The user has to use python Main.py command to run the program.

- If the program runs without any error the following options will be displayed –

1.Train the model.
2.Test sentence to translate.
3.Translate a Dutch document to English.
4.Translate an English document to Dutch.
5.Calculate the Jaccard coefficient.
6.Calculate Cosine similarity
7.For exit.

- Option 1 should be chosen to train the model for English to Dutch translation and vice-versa.
- Option 2 should be chosen to give user the choice to translate an English statement to Dutch statement and vice-versa.

The sentence to be translated can be entered in the given field.

- Option 3 should be chosen to translate input.txt to English and the result is stored in output.txt in the same directory.
- Option 4 should be chosen to translate input.txt to Dutch and the result is stored in output.txt in the same directory.
- Option 5 should be chosen to calculate the Jaccard's coefficient between the translated document and the actual translation i.e. output.txt and actual.txt.

- Option 6 should be chosen to calculate the Cosine similarity between the translated document and the actual translation i.e. output.txt and actual.txt.
- Option 7 should be chosen to exit the program GUI.