**Lab 1: Text comparison I**

**Tasks to do:**

1. Open the files lab1\_ex1.py and lab1\_ex2.py and read the code.
2. Open the file lab1\_ex3.py and read the code.
3. Calculate the frequency vector for each text (**algebra.txt**, **liter.txt**, **calculus.txt**)

and build the frequency matrix for all texts.

1. Calculate a distance matrix for all texts and analyze the texts.
2. Analyze the texts similarity using the distance matrix.

**Independent work:**

1. For comparison add two existing texts (**algebra2.txt** and **liter2.txt**).

Read, write and preprocess the texts.

1. Calculate the frequency matrix for all texts using the dictionary built from “**dictionary.txt**”
2. Calculate the frequency matrix for all texts using the dictionary built from “**general\_ dict.txt**”
3. Calculate distance matrix for all texts
4. Analyze and explain the results.