**instructions how to run the code:**

1. runs in PyCharm versions 3.7, 3.6 (that’s what we work with)
2. download the nltk library
3. put all the 3 files in the same folder
4. run the EX01 -its the an example for the code.

**The group members:**

Amit Tzimerman-304978273

Matan Choen-302608542

Yulia Shoot-311691372

Tanya Grabov-314597337

**Amit**-building the tokenization process ,the invert index and union all the code parts.

**Matan**-building the function that loads the tweets into 100 documents ,union all the code parts, fixing code parts.

**Yulia**-building the intersection process ,helping with the invert index function.

**Tanya**-building the normalization part, helping with the tokenization process.

Everybody-running the codes in their pc environment ,checking that it’s working.