

Document clustering Homework

The goal is to **cluster** the sentences of a book of your choice (eg from <https://www.gutenberg.org/>) using flair document embeddings.

1. Select book and split into sentences.
2. Use **flair document embeddings** for each of the sentences.
3. **Cluster** the sentences based on the embeddings. You can use any clustering toolkit (scipy?) and clustering algorithm and settings which you want.
4. Look at the results -- do they make sense somehow?
5. Experiment a bit with **different settings** to see the differences.