1. Choosing 10 job notices (200 tokens, California, Computer Science) 30min
2. Calculate the features based on fig 4 of survey paper \*\*\*
3. Manually extract keywords 1hr
4. Use YAKE’s demo for extracting keywords 1hr
5. Comparison between a chosen keyword vs its features which we have calculated 2hr

+ single doc: YAKE

+corpus: training a model on training data (hidden keyword)

1.Applying YAKE on our data

\* Reading hidden words paper

2. Reading a paper that uses training on a corpus

Comparison between 1 and 2

Applying \* to see whether it helps us

Figure out what features are we going to choose