Computational Communication Science 2 Week 1 - Lecture »Introduction«

Marthe Möller
Anne Kroon

A.M.Moller@uva.nl, @MartheMoller
a.c.kroon@uva.nl, @annekroon

April, 2022

Digital Society Minor, University of Amsterdam

Today

Introducing...the people

Introducing...the course

Text as Data

Analyzing songtexts

All course materials can be found at...

https://github.com/annekroon/CCS-2

Introducing...the people

Introducing...Marthe

dr. A. Marthe Möller

Assistant Professor Entertainment Communication

- Studying entertainment experiences in the digital space using:
  - Computational methods (e.g., ACA of user comments)
  - Experimental methods

@marthemoller |A.M.Moller@uva.nl |https://www.uva.nl/profiel/m/o/a.m.moller/a.m.moller.html

Introducing...Anne

dr. Anne Kroon

Assistant Professor Corporate Communication

- Research focus on biased AI in recruitment, and media bias regarding minorities
- Text analysis using automated approaches, word embeddings

@annekroon |a.c.kroon@uva.nl |http://www.uva.nl/profiel/k/r/a.c.kroon/a.c.kroon.html

Introducing...You

Your name?
Your background?
Your reason for taking this course?
Do you have a dataset you are working on?

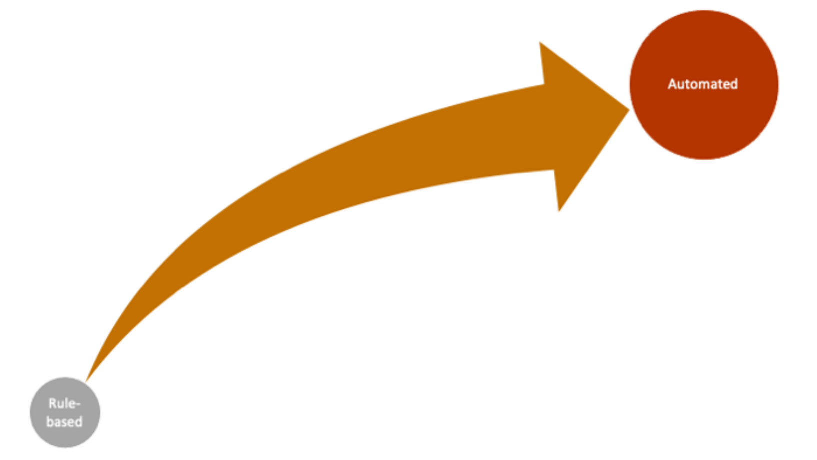
Introducing...the course

About CCS-2

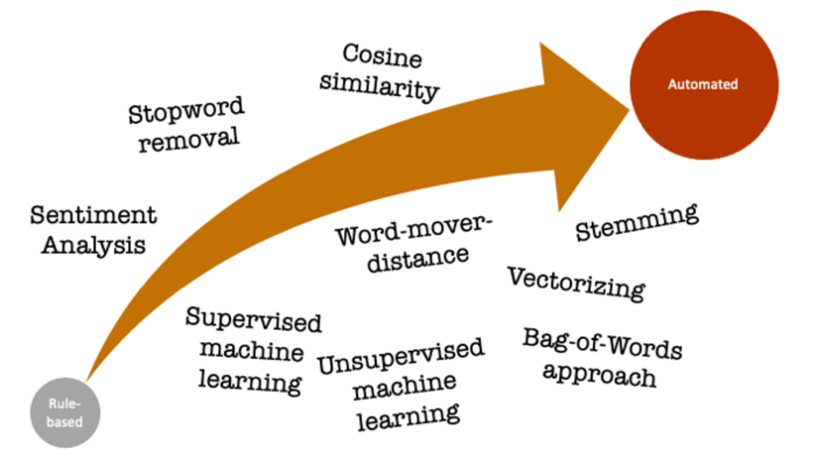
What is CCS-2?

- Next step after CCS-1
- How to use what you learned in CCS-1 for research?
  - Learn computational techniques (e.g. data vectorization, machine learning)
  - Learn how to use these techniques for research (e.g., content analysis)
- By the end of the course, you'll be prepared to do computational research in the Research Project

About CCS-2



About CCS-2


Cosine similarity
Automated
Stopword removal
Sentiment Analysis
Word-mover-distance
Stemming
Vectorizing
Supervised machine learning
Unsupervised machine learning
Bag-of-Words approach
Rule-based

About CCS-2

What will we do in this course?

- We discuss techniques in the lectures
- We practice with techniques in the tutorials
- Graded assignments to master the techniques:
  - Regular multiple choice questions (20%) about readings that use the techniques we discuss
  - Coding challenge (group assignment): Get more experienced with the techniques and build a recommender system
    - Report (20%)
    - Presentation (10%)
  - Take-home exam (50%) at the end of the course so you can show off what you learned
- We provide structure through the meetings and assignments, you do the (home-)work

About CCS-2

How to stay informed and where to find all the materials?
Regularly check:

- The course Canvas page
- Your email
- The course Github page

In addition, make sure that you read the course manual so that you know all the ins and outs of this course!

About CCS-2

How to contact Anne and Marthe?

We kunnen hier eventueel iets over zeggen: mogen ze mailen, zijn er spreekuren etc.?

Ready? Set? Go!

Without further ado...

...let's get started!

Text as Data

Text as Data

CCS-1: You learned how to...

- Work with Python, for example, you:
    - Store text in json-files, csv-files etc.
    - Work with texts in Python

Text as Data: Learning from text directly

Studying text can teach us a lot about human behavior:

What topics do people discuss on online cancer-related platforms?
(Sanders et al., 2020)

To what extent does content differ between online and print news?
(Burggraaff and Trilling, 2020)

What topics do people discuss in their movie reviews?
(Schneider et al., 2020)

Text as Data: Analzing text as a means

Studying text can give us information we can use to answer broader questions:

Analyze textual information about movies from IMDB to learn about the representation of women in movies
(Poma-Murialdo, 2019)

Automatically distinguish between reliable and unreliable online information about vaccines by investigating what characterizes reliable and unreliable texts
(Meppelink et al., 2021)

Text as Data: Combining text analysis with other methods

We can use data about text in combination with other methods:

Combining data about media content and survey data to investigate how media coverage affects citizens' trust in the EU (Brosius et al., 2019)

Analyzing songtexts

Molly Malone

What is the song "Molly Malone" about?

```

^^I^^IMollyMalone = "In Dublin's fair city where the girls are so pretty,
^^I^^II first set my eyes on sweet Molly Malone."
^^I^^I
^^I^^Iprint(MollyMalone)

^^I
```

```
^^I^^IIn Dublin's fair city, where the girls are so pretty, I first set my eyes on sweet Molly Malone.
^^I
```

Molly Malone

```
^^I^^I
^^I^^Iprint(type(MollyMalone))
^^I^^Iprint(len(MollyMalone))
^^I^^Iprint(MollyMalone[0])
^^I^^Iprint(MollyMalone[-1:])
^^I^^I
^^I
```

```
^^I^^I<class 'str'>
^^I^^I96
^^I^^II
^^I^^I.
^^I
```

Molly Malone

```
^^I^^I
^^I^^Iprint(Counter(MollyMalone.split()).most_common(13))
^^I^^I
^^I
```

```
^^I^^I[('In', 1), ("Dublin's", 1), ('fair', 1), ('city,', 1), ('
    where', 1), ('the', 1), ('girls', 1), ('are', 1), ('so', 1),
     ('pretty,', 1), ('I', 1), ('first', 1), ('set', 1)]
^^I
```

Molly Malone

```
^^I^^I
^^Istopwords = ['in', 'the', 'and', 'a','I', 'she', 'her', 'are']
^^Inostopwords = ' '

^^Ifor line in MollyMalone.split():
^^I^^Iif line not in stopwords:
^^I^^I^^Inostopwords=nostopwords+line+' '


^^Iprint(nostopwords)
^^I^^I
^^I
```

```
^^I In Dublin's fair city, where girls so pretty, first set my
    eyes on sweet Molly Malone.
^^I
```

Molly Malone

```
^^I^^I
^^I^^Iprint(Counter(nostopwords.split()).most_common(13))
^^I^^I
^^I
```

```
^^I^^I[('In', 1), ("Dublin's", 1), ('fair', 1), ('city,', 1), ('
    where', 1), ('the', 1), ('girls', 1), ('are', 1), ('so', 1),
     ('pretty,', 1), ('I', 1), ('first', 1), ('set', 1)]
^^I
```