Diskussion Recherche / Zwischenstand Projekte HS Experimentelles Arbeiten in der Sprachverarbeitung

> Nils Reiter nils.reiter@uni-koeln.de

> > December 8, 2022



Recherche Forschungsstand "Fake News Detection"

- Drei Dimensionen
 - Eigener Workflow
 - Stil der Papiere
 - Inhalt der Papiere

Recherche Forschungsstand "Fake News Detection"

Drei Dimensionen

- Eigener Workflow
- Stil der Papiere
- Inhalt der Papiere
- Was ist Ihnen aufgefallen?
- Was hat Sie überrascht?
- ► Was fanden Sie leicht/schwer?

Patronizing and condescending language (PCL)

Features

- 1. Preprocessing
- 2. Train/test-Split
- 3. Hudell-Definition (Lap)
- 4. Evaluation



Historical NER

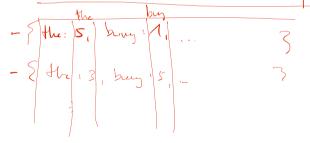
- Woltheise Klassifishing 1. NE? 7. Weden At

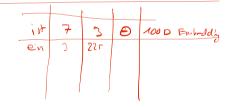


Features

Texthlassiphilatin

- Nontestwich





- 1. Get to know the phenomenon
 - ► Theory: Read literature about it
 - Practice: Annotate it (e.g., on paper!)

- 1. Get to know the phenomenon
 - ► Theory: Read literature about it
 - Practice: Annotate it (e.g., on paper!)
- 2. Establish gold standard

- 1. Get to know the phenomenon
 - Theory: Read literature about it
 - Practice: Annotate it (e.g., on paper!)
- 2. Establish gold standard
- 3. Establish experimental setup
 - Data structures to represent prediction result
 - Write a dummy model as a start
 - Write evaluation code and test it on the dummy model

- 1. Get to know the phenomenon
 - Theory: Read literature about it
 - Practice: Annotate it (e.g., on paper!)
- 2. Establish gold standard
- 3. Establish experimental setup
 - Data structures to represent prediction result
 - Write a dummy model as a start
 - Write evaluation code and test it on the dummy model
- 4. Establish a baseline and evaluate it
 - May require statistical data analysis

- 1. Get to know the phenomenon
 - Theory: Read literature about it
 - Practice: Annotate it (e.g., on paper!)
- 2. Establish gold standard
- 3. Establish experimental setup
 - Data structures to represent prediction result
 - Write a dummy model as a start
 - Write evaluation code and test it on the dummy model
- 4. Establish a baseline and evaluate it
 - May require statistical data analysis
- 5. The core: Train/develop your own model for prediction
 - Evaluate and analyse errors made by own model
 - Improve model and repeat

- 1. Get to know the phenomenon
 - Theory: Read literature about it
 - Practice: Annotate it (e.g., on paper!)
- 2. Establish gold standard
- 3. Establish experimental setup
 - Data structures to represent prediction result
 - Write a dummy model as a start
 - Write evaluation code and test it on the dummy model
- 4. Establish a baseline and evaluate it
 - May require statistical data analysis
- 5. The core: Train/develop your own model for prediction
 - Evaluate and analyse errors made by own model
 - Improve model and repeat
- 6. Document findings and publish them