

# Praktikum Data Mining

## **Merkmalsextraktion mit der Nicht-Negativen Matrixfaktorisierung**

Oliver Fessler

Maria Florusß

Stefan Seibert

Daniel Griebhaber

30. Juni 2014

# Durchführung des Versuchs

## RSS Nachrichten Feeds einbinden und parsen

Was für eine Datenstruktur liefert die Funktion zurück?

adsfhuial asdjfiown fjnisond fjsio jfoe sfjioa posieng

Wie kann auf den Titel und die Beschreibung des RSS-Feeds zugegriffen werden?

Wie kann auf den Titel und die Beschreibung des RSS-Feeds zugegriffen werden?

## Sammeln und speichern aller Worte der aktuellen Artikel aller eingebundenen Feeds

Erklären Sie den Ablauf und die Rückgabewerte der Funktionen `stripHtml(h)` und `separatewords(text)` und nehmen Sie diese in das File `newsfeatures.py` auf.

## Anzeige der Merkmale und der Gewichte

Geben Sie mindestens 3 Merkmale an und zu jedem Merkmal mindestens 3 Artikel, die das jeweilige Merkmal behandeln.

Mit den Parametern `m = 5` und `it = 10` ergeben sich folgende Merkmale:

1. {'obama', 'former', 'mcdonald', 'secretary', 'president', 'chief'} unter anderem mit den Artikeln
  - Obama to nominate former P&G CEO Bob McDonald as veterans secretary
  - From Big House to White House: Ex-Convicts To Be Honored By Obama Administration
  - Biden, Clinton, you're rich. Own it

- UK leaving EU 'would be bad news'
  - What Obama discovered about Iraq
2. {'european', 'including', 'tikrit', 'iraqi', 'inflation', 'across'} unter anderem mit den Artikeln
- ... as Iraq's Shiites answer call to arms
  - Iraqi army presses Tikrit assault as lawmakers scramble to fill posts
  - ISIS declares Islamic State amid battle for Tikrit ...
  - Group: ISIS 'crucifies' men in Syria
  - What are Namibia's 'fairy circles'?
3. {'korea', 'south', 'monday', 'hostile', 'american', 'military'} unter anderem mit den Artikeln
- North Korea says to try two detained U.S. citizens
  - Boeing, Airbus enter bids for \$1.38 billion South Korean refueling plane order: sources
  - N. Korea Preparing to Indict 2 American Tourists
  - Can North Korea take this Hollywood joke?
  - VIDEO: Italy finds 30 bodies in migrant boat
4. {'ukraine', 'ukrainian', 'separatists', 'russian', 'president', 'poroshenko'} unter anderem mit den Artikeln
- Ukraine's Poroshenko urges Putin to tighten borders after violence
  - Israel bombs Gaza after rocket attacks, Hamas gunman killed
  - Russia protests over shooting of cameraman in Ukraine
  - Separatists attack Ukraine base
  - By Numbers: Where Are the Hardest Places to Live in the U.S.?
5. {'family', 'honor', 'police', 'reports', 'young', 'saturday'} unter anderem mit den Artikeln

- Syria fighters hail declaration of Islamic 'caliphate'
  - Islamist rebels in Somalia kill three as Ramadan starts
  - Two police officers killed in bomb blasts near Cairo palace
  - Is a Vast Marine Sanctuary Any Use if You Cant Police It?
  - Gay Pride Parades From Around the World
6. {'first', 'india', 'economy', 'years', 'decades', 'minister'} unter anderem mit den Artikeln
- Exclusive: Facebook scores record 1 billion interactions for World Cup
  - India's Modi eyes first labor overhaul in decades to create jobs
  - Israel cranks up media campaign ahead of Iran nuclear deal deadline
  - Eyes on defense deals, Western powers rush to court India's Modi
  - Stocks set for quarterly gains, yen at five-week high