
Detektion von Fake News

Lars Möllerherm &
Clemens Vorsmann

June 4, 2019

Fakultät Physik

Definition und Motivation

■ Motivation:

- Fake News wurde zum Anglizismus des Jahres 2016 gewählt[2].
- Größer werdende Rolle in unserer Gesellschaft.
- Falsch-Informationen werden besser.
- Natürlicher "Bullshit"-Filter funktioniert nicht mehr.

■ Definition:

- Es existiert der BS-Detector von Daniel Sieradski[3].
- Beinhaltet Liste fragwürdigen URL's. Artikel mit Links auf diese URL's, werden als BS gekennzeichnet.
- Ist es möglich, anhand des Sprachmusters, diese Kategorisierung vorzunehmen?
- Mögliche Fake News Erkennung ohne URL Blacklist



[1]

Beschreibung des Datensatzes

	title	text	site_url	label
0	Muslims BUSTED: They Stole Millions In Gov't B...	Print They should pay all the back all the mon...	100percentfedup.com	fake
1	Re: Why Did Attorney General Loretta Lynch Plea...	Why Did Attorney General Loretta Lynch Plead T...	100percentfedup.com	fake
2	BREAKING: Weiner Cooperating With FBI On Hilla...	Red State : \nFox News Sunday reported this mo...	100percentfedup.com	fake
3	PIN DROP SPEECH BY FATHER OF DAUGHTER Kidnappe...	Email Kayla Mueller was a prisoner and torture...	100percentfedup.com	fake
4	FANTASTIC! TRUMP'S 7 POINT PLAN To Reform Heal...	Email HEALTHCARE REFORM TO MAKE AMERICA GREAT ...	100percentfedup.com	fake
5	Hillary Goes Absolutely Berserk On Protester A...	Print Hillary goes absolutely berserk! She exp...	100percentfedup.com	fake

	title	content	publication	label
0	House Republicans Fret About Winning Their Hea...	WASHINGTON — Congressional Republicans have...	New York Times	real
1	First, a Mixtape. Then a Romance. - The New Yo...	Just how is Hillary Kerr, the founder of ...	New York Times	real
2	Calling on Angels While Enduring the Trials of...	Angels are everywhere in the Mufiz family's ap...	New York Times	real
3	U.S. Plans to Step Up Military Campaign Agains...	ABU DHABI, United Arab Emirates — The Obama...	New York Times	real
4	272 Slaves Were Sold to Save Georgetown. What ...	WASHINGTON — The human cargo was loaded on ...	New York Times	real
5	Among Travelers and Commuters, the Homeless ST...	Wilson Silva said he knew the homeless situati...	New York Times	real

[4]

■ Fake News:

- <https://www.kaggle.com/mrisdal/fake-news>
- CC0: Public Domain License
- 25.Oct-25.Dec 2016
- 12999 Artikel

■ Real News:

- <https://www.kaggle.com/anthonymc1/fake-news-classifier-final-project>
- gathered in Kernel
- Oct-Dec 2016
- 15712 Artikel

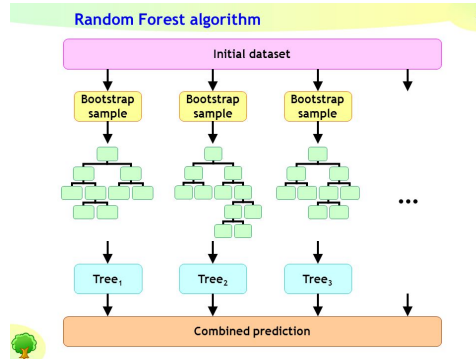
■ Information in beiden Datensätzen:

- Titel
- Inhalt
- Veröffentlichung
- URL

■ Nutzen von Natural Language Processing um

Mögliche Alternativ-Methode

- Random Forest Tree
- Vektor mit Worthäufigkeiten als Input
- Klassifizierung anhand des Blattes
- Trainingszeit geht linear mit Anzahl der Bäume (Parallelisierung)



Pavel Polishchuk



[Skeptiker] Schweiz. URL: <https://www.skeptiker.ch/was-sind-fake-news/>.



Anglizismus des Jahres. URL:
<http://www.anglizismusdesjahres.de/anglizismen-des-jahres/adj-2016/>.



B.S. Detector. URL: <https://github.com/selfagency/bs-detector>.



Fake News Classifier - Final Project. URL:
<https://www.kaggle.com/anthonymc1/fake-news-classifier-final-project>.