Cybersecurity & AI

Freie Universität Berlin

Two Fakeblock Classifiers - Suggestions

Friedrich Müller
Cybersecurity and AI Research Group,
Freie Universität Berlin

14.03.2022

Cybersecurity & AI

Freie Universität Berlin

- Keggle FakeNews dataset German
- GermanFakeNC

src_id	url	title	text	date	fake	src_name
838055	http://www.kleinezeitung.at//kultur...	US-Medien:...	Die Übernam...	2018-02-26 0...	0	news_csv
838056	http://www.kleinezeitung.at//kultur...	90. Oscars: N...	Verwechslun...	2018-02-26 0...	0	news_csv
838057	http://www.kleinezeitung.at//kultur...	Die TV-Tipps...	Achtung Abz...	2018-02-26 0...	0	news_csv
838067	http://www.kleinezeitung.at//leute/...	Jennifer Law...	Die Oscar-Pr...	2018-02-26 0...	0	news_csv
838069	http://www.kleinezeitung.at//leute/...	Chrissy Teige...	Das schwan...	2018-02-26 0...	0	news_csv
838144	http://www.kleinezeitung.at//intern...	Lehrer entgi...	55-Jähriger...	2018-02-26 0...	0	news_csv
838145	http://www.kleinezeitung.at//wirtsc...	Warum die T...	Aus Angst vo...	2018-02-26 0...	0	news_csv
838147	http://www.kleinezeitung.at//wirtsc...	...	Am Voraben...	2018-02-25 0...	0	news_csv
838148	http://www.kleinezeitung.at//kultur...	Dirigent Gus...	Gustav Kuhn...	2018-02-26 0...	0	news_csv
0	https://schluesselkindblog.com/201...	Schluesselki...	This Page Is...	2017-08-30	1	GermanFake...
1	http://blauerbote.com/2017/12/18/bi...	BILD-Journali...	Es ist schon...	2017-12-18	1	GermanFake...
2	http://blauerbote.com/2017/06/02/a...	Angela Merk...	Auch heute...	2017-06-02	1	GermanFake...
3	http://www.truth24.net/gruppenver...	Gruppenver...	Bad Soden /...	2018-02-17	1	GermanFake...
4	http://www.rapefugees.net/ozapft-i...	Ozapft Is! Tü...	Bereits der er...	2016-09-18	1	GermanFake...

Figure: merged datasets

Classifiers

- ► "Leudom" classifier (Dominik Leuzinger)
  - ► Title+Text concatinated
  - ► Stopword list
  - ► Stemming
  - ► Count Vectorization
  - ► Classifier: Catboost
    - ► (slightly worse than e.g. XGBoost, but 3.5 times faster)
    - ► (and some other advantages)
- ► Reis et al. classifiers
  - ► Title+Text concatinated (= stories?)
  - ► extensive state of the art textual feature extraction
  - ► Classifiers:
    Nearest Neighbors (KNN), Naive Bayes (NB), Random Forests(RF), Support Vector Machine with RBF kernel (SVM), and XGBoost (XGB)
  - ► (trained on 2282 BuzzFeed news articles related to the 2016 U.S.)

Performance Figures

Table 1. Results obtained for different classifiers w.r.t AUC and F1 score.

Classifier	AUC	F1
KNN	0.80±0.009	0.75±0.008
NB	0.72±0.009	0.75±0.001
RF	**0.85±0.007**	**0.81±0.008**
SVM	0.79±0.030	0.76±0.019
XGB	**0.86±0.006**	**0.81±0.011**

RF and XGB performed best.

Figure: Leudom / catboost

```
INFO:root:f1_score for testdata: 0.8681434599156118
INFO:root:precision_score for testdata: 0.8878101402373247
INFO:root:recall_score for testdata: 0.849329205366357
INFO:root:Classification report for testdata:
              precision    recall  f1-score   support

           0       0.99      0.99      0.99     11599
           1       0.89      0.85      0.87       969

    accuracy                           0.98     12568
   macro avg       0.94      0.92      0.93     12568
weighted avg       0.98      0.98      0.98     12568

INFO:root:Confusion matrix for testdata:
[[11495   104]
 [  146   823]]
```

Figure: Reis Classifiers

THANK YOU

friedrich.mueller@fu-berlin.de

References

https://dagshub.com/leudom/german-fake-news-classifier

Supervised Learning for Fake News Detection, Reis et al., 2019