A proposal for the Machine Learning Application
Research Paper

March 26, 2020

Optimising online documents for fact-checking

Group:

Alpha

Members:

Alexander Peikert	alexpeikert@uni-koblenz.de	218200812
Clemens Steinmann	csteinmann@uni-koblenz.de	218200209
Erwin Letkemann	erwinletkemann@uni-koblenz.de	218200352
Julian Dillmann	juliandillmann@uni-koblenz.de	218100919

WeST
Universität Koblenz-Landau
Germany

1 Introduction

1.1 Problem

With news media consumption rapidly growing on the web, it is fairly difficult to quickly find news sources that are genuine, due to the nature of the Internet. Easy access and possibilities to publish more news articles quickly overshadow efforts of proper fact checked journalism and in the worst case even distribute miss information. At the same time developing measures to fact check, in order to detect and prevent the spread of fake news is difficult.

Most methods in fact checking are still overall done manually. Hence fact checking is time and cost intensive, even though most organisations already focus on specific topics, like political statements. A verification on general topics in media outlets is generally disregarded due to even more intensive labour.

An automated detection system would be a good solution in latter case, but the bigger complexity of full news articles, on diverse topics, and no uniformity between different news providers, for example regarding the structure articles. Such a system seems impossible with today's technologies and the sheer amount of available and published data make manual verification for a bigger spectrum of topics impossible.

1.2 Our solution

In the course of the project we are going to develop a web-based application to help to optimise the preparations for future fact checking of news on a wider variety of topics.

We are evaluating English news article, by providing "Information nutritional labels" [1] and a ranking system to suggested news articles based on similar topics.

The the service is aimed at two different user groups with different benefits:

1. Mainly, our web service can be used by fact checkers as quick reference to find articles that are worth for further examination.

   At the same time our database can be used, by researcher, on the topic of computational detecting of fake news, by creating a tool to label news articles in order to filter down the mass of available data.

   The uniform structure of our solution provides an automated way to speed up the preparation phase of further work on news checking.

2. On the other side, ordinary web users can find similar articles for a given news article.

   The user can save time by using the provided labels as help to self judge the genuineness of an article before reading. Furthermore the suggested articles provide an opportunity to find other sources on the same topic.

   We hope to bring more awareness for people to reflect on the kind of news they make use of.

2 Related work

Related work on automated tools for fact checking focus on identification of "check worthy" statements or articles. ClaimPortal [6] is such an instance where tweets from several US politicians are analysed and assigned with a check worthiness score, in the regard if the factual claim in the post, if present, is of importance for the public.

Research on determining the check worthiness of political claims was also done with the help of nutritional labels and word embeddings [5].

The authors used existing methods in natural language processing to determine the nutritional labels and also used Word2Vec model and spaCy to determine key phrases in sentences. They hypothesized a correlation between the different implications of the labels and the veracity of claims, for example high emotions in political statements indicate an attempt to deceive the audience. Hence the statement should be check worthy.

All three of these were used in different combinations in several different ML algorithms to evaluate the performance on the Checkthat! 2018 [7] dataset. The paper concluded that the nutritional labels with word-embedding with stochastic gradient descend logless archived the best results, and outperformed the winning team from the Checkthat! 2018 challenge.

These results showcase that nutritional labelling could be of use in further research in computational detection of misinformation in a bigger range of topics or automated help tools for manual fact checking. Therefore we develop a solution to assist researchers and fact checkers to combat misinformation.

3 Our Application

3.1 Needed Components:

- Database
- Content extractor
- Content NLP (Natural language processing)
- Learning to rank algorithm for similar articles
- Web-application

3.2 General procedure

Given an input URL to an English news article, the applications' database will check whether the url has already been checked by our system.

1. If the news article is present in our database we will output the already calculated labels and if needed find the x-best matching articles in our database.

2. Otherwise the news article content extractor will crawl the html to extract the plain text of the news article and each of the labels will be evaluated with corresponding natural language processing methods. The result will be added to the database which leads us to 1.

3.3 Qualities of the components

- The database should hold: URL, metadata, labels, topic, published date

- Similarity between two news articles describes the general topic in our use case. Consequently the ranking of different articles on the same topic is done by comparing labels.

- We hypothesise certain combinations of label scores indicate check worthiness of articles or provide data for research on specific topic, for example high technicality and bad readability, emotional language and controversial topic.

3.4   Basic ideas & implementation ideas for labels.

- Factuality/ Opinion: We think the ratio between factual and opinion sentences in an article helps as the quick reference for check worthiness; [1] NLP to calculate the ratio of factual to opinion sentences over the article content.

- Emotion: We think that emotional language is used to influence the reader into believing certain stances; [3], [4] NLP of sentences and sum over the text will tell the ratio of positive and negative emotions.

- Controversy: We think that controversial topics are more likely to be used as clickbait and may not showcase factual statements; checking for the number of included topics for example from "Wikipedia:List of controversial issues".

- Credibility: As we are likely dealing with text from a big spectrum of topics, the origin of the article has to be considered too; checking the metadata with available resources like "Wikipedia:Reliable sources/Perennial sources" to find reliable sources.

- Technicality: We think an overall technical text could be used to distract from non factual statements; - .

- Virality: We think viral topics are more likely to have contradicting statements between sources; Use google trends for detection.

4   References

[1] Fuhr, Norbert et al. "An Information Nutritional Label for Online Documents." SIGIR Forum 51 (2018): 46-66.

[2] Sahu, Ishan & Majumdar, Debapriyo. (2017).Detecting Factual and Non-Factual Content in News Articles. 1-12. 10.1145/3041823.3041837.

[3] AFFIN Database Informatics and Mathematical Modelling, Technical University of Denmark (2011)

[4] Cambria, Erik et al. "The Hourglass of Emotions." COST 2102 Training School (2011).

[5] Cédric Lespagnol, Josiane Mothe, and Md Zia Ullah. 2019. Information Nutritional Label and Word Embedding to Estimate Information Check-Worthiness. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'19). Association for Computing Machinery, New York, NY, USA, 941–944. DOI:https://doi.org/10.1145/3331184.3331298

[6] Majithia, Sarthak & Arslan, Fatma & Lubal, Sumeet & Jimenez, Damian & Arora, Priyank & Caraballo, Josue & Li, Chengkai. (2019). ClaimPortal: Integrated Monitoring, Searching, Checking, and Analytics of Factual Claims on Twitter. 153-158. 10.18653/v1/P19-3026.

[7] Agez, Romain et al. "IRIT at CheckThat! 2018." CLEF (2018).