

UNIVERSITY OF TARTU
Institute of Computer Science
Conversion Master in IT Curriculum

Peeter Niidas

Debunking characterization of fact-checking articles

Master's Thesis (15 ECTS)

Supervisor(s): Shakshi Sharma, MSc
Rajesh Sharma, PhD

Tartu 2023

Debunking characterization of fact-checking articles

Abstract:

Nowadays, as we are facing fast growth and propagation of different information, it is vital to understand the characteristics of fact-checking articles to combat disinformation. This thesis aims to investigate debunking the characterisation of fact-checking articles by analysing their semantic and linguistic properties and social media influence. The thesis investigates the influence of negations, emotions, multilingualism and social media by using text analysis tools such as regex and multiple NLP libraries. Social media analysis goes deeper by detecting and analysing the dataset's URLs.

The results show some role of negations, emotions and overall tonality in fact-checking articles. The results shed light on the influence of mentioned linguistic and semantic features on their effectiveness. The thesis also investigates the influence of emotions and social media mentions of fact-checking articles. The thesis aims to provide preliminary information hidden in these connections and guide future research to improve understanding of content in fact-checking articles.

For future guidelines, this thesis proposes developing automated tools, improving media literacy education, exploring cross-platform analysis and evaluating the impact of fact-checking strategies. These strategies could enhance understanding of the situation of misinformation and contribute to developing a more effective fact-checking methodology. As most of the research in this field is fractured and every research team has followed their methodology, this thesis emphasises the importance of developing standardised techniques to speed up future research and achieve comparability.

Keywords:

Fact-checking, Fake News, Online Social Media, Rumour Detection

CERCS:

P170 Computer science

Faktikontrolli artiklite omaduste selgitamine

Lühikokkuvõte:

Erineva teabe hulga kiire kasvu ja leviku ajastul on valeinformatsiooniga võitlemisel oluline mõista fakte kontrollivate uudiste omadusi. Lõputöö eesmärk on uurida faktikontrolli uudistele iseloomulikke tunnuseid, analüüsides tekstis kasutatud semantilisi ja keelelisi tunnuseid ning sotsiaalmeedia mõju. Lõputöös uuriti eituste, emotsioonide, mitmekeelsuse ja sotsiaalmeedia mõju, kasutades tekstianalüüsi tööriistu nagu regex ja mitmed keeletötluse (NLP) teegid. Sotsiaalmeedia analüüsi osas tuvastati ja analüüsiti täiendavalt ka uudistes viidatud hüperlinkide mõju.

Analüüsi tulemused näitasid eituste, emotsioonide ja üldise toonaalsuse mõningast rolli fakte kontrollivates uudistes. Tulemused avavad tõeste ja väärte väidete tuvastamisel kasutatud keeleliste ja semantiliste tunnuste rolle. Lõputöös uuriti ka emotsioonide ja sotsiaalmeediale viitamise mõju faktikontrolli uudistele. Lõputöö eesmärk on anda ülevaade eelpool nimetatud seostes leiduvatest omadustest ja anda suuniseid tulevastele uuringutele parandamaks faktikontrolli uudiste sisust arusaamist.

Lõputöös on välja toodud ka suunised automatiseeritud tööriistade väljatöötamiseks, meediapädevuse hariduse parandamiseks, platvormidevahelise analüüsi uurimiseks ja faktikontrolli strateegiate mõju hindamiseks. Nimetatud suunised võimaldavad parandada arusaamist valeinformatsiooni levikust, et aidata välja töötada tõhusamaid faktide kontrollimise meetodikaid. Kuna antud teemal on tehtud mitmeid uurimustöid, kuid igas töös on kasutatud erinevat meetodikat, tuuakse lõputöös välja ka standardiseeritud tehnikate väljatöötamise olulisus tulevaste uuringute kiirendamiseks ja võrreldavuse saavutamiseks.

Võtmesõnad:

Faktikontroll, Libauudised, Sotsiaalmeedia, Libauudiste Tuvastamine

CERCS:

P170 Computer science

VI. Licence

Non-exclusive licence to reproduce thesis and make thesis public

I, **Peeter Niidas**,
(author's name)

1. grant the University of Tartu a free permit (non-exclusive licence) to:
reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, my thesis,

Debunking characterization of fact-checking articles,
(title of thesis)

supervised by Shakshi Sharma and Rajesh Sharma.
(supervisor's name)

2. I grant the University of Tartu the permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work from 06/05/2026 until the expiry of the term of copyright,
3. I am aware that the author retains the rights specified in points 1 and 2. .
4. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Peeter Niidas
09/05/2023