UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA
DIPARTIMENTO DI INFORMATICA, SISTEMISTICA E COMUNICAZIONE

Review of the Doctoral Thesis of Ms Aleksandra Nabożny, M.Sc.

Reviewer: Prof. Marco Viviani, University of Milano-Bicocca

Doctoral Thesis Title: Credibility Evaluation of Online Health Information using Human in the Loop Machine Learning

Summary

This Doctoral Thesis focuses on evaluating the credibility of medical content on the Web. In particular, the Doctoral Thesis aims to improve the process of evaluating *Online Health Information* (OHI) by experts, by creating an expert-supported, semi-automated system for capturing and tagging unreliable medical texts appearing on the Web.

To achieve this, three experiments were conducted to collect data with expert assessments of medical content. These experiments involved evaluating single sentences, both with and without considering the context of the entire medical text. The results indicated the difficulty experts faced when assessing sentences without context, leading to a second experiment that tested methods for enriching sentence context. This led to the identification of an efficient unit of text for evaluation, consisting of three consecutive sentences with keywords.

The Doctoral Thesis includes four separate articles describing the experiments and analyses conducted:

- The first article discusses the experiments and data analysis;
- The second article focuses on creating a dataset evaluated by domain experts from psychiatry, gynecology, cardiology, and pediatrics. This article also explores the detection of rhetorical patterns that mislead experts and distort credibility assessments;
- The third article presents filtering classifiers developed to enhance the efficiency of expert annotation;
- The fourth article examines the explanatory capabilities of the filtering classifier results.

The qualitative analysis performed in this Doctoral Thesis reveals the presence of repetitive rhetorical patterns in non-credible medical content, including patterns found in disinformation

and specific to popular science medical content. The created classifiers enable pre-filtration, which speeds up the identification of unreliable content by annotating experts. The explanatory capabilities of classifiers depend on the compression of input data, with better generalization limiting insight into semantic attributes and narrower generalization requiring separate classifiers for specific domains.

Considering the impossibility of having a large group of experts evaluate all published online data in real-time, the Doctoral Thesis focuses on maximizing the throughput of the expert-supported assessment system. Through the use of three-sentence text fragments as meaningful units for crowd-sourced data collection and the implementation of filtering classifiers, the throughput of the system is improved twofold. The analysis of filtering algorithms also aids in selecting parameters to provide desired feedback for end-users.

The Doctoral Thesis concludes with a qualitative analysis indicating that cognitive biases partially distort medical expert assessments, highlighting the need for further research in the psychology of disinformation to develop the envisioned system of a comprehensive evaluation of online medical content.

Overall Assessment

The Doctoral Thesis addresses a highly relevant issue, specifically focusing on medical misinformation within the context of the Human-in-the-Loop Machine Learning paradigm. It highlights the interdisciplinary nature of the problem, which is a positive aspect of the Doctoral Thesis.

Creating a dataset like the one presented in the Doctoral Thesis is undoubtedly valuable for identifying medical disinformation within the Human-in-the-Loop framework and can make a significant contribution to the scientific community working on this problem.

The involvement of medical experts from different disciplines in doctoral research activities increases the value and generalizability of the work.

Understanding the syntactic, semantic, and rhetorical structure of the new units of medical disinformation and investigating the relationship between manipulation types in the medical domain and cognitive biases are innovative aspects of the research compared to the existing literature.

The publication venues, including two B-class conferences, one Q1/Q2-category journal, and one Q1-category journal, demonstrate the interest, usefulness, and quality of the work conducted.

However, there are also some aspects that could have been improved. Given the nature of the problem and the numerous definitions of credibility that have been proposed so far, more

attention and detail could have been devoted to investigating these aspects. The provided definitions could have been supplemented with additional ones from the literature.

More reference could have been made to the fact that health misinformation is identified using both classification approaches and Information Retrieval, particularly in the context of Consumer Health Search tasks, where recent works are trying to consider salient text portions for retrieving relevant and credible medical information.

It would have been better to have quantitative results in addition to the qualitative outcomes, especially regarding the identification of optimal text portions for assessing the credibility of medical information. However, it is understood that in the Human-in-the-Loop paradigm, this may be more challenging compared to evaluating fully automated tools.

Overall, the organization and writing quality of the Doctoral Thesis could have been improved. At times, there is a lack of logical connections between sections, or introductory parts explaining what the reader will find in the subsequent subsections are missing. Some section titles are excessively long. Additionally, there are numerous typos and inconsistencies in typography, such as the use of italics, quotation marks, and citation styles, which could have been more carefully addressed in a Doctoral Thesis.

Having said that, in summary, the work is solid and valid. Regarding future developments, the most important aspect to consider is the incorporation of evaluation methods that add quantitative aspects in addition to the qualitative aspects.

Strengths and Weaknesses

Strengths of the Work:

- Focuses on evaluating the credibility of medical content on the web, addressing an important problem of medical misinformation.
- Aims to improve the evaluation process of Online Health Information (OHI) by experts.
- Identifies an efficient unit of text for evaluation consisting of three consecutive sentences with keywords.
- Reveals repetitive rhetorical patterns in non-credible medical content, including patterns found in disinformation and specific to popular science medical content.
- Develops filtering classifiers that enable pre-filtration, speeding up the identification of unreliable content.
- Involves medical experts from different disciplines increasing the value and generalizability of the work.
- Investigates the syntactic, semantic, and rhetorical structure of medical disinformation units and the relationship between manipulation types and cognitive biases.

- Published in reputable venues (conferences and journals) demonstrates the quality and interest of the work.

Weaknesses of the Work:

- Could have provided more attention and detail in defining credibility, considering the varied definitions proposed in the literature. Furthermore, sometimes the concepts of misinformation and disinformation are used interchangeably, even if they are not the same and even if the author herself stated this.
- Could have referenced the use of Information Retrieval in identifying health misinformation, particularly in the Consumer Health Search context.
- Qualitative analysis predominates, and more quantitative results could have been included.
- Organization and writing quality could have been improved, with better logical connections between sections and clearer introductions.
- Typos and inconsistencies in typography should have been addressed more carefully.
- Could have explored more ways to incorporate quantitative evaluation methods alongside qualitative aspects.

Detailed Comments

Importance of the Problem:

The Doctoral Thesis addresses a crucial issue concerning the evaluation of credibility in medical content on the web. Given the prevalence of medical misinformation, this is a highly relevant and timely topic. The research focuses on improving the evaluation process of Online Health Information (OHI) by experts, which is essential for ensuring accurate and trustworthy medical information is available to the public.

Interdisciplinary Approach:

The Doctoral Thesis acknowledges the interdisciplinary nature of the problem, which is a commendable aspect of the work. By involving medical experts from different disciplines such as psychiatry, gynecology, cardiology, and pediatrics, the research ensures a broader perspective and increases the generalizability of the findings. This interdisciplinary approach strengthens the validity and applicability of the proposed expert-supported system.

Experimental Design:

The Doctoral Thesis presents three well-designed experiments to collect data for expert assessments of medical content. The experiments include evaluating single sentences both with and without considering the context of the entire medical text. This approach provides

valuable insights into the challenges faced by experts and highlights the importance of context in assessing credibility accurately.

Identification of Efficient Evaluation Units:

The Doctoral Thesis successfully identifies an efficient unit of text for evaluation, consisting of three consecutive sentences with keywords. This is a significant contribution, as it helps streamline the evaluation process and allows for more efficient identification of unreliable medical texts. By focusing on meaningful text fragments, the proposed system can handle a larger volume of online content, thereby maximizing the throughput of the expert-supported assessment system.

Rhetorical Pattern Analysis:

The qualitative analysis performed in the Doctoral Thesis reveals the presence of repetitive rhetorical patterns in non-credible medical content. This finding is valuable, as it sheds light on the characteristics of disinformation and popular science medical content. Understanding these patterns allows for the development of effective filtering classifiers, enabling pre-filtration and faster identification of unreliable content. This aspect of the work contributes to the overall goal of improving the credibility evaluation process.

Limitations and Future Directions:

The Doctoral Thesis acknowledges some limitations, such as the qualitative nature of the analysis and the absence of quantitative results in certain aspects. This recognition of limitations is commendable, as it highlights areas for future improvement. The Doctoral Thesis could benefit from further quantitative analysis to complement the qualitative findings. Additionally, exploring different approaches, such as incorporating Information Retrieval methods, could enhance the evaluation of medical content credibility further.

Publication Quality:

The work's publication in conferences and journals of varying categories demonstrates its relevance and quality. The acceptance of the articles in reputable venues confirms the significance and interest of the research within the scientific community.

Writing and Organization:

While the work presents valuable findings, there are some areas where improvements could be made. The organization of the Doctoral Thesis could be enhanced by establishing clearer connections between sections and providing better introductions to subsections. Additionally, the writing quality, including typos and inconsistencies in typography, could be improved to ensure a more polished presentation of the research.

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA
DIPARTIMENTO DI INFORMATICA, SISTEMISTICA E COMUNICAZIONE

Overall Assessment

In conclusion, the Doctoral Thesis addresses an important problem in evaluating the credibility of medical content on the web. It demonstrates strengths in its focus, interdisciplinary approach, experimental design, identification of evaluation units, analysis of rhetorical patterns, and publication quality. However, improvements can be made in defining credibility, referencing related literature, incorporating quantitative analysis, and enhancing writing and organization. Overall, the work provides valuable insights into the evaluation of online medical content and contributes to the ongoing efforts in combating medical misinformation.

Distinction Assignment to the Doctoral Thesis

Based on the publication record and the reading of the manuscript, it is evident that the Doctoral Thesis is a pretty good and solid piece of work. The Doctoral Thesis has several strengths, such as its focus on a relevant problem, the use of experiments to collect data, the identification of efficient evaluation units, and the analysis of rhetorical patterns. However, there are also areas for improvement, including the need for quantitative analysis, better organization and writing quality, and further exploration of certain aspects. For the above-mentioned reasons, I do not feel like recommending this Doctoral Thesis for distinction.

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA
Firmato digitalmente da Marco Viviani
Data: 14.07.2023 23:23:59 CEST
Organizzazione: BICOCCAMILANO/1
2621570154

Milan, July 14, 2023

Prof. **Marco Viviani**
University of Milano-Bicocca
Department of Informatics, Systems, and Communication