

ANNOTATION PRACTICES IN MACHINE LEARNING RESEARCH ON DEPRESSION

Aleksandra Andrasz (a.andrasz@student.tudelft.nl)
Supervisors: Cynthia C. S. Liem, Andrew M. Demetriou

01 Background

An estimated 4,4% of people suffered from depression worldwide [1]. And the diagnosis and treatment remain complicated tasks.

Researchers, using machine learning, try to predict depression, optimal treatment and connected symptoms [2]

The models used in the area are mainly supervised learning models that rely on well-established training/verification data.

Annotation practices are of questionable quality [3]

02 Research Question

RQ: What are the data collection and reporting practices of annotations within Machine Learning Research surrounding depression prediction?

03 Methodology

1. Data Collection

Research papers were collected from Scopus from the 2nd of May until the 23rd of May.

Search keywords were divided into parts: depression, machine learning and others limiting the scope. Peer-reviewed English papers from 2013-2023 were included.

2. Data Synthesis

Papers were categorised into Computer Science, Medicine and Psychiatry, and Other domains. Each paper was reviewed on annotation practices and data availability. The results were collected and analysed per domain.

04 Results

Differences between domains

Domain	Papers	Human annotated
Computer Science	42	64.3%
Medicine and Psychiatry	33	80.6%
Other	22	54.5%

Table 1. Number of papers from different research domains, and percentage of human-annotated.

Reported Formal Instruction

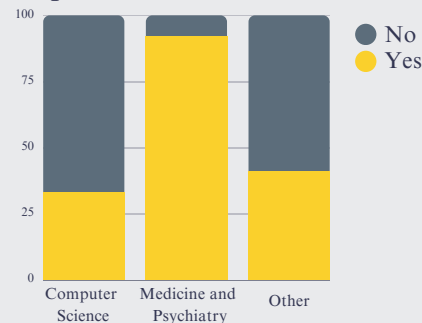


Figure 1. Percentage of papers that reported formal instructions.

Reported Training

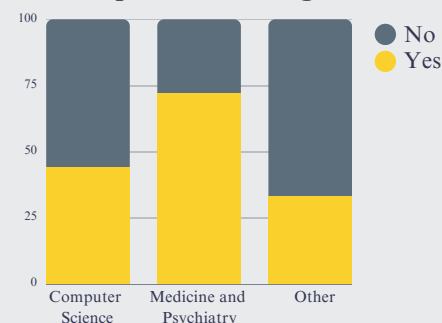
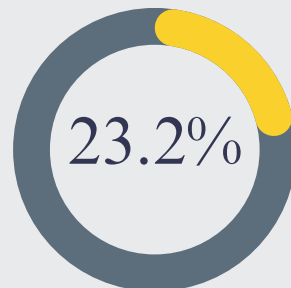


Figure 2. Percentage of papers that reported training.

Ground truth data

Availability of data



Ground truth data description

Label	Count
no	52
age	2
gender	4
age, gender	13
age, gender and other	22
gender, region	1
ethnicity, education, income	1

Table 2. Number of papers that reported demographic information.

05 Takeaways

- Reporting practices varied between domains. Medicine and Psychiatry papers exhibited distinct trends compared to others.
- Given the complexity of depression and the high risks associated with mistakes, researchers from the Computer Science domain could employ more rigorous reporting methods.
- Researchers could perform training and give formal instructions to the annotators.
- Issues across domains: data availability and information about people who were part of the dataset.
- Making data available might be a difficult task in depression research due to privacy concerns and ethical considerations.

06 Limitations

- Limited time: a period of 10 weeks, which put constraints on the analysis process.
- One person examining the data, with limited knowledge of depression and data collection.
- Biases: confirmation bias, selection bias.

07 Future work

- Searching could be done through multiple databases.
- It would be interesting to examine trends across the years.
- Investigating data quality in more detail could produce intriguing results.

References

- [1] World Health Organization. Depression and other common mental disorders: global health estimates. number-of-pages: 24.
- [1] Adrian B.R. Shatte, Delyse M. Hutchinson, and Samantha J. Teague. Machine learning in mental health: A scoping review of methods and applications, 7 2019
- [2] R. Stuart Geiger, Kevin Yu, Yanlai Yang, Mindy Dai, Jie Qiu, Rebekah Tang, and Jenny Huang. Garbage in, garbage out? do machine learning application papers in social computing report where human-labeled training data comes from? pages 325–336. Association for Computing Machinery, Inc, 1 2020.

How to make a research poster: A guide for students

Many technologies and breakthroughs would not be possible without research. It is important to keep members of the community informed about the latest updates. One way to do that is through research posters.

Introduction

Posters are popular method of presenting research findings in a concise and visually pleasing manner. They are commonly used in conferences and meetings. Start by introducing the subject of your research and/or your hypothesis. What are the questions about this topic that you want to answer? What new things can it contribute to the existing literature?

Objective

It is important for your readers to know what you want to achieve with your research. State this as clear as possible.

Methodology

Let people know how you did your study. Methods can vary depending on the subject or results you want to see. These methods can include:

- Interviews
- Surveys
- Comparison studies
- Experiments

You can also show studies of existing literature that were used as references.

Results/Findings

Results show the outcome of the research and should answer the question or hypothesis stated in the introduction.

- State what you've found from your study.
- You can also list your findings in bullets.

IMPORTANT!

Avoid using too much technical detail or using excessive jargon when presenting them.

Analysis

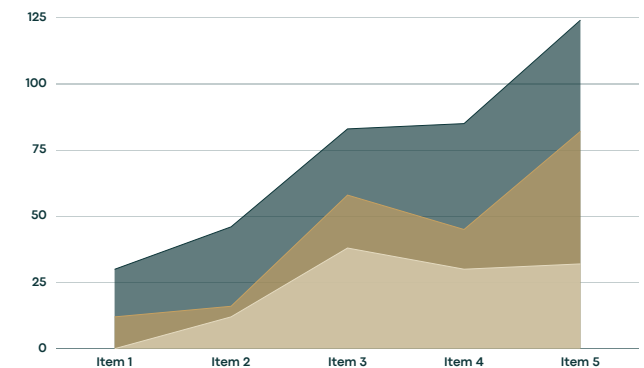
Expand on your findings by discussing what methods were used to analyze your data. It can get technical so keep it simple and direct to the point. Use bullets for emphasis. Include key graphs, tables, illustrations, and other images that support the study and show a visual analysis of the data. Make sure they are large enough to be seen from a distance but not clutter the poster.



Use illustrations to showcase your data in a visual form.



Write a caption that will clearly explain what this graphic is about and how it relates to the study.



Graphs are great in helping make numbers easier to understand.

Conclusion

Summarize your study and let the viewers know two to three key findings. You can also add a description of each that can give them an idea of what comes next. This section can also include any implications of the study, and if there are any actions or recommendations for future study.

Related literature

Research is often built on something that is already out there. Cite key references that you looked at while conducting your study.

Authors

Be proud of your work! Add the names of the people involved in this study. Don't forget to include titles and honorifics. We're proud of those too.

Affiliations

We're also proud of the institutions that we are with and support our research. Let's let them know by adding their names and logos here.