Presenter Notes: On the Dangers of AI Autism Diagnosis: ICDS Scholars Week 2024

Noah Duggan Erickson*

17 May 2024

Thanks

Many thanks to the multitude of faculty who have lent a hand in this project.

- Dr. Kristen Chmielewski (CHSS/RML) for helping to get the project off the ground
- Dr. Brian Hutchinson (CS) for his assistance with ML techniques
- Heather Cochran M.Ed. (Woodring) for her guidance in the project

1 Overview

- Getting started
- Technical, free to ask questions
- Updates since presentation given in HNRS, added visuals & ML progress

2 What is AIAD?

- Rigor involved in getting an ASD diagnosis
- A rapidly growing field in machine learning research
- Nexis search for `AI (AND) Autism (AND) diagno*` gave 96 results: **20** of them are from **2021**, **10** from **2022**.
- Trying... *trying...* to employ various machine learning techniques in the detection of ASD

*Department of Computer Science, College of Science and Technology, Western Washington University

3 Problems in General ML Research

- Taking a step back from AIAD for a second.
- What are the problem areas in general ML research?
- Models are growing faster than technology
- Increasing model size means increasing scale of data needed
- In the current hot topic of LLMs such as ChatGPT, where does the data come from? How is it filtered? What bias is intrinsic to the dataset?
- No such thing as a perfect filter. Demographics of who edits Wikipedia?
- In a **2020 UW** study, GPT-2 produced toxic responses to *at best* 17% of innocuous prompts, even after the best "conditioning" the study performed.

4 Disability Invisibility

- Won't spend too much time here...
- Treatment of disabled as sub-human, lesser
- In much of ML research, gender and race are frequently "protected characteristics". Disability often unmarked; *maybe* a footnote.
- Conventional AI bias metrics actively blind to ableism; "fairness", tying into WW-II masculinity(?)

5 Defining Autism

- How is autism defined in society?
- People who I shall not say their names aloud
- Indirect inhumanity
- AI ethics predicated on personhood; "viable knowers"

6 Current AIAD Literature

- What are researchers currently doing?
- A few main tracks, looking at two of them.

6.1 Description

- MRI
  - Most studies use the ABIDE dataset, an aggregation of 3D and 4D MRI brain scans from various institutions.
  - 78% accuracy from taking the average of a 2021 review
  - Variety of techniques, including CNNs and SVMs
- Assessments
  - Streamlining traditional assessments
  - Perpetuating the same biases as those assessments
  - Optimizing feature selection

6.2 CNN

- Tensorflow/Keras implementation of the CNN from that paper
- Paper claims X% accuracy
- Violin plot of predictions: Y-axis scale changes, but distributional equivalence remains.
- Accuracy plot: Dev set volatile
- Likely some sort of `nilearn` preprocessing I'm missing

6.3 Random Forest

- `sklearn` implementation of the pipeline from that paper
- Pipeline consists of KNN imputation of missing values, minmax scaling, and a random forest classifier
- Serious issue with original paper: Included sum of A-scores feature, defeating the purpose.
- Heavy selection bias, since existing assessments are used as the gold standard.
- Due to internals of Gini importance, age importance is inflated.
- Except yellow, colors represent one-hot groups
- Except age, all features are binaries

7 Problems in AIAD

- Conclusions, wrap up