Algorithmic decision-making in healthcare: Exploring the ethical considerations of using algorithms for healthcare decisions, such as patient diagnosis, treatment recommendations, or resource allocation.

Sunil Chomal, Jyoti Vasudev, Brendon Gory

CSCI E-184 Data Science and Artificial Intelligence Ethics, Governance, and Laws, Summer 2023

Harvard University Extension School

Prof. Bruce Huang

Introduction

- Ethical considerations using AI in healthcare
  - Everyone must be treated equal and fairly
  - Data and algorithms must be unbiased
  - AI must be explainable
- Treatment recommendations
  - Not every person currently receives same care (racial bias)
  - Reduce barriers imposed by socioeconomic separation
- Challenges and opportunities using AI in healthcare
  - Benefits
  - Drawbacks
  - Public perception
  - Recommendations

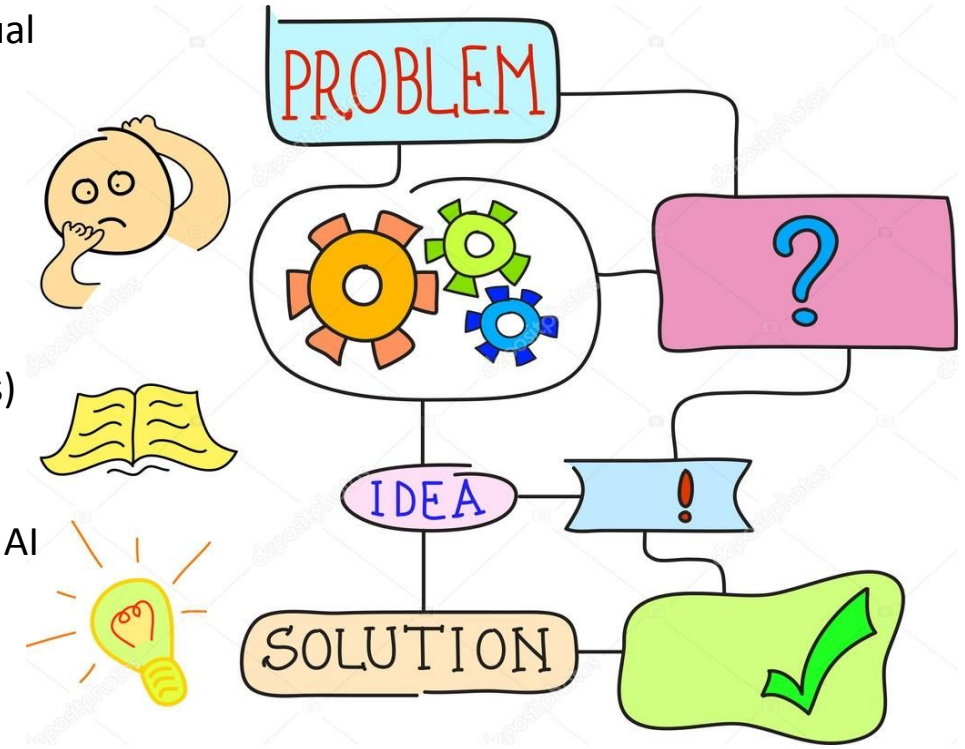


Image credit: https://st.depositphotos.com

Benefit: Promoting equity in healthcare

Addition of race in prediction for chronic kidney disease (CKG) unfairly more accurate for white patients

Lack of health insurance applied to machine learning models as a variable showed bias to underprivileged

Chance to reboot healthcare under the direction of “diverse group of stakeholders” and proper governance

Recognition of feedback loops where previous steps inform and bias the ensuing steps

Goal of P(Y|X): Same response “Y” regardless of variable(s) “X”

Everyone gets same treatment - no unintentional, institutionalized bias towards individuals or groups

Benefit: AI tools can lessen racial bias in pain treatment

Black and Hispanic people receive worse care on 40% of the department's care quality measures, especially in pain treatment

Black patients are under-treated for pain due to stereotypes, misinformation, and lack of empathy from health care providers

The historical roots of this issue may be in the legacy of slavery and medical experimentation

Pierson et al developed a deep learning model that can predict pain scores from knee radiographs with high accuracy

The model discovered image features that were associated with pain but were not captured by the conventional scoring system

The model reduced the racial gap in pain prediction by 47%, suggesting that it could improve the quality and equity of pain care

Benefit: Telemedicines / Telehealth using AI

High cost of healthcare, which increases year over year

Disparity of treatment between rural patients and rural patients

Growing crisis of burned out healthcare providers and difficulty replacing those roles (recent COVID-19 pandemic)

AI tools that can triage, gather data, and handle high-level questions allow healthcare workers to focus on their expertise

Automation of clinician workflows with technologies like NLP

Constant care with follow-up automations, reminders, and tireless data collection

Drawback: Difficulty with Public Trust in AI

People do not feel comfortable with results from ai because they feel the algorithms/data cannot be held accountable

Banerjee, Alsop, Jones, and Cardinal showed in their research starting with a simple model and constant education, transparency, and communication with patients helped increase trust in the treatment

For public trust, the AI cannot be presented as a “black box” where people perceive the technology as mystical and not founded on key principles

The difficulty of comprehending the ai models (interpretability) should not prevent the models from being explained (transparency)

If the models become more complex and data-intensive, healthcare will lose the opportunity to explain the tools because of the perception “it’s too difficult to know”

Drawback: How AI Models Could Fail to Detect Skin Cancer in Diverse Populations

Advanced medical equipment in developed countries mainly collects data from caucasian population, leading to under-representation of other ancestries

A study by Haenssle et al. showed that a machine-learning algorithm for detecting skin cancer performed better than dermatologists, but the data had limited diversity of skin types and ethnicities

A study by Adamson et al. revealed that most publicly available datasets for skin cancer diagnosis lacked information on skin type, race, or ethnicity, and had inconsistent and incomplete labels and annotations

Sufficient and quality data collection is essential for a reliable ai solution, otherwise AI models could be biased, inaccurate, or unsafe for diverse populations

AI models could potentially worsen the existing disparities instead of bridging the gap, if they are not trained on sufficient and representative examples

Drawback: When "Explainable" Is Not Enough

Introducing advanced technologies into clinical practice involves many different challenges

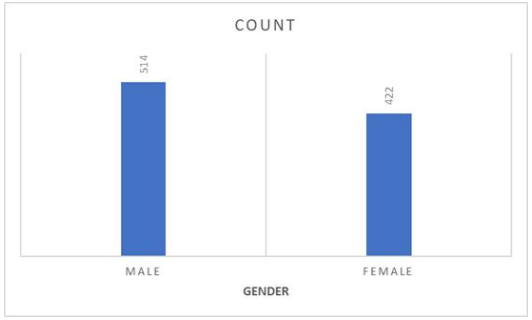
A study by Alison et al. showed that openness, concern and perceived benefit varied with gender, full-time employment status, healthcare choice and healthcare satisfaction

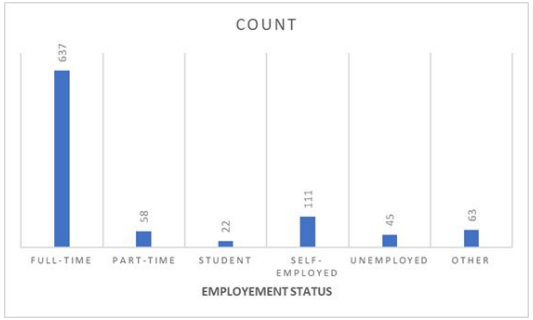
A UK study in 2021 concluded that people were open to use of app-based self monitoring, but overall there was a split in perceived benefits of such an approach

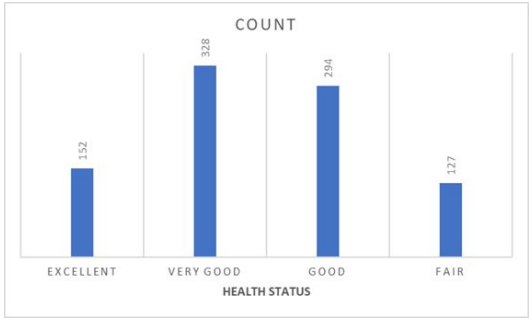
Healthcare is more than just understanding symptoms and explaining treatments, it involves emotion and empathy, which today, chatbots are incapable of and often are many layers below the surface

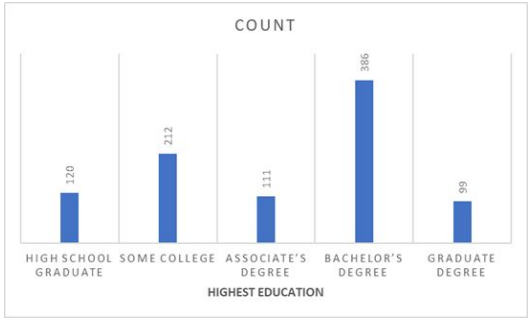
Use of AI technology in healthcare for patient care poses the risk of groups being alienated, due in part where technology is today, and the disparity in various socioeconomic parameters of the target audience

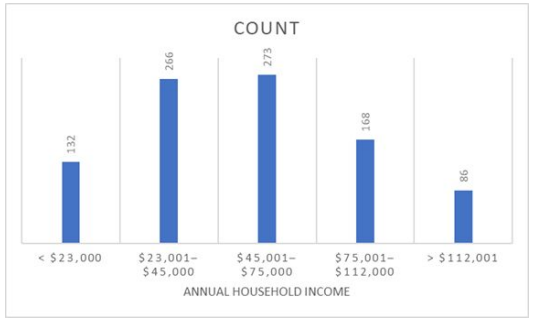
Drawback: When "Explainable" Is Not Enough

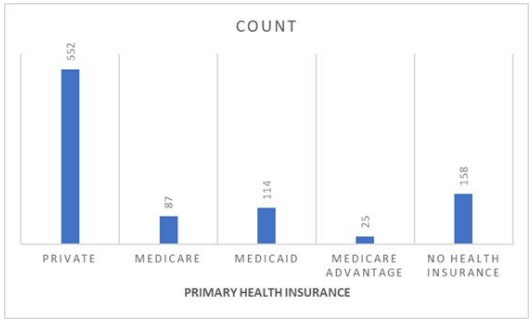


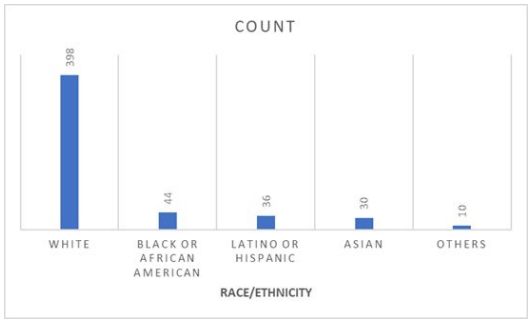


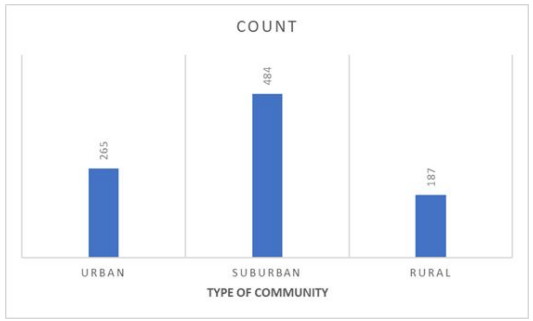












Bias: The data clearly has a bias towards white, male, educated participants who are in good health and have private insurance. This clearly biases the results to a specific demographic. However, to the credit of the authors, they have gone a step ahead and did a correlation analysis of openness, concern and benefit response to demographics

Drawback: When "Explainable" Is Not Enough

	Openness		Concern		Benefit	
	r	95% CI	r	95% CI	r	95% CI
Socio-demographics						
Age	−.12	[−.18, −.06]	.06	[.00, .12]	−.03	[−.09, .03]
Sex (1 = Male, 0 = Female)	.10	[.04, .16]	−.20	[−.26, −.14]	−.03	[−.09, .04]
Race (1 = White, 0 = Non-White)[a]	−.05	[−.11, .01]	.01	[−.05, .07]	−.08	[−.14, −.02]
Ethnicity (1 = Latino, 0 = non-Latino)	.06	[.00, .12]	−.09	[−.15, −.03]	−.02	[−.08, .04]
Household income	.07	[.01, .13]	−.08	[−.14, −.02]	.07	[.01, .13]
Community type	.06	[.00, .12]	−.06	[−.12, .00]	.01	[−.05, .07]
Employment status[b]	.17	[.11, .23]	−.18	[−.24, −.12]	.05	[−.01, .11]
Education	.04	[−.02, .10]	.03	[−.03, .09]	.01	[−.05, .07]
Health status and access						
Health status	.08	[.02, .14]	−.12	[−.18, −.06]	−.02	[−.08, .04]
Healthcare location[c]	.03	[−.03, .09]	−.01	[−.07, .05]	.02	[−.04, .08]
Healthcare choice[d]	.08	[.02, .14]	−.06	[−.12, .00]	.11	[.05, 17]
Health insurance (1 = Yes, 0 = No)	.09	[.03, .15]	−.10	[−.16, −.04]	.05	[−.01, .11]
Healthcare satisfaction (n = 735)	.11	[.04, .18]	−.07	[−.14, .00]	.14	[.07, .21]
Psychosocial variables						
Health System Trust Index	.27	[.21, .33]	−.27	[−.33, −.21]	.21	[.15, .27]
Trust in technology	.41	[.36, 46]	−.21	[−.27, −.15]	.41	[.36, .46]
Faith in technology	.38	[.32, .43]	−.10	[−.16, −.04]	.46	[.41, .51]
Conscientiousness	.02	[−.04, .08]	.11	[.05, .17]	.15	[.09, .21]
Agreeableness	.08	[.02, .14]	.11	[.05, .17]	.20	[.14, .26]
Extraversion	.08	[.02, .14]	−.12	[−.18, −.06]	.04	[−.02, .10]
Emotional stability	.08	[.02, .14]	−.06	[−.12, .00]	.07	[.01, .13]
Openness (trait-based)	.07	[.01, .13]	.07	[.01, .13]	.05	[−.01, .11]
Social conservatism	−.01	[−.07, .05]	−.10	[−.16, −.04]	.05	[−.01, .11]
Economic conservatism	−.06	[−.12, .00]	−.06	[−.12, .00]	.02	[−.04, .08]

N = 936 (except as noted for specific variables in Tables 1 and 2)

[a] Participants who selected any race other than White, or in addition to White, were classified as Non-White for purposes of this analysis

[b] 1 = full-time employment, 0 = all other options

[c] 1 = doctor office or private clinic, 0 = all other options

[d] 1 = great or some choice; 0 = little to no choice

This exploratory analysis of variables associated with openness, concern, and perceived benefit indicated that

- Socio-demographic and health variables were largely unrelated.
- There were modest relationships of age and sex to openness: older participants were less open, and
- Males were more open than females. Females also responded more negatively when presented with concerns.
- Full-time employment status was associated with greater openness and lower concern.
- People with greater healthcare choice and healthcare satisfaction perceived more significant benefits, and lower health status was associated with greater concern.

@Chomal, Vasudev, Gory

source: Alison et. al.

Public Perceptions - Dataset

subreddits:

subreddit	users
r/artificial	397k
r/artificialinteligence	233k
r/machinelearning	2.7m
r/chatGPT	2.6m
r/openAI	476k
r/GPT3	486k
r/datascience	967k

In these subreddits we query for submissions with the following keywords in the title of the submissions:

medicine, healthcare,

health, diagnosis,

patient, diagnosis,

doctor, pharma

The final dataset is a set of 981 submissions, across 7 subreddits and a total of 8964 comments, with a word count of 438,920

We have not collected any personal unique identifiers for redditors while collecting this data.

1. We chose a discussion from a community of 7.8 million members. We also chose keywords which are trending in various discussions (about 900+ submissions and nearly 9000 comments).
2. PRAW API allows us to get the reddit username of people posting the comments. We have ensured that we do not have these details in the data we have collected using this API.
3. To be in line with the terms & conditions of reddit developer apps API, we are not uploading the downloaded comments as part of this submission.

Public Perceptions - Word cloud & Sentiment Analysis

Word cloud does not throw up surprises, as ***AI***, ***data***, ***people***, ***healthcare***, and ***doctor*** are the most actively used terms.

What is interesting to note is that ***will***, and ***think*** are also trending showing that people are looking at AI in healthcare beyond just a passive tool, and the community at large is aware of its implications.

Sentiment Polarity	0.12	**range**: -1 to +1
Sentiment Subjectivity	0.48	**range**: 0 to 1

The general sentiment is better than neutral / slightly positive, but people seem to be biased / subjective in their opinions on this topic.

Public Perceptions - Findings

1. The comments are from subreddits related to AI, Machine Learning, GPT and Data Science, with query terms related to healthcare and medicine, which are dedicated to discussing and sharing news, articles, and opinions on advancements in AI in healthcare. Comments have different formats and purposes, ranging from asking questions, making statements, to sharing links to articles or videos.

2. AI fueled personalized medicine is a popular topic among the participants. There is also a debate and discussion on the role and value of AI versus human intelligence in medicine.

3. Ethical and social implications of AI for health are a prominent topic of discussion, as well as concerns about the trustworthiness, fairness, accountability, and transparency of AI in healthcare.

4. There is a high level of awareness and interest in the current state and future direction of AI in healthcare. Many comments are recommendations for courses or resources to learn AI skills specific to healthcare indicating a high demand and interest for education and training in this field.

5. Most comments are positive and optimistic, highlighting the potential benefits and opportunities of AI for improving diagnosis, treatment, research, and innovation in medicine.

6. Some comments are neutral and curious, asking questions or seeking information or advice on AI and healthcare. A comment from a user who claims to have built the closest to an A.I. Bayesian Brain with Human-like logic in healthcare was met with awe and skepticism.

7. Specific applications or examples of AI in healthcare are discussed, indicating the diversity and innovation in this field.

Recommendations

- Lessened bias in datasets
  - Ensure data are fair and unbiased
  - Data collection must be representative
- Lessened bias in algorithms
  - Considerations made for group and individual fairness
  - Algorithms must be explainable
- True governance over AI in healthcare
  - Private or public organizations designed to oversee best practices
  - Global, ethical consistency on AI use in healthcare
- Complete transparency
  - Make explainable “enough”
  - Public access to design, datasets and purpose of AI
- Standards that regulate AI
  - New regulations for AI like GDRP is for data
  - AI in healthcare held to higher accountability

Image credit: https://canstockphoto.com