Assessing the Algorithmic Fairness of Large Language Models across Health Care Use Cases

Sumanth Ratna
Yale University
sumanth.ratna@yale.edu

Dr. Jaan Lı
One Fact Foundation
Tartu University
jaan@onefact.org

PROJECT ABSTRACT. abstract here

BACKGROUND

as (Fiorella 2022).

color (Twomey et al. 2021),

(Heer and Moritz 2023) will be used to visualize all

For this, we have built the phenotype workflow (Zelko et al. 2023)

METHODS

We used the (Huang, Altosaar, and Ranganath 2020) model to analyze the data.

For the BM-25 algorithm, we use the following equation:

$$(D, Q) = \sum_{i=1}^{n} (k_1 + 1) \frac{f(q_i, D)}{k_1 \left(1 - b + b \ |D\frac{|}{a}\right) + f(q_i, D)}$$

where $f(q_i, D)$ is the frequency of term q_i in document D, $|D|$ is the length of document D, and a is the average document length in the collection.

RESULTS

t	1	2	3
y	0.3s	0.4s	0.8s

FIGURE 1. Table 1: test

FIGURE 2. Figure 1: The phenotype workflow can help practice data thinking to best validate definitions of health and disease, intellectual property, the market dynamics, etc.

REFERENCES

Fiorella, G. (2022, November 23). *How to Maintain Mental Hygiene as an Open Source Researcher.* https://www.bellingcat.com/resources/2022/11/23/how-to-maintain-mental-hygiene-as-an-open-source-researcher/

Heer, J., & Moritz, D. (2023). Mosaic: An Architecture for Scalable & Interoperable Data Views. *Private Communication and Early Access Provided to One Fact Foundation - Do Not Re-Cite.*

Huang, K., Altosaar, J., & Ranganath, R. (2020, November 28). *ClinicalBERT: Modeling Clinical Notes and Predicting Hospital Readmission.* https://doi.org/10.48550/arXiv.1904.05342

Twomey, C. R., Roberts, G., Brainard, D. H., & Plotkin, J. B. (2021). What We Talk about When We Talk about Colors. *Proceedings of the National Academy of Sciences*, *118*(39), e2109237118. https://doi.org/10.1073/pnas.2109237118

Zelko, J. S., Gasman, S., Freeman, S. R., Lee, D. Y., Altosaar, J., Shoaibi, A., & Rao, G. (2023, March 30). *Developing a Robust Computable Phenotype Definition Workflow to Describe Health and Disease in Observational Health Research.* https://doi.org/10.48550/arXiv.2304.06504