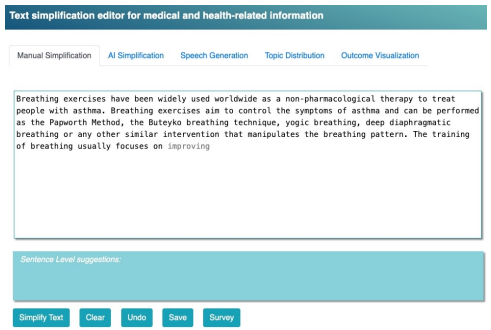
Autocomplete Using LLMs for Simplifying Health-related Texts

Nick Morgenstein
Computer Science Dept.
Pomona College
nmaa2019@mymail.pomona.edu

David Kauchak, PhD
Computer Science Dept.
Pomona College
david.kauchak@pomona.edu

Gondy Leroy, PhD.
Eller College of Management
University of Arizona
gondyleroy@arizona.edu



Autocomplete is a useful feature for simplification tools to help writers create simpler medical texts.

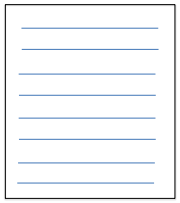
Challenge: generic LLMs have two key problems:

1) They are general-purpose and **not** tailored towards *simplicity* or the *medical domain*.

2) They always suggest autocompletions even if the quality of the suggestion isn't good

Acknowledgements
Research reported was supported by the National Institutes of Health (NIH) #R01LM011975. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Autocomplete task



Plain Language Adaptation of Biomedical Abstracts (PLABA)

"Diabetes mellitus is a common disease whose complications are severe."

previous words	word to predict
Diabetes	mellitus
Diabetes mellitus	is
Diabetes mellitus is	a
Diabetes mellitus is a	common

Evaluation

Randomly picked 200 sentences from PLABA: 3,755 test cases

Metrics:

- *Accuracy* – prediction match word
- *Word frequency* – measure simplicity
- *Word2vec similarity* – allows for near matches

Approaches

Baseline: GPT 3.5 Davinci

Fine-tuned: fine-tuned GPT 3.5 on 167K aligned sentence pairs from English Wikipedia and Simple English Wikipedia

Fine-tuned + repetitions: repeatedly queried the model, selected the most frequent suggestion

Results

	GPT 3.5	Fine-tuned
Accuracy	20.1%	21.1%
Word frequency	271B	3.23B
Word2vec	0.188	0.195

Fine-tuning makes the suggestions more accurate and simpler

# of repetitions	Accuracy	% predicted
1	16.3%	40%
2	18.2%	28%
3	33.7%	15%
4	51.5%	8%
5	70.5%	9%

Repetition is a strong predictor of quality and can be used to improve accuracy and to filter suggestions