ACL Paper Summary

Richard Padilla | Dr. Mazidi | CS 6301.M02 Spring ‘23

Title:

GPT-D: Inducing Dementia-related Linguistic Anomalies by Deliberate Degradation of Artificial Neural Language Models

Author list:

Changye Li , David Knopman , Weizhe Xu , Trevor Cohen , and Serguei Pakhomov

Affiliations:

Institute of Health Informatics, University of Minnesota 2
Mayo Clinic, Rochester, MN
Biomedical and Health Informatics, University of Washington
Pharmaceutical Care and Health Systems, University of Minnesota

Summarize problem addressed by the paper :

The authors in the paper are trying to generalize speech differences between a cognitively healthy individual and an individual with Alzheimer’s or dementia. They aim to generalize the differences by comparing the outputs of GPT-2 with a purposefully degraded version they created called GPT-D. Deep Learning techniques are really good at distinguishing between cognitively healthy individuals and individuals with Alzheimer’s on a limited set of data. However, there is no generalized method or research done to show that these techniques can differentiate healthy individuals from cognitively impaired individuals in generalized speech. Therefore, the authors set out to create such a generalization. The references were pretty exhaustive at over 40 references for the paper. A lot of the references were from the last five years, but went back as far as 1985. The references were definitely relevant to the paper. The references dealt with Alzheimer's, dementia, natural language processing, and language models.

Summarize prior work

Machine learning methods in the past have been particularly good at differentiating the speech patterns of healthy individuals from cognitively impaired individuals. This was done by training the BERT model on the Alzheimer’s/Dementia recognition through spontaneous speech challenge in 2020. This work was done in response to the lack of generalization of the DementiaBank, a corpus of responses based on the cookie theft photograph. The claims presented by the authors correspond pretty strongly to the data in the tables and figures. The perplexity between the models strongly differentiates cognitively impaired individuals from healthy ones.

ACL Paper Summary

Richard Padilla | Dr. Mazidi | CS 6301.M02 Spring ‘23

Unique contributions of this paper

The work in this paper extends the State of the art to a minor degree, at least technically. This is because the state of the art detection between healthy and cognitively impaired individuals was at 83.3%, and the perplexity between GPT-2 and GPT-D was 85%. This is a pretty small percent increase. So in terms of measurable output, the improvement found in this paper was not very large. The part of the paper that does extend the state of the art more than a minor degree is the ability of this method to generalize the speech differences between cognitively healthy individuals and individuals with Alzheimer’s/dementia. This has not been done before, and is pretty big since this method can be used outside of task-specific data; it can be used to evaluate everyday conversation

This work is significant to other researchers. Other researchers can take this knowledge and build upon it. Especially with the recent models of GPT 3.5 and GPT 4. It would be very interesting to see how the perplexities differ between these improved models, and how much of a boost they provide to the accuracy on ADDReSS. Other researchers will definitely be able to reproduce these results and build on them. The authors did a great job of explaining what they did and how they did it, and any other researchers will be able to reproduce their work for future analysis and experimentation. To reiterate, other researchers will be able to build upon this work by comparing the perplexities between more powerful language models like GPT 3.5 and GPT4, and artificially degraded versions of those models (e.g. GPT 3.5-D, GPT4-D).

How the authors evaluated their work

The authors tested the models on the ADReSS corpus. They performed cross validation, and looked at the performance on other corpuses as well. The authors compared their work to BERT, and BERT was indistinguishable from random guesses in determining healthy patients from cognitively impaired patients.

Number of citations the authors have received on Google Scholar:

Changye Li: 20, Trevor Cohen: 3945, Serguei Pakhomov: 6953 (most citations)
David Knopman: 0 , Weizhe Xu: 0

Conclusion of why you think their work was important.

The work in this paper was important because it helped create a method for determining Alzheimer’s/dementia speech patterns from cognitively normal speech patterns in a generalized way. The work expands the ability for machine learning methods to detect Alzheimer’s/dementia speech outside of narrowly defined tasks. The work also is able to replicate the speech patterns of AD patients in an emergent way, by artificially degrading the language model. They did not set out to implant rules or specifically

program Alzheimer's responses in the model, however they were able to generate a model that replicated the speech patterns of AD patients with general degradation.