Automatic Generation of Cancer Research Abstracts, AI vs Journal Articles

Dara Okeremi[1], Hanlin Zhu[2], Lan Luan[2]
[1]Rice University, Houston, TX, [2]The University of Texas at Austin, Austin, TX

Introduction: There is an overwhelming amount of journal articles for modern researchers to parse through. For instance, there have already been 168,168 cancer-related papers archived on PubMed this year. In order to keep up with this substantial amount of literature, there are emerging interests in applying artificial intelligence (AI) to facilitate paper reading and drafting of new scientific ideas. Here, we extend the application of the state-of-the-art automatic research assistants to the cancer field. Using training datasets composed of over 5,000 cancer-related journal papers abstracts, we evaluated AI-based background knowledge extraction and abstract writing. The best AI performance is rated to be on par with human writers through a survey to university cancer researchers. This automatic research assistant tool can potentially speed up scientific discovery and production by helping researchers to efficiently read existing papers, create new ideas and write up new discoveries.

Materials and Methods: Through Natural Language Processing, a branch of AI, machines were trained with biomedical datasets to read an input of any cancer-related paper. By extracting scientific words from the input and establishing relationships between them, machines created a language model that generated new passages by predicting the probability of the next word given new input. Two prominent language models that generate new passages were used: GPT-2[1] and PaperRobot[2]. After GPT-2 took an input, it modeled on the probability of the next word based on the previous words. In addition to the GPT-2 function that solely operated with a language model, PaperRobot used a reference distribution that took into account the input and a memory distribution that found the relationships between scientific terms in the knowledge graphs. Built on these developments, we fine-tuned the models of both GPT-2 and PaperRobot, trained these models on cancer-specific literature, and applied them to generate new passages. We evaluated their performance by surveys to faculty and students (approved by IRB of UT Austin) in cancer research on to what extent the automatically generated paragraphs (3 models, N=139) were similar to human-written paragraphs (N=26). All paragraphs were rated on a numeric scale of educational background (1–6) that reflected the scientific content and language.

Results and Discussion: We performed a Two-Way ANOVA on the questionnaire results and found differences in group means between the different models ($p<0.0001$), but no difference in group means between student versus faculty evaluators ($p=0.1092$). There was no interaction between models and evaluators ($p=0.7577$). We performed a Multiple Comparison test as shown in Figure 1 to determine which specific group differed from each other. All the relationships were statistically significant ($p<0.001$), except for the human and GPT-2 model ($p=0.4547$), as well as PaperRobot and CancerRobot, the fine-tuned version of PaperRobot ($p=0.9975$).

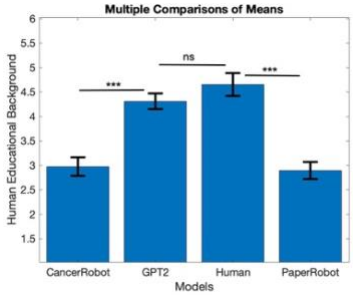

Figure 1. Bar Graph with error bars of Multiple Comparison Test. The range on the y-axis (1–6) corresponds to Middle School level to Faculty Member level. *** represents $p<0.001$ and ns represents no statistical significance.

Conclusions: We found that the quality of passages generated by GPT-2 and human-written journal articles are rated comparable. This result suggests that AI can be potentially used to automate draft generations in specific research fields such as cancer research that may speed up the scientific process of discovery and production.

Acknowledgements: This work was supported by the National Science Foundation through the Research Experience for Undergraduates program award #1461192.

References:

[1] Lee, J. S., & Hsiang, J. (2019). Patent Claim Generation by Fine-Tuning OpenAI GPT-2. *arXiv preprint arXiv:1907.02052*.
[2] Wang, Q., Huang, L., Jiang, Z., Knight, K., Ji, H., Bansal, M., & Luan, Y. (2019). PaperRobot: Incremental Draft Generation of Scientific Ideas. arXiv preprint arXiv:1905.07870.