Abstract

A summary of the contents of a book, article, or speech.

Reviews come in various forms—anonymous, open, and double blind, where reviewers are not revealed to the authors and authors are not revealed to reviewers. Whatever the process, act accordingly and with the highest moral principles. The cloak of anonymity is not intended to cover scientific misconduct. Do not take on the review if there is the slightest possibility of conflict of interest. Conflicts arise when, for example, the paper is poor and will likely be rejected, yet there might be good ideas that you could apply in your own research, or, someone is working dangerously close to your own next paper. Most review requests first provide the abstract and then the paper only after you accept the review assignment. In clear cases of conflict, do not request the paper. With conflict, there is often a gray area; if you are in any doubt whatsoever, consult with the Editors who have asked you to review.

Text mining with automatic extraction of key features is gaining increasing importance in science and particularly medicine due to the rapidly increasing number of publications.

Objectives: Here we evaluate the current potential of sentiment analysis and machine learning to extract the importance of the reported results and conclusions of randomized trials on stroke.

Methods: PubMed abstracts of 200 recent reports of randomized trials were reviewed and manually classified according to the estimated importance of the studies. Importance of the papers was classified as "game changer", "suggestive", "maybe" "negative result". Algorithmic sentiment analysis was subsequently used on both the "Results" and the "Conclusions" paragraphs, resulting in a numerical output for polarity and subjectivity. The result of the human assessment was then compared to polarity and subjectivity. In addition, a neural network using the Keras platform built on Tensorflow and Python was trained to map the "Results" and "Conclusions" to the dichotomized human assessment (1: "game changer" or "suggestive"; 0:"maybe" or "negative", or no results reported). 120 abstracts were used as the training set and 80 as the test set.

Results: 9 out of the 200 reports were classified manually as "game changer", 40 as "suggestive", 73 as "maybe" and 32 and "negative"; 46 abstracts did not contain any results. Polarity was generally higher for the "Conclusions" than for the "Results". Polarity was highest for the "Conclusions" classified as "suggestive". Subjectivity was also higher in the classes "suggestive" and "maybe" than in the classes "game changer" and "negative". The trained neural network provided a correct dichotomized output with an accuracy of 71% based on the "Results" and 73% based on "Conclusions" .

Conclusions: Current statistical approaches to text analysis can grasp the impact of scientific medical abstracts to a certain degree. Sentiment analysis showed that mediocre results are apparently written in more enthusiastic words than clearly positive or negative results [1].

ChatGPT-generated abstract

- Using ChatGPT to write a literature review on autologous fat grafting

A study aims to analyze the accuracy of human reviewers in identifying scientific abstracts generated by ChatGPT compared to the original abstracts. Participants completed an online survey presenting two research abstracts: one generated by ChatGPT and one original abstract. They had to identify which abstract was generated by AI and provide feedback on their preference and perceptions of AI technology in academic writing. This observational cross-sectional study involved surgical trainees and faculty at the University of British Columbia. The survey was distributed to all surgeons and trainees affiliated with the University of British Columbia, which includes general surgery, orthopedic surgery, thoracic surgery, plastic surgery, cardiovascular surgery, vascular surgery, neurosurgery, urology, otolaryngology, pediatric surgery, and obstetrics and gynecology. A total of 41 participants completed the survey. 41 participants responded, comprising 10 (23.3%) surgeons. Eighteen (40.0%) participants correctly identified the original abstract. Twenty-six (63.4%) participants preferred the ChatGPT abstract (p = 0.0001). On multivariate analysis, preferring the original abstract was associated with correct identification of the original abstract [OR 7.46, 95% CI (1.78, 31.4), p = 0.006]. Results suggest that human reviewers cannot accurately distinguish between human and AI-generated abstracts, and overall, there was a trend toward a preference for AI-generated abstracts. The findings contributed to understanding the implications of AI in manuscript production, including its benefits and ethical considerations [2]

[1]

Fischer I, Steiger HJ. Toward automatic evaluation of medical abstracts: The current value of sentiment analysis and machine learning for classification of the importance of PubMed abstracts of randomized trials for stroke. J Stroke Cerebrovasc Dis. 2020 Sep;29(9):105042. doi: 10.1016/j.jstrokecerebrovasdis.2020.105042. Epub 2020 Jun 23. PMID: 32807454.

[2]

Nabata KJ, AlShehri Y, Mashat A, Wiseman SM. Evaluating human ability to distinguish between ChatGPT-generated and original scientific abstracts. Updates Surg. 2025 Jan 24. doi: 10.1007/s13304-025-02106-3. Epub ahead of print. PMID: 39853655.