Research Proposal

Comparing the Use of Passive Voice in Academic vs. Non-Academic Writing on Data Science and Machine Learning

Zhanerke Myrzabekova

Writing Center Program, Nazarbayev University

WCS 230: Say What You Mean

Mr. Jeffrey Carney

4 December 2024

Comparing the Use of Passive Voice in Academic vs. Non-Academic Writing on Data Science and Machine Learning

This research project will examine how passive voice is implemented in academic and non-academic writing in the sphere of data science and machine learning. By analyzing both research papers and popular science articles, the study will investigate how authors present complex technical information. This project mainly focuses on the difference of use of passive voice in academic articles and popular science writing, and its effect on clarity and precision for a diverse audience.

The Scholarly Conversation

There is an ongoing debate around usage of passive voice in scientific writing, particularly regarding its impact on scientific practices and rhetorical goals. Banks (2017) investigated scientific journal articles from 1985 to 2015 and suggested that there has been a decline in using passive voice. That was true mostly for physical sciences rather than biological ones. He assumed that this shift might be influenced by the increase in mathematical modeling, which is one aspect of data science that tends to favor mental processes over material processes, thus prompting more usage of first person pronouns. As data science and machine learning are closer to physical sciences, I am going to concentrate on articles related to that sphere.

On the contrary, Ding (1998) argues that scientific writers prefer writing sentences in passive voice and avoid personal active subjects. Because passive sentences emphasize on what was completed rather than who did the action. Discrepancy between these articles suggest that further investigation is needed.

Furthermore, Ding (2001) identifies three main factors why authors tend to choose passive over active voice. Firstly, the agent of the action is not important or it is not necessary to know who did it. He shows the sentence “The tester is always connected to the machine” as an example. It is indeed true that most of the time when describing certain procedures, it is not obligated to mention the doer of action. Second reason is that passive voice helps to build more cohesive sentences by joining the first sentence to the second one. It is especially crucial, for instance, when describing how data science and machine learning work. Thirdly, engineering subjects mostly focus on objects.

Ding (2002) also points out that collaboration of scientists and falsifiability are the key factors which affect the usage of passive voice. By emphasizing materials, methods and results rather than the individual scientists performing the work, passive constructions contribute to a perception of objectivity.

These articles together illustrate the complex picture of passive voice in scientific writing. While one of them tends to analyze the decline in using passive voice, the others emphasize its importance in scientific papers.

Research Question

How does the use of passive voice differ between academic and non-academic writing on data science and machine learning? How it changed over recent decades, and what does this suggest about the norms in academic writing? Is passive voice more preferred in science writing compared to active?

Corpus

The corpus for this research project will primarily focus on data science and machine learning, consisting of both academic (scientific journals) and non-academic (popular websites,

technical reports) works. To increase comprehension of stylistic features and voice usage, studies from the last decades will also be included. However, since fields of computer science are evolving constantly, I will consider recent articles most of the time. This will include peer-reviewed articles and trustworthy non-academic findings differentiating by the usage of passive voice [560 words].

References

Banks, D. (2017). The extent to which the passive voice is used in the scientific journal article, 1985–2015. *Functional Linguistics*, *4*(1). https://doi.org/10.1186/s40554-017-0045-5

Ding, D. D. (1998). Rationality reborn: Historical roots of the passive voice in scientific discourse. In J. Battalio (Eds.), *Essays in the Study of Scientific Discourse: Methods, Practice, and Pedagogy* (pp. 117–135). Ablex Publishing Company.

Ding, D. D. (2001). Object-centered – How engineering writing embodies objects: A study of four engineering documents. *Technical Communication*, *48*(3), 297–308. http://www.jstor.org/stable/43090433

Ding, D. D. (2002). The passive voice and social values in science. *Journal of Technical Writing and Communication*, *32*(2), 137–154. https://doi.org/10.2190/efmr-bjf3-ce41-84kk