

Research Methodology

Artificial Intelligence Bootcamp 2023

Please download the papers below, and answer the questions on the following pages.

BioBERT: a pre-trained biomedical language representation model for biomedical text mining

Lee, Jinhyuk, et al. "BioBERT: a pre-trained biomedical language representation model for biomedical text mining." *Bioinformatics* 36.4 (2020): 1234-1240.

SciBERT: A pretrained language model for scientific text

Beltagy, Iz, Kyle Lo, and Arman Cohan. "SciBERT: A pretrained language model for scientific text." *arXiv preprint arXiv:1903.10676* (2019).

1. What are the strengths of both papers?

BioBERT

SciBert

2. What are the weaknesses of both papers?

BioBERT

SciBert

3. What are the most important aspects presented in both papers?

BioBERT

SciBert

4. Give a comparison of the description of the selected approach in each paper's methodology section. Which do you think is better? Why?

5. Have the authors of each paper been able to justify their approaches adequately?

BioBERT

SciBert

6. Compare the data used and the description of the data for both papers.

7. What issues have been highlighted by both authors?

BioBERT

SciBert

8. Compare the analysis of issues in both papers. Which paper's authors describe them in more detail?

9. Comment briefly on the clarity of presentation and organization of both papers.

BioBERT

SciBert

10. Comment briefly on the citing of sources determining the methodology of both papers.

BioBERT

SciBert

11. Was there any ambiguity in the explanation of the methodology of each paper? If so, what ambiguity was there?

BioBERT

SciBert

12. What could be added/removed/changed to improve the methodology for both papers?

BioBERT

SciBert