

Essay 3

For this paper, I'd like you to select a bioinformatics paper that involves clustering in a non-trivial way. After you read the paper, please write a 1-2 page analysis of the details of how it used clustering, as well as the implications of how they used clustering. (Note that you may want to do a bit more exploring and reading to understand the details of what they used, but the only formal reading assignment is the one paper.)

Some questions to get you thinking:

- Which clustering algorithm did the paper adopt, and how was it used in their research problem?
- If it was not one of the ones from class, please describe the method or algorithm in enough detail that I can understand what was used, as well as how. In the paper, do they describe it in enough detail for you to understand the assumptions and trade-offs, or do they present it as a black box? How did you learn enough to explain it in your summary, if they did not explain in sufficient detail?
- How would using a different clustering algorithm potentially impact their findings?
- Which implementation/code did they use, if that information is included, or did they code their own? If you can find more details on that implementation - can you find any reviews or critiques of that choice?
- How important was the clustering aspect of the code in the project?
- Knowing how many heuristics are used in clustering, do you find any issues in their results? Do you think an "optimal" clustering would have led to better or more accurate findings?

Again, these are just questions to get you thinking critically as you read, so you do not necessarily need to answer all of them. I'm trying to make this assignment flexible so you can learn about how clustering is used in your research area or in a novel way that is new for you, so please come discuss if you have any questions about the details.