

Research Assessment #8

Name: Aiswharyaa Lalgudi Nagarajan

Date: 15 November, 2024

Subject: ISM (Independent Study and Mentorship)

MLA citation:

Ouyang, Jia Fang, et al. "Deep Learning Models Will Shape the Future of Stem Cell Research." *Stem Cell Reports*, vol. 18, no. 1, 2023, pp. 6–12. doi:10.1016/j.stemcr.2022.11.007.

Assessment: Research Assessment #8

For my research assessment, I chose to do a research assessment on a paper titled "Deep Learning Models Will Shape the Future of Stem Cell Research" by researchers Ouyang and Fang. This research paper focused on deep learning, a subset of Machine learning, and how to use tools from here to create a digital model. The authors highlight the Deep Learning (DL) model's impact on stem cell research. They talk about datasets and tools that predict protein structures, construct cell differentiation trajectories, and identify the optimal conditions for the growth of the cells. The paper summarizes and highlights many different research papers that involve my topic. With each example, they introduced a new tool or database that could help my research, especially with the protein structures and the path of the cell's differentiation. Though they examined all these different papers, they communicated that deep learning will be crucial in stem cell biology in the future.

The research I read about in this paper applies to my study and my ISM Journey. My project aims to bridge the gap between artificial intelligence and stem cell biology. The essential tools and knowledge I gained from this paper would be beneficial in building my model. They talked about tools like AlphaFold and databases like Gene Expression Omnibus, which are precisely what I have been searching for in my journey so far. Also, their finding that bias will be a huge challenge when working

with deep learning was significant for me. It helped me be aware of the possible areas where errors could be in my research and how to fix them.

The paper discusses many topics but breaks the information into four parts: first, an introduction to DL, then the applications of DL; next, the challenges of DL; and finally, the future of DL in biology. They explain the shift that is currently happening from human analysis of data to the use of DL. They also talk about how researchers have used this so far in the example of AlphaFold and much more. They also discuss DL's limitations in the field, such as bias, interpretability, and data limitations. With my previous knowledge, I knew that DL was a subset of machine learning, but I didn't think it had so much application to stem cell biology. Moreover, I just learned that these previously made tools used DL to help with stem cell research. This new information can be added to the skills and tools I need to understand and research more. Also, I didn't know DL had limitations, including bias, which I had yet to learn applied to this field. Now that I know the limitations, I can consider them when creating my model.

This paper will be helpful when creating my computational model as it highlights many skills and tools I need to be familiar with. For example, the paper talked about DL models like Prescient, which are used for predicting cell fate. I am planning to use this as a foundation for building my model. The researchers also talked about a tool called TotalVi, which can combine datasets to create a single model of the cells. This is crucial for me as I will need to get datasets from multiple different places, so being able to put it all together will be valuable for me. The paper also talks about using a large and diverse group of datasets, so this is a principle I plan to use in my model as well. This portion of the paper helped me understand the need for high-quality training data so that the data processing is correct. All the tools and skills that relate to creating generative models for predicting cell behavior can be used in creating my model. All these tools will make it much easier for me to explore the different pathways the cells can travel.

This paper was exciting and motivating to read, as it completely related to my research topic and confirmed my previous ideas. Previously, I was focused on integrating artificial intelligence and biology, and this paper's conclusion confirmed the validity of my idea. Moreover, the overall content and organization of the paper are well-written, and I quickly understood it. The knowledge I gained from this helped me exponentially in helping me achieve my goals for ISM. The new tools, skills, and references I got from this paper can be directly applied in creating my project.