Short Story Proposal: Deep Learning in Proteomics

SJSU CMPE 258 Deep Learning - Professor Vijay Eranti

Haley Feng

Proposed Medium Title: A Gentle Introduction to Deep Learning in Proteomics

The use of deep learning in the study of proteins is no stranger to those in the field of proteomics. However, to those that don't have a background in biochemistry, it can be an overwhelming topic to grasp exactly the role machine learning plays in the drug discovery pipeline. The paper **"Deep Learning in Proteomics"** published from the scientific journal ***Proteomics*** in 2020 introduced their reader, the proteomics community, a comprehensive overview of what deep learning is and how it's applied in the analysis of proteomics data. Similarly, my proposal for the short story aims to introduce those in the field of data science to understand how we can utilize what we know in the field of proteomics. I will discuss some of the notable deep learning applications in the biotech industry and provide some basic background on what they are. These topics include:

- Retention Time Prediction
- MS/MS (Mass Spectrometry) Spectrum Prediction
- MHC (Major Histocompatibility Complex) Peptide Binding Prediction
- Protein Structure Prediction

I will also go over the framework and software in production and some of the direction and limitations researchers are currently facing with deep learning. By the end of my research, I hope to gain a better understanding of the topics listed and be able to give a high-level overview of how deep learning is used to transform protein data into novel biological insights.

Deep Learning in Proteomics

https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/epdf/10.1002/pmic.201900335

Wen, Bo, et al. "Deep Learning in Proteomics." *PROTEOMICS*, vol. 20, no. 21-22, 2020, p. 1900335., doi:10.1002/pmic.201900335.