BIOF 510 Final Project (Due Last Day of Class)

1. Problem Selection. First, choose a dataset that is either from your lab or relevant to your research. You should ensure that your data is reasonably suited for an ML project. For example, a dataset with only 10 samples is probably not a good choice.

2. Tell us about your data - we want to make sure you succeed, and the participation assignment this week will help us do that

3. Methods - Implement a method using algorithms we have covered in class. Some options:

- Deep Neural Networks
- CNN
- LSTM
- GANs
- Autoencoders/Unsupervised deep learning
- Graph-based deep learning (GNNs, graph attention networks, etc.)
- Transformers/other NLP techniques
- A method that involves self-supervised learning
- Reinforcement learning (covered in week 5 - an ambitious choice - ask me about it!)

4. Code: all of your code must be on a GitHub repository and clearly annotated. You must use object oriented programming (OOP). **If you are confused about OOP, please ask myself or Christina.**

You must provide access to all of your datasets unless you have discussed an alternative option with me. (40% of project grade)

5. Report: You must write a technical report describing your project. This report should be 3-5 pages, single spaced, not including references. (40% of project grade). This report should include:

- An introduction to your topic/problem
- A description of your datasets used
- A description of all methods
- A justification for all methods (why did you do the project in the way you did?)
- An overview of results
- A conclusion/future work section.

6. User guide/ReadMe: I must be able to easily use your method on the datasets you specify. Please provide a detailed user guide on how I can run your method. (5% of project grade)

7. Demonstration: You will present your project + results to myself and the TA via Zoom, and demonstrate how to use the method you designed. I recommend creating a brief slide presentation. (15% of project grade)