**Name** Matthew Feng

**Email** mattfeng@mit.edu

**Year** Sophomore (2021)

**Major** 6-2, 18

**1. Academic background**

**Biology:**

The extent of my biology experience includes 7.012, 9.00, and 9.01, the latter two of which only had incidental amounts of biology (action potentials, anatomy, etc.). As such, my knowledge of biology spans mostly the basics, without much depth in any particular aspect (such as cell biology or genetics). I’m very excited to learn the necessary biology that’s needed to complete the project we decide on, however!

**CS:**

I have much more experience in computer science than in biology – I took 6.006 last semester, and thus have a foundational knowledge of algorithms. I’m currently enrolled in 6.046 as well. I’m very interested in machine learning and deep learning, and so I have experience with TensorFlow, Keras, and other popular deep learning libraries. I’m most familiar with Python, but also have used R and MATLAB previously.

**Expertise:**

My expertise would lie in algorithms and machine learning theory, as well as experience manipulating data and creating visualizations for that data.

**2. My previous research experience**

I do not have previous research experience here at MIT, but I have worked at the National Institute of Standards and Technology previously, where I developed generative models for network traffic. Although unrelated to biology, through that project I learned much about data mining and processing, as well as using various Python libraries like scikit-learn, NumPy, SciPy, Matplotlib, etc. in order to understand the network traffic and build models similar to the data I was given.

**3. Areas of computational biology that interest me most**

I’m most interested in learning and exploring how to apply deep learning to genomics, which ranges from using recurrent neural network techniques to try to find the structure and meaning underlying genomic sequences, and building accurate predictive models such that we can predict with reasonable accuracy the effect of purposeful modification of certain symbols or bases along the sequence.

I’m also very interested in using computational techniques to understand cancer and cancer genomics, and to model the development of cancer for a full understanding of the process. I would also like to be able to take this fundamental understanding of the development of diseases and apply it to drug discovery so that we may find treatments as well.

**4. Preferred type of project**

I would like to work on a project that is more tool-building and analysis focused, one that explores data that we have and tries to derive new insights and understanding of that data, perhaps through new visual tools or techniques. Optimally, I would work on a project related to genomics, as that is the area I am currently most interested in, but I am definitely open to other ideas, as my scope of knowledge in biology is not wide enough for me to know all the interesting areas for research.