1. What is the problem you want to solve? Why is it an interesting problem?

Stock prediction. The environment is incredibly noisy and difficult to predict, making it more of a challenge than some other topics.

2. What data are you going to use to solve this problem? How will you acquire this data?

Stock price history and fundamental company data (earnings, revenue, etc.) Data will be gathered from online API’s like AlphaVantage and SEC-EDGAR (https://www.sec.gov/edgar/sec-api-documentation)

3. In brief, outline your approach to solving this problem. You might not know everything in advance, and this approach may change later. This might include information like:

a. Is this a supervised or unsupervised problem?

Supervised (data and outcome are basically labeled)

b. If supervised, is it a classification or regression problem?

Regression (price movement with magnitude)

c. What are you trying to predict?

Future price

d. What will you use as predictors?

Time history and fundamental data

e. Will you try a more “traditional” machine learning approach, a deep learning approach, or both?

Focus on deep learning and neural network-type approach

4. What will be your final deliverable? Will it be an application deployed as a web service with an API or a more robust web/mobile app.

Not sure

5. What computational resources would you need at a minimum to do this project? You may not have a very clear sense now but work with your mentor to come to an estimate. In real industrial applications, you’ll often be called upon to provide resource estimates at the beginning of a project.

Intend to use my laptop to get code working and to conduct initial training (AMD 5900X, 16GB RAM, NVIDIA 3070 GPU). May move to a cloud solution for training near the end of the project.