Summary (story: motivation->model->result->future directions):

The motivation of our projects is to build a model to predict the future return/trend of a basket of stocks so that financial company can use it to optimize the portfolios (e.g. balance between the return and risk) and even to develop the trading strategies.

We used the opening price data between 06/02/2012 and 06/02/2022 of the five stocks (AAPL, TSLA, AMD, SBUX, FB) from Yahoo to tune the LSTM model parameters and test on the most recent data. The five stocks we picked are from different sectors, which allow us to train a general model the fits a large variety of stocks. Testing data consists of the last 90 days while the price of the previous dates form the training set. The resulting plots shows that our predictions provide the correct trend with a lagging time in general. For relatively new stocks like TSLA, the volatility(beta) is relatively high. For stocks with relatively high (higher than the mean of the basket) historical beta, the prediction of the future 14 days (time-step) is better than the future 60 days. For stocks with relatively low (lower than the mean of the basket, such as APPL) historical beta, the prediction of the future 60 days (time-step) could be better.

For the future directions, we can first choose a single number metric such as f1-score, R-square, etc to iteratively tuning the model parameters for a lager basket of stocks (maybe NASDAQ 100). The metric can either measure the mean-square-error or even the lagging time since our model follows the real trend successfully. Then, we can use time series cross-validation to test some simple trading strategies such as finding the reverting moment, buy and sell daily at a fixed moment to see whether their return is significant in t-test comparing to the SPY return. In addition, the optimization part to build ETF, training a unique model for each sector may be better to optimize the return for a given input variance range (risk).