Portfolio Optimization

Daniel Hajialigol
Virginia Tech
6200 Heather Drive, Blacksburg, VA 24060
danielhajialigol@vt.edu

Abstract

It's no secret that many individuals have a common interest in the stock market, as many view it to be a potentially lucrative passive source of income. As a matter of fact, an average of 55% of Americans have invested in the stock market since 2009 (1). Although this is the case, making a good investment takes time, as keeping up with the latest news and trends can take a lot of bandwidth. As a result, there have been many models that have been used to solve this issue, whose primary objective is to predict the price of a stock. However, many of these models only take into account one stock, not a portfolio of stocks, which is a much more common scenario for an investor to be in. This problem, portfolio optimization, is the primary objective for this research project. Using the architecture in (2), this project's implementation saw returns up to 10% high than using a holding and selling strategy. However, the model has a lot of variability built in, making it an unstable source to base financial decisions off of in its current state.

1 Background Information

This section will outline the information needed in order to better understand the progress that has been for this project.

1.1 Data

The data that will be used for this project comes from `www.cryptodatadownload.com`, a website containing prices of cryptocurrencies. Each record corresponds to a particular time period, t. The data of interest for this project are (1) the high price, (2) the low price, and (3) the closing price of a given period. Each time period will correspond to a vector, v_t (known as the *price vector*), that contains the closing prices for all of the cryptocurrencies of interest at time step t. In this project, 12 of the cryptocurrencies that have the highest market value are inspected. The first cryptocurrency will be the currency the other crypto assets will be quoted in (which will be Bitcoin for this project), and this item has the property $v_{0,t} = 1 \ \forall t$ and will not be used as the input to the network policy. Amongst many factors, one factor that is important in optimizing a portfolio is to compare the closing prices of all of the assets of the current time period with all of the assets of the previous time period. This is captured by the following variable:

$$y_t = (1, \frac{v_{1,t}}{v_{1,t-1}}, \frac{v_{2,t}}{v_{2,t-1}}, \ldots, \frac{v_{m,t}}{v_{m,t-1}})$$

where $v_{i,t}$ corresponds to the closing price of the i^{th} asset at time step t. For the remainder of this paper, y_t will be considered the *price relative vector*. Armed with this information, it is now possible to get the portfolio value at a given time step t. This value,

$$p_t = p_{t-1} * y_t \times w_{t-1}$$

is a recursive function that takes in as input the following variables:

- p_{t-1}: The portfolio value at the previous time step
- y_t: The price relative vector
- w_{t-1}: The distribution of assets from the previous time step (will be referred to as the *portfolio vector*

This information allows us to define other important variables. One such variable is the logarithmic rate of return:

$$\rho_t = \frac{p_t}{p_{t-1}} - 1$$

Now we have the necessary information to compute the portfolio value after all time steps. This value is

$$p_f = p_0 * \prod_{t=1}^{t_f+1} y_t \times w_{t-1}$$

where

- p_0: Initial portfolio amount
- $t_f + 1$: Final time step

One could argue that this is a simple computation for a complex problem, but there is a reason why this could be the case: transaction costs are not considered in this formula. Whenever a trade is made, there is a cost associated with it that could be significant enough to cause a noticably decrease in the value of the portfolio. Thus, this project, inspired by (2), takes into account these transaction costs.

It is understood that the final portfolio value after accounting for transaction costs p_t will be a fraction of the portfolio value before the transaction costs $p_t^{'}$. This simple relationship is defined by $p_t = \mu_t p_t^{'}$. The logarithmic rate of return after factoring in transaction costs is now

$$\rho_t = \frac{p_t}{p_{t-1} - 1} = \frac{\mu_t p_t^{'}}{p_{t-1}} - 1 = \mu_t * y_t \times w_{t-1} - 1$$

The tensor $\mathbf{X}_t$, the input to the model (defined in the following section), is defined as follows:

$$\mathbf{X}_t = \left[\mathbf{V}_t^{lo} | \mathbf{V}_t^{hi} | \mathbf{V}_t\right]$$

where

$$\begin{aligned}
\mathbf{V}_t &= \left[v_{t-n+1} \oslash v_t \mid v_{t-n+2} \oslash v_t \mid \ldots \mid v_{t-1} \oslash v_t \mid \vec{1}\right] \\
\mathbf{V}_t^{hi} &= \left[v_{t-n+1}^{hi} \oslash v_t \mid v_{t-n+2}^{hi} \oslash v_t \mid \ldots \mid v_{t-1}^{hi} \oslash v_t \mid v_t^{hi} \oslash v_t\right] \\
\mathbf{V}_t^{lo} &= \left[v_{t-n+1}^{lo} \oslash v_t \mid v_{t-n+2}^{lo} \oslash v_t \mid \ldots \mid v_{t-1}^{lo} \oslash v_t \mid v_t^{lo} \oslash v_t\right]
\end{aligned}$$

where $\vec{1}$ is a vector of ones, and where $\oslash$ is the element-wise division operator.

Each $\mathbf{X}_t$ contains high, low, and closing prices across all cryptocurrencies from the previous n time periods, which is specified to be 50 in the paper. And since there are $m = 12 - 1$ assets/cryptocurrencies that will used, the dimensions of this tensor are defined as $\mathbf{X}_t \in \mathbb{R}^{3 \times 11 \times 50}$.

2 Architecture

This section discusses the architecture of the model implemented in this project. The first subsection will outline the environment, agent, and actions of this problem. The following subsection will then discuss the policy network (referred to as the model).

2.1 Agent, Environment, State, and Reward

Since this is a reinforcement problem, key variables, such as the agent, environment, state, and reward need to be clearly defined.

In this problem, the agent is the manager of the portfolio, who attempts to optimize the value of the portfolio. This agent acts in an environment, which is supposed to represent all of the market prices of the relevant cryptocurrencies. Of course keeping track of every cryptocurrency throughout all of its time is computationally expensive. To counter this issue, the agent only has stock information of the assets from the previous 50 time steps. This is exactly what X_t represents.

The goal of the agent is to maximize the portfolio value. In doing so, the agent must choose the optimal asset distribution. This is why the action at time step t, a_t, is simply just the portfolio vector w_t. Now having the environment and the action, we can now define the state, which is simply $s_t = (\mathbf{X}-_t, w_{t-1})$.

As stated above, optimizing the portfolio value is primary objective. Although this is the case, the agent does not have direct control over the initial portfolio value, p_0. Thus, the reward can't be the equation for the final portfolio value, p_{f_t}. Thus, the reward of the agent is to maximize the average of the logarithmic rate of return:

$$R(s_1, a_1, \dots, s_{t_f}, a_{t_f}, s_{t_f+1}) = \frac{1}{t_f} \sum_{t=1}^{t_f+1} ln(\mu_t * y_t) \times w_{t-1}$$

An interesting result of this rewards function is that all of the episodic rewards are equally important (2).

2.2 Policy Network

The policy network is used to approximate the "best" action (which is the distribution of money into the cryptocurrencies that will produce the highest value at a given time period) to take. This means that the network will produce the portfolio vector, w_t using the input tensor, $\mathbf{X}_t$ This network, a Convolutional Neural Network (CNN), is described below.

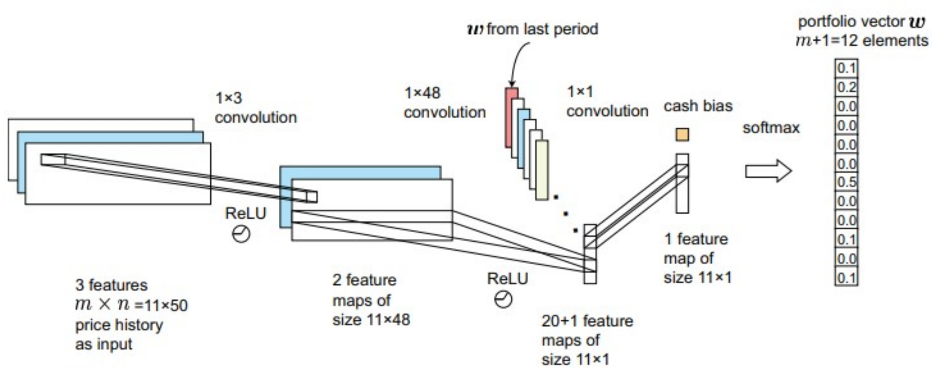

Figure 1: Policy Network implemented as a CNN

Each convolution operation involves a row vector (first dimension is always 1) from the feature map to make each row (corresponding to each cryptocurrency) independent of one another until the last layer. Each of these independent processes are referred to as Identical Independent Evaluators (IIE). When looking at all assets, they are referred to as the Ensemble Identical Independent Evaluators (EIIE).

3 Experiments

This section will outline the performance metrics and the results of this projects.

3.1 Evaluation Criteria

The evaluation metrics used in this project will be the same as the metric used in (2). These metrics are listed below:

- Final accumulated portfolio value (fAPV)
- Sharpe ratio (SR)
- Maximum Drawdown (MDD)

The fAPV is the ratio of the final portfolio value to the starting portfolio value. While this may not seem to be a bad metric, it doesn't take into account risk factors. This is where the SR measure comes in. The SR is the "average of the risk-free return by its deviation":

$$SR = \frac{\mathbb{E}_t[\rho_t - \rho_F]}{\sqrt{\underset{t}{var}(\rho_t - \rho_F)}}$$

where p_t and p_F are the periodic returns at time t and the rate of return of a risk-free asset, respectively. This is a better performance metric than the fAPV, but this metric treats movements in markets equally, which is not the case (as upward movements in the market result in profit and negative movements result in loss) (2). The third and final performance indicator, the MDD, is supposed to account for this. The MDD is the "biggest loss from a peak to a trough" (2):

$$D = \max_{\substack{\tau > t \\ t}} \frac{p_t - p_\tau}{p_t}$$

3.2 Results

Two different back tests (a period in which the model predicts portfolio values) were chosen for this project. The first back test was conducted from 01/01/2020 to 06/01/2020 and the second back test was spanned from 01/01/2021 to 06/01/2021. There were a couple of reasons why these time periods were selected. The interval for the first back test was chosen out of pure curiosity to determine how the model would be able to perform slightly before and slightly after the COVID-19 pandemic. The dates for the second back test was chosen to represent a more "normal" market, where prices were stable, relative to the first back test.

For the back tests, we will compare the EIIE with the Uniform Buy and Hold (UBAH) method, which uses a random, predefined distribution of investments in each assets, and only sells at the end of the time period. This method was chosen because it is very common technique that ordinary portfolio managers use; buy a handful of stocks and "wait and see" what the portfolio values is at the end of the time period.

Algorithm	Interval	MDD	fAPV	SR
CNN	01/01/2020 - 06/01/2020	0.427	1.404	0.0553
UBAH	01/01/2020 - 06/01/2020	0.599	1.311	0.0537
CNN	01/01/2021 - 06/01/2021	0.528	1.419	0.0474
UBAH	01/01/2021 - 06/01/2021	0.512	1.423	0.0492

Table 1: Performance Metrics across both Back Tests and Algorithms

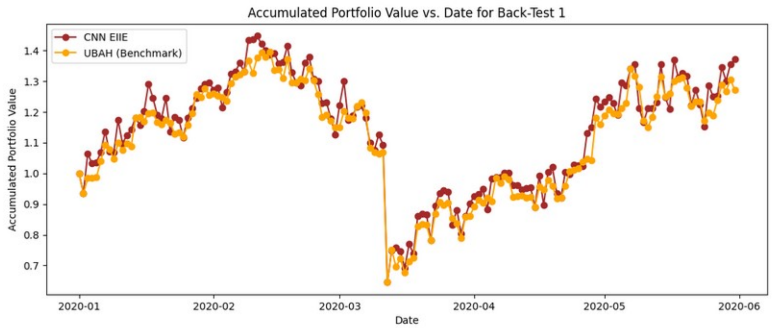

Figure 2: fAPV of the First Back Test

The results of the first back test are shown above. As is evident, the EIIE performs very closely to the UBAH approach, which was unexpected considering how the EIIE performed much better than the UBAH in (2) and because of the variation in the COVID-19 pandemic. Nonetheless, the EIIE technique ended up outperforming the UBAH method by almost 10% at the end of the time period. It is important to note that because neither technique outperformed the other longer than 9 days (02/05/2020 - 02/13/2020), it could very well be that the end result of the EIIE could've been in part due to the time period it ended at. The first two rows of Table 1 show the performance metrics of this back test. The fAPV of the EIIE is almost 10% higher than that of the UBAH technique. Although this is the case, the difference between the results of performance metrics for this project's implementation of (2) and the UBAH is much lower than the same difference in (2), most likely indicating there is an issue with this project's implementation of the policy network.

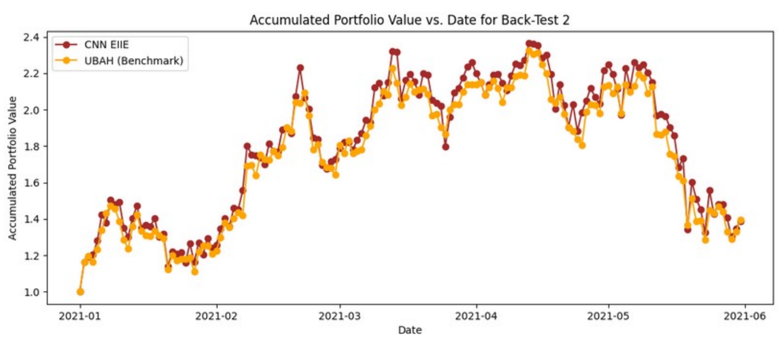

Figure 3: fAPV of the Second Back Test

As mentioned previously, another back test was conducted, with the results being showed above. The EIIE actually performs worse than the UBAH technique, as the fAPV of the EIIE is 0.04% lower than that of the UBAH. Comparing to the first back test, this was unexpected, as the volatility in the market of the second back is less than the volatility of the market in the first back test (as the market adjusted to the COVID-19 pandemic in the second back test). Even though this market is more stable than the market in the first back test, a potential reason for the weak performance could be the rise and fall of COVID-19 strains that arose in 2021, such as Delta and Omicron.

4 Conclusion

This project attempted to replicate the work in (2) to modern cryptocurrency prices. While the model performed better than a buying-and-holding strategy for some time periods, this result is unfortunately not achieved across many time periods. Thus, the next step in this project would be to identify what is incorrect in the project's implementation that is causing these results.

References

[1] J. Caporal, *How Many Americans Own Stock? About 145 Million – But the Wealthiest 10% Own More Than 80%*, Nov. 2021.

[2] Z. Jiang, D. Xu, and J. Liang, *A deep reinforcement learning framework for the financial portfolio management problem*, 2017.