Optimization of Cryptocurrency Investment Portfolios with Deep Learning

Sri Balaaji Natarajan Kalaivendan
Master's in Computing
Dublin City University
Dublin, Ireland
sri.natarajankalaivendan2@mail.dcu.ie

Prashanth Jayaraman
Master's in Computing
Dublin City University
Dublin, Ireland
prashanth.jayaraman2@mail.dcu.ie

***Abstract*—Recent years have seen a rise in the use of deep learning and reinforcement learning techniques for the active asset management task, with the goal of outperforming traditional benchmarks like the Buy-and-Hold approach [1]. This paper examines both Modern Portfolio Theory and reinforcement learning framework to active asset trading, highlighting the merits of each. The goal of financial portfolio optimization is to maximize returns while minimizing risk by redistributing money into multiple financial instruments at the same time. Reinforcement Learning (RL) and machine learning have been combined in this project to create a framework for a solution. One day's worth of Cryptocurrency trading data is presented in an open high low close format. Six Cryptocurrencies' historical trading data are used to train and test the built RL framework. We demonstrate that frequent rebalancing is a valuable risk management technique for asset types with significant volatility. Our findings contribute to the advancement of algorithmic finance. To extend these findings outside the realm of cryptocurrencies, further research is required.**

***Index Terms*—Portfolio Optimization, Cryptocurrency, Modern Portfolio Theory, Reinforcement Learning, Deep Learning**

I. Introduction

In this paper, rather than analyzing the basic qualities of cryptocurrencies, we focus instead on their technical properties, notably price movement and volume. Two characteristics namely decentralization and openness distinguish cryptocurrencies from traditional financial assets [5], making their market the ideal testing ground for our innovative machine-learning portfolio management studies. Without a central governing body, anyone can trade cryptocurrencies with minimal entry criteria, and cryptocurrency exchanges flourish. The abundance of low-volume currencies is a direct result [4]. Compared to regular markets, a lower amount of capital will be required to influence the pricing of these tiny marketplaces [5]. This will eventually enable trading machines to learn from their own market actions and capitalize on the resulting effects. In addition, market accessibility increases with openness. The majority of cryptocurrency exchanges have an application programming interface for getting market data and executing trading activities, and the majority of exchanges are open 24 hours a day, seven days a week, with no limits on the amount of trades. However, in practice transactions tend to be less frequent than regular occurrence. These nonstop markets provide much scope for investigating through machine learning algorithms.

When optimizing a financial portfolio, money is redistributed among various investment assets at a predetermined point in time in order to maximize profits while reducing losses. This research aims to develop a solution framework for this intricate financial engineering issue. The concepts behind Modern Portfolio Theory [2] which is closely related to the theory that a higher probability of a larger return is associated with a higher risk, and a higher probability of a smaller return is associated with a lower risk is used to build the basic framework. Given two portfolios with the same expected return, Modern Portfolio Theory predicts that investors will choose the portfolio with lower risk [2]. An investor will only take on more risk if the potential rewards are also greater [2]. The Reinforcement Learning framework is implemented using two machine learning methods: Convolutional Neural Network(CNN) and Long Short Term Memory (LSTM). Trading data for cryptocurrencies is used for all framework training and evaluation.

Reinforcement learning (RL) is a good framework for dealing with complex data and making hard decisions, such as when trading assets. In this framework, an agent i.e the investors takes actions in an environment based on what it knows about the state of that environment. Both the states visited and the actions taken give the agent rewards. In the case of trading assets, the state of the environment is the recent history of the assets. Actions are the transactions that are done to sell some of the assets the agent owns and get new ones. Rewards are scalar functions of the gains or losses the agent sees as a result of taking those actions. A person's portfolio is a list of all the assets he or she owns at any given time. This type of process is also called "portfolio management" because it involves managing a person's portfolio. The state and action spaces of RL algorithms are often immutable [3]. However, the cryptocurrency market is always expanding to accommodate additional assets [3]. Therefore, flexible state and action spaces are required to rapidly incorporate those assets into the process. Consequently, convolutional layers of neural networks are able to extract useful information regarding the prices of this particular set of assets.

Open, *High* , *Low*, and *Close* (OHLC) is the format used to display the price information for a specified time frame in the market [5]. *High* is the greatest price value achieved during the trading time, *Low* is the lowest price value achieved during the trading period, and *Close* is the closing price value achieved at the end of the trading period. Cryptocurrencies are subject to high-frequency trading, which means that the open price for a given period might not be the same as the closing price for the preceding period.

The purpose of this research is to create a Reinforcement Learning framework using Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM). The foundation of this system is Modern Portfolio Theory. Both the models are trained and tested with trading data of cryptocurrencies. CoinMarketCap API is used to fetch data of six different cryptocurrencies. The Dataset has Bitcoin, Ethereum, Litecoin, Bitcoin Cash, Monero, and Ripple data from 2013 to 2022 (present).The cryptocurrency market adopted a 24 hour trading cycle. The purpose of using a 24-hour window was to better understand cryptocurrency data by comparing it to that of other days and taking into account the natural swings in trading prices.As a result, here are our research questions:

- H1 : The optimising agent achieves a higher portfolio value than an equiweight agent.
- H2 : The LSTM agent outperforms the CNN agent in terms of risk-adjusted return (Sharpe ratio and Return over Maximum Drawdown(RoMaD)).
- H3 : The portfolio obtained by the optimising agent outperforms the portfolio obtained by Modern Portfolio theory in terms of Sharpe ratio.

The outline of our paper looks like this. We begin with an introduction to methods for optimising the portfolio and our research questions. In Section 2, we present a concise literature review, and in Section 3, we describe the methods, models, and evaluation metrics we use. Section 4 details our results, and Section 5 provides conclusions and future work.

II. Literature Review

A. Cryptocurrency Portfolio Management

"Follow-the-Winner", "Follow-the Loser", "Pattern-Matching", and "Meta-Learning" are four types of traditional portfolio management strategies [5]. The first two categories are based on previously built financial models, while parameter selections may be aided by machine learning techniques. The effectiveness of these strategies is determined by the models' applicability in various markets. "Pattern-Matching" algorithms select a portion of history that is comparable to the current circumstance and optimize the portfolio based on that history under certain market assumptions. The final class of methods, "Meta-Learning" aims to combine different classes of methods to improve performance [5].

In terms of AI Algorithms, Genetic algorithms uses a trial-and-error approach [6] to iterate over generations of solutions. The fitness of candidate solutions is evaluated in each generation, and the most fit solutions are allowed to breed. The fitness function, a problem-specific function that drives the optimization process, is the most important aspect of the method [6]. Lin and Liu demonstrated how to construct portfolios with good mean-variance efficiency using genetic algorithms. Their technique allowed them to calculate a near-optimal portfolio while taking into account the smallest transaction lots (Lin et al., 2007) [7]. Their methodology has the drawback of being static: while their model creates mean-variance optimized portfolios, they are only optimized for a short period of time since they lack trading logic. Aranha and Iba suggested a more dynamic method to portfolio optimization. They developed a hybrid model that combines a genetic algorithm with a local search technique. Portfolios can be rebalanced while taking trading expenses into account [8].

Gorgulho, Neves, and Horta offered another GA-driven active management technique. They suggested a stock selection investing simulator that makes use of technical indicators including the exponential moving average (EMA), the relative strength index (RSI), and the Moving Average Convergence Divergence oscillator (MACD) [9]. Their GA algorithm keeps track of a population of portfolios and executes weekly trades depending on technical indications [9].

Markowitz portfolio theory is examined here as a framework for constructing a cryptocurrency investment portfolio [10]. In order to enhance return or limit risk, a portfolio's asset allocation may be adjusted. Returns under the Markowitz model are assumed to follow a normal distribution. Using this premise, different investment portfolios can be easily compared based on only two criteria: standard deviation and mathematical expectation, which greatly reduces the complexity of the portfolio selection problem. Investing in a diverse selection of cryptocurrencies has the potential to provide better low-risk growth of capital, as demonstrated by examples of cryptocurrency-only portfolio construction [11]. In the paper [12], researchers employ a portfolio diversification method based on many models of portfolio development. On the basis of Modern Portfolio Theory, an optimal risk portfolio has been formulated, and the impact of cryptocurrency on the conventional investment portfolio of assets has been explored. The data reported in the research [13] indicate that the expected return on a cryptocurrency portfolio is higher than the return on a coin individually.

B. Deep learning in portfolio optimization

Deep reinforcement learning, which employs deep neural networks in both the actor and critic, has been shown to be effective in solving financial problems [14]. Machine learning models that are loosely influenced by the architecture and functions of human brains are known as deep learning (DL) models. They, like genetic algorithms, can uncover patterns in data without being explicitly instructed what to look for while only investigating a tiny portion of a large search area [14]. Thus this can discover and exploit relationships in data that are undetectable to any known financial theory due to their flexibility. The lack of technical indicators is a fundamental advantage of DL models over GA techniques. Human-selected elements from a vast variety of existing technical indicators are used in GA techniques. DL techniques, on the other hand, rely on automated feature learning [15]. This means they can use the raw pricing data to uncover previously undiscovered patterns. Jiang et al training's procedure is as follows[4]. The neural network used to train all coins is the same. Predictions for each cryptocurrency, on the other hand, are made separately. The high, low, and closing prices from the preceding 50 data points are used to make each forecast. To account for transaction fees (0.2 percent every transaction), predictions are graded using a reinforcement learning agent and the weights from the prior period. Finally, using a softmax function, the scores from the previous stage are transformed to weights, ensuring that the weights of the portfolio assets add up to one. Jiang, Xu, and Liang's results, on the face of it seem very good. Their strongest model, they claim, was able to achieve at least 4-fold returns in just 50 days. This conclusion cannot be explained by the spectacular bull run that cryptocurrency saw in 2017 when Bitcoin was employed as a monetary asset. Furthermore, the dates of their tests were from September 2016 to April 2017 [5]. Deng et al. adopted a similar design to Jiang et al (2017)[16]. Deep learning and reinforcement learning were also mixed by the authors. Deng et al. tested their method on stock indices and commodities futures, and the results were encouraging. Their findings revealed that the approach works best when applied to a market that is trending. However, because their model only traded one asset, it isn't a full-fledged portfolio optimization model. Deng used fuzzy clustering on the raw pricing data, which was a fundamental distinction between his and Jiang et al approach. This entails discretizing raw returns to fuzzy linguistic clusters in order to reduce the uncertainty of input data (Returns of 4.5 percent to 5.5 percent, for example, may be discretized as "extremely high."). The input vector becomes more resilient and less susceptible to slight changes as a result of this strategy. A pure DL model attempts to discriminate and discover significance between extremely near responses (such as +5.23 percent 15 and 5.24 percent) without fuzzifying. Fuzzifying, in other words, takes the average of comparable responses and trains the network to detect patterns in the robust fuzzified representations [16].

III. Methods

A. Data Description and Preprocessing

There are six different cryptocurrencies included in this list, and they were chosen not based on their correlation values but rather on their market cap and trading volume as an index as shown in Figure 1. The greater the volume of a cryptocurrency (referred to as a coin below), the more liquid it is on the market.The six cryptocurrencies chosen are Bitcoin, Ethereum, Litecoin, Bitcoin cash, Monero, Ripple.

#	Name	Price	Market Cap	Fully Diluted Mcap	Volume(24h)	Circulating Supply	Total Supply
1	Bitcoin BTC	$22,932.63	$438,154,501,159	$481,599,525,560	$35,187,599,812 1,534,345 BTC	19,105,593 BTC	19,105,593 BTC
2	Ethereum ETH	$1,619.73	$197,098,823,933	$197,098,823,933	$24,817,463,772 15,327,304 ETH	121,728,542 ETH	121,728,541.69 ETH
7	XRP XRP	$0.3531	$17,068,763,992	$35,307,548,687	$1,504,499,665 4,261,127,495 XRP	48,343,101,197 XRP	99,989,535,142 XRP
21	Litecoin LTC	$59.49	$4,213,020,529	$4,997,383,254	$593,017,869 9,967,917 LTC	70,815,806 LTC	84,000,000 LTC
27	Monero XMR	$157.39	$2,857,083,862	$2,857,083,862	$184,719,747 1,173,675 XMR	18,153,380 XMR	18,153,379.97 XMR
29	Bitcoin Cash BCH	$142.19	$2,721,571,080	$2,987,671,218	$558,508,947 3,925,696 BCH	19,129,613 BCH	19,129,612.5 BCH

Figure 1. Cryptocurrency MarketCap values and Volume in circulation. Sourced from https://coinmarketcap.com/

A sampling of the total trading days collected ranged from 3000 to 3500 because some cryptocurrencies like Bitcoin were launched in 2009, while other cryptocurrencies like Ethereum were launched much later. The 24-hour trading data was gathered using cryptocmd [18] and the Poloniex API [19] and this data was then pre-processed to fit the same timeline 2016-2022(Present).

B. Portfolio Management Using Modern Portfolio Theory

The foundation of Modern Portfolio Theory is the idea of the risk-return trade-off, which states that the likelihood of a higher return increases risk and vice versa [10]. The theory makes the assumption that investors are risk-averse, meaning that given a choice between two portfolios with equivalent expected returns, they will pick the one with lower risk. An investor will only take on more risk if the potential rewards are larger. This idea also emphasizes diversifying ones investments rather than putting all of your money into one asset. As a result, it may be possible to achieve lower risk than the portfolio with the lowest risk by optimizing allocation. The weighted average of the individual cryptocurrency return makes up the overall return. The correlation between the movements of the assets under consideration determines the overall risk, though. As a result, the total risk does not match the weighted average of the risk associated with certain assets. Modern Portfolio Theory simulates random weight vectors for the portfolio and the Annualized returns as well as the annualized volatility for the portfolio were computed for each weight vector produced.

- $$\text{Annualised Returns}_t = (w_{(t)}^T.M) * T) \tag{1}$$

  where,
  M = Mean Daily Returns for the individual assets considered.
  T = Total trading periods for a year (Cryptocurrency [24 Hours] - 365 days, Cryptocurrency [30Min] - 365*48 intervals)

- The standard deviation of the portfolio over a year beginning at the time t when the weight vector w is allocated to the portfolio is known as its annualized volatility.

$$\text{Annual Realized Volatility}_t = \sqrt{(w_{(t)}^T(C.w_{(t)}))} \quad (2)$$

where,
C= Calculated daily returns covariance matrix for all assets in the sample.

- The Sharpe ratio is a measurement of the excess returns over the risk-free rate of return per unit of portfolio risk [4]. It is determined by:

$$\text{Sharpe ratio} = \frac{R_a - R_f}{\text{Standard Deviation}_a} \quad (3)$$

How much of an advantage in terms of returns there is for being willing to put up with somewhat higher volatility in exchange for owning the riskier asset is quantified by this ratio.The standard deviation is the square root of the variance, and R_f is the annualized risk-free rate. Sharpe's formula was devised as a way to evaluate returns relative to risk for a portfolio (1966). The formula can be applied to the evaluation of both single assets and entire portfolios. Additionally, it is useful for contrasting the success of various financial holdings. Generally speaking, a Sharpe ratio above one is desirable. The Sharpe ratio is a measure of how much a portfolio of assets are outperforming the risk-free rate. Optimal portfolios with the highest Sharpe ratio and the lowest volatility are found.

C. Portfolio Management Using Reinforcement Learning

1) Building Price Tensor: Since only the changes in prices, and not the raw prices themselves, will determine the performance of the portfolio management, four lists containing the scraped raw price values for Open, High, Low, and Close (OHLC) over the entire time frame have been created for each cryptocurrency, and then normalized by the opening price. To achieve this uniformity, the lists have been normalized as follows:
Period t input tensor P_t has 4 components, denoted by ${P_t}^{Open}$, ${P_t}^{High}$, ${P_t}^{Low}$, and ${P_t}^{Close}$. The input tensor contains the OHLC prices (*Open*, *High*, *Low*, *Close*) for the previous n days. To make the features be the relative price changes, we divide the price of each dimension by ${P_t}^{open}$, the last open price. After normalization, all cryptocurrency prices will be on the same scale [20], and the resulting characteristics will accurately reflect any further price movement. All three of ${P_t}^{high}$, ${P_t}^{low}$, and ${P_t}^{close}$ use the same normalizing technique.

$$P_t = [{P_t}^{Open}, {P_t}^{High}, {P_t}^{Low}, {P_t}^{Close}]$$

- $$\text{Close Price List} = [\frac{P_{(t-n-1)}^{Close}}{P_{(t-n-1)}^{Open}}, ..., \frac{P_{(t-1)}^{Close}}{P_{(t-1)}^{Open}}]$$

- $$\text{High Price list} = [\frac{P_{(t-n-1)}^{High}}{P_{(t-n-1)}^{Open}}, ..., \frac{P_{(t-1)}^{High}}{P_{(t-1)}^{Open}}]$$

- $$\text{Low Price list} = [\frac{P_{(t-n-1)}^{Low}}{P_{(t-n-1)}^{Open}}, ..., \frac{P_{(t-1)}^{Low}}{P_{(t-1)}^{Open}}]$$

- $$\text{Open Price list} = [\frac{P_{(t-n)}^{Open}}{P_{(t-n-1)}^{Open}}, ..., \frac{P_{(t)}^{Open}}{P_{(t-1)}^{Open}}]$$

The neural network receives a price tensor as input at the end of period t with dimensions (x; y; z), where x is the number of features (OHLC in case of portfolio management), y is the number of non-cash assets, and z is the number of trading periods evaluated till time t. This tensor represents the 24-hour moving average of cryptocurrency prices and has the shape (4, 6, 1513).

2) Reinforcement Learning: The primary focus of reinforcement learning, one of the most fundamental type of machine learning, is on determining the best course of action for a software agent to take in a given environment so as to maximize a predefined criterion of cumulative reward and progress toward a desired outcome. The solution framework described here interprets the financial market as the environment in which portfolio optimisation software allocates capital among various assets at a given point in time. Using an approximation of the profit function, this goal-oriented RL framework trains the aforementioned neural networks (CNN and LSTM) to maximize profit. w_t is the result of an agent's decision at time t to rebalance their holdings across different investments.

$$Action_t = w_t = [w_{cash}; w_{asset1}; w_{asset2}; ...; w_{assetn}] \quad (4)$$

Reallocating the funds (i.e., buying or selling assets) takes into account the transaction cost, which is typically a constant commission fee involved for each cryptocurrency traded. In this way, at time t, the total transaction cost [19] is determined while the funds are redistributed using the updated weight vector. i.e w_t from weight vector at time *(t-1)* i.e. w_{t-1}

$$\text{Total Transaction cost}_{(t)} = \\ \text{Portfolio value}_{(t-1)} * \text{Transaction fee} * (w_t - w_{(t-1)}) \quad (5)$$

This transaction cost is deducted from the value of the stock portfolio, or the total amount of money, at each time step t.

$$\text{Portfolio Vector}_{(t)} = (\text{Portfolio value}_{(t-1)} * w_t) \\ - (\text{Total Transaction cost} + [0] * \text{Number of assets}) \quad (6)$$

The value of a portfolio as of time t is stored in a vector of size (number of assets + 1).

$$\text{Cumulative Portfolio Vector}_{(t)} = (\text{Portfolio value}_{(t-1)} * w_t) \\ - (\text{Total Transaction cost} + [0] * \text{Number of assets}) \\ * \text{Return rate}_{(t)} \quad (7)$$

$$\text{Final Weight vector}_{(t)} = \text{Portfolio vector}_{(t)} / \text{Cumulative Portfolio Vector}_{(t)} \quad (8)$$

If the agent has a positive balance at any given time, it will receive the asset's interest rate (represented by the return rate at time t) at that time and if it has negative balance it makes a payment. The value of the reward at time *t* is the percentage increase or decrease in the value of a portfolio and is thus tied to its overall worth which includes the *portfolio value(t-1)* and *total transaction cost(t)*.

$$Reward_{(t)} = (\text{Portfolio value}_{(t)} / \text{Cumulative Portfolio Vector}_{(t-1)}) - 1 \quad (9)$$

By factoring in w_{t-1} and *Portfolio Value*$_{t-1}$ as input to the agent at time *t*, we can fully capture this dependency. At a given instant in time *t*, *state(t)*, is defined by the price tensor at time *t*, weight vector at time *t-1* which is w_{t-1}) and Value of the Portfolio at Time *t-1* where $W0 = [1,0,0...0]^T$

$$State_{(t)} = (\text{Price Tensor}_{(t)}, w_{t-1}, \text{Portfolio value}_{(t-1)}) \quad (10)$$

The optimization agent for the portfolio and a starting point in which the assets are all given equal weights. Assuming this setup, the revised reward at time *t* is determined as follows:

$$\text{Adjusted Reward}_{(t)} = Reward_{(t)} - \text{baseline Reward}_{(t)}) * \alpha * max(w_{(t)}) \quad (11)$$

Where the maximum of the weight vector at time *t* is represented by the term α* *max(w(t))*. This metric is considered so that the representative doesn't put all of their money into one stock.

3) Convolutional Neural Network: The policy function instructs the agent on how to reallocate the weights among assets to maximize profit, given a specific state. CNN is used to craft the policy function for the first model. The input to this network is the price tensor *(price tensor(t))*, and the outputs are the weight vector *(w(t))*, the portfolio value at time *t (Portfolio value(t))*, and the adjusted reward at time *t (Adjusted Reward(t))*, as shown in **Figure 2**. The first convolution layer's activation function is ReLU for cryptocurrency data which is implemented using a brute-force approach. The second convolution layer has 20 filters and generates a vector of size (number of assets * 1 * 1).The output vector from before is stacked. The activation function used here is ReLU.

The portfolio vector of the previous time step is inserted before the last layer of the neural network in order to reduce transaction costs. A Portfolio Vector Memory (PVM) is used to store the network's output portfolio vector between time steps so that it can be used at the subsequent time step. Final convolution creates a unique vector with the same dimensions as the number of assets. After that, a bias toward cash is introduced, and softmax is used. This network has a unique feature: the network parameters are shared among all the assets, but they are processed individually. The softmax function guarantees that the weights used in the calculation are positive and add up to 1.

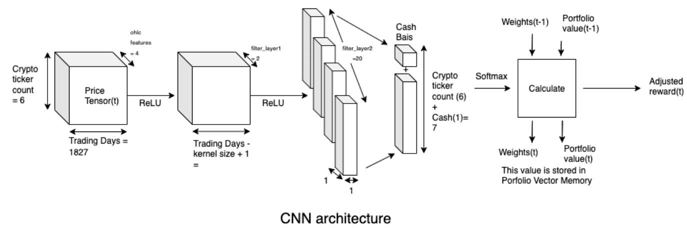


Figure 2. CNN Architecture Diagram

4) Long short term memory (LSTM) neural network: Policy design in the second model is accomplished with LSTM. If you want to process and make predictions based on sequential data, like the sequence of historical trading data, then LSTM is the way to go. **Figure 3** depicts the process flow for the LSTM network, which is very similar to the flow for the CNN network. The replacement of the series of two convolution layers in the RNN with a simple RNN unrolling layer is what ultimately distinguishes the two. The unrolling of a recurrent neural network is essentially a loop over successive time steps in which, in addition to the input at that time step, the network also takes in information remembered or passed on from the previous time step. Because of its chain-like structure, LSTM is highly dependent on sequences. Therefore, small recurrent subnets are responsible for collecting the prices of individual assets, and the structure of the ensemble network following the recurrent subnets is identical to that of the CNN's second half.

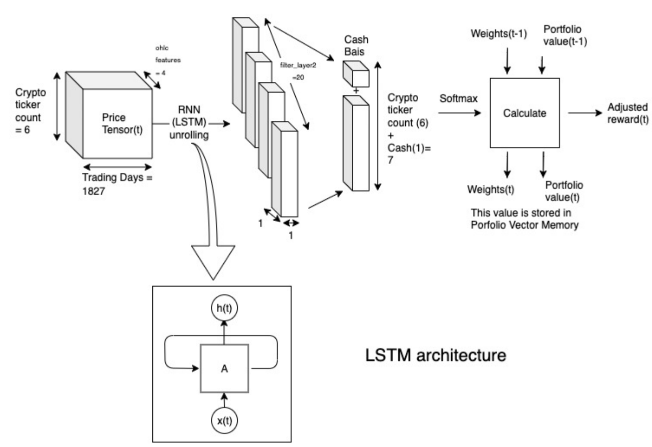


Figure 3. LSTM Architecture Diagram

D. Evaluation metrics

1) Benchmark Testing: The first metric used for evaluation is benchmark testing[17], which compares the results obtained by the constructed optimizing agent to those obtained by two other agents: one that allocates equal weights to all of the available assets, and another that does not invest in any of the assets. The result is the total value of the portfolio after all the steps in the test have been executed. While the hypothesis predicts superior performance from an optimizer agent over an equiweight agent, the latter fails to account for risk when calculating returns. For this reason, we also took into account the Sharpe Ratio and the Return over Maximum Drawdown, two additional metrics for assessing performance via historical data.

2) Backtesting : Backtesting is a term used in financial modeling to describe the process of evaluating the efficacy of a predictive strategy by observing its performance on historical data. It's an essential part of making a good trading system. Backtesting rely heavily on two assumptions [5]:

- First, there is no price movement (or "slippage") because all assets on the market are sufficiently liquid.
- The software trading agent's capital investment is negligible, so it has no effect on the market.

If the volume of trades in a given market is high enough, these two hypotheses come close to being true in practice. If the trading volume is high enough, these two assumptions are close to reality in the real world trading environment.To quantify return relative to risk, the Sharpe ratio is used [4]. When the current strategy generates a Sharpe ratio that is

- less than 1, it is sub-optimal and can be improved upon significantly.
- greater than 1, then it's decent.
- higher than 2, it is deemed good.
- higher than 3, it is deemed excellent.

There however is a risk of being misled when using the Sharpe Ratio as an evaluation tool because it uses standard deviation in the denominator, which is based on the assumption that financial returns follow a normal distribution [21]. Since it is unlikely that returns will be normally distributed, an additional metric called Return over Maximum Drawdown (RoMaD) was used because it is predicated on the idea that loss patterns observed over longer periods are the best surrogate for actual exposure. To achieve a return of Y percent, an investor is willing to tolerate a drawdown of X percent on occasion is the main ideology behind the method.

$$RoMaD = CPV/MDD \tag{12}$$

where,
CPV = Cumulative Portfolio Value
MDD = Maximum recorded decline from its highest point to its lowest point before recovering to a new high.

$$MDD = (PeakValue - TroughValue)/PeakValue \tag{13}$$

Downside risk over a given time period can be measured using *MDD*. Current strategy is deemed successful if

- it produces a RoMaD of less than 0.5 it is underperforming, with a lot of room for improvement.
- it's greater than 0.5, that's excellent.
- the value falls somewhere between 3 and 5, then it's acceptable.

IV. RESULTS

It has been observed that the price of one Bitcoin (BTC) is substantially higher when compared to the prices of other assets, and that this price has shown a great deal of variation over the past few years. Nevertheless, each of the other cryptocurrencies can be found within the same general range, which makes it challenging to comprehend the related shift in value over time by looking at the plot that was created. In light of this, a more effective method for determining the degree of volatility exhibited by each asset is to plot the daily returns of each asset. The term "daily returns" refers to the percentage change in the closing price of the current trading day in comparison to the closing price of the previous trading day. It has been observed that Ripple (XRP) and Bitcoin Cash (BCH) have demonstrated huge variation in their daily returns, whereas Bitcoin (BTC) has not demonstrated any extreme variation over the course of time.

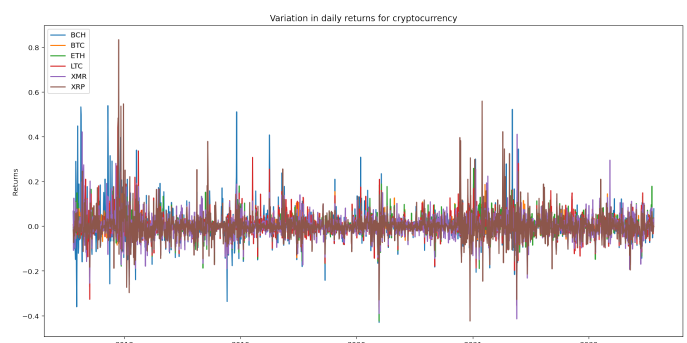


Figure 4. Variation in daily returns

A. Modern Portfolio Theory

An investor wants to make a lot of money while taking on a manageable amount of risk, the investor will seek to achieve the highest possible returns while maintaining the lowest possible standard deviation. As a result, the optimal choice of portfolio will be one that provides a combination of standard deviation and return that falls within the efficient set. We decided to settle on the following three key objective functions: Maximizing the Sharpe ratio, Maximizing the yearly return of the portfolio and Minimizing the annual volatility of the portfolio (standard deviation) [10]. After plotting randomly produced portfolios using a color map based on the Sharpe

Ratio, it can be seen in the fig 5 that a curving boundary is formed by the dots to the left of the cluster that was obtained. This cluster was obtained after plotting the random portfolios using the color map. The name given to this newly constructed boundary is the Efficient Frontier. The portfolio that has the least amount of risk (volatility) is represented by the green dot in the plot, and the portfolio that has the highest Sharpe Ratio is represented by the red dot. If a riskaverse investor were presented with two portfolios that produced the same returns, they would select the one that contained the greatest number of yellow dots, which indicated the lowest level of risk.

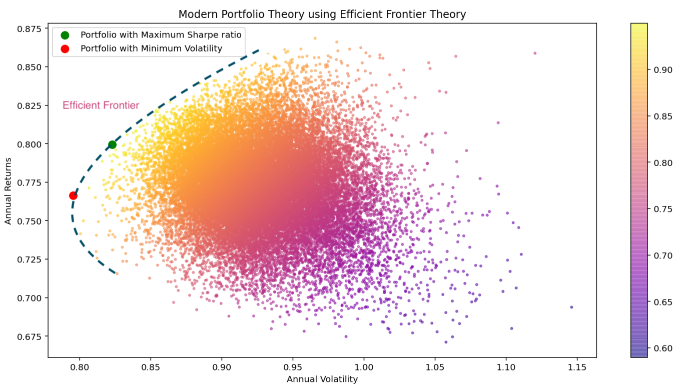
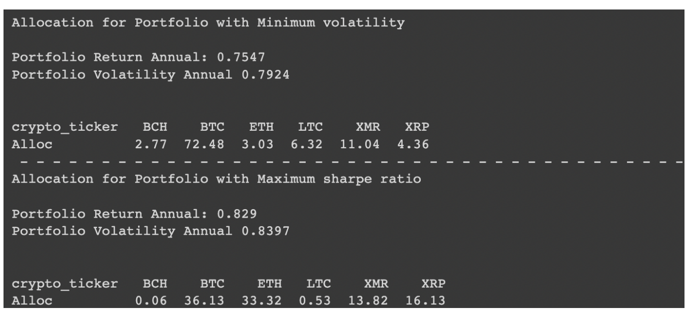


Figure 5. Efficient Frontier curve using Modern Portfolio Theory



```
Allocation for Portfolio with Minimum volatility

Portfolio Return Annual: 0.7547
Portfolio Volatility Annual 0.7924


crypto_ticker   BCH    BTC    ETH   LTC    XMR   XRP
Alloc          2.77  72.48  3.03  6.32  11.04  4.36
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Allocation for Portfolio with Maximum sharpe ratio

Portfolio Return Annual: 0.829
Portfolio Volatility Annual 0.8397


crypto_ticker   BCH    BTC    ETH   LTC    XMR    XRP
Alloc          0.06  36.13  33.32  0.53  13.82  16.13
```

Figure 6. MPT Allocation

Observing the weight vectors that have been assigned in figure 6 for the minimum volatility portfolio allocation, we can see that Bitcoin (BTC) has been given the highest weight. This is consistent with the fact that Bitcoin is one of the least volatile cryptocurrencies, as demonstrated by figure 4, and investing in Bitcoin carries the least risk possible. In order to achieve the best possible Sharpe Ratio portfolio allocation, Bitcoin (BTC) and Etherium (ETH) are given the greatest amount of weight. XMR and XRP are also given a considerable amount of weight, despite the fact that their price fluctuations are far larger than those of BTC (as can be seen in figure 4).

B. Reinforcement Learning

In order to better visualize the findings of the benchmark, cumulative portfolio values were shown for the optimizing agent, the equi-weight agent, and the case in which there was no investment done. It is clear from looking at figures 7 and 9 that the optimizing agent has the best performance for both LSTM and CNN. The plots show that the LSTM policy is giving larger returns when compared to the CNN policy. This is because LSTM is more accurate than CNN.

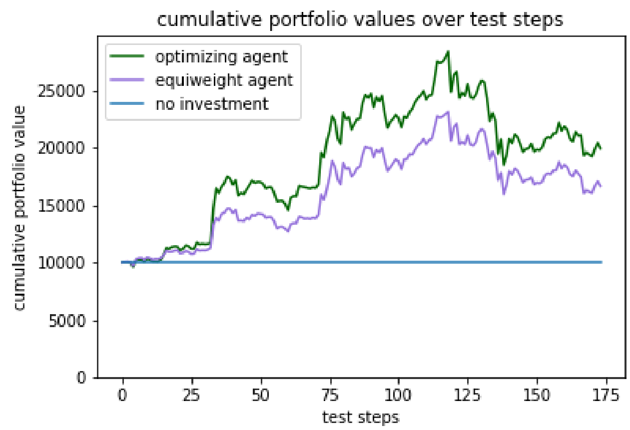


Figure 7. Cumulative Portfolio Value for CNN

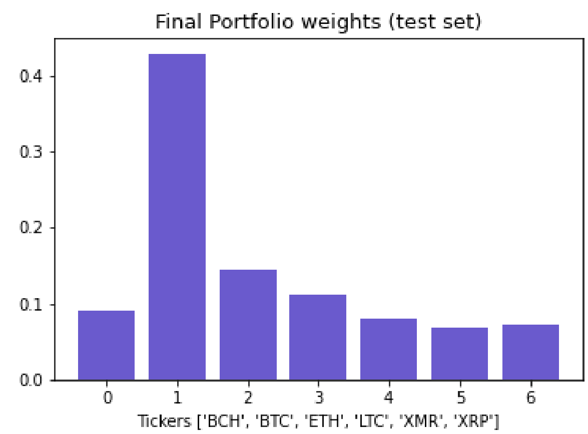


Figure 8. CNN Allocation

Calculations of Mean Sharpe Ratio and Mean RoMaD were performed in order to validate the results of the benchmark. As can be observed from the table, the MPT policy resulted in the lowest mean Sharpe ratio (1.22), while the LSTM strategy resulted in the highest (1.934). This is consistent with the results of the benchmark that were achieved. The results of the mean RoMaD point to the same conclusion, LSTM(1.77) performs better than CNN(1.463). Figure 8 shows that Bitcoin Cash, Ethereum, and Bitcoin were given the highest weights under the CNN policy. Figure 10 shows that practically all of the assets were given the almost equal weight under the LSTM. However, when taking into consideration the ultimate portfolio value after all of the testing processes, LSTM comes out on top by a significant margin.

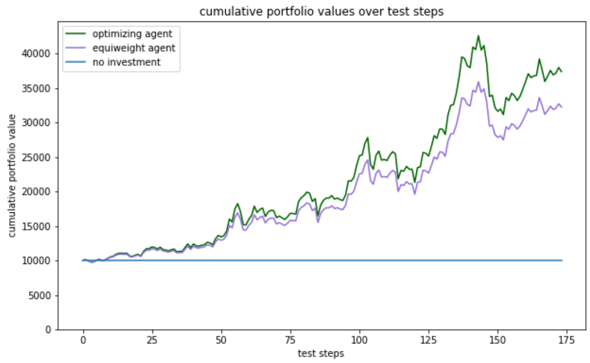


Figure 9. Cumulative Portfolio Value for LSTM

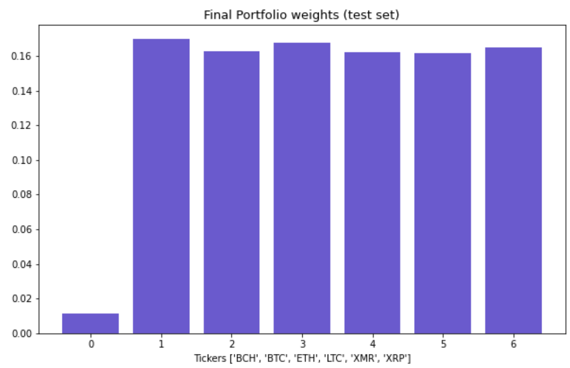


Figure 10. LSTM Allocation

V. Conclusion and Future Work

It has been demonstrated that models built with the Reinforcement Learning framework perform more efficiently than conventionally existing baseline methodologies such as Modern Portfolio Theory. LSTM, one of machine learning approaches that can be implemented using RL, has been demonstrated to perform significantly better than CNN in circumstances in which a large amount of sequential data is available. The methodologies that were used in project using Reinforcement Learning have yielded rather solid results, with a good balance between the returns on portfolio investments and the volatility of such returns.

Table I
Backtesting Results

Evaluation Metrics		
Model Type	Mean Sharpe Ratio	Mean RoMaD
CNN	1.631	1.463
LSTM	1.934	1.772
Modern Portfolio Theory	1.221	NaN

VI. Discussion

In the future, the models will be able to be trained using data collected every 30 minutes, and this will allow us to get useful insight on how the model functions when presented with more granular data while taking intraday traders into consideration. The modules that were developed can be incorporated into an automated trading system in order to sell and buy assets on the market in line with the optimal portfolio allocation that was determined. Because the policies that were learned using the RL trading framework were stable and showed good performance, the framework that was constructed can be used to generic reinforcement learning domains, such as recommendation systems, with just minor alterations.

Acknowledgment

We would like to thank Dr. Martin Crane (Head of School of Computing, DCU) for his guidance and assistance during the entire process. Additionally, we would like to thank Dublin City University for providing us with the chance to conduct this study.

References

[1] Felizardo, L. K., Paiva, F. C. L., de Vita Graves, C., Matsumoto, E. Y., Costa, A. H. R., Del-Moral-Hernandez, E., Brandimarte, P. (2022). Outperforming algorithmic trading reinforcement learning systems: A supervised approach to the cryptocurrency market. Expert Systems with Applications, 202, 117259.

[2] Kim, R. (2018). Efficient Frontier Portfolio Optimisation in Python. Retrieved 30 June 2022, from https://towardsdatascience.com/efficient-frontier-portfolio-optimisation-in-python-e7844051e7f

[3] Betancourt, C., Chen, W. H. (2021). Deep reinforcement learning for portfolio management of markets with a dynamic number of assets. Expert Systems with Applications, 164, 114002.

[4] Jiang, Z., Liang, J. (2017, September). Cryptocurrency portfolio management with deep reinforcement learning. In 2017 Intelligent Systems Conference (IntelliSys) (pp. 905-913). IEEE.

[5] Zhengyao Jiang, Dixing Xu, and Jinjun Liang A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem. Xi'an Jiaotong-Liverpool University, Suzhou, SU 215123, P. R. China.

[6] Kim, K., Han, I. (2000): "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index", Expert Systems with Applications

[7] Lin, C and Liu, Y (2008): "Genetic algorithms for portfolio selection problems with minimum transaction lots", European Journal of Operational Research

[8] Aranha, C. and Iba, H. (2009): "Using Memetic Algorithms To Improve Portfolio Performance In Static And Dynamic Trading Scenarios", Proceedings of the 11th Annual conference on Genetic and evolutionary computation

[9] Gorgulho, A., Neves, R. and Horta, N. (2011): "Applying a GA kernel on optimizing technical analysis rules for stock picking and portfolio composition", Expert Systems with Applications

[10] H. Markowitz, "Portfolio Selection", J. Fin., vol. 7, no. 1, pp. 77– 91, 1952. doi:10.1111/j.1540-6261.1952.tb01525.x

[11] A. Brauneis, and R. Mestel, “Cryptocurrency-portfolios in a mean-variance framework”, Fin. Res. Let., vol. 28, pp. 259–264, 2019. Doi:10.1016/j.frl.2018.05.008.

[12] Y. Andrianto, and Y. Diputra, “The Effect of Cryptocurrency on Investment Portfolio Effectiveness”, J. Fin. Accoun., vol. 5, no. 6, pp. 229-238, 2017. Doi:10.11648/j.jfa.20170506.14.

[13] N. Borri, “Conditional tail-risk in cryptocurrency markets”, J. Empir. Finance, vol. 50, pp. 1–19, 2019. Doi:10.1016/j.jempfin.2018.11.002.

[14] . Kim, T. W., Khushi, M. Portfolio optimization with 2D relative-attentional gated transformer. arXiv 2020. arXiv preprint arXiv:2101.03138.

[15] . Heaton, J., Polson, N. and Witte, J. (2016): “Deep learning for finance: deep portfolios”,Applied Stochastic Models in Business and Industry

[16] Deng, Y. et al. (2016): “Deep Direct Reinforcement Learning for Financial Signal”, IEEE Transactions on Neural Networks and Learning Systems (Volume: 28 , Issue: 3 March 2017)

[17] Chi Zhang, Corey Chen, Limian Zhang. Deep Reinforcement Learning for Portfolio Management. https://www-scf.usc.edu/zhan527/post/cs599/

[18] Scraping cryptocurrency data with CryptoCMD https://pypi.org/project/cryptocmd/0.3.2/

[19] Scraping cryptocurrency data with Poloneix API https://docs.poloniex.com/introduction

[20] Syu, J. H., Wu, M. E., Ho, J. M. (2020, October). Portfolio management system with reinforcement learning. In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 4146-4151). IEEE.

[21] Obeidat, S., Shapiro, D., Lemay, M., MacPherson, M. K., Bolic, M. (2018). Adaptive portfolio asset allocation optimization with deep learning. International Journal on Advances in Intelligent Systems, 11(1), 25-34.

Appendix

- Sri Balaaji Natarajan Kalaivendan: Conceptualization, Literature Review, Data pre-processing, Building RL Environment, Methodology, Visualization, Benchmark and Backtesting , Report - Review Editing.
- Prashanth Jayaraman: Conceptualization, Literature Review, Data Scraping , Methodology, Visualization, Results Validation, Report - original draft.

IEEE