Journal of Industrial and Systems Engineering
Vol. 15, No. 2, pp. 94-111
Spring (April) 2023

Optimizing Dynamic Portfolio Management in the Cryptocurrency Market Using Multi-Agent Deep Reinforcement Learning and the Fear and Greed Index

Mohammad Vahidpour[1], Amir Daneshvar[2*], Mohsen Amini Khouzani[3], Mahdi Homayounfar[4]

[1]*Department of Information Technology Management, Science and Research Branch, Islamic Azad University, Tehran, Iran*

[2]*Department of Industrial Management, Electronic Branch, Islamic Azad University, Tehran, Iran*

[3]*Department of Financial Engineering, Shahr-e-Qods Branch, Islamic Azad University, Shahr-e-Qods, Iran*

[4]*Department of Industrial Management, Rasht Branch, Islamic Azad University, Rasht, Iran*

Abstract

Selecting an appropriate investment portfolio is the foundation of modern financial theories and is particularly crucial in the highly diverse, complex and risky cryptocurrency market. Traditional models, such as the Markowitz model, are inadequate for addressing the dynamics of these financial environments, particularly in dealing with high computational complexity. Enhanced investment performance, and the fulfillment of diverse investor objectives hinge on selecting the right model and strategy. This study introduces a deep reinforcement learning environment capable of adapting to a dynamically changing state space with various assets. The proposed model adopts a multi-agent approach, with each agent assigned to a specific asset. It utilizes two DQN neural network models for action selection and LSTM neural network for predicting trend. For selecting appropriate actions, nine different indicators were utilized, including the Fear and Greed Index to identify market sentiment and other technical indicators. The dataset comprises 11 non-stable cryptocurrencies and one stablecoin to preserve capital value. Two different strategies were used to test the model. The cumulative profit and the Sharpe ratio were employed as evaluation metrics. The results indicate that the average profitability of the proposed model is 2.32 times higher, and the Sharpe ratio is 1.45 times greater than the buy-and-hold strategy. Additionally, the use of LSTM alongside DQN leads to more appropriate action selection, ultimately optimizing and enhancing the profitability of the investment portfolio.

Keywords: Cryptocurrency, Deep reinforcement learning, Long short-term memory, Machine Learning, Portfolio optimization, Fear and Greed Index

[*] Corresponding author
ISSN: 1735-8272,

1. Introduction

Portfolio selection is a fundamental concept in finance domain, with its initial studies often attributed to Markowitz in 1952. His groundbreaking research laid the groundwork for modern financial theories that emerged towards the end of the 20th century. (Jacob, 1974). Portfolio optimization entails the process of identifying and selecting the most suitable assets (stocks) and determining the optimal capital allocation. Employing optimization models in allocating risky assets require decisions under uncertainty. In essence, modern portfolio theory, optimization aims to maximize portfolio returns while minimizing risk (Lim et al, 2022). Effective management of the investment process, known as portfolio optimization, remains a crucial concern in financial matters. The methodology for constructing such a stock portfolio has intrigued to researchers.

Recent advancements in machine learning have increasingly led researchers to explore various techniques utilizing artificial intelligence, particularly in emerging fields such as financial markets. Traditional Markov chain models are insufficient due to the inherent volatility and complexity of these markets. One solution to this challenge is the application of reinforcement learning. Reinforcement learning, a subset of machine learning techniques, continuously evaluates performance and adjusts strategies based on past experiences and received feedback. This method enables the adaptation of strategies to change market conditions, providing dynamic and flexible responses to market fluctuations. The primary goal of reinforcement learning is to maximize rewards, making it highly applicable in financial markets for optimizing portfolio returns (Jiang & Liang, 2017). A significant challenge in reinforcement learning is the curse of dimensionality, which denotes the exponential growth of the state space as the number of dimensions increases. Despite this challenge, reinforcement learning enables decision-making in environments with partial information, resulting in a more efficient and effective decision-making process (Xiang and Foo, 2021).

In recent years, the financial applications of blockchain technology have experienced substantial growth and development among investors. Concurrently, the research community has devoted significant efforts to this field. The traditional approach to portfolio analysis proves inadequate for portfolios containing crypto assets, as it relies on the assumption of normally distributed returns (Saksonova et al., 2019). Cryptocurrencies exhibit distinct characteristics and behaviors, rendering traditional methods ineffective. Instead, predictive analysis of future returns relies on time series data and historical returns (Liu and Tsyvinski, 2021). Portfolio optimization enables investors to attain maximum returns while maintaining a desired level of risk by strategically combining cryptocurrencies (Mazanec, 2021). Selecting the appropriate strategy for constructing an optimal portfolio in cryptocurrency investment is paramount. A well-chosen strategy can effectively manage risk, enhance investment performance, and fulfill various investor objectives. Moreover, leveraging technology advancements, including artificial intelligence tools and data analysis facilitates the selection, updating, and adjustment of investment strategies in response to market fluctuations.

This study investigates a model for managing investors' portfolios in the cryptocurrency market through the application of artificial intelligence and deep reinforcement learning techniques while considering market sentiment. Drawing inspiration from contemporary portfolio optimization methods, a novel strategy is proposed to address existing gaps in the literature. The objective of this research is to employ deep reinforcement learning methods to dynamically adjust to changing market conditions, optimize capital allocation, and enhance automated decision-making processes. The innovation presented in this study encompasses several enhancements across various dimensions. Firstly, a tailored learning environment has been devised within the cryptocurrency market, equipped with adaptability to evolving market conditions. Within this environment, status, reward functions, and reinforcement learning algorithms have been designed to facilitate improved management. Moreover, a novel approach has been introduced in the design of the deep reinforcement learning architecture, dynamically adjusting to accommodate any number of assets. Two decision-making approaches are introduced: local and global. In the local approach, each asset is individually analyzed, and decisions are made based on their specific characteristics. Conversely, the global approach considers the collective impact of all assets to select the optimal strategy for maximizing portfolio profitability. Another innovative aspect lies in the utilization of two distinct machine learning algorithms. The LSTM neural network, adept at deciphering intricate time patterns, is employed for price prediction, while the DQN neural network is leveraged for determining requisite asset operations. To enhance decision-making, the Fear and Greed Index has been used. This index provides more informed investment strategies. Furthermore, the inclusion of both fixed assets and non-fixed currencies in this study serves to bolster capital preservation and portfolio management in financial markets.

2. Related works

In the following, examines studies that explore the use of cryptocurrencies and diversification in investment portfolios, investigating their impact on profits and risks. In study (Ma et al., 2020), examines the benefits of including a diverse portfolio using cryptocurrency. Due to the rapid growth of cryptocurrencies, investors can increase the yield of the portfolio and reduce the risk by diversifying. Additionally, the results have shown that when comparing two diverse portfolios with the same risk, the portfolio containing cryptocurrencies is supposed to have higher efficiency, and the Sharpe ratio is also improved. In this study, better results are obtained by diversifying with different currencies compared to using only one asset such as Bitcoin. In study (Chauhan and Arora, 2019), the significance of portfolio management in the realm of cryptocurrencies is explored, highlighting its capacity to optimize returns, manage risk, facilitate informed decision-making, and attain long-term financial objectives. Portfolio management entails allocating investments across various assets, diversifying the portfolio, and making informed decisions to mitigate investment risks. Through systematic evaluation of performance and adaptability to market dynamics, portfolio managers can enhance returns and shield the portfolio from potential losses. In study (Nozari et al., 2022), the modern portfolio theory (MPT) approach is employed to assess the influence of cryptocurrencies on portfolio performance. The objective is to ascertain the optimal allocation of cryptocurrency within a portfolio to enhance its overall performance. Factors such as standard deviation, rate of return, and the Sharpe ratio were taken into account to evaluate portfolio performance. The findings indicated that integrating cryptocurrencies into a portfolio can enhance its performance by yielding higher returns while mitigating risk. Furthermore, the results underscored the significance of precise asset allocation and diversification as pivotal factors to enhance portfolio performance.

In the following, discusses studies that explore the impact and efficacy of deep reinforcement learning techniques like DQN in optimizing investment portfolios within volatile and unstable markets. In study (Muniappan et al., 2024), examine a dynamic and adaptive approach for optimizing investment portfolios in volatile markets using a multi-agent DQN deep reinforcement learning strategy. Evaluation criteria such as MSE and standard deviation are employed in this analysis. The results obtained indicate that reinforcement learning, owing to its capacity to uncover optimal decision-making strategies through interaction with dynamic environments, can generate suitable returns based on risk even amidst market volatility. In study (Liu, 2023) investigates the application of DQN for portfolio management, employing the Sharpe ratio as the reward metric. Traditional mean-variance optimization (MVO) portfolio allocation strategies and NAV portfolio allocation (NPA) serve as benchmarks for comparison. The findings indicate that the strategy leveraging the Sharpe ratio outperforms conventional portfolio allocation methods based on metrics such as annual returns. This underscores the effectiveness of incorporating DQN reinforcement learning techniques in portfolio management, particularly in optimizing risk-adjusted returns compared to traditional approaches. In study (Tran et al., 2023), the discussion revolves around an approach aimed at optimizing the strategy of an automated trading system through deep reinforcement learning, with a particular focus on cryptocurrencies. The algorithm endeavors to leverage DDQN and Bayesian optimization techniques to achieve consistent positive returns. The system's performance is assessed against two key objective functions: cumulative return and Sharpe ratio. The Relative Strength Index (RSI) is utilized as part of the trading rules.

The following examines studies related to the Fear and Greed Index to evaluate market sentiments in the crypto currency market. In study (nozari et al., 2023) examines the fear and greed index and the price movement of cryptocurrencies. Data from four cryptocurrencies indicate a U-shaped relationship between collective investor sentiment and price concurrency. The results suggest that cryptocurrency investors should carefully consider the impact of market sentiment on price changes and adjust their risk management strategies accordingly. Investors are advised to diversify their portfolios thoughtfully and remain vigilant about negative events affecting cryptocurrencies. In the study (Johnson, 2023) examines the usefulness of the fear and greed index in making decisions about investing in the crypto currency market, the results show that this index can measure past performance and current sentiment and provide an insight into the current state of the market, but its application Alone, it is not enough to decide regarding future investments. In study (Let et al., 2022) the use of the Fear and Greed Index was investigated in comparison with the buy-and-hold strategy in the cryptocurrency market. The study focused on 20 cryptocurrencies. The results indicate that the fear and greed strategy is generally more profitable. However, not all cryptocurrencies are homogeneous, and the performance of these strategies can vary significantly from one cryptocurrency to another.

The following reviews the studies that have employed reinforcement learning methods to enhance the performance of cryptocurrencies portfolios. In study (Lim et al., 2022), aimed to offer an alternative approach to maximizing portfolio returns by employing reinforcement learning (RL) to address dynamic risks tailored to market conditions through dynamic portfolio rebalancing. The study explored four distinct methods for the proposed RL factor, which encompassed complete and incremental portfolio rebalancing, both with and without price prediction models leveraging technical indicators. Experiments conducted as part of the study demonstrated that a well-optimized RL agent, whether with or without an LSTM price prediction model focusing on technical indicators, could effectively utilize dynamic rebalancing with risk adjustments to enhance portfolio returns. In study (Zolfani et al., 2022), a multi-criteria cryptocurrency portfolio allocation model was introduced, leveraging the Promethee II method alongside eight criteria and nine cryptocurrencies. The primary objective of the study is to assess the effectiveness of the Promethee II asymmetric method for allocating cryptocurrency portfolios. To mitigate the uncertainty inherent in the problem, regression forecasting techniques derived from ARIMA, LSTM, and Random Forest Regression (RFR) models were employed. Risk-related measures such as SlideVaR, along with value at risk (VaR) and conditional value at risk (C-VaR), were utilized. Additionally, an asymmetric preference function was proposed to address the behavioral phenomenon of asymmetry in gains and losses within the model. In study (Betancourt, Chen, 2021), a dynamic approach was proposed to address the challenge posed by the emergence of multiple cryptocurrencies and their integration into portfolio management using deep reinforcement learning techniques. The proposed architecture employs the actor-critic reinforcement learning method in conjunction with the GRU neural network. The primary objective of the study was to maximize efficiency while minimizing transaction costs, with a focus on the cryptocurrency market dataset. The research findings indicate the attainment of a favorable average daily yield, suggesting the efficacy of the proposed methodology in managing portfolios amidst the dynamic landscape of cryptocurrencies. In study (Gu et al., 2021) the utilization of deep reinforcement learning techniques to enhance portfolio management performance within the cryptocurrencies market is explored. This research introduces a framework for a deep reinforcement learning model alongside neural network architecture. The model incorporates various financial indicators, including final price, highest and lowest prices, network value-to-transaction volume ratio, market value-to-realized value ratio, investment return, and volatility. Performance evaluation metrics such as absolute portfolio value, Sharpe ratio, and maximum withdrawal are employed to assess the effectiveness of the proposed model. In study (Huang et al., 2021), a deep reinforcement learning algorithm featuring an adaptive sampling strategy for continuous portfolio optimization is introduced. Deep reinforcement learning is employed to forecast price trends within the market. The objective of the proposed algorithm is to enhance generalizability, manage risk, and mitigate high computational costs by discerning asset correlations and minimizing computational resource requirements. The research employs Sharpe ratio and profit criteria to evaluate portfolio performance. The findings highlight the efficacy of RL-based methods in risk management. In study (Lucarelli and Borrotti, 2020), deep reinforcement learning in dynamic portfolio optimization within the cryptocurrency market is examined. The presented framework comprises local agents for each asset in the portfolio, updating local information using deep learning Q, while a global agent oversees the reward function and administers weighting and actions for each asset based on information provided by the local agents. The global reward function combines the Sharpe ratio and the net return of the portfolio through a linear combination. This framework underwent testing using a portfolio containing four cryptocurrencies. In study (Jiang and Liang, 2017), a model-free convolutional neural network (CNN) has been devised for the management of cryptocurrency portfolios encompassing 12 highly traded assets. The CNN framework is tailored to directly compute portfolio weights leveraging historical price data. Training employs a reinforcement learning approach aimed at maximizing cumulative efficiency, which serves as the reward function. The training process involves selecting multiple sets of initial weight values to ascertain optimal parameters. The research findings demonstrate that the model yielded favorable returns while mitigating overall risk.

These studies underscored the significance of portfolio management within the realm of cryptocurrencies, highlighting its pivotal role in portfolio optimization, risk management, and informed decision-making. Moreover, diversification of portfolios using cryptocurrencies has been shown to potentially enhance returns while concurrently mitigating portfolio risk. In assessing portfolio performance, factors such as return rate and standard deviation play crucial roles.

3. Problem definition

This section provides a comprehensive overview of the portfolio and its associated mathematical relationships. Subsequently, it delineates the trading section followed by an elaboration on the reinforcement learning environment utilized in this study, along with its various components and mathematical correlations.

3.1 Portfolio

In the realm of stock markets, the units representing the specific amounts of assets held by an investor are commonly referred to as shares. A portfolio can be considered as a collection of various financial assets, represented as a vector. Mathematically, a portfolio is expressed as Eq. (1) where N denotes the number of assets, and the portfolio vector signifies the proportion of capital invested in each asset, in this representation, $\omega_{i,t}$ Represents the weight of asset i in the portfolio at time t, indicating the relative allocation of the investor's capital across the diverse assets in the portfolio.

$$\omega_t = [\omega_{1,t} \quad \omega_{2,t} \quad \cdots \quad \omega_{N,t}] \tag{1}$$

Subject to:

$$\sum_{i=1}^{M} \omega_{i,t} = 1$$

The dynamic nature of financial markets leads to fluctuations in supply and demand over time, consequently giving rise to a time series. The market's dynamic behavior can be analyzed and modeled through the application of various technical and mathematical tools. The price of each asset at time t is denoted as P_t. In a portfolio, it is possible to represent the prices of assets as a vector in Eq. (2) Certainly, where the price of each asset at time t is represented as $p_{i,t}$.

$$P_t = [p_{1,t} \quad p_{2,t} \quad \cdots \quad p_{N,t}] \tag{2}$$

The fluctuation of prices over time holds significance as it mirrors the profit or loss incurred by an investor. However, directly employing prices in financial calculations is not informative. Hence, efficiency measures are used. Portfolio return, a critical metric, signifies the profit or loss experienced over a specific period, typically expressed as a percentage. Investors rely on this metric to assess their portfolio's performance. Various methods exist for calculating portfolio return, with one common approach being the simple return. The formula in Eq. (3) captures the percentage change in price between two time periods, offering insights into the growth or decline of an investment.

$$r_t = \frac{p_t}{p_{t-1}} - 1 \tag{3}$$

Within the framework of Markowitz's portfolio optimization theory, investment portfolio return plays a crucial role. It serves as a key metric for gauging the overall profit or loss of a portfolio, calculated using in Eq. (4), where N denotes the number of assets in the portfolio, r_i represents the positive or negative return of the *i-th* asset, and ω_i represents the weight or proportion of the *i-th* asset in the portfolio.

$$\sum_{i=1}^{N} r_i \quad \omega_i \tag{4}$$

3.2 Trading section

The term "Trading Section" is used to denote specific periods of time, which are further divided into distinct intervals. Owing to the dynamic nature of the market, prices undergo constant changes, resulting in variations in the portfolio's value at the commencement and conclusion of each period. Within the investment portfolio, the weight of its components adjusts after each period based on the changes in returns. Figure 1 illustrates the market movement during period t, where the portfolio's value is denoted by the symbol p, and its weight is represented by the symbol w. The weight changes in each period due to fluctuations in prices. At the outset of each period, investors may engage in buying or selling actions, leading to a shift in the distributed weight of portfolio components.

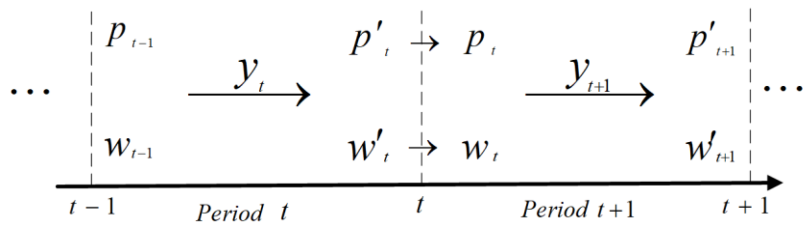

Figure 1: Portfolio adjustments over time intervals

3.3 Fear and Greed Index

The Fear and Greed Index, first introduced by CNNMoney, measures market sentiment based on the assumption that greed can drive prices higher than their worth, while fear can drive them lower. The index ranges from zero to 100, with values closer to zero indicating more fear and values closer to 100 indicating more greed. The sources used to measure this index include Bitcoin price in the last 30 and 90 days (25%), Bitcoin trading volume in the last 30 and 90 days (25%), social media activity about Bitcoin (15%), Bitcoin's share of the crypto currency market (10%), search trends related to Bitcoin (10%), and a survey of market participants (15%) (Johnson, 2023). This index is classified into five categories:

- 0 to 25: Extreme Fear
- 26 to 46: Fear
- 47 to 54: Neutral
- 55-75: Greed
- 76-100: Extreme Greed

3.4 Design reinforcement learning

Reinforcement learning is an artificial intelligence methodology wherein an agent acquires experience and rewards by interacting with the environment, subsequently enhancing its performance. The primary objective of reinforcement learning is to optimize the agent's performance to maximize cumulative rewards. In the following, the various components of the reinforcement learning model in the research are described, accompanied by their corresponding mathematical relationships.

3.4.1 Observation

Observation pertinent to the present state of the market involves using the prices of cryptocurrencies as input. In this study, analytical tools such as the simple price return outlined in Eq. (3).

3.4.2 Action

The action space denotes the set of actions that the agent can undertake. In this study, the action a_t for N assets contain the name of the chosen cryptocurrency denotes by C_i, the suggested weight of the action denotes by ω_i and the type of action that can be applied to each asset is indicated by the symbol a_i which encompasses three modes. The 'Buy' mode is used when the agent increases the weight of the Portfolio. The 'Hold' mode is applied when the agent does not change the weight, and the 'Sell' mode is employed when the agent reduces the weight of the portfolio. The mathematical relationships mentioned are expressed in Eq. (5).

$$a_t = \begin{bmatrix} C_1 & \omega_1 & a_1 \\ C_2 & \omega_2 & a_2 \\ \vdots & \vdots & \vdots \\ C_N & \omega_N & a_N \end{bmatrix} \tag{5}$$

3.4.3 State

A state in reinforcement learning refers to the current condition in which the agent is in it. State typically defined by a set of attributes that encapsulate the relevant information needed to make optimal decisions. In the Eq. (6), the state in this research is represented.

$$S_t = \begin{bmatrix} C_1 & \omega_{1,t} & RSI_{1,t} & ADX_{1,t} & Mom_{1,t} & HL_{1,t} & HO_{1,t} & LO_{1,t} & FG_t & r_{1,t} & \dots & r_{1,t-\tau} \\ C_2 & \omega_{2,t} & RSI_{2,t} & ADX_{2,t} & Mom_{2,t} & HL_{2,t} & HO_{2,t} & LO_{2,t} & FG_t & r_{2,t} & \dots & r_{2,t-\tau} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ C_N & \omega_{N,t} & RSI_{N,t} & ADX_{N,t} & Mom_{N,t} & HL_{N,t} & HO_{N,t} & LO_{N,t} & FG_t & r_{N,t} & \dots & r_{N,t-\tau} \end{bmatrix} \tag{6}$$

The state is defined as follows:

- N denotes the number of assets.
- τ denotes the length of the lookback window.
- C_i denotes the name of Assets.
- $\omega_{i,t}$ denotes the weights invested in asset i at time t.
- $RSI_{i,t}$ denotes the relative strength index of asset i at time t.
- $ADX_{i,t}$ denotes the average directional movement index of asset i at time t.
- $Mom_{i,t}$ denotes the momentum indicator of asset i at time t.
- $HL_{i,t}$ denotes the high and low price difference in a period t for asset i.
- $HO_{i,t}$ denotes the high and open price difference in a period t for asset i.
- $LO_{i,t}$ denotes the low and open price difference in a period t for asset i.
- FG_t denotes the crypto market Fear and Greed Index at time t.
- $r_{i,t}$ denotes the return of asset i at time t.

The difference between the maximum and minimum price within a specific period, known as the candlestick range, indicates the degree of price volatility during that interval. A larger range signifies greater market volatility or instability, which could signal the end of a trend and the beginning of a new one, thus serving as a potential buy or sell signal. The difference between the highest price and the opening price reflects buying pressure in the market; a larger value suggests prevalent positive sentiment and optimism. Similarly, the difference between the lowest price and the opening price indicates selling pressure; a larger value suggests dominant negative sentiment and market pessimism. Since the prices change in each period t, change in prices affects the weight of the portfolio. An increase in price causes an increase in weight and vice versa. Eq. (7) is used to calculate the updated weights.

$$W_{i,t} = W_{i,t-1}(1 + r_t) = W_{i,t-1}\left(1 + \frac{p_{i,t}}{p_{i,t-1}}\right) \tag{7}$$

3.4.4 Reward

The appropriate design of the reward relationship stands as one of the pivotal challenges and significant aspects in reinforcement learning. This is crucial as it profoundly influences the performance and learning capabilities of the agent. In reward calculation, the cumulative reward relationship, denoted by the symbol G, is employed to enable the agent to consider a history of decisions and rewards, thereby leveraging past experiences for future decisions. In Eq. ()**8**), the variable R signifies the reward, which can be derived from the Sample Return expressed in Eq. (3). The parameter γ, representing uncertainty, ranges between zero and one, signifying the significance of past rewards relative to immediate and future rewards. This parameter encapsulates the balance between historical experiences and the importance attributed to forthcoming outcomes in the learning process.

$$G_t = R_t - \gamma G_{t-1} = \sum_{k=0}^{\infty} \gamma^k R_{t-k} \quad (8)$$

4. Proposed system

In the proposed system, the RL framework is utilized to implement strategy within the cryptocurrency market. Within the RL process, the agent executes actions based on the current state and sends them to the environment. The environment, in turn, responds to these actions, resulting in dynamic changes that include updates to the state, rewards, and checks for final states corresponding to the action taken. Throughout the initial time until reaching the final state, the agent continuously executes a series of actions according to a policy, with data received from the environment contributing to the enhancement of its performance. The asset portfolio comprises N assets, where $N-1$ represent non-stable currencies and one represents a stablecoin like Tether, treated as fixed assets. Initially, the value of stablecoin equals the initial value, while the values of other assets are set to zero. As the algorithm progresses, based on the relationships defined within the model, the values of fixed assets are converted to variable assets and vice versa.

Figure 2 illustrates a visual depiction of the deep reinforcement learning portfolio management framework. This framework employs a multi-agent approach, wherein two parallel distinct deep neural networks are deployed for each currency. Specifically, the DQN neural network is utilized for training and action selection, while the LSTM neural network is employed to predict next state return value. Each local agent, implemented separately for every cryptocurrency, undergoes training using the DQN deep reinforcement learning neural network to effectively manage the asset portfolio by making informed decisions. To train the deep reinforcement learning algorithm, historical data is required. Therefore, a data warehouse is used to store raw data, which is then preprocessed, cleaned, normalized, and categorized. This data includes the previous state, action, new state, and the reward resulting from the action taken. These data help the model recognize market patterns and improve the accuracy of its predictions and decision-making.

The input of the DQN neural network consists of 9 variables derived from the latest state. These variables include the most recent return value, risk calculated from return values, Fear and Greed Index, RSI, ADX, Momentum, HL, and HO, as expressed in Eq. (9). The output of the network is the action variable, which has three possible states: buy, sell, and hold.

$$DQN\ State_{i,t} = [Return_{i,t} \quad Risk_{i,t} \quad RSI_{i,t} \quad ADX_{i,t} \quad MOM_{i,t} \quad HL_{i,t} \quad HO_{i,t} \quad LO_{i,t} \quad FG_t] \quad \textbf{(9)}$$

Furthermore, each agent is equipped with an LSTM neural network. Its input comprises return values, which encapsulate information regarding the asset's return history and is expressed in the Eq.)10). This data aids the neural network in recognizing diverse patterns and facilitating a more precise forecast for future performance. The output of this neural network entails the prediction of return values for the subsequent state of each asset. The primary objective of this neural network is to assist the portfolio management system in selecting the appropriate action based on assets prioritization.

$$LSTM\ Input_{i,t} = [r_{i,t} \quad r_{i,t-1} \quad \cdots \quad r_{i,t-\tau}] \tag{10}$$

Portfolio management is the action of continuous reallocation of a capital among various financial assets. This endeavor encompasses not only the initial selection of assets but also the continuous monitoring, evaluation, and adjustment of the portfolio to align with evolving market conditions, investment objectives, and risk levels. Algorithm 1 outlines the portfolio management framework utilizing reinforcement learning. A ϵ-greedy algorithm is employed to select actions from the three modes of buy, sell, and hold. The outputs of the DQN and LSTM neural networks are combined, and the optimal portfolio weights are optimized using Markowitz-based relationships in Eq.)11). The sum of the weights for all assets, both stable and non-stable, equals one. Therefore, the total weight of all non-stable assets must be less than or equal to one. v represents the defined permissible limit for portfolio adjustment. R_P denotes the return predicted by the LSTM network, while R_E represents the expected return of the portfolio.

$$max \sum_{i=1}^{N-1} \omega_{i,t}\, r_{i,t} \tag{11}$$

Subject to:

$\omega_{i,t} \geq 0$ *i=1, 2, …, N-1*

$$\sum_{i=1}^{N-1} \omega_{i,t} \leq 1$$

$R\_E_i \geq R\_P_i$ *i=1, 2, …, N-1*

$$\omega_{i,t} = \begin{cases} \omega_{i,t} > \omega_{i,t-1} \ and \ \omega_{i,t} \leq \omega_{i,t-1} + v & buy \\ \omega_{i,t} < \omega_{i,t-1} \ and \ \omega_{i,t} \geq \omega_{i,t-1} - v & sell \\ \omega_{i,t-1} & hold \end{cases}$$

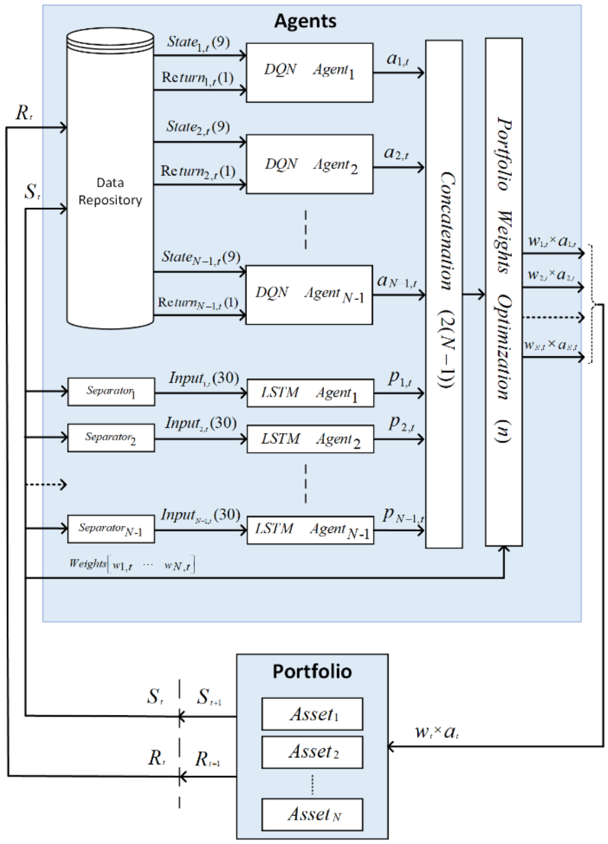

Figure 2: Deep Q-learning portfolio management framework

Employing stablecoin in portfolio management can serve as a strategy to preserve capital value against market fluctuations. Consequently, the weight of stable coin in the model equals the weight of the initial asset minus the total weight of variable non-stable assets, provided it does not descend below zero. Following this, normalization is applied to the final weights of all aggregated assets. The action is transmitted to the deep reinforcement learning network, and the subsequent state, along with the reward, is received. Leveraging the outcomes from the executed action, a descending gradient on the output of DQN is performed.

Algorithm 1: Deep Q-learning Portfolio Management Framework

1: **Inputs:** N assets, S_t States, R Returns for τ sliding windows
σ_e, exploration parameter
β, Minimum quorum for buying or selling
υ, Maximum quorum of weight change for each asset in portfolio
ϵ , Random action probability
2: **Initialize:** θ_0 Policy parameters, ω_0 initial allocation weights vector, D data repository
3: **for** episode =1 to P **do**
4: Global profit $G = 0$
5: **for** t=1 to T **do**
6: **for** n=1 to N-1 **do**
7: Calculate predict return based on LSTM and store in $P_{n,t}$
8: Following ϵ-greedy policy, select $a_{n,t}$ in A= {buy, hold, sell}:

$$a_{n,t} = \begin{cases} random & with\ probility\ \epsilon \\ arg\ max_a\ Q(s_{nt}, a_n; \theta_n) & otherwise \end{cases}$$

Revision of measures according to the integration of methods

$$a_{n,t} = \begin{cases} buy & a_{n,t} = buy\ and\ P_{n,t} > \beta \\ sell & a_{n,t} = sell\ and\ P_{n,t} < -\beta \\ hold & otherwise \end{cases}$$

Optimize portfolio weights based on Eq. 11

end for

Calculating the weight of stable coin:

$w_{N,t} = \max\,(\omega_0 - \sum_{i=1}^{N-1} w_{i,t}\ , 0)$

Normalize weights

Generate action $a_{n,t}$ consist of c_n asset name, $m_{n,t}$, $\omega_{n,t}$

Execute action a_{nt} and observe reward $R_{n,t+1}$ and state $s_{n,t+1}$

Store transaction $(S_{n,t}, S_{n,t+1}, a_{n,t}, R_{n,t+1})$ in D

Set $S_t = S_{t+1}$ and $R_t = R_{t+1}$

Accumulate profits $G_t = \sum_{i=1}^{N} R_{i,t}$

Sample random minibatch of transactions from D

for i=1 to N-1 **do**

$$y_{i,t} = \begin{cases} R_{i,t} & terminal\ S_{i,t} \\ R_{i,t} + \sigma_e \times max_a\ Q(s'_{i,t+1}, a'_i; \theta_n) & non-terminal\ S_{i,t} \end{cases}$$

Perform a gradient descent step on $(y_{i,t} - Q(s_{i,t}, a_{i,t;}\theta_n))^2$

end for

end for

end for

5. Experimental settings

In this section, the data utilized in the study, including cryptocurrencies, the historical training and testing periods, as well as the configurations of the neural network algorithm and reinforcement learning, are delineated.

5.1 Cryptocurrencies data

In this study, data from 2020 to 2024 and 11 cryptocurrencies were used in daily time frame. The website Alternative.me provides a daily Fear and Greed Index, which gauges the overall sentiment of the cryptocurrency market. In this research, we leverage their API to retrieve these index values and store them within a database.

Table 1 illustrates the historical period of data usage for network training and testing the proposed model. Data is collected as a scheduled task through the Coinmarketcap website API and stored in a MySQL database. Coinmarketcap is a leading cryptocurrency market data platform that plays a crucial role in providing real-time and historical data on various cryptocurrencies. Established in 2013, it has become a go-to resource for investors, traders, and enthusiasts seeking comprehensive information about the dynamic cryptocurrency market. Coinmarketcap provides essential metrics such as prices, trading volumes, market capitalization, and price charts for thousands of cryptocurrencies. The cryptocurrencies examined in this investigation comprise BTC, ETH, BNB, SOL, XRP, DOGE, ADA, TRX, LINK, BCH, LTC, and the stable coin of Tether.

5.2 Fear and Greed Index

The website Alternative.me provides a daily Fear and Greed Index, which gauges the overall sentiment of the cryptocurrency market. In this research, we leverage their API to retrieve these index values and store them within a database.

Table 1: Historical scope of the training and testing period

Milestones	Date	Status
Start	01 May 2020	Train
Finish	31 July 2023	
Start	01 August 2023	Test
Finish	31 May 2024	

5.3 Algorithm configuration

The mean squared error evaluation criterion is employed for neural networks with a learning rate of 0.001 and 100 training epochs in the configuration of LSTM and DQN neural networks. In reinforcement learning, the variable σ_e, or exploration parameter, is set to 0.9 during the training phase and to 0.1 during testing phase. The τ sliding windows is 30 days, and the transaction fee is equivalent to spot transactions at 0.1%.

6. Results

The covariance chart, presented in Figure 3, illustrates the degree of correlation between each cryptocurrency in the investment portfolio during the test period of this study. Investors can utilize this information to construct an optimal stock portfolio. The data indicate that BCH, SOL, and DOGE exhibited the highest variance. Additionally, cryptocurrencies with high covariance, such as SOL and LINK, tend to behave similarly during market fluctuations.

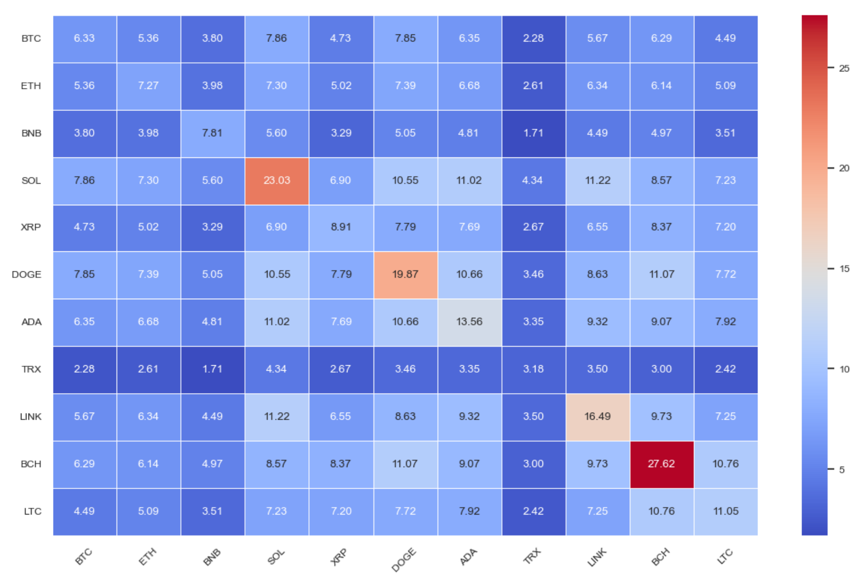

Figure 3: Covariance matrix of 11 non-stable cryptocurrencies during test period

Table 2 classifies the Fear and Greed Index into five distinct classes, detailing the number of samples in each class during the training and testing periods of the model. In the training period, most samples are concentrated in the Greed and Fear classes, indicating that these two classes are predominant. During the testing period, the Greed class has the highest number of samples compared to the other classes.

Table 2: Classification of Fear and Greed Index in two training and test period

Status	Extreme Greed	Greed	Neutral	Fear	Extreme Fear
Train	164	257	175	309	282
Test	45	165	55	40	0

Figure 4 illustrates the Fear and Greed Index during the test period. This index is designed to measure the prevailing market sentiments. During the test period, the minimum value was 30, the maximum value was 90, and the average value was 63.5, falling into the Greed class. The standard deviation of the values was 12.9.

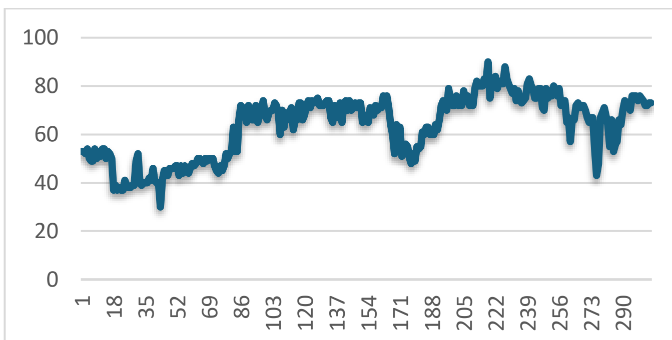

Figure 4: Fear and Greed Index during test period

Figure 5 illustrates the total weights of the investment portfolio for 11 non-stable cryptocurrencies during the test period. The average portfolio weight during this period was 0.47. The results indicate that the proposed model effectively adapted to varying market conditions. It maximized potential profits by increasing portfolio weights and mitigated potential losses by reducing portfolio weights, thereby capitalizing on available market opportunities.

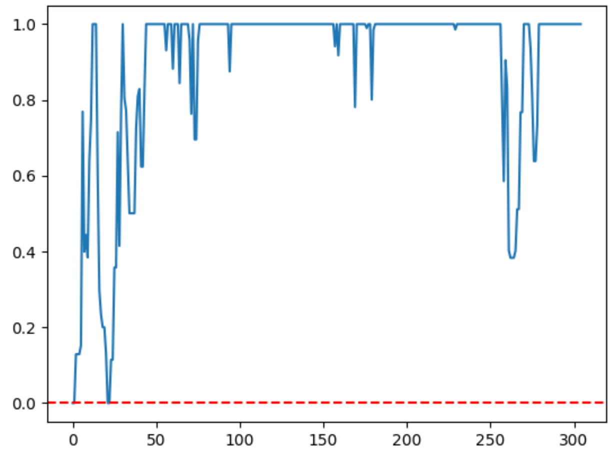

Figure 5: Sum of the weights of non-stable cryptocurrencies in the model`s portfolio during the test period

Figure 6 illustrates the changes in the weight of the investment portfolio for various cryptocurrencies, along with the stable currency represented by USD. Initially, the weight of USD is set to one, while all other assets have a weight of zero. Over time, the weight of all asset's changes in response to market conditions. In bullish market conditions, the model aims to invest more heavily in profitable cryptocurrencies. Conversely, in bearish conditions, the model focuses on preserving capital by shifting assets into the stable currency. The highest weights have been allocated to TRX, BNB, Link, and Doge, in that order. The average weight assigned to each asset is 8.15%.

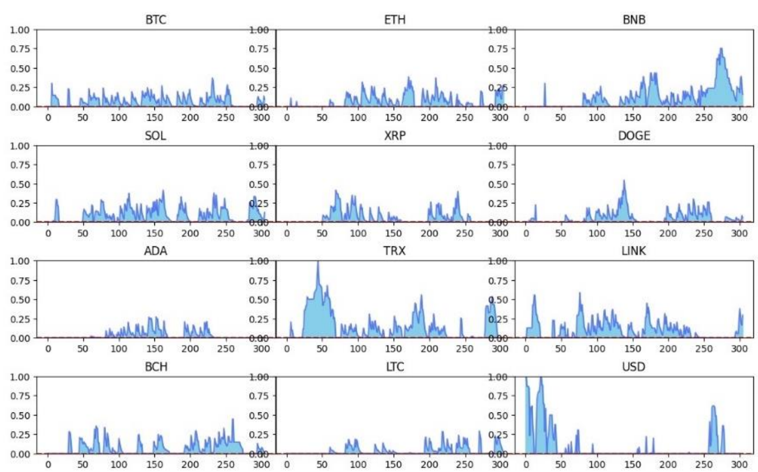

Figure 6: Portfolio weights changes for each currency in the proposed model

Figure 7 illustrates the cumulative returns of 11 different cryptocurrencies over the test period. The results indicate that the highest return was achieved by the cryptocurrency SOL, with a return of 33.86%. On the other hand, the largest loss was experienced by the cryptocurrency LINK, with a return of -7.4%. This loss was primarily due to negative price fluctuations at the beginning of the period. However, the model was able to recover from this initial loss and subsequently turn it into profitability.

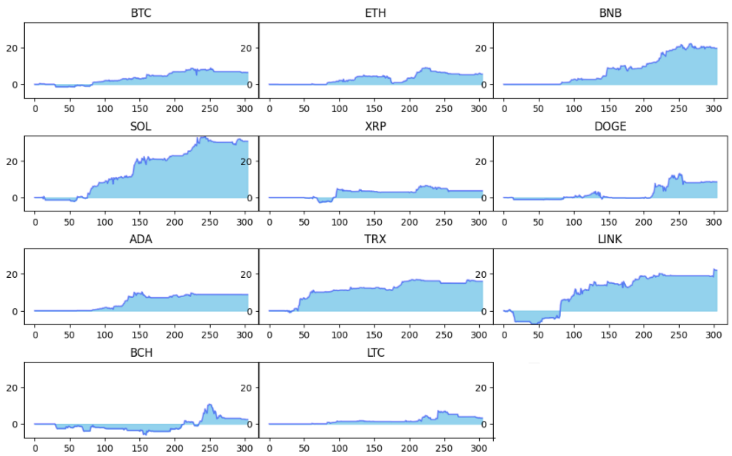

Figure 7: Accumulated profit in the proposed model for each non stablecoin

To evaluate the performance of the proposed model, two distinct strategies are compared. The buy-and-hold strategy is a straightforward and commonly used approach in financial markets. In this method, assets are purchased at the beginning of the evaluation period and held until the end of the period. The other strategy is the proposed model. Figure 8 illustrates the comparison of cumulative returns between these two strategies during the test period. The results show that the maximum return achieved with the proposed model is 146.2, compared to 84.4 for the buy-and-hold strategy. Additionally, the average return for the proposed model is 63.4, whereas it is 27.3 for the buy-and-hold strategy.

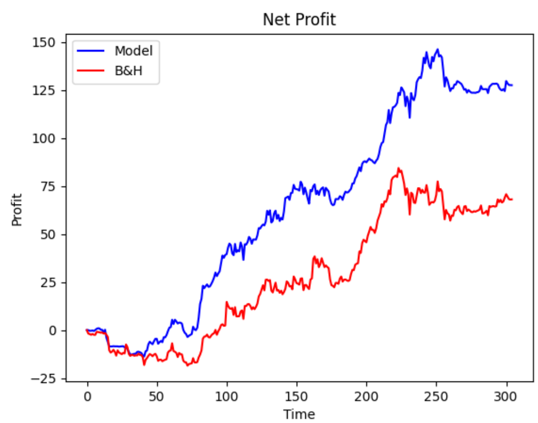

Figure 8: Portfolio cumulative profit for two strategies proposed model and B&H

7. Conclusions and discussion

This study proposes a cryptocurrency portfolio optimization framework using deep reinforcement learning. In this framework, LSTM is used to predict the returns of non-stable cryptocurrencies, while DQN is employed to select actions among buy, sell, or hold based on prevailing conditions. The findings from the study (Johnson, 2023) indicate that although the Fear and Greed Index has various advantages, it is not sufficient on its own for future investment decisions. Therefore, the proposed model incorporates the Fear and Greed Index alongside other financial indicators to consider market sentiment comprehensively such as RSI, ADX, Momentum, candlestick range (the difference between the highest and lowest prices), buying power (the difference between the highest price and the opening price), selling power (the difference between the lowest price and the opening price), the latest return value, and risk.

The covariance chart results show that cryptocurrencies exhibit different behaviors under varying market conditions. This diversity in the performance of different cryptocurrencies underscores the importance of managing and adjusting the investment portfolio based on changing market conditions. The results of this research complement the studies (Wang et al., 2024; Johnson, 2023; Let et al., 2022), demonstrating that the Fear and Greed Index is a crucial tool for analyzing market psychology. By identifying prevalent sentiments, it helps investors make more informed decisions.

Comparing the results of the non-stable cryptocurrencies portfolio weights chart with the Fear and Greed Index chart reveals that during periods of fear, the dominant market trend is bearish. Consequently, the model attempts to reduce the weight of non-stable cryptocurrencies and increase the weight of stable coin to prevent capital loss. Conversely, during periods of greed, the dominant market trend is bullish, and the model seeks to increase the weight of non-stable cryptocurrencies to maximize profits. Furthermore, the results comparing the cumulative profits of the two evaluated strategies show that the proposed model achieved an average growth of 132% more than the buy-and-hold strategy, and the Sharpe ratio of the proposed model is 45% higher than the other method. These findings demonstrate the model's ability to adapt to changing market conditions, manage the portfolio, and optimize the investment portfolio in non-stable cryptocurrency markets to maximize returns and reduce risk.

8. Future Works

In this study, the use of market psychology indicators such as the Fear and Greed Index and technical indicators like buying and selling power, relative strength, trend strength, and momentum was examined. Future research could incorporate other technical indicators such as volume indicators, moving averages, and volatility indices. Additionally, other methods of market psychology that analyze investor sentiments, behaviors, and attitudes could be utilized, such as social media analysis, financial and political news analysis, and search trend analysis.

The present study employs deep reinforcement learning in conjunction with LSTM for time series data analysis. However, other reinforcement learning methods such as Asynchronous Advantage Actor-Critic (A3C), Policy Gradient, Deep Deterministic Policy Gradient (DDPG), and Distributed Reinforcement Learning could also be used. Additionally, various types of neural networks such as Convolutional Neural Networks (CNNs), Graph Neural Networks (GNNs), and Quantum Neural Networks (QNNs) could be utilized. Comparing their results in terms of efficiency could provide valuable insights.

References

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.

Chauhan, V., & Arora, G. (2019). A review paper on cryptocurrency & portfolio management. In *2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC),* 60-62.

Gu, F., Jiang, Z., & Su, J. (2021). Application of Features and Neural Network to Enhance the Performance of Deep Reinforcement Learning in Portfolio Management. In *2021 IEEE 6th International Conference on Big Data Analytics (ICBDA)*, 92-97.

Huang, S. H., Miao, Y. H., & Hsiao, Y. T. (2021). Novel deep reinforcement algorithm with adaptive sampling strategy for continuous portfolio optimization. *IEEE Access*, *9*, 77371-77385.

Jacob, N. L. (1974). A limited-diversification portfolio selection model for the small investor. *The Journal of finance*, *29*(3), 847-856.

Jiang, Z., & Liang, J. (2017). Cryptocurrency portfolio management with deep reinforcement learning. In *2017 Intelligent systems conference (IntelliSys),* 905-913.

Johnson, J. (2023). How useful is the crypto Fear and Greed Index when making cryptocurrency investment decisions. *Available at SSRN 4407405*.

Lęt, B., Sobański, K., Świder, W., & Włosik, K. (2022). Investor Sentiment and Efficiency of the Cryptocurrency Market: The Case of the Crypto Fear & Greed Index. In *Eurasia Business and Economics Society Conference* (pp. 271-287). Cham: Springer Nature Switzerland.

Lim, Q. Y. E., Cao, Q., & Quek, C. (2022). Dynamic portfolio rebalancing through reinforcement learning. *Neural Computing and Applications*, *34*(9), 7125-7139.

Liu, Y., & Tsyvinski, A. (2021). Risks and returns of cryptocurrency. *The Review of Financial Studies*, *34*(6), 2689-2727.

Liu, Z. (2023). Reinforcement Learning in Portfolio Management with Sharpe Ratio Rewarding Based Framework. In *Proceedings of the 4th International Conference on Economic Management and Model Engineering, ICEMME 2022*, Nanjing, China.

Lucarelli, G., & Borrotti, M. (2020). A deep Q-learning portfolio management framework for the cryptocurrency market. *Neural Computing and Applications*, *32*, 17229-17244.

Ma, Y., Ahmad, F., Liu, M., & Wang, Z. (2020). Portfolio optimization in the era of digital financialization using cryptocurrencies. *Technological forecasting and social change*, 161, 120265.

Mazanec, J. (2021). Portfolio optimalization on digital currency market. *Journal of Risk and Financial Management*, *14*(4), 160.

Muniappan, A., Dalton, G. A., Prithiviraj, A., Jackulin, T., & Karthikayen, A. (2024). Intelligent systems and applications in engineering. *IJISAE*, 12(13), 386-395.

Nozari, H., & Nahr, J. G. (2022). The impact of blockchain technology and the internet of things on the agile and sustainable supply chain. International Journal of Innovation in Engineering, 2(2), 33-41.

Nozari, H., Tavakkoli-Moghaddam, R., Ghahremani-Nahr, J., & Najafi, E. (2023). A conceptual framework for Artificial Intelligence of Medical Things (AIoMT). In Computational Intelligence for Medical Internet of Things (MIoT) Applications (pp. 175-189). Academic Press.

Saksonova, S., & Kuzmina-Merlino, I. (2019). Cryptocurrency as an investment instrument in a modern financial market. *St Petersburg University Journal of Economic Studies*, *35*(2), 269-282.

Tran, M., Pham-Hi, D., & Bui, M. (2023). Optimizing Automated Trading Systems with Deep Reinforcement Learning. *Algorithms*, *16*(1), 23.

Xiang, X., & Foo, S. (2021). Recent advances in deep reinforcement learning applications for solving partially observable markov decision processes (pomdp) problems: Part 1—fundamentals and applications in games, robotics and natural language processing. *Machine Learning and Knowledge Extraction*, *3*(3), 554-581.

Zolfani, S. H., Taheri, H. M., Gharehgozlou, M., & Farahani, A. (2022). An asymmetric PROMETHEE II for cryptocurrency portfolio allocation based on return prediction. *Applied Soft Computing*, 131, 109829.