Mozaic

Beyond Traditional Models: Incorporating GMX Pools and Machine Learning for Superior Portfolio Performance

Abstract

This research investigates how machine learning can enhance cryptocurrency portfolio diversification and risk-adjusted returns by incorporating GMX liquidity pools that allow for additional yield. We explore advanced portfolio optimization techniques alongside machine learning models for predicting returns and yields from GMX pools in constructing portfolios with various risk profiles.

To account for the unique characteristics of these pools, we leverage feature selection to identify the most impactful factors influencing their yield and price. Our machine learning models integrate predictions from multiple algorithms, leading to significant performance improvements in the optimized portfolio compared to a benchmark. This is evidenced by the portfolio's higher cumulative return and superior Sharpe Ratio. Our simulations demonstrate that this data-driven approach, focused on relevant features, significantly improves portfolio performance compared to a benchmark. The optimized portfolio achieves both a higher cumulative return and a superior Sharpe Ratio, highlighting the benefits of integrating GMX pools and applying machine learning into cryptocurrency portfolio construction. Overall, this paper proposes a novel strategy for navigating the dynamic cryptocurrency market. It demonstrates the value of incorporating GMX pools and machine learning, particularly with a focus on relevant feature selection, for constructing robust cryptocurrency portfolios.

Keywords: GMX Pool; Cryptocurrency; Machine Learning; Gradient Boosting; Feature Selection; Portfolio optimization; Sharpe ratio

1. Introduction

Cryptocurrencies have upended finance, luring investors with high return potential despite their wild price swings. Traditional portfolio optimization methods struggle with crypto's unique characteristics, but machine learning offers a powerful solution. By analyzing vast amounts of data, machine learning models can predict returns and volatility, enabling investors to improve forecasts, estimate volatility more accurately, and develop dynamic asset allocation strategies for real-time risk-return optimization. However, machine learning's effectiveness hinges on data quality and algorithm selection. Further complexity arises with Decentralized Finance (DeFi) protocols like GMX pools. These liquidity pools offer passive income through yield farming, but integrating them into a portfolio requires careful consideration of their risk-return profiles and potential drawbacks like impermanent loss.

Building upon the foundational work of Markowitz (1952) on Modern Portfolio Theory (MPT), recent studies have explored the applicability of these principles in the context of cryptocurrencies. Works by Holovatiuk (2020) demonstrate the potential of cryptocurrencies to enhance portfolio diversification and improve the risk-return trade-off. Sahu et al. (2024) further emphasize the role of cryptocurrencies in portfolio optimization, highlighting their ability to improve portfolio efficiency through decorrelation with traditional assets. However, the unique characteristics of cryptocurrencies necessitate specialized approaches to portfolio optimization that go beyond the traditional MPT framework. Studies by Lorenzo and Arroyo (2023) explore these complexities and propose strategies to navigate the high volatility, serial dependence, and potential for extreme events inherent in cryptocurrency markets. Furthermore, Letho et al. (2022) examine diversification opportunities within emerging cryptocurrency markets.

In this context, machine learning (ML) techniques are increasingly being explored to understand the complexities of cryptocurrency markets and predict both returns and volatility (Amirzadeh et al. 2022, Tang et al. 2022, Wang et al. 2021). Studies by Poudel et al. (2023) demonstrate the effectiveness of various machine learning models like LSTM for price prediction of Bitcoin and Dogecoin and also the use of neural network algorithms to predict volatility. Similarly, Chen et al. (2020) explore the use of Random Forest algorithms for predicting Bitcoin volatility. However, a one-size-fits-all approach might not be ideal. Khan et al. (2023) highlight the need for cryptocurrency-specific models. Their findings suggest superior performance with smoothing techniques when predicting Bitcoin volatility compared to a neural network autoregressive model that outperforms for Ethereum. This aligns with the ongoing exploration of various ML techniques exemplified by Sharifi et al. (2022) focusing on support vector machines. These studies underscore the potential of ML to improve cryptocurrency market prediction, acknowledging the need for tailored models for cryptocurrency portfolios.

This paper explores the power of ML in portfolio optimization specifically designed for the cryptocurrency market. Our focus is on building efficient frontiers that strike the optimal balance between risk and return for investors. We delve into three prominent models: Black-Litterman, Hierarchical Risk Parity (HRP), and Mean-Variance Optimization. To account for the unique characteristics of GMX pools, we integrate predictions from multiple time series and machine learning models. This allows us to capture the unique characteristics of GMX liquidity pools. We leverage machine learning to identify the most impactful factors influencing yield and pool returns. This information is then fed into a diverse set of machine learning algorithms to find the best model for predicting returns and volatility. These predicted values are then fed into the portfolio optimization algorithms, as illustrated in Figure 1.

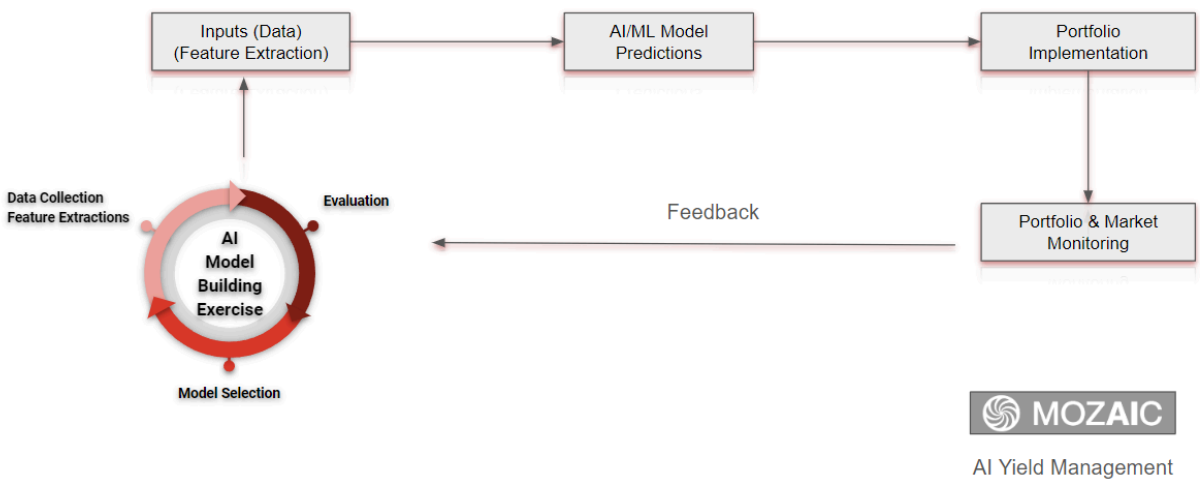

Figure 1 - AI Yield Management

By harnessing the power of machine learning for both prediction and optimization, these state-of-the-art techniques empower investors to make informed decisions. They can construct robust portfolios within the dynamic and evolving cryptocurrency market, aligning perfectly with their individual risk tolerance and return objectives.

The results are compelling. Our optimized portfolio, constructed using predicted returns and volatility from the machine learning models, achieves a daily return of 0.5%, surpassing the benchmark Bitcoin (BTC) return of 0.34%. Additionally, the portfolio boasts a superior Sharpe ratio, indicating a more attractive risk-adjusted return profile compared to simply holding BTC.

2. Yield Optimization Through AI

2.1 Data and Features

Our machine-learning models are trained on the data collected from various sources, including but not limited to:

- Cryptocurrency Prices: Daily closing prices are obtained from a reliable source for a significant historical period to capture market trends and fluctuations. We focus on a defined set of cryptocurrencies that form major liquidity pools on GMX.
- GMX Pool Data: Daily historical data, including pool prices and APY/yield information, is collected directly from the GMX platform.
- Additional Factors: Historical data on relevant DeFi protocol metrics (e.g., Total Value Locked, trading fees) is gathered from unspecified sources.
- Technical indicators are calculated based on historical GMX pool price data to capture both short-term and long-term price movements.
- Macroeconomic Factors: The overnight Fed Funds Rate, reflecting U.S. monetary policy and risk appetite in financial markets, is included to assess its potential impact on cryptocurrency and DeFi activity. Oil and other commodities, considered a major global economic influence, are also incorporated to understand their potential effect on cryptocurrency and DeFi investor sentiment. Furthermore, we gather historical data on a broader set of traditional factors that may influence GMX pool returns, yields, and volatilities.

2.2 ML forecasting algorithms

Training the Model for GMX Pool Prediction: To predict the yield and volatility of GMX pools, which are crucial for optimizing investment portfolios, machine learning algorithms are employed. These algorithms first identify the most important factors that influence pool performance. Then, they are trained on historical data separated into training, validation, and testing sets. The training set, spanning October 2023 to January 2024, provides the data for the algorithms to learn patterns. The testing set, February to March 2024, is used to evaluate the model's accuracy on unseen data.

Model Selection and Evaluation: A moving window approach is used, where the model is trained on a specific timeframe (weeks or months) and then predicts the following timeframe. This process is repeated by continually moving the window forward. To ensure the model's generalizability and avoid overfitting, a technique called multifold cross-validation is employed. Additionally, various machine learning algorithms, such as Gradient Boosting, LGBM, Lasso, and Linear regression, are explored. The final model selection is based on how well it performs on the testing set, measured by metrics like Root Mean Square Error (RMSE), Mean Square Error (MSE), and Mean Absolute Percentage Error (MAP APE), which evaluate the difference between predicted and actual values.

2.3 Portfolio Optimization

This research explores three advanced portfolio optimization models that move beyond the traditional mean-variance approach, namely, Black-Litterman, Hierarchical Risk Parity (HRP), and Mean-Variance with GARCH Models**.** The portfolio problem is defined as follows. If w is the weight vector of stocks with expected returns μ, then the portfolio return is equal to each stock's weight multiplied by its return, i.e $w^T\mu$. The portfolio risk in terms of the covariance matrix Σ is given by $w^T\Sigma w$. Portfolio optimization can then be regarded as a convex optimization problem, and a solution can be found using quadratic programming. If we denote the target return as μ^*, the precise statement of the long-only portfolio optimization problem is as follows:

$$
\begin{aligned}
&(w^T\Sigma w) \qquad (1)\\
s.t.\quad & w^T\mu \geq \mu^*\\
& w^T 1 = 1\\
& w_i \geq 0
\end{aligned}
$$

If we vary the target return, we will get a different set of weights (i.e a different portfolio) – the set of all these optimal portfolios is referred to as the efficient frontier.

The performance of each portfolio optimization model is evaluated using metrics such as Sharpe Ratio, Sortino Ratio, Maximum Drawdown, Calmar Ratio to compare their effectiveness in generating risk-adjusted returns for portfolios with and without GMX pools. The optimized portfolios are backtested on historical data to assess their performance under real-world market conditions. This will involve simulating portfolio behavior based on past price and yield data and evaluating achieved returns and risk metrics.

In a nutshell, we follow the procedure as shown in Figure 1 to perform our portfolio optimization. Various machine learning algorithms (e.g., Random Forests, LSTMs) are employed to predict future returns and volatilities for each GMX pool, leveraging both historical data, technical indicator values, and macroeconomic factors. Technical indicators capture potential trends and sentiment in the cryptocurrency market, while macroeconomic factors can provide insights into broader economic conditions that may influence investor behavior. The predicted returns, yields, and volatilities from the machine learning models are integrated into the portfolio optimization framework (Black-Litterman, HRP, Mean-Variance with GARCH). This approach allows the models to capture non-linear relationships and potentially improve the accuracy of risk and return estimates for GMX pools.

3. Results

Feature Selection and Model Evaluation: The first step involves identifying the key factors that influence the yield generated by a liquidity pool over a specific period. We then employ machine learning algorithms to predict this yield. To select the best-performing models, we evaluate their performance using metrics like Mean Square Error (MSE). For instance, as shown in Table 1 (not included here but assumed to present various models and their MSE), Gradient Boosting Regressor (GBR) emerges as the best model with the lowest MSE.

	Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
gbr	Gradient Boosting Regressor	9.0839	163.3279	12.7365	0.2155	0.6719	1.5599	5.3167
lightgbm	Light Gradient Boosting Machine	10.3807	206.7613	14.3408	-0.0817	0.7433	1.6515	0.3867
dt	Decision Tree Regressor	10.4390	201.5329	14.1806	-0.0937	0.7579	2.0252	5.7067
rf	Random Forest Regressor	10.3830	199.6089	14.0564	-0.1035	0.7516	1.9241	7.3933
ada	AdaBoost Regressor	10.8749	219.0463	14.6847	-0.1716	0.7886	1.9467	5.6100
xgboost	Extreme Gradient Boosting	10.7274	211.6512	14.3619	-0.2604	0.7714	2.0471	0.6700
lasso	Lasso Regression	12.8397	5331.6909	50.6127	-18.6038	0.8319	2.6455	0.0600
en	Elastic Net	13.2860	6387.9708	54.8236	-22.4407	0.8345	2.6071	0.0567
lar	Least Angle Regression	22.9211	13101.4262	78.3360	-39.5313	0.9564	2.1775	0.1033
lr	Linear Regression	26.2947	43706.4180	174.2071	-145.2548	1.0008	2.7951	0.1067
ridge	Ridge Regression	33.2523	91807.0704	243.7070	-294.4636	1.0356	3.2227	0.0433

Table 1 - Model comparison

Predicting Returns and Portfolio Construction: Following a similar approach, we identify the best model for predicting the returns of the liquidity pools. By employing these predicted returns and yields for GMX pools, we can utilize portfolio optimization algorithms to construct various risk-adjusted portfolios. This process allows us to create portfolios with different risk-return profiles.

Performance Analysis: The cumulative return for these portfolios, rebalanced at the end of each month, is plotted in Figures 2 and 3. The portfolios are optimized for returns in Figure 2, and for returns and yield in Figure 3. As you can see, the portfolio with the highest Sharpe ratio appears to outperform the others. This ratio, a measure of risk-adjusted return, suggests that this portfolio offers a better return relative to its level of risk.

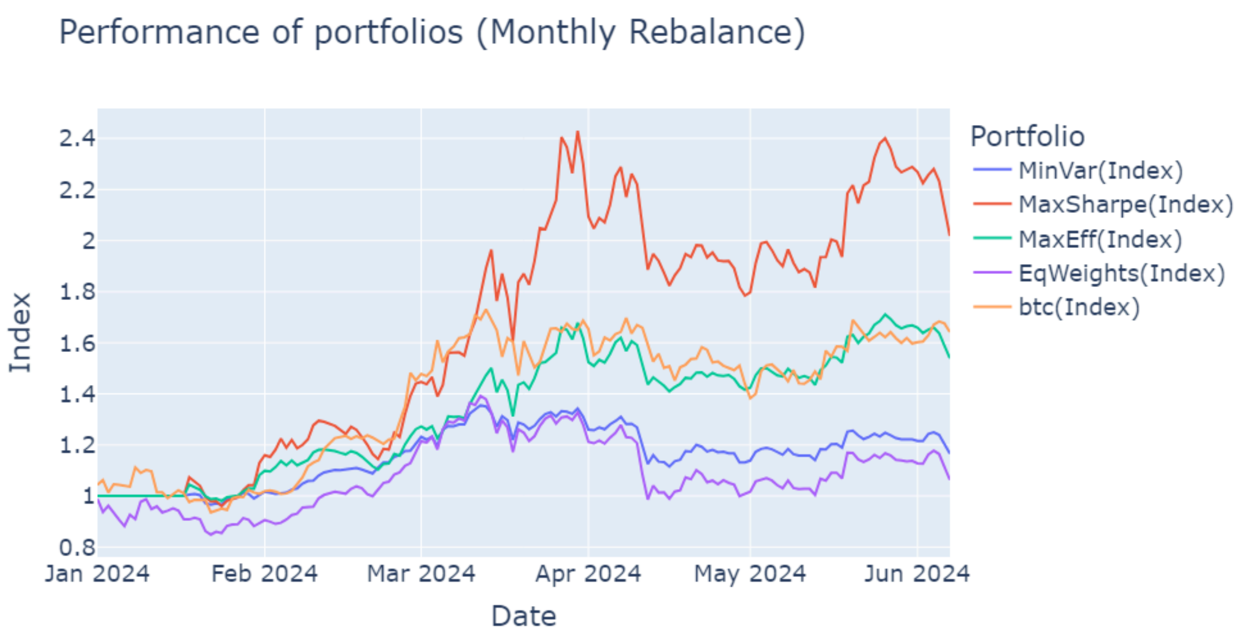

Figure 2 - Performance of portfolios (yield)

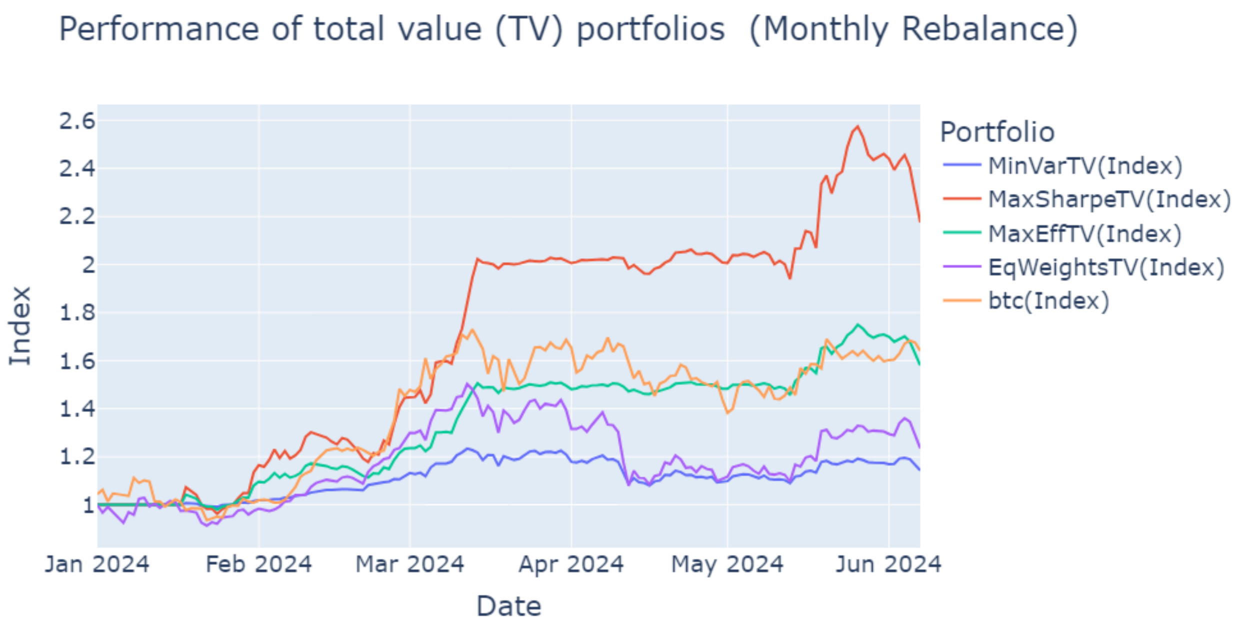

Figure 3 - Performance of total value (yield + price)

Benchmark Comparison: Finally, to gauge the effectiveness of our optimized portfolio, we compare its performance against a benchmark, such as Bitcoin (BTC), as shown in Table 2. This comparison helps us understand how our strategy stacks up against a commonly used investment in the cryptocurrency market.

Performance Metrics	Benchmark	Mozaic Strategy
Cumulative Returns	***62.36%***	***101.84%***
CAGR%	***137.77%***	***250.85%***
Sharpe	***2.11***	***2.29***
Smart Sharpe	***2.07***	***2.24***
Sortino	***3.5***	***3.89***
Max Drawdown	***-20.11%***	***-26.55%***
Volatility (Ann.)	***46.06%***	***63.51%***
Expected Daily %	***0.34%***	***0.50%***
Expected Monthly %	***8.41%***	***12.42%***
Kelly Criterion	***24.39%***	***-7.31%***
Daily Value-at-Risk	***-4.39%***	***-6.00%***

Table 2 - Performance Metrics

Our analysis reveals promising results for the machine learning-optimized portfolio. Here's a breakdown of the key findings:

Higher Expected Daily Return: Compared to the benchmark's 0.34% daily return, our portfolio boasts a higher expected daily return of 0.5%. This translates to potentially greater returns over time.

Superior Risk-Adjusted Returns: The Sharpe ratio, a metric that considers both returns and risk, is critical for investors. By achieving a higher Sharpe ratio than the benchmark, our portfolio demonstrates superior risk-adjusted returns. This result implies that for a given level of risk, our portfolio offers a more attractive return proposition.

In essence, these findings suggest that portfolio optimization with machine learning predictions has constructed a portfolio with potentially better returns and a more favorable risk-reward profile compared to the benchmark, which is predominately a bitcoin-denominated GMX pool.

4. Conclusion

This paper delves into the effectiveness of machine learning techniques in our portfolio optimization framework. We go beyond traditional methods like mean-variance optimization by exploring machine learning techniques that incorporate predicted returns and yields in the portfolio construction.

Machine Learning for Prediction: A key aspect of this research was the use of machine learning models to predict returns, yields, and volatilities for GMX pools. These models were trained on historical data, encompassing factors that potentially influence GMX pool behavior, including cryptocurrency prices, DeFi protocol metrics, and technical indicators (Li et al., 2019). By capturing complex relationships within the data, the machine learning models generated more accurate predictions, aligning with recent research on the effectiveness of machine learning for cryptocurrency return forecasting (Yoon & Baek, 2022).

Integrating Machine Learning into Optimization: The predicted returns and covariances from the machine learning models were seamlessly integrated into the portfolio optimization framework. This allowed the models to account for the unique risk-return profile of GMX pools, including potential impermanent loss. By maximizing the Sharpe Ratio, the optimization process identified the portfolio offering the best balance between risk and return.

Results and Significance: Our findings demonstrated that incorporating GMX pools and utilizing machine learning for return and covariance prediction can significantly improve portfolio performance. Compared to a benchmark portfolio, the optimized portfolio achieved a higher cumulative return while boasting a superior Sharpe Ratio. This suggests that the machine learning-driven approach effectively captured the potential benefits of GMX pools, leading to a more robust and risk-adjusted investment strategy for the dynamic cryptocurrency market.

References

Almeida, J. & Gonçalves, T. (2022). A Systematic Literature Review of Volatility and Risk Management on Cryptocurrency Investment: A Methodological Point of View. Risks. 10(5):107.

Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering, Journal of Computational and Applied Mathematics 365, 112395.

Holovatiuk, O. (2020). Cryptocurrencies as an asset class in portfolio optimisation. Central European Economic Journal, 7(54), 33-55.

Khan, F.U., Khan, F. & Shaikh, P.A. (2023). Forecasting returns volatility of cryptocurrency by applying various deep learning algorithms. Future Business Journal 9(25).

Letho, L., Chelwa, G. & Alhassan, A.L. (2022), "Cryptocurrencies and portfolio diversification in an emerging market", China Finance Review International, 12(1), 20-50. https://www.emerald.com/insight/content/doi/10.1108/CFRI-06-2021-0123/full/html

Li, F., Li, X., & Wang, S. (2019). A Survey on Cryptocurrency Network Security. IEEE Communications Surveys & Tutorials, 21(2), 1289-1313. https://ieeexplore.ieee.org/document/9996070

Lorenzo, J. M., & Arroyo, J. M. (2023). Online risk-based portfolio allocation on subsets of crypto assets applying a prototype-based clustering algorithm. Financial Innovation, 9(25), 232.

Markowitz, H. M. (1952). Portfolio selection. The Journal of Finance, 7(1), 77-91.

Poudel, S., Paudyal, R., Cankaya, B., Sterlingsdottir, N., Murphy, M., Pandey, S., Vargas, J., & Poudel, K. (2023). Cryptocurrency price and volatility predictions with machine learning. Journal of Marketing Analytics, 11(4), 642-660

Amirzadeh, R., Nazari, A., & Thiruvady, D. (2022). Applying artificial intelligence in cryptocurrency markets: A survey. Algorithms, 15(11), 428.

Sahu, S., Ochoa Vázquez, J.H., Ramírez, A.F., Kim, J.-M. (2024). Analyzing Portfolio Optimization in Cryptocurrency Markets: A Comparative Study of Short-Term Investment Strategies Using Hourly Data Approach. Journal of Risk and Financial Management, 17(3):125. https://doi.org/10.3390/jrfm17030125

Sharifi, A. M., Khalili Damghani, K., Abdi, F., & Sardar, S. (2022). A hybrid model for predicting bitcoin price using machine learning and metaheuristic algorithms. Journal of applied research on industrial engineering, 9(1), 134-150

Tang, Y., Song, Z., Zhu, Y., Yuan, H., Hou, M., Ji, J., ... & Li, J. (2022). A survey on machine learning models for financial time series forecasting. Neurocomputing, 512, 363-380.

Wang, X., Zhang, H., Zhang, Y., Wang, M., Song, J., Lai, T., & Khushi, M. (2021). Learning nonstationary time-series with dynamic pattern extractions. IEEE Transactions on Artificial Intelligence, 3(5), 778-787.

Yoon, S., & Baek, E. C. (2022). A hybrid machine learning model for cryptocurrency return forecasting. Expert Systems with Applications, 192, 116282. https://www.sciencedirect.com/science/article/abs/pii/S2214635022000673