ML Strategy for Cryptocurrency Portfolio Management

Yeyuxi Yi, Ian Chen-Adamczyk

May 10, 2025

Abstract :

Cryptocurrency portfolio management presents unique challenges due to the market's inherent volatility, non-stationarity, and strong sentiment-driven dynamics. Traditional mean-variance optimization techniques often fail to capture the rapid regime shifts and nonlinear patterns observed in crypto asset prices. In this study, we develop a machine learning-driven portfolio strategy leveraging Long Short-Term Memory (LSTM) models to predict price movements and incorporating sentiment analysis to better account for market psychology. Our results show that the machine learning-enhanced portfolios consistently outperform traditional Markowitz portfolios in terms of Sharpe ratio and cumulative returns. Notably, the meta-portfolio, integrating multiple forecast models, exhibited the most significant performance improvements (with a Sharpe ratio of 27.10 in comparison to a dynamically-rebalanced benchmark with Sharpe ratio 16.60), particularly during periods of dynamic market conditions. However, we also observed challenges related to model underperformance during unexpected price surges and potential overfitting when incorporating sentiment data. These findings highlight the promise and limitations of using advanced machine learning techniques for dynamic cryptocurrency portfolio optimization. Future work will aim to enhance model robustness through hybrid forecasting methods and improved regularization techniques.

Index Terms: Cryptocurrency, LSTM, Portfolio Management, Financial Time Series Forecasting, Sentiment Analysis, Sharpe Ratio

1 Introduction

1.1 Motivation

Cryptocurrency is a digital currency that is exchanged between peers without the need for a third party, such as banks. It enables consumers to digitally connect directly through a transparent process, showing the financial amount but not the identities of the people conducting the transaction. The network consists of a chain of computers, all of which are required to approve a cryptocurrency exchange and prevent the duplication of the same transaction (Northeastern University, D'Amore-McKim School of Business, 2023). It is also decentralized, meaning that there is no central authority or entity that owns it; instead, it operates across a network of participants.

Cryptocurrency has been the subject of growing attention, due in part to its rapidly increasing and highly volatile exchange rate to other currencies. Despite the hype surrounding cryptocurrency, it is difficult to identify the factors that drive its growth and influence its value. For example, Bitcoin's decentralized structure, based on the contributions of its users rather than a central authority, implies that the dynamics of its economy may be strongly driven by social factors, which consist of interactions between market participants (Garcia et al., 2014).

Aswath Damodaran, an NYU finance professor, discussed Bitcoin in a CNBC interview. "It's a pricing case study, not a valuation," Damodaran said. "Whether Bitcoin is a currency or a collectible, it cannot be valued. It can be priced. The question we can ask is, '$60,000 a fair price for Bitcoin?'" Hence, in a space where the valuation approach does not work, technical indicators and sentiment analysis become critical.

Machine-learning models have demonstrated notable success in several key areas of portfolio management: return forecasting, risk modeling, dynamic allocation, and factor selection. In return forecasting, Fischer and Krauss (2018) employed Long Short-Term Memory (LSTM) networks to predict next-day stock returns on the S&P 500, achieving a pre-cost Sharpe ratio of 5.8—substantially outperforming memory-free classifiers like random forests or logistic regression. In dynamic allocation, Dixon, Halperin, and Bilokon (2020) surveyed RL applications in finance, noting that deep-RL algorithms can optimize portfolios in continuous action spaces, balancing risk and return objectives without explicit parametric assumptions.

Given the successful application of machine learning in traditional finance, we aim to apply similar methodologies to the cryptocurrency market, which presents unique challenges due to its noise and sentiment-driven dynamics.

Classical mean–variance optimization is notoriously sensitive to estimation error: small inaccuracies in the inputs—namely the expected return vector and covariance matrix—can lead to dramatically different portfolio weights and a suboptimal efficient frontier. Machine-learning methods can mitigate this by leveraging large, high-dimensional feature sets and nonlinear function approximators to produce more accurate, out-of-sample forecasts of returns and risks. When these ML-derived estimates replace naive historical averages in the Markowitz framework, the resulting "ML-enhanced" efficient frontier lies strictly above the traditional one, offering higher expected return for the same level of risk. In the context of cryptocurrency markets—where return distributions are heavy-tailed, non-stationary, and driven by complex on-chain and off-chain signals—ML models (e.g., deep neural networks, gradient-boosted trees) have shown superior predictive power compared to linear benchmarks. Empirical studies confirm these theoretical gains. For example, Fidelity's "Applying Machine Learning Portfolio Modeling to Bitcoin" report shows that ML-based return and risk forecasts for a Bitcoin-only universe generate an efficient frontier with both higher returns and lower volatility compared to a historical-mean approach. Hence, we want to further enhance this approach to construct a machine learning-guided cryptocurrency portfolio that outperforms a historical-mean guided portfolio.

Cryptocurrency assets are notorious for their outsized return potential—but that upside comes hand-in-hand with extreme volatility and regime shifts that defy traditional valuation frameworks. Standard econometric pricing models (CAPM, APT, even multifactor regressions) assume relatively stable correlations, normally distributed returns, and well-behaved risk premia—assumptions routinely violated in crypto markets. Empirical studies show that during rapid bull runs or sudden crashes, correlations among major coins spike toward one, only to decorrelate in calmer periods. Nevertheless, when constructing a portfolio of diversified crypto assets—spanning large-cap coins, mid-caps with robust DeFi ecosystems, and even select utility tokens—raw volatility can be smoothed, idiosyncratic rallies can be captured, and, in many backtests, return profiles can be achieved that lie above the traditional efficient frontier derived from equities and bonds. In other words, even though individual coins may swing wildly, a carefully balanced basket can deliver attractive risk-adjusted returns—provided adequate tools are used to manage those swings.

That's where machine learning comes in. Unlike static quant models, ML techniques can ingest not just price and volume histories but also "soft" signals like market sentiment. Non-technical indicators often lead price moves: a sudden surge in positive sentiment on Twitter or a spike in stablecoin minting on Ethereum can foreshadow fresh buying pressure. Training models such as LSTMs, by recognizing these patterns through time-series architectures, gain a forecasting edge that pure econometric regressions miss.

Finally, by plugging these ML-enhanced forecasts of expected returns and dynamic risk estimates into a mean–variance optimizer, the efficient frontier is effectively expandeed. Forecasted returns calibrated with real-time sentiment and network data allow the optimizer to tilt toward coins poised for short-term rallies, while the improved covariance estimates (learned via clustering or copula-based methods) keep overall portfolio volatility in check. The resulting portfolio will have a ex-ante Sharpe ratio sits noticeably higher than those of standard benchmarks.

1.2 Contribution

In this study, we develop a machine learning-driven portfolio management strategy specifically tailored for the volatile and sentiment-driven nature of cryptocurrency markets. Our primary contribution lies in integrating Long Short-Term Memory (LSTM) models for price prediction with sentiment analysis to optimize portfolio performance.

First, we construct a novel portfolio optimization approach that blends machine learning forecasts with classical mean-variance optimization. Specifically, we use LSTM models to predict future prices of major cryptocurrencies, leveraging both technical indicators and sentiment data.

To enhance prediction accuracy, we perform hyperparameter tuning using Optuna and compare the LSTM models with traditional benchmarks, including trivial lag-one prediction.

Second, we analyze the significance of sentiment features, particularly the Fear and Greed (F&G) Index, within our forecasting framework. Through XGBoost feature importance analysis and ablation studies, we determine that sentiment data can improve predictions for some assets, though it may also lead to overfitting, especially in sentiment-driven cryptocurrencies like Dogecoin.

Third, we develop a meta-portfolio that dynamically blends multiple model predictions, including both static and dynamic mean-variance approaches. This meta-portfolio adapts to varying market conditions by adjusting weights based on recent performance, aiming to mitigate risks associated with sudden market shifts.

Our blended portfolio model integrates LSTM predictions with traditional return estimations, producing a dynamically rebalanced portfolio that effectively captures short-term price movements and sentiment-driven trends.

Additionally, we evaluate our model's performance against traditional mean-variance optimized portfolios, buy-and-hold strategies, and dynamically rebalanced portfolios. Our results demonstrate that the meta-portfolio consistently outperforms these benchmarks, achieving higher Sharpe ratios and cumulative returns, especially during periods of dynamic market conditions.

Despite the improved overall performance, we also identify limitations where individual LSTM models underperform during unprecedented price surges, primarily due to the model's tendency to underpredict extreme bullish trends.

Last but not the least, We highlight the challenge of overfitting when incorporating sentiment features, particularly in assets with volatile sentiment dynamics. This finding points to the need for robust regularization techniques to maintain model stability.

We also discuss the inherent difficulty of modeling rapid market rallies using historical data alone, suggesting the integration of trend-following techniques as a potential area for future improvement.

Through this work, we demonstrate the potential of combining machine learning predictions with traditional portfolio optimization techniques to construct a more resilient and adaptive cryptocurrency portfolio. The insights from our performance analysis provide a foundation for further research into hybrid modeling approaches that account for both statistical and sentiment-driven market dynamics.

1.3 Related Works

Many articles have evaluated the performance of cryptocurrency assets with regards to portfolio management. One study found that cryptocurrency asset portfolios could be selected to minimize risk and even outperform the S&P 500 (Hu et al., 2021). Though failing tests of normality, the returns for common cryptocurrencies fit a Cauchy distribution, allowing for accurate measures of value-at-risk (Hrytsiuk et al., 1909). Across a variety of decision-making processes and investor-defined criteria, Bitcoin and Ethereum emerged as consistently safe options compared to other cryptocurrencies (Maghsoodi, 2023). This mirrored the experience of the authors who, in optimizing a cryptocurrency portfolio to minimize risk, found that Bitcoin dominated the allocated weights (Hrytsiuk et al., 1909). In general, it seems that cryptocurrency portfolios offer high returns along with high risk, so managing such a portfolio requires frequently rebalancing (Saksonova and Kuzmina-Merlino, 2019).

Machine learning is a promising technique for portfolio management, in general and when applied to cryptocurrency assets. One successful "model-less" approach involved training a neural network on price data inputs to directly predict portfolio weights to optimize returns over the following period (Jiang and Liang, 2017). Though the model outperformed many of the tested benchmarks with increased returns and a better Sharpe ratio, the authors noted difficulties with the training data being highly time-dependent and newly-emergent cryptocurrencies having limited price history data. Another approach included technical financial indicators in the input data as well as the cryptocurrency "Fear and Greed" sentiment index (Vahidpour et al., 2023). In this model, LSTM models were trained for each individual asset to provide price predictions. These predictions were included with the rest of the inputs to a global DQN agent that controlled the weights of the portfolio. The authors selected multiple "non-stable" assets, managed by the agent, and one "stable" asset to ensure the entire portfolio was allocated. The authors found that market sentiment had a strong correlation with the agent's decisions, noting that market fear was correlated with less being allocated to non-stable assets, while market greed was correlated with the reverse being true (Vahidpour et al., 2023).

2 Preliminaries

2.1 Portfolio Optimization

We consider a portfolio of N assets, $1 \leq i \leq N$, for T trading periods from t_{j-1} to t_j, $1 \leq j \leq T$. The price of asset i at time t_j is $p_{i,j}$, and the return $r_{i,j}$ of asset i during trading period j is

$$r_{i,j} = \frac{p_{i,j}}{p_{i,j-1}} - 1$$

For Markowitz portfolio theory, the returns distributions of these assets can be modeled as multivariate normal $N(\vec{\mu}, \Sigma)$; the mean and covariance $\vec{\mu}$ and Σ can be estimated from historical returns.

We will also consider a portfolio with weights $w_{i,j}$ for each asset and trading period:

$$\{w_{i,j} | 1 \leq i \leq N, 1 \leq j \leq T\}$$

The return of such a portfolio for trading period j is

$$r_j = \sum_{i=1}^{N} w_{i,j} r_{i,j}$$

and the cumulative return is

$$Return = \prod_{j=1}^{T} (r_j + 1)$$

The realized Sharpe ratio indicates annualized, risk-adjusted returns, and is given by

$$Sharpe\ Ratio = \frac{Return^{365/T} - 1}{\sqrt{365} SD(r_j)}$$

$$SD(r_j) \approx \sqrt{\frac{\sum_{j=1}^{T} r_j - \bar{r}}{T - 1}}$$

where R_f is the risk-free rate. The risk-free rate for cryptocurrencies is considered to be 0, corresponding to the return of the Tether stablecoin in USD.

We will consider the portfolio performance for a train/test split $1 < S < T$: our train set includes trading periods $1 \leq j < S$, while our test set includes trading periods $S \leq j < T$. For any trading period, all information prior to and including the start of the period is available for determinining the portfolio weights. In particular, our training set will be from January 1, 2022 to June 30, 2024; our testing set will be from June 1, 2024 to December 31, 2024. This choice is made according to the availability of cryptocurrency data through the Binance API, and to avoid potentially confounding factors from the pandemic.

2.2 Cryptocurrency Market

The cryptocurrency market is primarily dominated by Bitcoin, but includes a wide variety of altcoins, which can range from viable assets that implement substantially different features from Bitcoin, to memecoins whose value depends mainly on popular trends. Cryptocurrency indices like the Coinbase 50 list assets with the largest market share and most near-term stability.

A stablecoin is a cryptocurrency that is designed to have a constant price with respect to a national currency. For example, Tether (USDT) has a practically constant price of 1.00 USD. For this paper, cryptocurrency prices will be given in USDT instead of USD.

2.2.1 Fear and Greed Index

The Fear and Greed Index (F&G) is a sentiment score calculated daily by Alternative.me by looking at volatility (25%), market momentum and volume (25%), social media sentiment (15%), Bitcoin dominance (10%), Google trends (10%), and survey responses (15%). A high F&G score (60-100) indicates a bullish market and market movement to smaller and more volatile cryptocurrencies, whereas a lower F&G score (0-40) indicates a bearish market and market movement to more established and stable cryptocurrencies (especially Bitcoin).

2.3 ML Forecasting

This paper primarily considers price prediction through LSTM regression. For asset i and trading period j, we compile a feature vector of technical indicators $\vec{v}_{i,j}$ as well as the target price for the following period $p_{i,j+1}$. Each LSTM model will predict the price using the features for τ preceding periods:

$$\text{LSTM} : \left[\vec{v}_{i,j-\tau+1}, \ldots, \vec{v}_{i,j-1}, \vec{v}_{i,j}\right] \mapsto \hat{p}_{i,j+1}$$

The above prediction will be trained for $1 \leq j < S$ and tested for $S \leq j < T$. As a benchmark, we will consider the trivial price prediction, which simply predicts the next period's price by using the current period's price at lag one:

$$\text{trivial} : \left[\vec{v}_{i,j-\tau+1}, \ldots, \vec{v}_{i,j-1}, \vec{v}_{i,j}\right] \mapsto p_{i,j}$$

We use Python TensorFlow sequential models to implement these LSTM models. Because LSTM models have many configurable hyperparameters, we use Optuna, a Bayesian optimization library, to streamline hyperparameter tuning.

3 Methodology

3.1 Data Collection and Pre-Processing

Show a figure of the asset prices. Describe the types of returns and shifts that are apparent, and why Bitcoin looks so weird over the years. Asset correlations.

We aggregated data for five major cryptocurrencies—Bitcoin (BTC), Dogecoin (DOGE), Ethereum (ETH), Solana (SOL), and Ripple (XRP)—spanning 01/2022 to 03/2025. We selected those 5 coins, not only there are the top 5 largest market cap coins in current(05/2025) cryptocurrency market, but also their total weight is over 90% that captures the majority of the market movement. Price and volume series were retrieved from the Binance REST API at daily frequency, ensuring consistency across symbols. There were no instances of missing data at 1-day frequencies for the duration of the training and testing periods. Daily F&G data since 2020 was downloaded from the Alternative.me API.

Besides data collection, we have also computed key technical indicators designed to capture momentum, trend, volatility, and order-flow dynamics:

- Moving Average (MA): smooths price series over a fixed window to identify underlying trend direction and support/resistance levels.
- Relative Strength Index (RSI): identifies overbought/oversold conditions by comparing recent gains to losses.
- ADX (Average Directional Index): quantifies the strength of the prevailing trend, helping distinguish between trending and ranging regimes.
- Momentum (MOM): measures the change in closing price over a fixed look-back, signaling accelerations in buying or selling pressure.
- Volatility (Risk): computed as the rolling standard deviation of log returns, it serves as a proxy for short-term risk and drawdown potential.
- Slope: the linear-regression slope of price over the look-back window, providing a compact trend estimate.
- Buy-Pressure Ratio & Trades-per-Volume: derive from order-book and trade data to approximate net buying interest and liquidity intensity.

By incorporating this diverse, normalized feature set into our ML models, we capture not only where prices are headed but also how strongly, with what volatility, within what trend regime, and on what underlying order-flow dynamics.

At the data processing stage, F&G scores were discretized into five categorical regimes (Extreme Fear, Fear, Neutral, Greed, Extreme Greed) and one-hot encoded. Finally, all continuous features were standardized (zero mean, unit variance) to prevent scale dominance.

Figures 1 and 2 show a preliminary view about our data. The price chart of BTC vs. DOGE visually demonstrates their co-movement during bullish phases, while DOGE exhibits sharper, sentiment-driven spikes. By including these visualizations, we contextualize the interactions among cryptocurrencies, validating our choice of features to better model cross-asset dynamics and market sentiment.

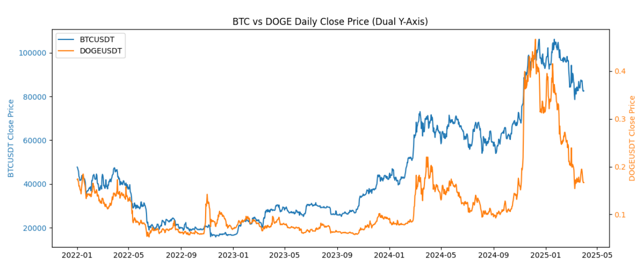

Figure 1. Bitcoin and Dogecoin price from 2022 to 2025

The correlation matrix captures inter-crypto relationships. The matrix reveals that Bitcoin (BTC) has the highest correlation with Solana (SOL) and Dogecoin (DOGE), suggesting a shared market sentiment and potential risk clustering during volatile periods. Ripple (XRP), on the other hand, shows comparatively lower correlations with Ethereum (ETH) and other coins, reflecting its unique revenue-backed characteristics.

	BTC	DOGE	ETH	SOL	XRP
BTC	1.00	0.85	0.79	0.94	0.77
DOGE	0.85	1.00	0.68	0.82	0.80
ETH	0.79	0.68	1.00	0.87	0.41
SOL	0.94	0.82	0.87	1.00	0.61
XRP	0.77	0.80	0.41	0.61	1.00

Figure 2. Correlation Matrix of all five assets

3.2 *Sentiment Feature Importance*

Figures 3 and 4 below illustrate the movement of F&G score and the corresponding market conditions. Noticeably, Figure 4's alternating red, gray, and green bands make it visually clear that sentiment shifts precede price moves: sustained green zones often herald breakouts, long red stretches coincide with steep sell-offs, and neutral gray patches mark consolidation phases. Sentiment–price divergences—where prices stabilize or tick up under lingering bearish shading—flag contrarian entry points. By encoding the timing, color intensity, and duration of these regimes as features, ML models can leverage leading sentiment signals to improve short-term price forecasts and portfolio decisions.

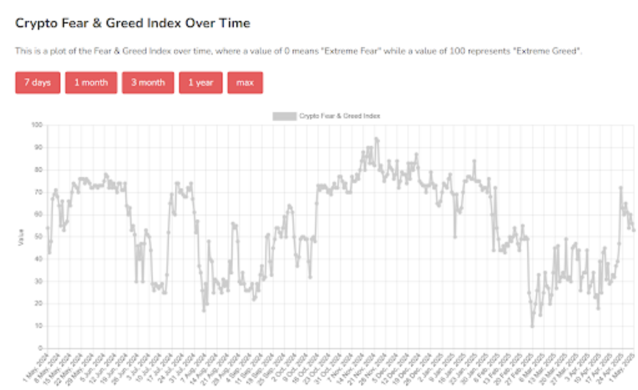

Figure 3. Fear&Greed Index from from 05/2024 to 05/2025

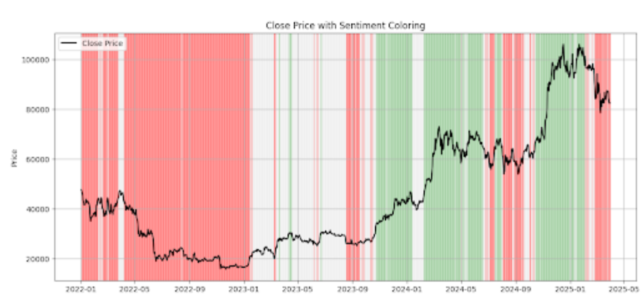

Figure 4. BTC closing prices from 2022-2025 with (F&G < 45) highlighted in red, and (F&G > 65) highlighted in green.

We use two methods to determine the importance of sentiment as a feature: XGBoost feature importance, and ablation study. XGBoost, or Extreme Gradient Boosting, is a fast, scalable tree-based ensemble algorithm optimized for predictive accuracy. XGBoost quantifies how much each input variable contributes to a model's predictive performance by measuring its gain, the average improvement in the training objective when the feature is used.

For the ablation study, we train two LSTMs for each asset, with one that includes sentiment as a feature and one that excludes sentiment.

We expect these two methods to provide similar results because both approaches aim to quantify the impact of sentiment on price prediction accuracy.

XGBoost feature importance directly evaluates the contribution of each feature to the model's predictive power by measuring the average gain when the feature is used for decision splits. This method highlights features that consistently improve model performance, offering a clear indication of which variables are most influential.

On the other hand, the ablation study compares two versions of the LSTM model: one with sentiment features and one without. By observing the difference in prediction accuracy, we can infer the practical value of including sentiment data. If sentiment is indeed a crucial predictor, the model that incorporates it should significantly outperform the model without it.

3.3 LSTM Forecast

In order to build the LSTM model, we used two phases of Optuna hyperparameter tuning. For the first phase, we trained the model on synthetic sinusoidal data and used hyperparameter tuning on the activation functions, optimizer, and layering configuration. This phase confirmed that the TensorFlow model and data pre-processing pipeline was properly configured to capture the desired temporal trends. In particular, we confirmed that the scaling of the data was appropriate; that a single hidden LSTM layer was sufficient; and that the 'adam' optimizer with mean-squared error (MSE) loss and a Dense linear output layer provided adequate results. For the second phase, we fine-tuned more nuanced hyperparameters for each specific asset. The hyperparmeters included the lookback window size τ, the dropout rate, the learning rate, and batch size.

3.4 Optimization Program

We considered optimization programs based on the max-Sharpe mean-variance optimization program. Below is a standard formulation:

$$\max_{(\vec{w})_i} \frac{\vec{w}^T\left(\vec{\mu} - R_f\vec{1}\right)}{\sqrt{\vec{w}^T\boldsymbol{\Sigma}\vec{w}}}$$

$$\text{subject to } \; w_i \geq 0 \;\; (1 \leq i \leq N) \;\; \text{and} \;\; \sum_{i=1}^{N} w_i = 1$$

where $\vec{\mu}$ and $\boldsymbol{\Sigma}$ are estimated from historical returns.

To optimize this program, the pyportfolio implementation is used with positive but negligible $R_f = 0.0001$, as the implementation is not designed to accomodate zero or negative risk-free rates. This optimization problem can also become infeasible for the pyportfolio solver when every asset's expected return falls below the risk-free rate, or when the covariance matrix is singular. This can occur for some trading periods in the course of dynamic rebalancing, in which case the portfolio weights will be set to those of the previous period.

3.4.1 Benchmarks We consider two benchmarks: buy and hold max-Sharpe, and dynamically-rebalanced max-Sharpe. For buy and hold max-Sharpe, the portfolio weights are optimized using mean-variance optimization, with estimated return means and covariances calculated from the train set ($\sim$ 1000 entries) . Dynamically-rebalanced max-Sharpe works similarly, but the estimated means and covariances are instead calculated from a rolling window of constant (though arbitrary) length W. In practice, W will range from 30 to 720. Both benchmarks are selected because they rely entirely on historical data and maximize the Sharpe ratio. We note that due to the volatility of cryptocurrencies, estimates from the train set can become outdated, since cryptocurrency trends can shift in the course of a few months. Thus, we use rolling estimates in the second benchmark, to emphasize recency in the historical data.

3.4.2 Blended Portfolio We modify the max-Sharpe optimization program by using rolling windows, and augmenting the historical data to include the LSTM prediction. To calculate the weights for trading period j, $\vec{\mu}$ and $\boldsymbol{\Sigma}$ are estimated from the historical prices:

$$p_{i,j-W}, p_{i,j-W+1}, \ldots, p_{i,j-1}, (1-\alpha)\overline{p}_i + \alpha\hat{p}_{i,j}$$

where $\overline{p}_i$ is the rolling mean for asset i from t_{j-W} to t_{j-1} inclusive. Thus, decreasing W increases the influence of recent prices, and increasing α increases the influence of LSTM predictions.

3.4.3 Meta Portfolio Since the blended portfolio has multiple parameters (α, W), it is not immediately obvious which parameters will be optimal for the test set. Thus, we now consider a portfolio of parameterized portfolios. We create a dynamically-rebalanced max-Sharpe portfolio based on the returns $r_{I,j}$ of each portfolio I for trading period j , with $1 \leq I \leq M$ and M the number of portfolios. A meta-portfolio essentially monitors the performance of each portfolio, adjusting to improving or degrading performance on the test set and avoiding correlated risk.

4 Experiments

4.1 Sentiment Feature Importance

Figures 5-7 capture the features importance ranking of different types of cryptocurrency, leading to insightful interpretations. As a hype-driven memecoin, DOGE's price is overwhelmingly governed by peaks of exuberant sentiment. When the Fear&Greed Index hits "Extreme Greed," it encapsulates the social-media hype and retail mania that drive its rallies—dwarfing the role of more "fundamental" momentum or volatility indicators. Bitcoin's broad market status makes it sensitive to macro momentum and volatility regime shifts. Severe bearish sentiment spikes (e.g. "Extreme Fear") rank highly—reflecting flight-to-safety episodes—but routine sentiment signals (greed/neutral) contribute less once you control for technical trend measures. For a utility- and revenue-oriented token like XRP, traditional volatility and trend indicators dominate price prediction. Bearish extremes in sentiment still pro-

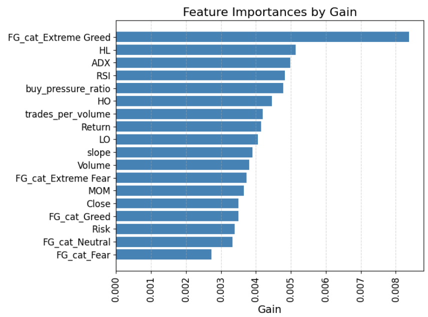

Figure 5. Features Importance Ranking for XGBoost of future return (DOGE)

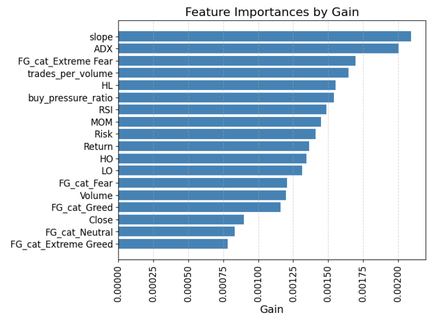

Figure 6. Features Importance Ranking for XGBoost of future return (Bitcoin)

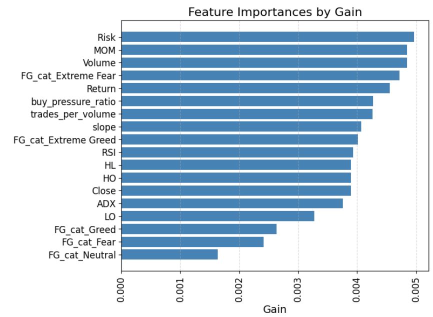

Figure 7. Features Importance Ranking for XGBoost of future return (Ripple)

vide early warning (hence the high rank of Extreme Fear), but overall the network's on-chain health and technical momentum carry more explanatory power than everyday sentiment fluctuations.

The ablation study aimed to evaluate the effect of incorporating Fear & Greed (F&G) sentiment features in LSTM models predicting next-day close prices. Contrary to our expectations, the results indicated below that in some cases, removing the F&G features led to improved model performance.

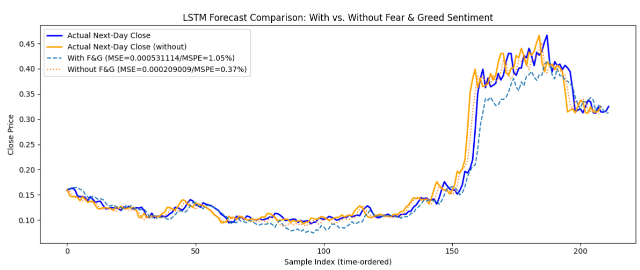

Figure 8. MSE comparison of LSTMS with and without F&G for Dogecoin

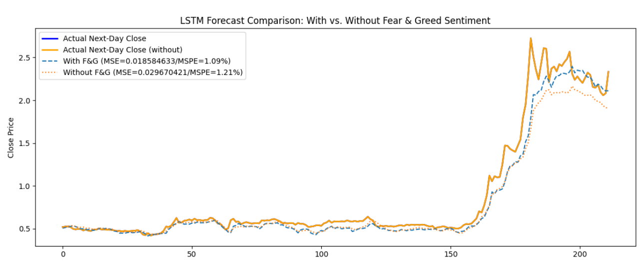

Figure 9. MSE comparison of LSTMS with and without F&G for XRP

Figures 8 and 9 demonstrate a result that contradicts the XGBoost Feature Importance rankings, showing that the model without the F&G sentiment features sometimes performs better than the model that includes them. Specifically, in the case of Dogecoin, where the price is predominantly driven by sentiment, the model without F&G achieved a lower Mean Squared Error (MSE) and Mean Squared Percentage Error (MSPE). This indicates that excluding sentiment features actually enhanced prediction accuracy.

Conversely, in the case of XRP, where the price is less influenced by sentiment, the model without F&G also achieved better results.

One plausible reason for this outcome is overfitting. The F&G sentiment features might lead the model to overfit to noise rather than capturing the actual trend when predicting short-term price changes. This effect is particularly noticeable in Dogecoin, where sentiment showed a leading gain in XGBoost, suggesting that the model might be capturing short-lived patterns that do not generalize well.

While sentiment analysis can be valuable, incorporating it without adequate regularization can lead to overfitting during the training period, making the model less applicable to out-of-sample data.

4.2 LSTM Forecast

We trained LSTM models on each cryptocurrency asset, developing two distinct models per asset: one to predict the close price and another to predict the 7-day moving average (MA(7)). The rationale behind the latter approach was to assess model performance in a less noisy environment, as the moving average tends to smooth out short-term fluctuations. Table 1 presents the performance of our trained LSTMs, comparing them to a trivial benchmark that simply uses a one-day lag as the prediction.

By comparing the Mean Squared Error (MSE) between the models and their benchmarks, we observed that our models systematically underperform the benchmark across most assets. Upon examining the predicted and actual value graphs, we discovered a consistent pattern: all of our models tend to underpredict the actual value. One example is the figure below showing the consistency of underprediction during the price surge.

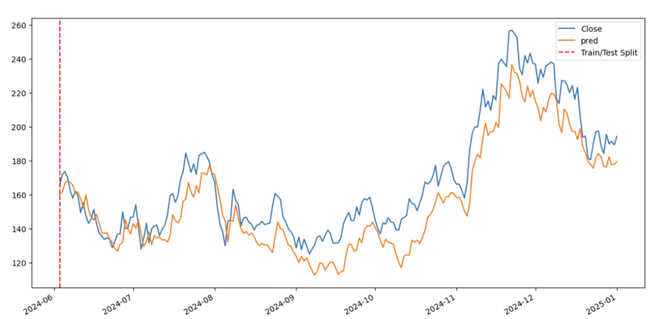

Figure 10. Solana Predicted value vs Actual value

One plausible explanation for this discrepancy is the market dynamics during the test period. Specifically, the prices of all the cryptocurrencies we studied surged to unprecedented heights—levels that were not represented in the training data. As a result, the LSTM models, having learned from relatively stable or moderate price movements, became inherently conservative and exhibited a kind of prediction inertia. In essence, the models were "reluctant" to predict such high values, leading to systematic underprediction during a highly bullish market phase. One example is the figure below showing Bitcoin highest price in the test period is almost double as the highest price in the training period.

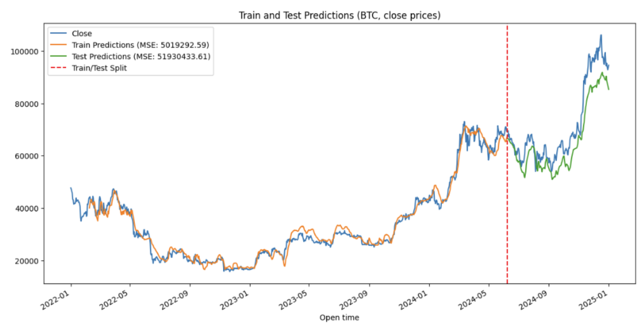

Figure 11. Bitcoin Predicted value vs Actual value

This insight suggests that the models' tendency to underfit rapidly rising trends may stem from the inherent challenge of generalizing from past data that lacks comparable price spikes. Moving forward, incorporating dynamic adjustment techniques or using hybrid models that blend LSTM predictions with trend-following components might mitigate this issue.

Figure 12. Bitcoin close prices compared to the LSTM forecast for the train and test set. The above graph shows predictions for the immediate target, and the below graph shows predictions for the smoothed target.

(a) Direct close price target **(b)** Smoothed MA(7) target

Figure 13. LSTM forecast residuals for ETH vs. one-day price movement. The left figure shows a correlation between price movement and residual error, while the right figure shows less of a correlation.

Figure 13 shows correlated clustering of residuals for the direct close price target, and relatively uncorrelated clustering of the residuals for the smoothed moving-average target, with respect to the one-day movement of the price into the target day. The correlation of the residuals could have implications, going beyond the MSE score, for how inaccurate predictions impact the optimization program. In this case, the prediction will likely underpredict the next day price for upward price movements, and overpredict the next day price for downward price movements, in both cases underestimating the magnitude of each movement.

To simplify the optimization program in the next section, we use only the direct close price forecasts instead of the moving-average smoothed forecasts.

4.3 Portfolio Performance

We compared six portfolios, two of which incorporate ML predictions. The "Test Hold Portfolio" is a buy-and-hold portfolio using max-Sharpe weights based on $\vec{\mu}$ and $\mathbf{\Sigma}$ estimated from the test set. It represents the maximum Sharpe ratio possible on the test set without any rebalancing. "Train Hold Portfolio" analagously uses estimates from the train set, and represents the maximum Sharpe ratio based on assumptions from historical data. The "Dynamically-Rebalanced Mean Variance" portfolios use values on a rolling one-year window and a rolling 30-day window. "Close Price Prediction Portfolio" implements the blended portfolio model with specific values of α and W, and "Meta Portfolio" optimizes the blended portfolios for $\alpha \in \{0, 0.1, 0.5\}$ and $W \in \{30, 80, 240, 720\}$. Figure 14 shows the cumulative returns for all five portfolios. Table 2 lists the portfolio performance metrics, as well as the corresponding metrics for the assets.

The test hold portfolio radically outperforms the one-year dynamically-rebalanced portfolio and train hold portfolio, so long-term historical trends do not appear to be effective, while the blended price prediction portfolio with $W = 30$ and the 30-day dynamically-rebalanced portfolio illustrate the importance of short-term trends. The meta portfolio does outperform all of the other portfolios. Only the meta portfolio and the blended portfolio perform close to the level of the XRP asset, illustrating how these two portfolio methods are particularly adept at exploiting XRP's sudden growth.

5 Conclusion

Our machine learning-driven portfolio management strategy for cryptocurrencies demonstrates promising results compared to traditional techniques, particularly the classical mean-variance op-

Table 1
LSTM Model Performance With Tuned Hyperparameters.
MSE: Mean Squared Error; **Benchmark**: Trivial Lag-One Prediction; **Layers**: input layer, one LSTM layer, output layer

Asset	Target	Layer Size	Dropout	LR	Batch	Lookback	Epochs	MSE (LSTM)	MSE (Benchmark)
BTC	MA(7)	136	0.00491	0.000267	16	45	16	8.5684E+06	5.3041E+05
BTC	CLOSE	172	0.03248	0.000081	64	38	21	2.4489E+07	3.4965E+06
XRP	MA(7)	177	0.00867	0.001078	64	22	31	8.7010E-03	9.6810E-04
XRP	CLOSE	169	0.00174	0.000742	16	12	19	1.3244E-02	5.3863E-03
DOGE	MA(7)	135	0.00040	0.000571	16	20	14	4.4000E-05	3.2321E-05
DOGE	CLOSE	147	0.00451	0.004537	32	15	33	4.9100E-04	1.4930E-04
ETH	MA(7)	129	0.00358	0.009762	32	51	19	3.0742E+03	1.5456E+03
ETH	CLOSE	169	0.00606	0.000164	32	48	37	1.2552E+04	9.9347E+03
SOL	MA(7)	159	0.00061	0.001543	64	59	31	4.3011E+01	6.6446E+00
SOL	CLOSE	68	0.00175	0.001486	32	14	16	9.6078E+01	4.7900E+01

Table 2
Portfolio Performance for Test Period (2024/06/01 to 2025/01/01).
Description: Placeholder for portfolio performance metrics from June 2024 to January 2025.

Portfolio	Cumulative Return	Annualized Return	Annualized Volatility	Sharpe Ratio
Test Hold	4.891592	16.61119	0.986739	16.83443
Hold	1.31903	0.649271	0.779784	0.83263
Dynamic MVO (365)	1.654721	1.484393	0.81347	1.824767
Dynamic MVO (30)	5.010041	17.389277	1.044925	16.641644
Close Prediction	5.929575	23.934459	1.053663	22.715483
Meta	6.5945	29.213764	1.075953	27.151514
BTCUSDT	1.385676	0.797679	0.502432	2.059156
DOGEUSDT	2.223719	3.207848	0.98225	5.763514
ETHUSDT	0.943895	-0.098611	0.641963	0.162567
SOLUSDT	1.251643	0.497175	0.781126	1.31012
XRPUSDT	4.745321	15.441657	0.985686	25.30047

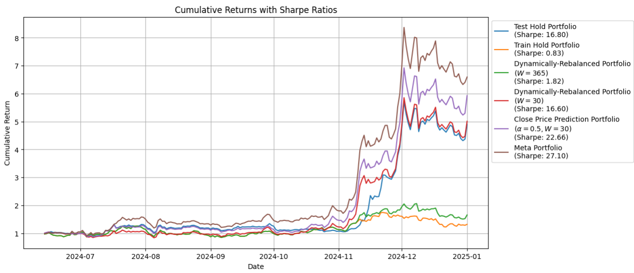

Figure 14. Cumulative returns for portfolio weights on the test set (2024/06/01 to 2025/01/01)

timization. By incorporating LSTM-based price forecasts and sentiment-driven insights, our dynamic portfolio consistently achieved a higher Sharpe ratio and cumulative return than the traditional Markowitz approach. Notably, the meta-portfolio, which combines multiple machine learning forecasts, significantly outperformed both static buy-and-hold strategies and traditional dynamic rebalancing.

However, our results also reveal some limitations and challenges. Despite the improved performance of our meta-portfolio, individual LSTM models occasionally underperformed compared to simpler benchmarks, particularly during periods of unprecedented price surges. This underperformance was primarily due to the models' tendency to underpredict during bullish phases, likely caused by the training data lacking comparable high-price movements. Furthermore, the inclusion of sentiment features sometimes led to overfitting, especially in the case of highly sentiment-driven assets like Dogecoin. These findings suggest that while machine learning can enhance portfolio optimization, the models must be carefully tuned to account for rapid market shifts and noise sensitivity.

Future work could focus on incorporating more robust regularization techniques and exploring hybrid models that blend LSTM predictions with trend-following or momentum-based adjustments to better capture extreme market conditions without overfitting.

Acknowledgements

Over the course of the seminar, this research received support from Eric Balkanski, Assistant Professor, and Cherlin Zhu, PhD student in the IEOR department at Columbia University.

Code

All code used in the experiments is available on the GitHub repository, `https://github.com/iancadev/crypto`.

References

Dixon, Matthew, Igor Halperin, and Paul Bilokon (July 2020). "Applications of Reinforcement Learning". In: pp. 347–418. ISBN: 978-3-030-41067-4. DOI: 10.1007/978-3-030-41068-1_10.

Fischer, Thomas and Christopher Krauss (2018). "Deep learning with long short-term memory networks for financial market predictions". In: *European Journal of Operational Research* 270.2, pp. 654–669. ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2017.11.054. URL: https://www.sciencedirect.com/science/article/pii/S0377221717310652.

Garcia, David et al. (2014). "The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy". In: *Journal of The Royal Society Interface* 11.99, p. 20140623. DOI: 10.1098/rsif.2014.0623. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rsif.2014.0623. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2014.0623.

Hu, Yuan, W. Lindquist, and Frank Fabozzi (Apr. 2021). "Modeling Price Dynamics, Optimal Portfolios, and Option Valuation for Cryptoassets". In: *The Journal of Alternative Investments* 24, jai.2021.1.133. DOI: 10.3905/jai.2021.1.133.

Ijadi Maghsoodi, Abtin (Oct. 2022). "Cryptocurrency portfolio allocation using a novel hybrid and predictive big data decision support system". In: *Omega* 115. DOI: 10.1016/j.omega.2022.102787.

Jiang, Zhengyao and Jinjun Liang (Sept. 2017). "Cryptocurrency Portfolio Management with Deep Reinforcement Learning". In: DOI: 10.1109/IntelliSys.2017.8324237.

Northeastern University, D'Amore-McKim School of Business (2023). *Guide to the Rise of Cryptocurrency, Digital Currency and Bitcoin*. https://onlinebusiness.northeastern.edu/masters-in-finance-msf/knowledge/guide-to-the-rise-of-cryptocurrency-digital-currency-and-bitcoin/. Accessed: 2025-05-10.

Petro, Hrytsiuk, Tetyana Babych, and Larysa Bachyshyna (Jan. 2019). "Cryptocurrency portfolio optimization using Value-at-Risk measure". In: DOI: 10.2991/smtesm-19.2019.75.

Saksonova, S. and I. Kuzmina-Merlino (2019). "The Principles of Creating a Balanced Investment Portfolio for Cryptocurrencies". In: *Reliability and Statistics in Transportation and Communication*. Ed. by I. Kabashkin, I. Yatskiv (Jackiva), and O. Prentkovskis. Cham: Springer International Publishing, pp. 714–724.

Vahidpour, M. et al. (2023). "Optimizing Dynamic Portfolio Management in the Cryptocurrency Market Using Multi-Agent Deep Reinforcement Learning and the Fear and Greed Index". In: *Journal of Industrial and Systems Engineering* 15.2, pp. 94–111.