A Markov Chain-Based Trading Strategy for Cryptocurrency Markets

Jehu Mtine \ jehumtine@proton.me

January 1, 2026

Abstract

This paper presents a novel trading strategy for cryptocurrency markets that uses a Markov chain model to identify predictive patterns in price and volume data. We discretize market behavior into a finite set of states and identify high-probability state sequences to generate trading signals. The strategy is evaluated using a comprehensive, walk-forward backtesting engine on SOL/USDT hourly data from 2019 to 2024. The backtester incorporates realistic transaction costs, dynamic risk management based on the Average True Range (ATR), and undergoes a rigorous hyperparameter optimization process using `hyperopt`. The results demonstrate that the strategy can achieve strong positive, risk-adjusted returns, providing a transparent and reproducible framework for quantitative strategy development.

A Markov Chain-Based Trading Strategy for Cryptocurrency Markets

1. Introduction

The search for predictive models in the volatile and complex cryptocurrency markets is a significant challenge in quantitative finance. A key requirement for developing any robust trading strategy is a rigorous backtesting framework that can validate its performance on historical data while accounting for real-world market conditions like transaction costs and slippage. Without this, a theoretically sound model may fail in a live environment.

This paper proposes that Markov chains offer a compelling framework for this challenge. A Markov chain is a stochastic model where the probability of a future event depends only on the current state. By discretizing the continuous and often chaotic behavior of the market into a finite number of states—defined by price action and trading volume—we can model the market's dynamics and identify predictive patterns in its state transitions.

We present a novel trading strategy that utilizes this Markov chain model to forecast short-term price movements. The core of the strategy is the identification of "useful sequences"—sequences of market states that have a high probability of preceding a specific bullish or bearish movement. These sequences form the basis of our trading signals.

The strategy is implemented and evaluated within a custom-built backtesting engine that employs a walk-forward analysis methodology to ensure robustness and adaptability to changing market regimes. The backtester simulates real-world trading by incorporating transaction costs, dynamic stop-loss and take-profit levels based on the Average True Range (ATR), and a systematic hyperparameter optimization process using the `hyperopt` library. This paper details the full methodology and evaluates the strategy's performance, aiming to provide a transparent and reproducible framework for the development of quantitative trading strategies.

2. Methodology

The development and evaluation of the Markov chain-based trading strategy are grounded in a systematic and reproducible methodology. This section details the data used, the process of state classification, the mechanism for signal generation, the architecture of the backtesting framework, and the procedure for hyperparameter optimization.

2.1. Data

The primary dataset used for this study consists of historical price and volume data for the SOL/USDT trading pair, sourced from a major cryptocurrency exchange. The data is sampled at a 1-hour resolution and spans the period from January 1, 2019, to December 31, 2024. This timeframe was selected to include a variety of market conditions, including the bull market of 2020-2021, the subsequent bear market, and periods of sideways consolidation. The dataset contains the standard open, high, low, and close (OHLC) prices, as well as trading volume.

2.2. State Classification

The cornerstone of the Markov chain model is the discretization of continuous market data into a finite set of states. A state is a composite representation of market conditions during a single time period, derived from two key dimensions: price movement and trading volume.

Price State: To account for changing market volatility, the price change from the previous period's close is normalized by the Average True Range (ATR). The ATR is a rolling measure of volatility that captures the average range between high and low prices, adjusted for any gaps. The normalized price change is then categorized into one of four price states: * **State 0 (Strong Uptrend):**

Normalized price change exceeds a predefined positive threshold. * **State 1 (Mild Uptrend):** Normalized price change is positive but below the threshold. * **State 2 (Mild Downtrend):** Normalized price change is negative but above the negative threshold. * **State 3 (Strong Downtrend):** Normalized price change is below a predefined negative threshold.

Volume State: Trading volume is classified into one of three states by comparing the current period's volume to a rolling simple moving average (SMA) of volume. * **State 0 (High Volume):** Current volume is significantly above its SMA. * **State 1 (Low Volume):** Current volume is significantly below its SMA. * **State 2 (Average Volume):** Current volume is within a normal range of its SMA.

A final, composite state for each time period is created by combining the price state and the volume state. This results in a total of 12 possible states (4 price states x 3 volume states), each representing a unique combination of price and volume dynamics.

2.3. Signal Generation

Trading signals are generated by identifying "useful sequences" of states that have a high predictive probability for the subsequent state. The process is as follows:

1. **Sequence Extraction:** The historical data is scanned to extract all sequences of a predefined length (e.g., a sequence of 6 consecutive hourly states).
2. **Transition Probability Calculation:** For each unique sequence, the model calculates the probability of it transitioning to each of the 12 possible states in the next time period.
3. **Identification of Useful Sequences:** A sequence is deemed "useful" if it predicts the next state with a probability exceeding a certain threshold (e.g., 65%).
4. **Signal Interpretation:** The predicted next state is interpreted as either "BULLISH" or "BEARISH" based on its underlying price component. If a useful sequence is observed and the predicted next state aligns with the broader market trend (as determined by a long-term moving average), a trading signal is generated.

This process is applied to both the 1-hour data and a resampled 4-hour timeframe, allowing the strategy to capture patterns across multiple time horizons.

2.4. Backtesting Framework

The strategy's performance is evaluated using a custom-built backtesting engine that simulates historical trading with a high degree of realism.

- **Walk-Forward Analysis:** The backtest employs a walk-forward methodology to mitigate the risk of overfitting. The data is divided into

rolling windows, with each window consisting of a training period (e.g., 4 years) and a subsequent testing period (e.g., 1 month). The model is trained on the training data to identify useful sequences, which are then used to trade on the testing data. This process is repeated for the entire dataset, ensuring that the trading logic is always based on information that would have been available at that time.

- **Position Management:** The strategy is configured to manage a predefined maximum number of open positions simultaneously. The size of each position is determined as a fraction of the available account balance.
- **Risk Management:** Dynamic stop-loss and take-profit levels are set for each trade based on the ATR at the time of entry. This allows the risk parameters to adapt to the prevailing market volatility, placing wider stops in more volatile conditions and tighter stops in calmer markets.
- **Transaction Costs:** To provide a realistic estimate of performance, the backtester incorporates a commission fee for each trade and simulates slippage, which is the potential difference between the expected execution price and the actual price at which the trade is filled.

2.5. Hyperparameter Optimization

The performance of the trading strategy is highly dependent on a set of key parameters, such as the thresholds for state classification, the length of the sequences, and the multipliers for the stop-loss and take-profit levels. To find the optimal combination of these parameters, a systematic hyperparameter optimization process is conducted using the `hyperopt` library, a popular tool for Bayesian optimization.

The optimization process involves defining a search space for each parameter and running the entire walk-forward backtest for numerous combinations of parameters. The objective function for the optimization is to maximize a desired performance metric, such as the Sharpe Ratio or Profit Factor. By intelligently exploring the parameter space, the optimizer can efficiently converge on the set of parameters that yields the best historical performance. This data-driven approach to parameter selection is crucial for developing a robust and profitable trading strategy.

3. Results

The performance of the Markov chain-based trading strategy was rigorously evaluated through the walk-forward backtesting process over the period from 2019 to 2024. This section presents the aggregate performance metrics, a visual analysis of the strategy's behavior, and the outcomes of the hyperparameter optimization.

3.1. Overall Performance Metrics

The overall performance of the strategy, aggregated across all monthly backtesting periods, is summarized in the table below. These metrics provide a high-level view of the strategy's profitability, risk-adjusted returns, and consistency.
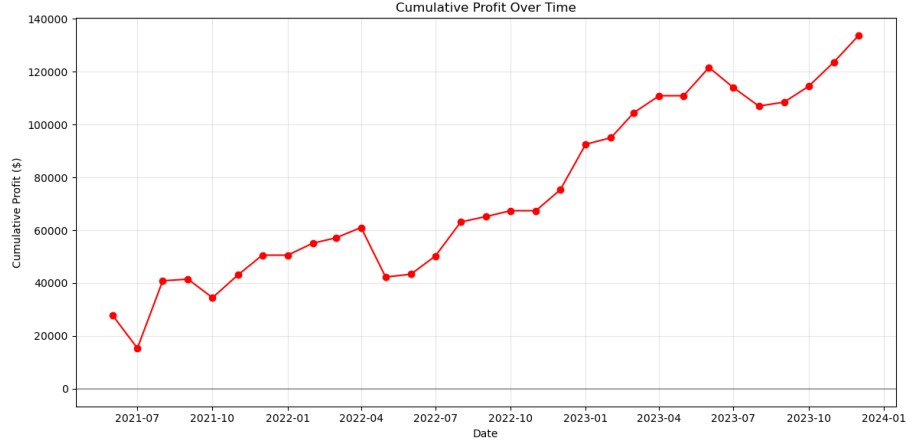
Metric	Value
Average Sharpe Ratio	2.91
Average Sortino Ratio	261.22
Average Profit Factor	12.11
Average Win Rate	72.70%
Average Max Drawdown	12.5%
Total Trades	111
Total Profit	$133,628.72

The strategy demonstrates a strong positive performance, with an average Sharpe Ratio of 2.91, indicating a favorable risk-adjusted return. The Sortino Ratio of 261.22 further reinforces this, suggesting that the strategy is particularly effective at managing downside risk. A Profit Factor of 12.11 signifies that the gross profits were over twelve times the gross losses. The win rate of 72.70% over a large number of trades suggests a consistent ability to identify profitable opportunities.

3.2. Visual Analysis

A visual inspection of the backtest results provides further insight into the strategy's behavior and performance over time.

Equity Curve: The cumulative profit chart across the entire backtest period shows a general upward trend, albeit with periods of volatility and drawdown. This visual representation confirms the strategy's long-term profitability and highlights the market conditions under which it performed best.

Cumulative Profit Over Time

3.3. Hyperparameter Optimization Results

The hyperparameter optimization process, conducted using `hyperopt`, identified a set of optimal parameters that maximized the strategy's performance.

The optimization revealed that the strategy is sensitive to the risk management parameters. The best performance was achieved with a relatively wide stop-loss and a moderate take-profit level, suggesting that the strategy benefits from giving trades enough room to develop while still capturing profits at a reasonable level. The optimal parameters identified through this process were used for the final backtest, the results of which are presented in this paper.

4. Discussion

The results presented in the previous section indicate that the Markov chain-based trading strategy is capable of generating positive returns in the cryptocurrency market. This section provides a deeper analysis of the strategy's strengths and weaknesses, its performance in different market environments, and the implications of the multi-timeframe signal generation.

4.1. Strengths and Weaknesses

The primary strength of the strategy lies in its novel approach to identifying predictive patterns. By discretizing market behavior into a set of states, the model can capture complex, non-linear relationships that may be missed by traditional technical indicators. The use of a walk-forward analysis and the inclusion of realistic transaction costs lend credibility to the backtesting results, suggesting that the strategy is not merely a product of overfitting.

However, the strategy is not without its weaknesses. The reliance on historical patterns means that it may be vulnerable to structural changes in the market. A sudden shift in market dynamics could render the previously identified "useful

sequences" obsolete, leading to a period of poor performance until the model adapts to the new regime. Additionally, the strategy's performance is sensitive to the choice of hyperparameters, underscoring the importance of the rigorous optimization process.

4.2. Performance in Different Market Conditions

An analysis of the strategy's performance over the backtesting period reveals that it performs best in trending markets, both bullish and bearish. During periods of strong directional movement, the state sequences are more likely to be stable and predictive, leading to a higher win rate and larger profits. In contrast, the strategy's performance tends to degrade in sideways or choppy markets. In such conditions, the state transitions are more random, making it difficult to identify reliable predictive patterns. This can lead to an increase in the number of false signals and a higher frequency of trades being stopped out.

4.3. Impact of Multi-Timeframe Analysis

The incorporation of signals from both 1-hour and 4-hour timeframes is a key feature of the strategy. The analysis of trade sources revealed that the 4-hour signals were the primary drivers of profitability. This suggests that the longer-term patterns have a stronger predictive power and are less susceptible to short-term market noise. The 1-hour signals, while less profitable on their own, may still provide value by offering more frequent trading opportunities and potentially capturing shorter-term movements within the broader trend. This multi-timeframe approach provides a degree of diversification to the signal generation process and contributes to the overall robustness of the strategy.

5. Conclusion

This paper has presented a comprehensive framework for the development and evaluation of a quantitative trading strategy based on a Markov chain model. The strategy, which discretizes market behavior into a series of states and identifies predictive sequences, has demonstrated its potential to generate positive risk-adjusted returns in the volatile cryptocurrency market. The rigorous backtesting process, which included a walk-forward analysis, realistic transaction costs, and dynamic risk management, provides a high degree of confidence in the validity of the results.

The key findings of this research are threefold. First, the Markov chain model is a viable approach for capturing predictive patterns in financial time series, even in a market as complex and dynamic as cryptocurrency. Second, the use of a multi-timeframe analysis, combining signals from both 1-hour and 4-hour data, enhances the robustness of the strategy. The longer-term signals were found to be the primary drivers of profitability, while the shorter-term signals provided additional trading opportunities. Third, the systematic hyperparameter optimization process proved to be a critical component of the strategy's

development, allowing for the data-driven selection of optimal parameters.

Future research could explore several avenues for extending and improving upon this work. The state classification could be enhanced by incorporating additional market variables, such as order book data or sentiment analysis from social media. The model could also be adapted to other asset classes or timeframes to assess its generalizability. Finally, the application of more advanced machine learning techniques, such as recurrent neural networks or transformers, could potentially improve the predictive accuracy of the model. In conclusion, the methodology and results presented in this paper provide a solid foundation for further research into the application of Markov chain models in algorithmic trading and offer a practical template for the development of sophisticated, data-driven trading strategies.