Hidden Markov Model for Bitcoin Price Dynamics Modeling

Leveraging HMM for Trend-Following Strategies in Bitcoin

By: Brayan Mercado Fuentes

- DISCLAIMER -

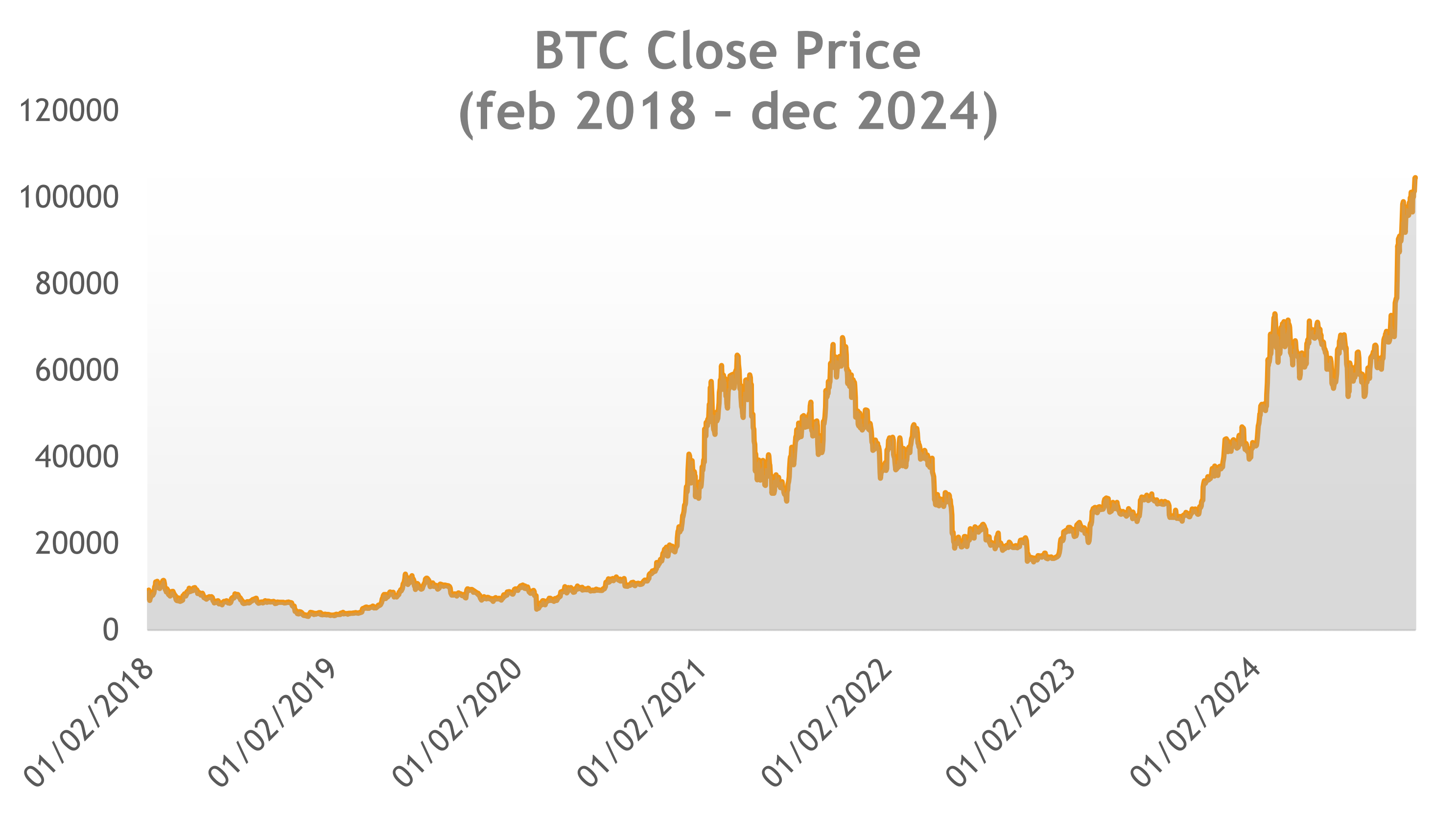
This project explores the potential application of Hidden Markov Models (HMM) to identify short-term trends in Bitcoin price movements and generate trading signals. Even if the findings suggest optimal results, this approach should never be used as the sole method for determining profitable trading signals. Trading decisions inherently involve significant risks and require consideration of diverse factors, including market conditions, fundamental analysis, and other technical indicators. This HMM should be regarded as a supplementary tool within a comprehensive trading strategy and not as a standalone predictor or guarantee of financial success.

Content Outline

Chapter 1: Introduction
1.1 Background and Motivation

- Cryptocurrencies, a revolutionary asset class, have emerged as a significant innovation in financial markets. Among these, Bitcoin, the first decentralized digital currency has garnered widespread attention due to its unique characteristics, such as decentralized governance, limited supply, and high liquidity. Despite its growing adoption as a store of value and medium of exchange, Bitcoin exhibits extreme price volatility compared to traditional assets, making it an intriguing subject for financial modeling and trading strategies.

Chapter 1: Introduction
1.1 Background and Motivation



*In the last 6 years the Bitcoin price has increased over 1062%

- Bitcoin's price movements are driven by a combination of factors, including macroeconomic trends, market sentiment, and speculative behavior. This complexity introduces significant challenges for traditional financial models, which often assume linearity or stationarity in price dynamics.
- The unpredictable nature of Bitcoin's price volatility presents both risks and opportunities: while it deters some investors, it also offers significant potential to capitalize on price inefficiencies and trend following.

Chapter 1: Introduction
1.2 Problem Statement

- The primary focus of this project is to address Bitcoin's price dynamics complexity by answering the following research problem:

How can a Hidden Markov Model (HMM) be effectively employed to identify short-term trends in Bitcoin price movements and generate profitable trading signals?

Chapter 1: Introduction
1.3 Research Objectives

- **Development of a Robust Discrete Hidden Markov Model (HMM):** Design and implement a discrete HMM tailored to Bitcoin's price data, with a focus on short-term trends. The model will utilize discrete observation symbols to capture the state dynamics.

- **Estimation of Hidden States and Probabilities:** Identify and estimate hidden market states (e.g., uptrend, downtrend, and sideways trends) based on observed sequences of discretized features. Allow the HMM to learn the underlying transition and emission probabilities governing these states.

Chapter 2: Literature Review

2.1 Financial Time Series Analysis

- The analysis and prediction of financial time series, such as asset prices, have long been critical areas of study in quantitative finance. Traditional statistical models have dominated this domain, offering tools for analyzing trends, volatility, and patterns in financial data. However, these methods often face limitations in capturing the complexities of nonlinear and nonstationary behaviors inherent in modern financial markets. In the following slides a comparison table for most popular financial models is shown as a general review of the concepts, strengths and weaknesses.

Chapter 2: Literature Review
2.1.1 Traditional Models & Limitations of Traditional Models

MODEL	STRENGTHS	WEAKNESSES
Linear Regression	- Easy to implement and interpret. - Useful for modeling relationships between variables. - Works well with linear trends.	- Assumes linearity and normally distributed errors. - Sensitive to outliers and multicollinearity.
ARIMA (Autoregressive Integrated Moving Average)	- Effective for time series data with trends and seasonality. - Provides forecasts based on past values and error terms.	- Requires stationary data. - Can struggle with complex, non-linear dependencies or sudden shocks.

Chapter 2: Literature Review
2.1.1 Traditional Models & Limitations of Traditional Models

MODEL	STRENGTHS	WEAKNESSES
Markov Models	- Suitable for modeling sequential data and state transitions. - Explains probabilities of transitions in a system.	- Assumes the Markov property. - Limited in modeling long-term dependencies.
GARCH (Generalized Autoregressive Conditional Heteroskedasticity)	- Excellent for modeling and forecasting volatility. - Handles time-varying variance effectively.	- Limited to modeling volatility, not direct financial outcomes. - Sensitive to model misspecifications.
VAR (Vector AutoRegression)	- Captures relationships between multiple time series. - Good for multivariate time series forecasting.	- Assumes stationarity of time series. - Computationally intensive for high-dimensional data.

Chapter 2: Literature Review
2.1.1 Traditional Models & Limitations of Traditional Models

MODEL	STRENGTHS	WEAKNESSES
Random Forest	- Handles non-linearity and high-dimensional data well. - Robust to overfitting in most cases. - Offers feature importance insights.	- Can be computationally intensive. - Less interpretable compared to simpler models. - May require parameter tuning.
Neural Networks	- Highly flexible for capturing complex patterns. - Suitable for large datasets with non-linear relationships. - Can integrate unstructured data like text or images.	- Computationally expensive. - Prone to overfitting with insufficient data. - Requires careful tuning of architecture and parameters.

Chapter 2: Literature Review
2.2 Hidden Markov Models

- 2.2.1 Theoretical Foundations of Markov Models and HMM

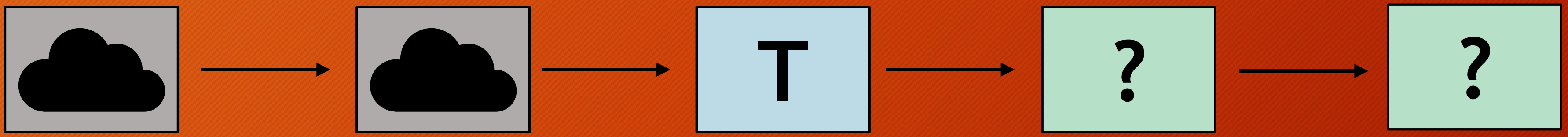
A Markov model is a model used to represent systems that transition from one state to another, fundamental part of any Markov model is that the future state of a process depends only on its current state and not on the sequence of past states.

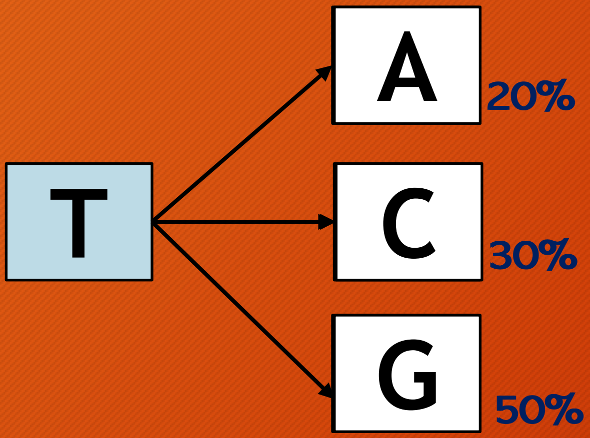
This arises because the state is conditionally independent of the history of states, given our immediate predecessor, there's no dependence on the previous history that isn't completely contained in the information that comes from just previous state.

Chapter 2: Literature Review
2.2.1 Theoretical Foundations of Markov Models and HMM

Basic Example of "The Markov Property"

Let's assume the following distribution of letters in a word:





There's some distribution over the letters that follow, so given that we have a "T" previously, the next letter can be an "A", "C" or "G", and there's going to be some probability of what that is. If we're just going to predict the maximum probability next letter (ceteris paribus), then "G" is the most likely letter to come next with a 50% probability.

Chapter 2: Literature Review
2.2.1 Theoretical Foundations of Markov Models and HMM

Key Concepts of a Markov Model

- **States:** The distinct configurations or conditions that the system can occupy or "be in".
- **Transition Probabilities:** The likelihood of moving from one state to another, often represented in a **transition matrix.**
- **Markov Property:** The future state depends only on the present state, not on the history of past states.

$$P(S_{t+1} = j \mid S_t, S_{t-1}, \dots, S_0) = P(S_{t+1} = j \mid S_n = i)$$

A full probabilistic description of a given system would require knowledge of the current state and all previous states. But not so, if we represent it as a Markov Model, due to this property.

Chapter 2: Literature Review
2.2.1 Theoretical Foundations of Markov Models and HMM

Key Concepts of a Markov Model

- **Transition Matrix:** Represents the probabilities of transitioning from one state to another in a stochastic process.

If we assume that our transition probabilities don't vary over time, then:

$$A = \{a_{ij}\} = P(S_t = j \mid S_{t-1} = i) \quad 1 \leq i, j \leq N$$

some constraints…

$$a_{ij} \geq 0$$

- Any transition probability has got to be greater or equal to zero.

$$\sum_{j=1}^{N} a_{ij} = 1$$

- The sum of all possible destinations (transitions) from state i to state j, must sum up to 1.

Chapter 2: Literature Review
2.2.1 Theoretical Foundations of Markov Models and HMM

Hidden Markov Model

The nature of an observable or discrete Markov Model is being able to observe in which state is the system in, that means that the output is the sequence of states themselves. For example, weather conditions or rolling a dice, even traffic lights! Those are observable states.

But there are lots of cases where we can't observe the state we are interested in. We can only see the effect of being in the state. For example, we cannot directly "see" someone's emotional state, but we might infer it from their words or actions.

Hidden Markov Models (HMMs) are one kind of Markov Models i.e., statistical models, designed to infer hidden states from observed sequences.

Chapter 2: Literature Review
2.2.1 Theoretical Foundations of Markov Models and HMM

Hidden Markov Model

- A Basic Discrete HMM has:
  - *N* states
    - They are hidden, but they typically correspond to something we know about the world.
    - States are generally connected, i.e., ergodic.
      - Ergodicity: In a ergodic Markov Model, every state is reachable from every other state in the long run.

$$S = \{ s_1, s_2, s_3, s_4, s_5, \dots, s_n \}$$

  - *M* observable symbols (features)
    - They are what can be observed
    - They are a discrete alphabet

$$V = \{ v_1, v_2, v_3, v_4, v_5, \dots, v_m \}$$

- State transition matrix "*A*"

$$A = \{a_{ij}\} = P(S_t = j \mid S_{t-1} = i) \quad 1 \le i, j \le N$$

$$a_{ij} \ge 0$$

$$\sum_{j=1}^{N} a_{ij} = 1$$

- Emission (Observation) probabilities

$$B = \{b_j(o)\}$$

$$b_j(o) = P(O\,t = o\,t \mid S\,t = s)$$

- Initial state probabilities:

$$\pi = \{\pi_i\}$$

$$\pi_i = P(S_0 = i) \quad 1 \le i \le N$$

Chapter 2: Literature Review
2.2.2 Advantages of HMMs over Traditional Models

- **Captures Hidden States:** Models unobservable factors influencing the data.
- **Handles Non-Stationarity:** Can adapt to data that changes over time without needing it to be stationary.
- **Captures Regime Shifts & Nonlinearity:** Can model different regimes and non-linear behaviors.
- **Models Complex Dependencies:** Handles complex relationships, in a simple way.
- **Stochastic Process Representation:** Treats the data as a random process, capturing uncertainty more effectively.
- **Lower Complexity and Computational Power Use:** HMMs are generally simpler to implement and interpret compared to Neural Networks (NNs) or Random Forests (RF).

It is certainly not the optimal choice, as neural networks and other deep learning and machine learning models typically outperform it.
However, **given my current capabilities, it is the most suitable option.**

Chapter 3: HMM Specification & Mathematical Formulation

Chapter 3: HMM Specification & Mathematical Formulation

- This chapter contains a detailed mathematical description of the underlying model.
- While this section provides an in-depth view of the model's structure, it is not necessary for a review at this stage.
- I have included it for those who may be interested in understanding the technical foundations and formulation behind the model.
- Please feel free to focus on the strategic and practical aspects of the presentation, which are outlined in the previous chapters.

GO TO CONTENT OUTLINE

GO TO CHAPTER 4: DATA AND METHODOLOGY

Hidden Markov Model for Bitcoin Price Dynamics Modeling

Leveraging HMM for Trend-Following Strategies in Bitcoin

Chapter 3: HMM Specification & Mathematical Formulation

By: Brayan Mercado Fuentes

1. Model Definition

Hidden States

Let the set of hidden states $S = \{s_1, s_2, s_3, s_4, s_5\}$, where:

- S1 = Strong Bull Market
- S2 = Moderate Bull Market
- S3 = Sideways/Neutral Market
- S4 = Moderate Bear Market
- S5 = Strong Bear Market

These states represent the latent market regimes that are not directly observable.

Observable Variables

At time t, the observable features vector $O_t = \begin{bmatrix} Log\ Return_t, \\ RSI_t, \\ MACD_t, \\ SMA\ Crossover_t, \\ Volume_t, \\ FnG\ Index_t \\ \cdots \end{bmatrix}$

For simplicity, we assume that O_t is a d-dimensional vector.

1. Model Definition

Relationship Between Hidden States and Observables

The HMM assumes:

The hidden state at time t, St , depends only on the hidden state at time t−1:

$$P(S_t \mid S_{t-1}, S_{t-2}, \dots, S_1) = P(S_t \mid S_{t-1})$$

The observable variable Ot depends only on the current hidden state St

$$P(O_t \mid S_t, S_{t-1}, S_{t-2}, \dots, S_1) = P(O_t \mid S_t)$$

Thus, the joint probability of the sequence of hidden states S=(S1,S2,…,ST) and observations O=(O1,O2,…,OT) is given by:

$$P(S, O) = P(S_1) \prod_{t=2}^{T} P(S_t | S_{t-1}) \prod_{t=1}^{T} P(O_t | S_t)$$

2. Transition Probability

Transition Matrix

Define the transition probability matrix A=[aij], where:

$$a_{ij} = P\left(S_t = s_j \middle| S_{t-1} = s_i\right) \quad 1 \leq i,j \leq N$$

$$A = \{a_{ij}\} = \begin{bmatrix} P(S_1 \to S_1) & P(S_1 \to S_2) & P(S_1 \to S_3) & P(S_1 \to S_4) & P(S_1 \to S_5) \\ P(S_2 \to S_1) & P(S_2 \to S_2) & P(S_2 \to S_3) & P(S_2 \to S_4) & P(S_2 \to S_5) \\ P(S_3 \to S_1) & P(S_3 \to S_2) & P(S_3 \to S_3) & P(S_3 \to S_4) & P(S_3 \to S_5) \\ P(S_4 \to S_1) & P(S_4 \to S_2) & P(S_4 \to S_3) & P(S_4 \to S_4) & P(S_4 \to S_5) \\ P(S_5 \to S_1) & P(S_5 \to S_2) & P(S_5 \to S_3) & P(S_5 \to S_4) & P(S_5 \to S_5) \end{bmatrix}$$

Where:

$P(Si \to Sj)$ represents the probability of transitioning from state Si to state Sj.

The matrix A satisfies: $a_{ij} \geq 0 \quad \text{and} \quad \sum_{j=1}^{5} a_{ij} = 1$

3. Emission Probability

Since the features are discretized, the emission probabilities are the probabilities of observing a particular discrete feature vector Ot given the hidden state St. We define the emission probability as:

$$P(\,O_t\ \mid S_t = s_i\,) = P(O_{t,1} = o_{t,1}, O_{t,2} = o_{t,2}, \ldots, O_{t6}|S_t = s_i),$$

where ot, j is the observed category for the j-th feature at time t. In the discrete setting, this is simply a multinomial distribution over the possible categories for each feature.

For simplicity, we assume that the features are conditionally independent given the hidden state St . Therefore, the emission probability factorizes:

$$P(\,O_t\ \mid S_t = s_i\,) = P\big(O_{t,1}\big|S_t = s_i\big) * P\big(O_{t,2}\big|S_t = s_i\big) *, \ldots, * P(O_{t6}|S_t = s_i),$$

Each conditional probability $P(Ot,j|St{=}si)$ is modeled as a multinomial distribution over the categories of feature j.

Define:

$$P\big(O_{t,j} = o\big|S_t = s_i\big) = p_{ij}^{(o)}$$

where $pij(o)$ is the probability of observing category o for feature j given that the system is in hidden state si.

Thus, the emission probabilities are:

$$P(O_t|S_t = s_i) = \prod_{j=1}^{6} P\big(O_{t,j}\big|S_t = s_i\big) = \prod_{j=1}^{6} p_{ij}^{(O_{i,j})}$$

4. Observation Model
The likelihood of the entire observation sequence O=(O1,O2,…,OT) given the hidden state sequence S is:

$$P(O|S) = \prod_{t=1}^{T} P(O_t|S_t)$$

which combines the emission probabilities for each timestep.

5. Parameter Estimation (Baum-Welch Algorithm)
To estimate the parameters λ=(A,{pij(o)},π), we use the Baum-Welch algorithm, an Expectation-Maximization (EM) method.

6. Inference (Viterbi Algorithm)
To estimate the most likely sequence of hidden states S=(S1,S2,…,ST) given the observation sequence O, we use the Viterbi Algorithm.

Chapter 4: Data and Methodology
4.1 Data Description

This chapter outlines the data sources, preprocessing methods, and actual development of the model in Anaconda3 environment.

- The primary dataset consists of historical Bitcoin price data collected from cryptodatadownload.com on their Bitstamp data map collection.

  From this source that provides daily OHLC and Volume BTC/USD raw data from 2014-11-28 to 2024-12-16; I've calculated the model features:

**Most features or technical analysis indicators were calculated using pandas_ta library*

- **Logarithmic Returns:** is a measure of the percentage change in the price of an asset, expressed using the natural logarithm. It is widely used in finance for analyzing returns because it ensures time consistency and is additive over time.

$$R_{log} = ln\left(\frac{P_t}{P_{t-1}}\right)$$

Chapter 4: Data and Methodology
4.1 Data Description

- **Relative Strength Index:** RSI is a momentum oscillator used to measure the speed and change of price movements. It oscillates between 0 and 100 and helps identify overbought or oversold conditions in the market.

$$RSI = 100 - \frac{100}{1 + RS}$$

$$RS = \frac{Average\ Gain\ over\ n\ periods}{Average\ Loss\ over\ n\ periods}$$

- **Moving Average Convergence Divergence:** MACD is a trend-following momentum indicator that shows the relationship between two moving averages of an asset's price.

$$MACD = EMA_{short} - EMA_{long}$$

Chapter 4: Data and Methodology
4.1 Data Description

- **Simple Moving Average (SMA):** is the average of a selected range of prices (closing prices, typically) over a defined number of periods.

$$SMA = \frac{1}{n}\sum_{i=1}^{n} P_i$$

- **Volumen Change:** Volume change measures the change in trading volume between two periods, often used to gauge market activity and investor interest.

$$Volume\ Change\ (\%) = \frac{Volume_t - Volume_{t-1}}{Volume_{t-1}} * 100$$

Chapter 4: Data and Methodology
4.1 Data Description

- **Crypto Fear & Greed Index:** this unofficial Fear & Greed Index by alternative.me, gathers data from five different sources. Each data point is valued the same as the day before in order to visualize a meaningful progress in sentiment change of the crypto market.

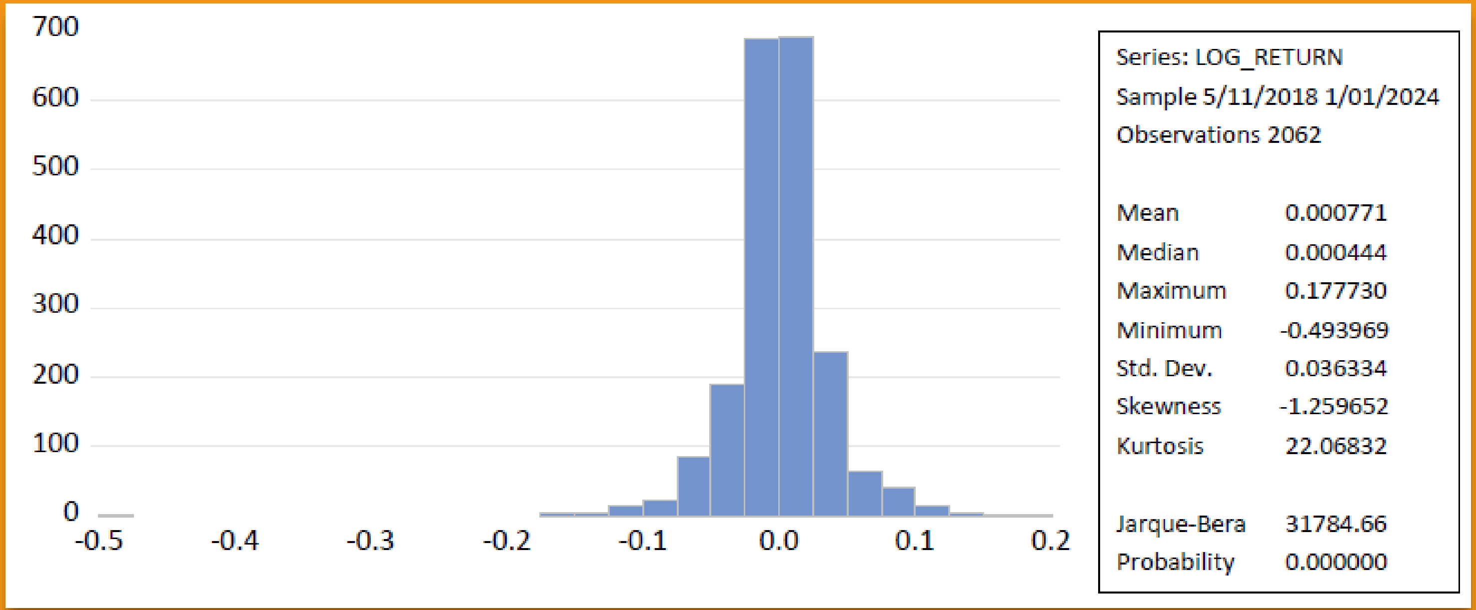
The composite index:

- **Volatility (25%):** measuring the current volatility and max. drawdowns of bitcoin and compare it with the corresponding average values of the last 30 days and 90 days.
- **Market Momentum/Volume (25%):** measuring the current volume and market momentum (again in comparison with the last 30/90-day average values) and put those two values together.
- **Social Media (15%):** gathers and count posts on various hashtags for Bitcoin and check how fast and how many interactions they receive in certain time frames. An unusual high interaction rate results in a grown public interest in the coin.
- **Surveys (15%)…**
- **Dominance (10%)…**
- **Trends (10%)…**

For more details review their website at https://alternative.me/crypto/fear-and-greed-index/

Chapter 4: Data and Methodology
4.1 Data Description

- The described and used data exhibits Bitcoin daily price from February 2018 to December 2024, having 2511 observations in total.
- Before developing the HMM I used a linear regression model between Bitcoin and each feature, just to analyze their statistical properties, linear dependence or multicollinearity patterns.
- The histogram represents the distribution of Bitcoin's log returns and it shows a strong peak around 0, indicating that most log returns are close to 0.
- The Jarque-Bera statistic is extremely high (31784.66), with a probability value of 0.000000. This strongly suggests that the log returns do not follow a normal distribution.



Chapter 4: Data and Methodology
4.1 Data Description

Correlation

	CLOSE	FNG_VALUE	LOG_RETU	MA50	MA100	MACD	OPEN_CLO	VOLUME_C
CLOSE	1.000000	0.285988	0.023525	0.965725	0.918548	0.216191	0.035303	-0.008167
FNG_V	0.285988	1.000000	0.050615	0.099164	-0.017129	0.729675	0.033473	-0.004883
LOG_R	0.023525	0.050615	1.000000	-0.029514	-0.034920	0.072799	0.789282	-0.030550
MA50	0.965725	0.099164	-0.029514	1.000000	0.978754	-0.036289	-0.031165	-0.006398
MA100	0.918548	-0.017129	-0.034920	0.978754	1.000000	-0.142742	-0.037610	-0.005949
MACD	0.216191	0.729675	0.072799	-0.036289	-0.142742	1.000000	0.089398	-0.000366
OPEN	0.035303	0.033473	0.789282	-0.031165	-0.037610	0.089398	1.000000	-0.034440
VOLUM	-0.008167	-0.004883	-0.030550	-0.006398	-0.005949	-0.000366	-0.034440	1.000000

- This correlation matrix provides valuable insights into the relationships between key variables in Bitcoin trading. The Close Price (CLOSE) shows a strong positive correlation with both the 50-day and 100-day moving averages (MA50: 0.97, MA100: 0.92), highlighting their dependence on price trends. The Fear and Greed Index (FNG_VALUE) has a moderate correlation with the MACD (0.73), suggesting that market sentiment influences price momentum. Log Returns (LOG_RET) and Volume (VOLUME_C), however, exhibit negligible correlations with most variables, indicating limited interaction within this dataset. Additionally, the two moving averages (MA50 and MA100) are almost identical in behavior, as reflected by their very high correlation (0.98). These insights can guide the identification of key predictors and trends for Bitcoin price analysis.

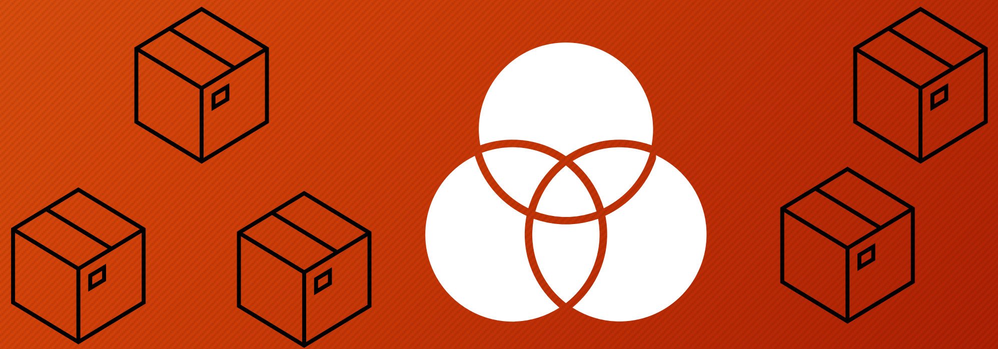
Chapter 4: Data and Methodology
4.2 Function Definitions

- The following section corresponds to the function definitions in the code for the HHM, the code as said before is written in Anaconda3 Environment with Python 3.9 the libraries used for the code are pandas, pandas' technical analysis, numpy, matplotlib, mplfinance and hmmlearn for the actual HMM development.
- The Functions are clustered into 6 different categories:
  - load_and_prepare_data
  - calculate_technical_indicators
  - discretize_features
  - plot_ohlc_with_indicators (optional)
  - train_hmm
  - plot_hidden_states
- Since most of the categories have been dealt with, it just remains to explain how the discretizing of the given features works for this model.

Chapter 4: Data and Methodology
4.2 Function Definitions

- **Discretize Features**

It means converting continuous data or numerical variables into discrete categories or bins. This process involves dividing the range of a continuous variable into intervals and assigning each interval a specific label, often to simplify the data or make it more suitable for certain types of analysis, such as categorical machine learning models or visualizations.



Chapter 4: Data and Methodology
4.2 Function Definitions

1. **Log Return:**

Log returns are often normally distributed. A simple way to discretize this feature is by splitting the range into several bins using fixed intervals that reflect typical market movements.

The bins are going to be:

- Bin 1: Extreme Negative Log Returns = $B_1 = \{x \in \mathbb{R}: -\infty < x \leq -5\%\}$
- Bin 2: Negative = $B_2 = \{x \in \mathbb{R}: -5\% < x \leq -0.5\%\}$
- Bin 3: Low/Neutral = $B_3 = \{x \in \mathbb{R}: -0.5\% < x \leq 0.5\%\}$
- Bin 4: Positive = $B_4 = \{x \in \mathbb{R}: 0.5\% < x \leq 5\%\}$
- Bin 5: Positive = $B_5 = \{x \in \mathbb{R}: x > 5\%\}$

Chapter 4: Data and Methodology
4.2 Function Definitions

2. RSI (Relative Strength Index):

RSI typically ranges from 0 to 100. To discretize this, I mapped the RSI value into predefined categories that usually represent different market conditions:

The bins are going to be:

- Bin 1: Oversold = $B_1 = \{x \epsilon \mathbb{R}: x < 30\}$
- Bin 2: Neutral = $B_2 = \{x \epsilon \mathbb{R}: 30 \leq x < 50\}$
- Bin 3: Bullish = $B_3 = \{x \epsilon \mathbb{R}: 50 \leq x \leq 65\}$
- Bin 4: Overbought = $B_4 = \{x \epsilon \mathbb{R}: x > 65\}$

Chapter 4: Data and Methodology
4.2 Function Definitions

3. MACD (Moving Average Convergence Divergence):

The MACD values can be discretized based on the MACD signal line and the

The bins are going to be:

- Bin 1: Strong Bearish = (MACD $<$ Signal Line, MACD significantly negative)
- Bin 2: Weak Bearish = (MACD $<$ Signal Line, but close to zero)
- Bin 3: Neutral =(MACD $\approx$ Signal Line, near zero)
- Bin 4: Weak Bullish (MACD $>$ Signal Line, but close to zero)
- Bin 5: Strong Bullish (MACD $>$ Signal Line, MACD significantly positive)

Chapter 4: Data and Methodology
4.2 Function Definitions

4. 50 and 100 SMA Crossover Strategy:

The crossover of the 50-day and 100-day moving averages typically indicates a change in trend direction. To discretize the feature into categories based on the direction and strength of the crossover.

- Bin 1: Strong Bearish (50-day SMA below 100-day SMA, and the gap is widening)
- Bin 2: Weak Bearish (50-day SMA below 100-day SMA, but the gap is narrowing)
- Bin 3: Neutral (50-day SMA near 100-day SMA)
- Bin 4: Weak Bullish (50-day SMA above 100-day SMA, but the gap is narrowing)
- Bin 5: Strong Bullish (50-day SMA above 100-day SMA, and the gap is widening)

Chapter 4: Data and Methodology
4.2 Function Definitions

5. Volume in USD:

Volume is often highly variable, and it can be useful to discretize it into categories that capture relative trading activity. A simple approach is to use quantiles or fixed thresholds based on historical average volume to determine "high" and "low" volumes.

- Bin 1: Low Volume (Below 25th percentile of historical volume)
- Bin 2: Below Average Volume (25th–50th percentile)
- Bin 3: Average Volume (50th-75th percentile)
- Bin 4: Above Average Volume (75th–90th percentile)
- Bin 5: High Volume (Above 90th percentile)

Chapter 4: Data and Methodology
4.2 Function Definitions

6. Fear and Greed Index:

The Fear and Greed index is already discrete, so I can directly map it to categories based on the index values.

- Bin 1: Extreme Fear (0–25)
- Bin 2: Fear (26–46)
- Bin 3: Neutral (47–54)
- Bin 4: Greed (55–75)
- Bin 5: Extreme Greed (76–100)

Chapter 4: Data and Methodology
Model Training

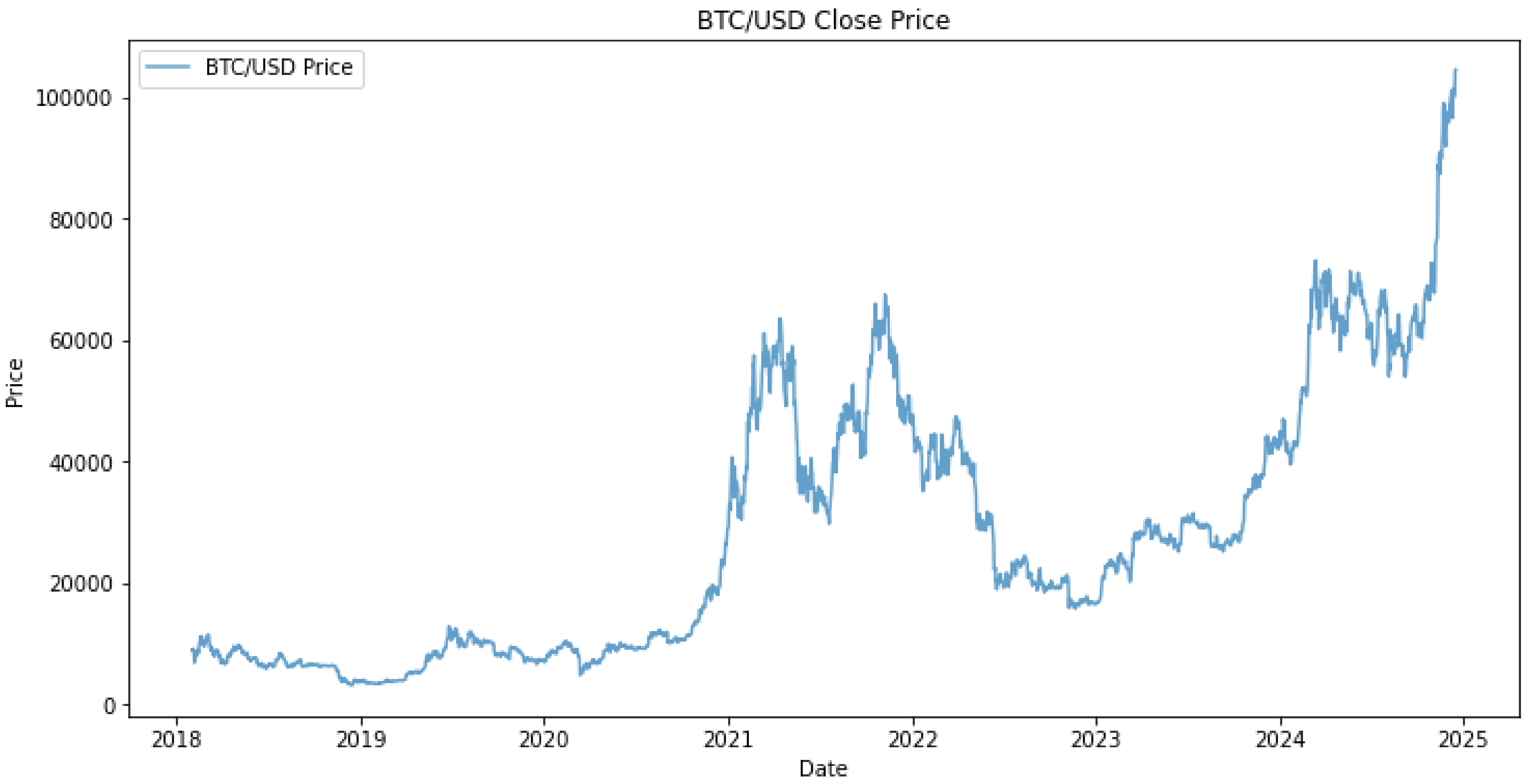
- For the model training I used the MultinomialHMM library in the hmmlearn github repository.
- I configured the model for the optimal number of hidden states (even if in the beginning it were only 5 hidden states), by evaluating the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) for determining the best model, which was 8 hidden components.
- The maximum number of iterations the model will use when fitting the data. Here, the model is set to perform up to 500 iterations during the training process. In practice, the model may converge (i.e., reach an optimal set of parameters) before completing all 500 iterations, but this is the upper bound.
- The random seed for reproducibility. By setting random state = 42, you ensure that the results of the model fitting process can be replicated.

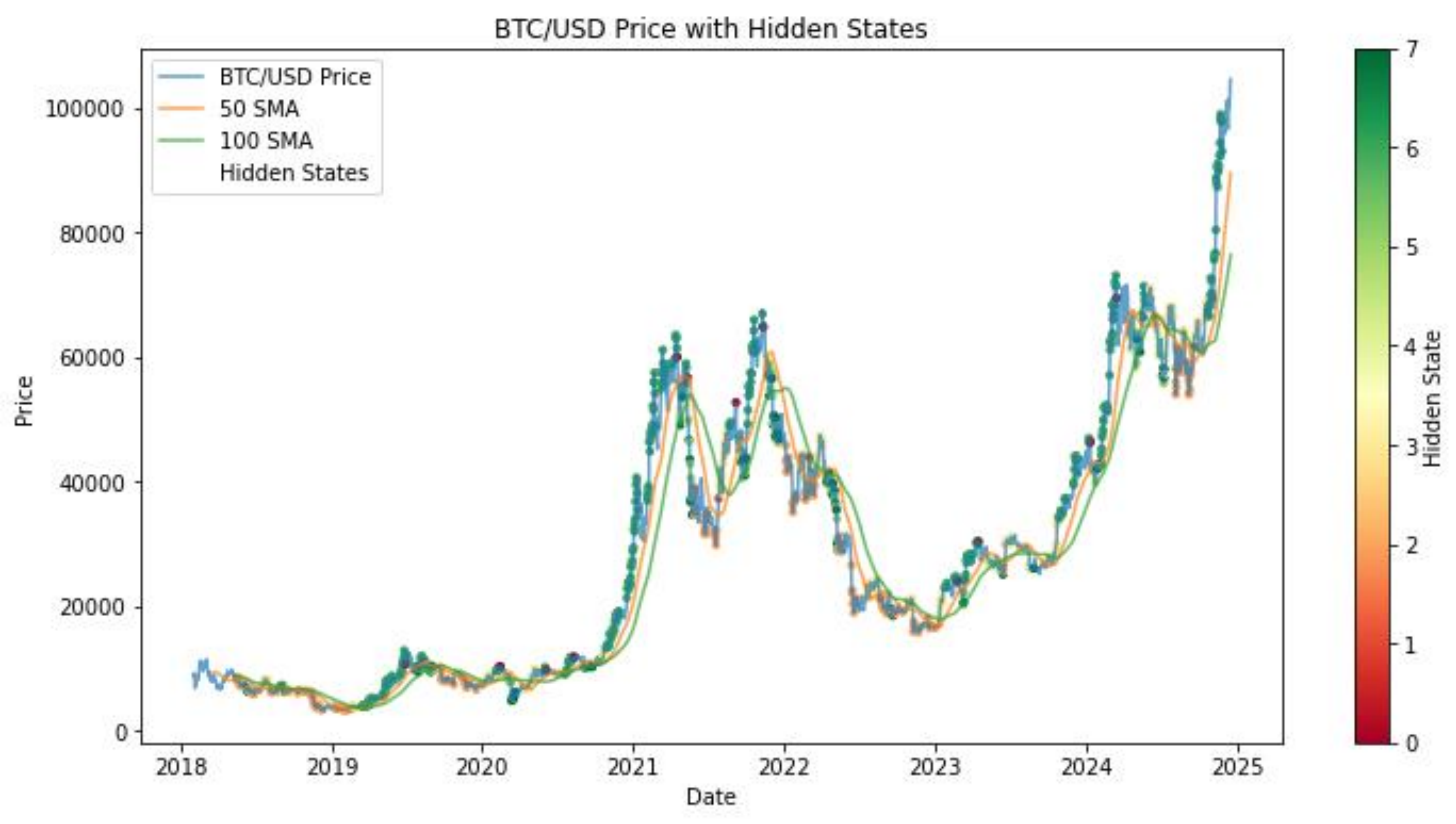
Chapter 4: Data and Methodology
Model Training

From here, I trained the model using the fit() function and then proceeded to used the predict() method to infer the most likely sequence of hidden states given the observations in my data.

The model uses the trained parameters (transition probabilities and emission probabilities) to estimate which hidden state is most likely at each point in the sequence of observations. This results in a sequence of hidden states that correspond to the observed data.

Also, we had to know exactly how the model is classifying the observations for each state, for this we can use the summary statistics of the states for each feature.


BTC/USD Close Price
BTC/USD Price
Price
100000
80000
60000
40000
20000
0
2018
2019
2020
2021
2022
2023
2024
2025
Date


BTC/USD Price with Hidden States
BTC/USD Price
50 SMA
100 SMA
Hidden States
Price
Date
Hidden State
0
20000
40000
60000
80000
100000
2018
2019
2020
2021
2022
2023
2024
2025
0
1
2
3
4
5
6
7

Chapter 4: Data and Methodology
Model Training

State-wise Summary for LOG-RETURN:

hidden_state	log_return mean	log_return std	log_return min	log_return max
0	-0.016647165	0.039033775	-0.097210181	0.01799087
1	0.017746916	0.03142987	-0.036293395	0.084097492
2	-0.006282753	0.03813108	-0.169041141	0.112530172
3	0.004060177	0.028894908	-0.098280257	0.123665419
4	-0.00832438	0.02689091	-0.139218715	0.046549623
5	0.005670662	0.038324308	-0.070227735	0.149559387
6	0.01115793	0.036113735	-0.147043425	0.17772991
7	-0.017732615	0.067343938	-0.49396946	0.087398235

Chapter 4: Data and Methodology
Model Training

State-wise Summary for RSI:

hidden_state	RSI mean	RSI std	RSI min	RSI max
0	59.47893542	8.190242267	45.1247483	70.24621826
1	60.73460535	4.250522144	54.52862539	70.61124431
2	36.74921779	7.96782779	9.63298186	54.53305351
3	63.64403349	9.185423266	37.63162503	89.2630629
4	40.18405802	6.807324908	18.93472229	52.21782052
5	39.4009315	6.398336956	26.2751642	50.7047298
6	70.73195005	8.753474469	48.84838534	89.2811718
7	38.75278083	7.687551863	15.67241977	60.92190377

Chapter 4: Data and Methodology
Model Training

State-wise Summary for Fear & Greed Index:

hidden_state	fng_value mean	fng_value std	fng_value min	fng_value max
0	65.07142857	17.9291953	17	83
1	39.9375	11.1861149	25	59
2	24.56730769	9.975627298	6	74
3	48.81184669	15.00743868	10	78
4	48.64393939	11.48123765	13	79
5	28.11764706	11.66946745	5	67
6	72.37007874	14.16526691	17	95
7	28.38372093	11.64693204	8	57

Chapter 4: Data and Methodology
Model Training

- **State 0 (Bullish to Bearish Turnover), i.e. SELL!**
  - Log Return: Mean = -1.66%, Std Dev = 3.90% → The negative mean suggests a tendency towards small losses or consolidations. The high standard deviation indicates significant volatility.
  - RSI: Mean = 59.48, Std Dev = 8.19 → RSI is generally neutral to slightly bullish, indicating a market that is neither overbought nor oversold.
  - Volume: Mean = 129,621,632, Std Dev = 86,717,236 → High volume, with significant fluctuations, indicating active market participation.
  - FNG Index: Mean = 65.07, indicating "Greed" which suggests bullish sentiment, though the market is also transitioning, hence the potential turning point.

Chapter 4: Data and Methodology
Model Training

- **State 1 (Weak Bullish to Neutral)**
  - Log Return: Mean = 1.77%, Std Dev = 3.14% → Positive mean log return with moderate volatility, suggesting a weak bullish market.
  - RSI: Mean = 60.73, Std Dev = 4.25 → Slightly bullish RSI, indicating the market is in a favorable condition but not overly extended.
  - Volume: Mean = 104,003,417, Std Dev = 78,397,523 → Above-average trading volume, but not excessive, implying moderate market interest.
  - FNG Index: Mean = 39.94, indicating "Fear", signaling that the market sentiment is cautious and may lean bearish in the near future.

Chapter 4: Data and Methodology
Model Training

- **State 2 (Neutral to Bearish)**
  - Log Return: Mean = -0.63%, Std Dev = 3.81% → Slightly negative log return with high volatility, suggesting an unstable or neutral market condition.
  - RSI: Mean = 36.75, Std Dev = 7.97 → RSI indicates oversold conditions, signaling potential for a rebound but also ongoing bearish pressure.
  - Volume: Mean = 77,153,314, Std Dev = 67,819,404 → Above-average volume with high volatility, showing that the market is still quite active.
  - FNG Index: Mean = 24.57, indicating "Fear", suggesting that investor sentiment is still bearish, which aligns with the bearish indicators from the other features.

Chapter 4: Data and Methodology
Model Training

- **State 3 (Bullish to Bullish with Weak Pullbacks), i.e. HODL**
  - Log Return: Mean = 0.41%, Std Dev = 2.89% → Positive mean log return with relatively low volatility, indicating a steady bullish trend.
  - RSI: Mean = 63.64, Std Dev = 9.18 → Strong bullish RSI, indicating that the market is in an overbought phase but is still maintaining upward momentum.
  - Volume: Mean = 104,003,417, Std Dev = 78,397,523 → High volume suggests that there is solid market participation, further supporting the bullish trend.
  - FNG Index: Mean = 48.81, indicating "Neutral", showing that investor sentiment is relatively balanced despite the bullish market conditions.

Chapter 4: Data and Methodology
Model Training

- **State 4 (Weak Bearish to Neutral)**
  - Log Return: Mean = -0.83%, Std Dev = 2.69% → Slightly negative log return with relatively low volatility, suggesting a mild bearish phase.
  - RSI: Mean = 40.18, Std Dev = 6.81 → RSI in the oversold region, indicating that the market is under bearish pressure but not yet in extreme conditions.
  - Volume: Mean = 59,617,313, Std Dev = 45,404,261 → The market shows above-average volume, but with high volatility.
  - FNG Index: Mean = 48.64, indicating "Neutral", which suggests investor sentiment is neither overly fearful nor greedy.

Chapter 4: Data and Methodology
Model Training

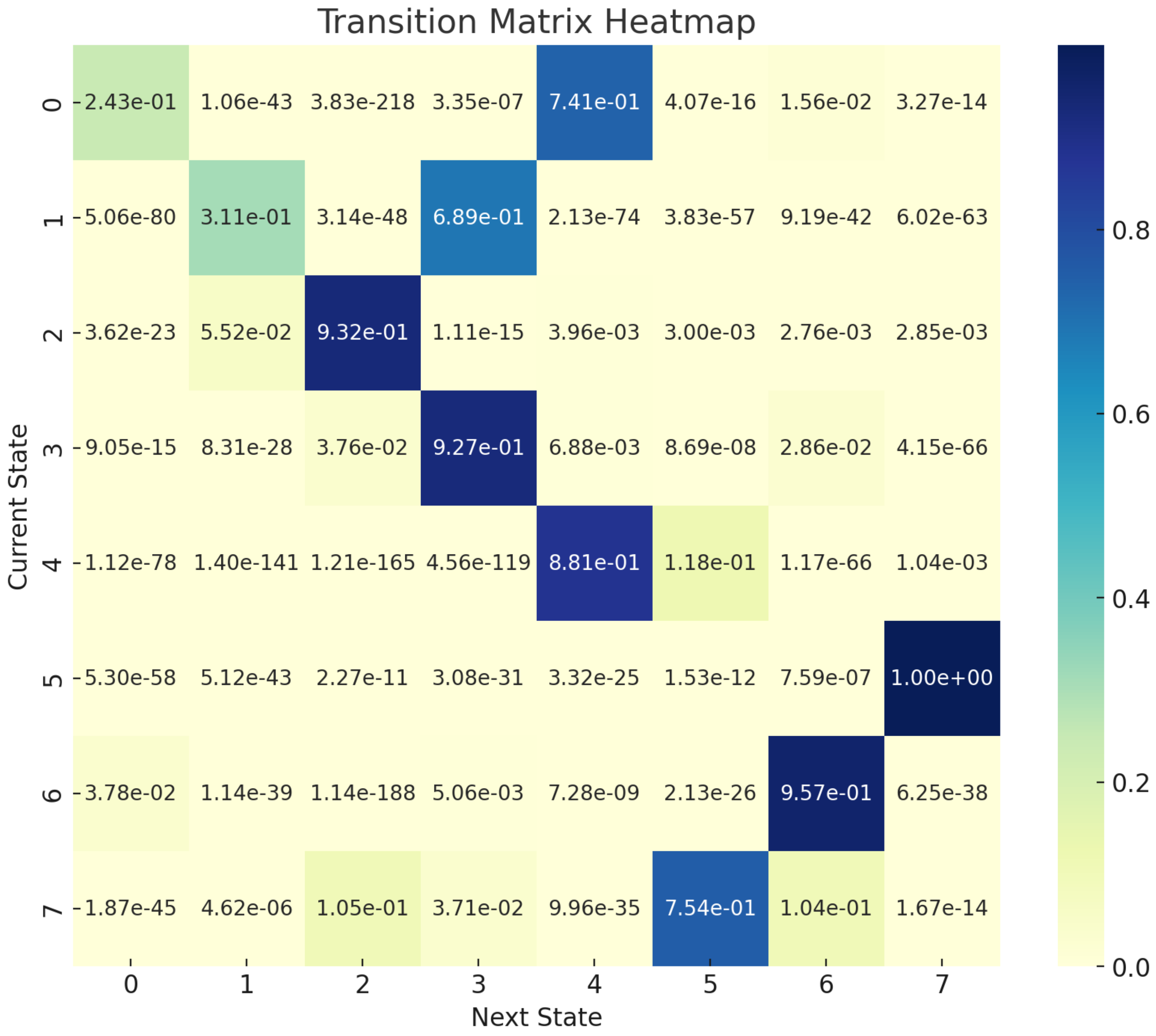
- **State 5 (Bullish Trend Start with Strong Momentum)**
  - Log Return: Mean = 0.57%, Std Dev = 3.83% → Positive log return with high volatility, indicating a strong bullish trend.
  - RSI: Mean = 39.40, Std Dev = 6.40 → RSI in a bullish zone but potentially under some bearish pressure.
  - Volume: Mean = 106,249,121, Std Dev = 88,548,648 → High volume signals significant market participation during this phase.
  - FNG Index: Mean = 28.12, indicating "Fear", suggesting a bearish sentiment despite bullish market activity.

Chapter 4: Data and Methodology
Model Training

- **State 6 (Strong Bullish Transition), i.e. FRENEZY**
  - Log Return: Mean = 1.12%, Std Dev = 3.61% → Positive log returns with moderate volatility, pointing to a sustained bullish trend.
  - RSI: Mean = 70.73, Std Dev = 8.75 → Overbought RSI, confirming strong bullish momentum.
  - Volume: Mean = 173,650,206, Std Dev = 178,029,845 → Extremely high volume, indicating massive market participation during this bull phase.
  - FNG Index: Mean = 72.37, indicating "Greed", confirming that market sentiment is strongly bullish.

Chapter 4: Data and Methodology
Model Training

- **State 7 (Bearish Transition to Bullish)**
  - Log Return: Mean = -1.77%, Std Dev = 6.73% → A larger negative log return with high volatility, suggesting a strong transition from bearish to bullish.
  - RSI: Mean = 38.75, Std Dev = 7.69 → RSI in the oversold region, suggesting that the market is near the bottom of a bearish phase and could be set for a reversal.
  - Volume: Mean = 131,058,556, Std Dev = 154,620,790 → High volume with significant fluctuations, suggesting that the market is in transition.
  - FNG Index: Mean = 28.38, indicating "Fear", which aligns with the bearish sentiment as the market prepares to reverse.

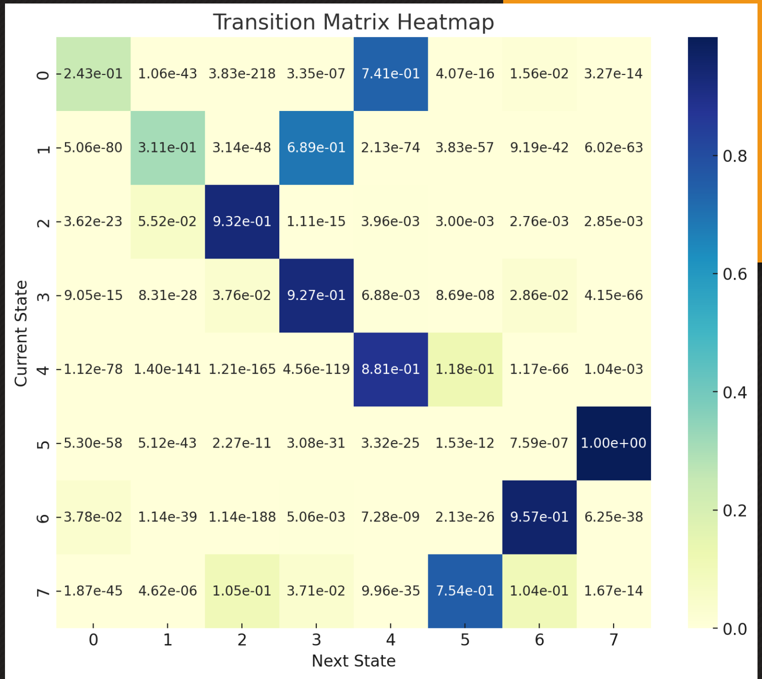


Transition Matrix Heatmap

The color intensity indicates the likelihood of transitioning from one state to another, with the colored areas showing higher probabilities. The values in each cell represent the exact transition probabilities between the current state (rows) and the next state (columns).

For instance: The highest transition probabilities are concentrated on the diagonal, indicating that states tend to stay within themselves (e.g., from state 0 to state 0, or from state 7 to state 7).The off-diagonal values are much smaller, reflecting the less probable transitions between states.

Transition Matrix Heatmap



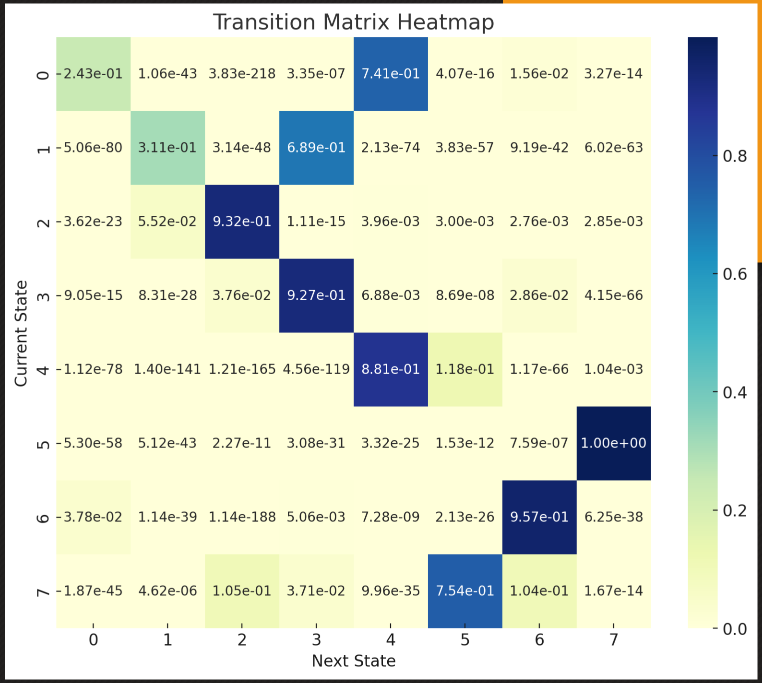
Row 0 (State 0):
Transition from state 0 (Bullish to Bearish Transition):

- The probability of staying in state 0 is **0.2431**.
- Transitioning to state 4 (Weak Bearish to Neutral) is **0.7413**, indicating that this is more likely after a bearish turning point.
- Very small probabilities exist for the other states.

Row 1 (State 1):
Transition from state 1 (Weak Bullish to Neutral):

- **0.3109** probability of staying in state 1.
- **0.6891** probability of moving to state 3 (Bullish to Bullish with Weak Pullbacks), suggesting that the market could transition.
- Very small probabilities for other transitions.

Transition Matrix Heatmap



Row 2 (State 2):
Transition from state 2 (Neutral to Bearish)

- **0.9322** probability of staying in state 2.
- Small probability of moving to other states.

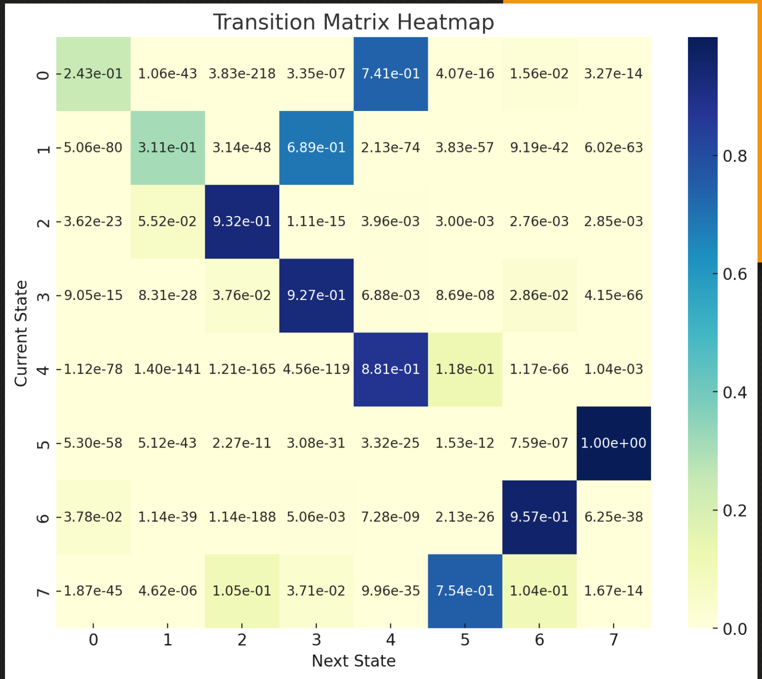
Row 3 (State 3):
Transition from state 3 (Bullish to Bullish with Weak Pullbacks):

- **0.9269** probability of staying in state 3.
- **0.0376** probability of moving to state 2 (Neutral to Bearish) ,suggesting a possible return to a milder neutral or bearish state

Row 4 (State 4):
Transition from state 4 (Weak Bearish to Neutral):

- **0.8807** probability of staying in state 4,
- **0.1182** probability of moving to state 5 (Bullish Trend Start with Strong Momentum), signaling a slight chance of the market moving into a bullish phase.

Transition Matrix Heatmap



Row 5 (State 5):
Transition from state 5 (Bullish Trend Start with Strong Momentum):

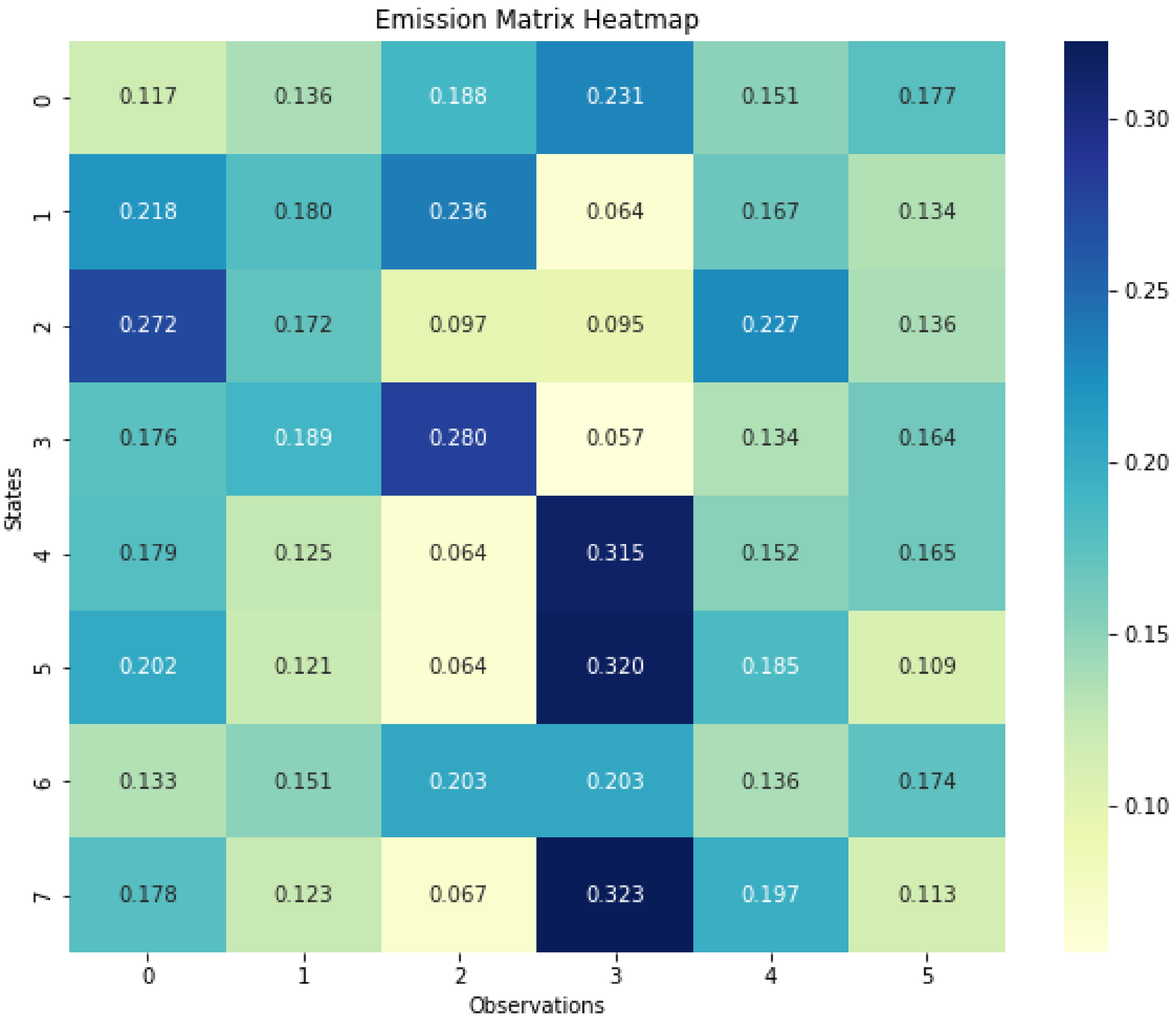
- **0.9999** probability of transitioning to state 7 (Bearish Transition to Bullish), indicating that, this state is likely the price movement when even during a bearish trend, they may present some uptrends, and then a turning point.

Row 6 (State 6):
Transition from state 6 (Strong Bullish Transition):

- **0.9571** probability of staying in state 6, indicating a strong bullish continuation.

Row 7 (State 7):
Transition from state 7 (Bearish Transition to Bullish):

- **0.7537** probability of transitioning to state 5 (Bullish Trend Start with Strong Momentum),
- **0.1050** probability of moving to state 2 (Neutral to Bearish),
- **0.1040** probability of moving to state 6 (Strong Bullish Transition)



Emission Matrix Heatmap

The color intensity indicates the likelihood of observing a specific event given a hidden state, with the colored areas showing higher probabilities. The values in each cell represent the exact emission probabilities between the current state (rows) and the observation “feature” (columns).

These are not as relevant as the transition probabilities, since the model is the one classifying them, so maybe for us it does not make sense, but for the model, does.

Chapter 5: Model Implementation, Results & Limitations.

- The implementation of a Hidden Markov Model (HMM) for Bitcoin price dynamics involves not only identifying the hidden states but also leveraging these states to create actionable trading signals.
- By associating market states with specific trends and sentiments, we can develop a systematic trend-following strategy.
- This section outlines the methodology for transforming HMM-derived states into trading signals and integrating these signals into a trading framework.

Chapter 5: Model Implementation, Results & Limitations.

Each hidden state identified by the HMM represents a unique market condition, characterized by specific attributes such as log returns, Relative Strength Index (RSI), trading volume, and sentiment indicators like the Fear & Greed (FNG) Index.

- The strategy I’ll apply is the following:
  - If the model state results to be in SATE 0, it results in a "Sell" Signal, assigned to potential bearish trend reversal to mitigate losses or take advantage of downward trends through short positions.

  - If the model state results to be in SATE 7, it results in a “Buy" Signal, assigned to a potential bullish trend reversal to capitalize on upward trends.

  - If the model state results to be in SATE 1, it results in a “Hold" Signal, assigned to a neutral state to avoid unnecessary trades during periods of uncertainty, in practice, I would use more tools to determine if I should exit or hold during this time, and always use a stop-loss.

  - If the model state results to be in SATE 2, it results in a “Sell" Signal with a more cautions approach going to CASH due to the slight persistent downtrends, assigned to a neutral-bearish state to avoid unnecessary risks.

Chapter 5: Model Implementation, Results & Limitations.

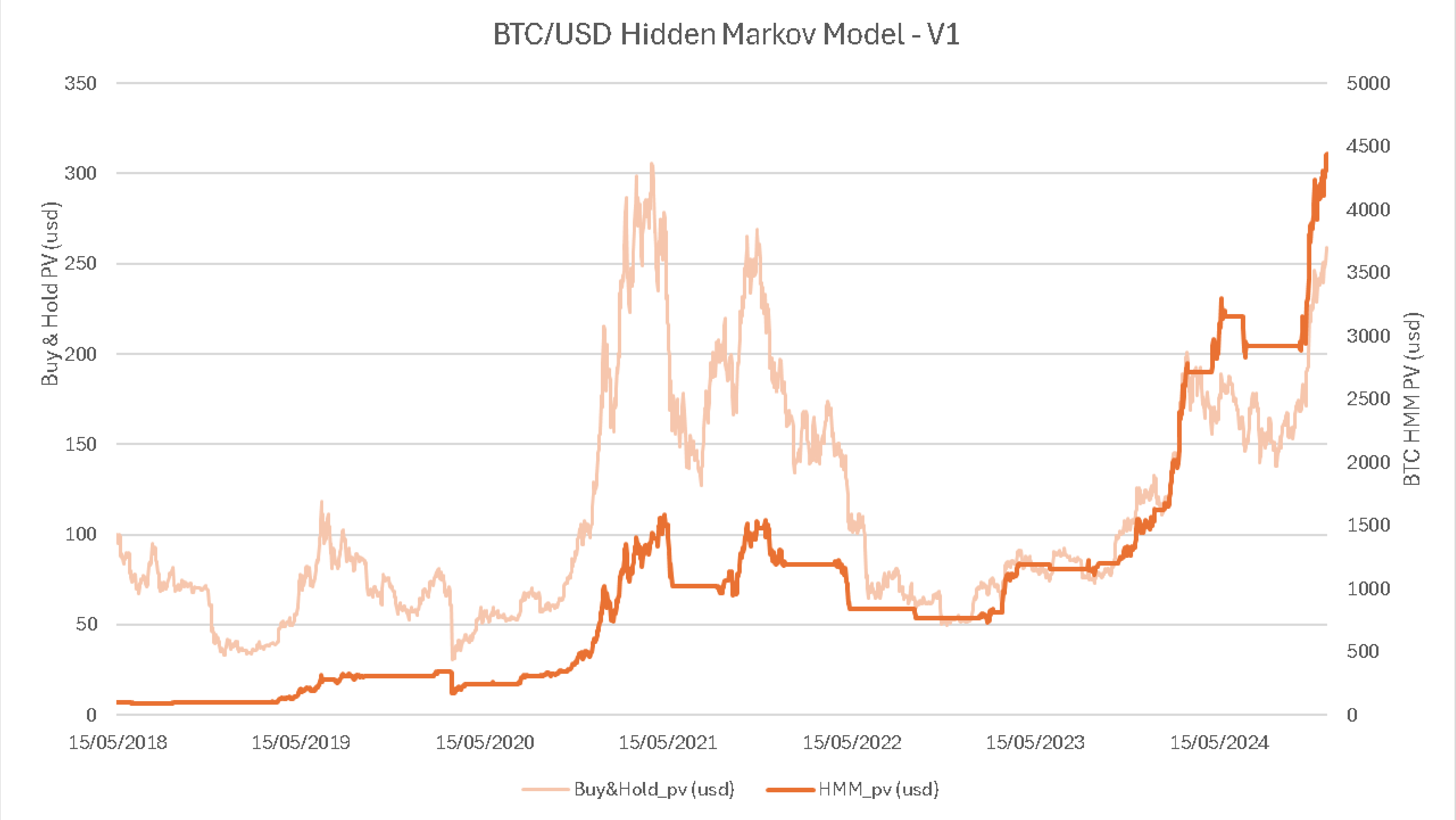
- If the model state results to be in SATE 3, it results in a “Buy" Signal, assigned to a potential bullish trend to capitalize on upward trends, in practice, I would use more tools to determine if I should buy or stay out of the market since this state tends have pullbacks and have persistent low positive log returns. If the state keeps going back and forth from state 3 to 2, then it’s preferable to go to cash.
- If the model state results to be in SATE 4, it results in a “Sell" Signal with a more cautions approach going to CASH due to the slight persistent downtrends, assigned to a neutral-bearish state to avoid unnecessary risks.
- If the model state results to be in SATE 5, it results in a “Buy" Signal with more cautions approach due to the past recent downtrend, assigned to the possible start of an uptrend with some risk of pullback. I would use more tools to determine if I should enter the position during this time and always use a stop-loss.
- If the model state results to be in SATE 6, it results in a “Buy" Signal with more cautions approach due to the past recent uptrend, assigned to the possible continuation of an uptrend.
- If the model cannot specify the state results, then I use the las result following the Markovian Property.

Chapter 5: Model Implementation, Results & Limitations.

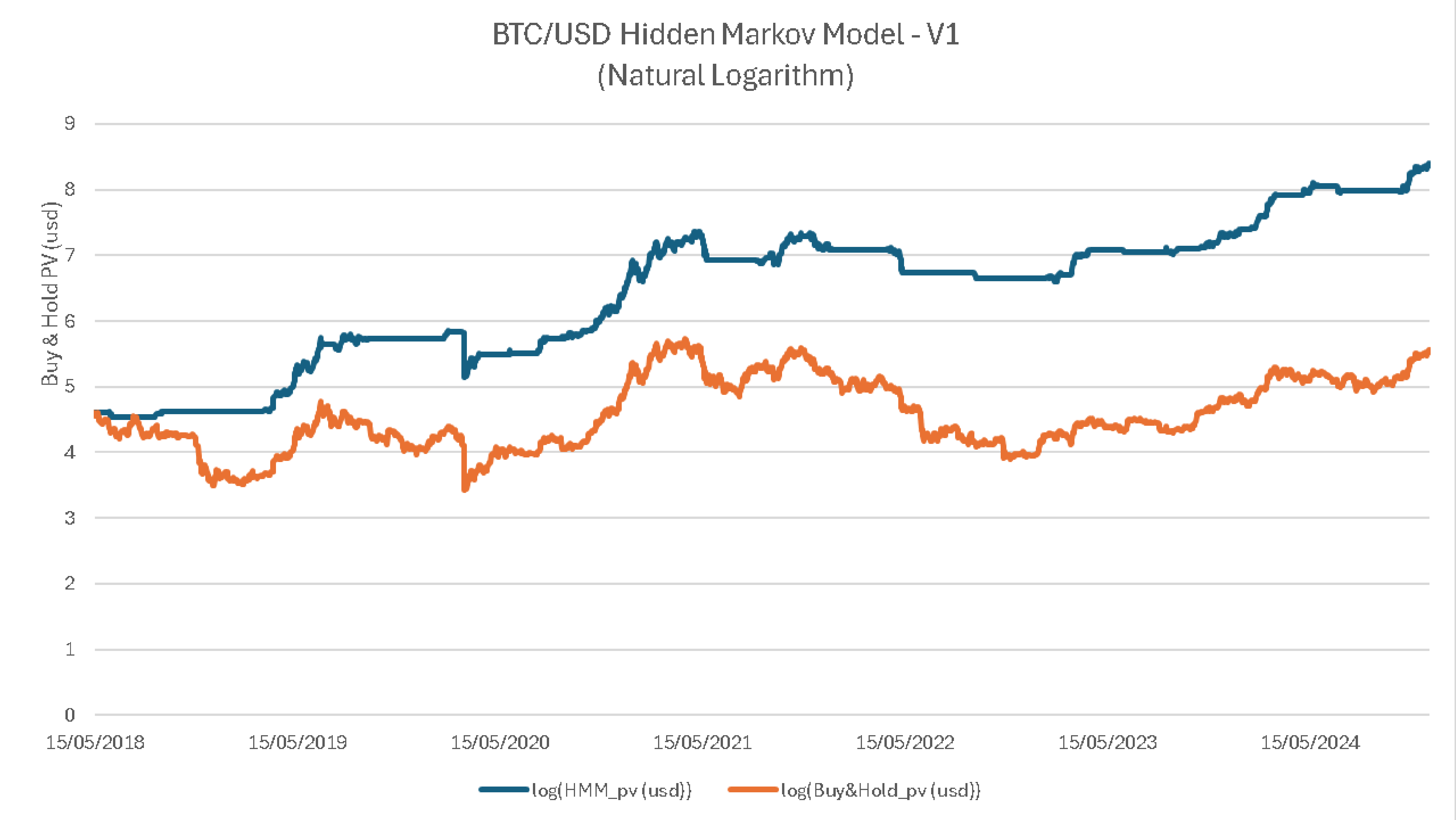
Always keep in mind that sometimes the model could mis-specify the current state; that's why this model should only be used as a complementary tool in a systematic trading system for Bitcoin or its derivatives.

Hidden State	Market Regime	Trade Signal
0	Bullish to Bearish Transition	Sell -> Cash
1	Weak Bullish to Neutral	Hold
2	Neutral to Bearish	Sell -> Cash
3	Bullish to Bullish with Weak Pullbacks	Buy* / Sell*
4	Weak Bearish to Neutral	Sell -> Cash
5	Bullish Trend Start with Strong Momentum	Buy
6	Strong Bullish Transition	Buy
7	Bearish Transition to Bullish	Buy

* = I don't like pullbacks, so I'll sell, but another strategy could be buying.



BTC/USD Hidden Markov Model - V1
Buy & Hold PV (usd)
BTC HMM PV (usd)
0
50
100
150
200
250
300
350
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
15/05/2018
15/05/2019
15/05/2020
15/05/2021
15/05/2022
15/05/2023
15/05/2024
Buy&Hold_pv (usd)
HMM_pv (usd)


BTC/USD Hidden Markov Model - V1
(Natural Logarithm)
Buy & Hold PV (usd)
9
8
7
6
5
4
3
2
1
0
15/05/2018
15/05/2019
15/05/2020
15/05/2021
15/05/2022
15/05/2023
15/05/2024
log(HMM_pv (usd))
log(Buy&Hold_pv (usd))

Chapter 5: Model Implementation, Results & Limitations.

Stand alone the strategy presented in the las slide, outperformed a Buy & Hold Traditional strategy, here are some metrics to consider:

- **Cumulative Return:**
  - Buy & Hold = 158.73%
  - BTC HMM = 4341.10%
- **CAGR:**
  - Buy & Hold = 15.27%
  - BTC HMM = 76.32%
- **Sharpe Ratio:**
  - Buy & Hold = 0.029
  - BTC HMM = 0.077
- **Max Drawdown:**
  - Buy & Hold = -49.39%
  - BTC HMM = -49.39% - The model did not classify correctly the state of this observation, else, the max drawdown would have been -15%

Everyone says that everyone can be wealthy and rich in Excel, but if you would like to test these results for yourself, feel free to ask for the data set at bmercadof0@gmail.com

Chapter 5: Model Implementation, Results & Limitations.

The model presented throughout this document demonstrates a foundational implementation of a Hidden Markov Model (HMM) applied to Bitcoin price dynamics. Through a series of tests, the optimal model configuration identified eight hidden states, which were interpreted based on the author's expertise. However, significant challenges emerged, particularly in the accurate classification of States 5 and 7.State 5 was interpreted as signaling the "Start of a Bullish Trend with Strong Momentum," while State 7 was identified as a "Bearish Transition to Bullish." Despite these definitions, the model exhibited limited predictive accuracy in consistently distinguishing these states. It appeared to attempt classification but lacked the confidence to reliably differentiate them, often misclassifying transitions between the two. This was particularly evident for State 7, as its transitional nature complicated the model’s ability to determine whether it had entered State 7 or remained in State 5.Another notable limitation of the model lies in its operational simplicity. While the strategy presented is straightforward and effective under specific conditions, it frequently results in the model being out of the market, even during periods when profitable opportunities were available. This highlights the trade-off between model simplicity and the potential to capture complex market dynamics.

Chapter 5: Model Implementation, Results & Limitations.

Further improvements to the model could address these challenges by enhancing the selection and integration of predictors. Identifying more robust observations, features, and exogenous variables could improve the model's ability to discern hidden states accurately and respond to market shifts more effectively. Additionally, refining the discretization process for these predictors could yield a more precise representation of the underlying data, further enhancing the model's predictive power.

Such improvements would likely result in a more nuanced and responsive framework, capable of better adapting to the intricacies of Bitcoin price movements and reducing the likelihood of missed market opportunities.

References

- by djp3 [by djp3]. (2020). Hidden Markov Models Playlist [Video]. By Djp3. Retrieved December 10, 2024, from https://youtube.com/playlist?list=PLix7MmR3doRo3NGNzrq48FItR3TDyuLCo&feature=shared
- L.R. Rabiner A. tutorial on hidden Markov models on selected applications in speech recognition. In A. Waibel and K. – F. Lee, editors, Readings in Speech Recognition, pages 267 – 296. Kaufmann, San Mateo, CA, 1990.
- Crypto Data Download. (2024). Bitstamp Bitcoin Exchange Data. [database] cryptodatadownload.com. Retrieved December 20, 2024, from https://www.cryptodatadownload.com/data/bitstamp/
- Alternative. (2024). Crypto Fear & Greed Index. [database] alternative.me. Retrieved December 20, 2024, from https://alternative.me/crypto/fear-and-greed-index/
- Hamdy A. Taha, A. [Hamdy A. Taha]. (2012). Operations Research (9th ed.) [Physical Book]. PEARSON EDUCATION. Pages 571 to 592. ISBN: 978-607-32-0796-6
- ChatGPT - Financial Model Comparison. (s. f.). ChatGPT. https://chatgpt.com/share/67663ec5-acc0-8004-ac2d-7e8dadbf1f53