VNU Journal of Science: Mathematics – Physics, Vol. 40, No. 4 (2024) 47-63

VNU Journal of Science: Mathematics - Physics

Journal homepage: https://js.vnu.edu.vn/MaP

Original Article

Application of Markov Chains in Stock Price Trend Forecasting

Ha Thanh Bac*, Vo Thi Thanh Thu, Vo Doan Xuan Tien, Ho Ngoc Thien, Le Thanh Binh

Quy Nhon University, 170 An Duong Vuong, Quy Nhon, Binh Dinh, Vietnam

Received 20 May 2024
Revised 19 August 2024; Accepted 10 December 2024

Abstract: In this work we explore the application of Markov chains in forecasting stock price trends. Markov chains let a stochastic process that transitions from one state to another are based on probabilities offer a valuable framework for analyzing sequential data such as stock prices. By modeling the state transitions of stock prices, we aim to predict future price movements and identify potential trends. Through empirical analysis and evaluation, we demonstrate the effectiveness of Markov chains in stock price trend forecasting across various stocks. This work is a contributionin the growing body of literature on quantitative methods in financial forecasting and provides insights into the practical application of Markov chains in the stock market domain.

Keywords: Markov chain, stock price forecasting, time series.

1. Introduction

Nowadays, the stock market has become an attractive destination, attracting numerous investors from individuals to institutions because of its high profits, but with hidden risks. Therefore, forecasting market trends is always a top priority for individuals and organizations analyzing securities. Machine learning techniques, with their ability to process big and complex data and make accurate forecasts, are becoming useful tools in forecasting stock market prices. In this work we aim to build a model that forecasts stock market trends, to assist investors and financial professionals in making smart investment decisions and optimizing returns.

In the past, there have been many domestic and foreign researchers who have proposed different methods to improve the ability to forecast stock trends. Some results from recent research works can be listed as follows:

* Corresponding author.
E-mail address: bachathanh1@gmail.com

https//doi.org/10.25073/2588-1124/vnumap.4941

In a work [1] Sun and Zhao used machine learning models (multiple linear regression, random forest, and LSTM) to forecast closing prices based on NASDAQ ETF data and data of statistical factors. The author's empirical process shows that the multiple linear regression model is consistent with stock price forecasts.

Lv and Jiang [2] proposed a predictive model with the RBF multiview neural network. With this model, it is possible to take full advantage of both the internal information provided by the correlation between each mode and the distinctive characteristics of each mode to form independent sample information. Since then, the authors have also shown the feasibility of the model.

Khoa et al., [3] used the Support Vector Regression (SVR) model on the CAPM platform to forecast profiting individual stocks and identify factors affecting the error in the forecast. Through the experimental process on the dataset collected from listed companies on the Ho Chi Minh City Stock Exchange from December 2012 to September 2020, the research also shows that the SVR model is more effective than CAPM.

Markov chain is a model that can be effectively applied to stock price forecasting. This model has proven its efficacy across various fields such as economics, healthcare, etc. In the stock market, Markov chains have been widely used, showing following results:

Wang and Theobald [4] used a Markov regime-switching model to study stock return expectations and variances in six emerging East Asian markets from 1970 to 2004. They found two-state transitions in Malaysia, Philippines, and Taiwan; and three-state transitions in Indonesia, South Korea, and Thailand.

Hoa and Huong [5] applied a Markov-switching EGARCH model to examine the dynamic linkages between exchange rates and stock market volatility in ASEAN markets from 2005 to 2013. The authors showed a significant relationship between stock and foreign exchange markets, with stock return volatility responding asymmetrically to the foreign exchange markets.

Dar et al., [6] analyzed and forecasted Tata Consultancy Services (TCS Ltd.) stock prices in the Indian market using a Markov chain model. Using data from 2020 to 2022, they identified long-term trends and predicted future market states, achieving relatively good results.

Recognizing the high potential and applicability of the Markov model in the stock market, we will analyze, evaluate, and apply the Markov chain model to assess characteristics, predict trends, and forecast the closing prices of stocks listed on the Vietnamese stock market.

2. Markov Chain Theory

The primary focus of this work is the Markov chain. In this section, we present the essential theories of Markov chains to provide a foundation for their application in stock price forecasting. The knowledge presented here is referenced from [7-9].

2.1. Time Homogeneous Discrete Time Markov Chains

Stochastic process is a set of random variables indexed by time, typically denoted as $\{X_n\}$. The set of all possible values that these random variables can take is called the state space. Furthermore, if $n = 0, 1, 2, ...$ then we have a discrete time stochastic process, whereas if $n \in [0, \infty)$ and a continuous time stochastic process.

Consider a discrete time stochastic process $\{X_n, n = 0, 1, 2, ...\}$ with a state space S that is countable. This process is called a Markov chain if it satisfies the Markov property, which means that the conditional distribution of the future state of the process depends only on the current state of the process and is independent of the past:

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = P(X_{n+1} = j | X_n = i) \coloneqq p_{ij},$$

for all $i, j, i_0, \dots, i_{n-1} \in S$ và $n \geq 0$. The *transition probabilites* of the Markov chain satisfy the following conditions:

$$0 \leq p_{ij} \leq 1, \forall i, j \in S;$$

$$\sum_{j \in S} p_{ij} = 1, \forall i \in S.$$

If p_{ij} does not depend on n, we call the Markov chain time homogeneous. In this paper, we only consider time homogeneous Markov chains. The matrix $\boldsymbol{P} = (p_{ij})$ is called the *transition probability matrix*.

The probability that the Markov chain will be in state j at the n^{th} step starting from state j is given by

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i).$$

Then, the matrix $\boldsymbol{P}^{(n)} = \left(p_{ij}^{(n)}\right)$ is called the n-step transition probability matrix.

Theorem 2.1. (Chapman-Kolmogorov equation) Let $\{X_n\}$ be a Markov chain with state space S and transition matrix $\boldsymbol{P}$. Then, for non-negative integers n, m, we have

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)}, \tag{1}$$

or, in matrix notation, $\boldsymbol{P}^{(n+m)} = \boldsymbol{P}^{(n)} \boldsymbol{P}^{(m)}$.

From the theorem above, we can easily prove that $\boldsymbol{P}^{(n)} = \boldsymbol{P}^n$.

2.2. State Distributions

The *state distribution* of the chain at time n is given by the formula

$$\alpha_j^{(n)} = P(X_n = j);\ n = 0,1,2, \dots;\ j \in S.$$

Let $\boldsymbol{\alpha}^{(n)} = \left(\alpha_j^{(n)}, j \in S\right)$, then $\boldsymbol{\alpha}^{(0)}$ is called the initial distribution of the chain.

Theorem 2.2. For every n, the distribution $\boldsymbol{\alpha}^{(n)}$ sastifies

$$\boldsymbol{\alpha}^{(n)} = \boldsymbol{\alpha}^{(0)} \boldsymbol{P}^{(n)}. \tag{2}$$

2.3. Classification of States

2.3.1. Communicating States and Class Structure

We say state j is accessible from state i, written as $i \to j$, if $p_{ij}^{(n)} > 0$ for some n. If $i \to j$ and $j \to i$, we say i communicates with j and write $i \leftrightarrow j$.

Communication is an equivalence relation. That is, it has the following properties:

− *reflexive*: $i \leftrightarrow i$, for all i;

− symmetric: if $i \leftrightarrow j$ then $j \leftrightarrow i$;

− transitive: if $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$.

Therefore, the states of a Markov chain can be partitioned into *communicating classes* such that only members of the same class communicate with each other. That is, two states i and j belong to the same class if and only if $i \leftrightarrow j$.

If the entire state space S is one communicating class, we say that the Markov chain is ***irreducible***.

2.3.2. Periodicity

We say that a state $i \in S$ has period d_i, where

$$d_i = \text{gcd}\left\{n \in \{1, 2, ...\}: p_{ii}^{(n)} > 0\right\}.$$

If $d_i = 1$, then i is called *aperiodic*.

Theorem 2.3. All states in a communicating class have the same period. In other words, if $i, j \in S$ are such that $i \leftrightarrow j$, then $d_i = d_j$.

2.3.3. Recurrence and Transience

We say that a state i is *recurrent* if

$$P(X_n = i \text{ for infinitely many } n) = 1.$$

We say that i is *transient* if

$$P(X_n = i \text{ for infinitely many } n) = 0.$$

We can easily show that every state is either recurrent or transient.

Theorem 2.4. Let $i, j \in S$ be such that $i \leftrightarrow j$. If i is recurrent, then j is also recurrent; while if i is transient, then j is also transient.

For this reason, we can refer to a communicating class as a "recurrent class" or a "transient class". If a Markov chain is irreducible, we can refer to it as a "recurrent Markov chain" or a "transient Markov chain".

The *first passage time* to state i is defined as follows

$$T_i = \inf\{n \geq 1: X_n = i\}.$$

We then have the expected return time to state i:

$$\mu_i = E(T_i | X_0 = i).$$

Consider a recurrent state i. The state i is said to be *positive recurrent* if $\mu_i < \infty$, or *null recurrent* if $\mu_i = \infty$. The following facts can be proven:

In a recurrent class, either all states5 are positive recurrent or all states are null recurrent.

An irreducible Markov chain with a finite number of states is positive recurrent.

2.4. Invariant Distributions

Let $\boldsymbol{\pi} = (\pi_i)$ be a distribution on S, in that $\pi_i \geq 0$ for all $i \in S$ and $\sum_{i \in S} \pi_i = 1$. We call $\boldsymbol{\pi}$ an invariant distribution if

$$\pi_j = \sum_{i \in S} \pi_i p_{ij} \text{ for all } j \in S,$$

or, equivalently, if $\boldsymbol{\pi} = \boldsymbol{\pi P}$. The invariant distribution is also called equilibrium distribution or stationary distribution.

Theorem 2.5. An irreducible Markov chain has a unique invariant distribution $\boldsymbol{\pi}$ if and only if it is positive recurrent. Moreover, this distribution satisfies

$$\pi_i = \frac{1}{\mu_i}, \text{ for all state } i,$$

where μ_i is the expected return time to state i.

2.5. Convergence to Equilibrium

We shall investigate the limiting behaviour of the n-step transition probabilities $p_{ij}^{(n)}$ as $n \to \infty$. The following is one of the main results in Markov chains theory.

Theorem 2.6. (Convergence to equilibrium) Let $\{X_n\}$ be an irreducible and aperiodic Markov chain. Then for any initial distribution $\boldsymbol{\lambda}$, we have that $P(X_n = j) \to \frac{1}{\mu_j}$ as $n \to \infty$, where μ_j is the expected return time to state j. In particular:

Suppose $\{X_n\}$ is positive recurrent. Then the Markov chain has a unique invariant distribution $\boldsymbol{\pi}$ given by $\pi_i = \frac{1}{\mu_i}$, so $P(X_n = j) \to \pi_j$ for all j.

Suppose $\{X_n\}$ is null recurrent or transient. Then $P(X_n = j) \to 0$ for all j.

2.6. Ergodic Theorem

The previous theorem delves into the asymptotic behavior of $P(X_n = j)$ which signifies the probability of the Markov chain occupying state j at some distant time n. Also, one can analyze the long-run amount of time spent in state j, effectively averaging its behavior over the long haul. In mathematical jargon, the term "ergodic" encapsulates notions pertaining to the long-term proportion of time.

Denote by $V_i(n)$ the number of visits to i before n:

$$V_i(n) = \sum_{k=0}^{n-1} 1\{X_k = i\}.$$

Then $\frac{V_i(n)}{n}$ is the proportion of time before n spent in state i, and its limiting value (if it exists) is the long-run proportion of time spent in state i.

Theorem 2.7. (Ergodic theorem) Let $\{X_n\}$ be an irreducible Markov chain. Then for any initial distribution $\boldsymbol{\lambda}$, we have

$$P\left(\frac{V_i(n)}{n} \to \frac{1}{\mu_i} \text{ as } n \to \infty\right),$$

where μ_i is the expected return time to state i.

2.7. Estimating the Transition Matrix

If we have the transition probability matrix of the Markov chain, we can perform various analyses. However, in practice, we often do not know this matrix in advance, so we need to estimate it based on observations of the process under study. Suppose we are studying a Markov chain with m states. In that case, the parameters that we need to estimate are the m^2 elements p_{ij} of the matrix $\boldsymbol{P}$, where

$$p_{ij} = P(X_{t+1} = j | X_t = i).$$

Suppose we have observed a random sample $(x_0, x_1, \dots, x_n)$. These are realizations of the random variables $(X_0, X_1, \dots, X_n)$. Using the maximum likelihood estimation (MLE) method, we obtain the following maximum likelihood estimate:

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_j n_{ij}},$$

where n_{ij} is the number of transitions from i to j in $(x_0, x_1, \dots, x_n)$.

When estimating the transition matrix, it's possible that some transitions do not occur in our data, leading to some estimate being zero or divisions by zero during the computation process. To overcome this, we use Laplace smoothing method by adjusting the estimate as follows:

$$\hat{p}_{ij} = \frac{n_{ij} + \alpha}{\sum_j n_{ij} + m\alpha},$$

where α is smoothing parameter. In this work, we choose $\alpha = 0.001$ as it provides a balance between maintaining the integrity of the original data and preventing zero probabilities or divisions by zero. This small value of α introduces only a slight adjustment to the estimates, ensuring that the resulting transition matrix remains as accurate as possible. In practice, α can be chosen as any small number, such as 0.001 or 0.003, depending on the specific needs of the analysis. The key is to select a value that provides sufficient smoothing without overly distorting the original data.

3. Methodology

To utilize the Markov chain model, we make the following assumptions about the stock market:

Daily stock prices follow a Markov process, meaning that future stock prices depend only on the current price and are independent of past prices.

This process is time homogeneous, implying that the transition probabilities between states do not depend on the timing of the transition.

In reality, due to the complexity and randomness of the stock market, these assumptions are difficult to satisfy. However, numerous studies have demonstrated the effectiveness of Markov chains in modeling stock prices. Therefore, these assumptions are necessary for us to apply this model.

The result of this work is presented as a web application using the Streamlit package in the Python programming language. Users of the application will select stock tickers of companies in Vietnam by industry groups (banking, securities, petroleum, etc.) and a model for stock price forecasting (currently including Markov chain and ARIMA model). The application will provide information such as historical data, line charts, candlestick charts, etc., and utilize the selected model to forecast prices, offering useful insights to users during the investment process.

3.1. Data Collection

We collected stock data from Yahoo Finance. The data was collected from January 2022 to real-time during the application's runtime. Some stock tickers only have data from around 2023. This data is divided into training set and test set in an 8:2 ratio.

Based on this data, two Markov chain models will be constructed to address two problems:

Problem 1: Investigating the trends (upward, downward, flat) of stock prices.

Problem 2: Forecasting the closing price of stocks in the next trading session.

3.2. Problem 1: Investigating the Trends of Stock Prices

First, we define the states. Following the method in [10, 11], the states regarding price trends are determined based on the closing price difference between two consecutive days. Let y_i be the closing price on day i. Let $d_i = y_i - y_{i-1}$. Then, the closing price of a particular day will be assigned the state *up* if $d_i > 0$, *flat* if $d_i = 0$, and *down* if $d_i < 0$. Consider the Markov chain consisting of these 3 states. The transition probability matrix $\boldsymbol{P}$ will be estimated from training data using MLE method.

After estimating $\boldsymbol{P}$**,** we carry out the following tasks:

Checking properties of the Markov chain: irreducibility, aperiodicity, recurrence.

Forecasting the trends (states) for the next few days.

Calculating invariant distribution of the chain.

Calculating the following quantities of the chain: expected return times, expected hitting times, expected numbers of visits.

3.3. Problem 2: Forecasting the Closing Price of Stocks in the Next Trading Session

3.3.1. Defining States

Similar to the method in [12], in this problem, we define states based on ranges of the *daily return rate*:

$$r_t = \frac{y_t - y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1.$$

The return rate is used instead of price difference as in problem 1 because relative quantities are often more intuitive and common in finance. Importantly, the states will not be affected by the trend of stock prices over time.

For each stock ticker, we find the minimum daily return rate r_{min} and the maximum r_{max} from the training data. We divide each interval $[r_{min}; 0)$ and $[0;\ r_{max}]$ into 3 equal-length sub-intervals. Then, we will have 6 intervals corresponding to 6 states. There are 3 increasing states labeled as U1, U2, U3 and 3 decreasing states labeled as D1, D2, D3. In this notation, U=Up (increase), D=Down (decrease), and the number following indicates the magnitude (small, medium, large). The states U3 and D3 are extended to infinity to accommodate data from the test set outside of $[r_{min};\ r_{max}]$. An example of states for the stock code BID of Joint Stock Commercial Bank for Investment and Development of Vietnam - BIDV is presented in the table below, which has $r_{min} = -7.49\%$ and $r_{max} = 6.53\%$.

Table 1. Definition of states based on the daily return rate (%) of stock ticker BID

D3	D2	D1	U1	U2	U3
$(-\infty; -4.98)$	$[-4.98; -2.49)$	$[-2.49; 0)$	$[0; 2.17)$	$[2.17; 4.34)$	$[4.34; \infty)$

3.3.2. Sliding Window Method

To predict the closing price for the next trading session, we use the sliding window method. The idea is to predict the price for day n, we will use data from the previous m days to estimate the transition probability matrix $\boldsymbol{P}$. The size of the window m is a hyperparameter of the model and will be selected through blocked cross-validation. To predict for the next day $n + 1$, the window will shift by one day, leading to an update of the transition probability matrix.

Based on the state of the previous day (today), we can predict the state of the next day (tomorrow) using the formula

$$\boldsymbol{\alpha}^{(1)} = \boldsymbol{\alpha}^{(0)}\boldsymbol{P},$$

where the vector $\boldsymbol{\alpha}^{(0)}$ has a value of 1 at the position corresponding to today's state, and all other positions are 0. The predicted state for tomorrow is the state with the highest probability (position with the largest element in vector $\boldsymbol{\alpha}^{(0)}$). In case multiple states have equal probabilities, one of them is randomly chosen. Then, based on the definition of this state and the closing price of today, we can calculate the closing price range for tomorrow. The final predicted value is the midpoint of this range.

3.3.3. Choosing Window Size

To determine the optimal window size for prediction, we used the blocked cross-validation method (similar to [12]). For each stock ticker, we tested window sizes from 5 to 30. This process was carried out on the training dataset following these steps (illustrated with a window size of 5):

i) The transition probability matrix $\boldsymbol{P}$ was estimated based on data from day 1 to day 5;

ii) With matrix $\boldsymbol{P}$, a price prediction for day 6 was made. This prediction was saved for evaluation at the end of the procedure;

iii) The sliding window was shifted to contain data from day 7 to day 11 (excluding day 6 already used for validation) and continued to predict for day 12;

iv) These steps were repeated until the end of the training set;

v) The predictions were evaluated using the Root Mean Squared Error (RMSE) index:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2},$$

where y_t and $\hat{y}_t$ are the actual price and the predicted price of the day being validated at time t.

During the process, each data point was only used once, either for training or for validating. The optimal window size ms for each stock ticker is the size with the lowest RMSE index.

4. Results and Discussion

4.1. Data Visualization

As mentioned, the results of this work have been presented in a web application built using the Python programming language, accessible at the following address: https://stock-price-forecasting-web-app.streamlit.app

The interface of the web app comprises a sidebar and a main page. Within the sidebar, users can find selection boxes to choose the industry group, stock ticker, and model for forecasting stock prices. The current version of our web application includes the following stock tickers:

Table 2. List of stock tickers (categorized by industry group)

Industry Group	*Banking*	*Securities*	*Electronics*	*Petroleum*	*Public Investment*	*Steel*
Stock Tickers	BID	AGR	BTP	ASP	C47	DTL
	VCB	APG	CHP	CNG	CII	HMC
	TCB	BSI	DRL		CTD	HPG
	CTG		CAV		CTI	HSG

On the main page, there are two tabs: The Data tab displays historical data and charts, while the Forecast tab is for model predictions.

From here on, unless stated otherwise, the results presented are based on the BID stock ticker. In the Data tab, besides historical data, we display several charts to help users grasp the trends and movements of the current stock price. First, we plot a line chart of the closing price along with the moving average line - a popular method for smoothing out price fluctuations, enabling better observation of market trends.

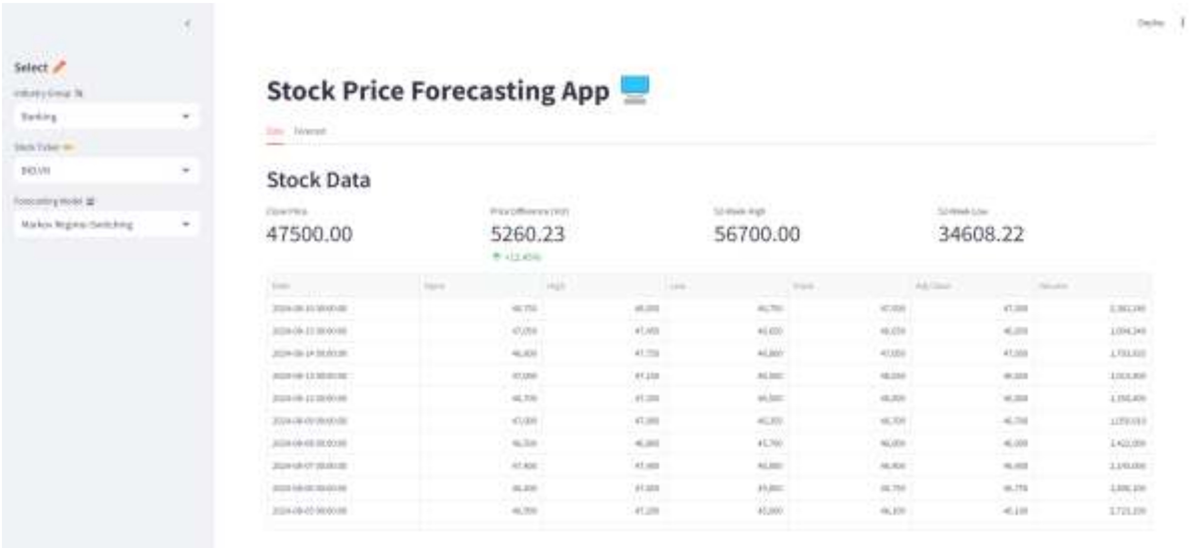

Figure 1. Basic interface of the web application.

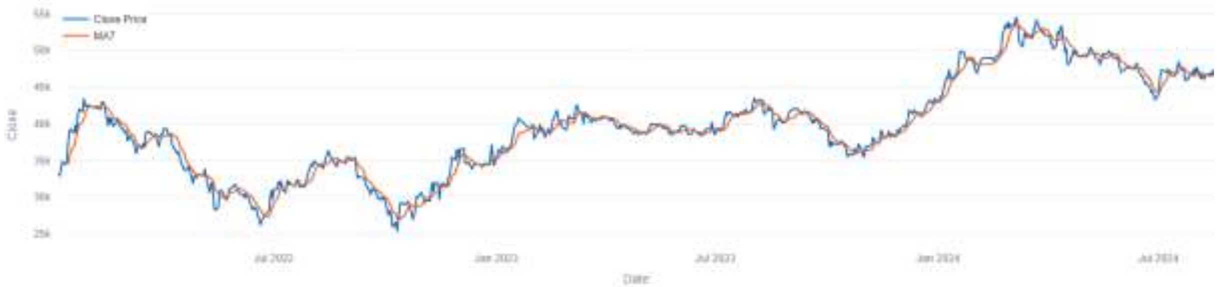

Figure 2. Line chart of closing price with moving average.

We use candlestick charts to represent key stock price information for each trading session, including the opening, high, low, and closing prices.

Figure 3. Candlestick chart of stock prices.

Next, we create a bar chart to display the trading volume of stocks for each session. This bar chart illustrates the market activity level for each day, providing insights into the fluctuation and nature of trading volume.

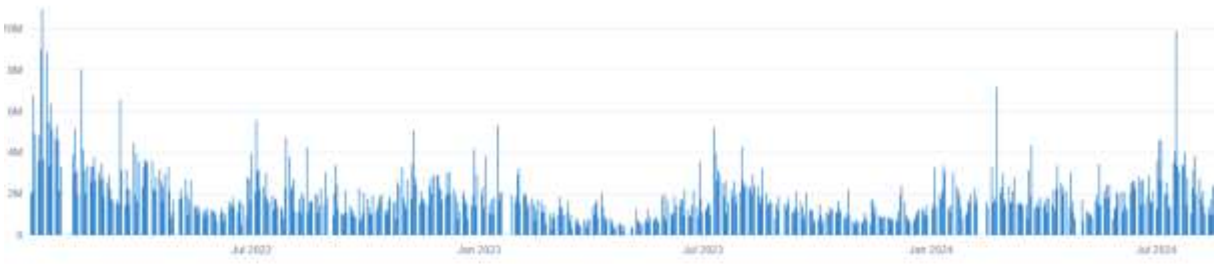

Figure 4. Bar chart of trading volume.

4.2. Stock Price Forecasting

For the Forecast tab, we will utilize the user-selected model to provide valuable insights and forecasts regarding the stock price in the upcoming sessions.

4.2.1. Markov Chain Model

Problem 1: Investigating the trends of stock price

To estimate the transition matrix for the Markov chain, we calculate price differences and assign states to each day as outlined earlier. With the help of the *pydtmc* library in Python, we obtain the Markov chain from the training data and visualize it as follows:

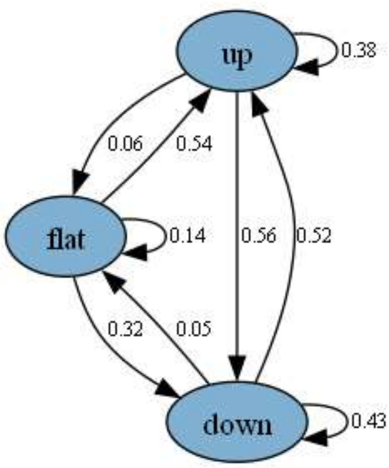

Figure 5. Markov chain of stock price trends.

Looking at the graph, we can easily observe the transition probabilities of the chain. For instance, the probability of the stock price being in an increasing trend today and transitioning to a decreasing trend tomorrow is 0.56. Next, we examine the important properties of the Markov chain and obtain the results as follows:

Table 3. Results of checking some properties of the Markov chain

Irreducibility	Aperiodicity	Recurrent states	Transient states
True	True	Up, Flat, Down	

So, the Markov chain we just estimated has some very "nice" properties: it's irreducible, aperiodic, and recurrent. Moreover, because it has a finite number of states, the chain is positive recurrent. Now we try using this Markov chain to predict states for some future days and compare with actual data from the test set.

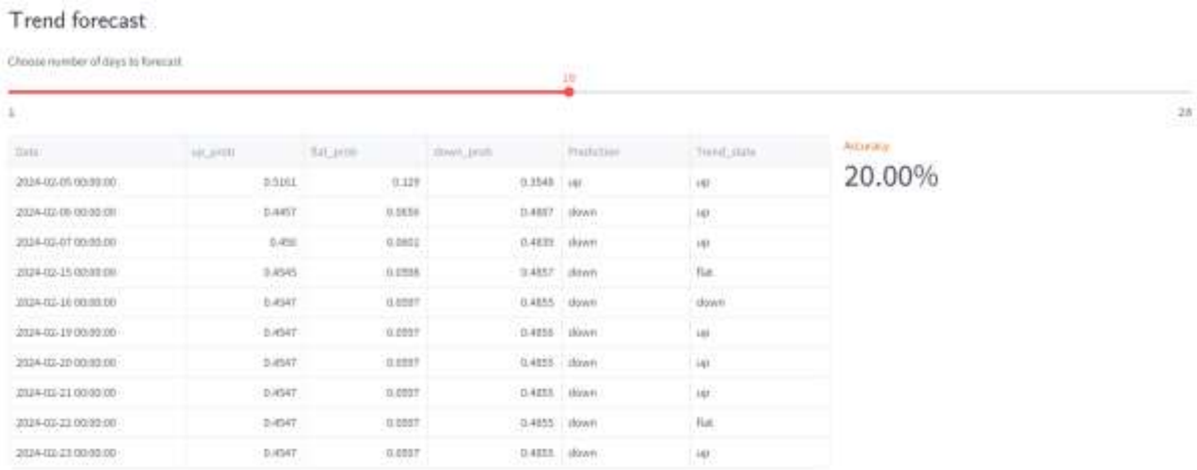

Figure 6. Predicted price trends for the next 10 days.

Here, each day's prediction is based on the state with the highest probability. We notice that prediction accuracy isn't very high. From the second day onwards, the Down state has the highest probability (around 0.48), but it's not significantly different from the Up state (probability 0.45). Additionally, probabilities stabilize over time due to the irreducibility and aperiodicity of the Markov chain, leading transition probability rows to converge to the invariant distribution (equilibrium distribution). Consequently, the Down state becomes the predominant prediction after several days. Thus, we advise users not to solely rely on the final prediction. Instead, consider state probabilities and additional information provided for well-informed investment decisions.

According to the Ergodic theorem, the invariant distribution is the long-run proportion of time the chain spends in each state:

Table 4. Invariant distribution of the Markov chain

Up	*Flat*	*Down*
0.4547	0.0597	0.4855

In the long run, we expect approximately 45% of the time the chain will be in the Up state, 6% of the time in the Flat state, and 48% of the time in the Down state. We will compare this prediction with the actual time proportions in the test data.

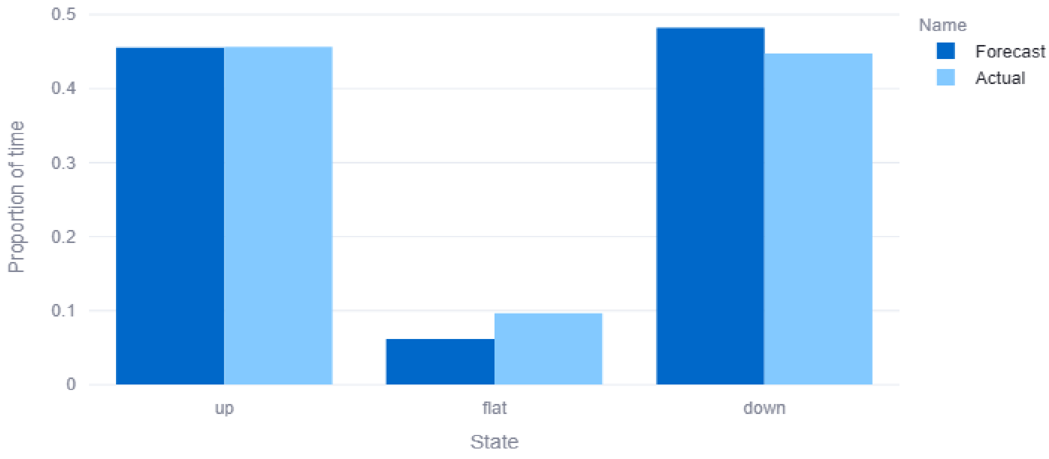

Figure 7. Comparison of long-run proportion of time between forecast and actual data.

The model's predictions closely match the actual data, aiding investors in decision-making for various stocks. Long-run time proportions help identify industries or companies with growth potential for effective investment. Next, we present additional characteristics of the BID stock price calculated using the Markov chain model:

Table 5. Expected return times

Up	Flat	Down
2.1951	16.0714	2.0737

Here, investors can expect that the closing price will increase again about 2 days after the last price hike. The relatively high average time to return to the Flat state (around 16 days) reflects the random nature and continuous fluctuations of the stock market. Moreover, the Down state also has a return time of about 2 days, indicating that stock prices are likely to alternate between increases and decreases in consecutive trading sessions.

Table 6. Expected hitting times

	Up	Flat	Down
Up	0	17.5096	1.8366
Flat	1.8856	0	2.3146
Down	1.9173	17.7063	0

Starting from the Up state, the chain transitions to the Down state take place during 1 to 2 days, and to the Flat state, during approximately 17-18 days. Flat exhibits a longer time to reach, whereas Up and Down have shorter times, further reinforcing the market's quick response to continuous price trend reversals.

Table 7. Expected numbers of visits in 5 days

	Up	Flat	Down
Up	2.208	0.3114	2.4806
Flat	2.349	0.4009	2.2501
Down	2.3345	0.2992	2.3662

Table 7 indicates that, starting from the Up state, over a 5-day period, the stock price tends to stay in the Up state for around 2.208 days, in the Flat state for approximately 0.3114 days, and in the Down state for roughly 2.4806 days. Over 5 days, which can be understood as a trading week, the number of times the stock price increases and decreases is almost equal. In our analysis of other bank stocks, we observed comparable outcomes: the prevalence of Up and Down states, each comprising roughly 45% in the long run. This equilibrium is reflected in other metrics for these states, portraying a market characterized by randomness and volatility, with stability being a rare occurrence.

Let's explore the attributes of firms in the electricity sector, using BTP from the Ba Ria Thermal Power Joint Stock Company as an example. Other companies within the sector exhibited comparable metrics.

Table 8. Expected return times of BTP ticker

Up	Flat	Down
2.7255	3.8611	2.6731

Table 9. Expected hitting times of BTP ticker

	Up	Flat	Down
Up	0	3.8456	2.8783
Flat	2.8768	0	2.7192
Down	2.9766	3.7863	0

Table 10. Expected numbers of visits in 5 days of BTP ticker

	Up	Flat	Down
Up	1.8834	1.2903	1.8264
Flat	1.8279	1.2862	1.8859
Down	1.7912	1.3056	1.9031

In contrast to the banking sector, metrics across states in the electricity sector appear more evenly distributed. Expected return times and numbers of visits to each state – Up, Flat, and Down – show minimal variation. This suggests that the electricity stock market may exhibit relatively more stability and less volatility compared to banking. Notably, the Flat state is expected to occur at least once within a trading week, indicating a period of market stability before potential shifts. This frequent occurrence of the Flat state may imply the influence of stabilizing factors, such as risk management policies or external factors, in maintaining market stability in the electricity sector.

The model's predictions generally align with reality. Banking stocks tend to be highly volatile due to their sensitivity to economic and financial fluctuations, including global market conditions, interest rates, exchange rates, and credit risks. Conversely, the electricity sector exhibits greater stability, driven by the essential nature of electricity services. This stability in demand leads to steadier revenue and profits for electricity companies, which typically operate with larger capital structures and face less volatility compared to banks.

Problem 2: Forecasting the closing price of stocks in the next trading session

To assess the model's effectiveness, we plot a line chart comparing predicted prices with actual prices and calculate evaluation metrics. In addition to Root Mean Squared Error (RMSE), we also utilize Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|,$$

with y_t representing the actual closing price and $\hat{y}_t$ denoting the predicted closing price for the t-th testing day, we present the results for several stock tickers below.

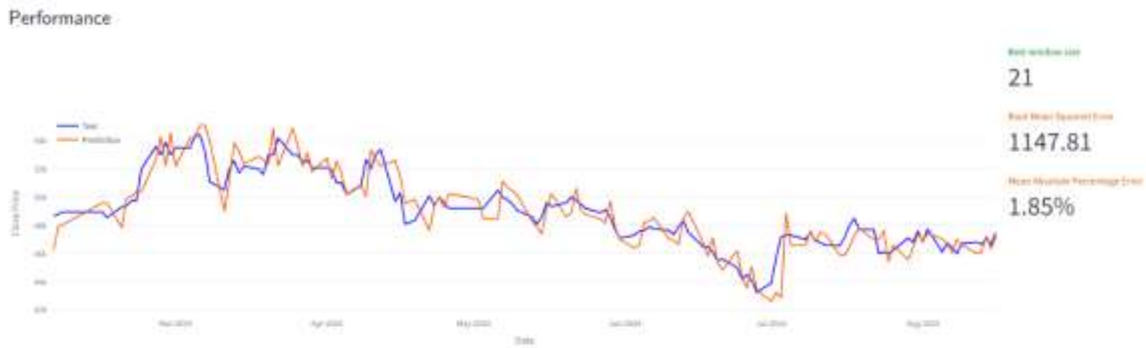

Figure 8a. Ticker BID - Joint Stock Commercial Bank for Investment and Development of Vietnam

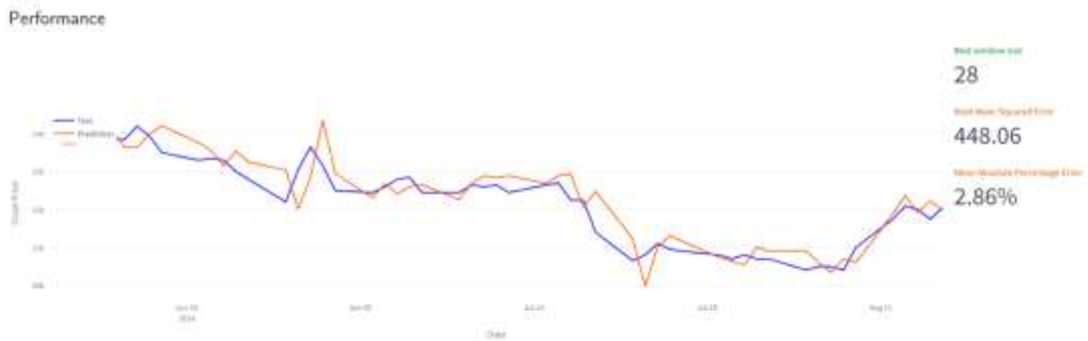

Figure 8b. Ticker APG - APG Securities Joint Stock Company.

Figure 8c. Ticker ASP - An Pha Petroleum Group Joint Stock Company.
Figure 8. Model performance for some stock tickers.

The optimal window size typically falls within the range of 20 - 30 days, this is meaning that the model needs data from the previous 20 - 30 days to make predictions for the next day. The model's predictions are relatively accurate with an average deviation of about 2-3%.

For comparison, our web app also integrates popular time series models, including ARIMA built using the Auto ARIMA method and Facebook Prophet. The performance of these models on some stock tickers is shown in the table below:

Table 11. Comparison of the performance of different models on the test set

Model	Multi-day Prediction	MAPE for some stock tickers				
		BID	APG	ASP	BTP	**DTL**
Markov Chain	No	1.85%	2.86%	1.79%	1.83%	2.67%
Auto ARIMA	Yes	5.07%	14.79%	5.80%	6.34%	18.24%
Facebook Prophet	Yes	7.49%	10.63%	10.07%	14.77%	6.59%

We observe that the Markov chain model has significantly lower MAPE compared to the other two models. Both the ARIMA and Facebook Prophet models have the capability to predict multiple days ahead, which allows investors to plan long-term investments and make strategic decisions for the future. The step-by-step forecasting process works by initially predicting the value for the next day using the model trained on historical data. This predicted value is then added to the current dataset, and the updated data, including the new prediction, is used to forecast the subsequent day. This iterative process continues, generating forecasts for the desired number of days. This feature is particularly valuable for investors who are strategizing for long-term investment opportunities. By leveraging these models, investors can make informed decisions about future market trends, plan investment strategies, and manage their portfolios with a longer-term perspective. The ability to project several days ahead provides a comprehensive view that aids in identifying potential opportunities and risks over extended periods. Nevertheless, it is important to note that the accuracy of these multi-day forecasts will diminish as the forecast horizon extends. This decrease in accuracy arises from the compounding of forecast errors over time. In volatile markets such as the stock market, where conditions can change rapidly and unpredictably, these errors can accumulate, leading to less reliable predictions. This inherent limitation means that while ARIMA and Facebook Prophet are useful for short to medium-term forecasts, their predictions may become less dependable as they project further into the future.

In contrast, the Markov chain model is designed to provide forecasts for only the next day. This limitation constrains its usefulness in long-term trading and investment strategies, as it does not account for longer-term trends or shifts in market conditions. However, this does not render the Markov Chain model obsolete because we partially address this by providing insights into the long-run stock characteristics, as discussed in Problem 1. With the sliding window approach, the model updates its data daily, thereby minimizing the impact of historical discrepancies and reducing bias. By focusing on the most recent data, the sliding window approach ensures that the model remains responsive to current market conditions and provides relatively accurate predictions for the next trading session. While it cannot offer long-term forecasts, this method enhances the Markov chain model’s reliability in predicting short-term movements and provides valuable insights into the immediate future of stock prices.

The Markov chain model can be employed to analyze the behavior of stock markets [4], assess the impact of various factors on market volatility as in [5] and examine transitions between different market states to evaluate future potential [6]. These are very valuable characteristics, yet they require investors to have a certain level of knowledge to effectively apply them in strategic planning and making informed investment decisions. While this study provides similar insights to those found in existing research, we go a step further by offering specific predictions for stock prices. This approach provides a more intuitive basis for decision-making, particularly for novice investors who may lack experience. By giving them concrete projections, our research helps bridge the gap between complex market analysis and practical, actionable investment strategies.

Besides the limitation of predicting only one day ahead, the Markov chain model's reliance on states for price prediction restricts its flexibility in capturing complex market dynamics. To enhance this, increasing the number of states could be beneficial, but it must be balanced with available data. Too many states may lead to issues related to data scarcity. If some states are only represented in a few data points, the model's generality and effectiveness will be reduced. We've opted for 6 states in our Markov chain to maintain this balance.

Some of our future research directions include:

- *Enhancing state definitions* in the Markov chain model by increasing the number of states or changing the variables used, to better capture market complexities.

- *Exploring Markov chain extensions* like the Markov Regime-Switching model, which models market changes over different periods.

- *Incorporating additional features* such as trading volume, market sentiment from news, and macroeconomic factors to enrich the model.

- *Evaluating the model's generalizability* by conducting studies across various financial markets and sectors.

4. Conclusion

In this work, we studied the theory of Markov chains and applied the Markov chain model to forecast stock prices. We conducted experiments on data from several stock tickers listed on the Vietnamese stock market and evaluated the results. The Markov chain model provided valuable insights into stock price state probabilities and transitions. We estimated the transition probability matrix and calculated key metrics like the invariant distribution, mean return time, mean hitting time, and mean number of visits per state. These metrics help investors understand both short-term and long-term stock price behaviors for informed investment decisions. Using daily return-based states, the Markov model predicted the next session's closing prices with a promising error margin of about 2-3% through a sliding window approach. However, there are some limitations included predicting only one day ahead and lacking flexibility in capturing complex market dynamics, that we aim to address in future research.

References

[1] Z. Sun, S. Zhao, Machine Learning in Stock Price Forecast, E3S Web of Conferences, Vol. 214, 2020, pp. 02050, https://doi.org/10.1051/e3sconf/202021402050.

[2] B. Lv, Y. Jiang, Prediction of Short-Term Stock Price Trend Based on Multiview RBF Neural Network, Computational Intelligence and Neuroscience, 2021, https://doi.org/10.1155/2021/8495288.

[3] B. T. Khoa, T. T. Huynh, T. D. Tung, N. V. Duc, Factor Affecting the Error in Individual Stock's Return Forecasting: Appling Machine Learning with Spark MLlib, Ho Chi Minh City Open University Journal of Science: Engineering and Technology, Vol. 17, No. 1, 2021, pp. 5-19, https://doi.org/10.46223/Hcmcoujs.Tech.vi.17.1.2245.2022 (in Vietnamese).

[4] P. Wang, M. Theobald, Regime-Switching Volatility of Six East Asian Emerging Markets, Research in International Business and Finance, Vol. 22, No. 3, 2008, pp. 267-283, https://doi.org/10.1016/j.ribaf.2007.07.001.

[5] N. T. L. Hoa, L. T. T. Huong, Dynamic Linkage Between Exchange Rates and Stock Market Volatility in ASEAN Emerging Markets, Journal of Integration and Development, Vol. 17, No. 27, 2014, pp. 31-35 (in Vietnamese).

[6] G. F. Dar, T. R. Padi, S. Rekha, Stock Price Prediction Using a Markov Chain Model: A Study for TCS Share Values, Advances and Applications in Statistics, Vol. 80, 2022, pp. 83-101, https://doi.org/10.17654/0972361722068.

[7] N. D. Tien, Probability Models and Applications, Part 1 - Markov Chains and Applications, Hanoi National University Publishing House, Hanoi, 2001 (in Vietnamese).

[8] J. R. Norris, Markov Chains, Cambridge University Press, 1998.

[9] S. M. Ross, Introduction to Probability Models, Academic Press, 2014.

[10] D. N. Choji, S. N. Eduno, G. T. Kassem, Markov Chain Model Application on Share Price Movement in Stock Market, Computer Engineering and Intelligent Systems, Vol. 4, No. 10, 2013, pp. 84-95.

[11] F. O. Mettle, E. N. B. Quaye, R. A. Laryea, A Methodology for Stochastic Analysis of Share Prices as Markov Chains with Finite States, SpringerPlus, Vol. 3, 2014, pp. 1-11, https://doi.org/10.1186/2193-1801-3-657.

[12] M. Aronsson, A. Folkesson, Stock Market Analysis with a Markovian Approach: Properties and Prediction of OMXS30, Dissertation, 2023, https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-341862 (accessed on: March 1st, 2024).