Dean&Francis

ISSN 2959-6157

Limitations in Predicting Stock Prices Based on Markov Chains

Shichao Zhang

Beijing 21st Century High School, Beijing, China

*Corresponding author: shichaozhang@bj21cs.com

Abstract:

Markov chains have applications in several fields, including but not limited to physics, biology, engineering, economics, and computer science. In the field of computer science, particularly in the areas of algorithm design, it plays a role in complexity theory, network science, and artificial intelligence. In today's chaotic and complex economic situation, this paper predicts the stock prices of real stocks based on the underlying model of Markov chains. This paper uses the basic Markov chain's properties, and a new transfer probability vector is derived by multiplying a given initial transfer probability vector with a multi-step probability transfer matrix. The new probability transfer vector is tested to see if it is consistent with the real situation. Repeat the above steps to make multiple predictions. After the prediction results of two different stocks, the result is that the stock price prediction based on the Markov chain is difficult to have high accuracy and can have relatively good prediction results in the short term. However, as the period increases and the factors affecting the stock price increase, the prediction results show more deviations. The result can make more reference for improving the prediction of the stock price prediction model, according to the above problems to make targeted improvements or add more models, to improve the prediction accuracy.

Keywords: Stock Price Prediction; Markov chain; Limitation

1. Introduction

As the world economy continues to grow and change, there are many different investment options available to every individual, such as investments, stocks, futures, and bonds. Among them, stocks are preferred by most people with the advantage of lower thresholds and high liquidity. At the same time stocks, as an important part of the system of market economy, have the function of optimizing asset allocation and diversifying risks [1]. Forecasting stock prices has always been the pursuit of many researchers, enterprises, and stockholders. However, due to the uncertainty of the factors leading to the rise and fall of stocks,

ISSN 2959-6157

market environment, company's operating conditions, and exchange rate changes, many people are unable to maintain a stable and sustainable income, as well as the stock's nonlinear system, non-stationarity, stochastic and complexity and many other characteristics, which makes many scientists and researchers have been committed to constructing more and more optimized models to predict the stock price, Monte Carlo simulation, and DCF model. At the same time, the ARCH model, GARCH model, SVR rolling model, BP neural network, time series model ARIMA, random forest, support vector regression, and other methods are derived, which have different improvements for the prediction accuracy of stocks [2].

In addition to this, multiple models are combined to deal with various factors simultaneously, and this combination of models greatly increases the accuracy of the predictions. Chen and Huang used the MC_SVR rolling model for stock price prediction research, based on the prediction results of ten stocks, the model shows that it has stronger generalization ability and higher prediction accuracy. This paper is based on random wandering regression - Markov chain - for prediction, however, there are still many blind spots in Markov chain prediction [3].

In this paper, the stock prices of two stocks with different characteristics in the same period are used for prediction based on the Markov chain regression model. Firstly, the stock price data of the given period is counted, and based on this data, the one-time transfer probability is derived thus constructing the transfer probability matrix, and then the prediction transfer matrix is derived through the C-K equation. Eventually, a multi-step prediction is started from the last stock price state in the period, and then the multi-step prediction results are used to compare with the actual real stock price, and finally test whether the predicted stock price is in line with the actual stock price. The results of the two forecasts are finally compared. This result can help to further optimize the stock price prediction model for more accurate prediction.

2. Theory of Markov Chain

A Markov chain is a stochastic process with a "memoryless property", which means that the future state of a state depends only on the current state, not on a sequence of previous states. In a Markov chain, a system can transition between a series of states, and these transitions are determined by a set of probabilistic rules. The characteristics of Markov chains is shown as [4]

$$P\{X_n = j \mid X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \cdots, X_0 = i_0\} = P\{X_n = j \mid X_{n-1} = i_{n-1}\} \tag{1}$$

In Eq. (1), $X_n = j$ represents the state of the system at moment n as j. Also, equation (1) exhibits the nature of Markov chains which is memoryless. $P\{X_n = j \mid X_{n-1} = i_{n-1}\}$ is the one-step transfer probability in the Markov chain, denoted as P_{ij}, which represents the probability of the next transfer from state i to j. The probability of transferring from state i to state j after n steps is denoted as $P_{ij}^{(n)}$. The matrix consisting of a one-step probability transfer is called the probability transfer matrix P, i.e.,

$$\mathrm{P} = \begin{pmatrix} \mathrm{P}_{00} & \cdots & \mathrm{P}_{0m} \\ \vdots & \ddots & \vdots \\ \mathrm{P}_{n0} & \cdots & \mathrm{P}_{nm} \end{pmatrix} \tag{2}$$

Let the initial state probability vector be $\pi(0)$ and the nth state probability vector be $\pi(n)$, P^n represents the n-th power of the one-step transfer probability. And because of the memoryless nature of Markov chains, the current state is only related to the previous step state, which can be obtained as follows in equation [5]

$$\pi(n) = \pi(0) \bullet P^n \tag{3}$$

When using Markov chains to predict stock prices, the data needs to be tested for "Markovianity". If the research and analysis based on the Markov chain is not tested for Markovianity before operation, the results of the experiment and analysis will end up with great errors and uncertainties, and the conclusions drawn will also be very unscientific and uncritical. The following gives the specific theorem of Markov property test.

Firstly, assume that the data under discussion has m states, denote by f_{ij} the frequency of transfer from the original state i to state j in the time series $x_1, x_2, x_3, \cdots, x_n$, $i, j \in S$. The value obtained by dividing the jth column of the transfer frequency matrix by the sum of the rows and columns is denoted as the "marginal probability" and is given by $P_{\bullet j}$ [6]

$$P_{\bullet j} = \frac{\sum_{i=1}^{m} f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}} \tag{4}$$

Then the χ^2-distribution of the statistic $\chi^2 = 2\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij} \left| log \frac{P_{ij}}{P_{\bullet j}} \right|$ with degree of freedom $(m-1)^2$

is the limiting distribution. Where $P_{ij} = \frac{f_{ij}}{\sum_{j=1}^{i} f_{ij}}$, and given the significance level α, $\{x\}$ is consistent with the characterization of Markovianity if $\chi^2 > \chi_\alpha^2\left((m-1)^2\right)$ and otherwise it is not consistent with Markovianity.

3. Application

In this paper, the author takes Lululemon stock as an example and analyze the stock price data collected from May 6, 2024 to July 1, 2024 on Yahoo Finance to predict the trend of the stock price using Markov chain model. A total of 42 time series data were collected in this experiment to analyze the daily stock price fluctuation and trend, see Table 1.

Step 1: Collecting stock data. This paper uses the stock of Lululemon and collects the daily closing price, maximum price, and chain growth rate from 2024/5/6-2024/7/7.

Table 1. Lululemon Stock data table 2024/5/6-2024/7/5

Date (2024)	Closing price	Highest price	Sequential Growth Rate	State	Date	Closing price	Highest price	Sequential Growth Rate	State
5/6	350.24	357.90	-1.38%	S_3	6/5	308.27	308.76	-4.98%	S_3
5/7	349.85	354.16	-11.98%	S_2	6/6	323.03	337.76	3.93%	S_4
5/8	345.61	349.02	13.49%	S_5	6/7	317.86	329.73	-3.46%	S_3
5/9	352.95	355.20	-7.67%	S_3	6/10	318.26	322.03	0.15%	S_4
5/10	352.96	353.38	6.99%	S_4	6/11	318.04	319.70	-4.02%	S_3
5/13	347.16	355.44	-8.03%	S_3	6/12	309.81	321.95	3.11%	S_4
5/14	352.35	353.39	11.04%	S_5	6/13	307.49	309.14	-1.96%	S_3
5/15	346.85	354.53	-9.83%	S_2	6/14	306.01	308.00	0.59%	S_4
5/16	338.28	347.68	10.90%	S_5	6/17	312.91	313.22	-3.14%	S_3
5/17	334.95	338.75	-9.89%	S_2	6/18	313.23	315.50	4.42%	S_4
5/20	327.07	334.21	9.16%	S_5	6/20	310.77	314.66	-3.17%	S_3
5/21	322.98	328.29	-10.76%	S_2	6/21	311.82	313.45	0.05%	S_4
5/22	299.63	307.92	13.73%	S_5	6/24	312.28	316.53	1.41%	S_4
5/23	299.74	303.66	-12.73%	S_2	6/25	309.07	317.21	-0.18%	S_3
5/24	303.01	306.92	12.80%	S_5	6/26	304.81	309.88	3.18%	S_4
5/28	295.25	304.42	-11.77%	S_2	6/27	308.30	308.44	-3.31%	S_3
5/29	298.54	301.45	10.55%	S_5	6/28	298.70	303.25	1.64%	S_4
5/30	302.90	307.38	-7.50%	S_3	7/1	302.36	302.58	-2.05%	S_3
5/31	311.99	313.33	7.04%	S_4	7/2	301.67	304.39	2.15%	S_4
6/3	306.62	314.13	-6.48%	S_3	7/3	300.32	302.59	-2.82%	S_3
6/4	306.78	309.28	4.42%	S_4	7/5	298.14	303.14	-4.98%	S_3

ISSN 2959-6157

Table 2 Transfer of probability of Sequential Growth Rate

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	0	0	0	0	0	0
S_2	0	0	0	0	6	0
S_3	0	1	1	12	1	0
S_4	0	0	12	1	0	0
S_5	0	5	2	0	0	0
S_6	0	0	0	0	0	0

Step 2: Defining the state space. This paper adopts the mean-mean-variance grouping method, which is one of the most used methods, it is through the calculation of the mean and variance of the indicator values, centered on the mean, and grouped by the variance as the main criterion. Assuming that the sequence of indicator values is $x_1, x_2, x_3, \cdots, x_n$, the mean value is $\bar{x}$,which is calculated as follows [7]

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} \quad (5)$$

The variance is S and is calculated as follows:

$$S = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(x_t - \bar{x})} \quad (6)$$

Applied to the current data collection, the indicator values were grouped into six groups, which are $S_1:(-\infty, -15.90\%]$, $S_2:(-15.90\%, -8.53\%]$, $S_3:(-8.53\%, -0.16\%]$, $S_4:(-0.16\%, 7.37\%]$, $S_5:(7.37\%, 14.74\%]$, and $S_6:(14.74\%, +\infty)$ [8].

Step 3: To construct the transfer matrix. The transfer of the price growth rate of the stock can be derived from Table 1, which leads to Table 2. From Table 2, the transfer probability matrix f_{ij}, p_{ij} can be constructed, i.e.,

$$f_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 0 \\ 0 & 1 & 1 & 12 & 1 & 0 \\ 0 & 0 & 12 & 1 & 0 & 0 \\ 0 & 5 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, p_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1/15 & 1/15 & 12/15 & 1/15 & 0 \\ 0 & 0 & 12/13 & 1/13 & 0 & 0 \\ 0 & 5/7 & 2/7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (7)$$

Step 4: Testing for Markovianity. The marginal transfer probabilities can be derived from p_{ij}, which is the sum of the transfer probabilities for each column, which are $p_1 = 0, p_2 = \frac{82}{105}, p_3 = \frac{301}{273}, p_4 = \frac{122}{65}, p_5 = \frac{1}{15}, p_6 = 0$, respectively.

Table 3. Calculation table for statistic $\chi^2 = 2\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}\left|log\frac{P_{ij}}{P_{\bullet j}}\right|$

State	$f_{i1}\left\|log\frac{P_{i1}}{P_{\bullet 1}}\right\|$	$f_{i2}\left\|log\frac{P_{i2}}{P_{\bullet 2}}\right\|$	$f_{i3}\left\|log\frac{P_{i3}}{P_{\bullet 3}}\right\|$	$f_{i4}\left\|log\frac{P_{i4}}{P_{\bullet 4}}\right\|$	$f_{i5}\left\|log\frac{P_{i5}}{P_{\bullet 5}}\right\|$	$f_{i6}\left\|log\frac{P_{i6}}{P_{\bullet 6}}\right\|$	Total
S_1	0	0	0	0	0	0	
S_2	0	0	0	0	−1.9656	0	−1.9656
S_3	0	−1.0356	−1.2184	−4.4604	0	0	−6.7144
S_4	0	0	−0.9300	−1.3871	0	0	−2.3171

S_5	0	−1.5305	−1.7102	0	0	0	−3.2407
S_6	0	0	0	0	0	0	0
Total	0	−2.5661	−3.8586	−5.8475	−1.9656	0	−14.2378

According to Equation $\chi^2 = 2\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}\left|log\left(P_{ij} / P_{\cdot j}\right)\right|$, one can calculate the Chi-Square Statistic $\chi^2 = 2\bullet(-14.2378) = -28.4756$ according to Table 3. Since the Chi-Square statistic should not be negative, people need to take the above result to its absolute value $\chi^2 = 28.4756$. According to the look up table, the Chi-square critical value of significance level $\alpha = 0.05$ is about 37.652 with degree of freedom $(m-1)^2 = 25$. Since the calculated value of χ^2 (28.4756) is less than the critical value (37.652), this set of data is consistent with the Mahalanobis nature [9].

Step 5: To construct the multi-order transfer probability matrix. By using the Chapman-Kolmogorov equation, the state transfer probability matrices of order 2-4 are computed

$$P_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0.714 & 0.286 & 0 & 0 & 0 & 0 \\ 0.053 & 0.053 & 0.702 & 0.058 & 0.067 & 0 \\ 0.062 & 0.062 & 0.809 & 0.067 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.286 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{8}$$

$$P_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.286 & 0 \\ 0.094 & 0.066 & 0.615 & 0.051 & 0.053 & 0 \\ 0.054 & 0.054 & 0.710 & 0.059 & 0.062 & 0 \\ 0.204 & 0.082 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{9}$$

and

$$P_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0.204 & 0.082 & 0 & 0 & 0 & 0 \\ 0.079 & 0.056 & 0.539 & 0.045 & 0.066 & 0 \\ 0.091 & 0.065 & 0.622 & 0.052 & 0.054 & 0 \\ 0 & 0 & 0 & 0 & 0.082 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{10}$$

Based on the state on July 5 as the initial state vector, the growth rate on July 5 is -4.98% (state is S_3), so the initial state vector is set to $[0,0,1,0,0,0]$. Because of Eq. (3), the stock price after July 5 can be predicted by multiplying the initial state vector and the first order state transfer probability matrix. It is found that the relation $\pi(1) = \pi(0)P = [0,0,0.9231,0.0769,0,0]S_3$ and $\pi(2) = \pi(1)P = [0,0.0667,0.0667,0.8,0.0667,0]S_4$. Repeat this step to get the Table 4 shown as follows.

ISSN 2959-6157

Table 4. Transfer probabilities

$\pi(3)$	$\pi(4)$	$\pi(5)$	$\pi(6)$	$\pi(7)$	$\pi(8)$	$\pi(9)$	$\pi(10)$	$\pi(11)$	$\pi(12)$	$\pi(13)$	$\pi(14)$
S_3	S_4	S_3	S_4	S_3	S_4	S_3	S_4	S_3	S_4	S_3	S_4

The predicted and actual results have been between states S and X, resulting in the inability to accurately predict state intervals due to smaller gaps between predicted and actual results many times. So, in this paper, based on the characteristics of the stock (Lululemon) of the first test, high price, consistently low price and high floating frequency, another stock with lower price and relatively stable price trend and floating frequency is chosen for the second test. The result is shown in Table 5.

Table 5. Stock data table 2024/7/5-2024/7/25

Date (2024)	Closing price	Highest price	Sequential Growth Rate	State	Date	Closing price	Highest price	Sequential Growth Rate	State
7/8	294.03	301.91	-1.37%	S_3	7/17	292.68	293.94	0.31%	S_4
7/9	289.87	293.70	-1.41%	S_3	7/18	285.13	295.35	-2.58%	S_3
7/10	288.08	290.21	-0.62%	S_3	7/19	280.24	287.00	-1.72%	S_3
7/11	289.20	295.50	0.39%	S_4	7/22	285.00	286.43	1.70%	S_4
7/12	291.06	296.38	0.64%	S_4	7/23	281.37	287.25	-1.28%	S_3
7/15	283.72	290.00	-2.52%	S_3	7/24	272.06	281.00	-3.31%	S_3
7/16	291.76	291.87	2.83%	S_4	7/25	247.32	262.00	-9.09%	S_2

In this paper, stock data of EDU (New Oriental Education & Technology Group Inc.) from 2024/5/6-20247/5 with high price, closing price, and YoY growth rate are collected from Yahoo finance, and the raw data are shown in Table 6.

Table 6 EDU Stock data table 2024/5/6-2024/7/5.

Date (2024)	Closing price	Highest price	Sequential Growth Rate	State	Date	Closing price	Highest price	Sequential Growth Rate	State
5/6	83.99	85.17	-0.89%	S_3	6/5	76.63	76.72	-0.52%	S_3
5/7	81.64	82.42	-2.80%	S_3	6/6	78.35	78.97	2.24%	S_5
5/8	82.32	84.67	0.83%	S_3	6/7	77.53	78.53	-1.05%	S_3
5/9	83.98	84.27	2.02%	S_5	6/10	78.50	79.76	1.25%	S_4
5/10	84.17	85.15	0.23%	S_3	6/11	76.19	79.00	-2.95%	S_3
5/13	87.68	87.75	4.17%	S_3	6/12	76.48	76.95	0.38%	S_3
5/14	88.62	89.33	1.07%	S_4	6/13	76.55	78.30	0.09%	S_3
5/15	87.24	88.53	-1.56%	S_3	6/14	75.18	76.90	-1.79%	S_3
5/16	83.74	84.88	-4.01%	S_2	6/17	73.72	75.36	-1.94%	S_3
5/17	82.17	84.15	-1.87%	S_3	6/18	73.00	73.95	-0.98%	S_3
5/20	81.37	83.50	-0.97%	S_3	6/20	71.98	72.77	-1.40%	S_3
5/21	79.75	80.21	-1.99%	S_3	6/21	71.34	71.53	-0.89%	S_3

5/22	79.98	81.13	0.29%	S_3	6/24	76.76	77.63	7.59%	S_6
5/23	79.21	79.80	-0.96%	S_3	6/25	76.45	77.64	-0.40%	S_3
5/24	78.12	79.15	-1.38%	S_3	6/26	77.36	78.04	1.19%	S_4
5/28	81.38	82.48	4.17%	S_5	6/27	76.42	76.89	-1.21%	S_3
5/29	84.44	85.37	3.76%	S_5	6/28	77.73	78.87	1.72%	S_4
5/30	83.78	84.60	-0.78%	S_3	7/1	79.00	79.65	1.63%	S_4
5/31	79.93	82.21	-4.59%	S_2	7/2	80.18	80.37	1.49%	S_4
6/3	79.43	81.40	-0.63%	S_3	7/3	82.38	83.34	2.74%	S_5
6/4	77.03	79.01	-3.02%	S_2	7/5	79.56	82.37	-3.43%	S_2

Step 1: Defining the state space. This paper adopts the mean-mean-variance grouping method, which is one of the most commonly used methods, it is through the calculation of the mean and variance of the indicator values, centered on the mean, and grouped by the variance as the main criterion. Based on Eq. (5) and Eq. (6), six groupings can be derived, which are $S_1:(-\infty,-5.490\%]$, $S_2:(-5.49\%,-2.701\%]$, $S_3:(-2.701\%,0.088\%]$, $S_4:(0.088\%,2.789\%]$, $S_5:(2.789\%,5.678\%]$, and $S_6:(5.678\%,+\infty)$.

Step 2: To construct the transfer matrix. The transfer probability matrix can likewise be constructed in the same steps as previously. Namely,

$$f_{ij}=\begin{pmatrix}0&0&0&0&0&0\\0&0&3&0&0&0\\0&3&14&4&3&1\\0&0&3&2&1&0\\0&1&3&0&0&0\\0&0&1&0&0&0\end{pmatrix},\ p_{ij}=\begin{pmatrix}0&0&0&0&0&0\\0&0&1&0&0&0\\0&3/25&14/25&4/25&3/25&1/25\\0&0&1/2&1/3&1/6&0\\0&1/4&3/4&0&0&0\\0&0&1&0&0&0\end{pmatrix}. \tag{11}$$

Step 3: Testing for Markovianity. The marginal transfer probabilities can be derived from p_{ij}, which is the sum of the transfer probabilities for each column, which are $p_1=0, p_2=\frac{37}{100}, p_3=\frac{281}{100}, p_4=\frac{37}{75}, p_5=\frac{43}{150}, p_6=\frac{1}{25}$, respectively. Consistent with the previous steps, the data is tested for Markovianity based on the calculated marginal transfer probabilities. Due to $\chi^2=1.0157<37.652$, the data can be assumed to satisfy Markovianity at a given significance level $\alpha=0.05$ [10].

Step 4: To construct the multi-order transfer probability matrix. By using the Chapman-Kolmogorov equation, the state transfer probability matrices of order 2-14 are computed. Namely,

$$P_2=\begin{pmatrix}0&0&0&0&0&0\\0&0.12&0.56&0.16&0.12&0.04\\0&0.0972&0.6436&0.1429&0.067&0.0224\\0&0.1017&0.5717&0.1911&0.1156&0.02\\0&0.09&0.67&0.12&0.09&0.03\\0&0.12&0.56&0.16&0.12&0.04\end{pmatrix}, \tag{12}$$

$$P_3=\begin{pmatrix}0&0&0&0&0&0\\0&0.0972&0.6436&0.1429&0.09387&0.0224\\0&0.1007&0.6219&0.1506&0.1011&0.0257\\0&0.0975&0.6240&0.1552&0.1005&0.0229\\0&0.1029&0.6277&0.1472&0.1004&0.0268\\0&0.0972&0.6436&0.1429&0.0939&0.0224\end{pmatrix}, \tag{13}$$

ISSN 2959-6157

and also others

$$P_{14} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0.625 & 0.15 & 0.1 & 0.025 \\ 0 & 0.1 & 0.625 & 0.15 & 0.1 & 0.025 \\ 0 & 0.1 & 0.625 & 0.15 & 0.1 & 0.025 \\ 0 & 0.1 & 0.625 & 0.15 & 0.1 & 0.025 \\ 0 & 0.1 & 0.625 & 0.15 & 0.1 & 0.025 \end{pmatrix}. \quad (14)$$

These results show a gradual stabilization of the state transfer vectors as the order increases, indicating that the system gradually reaches a steady state distribution (see Table 7).

Table 7. EDU Stock data table 2024/7/8-2024/7/25.

Date (2024)	Closing price	Highest price	Sequential Growth Rate	State	Date	Closing price	Highest price	Sequential Growth Rate	State
7/8	79.10	79.99	-0.57%	S_3	7/17	72.08	73.86	-2.23%	S_3
7/9	80.36	80.50	1.59%	S_4	7/18	71.78	73.27	-0.42%	S_3
7/10	79.80	81.85	-0.70%	S_3	7/19	74.05	74.06	3.17%	S_5
7/11	76.20	79.59	-4.52%	S_2	7/22	76.32	77.91	3.06%	S_5
7/12	77.54	79.65	1.76%	S_4	7/23	75.61	76.01	-0.93%	S_3
7/15	74.25	76.65	-4.24%	S_2	7/24	74.39	75.67	-1.61%	S_3
7/16	73.72	74.95	-0.71%	S_3	7/25	70.33	73.82	-5.45%	S_2

4. Conclusion

In this paper, based on the fundamental properties of Markov chains, an attempt has been made to predict the stock prices of two stocks using the prediction of multiple probability transfer vectors. Although the prediction of each stock shows good prediction results in a shorter period, once the prediction period is lengthened, many prediction errors and biases occur. When the first stock was predicted, the prediction results were inaccurate due to a little bias many times because the predicted multiple probability transfer vectors appeared to oscillate between the S3 and S4 states. Thus, based on the nature of the first stock, the author used a second stock with a lower price and a calmer trend for the prediction. Although the forecasts improved in both the long and short term, there was still a significant amount of bias especially in the long-term forecast range. Accuracy is very important for a stock price forecasting model, especially for the long-time horizon. Many existing stock price forecasting models have different advantages and characteristics, and there are some disadvantages to combining more forecasting models or using more powerful algorithms to support computers or AI to complete the forecasting. This paper hopes to make more targeted adjustments to the Markov chain-based prediction model based on the results of this paper so that its accuracy and reliability can be further improved.

References

[1] Zhiqian Liu. Markov chain-based forecasting of stock price fluctuations. Cooperative economy and science and technology, 2021, (24): 64-67.

[2] Yufen Zhang & Yalin Zhu. A study of Markovianity and its testing methods.value engineering. 2012, (02): 312-313.

[3] Yang Zhenpei. A Study of Stock Price Prediction Models Based on Kernel Clustering Localized Sliced Inverse Regression and Bayesian Model Averaging. Highlights in Business, Economics and Management, 2024, 38: 157–66.

[4] Wei Zhang. Research on rainfall prediction in Zhengzhou based on Markov chain (Master's Thesis, North China University of Water Conservancy and Hydropower), 2019.

[5] Motahari, H., Noudehi, M. G., Jamali, T., & Jafari, G. R. Generalization of Chapman-Kolmogorov equation to two coupled processes, 2024.

[6] IaconoM., LevinsonD., El-GeneidyA., & WasfiR. A Markov Chain Model of Land Use Change. TeMA - Journal of Land Use, Mobility and Environment, 2015, 8(3): 263-276.

[7] Wu, G. The first, second, third and fourth order Markov chain analysis on the amino-acid sequence of human dopamine β-hydroxylase. Molecular Psychiatry, 2000, 5(4), 448–451.

[8] Mallak, Saed; Kanan, Mohammad; Al-Ramahi, Nidal; Qedan, Aya; Khalilia, Hadi; and Khassati, Ahmad. Using Markov Chains and Data Mining Techniques to Predict Students' Academic Performance. Information Sciences Letters, 2023,

12(9): 2073–83.
[9] Kim, D. J., Sur, J. M., & Cho, H. U. A long term expected risk estimation of maritime accidents through Markov chain approach and probabilistic risk matrix. The Asian Journal of Ship and Logistics, 2023, 39(3), 1-12.
[10] Engbring, K., Boriskovsky, D., Roichman, Y., & Lindner, B. A nonlinear fluctuation-dissipation test for Markovian systems. Physical Review X, 2023, 13(2), 021034.