Statistics and Applications {ISSN 2454-7395 (online)}
Volume 23, No. 1, 2025 (New Series), pp 249–272
https://www.ssca.org.in/journal

Three-State Markov Probability Distributions for the Stock Price Prediction

Tirupathi Rao Padi[1], Sarode Rekha[2] and Gulbadin Farooq Dar[3]
[1,2]*Department of Statistics Pondicherry University, Puducherry-605014*
[3]*Department of Statistics, St. Joseph's University, Bangalore-560027*

Received: 05 December 2023; Revised: 09 April 2024; Accepted: 10 June 2024

Abstract

This paper focuses on predicting the movement of State Bank of India (SBI) stock prices using the Markov model, a challenging task in financial markets. It comprises two main sections: Firstly, it formulates probability distributions for various states using Markov model parameters, deriving Pearson's coefficients like average, variance, skewness, and kurtosis. Secondly, real-time SBI data is gathered and divided into five datasets representing each business day. Numerical calculations are performed using R software, computing parameters such as transition probability matrix (TPM) and initial probability vector (IPV) for each dataset. Expected returns and closing price predictions are determined, validated through the Chi-square test for goodness of fit, and assessed for robustness using Akaike information criterion (AIC) and Bayesian information criterion (BIC). The model is designed to facilitate optimal investment strategies and could benefit from user-friendly digital interfaces for traders. It explores indicators such as timing for buying/selling, probability of price movements, expected gains/losses, and estimated closing prices to enhance understanding of SBI's market behaviour in the Indian context.

Key words: Markov model; Share price; Probability distribution; Transition probability; Initial probability.

AMS Subject Classifications: 62K05, 05B05.

1. Introduction

Mathematical and stochastic modeling are pivotal tools in unraveling the intricacies of the stock market and projecting its future trends. This paper is dedicated to crafting a three-state Markov probability distribution model to analyse stocks and anticipate their forthcoming price fluctuations.

Certainly, at the core of every nation's economic structure lies an indispensable link with its stock market. This intricate connection serves as the lifeblood of the country's financial stability, embodying a sophisticated network where a diverse spectrum of individuals and entities converge. Participants engage in a multifaceted interplay of buying and

Corresponding Author: Sarode Rekha
Email: statisticsrekha@gmail.com

selling an extensive array of financial instruments within this bustling marketplace. This dynamic interaction propels economic activities and nurtures an environment teeming with opportunities for trading and investment.

Amidst this whirlwind of financial transactions, traders emerge as the linchpin, fueled by the relentless pursuit of optimal outcomes. Their endeavours are marked by a meticulous analysis of market trends, a process that involves an exhaustive examination of historical data, intricate technical analyses, and a keen understanding of global economic influences. Armed with this knowledge, traders navigate the intricate pathways of the market, making calculated and strategic decisions.

Crucially, these traders are not guided by mere intuition but rather by a commitment to informed decision-making. They implement sophisticated risk management strategies, diligently assessing potential risks and rewards. Their objective is crystal clear: to optimise profits and minimise losses. Every move within this dynamic financial landscape is a result of careful consideration, a balance between seizing opportunities and mitigating risks.

In the grand tapestry of the stock market, the significance of this strategic decision-making cannot be overstated. It not only influences individual financial destinies but also ripples through the larger economic fabric of the nation. The stock market becomes a barometer, reflecting the collective confidence and sentiment of investors, thereby shaping the economic trajectory of the entire country.

In essence, the stock market embodies more than just financial transactions; it symbolizes the aspirations, strategies, and challenges of a nation's economic journey. Armed with their expertise and insights, traders play a pivotal role in shaping this intricate landscape, where every decision made resonates far beyond individual portfolios, weaving into the intricate tapestry of a nation's economic prosperity.

A Markov regime ARCH model used to investigate and analyse the volatility within market behaviour (Cai (1994)). There is sufficient evidence on the usage of Wiener–Hopf results for solving the option pricing problems with the Markov processes (Jobert and Rogers (2006)). When applied to forecast data from the stock market, the HMM with fuzzy model innovation produced results that were more accurate than those from forecasting models like ARIMA, ANN, *etc.* (Hassan (2009)). A flexible Mixed HMM approach that considers temporal and spatial variability. This method is adaptable because it can handle the distinctive features of financial time series data, such as asymmetry, kurtosis, and unobserved heterogeneity (Dias *et al.* (2010)). HMM and support vector machines were used to predict the movement of the stock price (Rao and Hong (2010)). The stock price dynamics were examined through a semi-Markov return model (D'Amico and Petroni (2012)). A finite state Markov chain model was used to evaluate share price movements in the share market (Choji *et al.* (2013)). The utilization of a Markov-switching using GARCH approach has provided a method for predicting the volatility in the Tehran Stock Exchange-TSE (Abounoori *et al.* (2016)). The Nigerian Stock Exchange market has utilized the Markov chain model for analysing its behaviour (Adesokan *et al.* (2017)). The Markov chain model was used to forecast the stock price movement of the Taiwanese company High Tech Computer (Huang *et al.* (2017)). The Markov chain is used in forecasting the behaviour of the Nepal Stock Exchange Index (Bhusal (2017)). The Markov chain model is used to predict the stock market trend in the context of the Indian stock market (Padi *et al.* (2022)). The HMM was utilised to properly comprehend the financial factors in the stock market, and the results were more helpful for portfolio managers in making the best choices (Dar *et al.* (2022)). The impact of international trade on the share prices of the Industrial Bank of Korea was assessed through the utilization of stochastic prediction modelling (Dar *et al.* (2023)).

Numerous studies have predominantly concentrated on classical methodologies for either developing new models or applying existing Markov models to forecast market behaviour. However, there exists a dearth of research on deriving probability distributions for sequences of states and estimating parameters through predictive modeling, specifically tailored to Markov processes. Delving into the probability distributions of transitional states can furnish more precise information inputs. The parametric estimation within the Markov model and its extension into probability distributions have been largely overlooked by probability researchers.

In response to this research gap, our study underscores the importance of Markov modeling in formulating probability distributions by constructing the Markov model based on parameters such as TPM and IPV. We have mathematically derived explicit relationships for various statistical measures using these formulated probability distributions. Focusing on three states - *Rise State*, *Stable State*, and *Fall State* - of SBI shares, our General Markov model entails two key parameters: TPM, governing transitions among states, and IPV, describing the likelihood of each state's initial occurrence. Our primary objective is to establish probability distributions separately for *Rise State*, *Stable State*, and *Fall State* across all segregated data sets for different business days. We have derived explicit mathematical relationships for diverse statistical measures and Pearson's coefficients. Sensitivity analysis has been conducted by determining Markov model parameters, obtaining probability distributions, and analysing statistical measures to gain a comprehensive understanding of SBI share price behaviour. Additionally, our model encompasses an additional study where expected returns and closing prices of SBI are computed using the formulas outlined in Section 2.7.

2. Stochastic model

The Markov model is a type of mathematical model that focuses on predicting the next event based on the event that happened just before it, without considering events from a long time ago. This means, it doesn't have a memory of past events beyond the most recent one. The schematic diagram for the model is placed below.

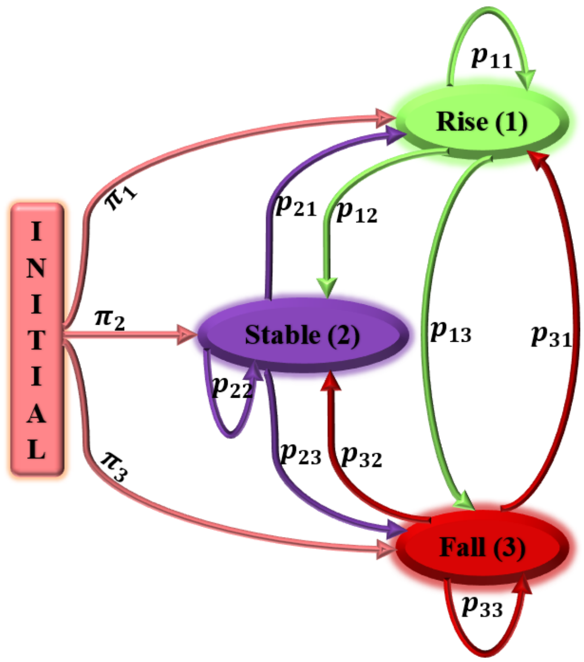

Figure 1: Schematic Diagram of Three-State Markov Model

In this study, the main aim is to figure out the likelihood of different states happening. These states are divided into three categories: *Rise State*, *Stable State*, and *Fall State*. The Markov model consists of two key parameters namely TPM and IPV.

2.1. Transition probability matrix (TPM)

A Transition Probability Matrix (TPM) is often called a Stochastic Matrix. It is defined as

$$P = \begin{array}{cc} & Y_n \\ Y_{n-1} & \left(P_{jk}\right) \end{array} \qquad \forall j, k = 1, 2, 3$$

$P\{Y_n = k/Y_0 = 1, Y_1 = 2, ..., Y_{n-1} = j\} = P[Y_n = k/Y_{n-1} = j] = P_{jk}$ be the transition probability from jth state to kth state. Every TPM must satisfy the following conditions like,

- The matrix must possess equal numbers of rows and columns; *i.e.*, TPM is a squared matrix.
- Each element within the matrix must represent a probability; *i.e.*, $P_{jk} \geq 0$.
- The sum of each row must be equivalent to one; *i.e.*, $\sum_{k=1}^{3} P_{jk} = 1, \forall j, k = 1, 2, 3.$

It earns the label “Doubly Stochastic Matrix” when the sums of both its each row and each column are equal to one.

2.2. Initial probability vector (IPV)

The initial probability vector determines the chance of happening in a particular state. It is denoted by π.

$$\pi = (\pi_1, \pi_2, \pi_3)$$

2.3. Notations and terminology

π_k : Initial probability for the k^{th} state, $\pi_k \geq 0$; for all k=1,2,3; $\sum_{k=1}^{3} \pi_k = 1$; $\pi_k = \frac{n_k}{n}$; $n = \sum_{k=1}^{3} n_k$, *i.e.*, Total number of observations considered for the study in the specific business day

p_{jk}: The transition probability between states j and k represents the likelihood of moving from state j to state k in a given system or process.
$i.e.$, $P\{Y_n = k/Y_{n-1} = j\} \geq 0$; $0 \leq p_{jk} \leq 1$ and $\sum_{k=1}^{3} p_{jk} = 1 \forall j = 1, 2, 3.$

j : Origin state

k : Destination state

y_t: Share price of the SBI on t^{th} day

Δy_t: $y_t - y_{t-1}$; The difference between the current day (t) share price and previous day (t-1) share price in SBI

dy_t: Derivative of the share price's return at time 't'; $dy_t = \frac{\Delta y_t}{y_{t-1}}$

R : *Rise State* occurs in SBI; $R = \left(dy_t \geq \mu + \frac{3\sigma}{\sqrt{n}}\right)$

S : *Stable State* occurs in SBI; $S = \left(\mu - \frac{3\sigma}{\sqrt{n}} < dy_t < \mu + \frac{3\sigma}{\sqrt{n}}\right)$

F : *Fall State* occurs in SBI; $F = \left(dy_t \leq \mu - \frac{3\sigma}{\sqrt{n}}\right)$

μ: Mean of dy_t

σ: standard deviation of dy_t

n : Total number of observations in the business day

m : Number of estimated values for testing the goodness of fit

O_i: Observed share value on i^{th} day; i=1,2, ..., n

E_i: Estimated share value on i^{th} day; i=1,2, ..., n

v : Number of parameters in the study of specific business day

$Y(\omega_1)$: Number of times *Rise State* occurs, $[Y(\omega_1) = y] = 0, 1$

$Y(\omega_2)$: Number of times *Stable State* occurs, $[Y(\omega_2) = y] = 0, 1$

$Y(\omega_3)$: Number of times *Fall State* occurs, $[Y(\omega_3) = y] = 0, 1$

2.4. Probability distribution and some statistical measures for ***Rise State***

2.4.1. The probability distribution for ***Rise State***

Let us consider a random variable denoted by $Y(\omega_1) = y$ which represents the happening of the *Rise State*. This variable can assume values 0 and 1, where '0' signifies its absence of the *Rise State* and '1' signifies its presence of the *Rise State*.

$$P[Y(\omega_1) = y] = \begin{cases} \sum\limits_{k=1}^{3}\sum\limits_{j=2}^{3} \pi_k p_{kj} & ;\text{for } y = 0 \\ \sum\limits_{k=1}^{3} \pi_k p_{k1} & ;\text{for } y = 1 \\ 0 & ;\text{otherwise}(y \geq 2) \end{cases} \tag{1}$$

2.4.2. Statistical measures for ***Rise State***

The Average Occurrence of *Rise State*

$$\mu_R = \sum_{k=1}^{3} \pi_k p_{k1} \tag{2}$$

The Variance of a *Rise State*

$$\sigma_R^2 = \mu_R^2\Big(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\Big) + (1-\mu_R)^2\Big(\sum_{k=1}^{3} \pi_k p_{k1}\Big) \tag{3}$$

The Third Central Moment for *Rise State*

$$\mu_{3R} = -\mu_R^3\Big(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\Big) + (1-\mu_R)^3\Big(\sum_{k=1}^{3} \pi_k p_{k1}\Big) \tag{4}$$

The Coefficient of skewness for *Rise State*

$$\begin{aligned}\beta_{1R} =& \Big[-\mu_R^3\Big(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\Big) + (1-\mu_R)^3\Big(\sum_{k=1}^{3} \pi_k p_{k1}\Big)\Big]^2 \times \\ & \Big[\mu_R^2\Big(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\Big) + (1-\mu_R)^2\Big(\sum_{k=1}^{3} \pi_k p_{k1}\Big)\Big]^{-3}\end{aligned} \tag{5}$$

Coefficient of Kurtosis for *Rise State*

$$\beta_{2R} = \Big[\mu_R^4\Big(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\Big) + (1-\mu_R)^4\Big(\sum_{k=1}^{3} \pi_k p_{k1}\Big)\Big]\Big[\mu_R^2\Big(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\Big) + (1-\mu_R)^2\Big(\sum_{k=1}^{3} \pi_k p_{k1}\Big)\Big]^{-2} \tag{6}$$

The Coefficient of Variation for *Rise State*

$$C.V_R = \Big[\mu_R^2\Big(\sum_{k=1}^{3}\sum_{j=2}^{3} \pi_k p_{kj}\Big) + (1-\mu_R)^2\Big(\sum_{k=1}^{3} \pi_k p_{k1}\Big)\Big]^{1/2}\Big[\sum_{k=1}^{3} \pi_k p_{k1}\Big]^{-1}\% \tag{7}$$

2.4.3. Moment generating function for *Rise State*

$$M_{YR}(t) = \Big(\sum_{k=1}^{3}\sum_{j=2}^{3}\pi_k p_{kj}\Big) + e^t\Big(\sum_{k=1}^{3}\pi_k p_{k1}\Big) \tag{8}$$

2.4.4. Characteristic function for *Rise State*

$$\phi_{YR}(t) = \Big(\sum_{k=1}^{3}\sum_{j=2}^{3}\pi_k p_{kj}\Big) + e^{it}\Big(\sum_{k=1}^{3}\pi_k p_{k1}\Big) \tag{9}$$

2.4.5. Probability generating function for *Rise State*

$$P_{SR}(t) = \Big(\sum_{k=1}^{3}\sum_{j=2}^{3}\pi_k p_{kj}\Big) + S\Big(\sum_{k=1}^{3}\pi_k p_{k1}\Big) \tag{10}$$

2.5. Probability distribution and some statistical measures for *Stable State*

2.5.1. The probability distribution for *Stable State*

Let us consider a random variable denoted by $Y(\omega_2) = y$ which represents the happening of the *Stable State*. This variable can assume values 0 and 1, where '0' signifies its absence of the *Stable State* and '1' signifies its presence of the *Stable State*.

$$P[Y(\omega_2) = y] = \begin{cases} \sum\limits_{k=1}^{3}\sum\limits_{j=1,j\neq 2}^{3}\pi_k p_{kj} & ; \text{for } y = 0 \\ \sum\limits_{k=1}^{3}\pi_k p_{k2} & ; \text{for } y = 1 \\ 0 & ; \text{otherwise}(y \geq 2) \end{cases} \tag{11}$$

2.5.2. Statistical measures for *Stable State*

The Average Occurrence of *Stable State*

$$\mu_S = \sum_{k=1}^{3}\pi_k p_{k2} \tag{12}$$

The Variance of a *Stable State*

$$\sigma_S^2 = \mu_S^2\Big(\sum_{k=1}^{3}\sum_{j=1,j\neq 2}^{3}\pi_k p_{kj}\Big) + (1-\mu_S)^2\Big(\sum_{k=1}^{3}\pi_k p_{k2}\Big) \tag{13}$$

The Third Central Moment for *Stable State*

$$\mu_{S3} = -\mu_S^3\Big(\sum_{k=1}^{3}\sum_{j=1,j\neq 2}^{3}\pi_k p_{kj}\Big) + (1-\mu_S)^3\Big(\sum_{k=1}^{3}\pi_k p_{k2}\Big) \tag{14}$$

The Coefficient of Skewness for *Stable State*

$$\beta_{1S} = \Big[-\mu_S^3 \Big(\sum_{k=1}^{3} \sum_{j=1, j\neq 2}^{3} \pi_k p_{kj} \Big) + (1-\mu_S)^3 \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big) \Big]^2 \Big[\mu_S^2 \Big(\sum_{k=1}^{3} \sum_{j=1, j\neq 2}^{3} \pi_k p_{kj} \Big) (1-\mu_S)^2 \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big) \Big]^{-3} \tag{15}$$

Coefficient of Kurtosis for *Stable State*

$$\beta_{2S} = \Big[\mu_S^4 \Big(\sum_{k=1}^{3} \sum_{j=1, j\neq 2}^{3} \pi_k p_{kj} \Big) + (1-\mu_S)^4 \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big) \Big] \Big[\mu_S^2 \Big(\sum_{k=1}^{3} \sum_{j=1, j\neq 2}^{3} \pi_k p_{kj} \Big) (1-\mu_S)^2 \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big) \Big]^{-2} \tag{16}$$

Coefficient of variation for *Stable State*

$$C.V_S = \Big[\mu_S^2 \Big(\sum_{k=1}^{3} \sum_{j=1, j\neq 2}^{3} \pi_k p_{kj} \Big) + (1-\mu_S)^2 \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big) \Big]^{1/2} \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big)^{-1} \% \tag{17}$$

2.5.3. Moment generating function for ***Stable State***

$$M_{YS}(t) = \Big(\sum_{k=1}^{3} \sum_{j=1, j\neq 2}^{3} \pi_k p_{kj} \Big) + e^t \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big) \tag{18}$$

2.5.4. Characteristic function for ***Stable State***

$$\phi_{YS}(t) = \Big(\sum_{k=1}^{3} \sum_{j=1, j\neq 2}^{3} \pi_k p_{kj} \Big) + e^{it} \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big) \tag{19}$$

2.5.5. Probability generating function for ***Stable State***

$$P_{SS}(t) = \Big(\sum_{k=1}^{3} \sum_{j=1, j\neq 2}^{3} \pi_k p_{kj} \Big) + S \Big(\sum_{k=1}^{3} \pi_k p_{k2} \Big) \tag{20}$$

2.6. Probability distribution and some statistical measures for ***Fall State***

2.6.1. The Probability distribution for ***Fall State***

Let us consider a random variable denoted by $Y(\omega_3) = y$ which represents the happening of the *Fall State*. This variable can assume values 0 and 1, where '0' signifies its absence of the *Fall State* and '1' signifies its presence of the *Fall State*.

$$P[Y(\omega_3) = y] = \begin{cases} \sum\limits_{k=1}^{3}\sum\limits_{j=1}^{2} \pi_k p_{kj} & ;\text{for } y = 0 \\ \sum\limits_{k=1}^{3} \pi_k p_{k3} & ;\text{for } y = 1 \\ 0 & ;\text{otherwise}(y \geq 2) \end{cases} \tag{21}$$

2.6.2. Statistical measures for *Fall State*

The Average Occurrence of *Fall State*

$$\mu_F = \sum_{k=1}^{3} \pi_k p_{k3} \tag{22}$$

The Variance of a *Fall State*

$$\sigma_F^2 = \mu_F^2\Big(\sum_{k=1}^{3}\sum_{j=1}^{2} \pi_k p_{kj}\Big) + (1-\mu_F)^2\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big) \tag{23}$$

The Third Central Moment for *Fall State*

$$\mu_F^3 = -\mu_F^3\Big(\sum_{k=1}^{3}\sum_{j=1}^{2} \pi_k p_{kj}\Big) + (1-\mu_F)^3\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big) \tag{24}$$

The Coefficient of Skewness for *Fall State*

$$\begin{aligned}\beta_{1F} =& \Big[-\mu_F^3\Big(\sum_{k=1}^{3}\sum_{j=1}^{2} \pi_k p_{kj}\Big) + (1-\mu_F)^3\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big)\Big]^2 \times \\ & \Big[\mu_F^2\Big(\sum_{k=1}^{3}\sum_{j=1}^{2} \pi_k p_{kj}\Big) + (1-\mu_F)^2\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big)\Big]^{-3}\end{aligned} \tag{25}$$

Coefficient of Kurtosis for *Fall State*

$$\beta_{2F} = \Big[\mu_F^4\Big(\sum_{k=1}^{3}\sum_{j=1}^{2} \pi_k p_{kj}\Big) + (1-\mu_F)^4\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big)\Big]\Big[\mu_F^2\Big(\sum_{k=1}^{3}\sum_{j=1}^{2} \pi_k p_{kj}\Big) + (1-\mu_F)^2\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big)\Big]^{-2} \tag{26}$$

The Coefficient of Variation for *Fall State*

$$C.V_F = \Big[\mu_F^2\Big(\sum_{k=1}^{3}\sum_{j=1}^{2} \pi_k p_{kj}\Big) + (1-\mu_F)^2\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big)\Big]^{1/2}\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big)^{-1}\% \tag{27}$$

2.6.3. Moment generating function for *Fall State*

$$M_{YF}(t) = \Big(\sum_{k=1}^{3}\sum_{j=1}^{2} \pi_k p_{kj}\Big) + e^t\Big(\sum_{k=1}^{3} \pi_k p_{k3}\Big) \tag{28}$$

2.6.4. Characteristic function for ***Fall State***

$$\phi_{YF}(t) = \Big(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\Big) + e^{it}\Big(\sum_{k=1}^{3}\pi_k p_{k3}\Big) \tag{29}$$

2.6.5. Probability generating function for ***Fall State***

$$P_{SF}(t) = \Big(\sum_{k=1}^{3}\sum_{j=1}^{2}\pi_k p_{kj}\Big) + S\Big(\sum_{k=1}^{3}\pi_k p_{k3}\Big) \tag{30}$$

2.7. Predictions of returns on income

2.7.1. Expected returns on SBI shares

The explicit mathematical relation for computing expected share returns

$$[E.S.R]_{3\times 1} = [p_{jk}]^{n}_{3\times 3}[M.S]_{3\times 1}; \forall\, n = 1, 2, ... \tag{31}$$

E.S.R =Expected share price returns
P^n = Limiting Probability Matrix (Computed using TPM)
M.S =Mean state

2.7.2. Prediction of closing prices of SBIs shares

The explicit mathematical relation for predicted Closing prices of SBI shares

$$P.S.P = (Y_{Rt} \times \mu_R) + (Y_{St} \times \mu_S) + (Y_{Ft} \times \mu_F) \tag{32}$$

where,

Y_{Rt} = Expected closing price of the SBIs share on the current day for the *Rise State*
Y_{St}= Expected closing price of the SBIs share on the current day for the *Stable State*
Y_{Ft}= Expected closing price of the SBIs share on the current day for the *Fall State*
μ_R= Average chance for occurrence of the *Rise State*
μ_S = Average chance for occurrence of the *Stable State*
μ_F = Average occurrence for occurrence of the *Fall State*

2.8. Validation of the model

2.8.1. Testing for model's goodness of fit

The Chi-Square test statistic, denoted as χ^2 , is utilized to assess the goodness of fit between observed and expected categorical data. In the context of comparing observed (original) and expected (predicted) share prices, the formula for χ^2 is:

$$\chi^2 = \sum_{i=1}^{m}\frac{[O_i - E_i]^2}{E_i} \sim \chi^2_{m-1} \tag{33}$$

2.8.2. Computation of AIC and BIC

The formulas for calculating AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) are as follows:

$$AIC = -2\, loglikelihood + 2v \tag{34}$$

$$BIC = -2\, loglikelihood + v\; 2logn \tag{35}$$

3. Data description for the developed model

Figure 2 provides a clear depiction of the data and methodology employed in the current study. It delineates the detailed procedures utilized to assess the results with precision and thoroughness.

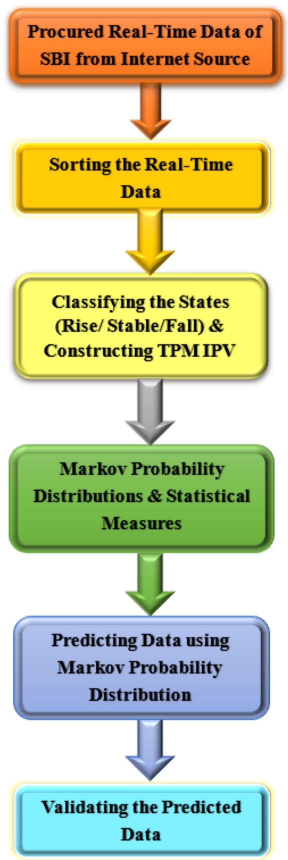

Figure 2: Three-State Markov model data description & methodology flow chart

3.1. Data source and organization of the data

The detailed description of Figure 2; in order to utilise the Markov model that was developed, real-time data on the closing prices of SBI (State Bank of India) stocks was considered. This real-time data, crucial for evaluating the model's effectiveness, consisted of 251 observations collected over a period spanning from 2^{nd} May 2022 to 5^{th} May 2023.

These observations were meticulously sourced from the renowned financial platform, Yahoo Finance, accessible via the internet link (https://in.finance.yahoo.com).

The dataset, which served as the foundation for this analysis, was specifically focused on the closing prices of SBI stocks. Closing prices, in the context of stock market analysis, represent the final prices at which a stock trades during a regular trading session. These prices are often used to assess the overall performance of a particular stock.

This dataset, constituting 251 data points, is of paramount importance for evaluating the Markov model's predictive capabilities in real-world scenarios. It forms the basis upon which the model's predictions and effectiveness in forecasting SBI stock prices are tested and validated. The historical closing prices, meticulously organized and structured, were compiled into a comprehensive sample data template, as detailed in Table 1. This template serves as the primary reference for the subsequent analysis and assessment of the Markov model's accuracy and reliability in predicting the closing prices of SBI stocks during the specified period.

Table 1: SBI's sample data matrix

S.No.	Date	Closing Price
1	02-05-2022	491
2	04-05-2022	479.649994
3	05-05-2022	480
⋮	⋮	⋮
249	03-05-2023	570.5
250	04-05-2023	580
251	05-05-2023	576.5

3.2. Data formulation

In light of the observed influence of market seasonality on closing prices concerning specific weekdays, the 251 collected observations were categorized based on business days (Monday, Tuesday, Wednesday, Thursday, and Friday). The Sample data matrix of Mondays data placed in Table 2. Remaining business days also done like Mondays data.

Table 2: SBI's Monday data matrix

S.No.	Date	Closing Price	Returns	State	Transition
1	02-05-2022	491	-	-	-
2	09-05-2022	475.899994	-0.03075	F	-
3	16-05-2022	455	-0.04392	F	FF
⋮	⋮	⋮	⋮	⋮	⋮
49	10-04-2023	526.299988	-0.00085	S RS	
50	17-04-2023	544	0.033631	R	SR
51	24-04-2023	554.599976	0.019485	R	RR

To delve deeper into this segmentation and its impact, individual sensitivity studies were undertaken for each business day. Prior to these studies, the data from all five datasets were pooled together, facilitating comprehensive analysis. Within each dataset, a meticulous classification was performed, focusing on the transient state of returns. This systematic approach allowed for a detailed exploration of how market dynamics and price fluctuations varied across different weekdays, shedding light on the intricate relationship between market behaviour and specific business days.

3.3. Data disclosure

The states are determined according to the values of dY_t and are categorized into three distinct types: *Rise*(State-1) when the condition $dY_t \geq \mu + \frac{3\sigma}{\sqrt{n}}$ is met, *Stable* (State-2) when the condition $\mu - \frac{3\sigma}{\sqrt{n}} < dY_t < \mu + \frac{3\sigma}{\sqrt{n}}$ is satisfied, and *Fall* (State-3) when the condition $dY_t \leq \mu - \frac{3\sigma}{\sqrt{n}}$ holds true. In these definitions, μ represents the mean, σ represents the standard deviation of dY_t, and n signifies the number of observations within the segregated dataset.

Classification of states for Monday, Tuesday, Wednesday, Thursday, and Friday are placed in the Figures 3, 4, 5, 6, and 7 respectively.

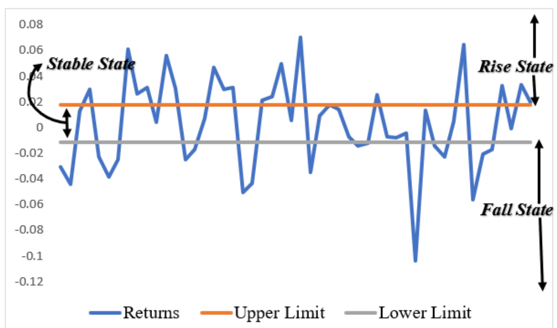

Figure 3: Classification of states in Monday data

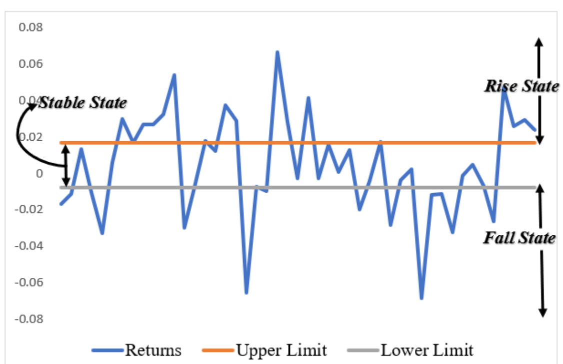

Figure 4: Classification of states in Tuesday data

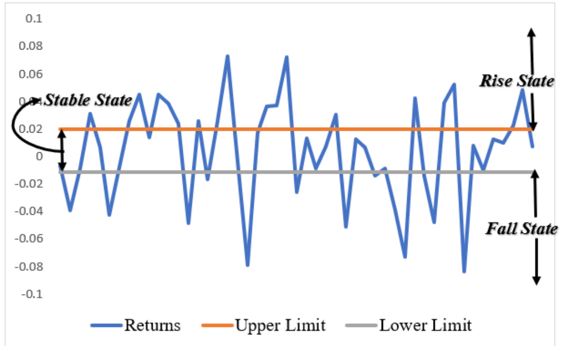

Figure 5: Classification of states in Wednessday data

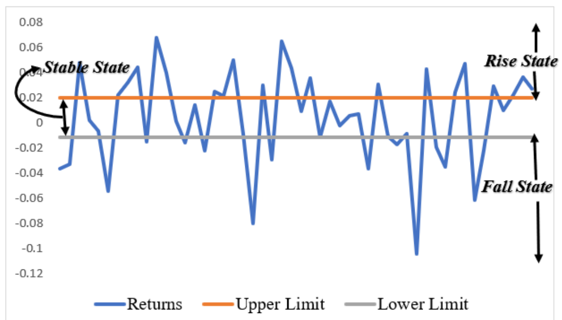

Figure 6: Classification of states in Thursday data

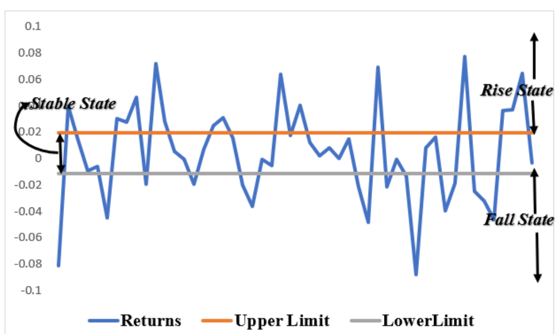

Figure 7: Classification of states in Friday data

Markov model is a composition of TPM, and IPV, which are computed with real-time data through R programming. Separate probability distributions, and statistical characteristics like average, variance, third central moments, skewness, kurtosis *etc.* are obtained for all segregated data sets. However, we have considered the averages for computing the predicted closing prices. The expected returns for SBI of all data sets are calculated by a formula as in section 2.7.1. We have obtained the predicted values (about 10 observations) of expected returns using the notion of sections 2.7.1 and 2.7.2. The developed Markov model is validated with the Chi-Square test for all data sets individually. AIC and BIC are also computed for each data set separately for the model's goodness of fit.

4. Results and discussion

The below parameters are placed in sections 4.1 and 4.2 which are computed by the above methodology.

4.1. Transition probability matrix (TPM) for SBI share closing prices

The explored TPM for Monday, Tuesday, Wednesday, Thursday, and Friday sets are as follows.

4.1.1. Transition behaviour of the market from monday to friday

The explored TPMs from Monday to Friday data placed in below Table 3.

Table 3: Transition Probabilities for all Business Days in a Week

Day	Transition Probabilities								
	RR	RS	RF	SR	SS	SF	FR	FS	FF
Monday	0.4706	0.2353	0.2941	0.4286	0.3571	0.2143	0.2222	0.2778	0.5
Tuesday	0.625	0.1875	0.1875	0.3125	0.375	0.3125	0.1429	0.5	0.3571
Wednesday	0.4445	0.2222	0.3333	0.3529	0.3529	0.2941	0.3077	0.5385	0.1538
Thursday	0.45	0.3	0.25	0.1538	0.3077	0.5385	0.625	0.1875	0.1875
Friday	0.4	0.4	0.2	0.1667	0.5	0.3333	0.375	0.25	0.375

The graphical representation of the above table is placed in Figure 8. It gives a clear interpretation of the transition behaviour of all business days in a week.

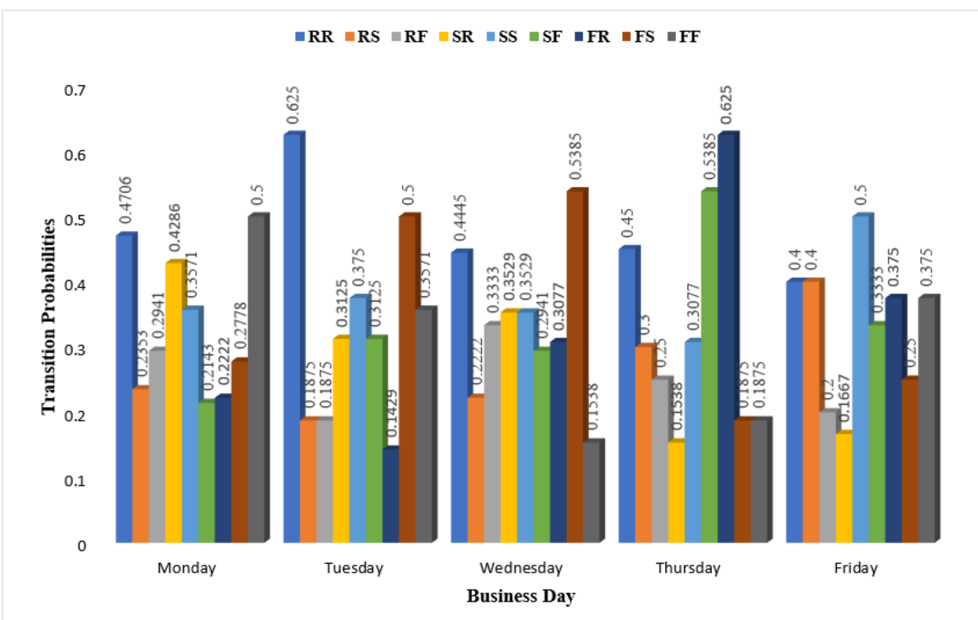

Figure 8: Transition probabilities for all business days in a week

From the above Table 3 and Figure 8 transition probabilities on Monday data set, it is observed that *Fall State* in the current day given that *Fall State* in the previous day is having the highest likelihood (50%); similarly *Rise State* in the current day given that *Rise State* in the previous day is having second highest likelihood (47.06%); *Rise State* in the current day given that *Stable State* in the previous day is having third highest likelihood (42.86%); and *Fall State* in the current day given that *Stable State* in the previous day having least likelihood (21.43%).

Based on the transition probabilities gleaned from Table 3 and Figure 8 of the Tuesday data set, it is observed that *Rise State* in the current day given that *Rise State* in the previous day is having the highest likelihood (62.5%); similarly *Stable State* in the current day given that *Fall State* in the previous day is having second highest likelihood (50%); *Stable State* in the current day given that *Stable State* in the previous day is having third highest likelihood (37.5%); and *Rise State* in the current day given that *Fall State* in the previous day having least likelihood (14.29%).

In analysing the transition probabilities extracted from Table 3 and Figure 8 of the Wednesday data set, it is observed that *Stable State* on the current day given that *Fall State* in the previous day is having the highest likelihood (53.85%); similarly *Rise State* in the current day given that *Rise State* in the previous day is having second highest likelihood (44.45%); *Stable State* in the current day given that *Stable State* in the previous day and *Rise State* in the current day and *Stable State* in the previous day both are having third highest likelihood (35.29%); and *Fall State* in the current day given that *Fall State* in the previous day having least likelihood (15.38%).

Examining the transition probabilities sourced from Table 3 and Figure 8 of the Thursday data set, it is observed that *Rise State* in the current day given that *Fall State* in the previous day is having the highest likelihood (62.5%); similarly, *Fall State* in the current day given that *Stable State* in the previous day is having the second highest likelihood (53.85%); *Rise State* in the current day given that *Rise State* in the previous day is having third highest likelihood (45%); and *Rise State* in the current day given that *Stable State* in the previous day having least likelihood (15.38%).

From the above Table 3 and Figure 8 transition probabilities of Friday data set, it is observed that *Stable State* in the current day given that *Stable State* in the previous day is having the highest likelihood (50%); similarly *Rise State* in the current day given that *Rise State* previous day and *Stable State* in the current day and *Rise State* in the previous day are having second highest likelihood (40%); *Fall State* in the current day given that *Fall State* in the previous day and *Rise State* in the current day and *Fall State* in the previous day are having third highest likelihood (37.5%); and *Fall State* in the current day given that *Rise State* in the previous day having least likelihood (20%).

These findings highlight distinct patterns in SBI's share prices throughout the week, indicating varying transient behaviours. This information can be invaluable for portfolio managers, enabling them to assess how SBI's share prices transition between *Rise*, *Stable*, and *Fall* states each day. These indicators provide crucial insights, allowing managers to strategize effectively, capitalize on profit opportunities, and implement corrective measures to mitigate losses.

4.2. Initial probability vector (IPV) for SBI share prices

After a thorough process of the real-time data, the IPVs of the *Rise*, *Stable*, and *Fall* states are obtained.

4.2.1. Indicators of *Rise*, *Stable*, and *Fall* States on Monday to Friday

The indicators on the chances of *Rise*, *Stable*, and *Fall* states the data under study are as below.

Table 4: Initial Probabilities for all Business Days in a Week

Day	Initial Probabilities		
	Rise	*Stable*	*Fall*
Monday	0.36	0.28	0.36
Tuesday	0.3617	0.3404	0.2979
Wednesday	0.3673	0.3673	0.2654
Thursday	0.42	0.26	0.32
Friday	0.30	0.38	0.32

The graphical representation of the above Table 4 is placed in Figure 9.

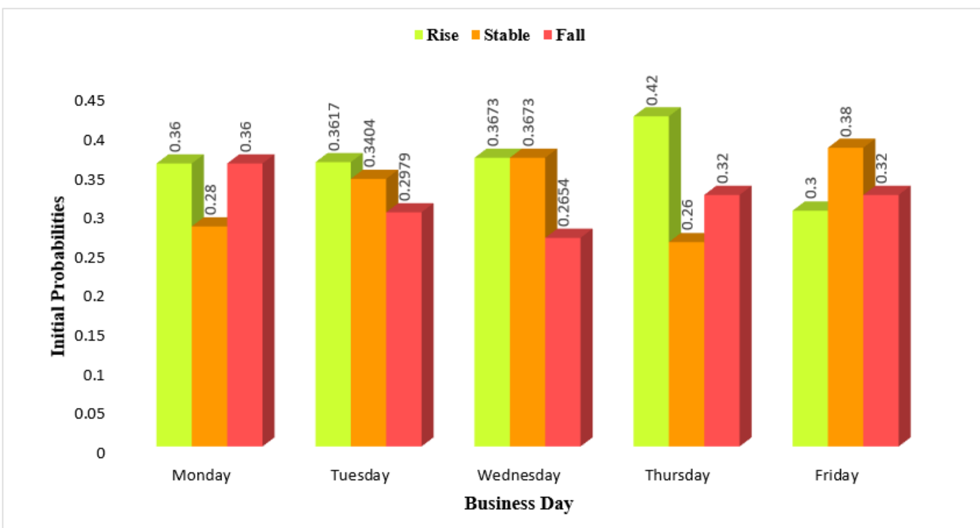

Figure 9: Initial probabilities for all business days in a week

From the above Table 4 and Figure 9, it is observed that on Monday, the likelihood of both *Rise State* and *Fall State* is equal at 36%. From Tuesday to Thursday, *Rise State* consistently has a higher likelihood than the other states, with Thursday having the highest probability at 42%, followed by Wednesday at 36.73%. Conversely, on Friday, the likelihood of the *Rise State* drops to its lowest at 30% compared to the other states. This suggests a strategy for short-term traders: consider selling shares during the middle of the week when the probability of a price increase is notably higher.

4.3. Probability distributions for *Rise*, *Stable*, and *Fall* states

The probability distributions for *Rise*, *Stable*, and *Fall* states of all business days in a week (Monday, Tuesday, Wednesday, Thursday, and Friday) are as in Table 5.

Table 5: Probability distributions of all states

Day	Chance of happening of the state		
	Rise	*Stable*	*Fall*
Monday	0.3694	0.2847	0.3459
Tuesday	0.375	0.3444	0.2806
Wednesday	0.3745	0.3541	0.2713
Thursday	0.429	0.266	0.305
Friday	0.3033	0.39	0.3067

Figure 10 illustrates the occurrence of *Rise*, *Stable*, and *Fall* states graphically. It provides a visual representation of the frequency of each state over the observed period.

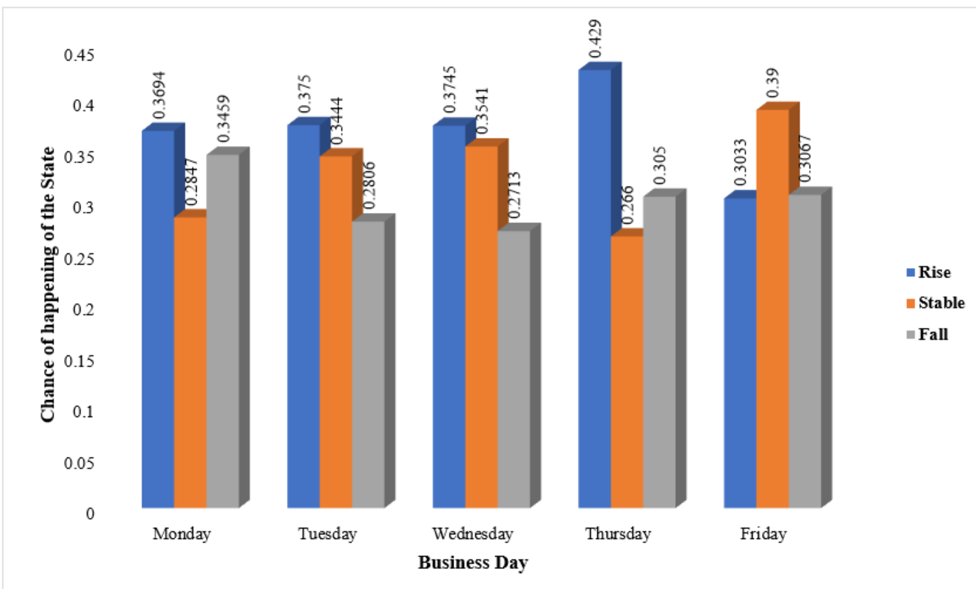

Figure 10: Chance of happening of *Rise*, *Stable*, and *Fall* states

According to the data presented in Table 5 and Figure 10, there is a noticeable trend in the occurrence of *Rise* and *Fall* states across different days of the week. Specifically, the likelihood of the *Rise State* is highest on Thursdays, closely followed by Tuesdays. Similarly, Fridays exhibit the highest probability of the *Rise State*, with Wednesdays following closely behind. In contrast, the *Fall State* is more likely to occur on Mondays, with Fridays showing the next highest probability.

This information suggests certain patterns or tendencies in market behaviour throughout the week. Traders may find it useful to be aware of these tendencies when making decisions about trading strategies, timing of trades, and risk management. For instance, understanding that Thursdays often have a higher chance of experiencing the *Rise State* could influence traders to adjust their positions accordingly or anticipate potential market movements. Similarly, knowledge of increased *Fall State* occurrences on Mondays might prompt traders to exercise caution or implement specific risk mitigation measures at the beginning of the trading week. Overall, awareness of these patterns can help traders make more informed decisions and navigate market dynamics more effectively.

Statistical measures/characteristics are useful in understanding the behaviour of the probability distributions.

4.4. Discussion on statistical measures

In order to have a better understanding of the model behaviour and the probability distributions, the statistical measures are computed and placed in Table 6.

Table 6: Statistical measures for *Rise*, *Stable* and *Fall* states

State	Statistical Measure	Monday	Tuesday	Wednesday	Thursday	Friday
Rise State	Average	0.3694	0.375	0.3745	0.429	0.3033
	Variance	0.2329	0.2344	0.2343	0.245	0.2113
	3rd central moment	0.0608	0.0586	0.0588	0.0348	0.0831
	Beta -1	0.2928	0.2667	0.2687	0.0823	0.7321
	Beta -2	1.2928	1.2667	1.2687	1.0823	1.7321
	C.V.	130.652	129.099	129.223	115.369	150.953
Stable State	Average	0.2847	0.3444	0.3541	0.266	0.39
	Variance	0.2036	0.2258	0.2287	0.1952	0.2379
	3rd central moment	0.0877	0.0703	0.0667	0.0914	0.0523
	Beta -1	0.9104	0.4288	0.3721	1.1218	0.2034
	Beta -2	1.9104	1.4288	1.3721	2.1218	1.2035
	C.V.	158.505	137.966	135.045	166.114	125.064
Fall State	Average	0.3459	0.2806	0.2713	0.305	0.3067
	Variance	0.2262	0.2019	0.1977	0.212	0.2126
	3rd central moment	0.0697	0.0886	0.0904	0.0827	0.0822
	Beta -1	0.4199	0.954	1.0582	0.7175	0.7032
	Beta -2	1.4199	1.954	2.0582	1.7175	1.7032
	C.V.	137.519	160.124	163.885	150.953	150.361

4.4.1. Discussion on the results

The results presented in Table 6 indicates that the *Rise State* is more frequently observed from Monday to Thursday compared to the *Stable* and *Fall* states. Specifically, there is a higher probability of the *Rise State* occurring during these days. Furthermore, Thursday stands out as the day with the highest likelihood for the *Rise State* in comparison to other business days.

Conversely, the *Stable State* exhibits a higher probability of occurrence on Fridays, suggesting a distinct pattern at the end of the week. This observation implies that different states (*Rise*, *Stable*, and *Fall*) exhibit varying likelihoods on different days, providing valuable insights into the underlying patterns of the data. The below Figure 11 shows the graphical representation of this content.

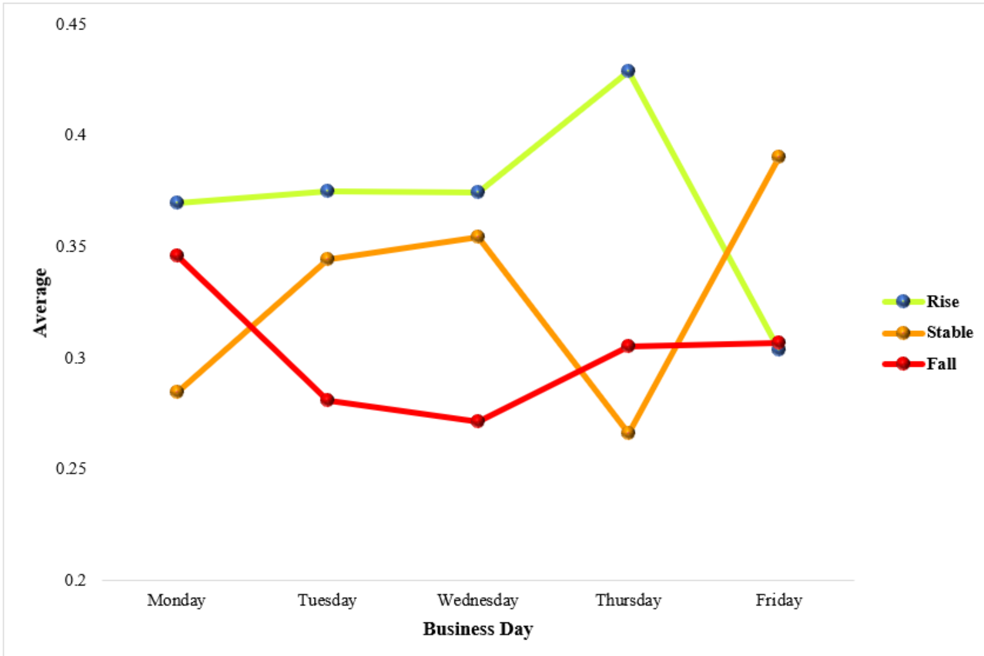

Figure 11: Averages of *Rise*, *Stable*, and *Fall* states in different days in a week

The analysis of the provided Table 6 reveals interesting patterns in the variability (variance) of different states (*Rise*, *Stable*, and *Fall*) across the weekdays. In the *Rise State*, similar variances are observed from Monday to Thursday, indicating consistent behaviour during these days. However, on Friday, there is a notable decrease in variance, suggesting a more stable trend compared to the preceding days.

In the *Stable State*, the highest variance is observed on Friday, signifying fluctuations, and unpredictability in the stock market towards the end of the week. Conversely, Thursday stands out with the least variance in this state, indicating a more stable and predictable market behaviour on that day.

For the *Fall State*, high variance is noted on Monday, suggesting significant fluctuations at the beginning of the week. In contrast, Wednesday exhibits the least variance in this state, indicating a relatively calmer and more predictable market environment.

Interestingly, the data emphasizes that Thursday is characterized by the least variance across all states (*Rise*, *Stable*, and *Fall*). This suggests that Thursdays tend to have a more stable market behaviour, making them potentially favourable for certain investment strategies.

These observations provide valuable insights for investors, indicating specific days of the week when the stock market is either more stable or prone to fluctuations. Investors could potentially use this information to inform their trading decisions, adapting their strategies based on the observed patterns of variance in different market states across weekdays.

The positive skewness indicated by the non-negative third central moment across all states (*Rise*, *Stable*, and *Fall*) implies that in the stock market, there are more frequent occurrences of small or moderate gains. These modest gains are a common feature, suggesting relative stability in stock prices. However, the presence of occasional significant upward shifts in stock prices, although infrequent, contributes to the overall positive skewness.

For investors, this pattern highlights the regularity of stable or moderately positive market movements, punctuated by occasional notable upticks. Recognizing these infrequent but substantial positive shifts is vital for investors seeking opportunities for significant profits. However, it also underscores the need for prudent risk management, as these occasional large movements can result in substantial losses if not carefully navigated. Understanding this skewed distribution is essential for making informed investment decisions in the stock market.

The kurtosis values being less than three for all states (*Rise*, *Stable*, and *Fall*) on every business day indicate a platykurtic distribution in the stock market.

The observation of the lowest coefficient of variation in the *Rise State* on Thursday (115.369) implies that this particular day showcases a remarkable consistency and stability in stock market performance, graphically it is presented in Figure 12.

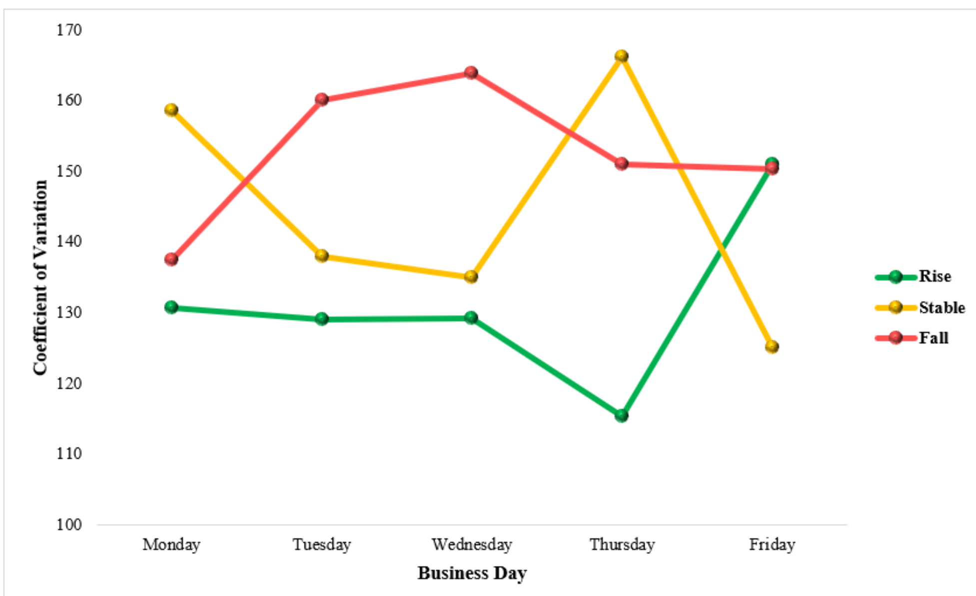

Figure 12: Coefficient of variation for *Rise*, *Stable*, and *Fall* states in different days in a week

Hence, Figure 12 may advise to short-term traders that Thursday might be an opportune day to consider selling stocks to maximise returns. The lower coefficient of variation indicates reduced volatility and fluctuations, indicating a more predictable market environment. This stability can provide short-term traders with confidence in making strategic decisions, potentially leading to optimal returns on their investments. Understanding these patterns in the coefficient of variation aids traders in identifying favourable moments for executing trades and capitalizing on market stability.

4.5. Expected (predicted) returns for SBI's shares

The expected returns computed for all days in a week are separately computed using the formula mentioned in Section 2.7.1.

4.5.1. Expected returns for SBIs all business days data of all states

The given below are the expected SBI share price returns in 10 business days due to *Rise*, *Stable*, and *Fall* states.

Table 7: Expected returns for *Rise*, *Stable*, and *Fall* states

State	Day	Monday	Tuesday	Wednesday	Thursday	Friday
Rise	t=1	0.009207387	0.015662265	0.003799441	0.008274521	0.013172060
	t=2	0.004862423	0.009527704	0.004649036	0.003889552	0.005524245
Stable	t=1	0.010790754	0.002533674	0.002325153	0.013513428	0.001853549
	t=2	0.006335824	0.004617461	0.004317823	0.006204354	0.002929969
Fall	t=1	-0.006832183	-0.003926799	0.007331084	0.016880182	0.004984202
	t=2	0.001627426	0.002101875	0.003548923	0.005802842	0.006345211

Analysing the resulted Table 7, it is observed there is expected returns of *Fall State* on Monday and Tuesday are negative, it may indicate to the traders there is a risk factor involved in share market, so these results may advise to the short-term traders to adopt risk tolerance and portfolio strategy to overcome the loss on investment.

4.5.2. Estimated (Predicted) closing prices of SBI shares

The SBI's closing prices are predicted using the linearity formula which is placed in Section 2.7.2 The predicted share prices are

Table 8: Predicted closing prices

Week	Monday	Tuesday	Wednesday	Thursday	Friday
First	556.8796	578.5471	572.9164	582.9601	579.2679
Second	559.2062	570.899	575.3413	585.924	582.0315

Figure 13 depicts the observed and predicted closing prices of SBI shares. These forecasts are valuable for short-term traders, enabling them to discern patterns in SBI share prices and make informed decisions for trading in the upcoming week.

4.6. Validation of the model

4.6.1. Chi-square test

The Markov model developed was assessed using the Chi-Square test for goodness of fit, considering both expected and observed values over two weeks (business days only). The test's null hypothesis (H_0) posits that the developed model fits the data well, meaning there is no significant difference between the observed and expected closing prices of SBI shares.

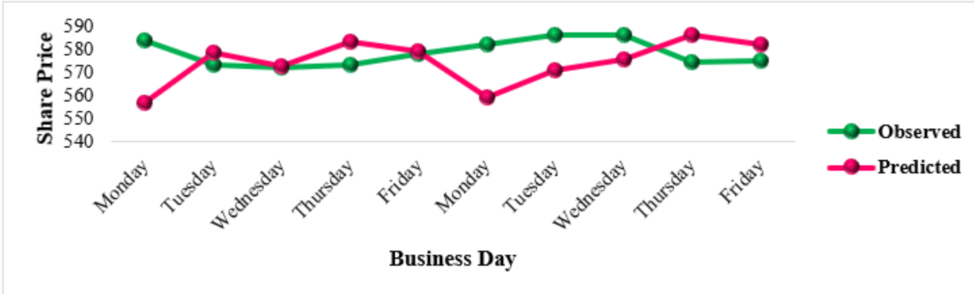

Figure 13: Predicted Closing Prices of First and Second Week

The calculated probability value (p-value) for SBI is 0.9719 with 9 degrees of freedom. The result indicates that the stated hypothesis is not rejected, confirming that the developed model aligns with the data.

4.6.2. AIC and BIC

Additionally, the model's robustness was evaluated using the Akaike information criterion (AIC) and Bayesian information criterion (BIC).

The AIC and BIC were calculated using the formulae mentioned in the Section 2.8.2. For Monday, Tuesday, Wednesday, Thursday, and Friday, the AIC values are 70.80734, 57.5237, 63.0746, 69.68, and 64.1920, while the corresponding BIC values are 91.6174, 75.8102, 81.7867, 90.49, and 83.1102. These findings indicate that the AIC and BIC values are lowest for Tuesday data, followed by Wednesday's data. Consequently, the results obtained from the developed model affirm that the upward trend in share value is notably more consistent during the middle of the week.

Acknowledgements

We are thankful to the DST INSPIRE. Ms. Sarode Rekha is supported by INSPIRE fellowship (IF190741) of the Department of Science and Technology (DST), Government of India.

Conflict of interest

The authors do not have any financial or non-financial conflict of interest to declare for the research work included in this article.

References

Abounoori, E., Elmi, Z. M., and Nademi, Y. (2016). Forecasting Tehran stock exchange volatility; markov switching garch approach. *Physica A: Statistical Mechanics and its Applications*, **445**, 264–282.

Adesokan, I., Ngare, P., and Kilishi, A. (2017). Analyzing expected returns of a stock using the markov chain model and the capital asset pricing model. *Applied Mathematical Sciences*, **11**, 2777–2788.

Bhusal, M. K. (2017). Application of markov chain model in the stock market trend analysis of Nepal. *International Journal of Scientific & Engineering Research*, **8**, 1733–1745.

Cai, J. (1994). A markov model of switching-regime arch. *Journal of Business & Economic Statistics*, **12**, 309–316.

Choji, D. N., Eduno, S. N., and Kassem, G. T. (2013). Markov chain model application on share price movement in stock market. *Computer Engineering and Intelligent Systems*, **4**, 84–95.

Dar, G. F., Ahn, Y.-H., Dar, Q. F., and Ma, J.-H. (2023). Impact of international trade on the share prices of the industrial bank of Korea using stochastic prediction modeling. *Journal of Global Business and Trade*, **19**, 71–90.

Dar, G. F., Padi, T. R., Rekha, S., and Dar, Q. F. (2022). Stochastic modeling for the analysis and forecasting of stock market trend using hidden markov model. *Asian Journal of Probability and Statistics*, **18**, 43–56.

Dias, J. G., Vermunt, J. K., and Ramos, S. (2010). Mixture hidden markov models in finance research. In *Advances in Data Analysis, Data Handling and Business Intelligence: Proceedings of the 32nd Annual Conference of the Gesellschaft für Klassifikation eV, Joint Conference with the British Classification Society (BCS) and the Dutch/Flemish Classification Society (VOC), Helmut-Schmidt-University, Hamburg, July 16-18, 2008*, pages 451–459. Springer.

D'Amico, G. and Petroni, F. (2012). A semi-markov model for price returns. *Physica A: Statistical Mechanics and its applications*, **391**, 4867–4876.

Hassan, M. R. (2009). A combination of hidden markov model and fuzzy model for stock market forecasting. *Neurocomputing*, **72**, 3439–3446.

Huang, J.-C., Huang, W.-T., Chu, P.-T., Lee, W.-Y., Pai, H.-P., Chuang, C.-C., and Wu, Y.-W. (2017). Applying a markov chain for the stock pricing of a novel forecasting model. *Communications in Statistics-theory and Methods*, **46**, 4388–4402.

Jobert, A. and Rogers, L. C. (2006). Option pricing with markov-modulated dynamics. *SIAM Journal on Control and Optimization*, **44**, 2063–2078.

Padi, T. R., Dar, G. F., and Rekha, S. (2022). Stock market trend analysis and prediction using markov chain approach in the context of Indian stock market. *IOSR Journal of Mathematics*, **18**, 40–45.

Rao, S. and Hong, J. (2010). Analysis of hidden markov models and support vector machines in financial applications. *University of California at Berkeley: Berkeley, CA, USA*.