

**STOCK PREDICTION OF BANGLADESH SHARE MARKET**

**BY**

**MD. ALMAS HOSSAIN**

**ID: 152-15-5635**

**AND**

**MD. ARIFUL ISLAM SAIKAT**

**ID: 152-15-5616**

This Report Presented in Partial Fulfillment of the Requirements for the Degree  
of Bachelor of Science in Computer Science and Engineering

Supervised By

**SHEIKH ABUJAR**

Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

**ABDUS SATTAR**

Asst. Professor

Department of CSE

Daffodil International University



**DAFFODIL INTERNATIONAL UNIVERSITY**

**DHAKA, BANGLADESH**

**MAY 2019**

## APPROVAL

This Project/internship titled “**Stock Prediction of Bangladesh Share Market**”, submitted by Md. Almas Hossain, ID No: 152-15-5635 and Md. Ariful Islam Saikat, ID No: 152-15-5616 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 3<sup>rd</sup> May, 2019.

### BOARD OF EXAMINERS



**Dr. Syed Akhter Hossain**  
**Professor and Head**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Chairman**



**Dr. Md. Ismail Jabiullah**  
**Professor**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Internal Examiner**



**Dr. Sheak Rashed Haider Noori**  
**Associate Professor & Associate Head**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Internal Examiner**



**Dr. Dewan Md. Farid**  
**Associate Professor**

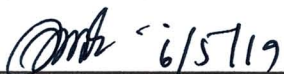
Department of Computer Science and Engineering  
United International University

**External Examiner**

## DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Sheikh Abujar, Lecturer, Department of CSE, Daffodil International University**. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

**Supervised by:**

  
\_\_\_\_\_

**SHEIKH ABUJAR**

Lecturer

Department of CSE

Daffodil International University

**Co-Supervised by:**

  
\_\_\_\_\_

**ABDUS SATTAR**

Asst. Professor

Department of CSE

Daffodil International University

**Submitted by:**

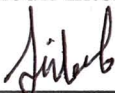
  
\_\_\_\_\_

**MD. ALMAS HOSSAIN**

ID: 152-15-5635

Department of CSE

Daffodil International University

  
\_\_\_\_\_

**MD. ARIFUL ISLAM SAIKAT**

ID: 152-15-5616

Department of CSE

Daffodil International University

## ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **SHEIKH ABUJAR, LECTURER**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “*Data Mining*” to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Syed Akhter Hossain**, Professor and Head, Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

## **ABSTRACT**

The stock market plays several roles in our economy. It works as a mid-part between surplus and deficit units of the economy. But considering Bangladesh stock market, it is not yet to play its potential role as a medium for financing long term investment. But however, a fierce downtrend has been trying to overcome successfully after the hangover of an unsavory historical past which saw a dramatic immersion in the last half of 1996 and beginning of 2011. A number of actions from the regulatory bodies were the main reason of losing public confidence from the stock market.

To improve this situation and to bring back the public confidence to stock market, we are going to implement a project “Stock Prediction of Bangladesh Share Market” which can predict the stock market. By the help of “Hidden Markov Model (HMM)” using Data Mining, we are going to forecast or predict future share price. Main Outcome of this project is to make people interest in Bangladesh Share Market and help them to invest their valuable money in the right place.

# TABLE OF CONTENTS

<b>CONTENS</b>	<b>PAGE</b>
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
<b>CHAPTER</b>	
<b>CHAPTER 1: INTRODUCTION</b>	<b>1-2</b>
1.1 Introduction	1
1.2 Motivation	1
1.3 Research Questions	1-2
1.4 Expected Outcome	2
1.5 Report Layout	2
<b>CHAPTER 2: BACKGROUND</b>	<b>3-5</b>
2.1 Introduction	3
2.2 Related Works	3-4
2.3 Research Summary	4-5
2.4 Scope of the Problem	5
2.5 Challenges	5

<b>CHAPTER 3: RESEARCH METHODOLOGY</b>	<b>6-23</b>
3.1 Introduction	6
3.2 Research Subject and Instrumentation	6
3.2.1 Proposed System Architecture	6-7
3.3 Data Collection Procedure	7-9
3.4 Program flow Diagram for the Proposal System	9-10
3.5 Implementation Requirements	10
3.5.1 Approach	10
3.5.2 Pseudo code of the calculation	10-16
3.5.3 Markov Model	16-17
3.5.4 Method	17-19
3.5.5 Implementation	19-22
3.5.6 Platform used for the system	22-23
<b>CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION</b>	<b>24-26</b>
4.1 Introduction	24
4.2 Experimental Result	24
4.3 System Testing	24-25
4.4 Discussion	26

**CHAPTER 5: SUMMARY, CONCLUSION, LIMITATION AND IMPLIATION FOR FUTURE RESEARCH** 27

5.1 Summary of the Study 27

5.2 Conclusion 27

5.3 Limitations 27

5.4 Implication for Future Work 27



## **LIST OF TABLES**

<b>TABLE</b>	<b>PAGE NO</b>
Table 4.1: Closest Predicted Results	<b>24</b>
Table 4.2: System Testing	<b>25</b>
Table 4.3: Result Comparing	<b>26</b>

## **LIST OF FIGURES**

<b>FIGURES</b>	<b>PAGE NO</b>
Figure-3.1: Proposed System Architecture	6
Figure-3.2: Dhaka Stock Exchange (DSE) Official Website	8
Figure-3.3: Data set of ACI Ltd.	8
Figure-3.4: Flow Chart Diagram	9
Figure-3.5: Transaction Matrix	18
Figure-3.6: Transaction Nodes	18
Figure-3.7: System Login Page	20
Figure-3.8: System Admin Panel	20
Figure-3.9: User View of Dashboard	21
Figure-3.10: Opening Price of a company	21
Figure-3.11: Closing Price of a company	22
Figure-3.12: Sublime	23
Figure-3.13: Xampp	24
Figure -A1: Plagiarism Report	30
<b>REFERENCES</b>	<b>28-29</b>
<b>APPENDIX</b>	<b>30</b>

# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

Stock or share market is a system where those entrepreneurs who do need fund business are get those money from small institutions at low costs [1]. In 2010, Bangladesh stock market was considered a great year for investors. Although this record fall in December in the same year. It created so much panic among those people who had put their money in stocks market. There was rarely any investor who losses his money in 2010. Bangladesh stock market fall again in the start of 2011. The bullish trend of market suddenly turned to rugged that is more than expected which ultimately made the market coming down and reduced confidence level which is inimically affected to the trading nature and attitude of retail investors.

Though looking back at the previous year, it feels Bangladesh's stock market was quite active and the financial sector's performance is really good and managed to exceed expectations. To stable this situation more flawlessly and to insure more involvement of the people in stock market, we are developing a system which is based on Data Mining. It could forecast future price of a stock. People will be benefited from it and will show interest in investing their money in stock market.

### 1.2 Motivation

As a developing country, Bangladesh's share market takes a big place of its economy. So to move the economy much further and to set in a stable position, our share market has to be stable and uprising. For do this, more involvement of people are needed. We are developing a system which can help the people of Bangladesh to engage more with it. We are using Data Mining for the best result of it. So that people will be benefited from it and help to stable the economy of Bangladesh.

### 1.3 Research Questions

If we look around there are already has many systems based on stock prediction. If we consider Bangladesh share market, those systems are not that much reliable. From those systems, some of are not accurate and some of are much costly to get the data of prediction. This data is not

public, so everyone offers the data in exchange for a good amount of money. So how we will offer much flexible system that can be easily accessible with a minimal level of privacy and security. There are many ways to do forecasting or predict using different methods. How we choose the best method in the circumstances of Bangladesh and apply it perfectly is the biggest question. How the dataset work on that method is also an important question. In which way results should be shown as people can easily understood that.

## **1.4 Expected Output**

Our goal is to predict the price as accurate as possible. We will predict the next day stock price based on previous stock transaction data like, opening and closing price. There is no possible way that our result will be 100% accurate but we can push the accuracy as higher as the result is reliable. For more accurate forecasting we will show four result instead of one result. Also show the percentage of accuracy of each prediction. We will maintain proper privacy and security of the system. Also this system will be flexible for accessing.

## **1.5 Report Layout**

This report consists of five chapters, and this section provides insight of all five chapters.

1. Chapter one provides introduction, motivation and expected outcome of the study.
2. Related research work is discussed on chapter two. It also provides challenges of the research.
3. In chapter three, Methodology and System design for the Proposed System is provided.
4. Chapter four of this document describe our proposed system implementation and testing.
5. Lastly, chapter five is on conclusion, limitations, future study.

## **CHAPTER 2**

### **BACKGROUND**

#### **2.1 Introduction**

Data mining is to identify and establish patterns and relationships in the process of sorting through large number of data sets to solve problems via data analysis. Data mining tools allow to predict future trends [10].

Hidden Markov Model (HMM) is a method of predict future data based on Data Mining. Hidden Markov Model (HMM) is a measurable state machine. It contains some fixed number of states. For modelling a time series of multivariate observations, it gives a probabilistic frame. In the beginning of the 1970's, first introduction of Hidden Markov Models (HMM) was as a special tool or item in speech recognition. In last several years for its strong mathematical structure, HMM has become popular and also for use in a wide range of applications, its theoretical basis works so well which is based on statistical methods. Researchers proposed HMM as a predictor or classifier for DNA sequence analysis [12], Speech signal recognition [13, 14, 15], Natural language domains, Handwritten characters recognition [11] etc. in recent years. It is clearly proved that HMM is a very powerful and useful tool for the various applications. Some of the advantages of HMM can be shortened as:

- Hidden Markov Model (HMM) has very strong and powerful statistical foundation.
- It can strongly handle and manipulate new data.
- It is efficient for computation to evaluate and develop. It mostly due to the existence of established algorithms which is well trained.
- It has the efficient ability to predict similar patterns [17].

#### **2.2 Related Works**

Forecasting any data or information is not a big thing now. It's become so simple now-a-days with the help of data mining. There are many methods and models are developed with the outgrowth of data science. Many methods and models are established to predict or forecast future data. Some are Auto regression moving average (ARMA) and the multiple regression models, Artificial Neural Networks (ANN), Fuzzy logic, Hidden Markov Model (HMM), hybridization of Fuzzy system and ANN and also Support vector machines.

The paper [16] describe, they took a sample share of Oil India Ltd. They considered daily close value data for two months period. Took three observing symbols I, N and D. The symbol I- indicates Increasing, N-indicates No change and D-indicates Decreasing. If nth day's close value – (n-1) th day's close value  $> 0$ , then observing symbol is I. If nth day's close value – (n-1)th day's close value  $< 0$ , then observing symbol is D. If nth day's close value – (n-1)th day's close value = 0, then observing symbol is N. Also assumed seven hidden states. From that hidden state sequences, they produce the observations. The various probability values of transaction probability matrix (tpm) and emission probability matrix (epm) for difference in one day, two days and three days close values are calculated. From this states and MATLAB software generated sequence, it can predict future vales of stock market.

In the paper [18], it developed an HMM based tool which is for time series forecasting. For an example it is used for the stock market forecasting. While implementing the HMM, choice of the number of states and the model and also observation symbol become a tiring task. For example along with 4 states, it has used left-right HMM. It considered four input items for a stock which were respectively the opening price, the closing price, the highest price and the lowest price. The target price is considered from the next day's closing price associated with the four input features. It used daily data of approximately past one and a half years to train the model. For efficiency testing of the model, it used last three month's data of recent date.

Nguyet [19], predict stock prices using the HMM and compare the predictions with the real stock prices, and with the results of using the historical average return method. He used S&P 500 monthly prices from January 1950 to November 2016. He divided the process into three steps. First step is to adjust the HMM's parameters using training data and from the observation of the data set the probability is calculated. Then, he will find a "similar data" set in the past that has a similar behavior to that of the training data. Last step is to use the difference of stock prices in the last month of the founded sequence with the price of the running month to predict future stock prices.

## **2.3 Research Summary**

In our study, we are trying to make it as simple as possible. There are already lots of prediction sites available in the internet. But from those sites, some are just fake and irrelevant. They make those sites for people's harassment. People don't get any benefit from those sites. Also

some of the sites show real prediction but those site charge lots of money. So overall, it is difficult for people to get easy access to the predicted data.

We will build the system in web because people could easily access to the system from anywhere and from any devices like computer, mobile etc. We will use HTML, CSS, PHP, JavaScript, Ajax to develop our project. We will maintain the security and privacy of the system. We will separate the web view of the system for admin and user so that the privacy maintained properly.

## **2.4 Scope of the Problem**

After we have gone through the background study and brainstorming discussion, we have discovered that in Bangladesh's perspective, the prediction of stock price could not be fully relying on the transection of the stock data. There are so many factors are involved. To get the accurate result of the prediction, we have to consider those factors. Otherwise the prediction accuracy could not be that so high.

## **2.5 Challenges**

Our main challenge is to forecast the stock price accurately. The accuracy of the prediction might not be 100% but we have to maintain the accuracy as close as possible. For more accuracy we have to collect more data. If we could collect more data, the prediction will be more useable.

It will be a difficult task to collect company's share data. As it is not available directly we have to collect the data manually. As we will calculate the prediction based on one year data, it will not be that easy to collect it manually.

We can't predict all company's share because it will be a time consuming work to do. We will just select top 5 company's share of Bangladesh. We will collect their data and analysis those data. There are lots of company in share market.

Bangladesh's share market is different from other country's share market. The pricing of Bangladesh share market is mostly vary with different types of factors. Many impactors are causing the price HIGH and LOW. Those factors or impactors doesn't follow any rule or model. For this problems prediction could be more hard to calculate.

## CHAPTER 3

### RESEARCH METHODOLOGY

#### 3.1 Introduction

Methodology is how will be the work done and which way the work will be done. System design is consisting of some approaches taken to work out specific options, demands, expectation by communication with the system users.

To predict or anticipate data and information about future is an interesting thing and if it is done in Stock market area where past data could be used to anticipate and predict, it becomes more interesting to work. We sought out the process to predict the current future stock prices for five given companies using the stochastic method called Markov Chains. This application is based on HTML, CSS, PHP, JavaScript, Ajax and MySQL.

#### 3.2 Research Subject and Instrumentation

We are researching about stock prediction which is based on Data mining. For this we have to collect some dataset to implement data mining method. We are using markov chain model to predict the stock price.

##### 3.2.1 Proposed System Architecture

In our proposed system, at first we collect the data from DSE official website then insert this into the database. Based on the data of database, we do the calculation for prediction. Then we show the prediction result to our web site.

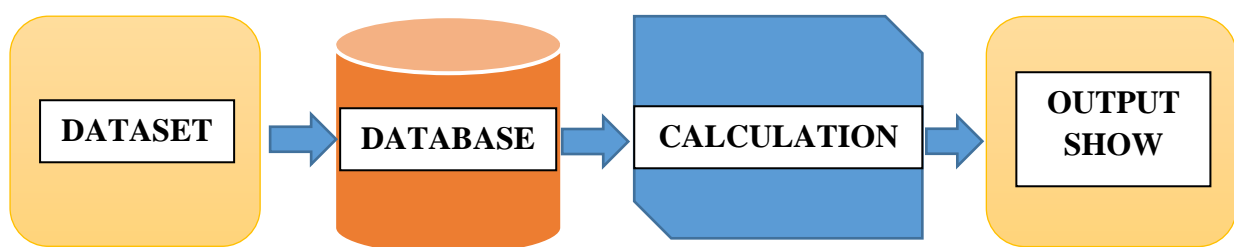


Figure -3.1: Proposed System Architecture

In Figure -3.1, Stock prediction of Bangladesh share market consist of some dataset which used for the calculation of the prediction came via database.



## **DATASET**

We collected this stock data from Dhaka Stock Exchange (DSE) official website. This data set contains data from 1<sup>st</sup> January, 2018 to 31<sup>st</sup> December, 2018.

## **DATABASE**

From the dataset, we picked just four items. We picked Opening price of the stock, Closing price of the stock, HIGH price and LOW price. Then we normalized the data as it required for calculation purpose. We inserted few new columns like, moving average (which is the average of recent 3 day opening/closing price), moving difference (which is the difference between opening/closing price and moving average). Then we inserted the data to the database of the system. For every company, we created new table so that we can easily do our calculation. Our database are aci, pran, beximco, brac bank and gp.

## **CALCULATION**

We did the main calculation in our PHP file. We used different formulas to simplify our calculation. We set function for every small calculation. From the normalized data we do the actual calculation from it. Our whole project and prediction are dependent on the calculation. So we did the calculation carefully.

## **OUTPUT SHOW**

After all calculation in our PHP file, we show the result on our webpage. We separated the four resulted value of the opening prediction and the closing prediction, so that people can easily differentiate the result. We also show the probability of the predicted price. People can easily interact with our webpage.

### **3.3 Data Collection Procedure**

We collected the stock data of five big companies of Bangladesh. Companies are,

- ACI Ltd. (ACI)
- Agricultural Marketing Company Ltd. (AMCL-(PRAN))
- Bangladesh Export and Import Company Ltd. (BEXIMCO)
- BRAC Bank Ltd. (BRACBANK)
- Grameenphone Ltd. (GP)

We collected this stock data from Dhaka Stock Exchange (DSE) official website. This data set contains data from 1<sup>st</sup> January, 2018 to 31<sup>st</sup> December, 2018.



Figure -3.2: Dhaka Stock Exchange (DSE) Official Website

In Figure -3.3, it shows the Dhaka stock market website. It contains various information regarding to stock exchange. It represents overall summery of the particular day's stock market along with companies information and their stock prices. Also shows how much company's price goes up and down, what is the recent status of the companies.

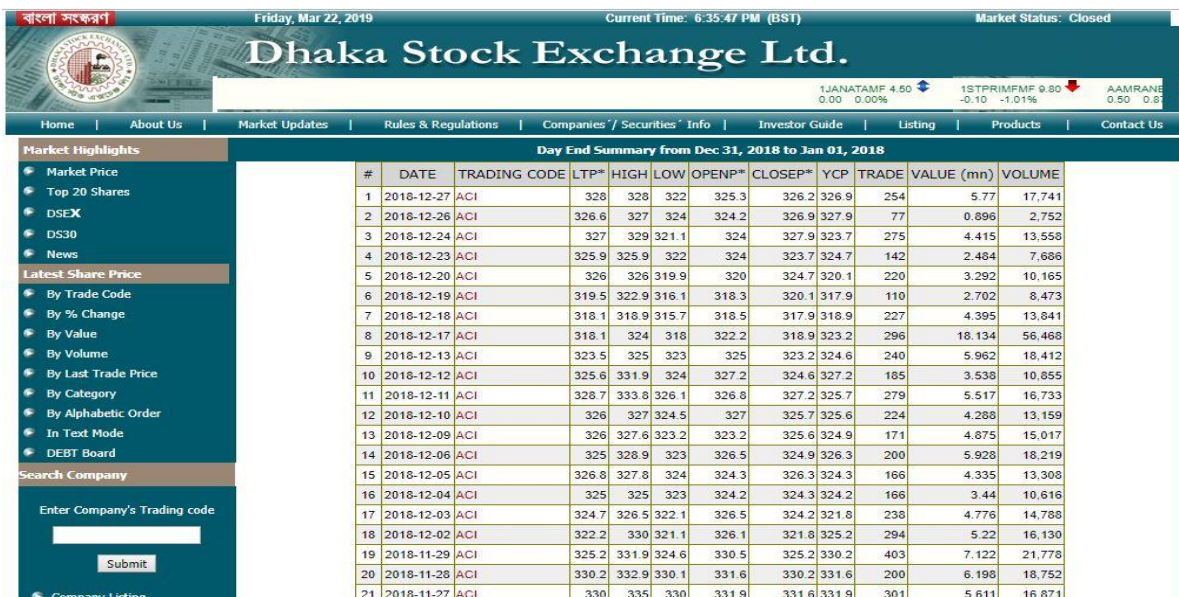


Figure -3.3: Data set of ACI Ltd.

In Figure -3.4, it shows the data of ACI Ltd. between a fixed interval of dates. Here it provides this information, Date, Trading Code of the company stock, LTP = Last Trading Price, High price, Low price, OPENP = Opening Price of that stock of the day, CLOSEP = Closing Price of that stock of the day, YCP = Yesterday's Closing Price, Trade, Value and the Volume of the stock.

### 3.4 Program flow Diagram for the Proposed System

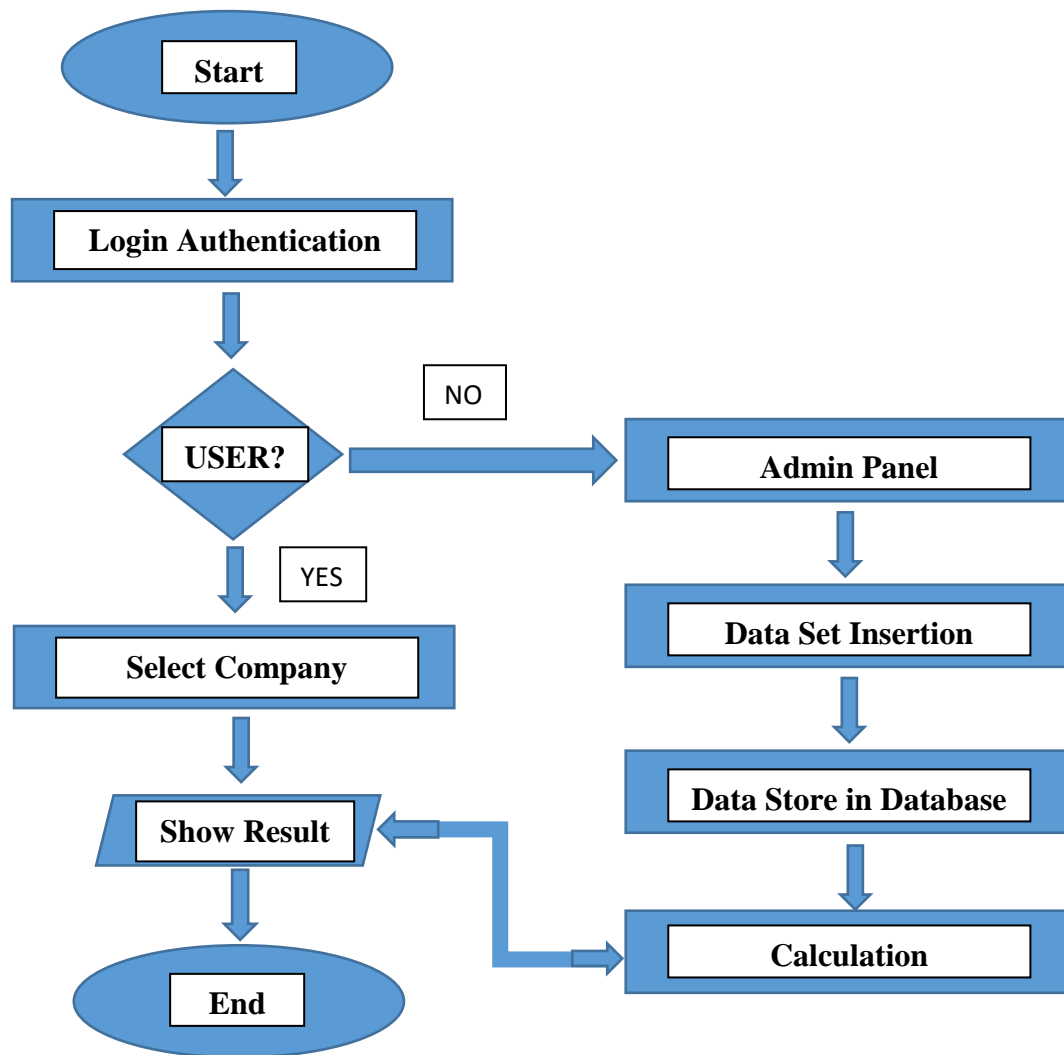


Figure -3.4: Flow Chart Diagram

In Figure -3.2, flow chart of the system is shown. People have to login to enter the website. We use user type in our database for user authentication, if the user type is 1, it will be admin and if its 2, it will be user. So if he or she is a user, it will take to the user panel and if not then it

will sent him or her to admin panel. In admin panel, admin can input the data into database. In user panel, user can select particular company to see that company's predicted price. After selecting the company, it fetch the calculated data which is calculated under that company's transaction values. He or she will see a new page which contains four result each for opening and closing. After finishing his or her work, he or she could log out from the system.

### 3.5 Implementation Requirements

We implement it in many steps. So that it could be easy for us.

#### 3.5.1 Approach

- We analyzed a year's worth of stock portfolio for five of companies in this project. We applied Markov model and moving averages to the data in prospect to forecast the stock prices for the near future.
- Firstly, to build an approximate appreciation of the data, we applied moving averages. Then we found a difference data set and on that data set we will apply Markov Chains model. We also took the difference of the closing price and the moving average price. These differences were used what we would apply Markov Chains to.
- We then created the transition matrix and calculated the firm state probabilities to make the predictions by applying Markov Chain properties to the data set.

#### 3.5.2 Pseudo code of the calculation

##### Basic Calculation for Normalization of Dataset:

Before do our main calculation, we have to normalize our dataset as per our requirement. It is very important because our farther calculation is depend on this data normalization. We do the average of opening and closing price of recent 3 days and also do subtraction from opening price to moving price. Here is the formula,

1. *MOVING\_AVERAGE [] = NULL*
2. *for i = 1; i <= data.length; i++*
3. *MOVING\_AVERAGE [i+2] = (Recent 3 days opening price)*
4. *OPENING\_PRICE [] = NULL*
5. *for i = 1; i <= data.length; i++*

6.  $MOVING\_DEFERRENCE [i+2] = OPENING\_PRICE [i+2] - MOVING\_AVERAGE [i+2]$
7.  $MOVING\_AVERAGE [] = NULL$
8. *for*  $i = 1; i \leq data.length; i++$
9.  $MOVING\_AVERAGE [i+2] = (Recent\ 3\ days\ closing\ price)$
10.  $CLOSEING\_PRICE [] = NULL$
11. *for*  $i = 1; i \leq data.length; i++$
12.  $MOVING\_DEFERRENCE [i+2] = CLOSEING\_PRICE [i+2] - MOVING\_AVERAGE [i+2]$

### Method for Proposed System Design:

For entering our main calculation, first we have to calculate the binning value. We read the total transition day which is consider as database rows of a company in here and do binning from it. Here is the formula,

1.  $COUNT = STOCK\_DATA\_OF\_A\_COMPANY$
2.  $SQRT = sqrt (COUNT)$
3.  $MAX = \max (STOCK\_DATA\_OF\_OPEING\_OR\_CLOSING\_DIFFERENCE)$
4.  $MIN = \min (STOCK\_DATA\_OF\_OPEING\_OR\_CLOSING\_DIFFERENCE)$
5.  $MEAN = (MAX-MIN) / SQRT$
6. *While*  $MIN \leq MAX$
7.  $BIN [] = MIN + MEAN$
8.  $MIN += MEAN$

After finishing the binning, we count the frequency of the moving difference price value based on the bin range. Here is the formula,

1.  $COUNT = STOCK\_DATA\_OF\_A\_COMPANY$
2.  $FREQUENCY [] = 0$
3.  $MOVING\_DIFFERENCE [] = STOCK\_DATA\_MOVING\_DIFFERENCE$

4. *for*  $i = 1; i < COUNT; i++$
5.     *for*  $j = 1; j < count (BIN); j++$
6.         *if*  $BIN [j] \leq MOVING\_DIFFERENCE [i] \ \&\& \ MOVING\_DIFFERENCE [i] < BIN [j+1]$
7.              $FREQUENCY [j] ++$

After calculating the frequency, we calculate the cumulative frequency of that frequency. Formula is,

1.  $CUMULATIVE\_FREQUENCY [] = 0$
2. *for*  $i = 1; i < count (FREQUENCY); i++$
3.     *for*  $j = 1; j \leq i; j++$
4.          $CUMULATIVE\_FREQUENCY [i] += FREQUENCY [j]$

We create five quartile range with the help of cumulative frequency. Depend on this quartile range we select the bin value. Calculation is based on bellow formula,

1.  $N = end (CUMULATIVE\_FREQUENCY)$
2.  $N1 = N/4$
3.  $N2 = N/2$
4.  $N3 = (N*3)/4$
5.  $RANGE [] = BIN [0]$
6. *for*  $i = 1; i < count (CUMULATIVE\_FREQUENCY); i++$
7.     *if*  $CUMULATIVE\_FREQUENCY [i] \leq N1 \ \&\& \ N1 < CUMULATIVE\_FREQUENCY [i+1]$
8.          $RANGE [i] = (N1 - CUMULATIVE\_FREQUENCY [i]) < (CUMULATIVE\_FREQUENCY [i+1] - N1) ? BIN [i] : BIN [i+1]$
9.     *if*  $CUMULATIVE\_FREQUENCY [i] \leq N2 \ \&\& \ N2 < CUMULATIVE\_FREQUENCY [i+1]$
10.          $RANGE [i] = (N2 - CUMULATIVE\_FREQUENCY [i]) < (CUMULATIVE\_FREQUENCY [i+1] - N2) ? BIN [i] : BIN [i+1]$
11.     *if*  $CUMULATIVE\_FREQUENCY [i] \leq N3 \ \&\& \ N3 < CUMULATIVE\_FREQUENCY [i+1]$

12.  $RANGE [i] = (N3 - CUMULATIVE\_FREQUENCY [i]) < (CUMULATIVE\_FREQUENCY [i+1] - N3) ? BIN [i] : BIN [i+1]$
13.  $RANGE [i] = end (BIN)$

From the quartile range, we then named four state p1, p2, p3, p4 of moving difference based on the bin range. Formula is bellow,

1.  $MOVING\_DIFFERENCE [i] = STOCK\_DATA\_MOVING\_DIFFERENCE [i]$
2. *foreach*  $MOVING\_DIFFERENCE$  as  $key \Rightarrow value$
3. *if*  $RANGE [0] \leq value \ \&\& \ value < RANGE [1]$
4.  $P\_STATE [key] = "P1"$
5. *elseif*  $RANGE [1] \leq value \ \&\& \ value < RANGE [2]$
6.  $P\_STATE [key] = "P2"$
7. *elseif*  $RANGE [2] \leq value \ \&\& \ value < RANGE [3]$
8.  $P\_STATE [key] = "P3"$
9. *elseif*  $RANGE [3] \leq value \ \&\& \ value < RANGE [4]$
10.  $P\_STATE [key] = "P4"$
11. *else*
12.  $P\_STATE [key] = NULL$

After set those state, we also consider transition states. Here is the formula,

1. *for*  $key = 0; key < count (P\_STATE) - 1; key++$
2. *for*  $i = 1; i \leq 4; i++$
3. *for*  $j = 1; j \leq 4; j++$
4. *if*  $P\_STATE [key] == "P" . i \ \&\& \ P\_STATE [key + 1] == "P" . j$
5.  $TRANSITION\_STATES [key + 1] = "P_" . i . j$

Then we have to count the state and see how many state established based on P state. Formula,

1.  $STATE\_COUNT = array ( 'P1' \Rightarrow 0, 'P2' \Rightarrow 0, 'P3' \Rightarrow 0, 'P4' \Rightarrow 0,$

```

        'P_11' => 0, 'P_12' => 0, 'P_13' => 0, 'P_14'
=> 0,
        'P_21' => 0, 'P_22' => 0, 'P_23' => 0, 'P_24'
=> 0,
        'P_31' => 0, 'P_32' => 0, 'P_33' => 0, 'P_34'
=> 0,
        'P_41' => 0, 'P_42' => 0, 'P_43' => 0, 'P_44'
=> 0
    )
2. foreach P_STATE as key => value
3.   if value == "P1"
4.     STATE_COUNT ["P1"] ++
5.   elseif value == "P2"
6.     STATE_COUNT ["P2"] ++
7.   elseif value == "P3"
8.     STATE_COUNT ["P3"] ++
9.   elseif value == "P4"
10.    STATE_COUNT ["P4"] ++

11. foreach TRANSITION_STATES as key => value
12.   for i = 1; i <= 4; i++
13.     for j = 1; j <= 4; j++
14.       if value == "P_" . i . j
15.         STATE_COUNT ["P_" . i . j] ++

```

Then we have to create transition matrix from state count. Formula is,

```

1. TRANSITION_MATRIX = array (
    array(
        STATE_COUNT ["P_11"] /
        STATE_COUNT ["P1"],
        STATE_COUNT ["P_12"] /
        STATE_COUNT ["P1"],
        STATE_COUNT ["P_13"] /
        STATE_COUNT ["P1"],

```



```

STATE_COUNT ["P_14"] /
STATE_COUNT ["P1"]
),
array(
STATE_COUNT ["P_21"] /
STATE_COUNT ["P2"],
STATE_COUNT ["P_22"] /
STATE_COUNT ["P2"],
STATE_COUNT ["P_23"] /
STATE_COUNT ["P2"],
STATE_COUNT ["P_24"] /
STATE_COUNT ["P2"]
),
array(
STATE_COUNT ["P_31"] /
STATE_COUNT ["P3"],
STATE_COUNT ["P_32"] /
STATE_COUNT ["P3"],
STATE_COUNT ["P_33"] /
STATE_COUNT ["P3"],
STATE_COUNT ["P_34"] /
STATE_COUNT ["P3"]
),
array(
STATE_COUNT ["P_41"] /
STATE_COUNT ["P4"],
STATE_COUNT ["P_42"] /
STATE_COUNT ["P4"],
STATE_COUNT ["P_43"] /
STATE_COUNT ["P4"],
STATE_COUNT ["P_44"] /
STATE_COUNT ["P4"]
)
)
)

```

We also calculate the probability of the prediction. So based on that transition matrix we calculate a probability of the predicted result. Here is the probability formula,

1.  $Q [1] = TRANSITION\_MATRIX$
2.  $for\ i = 2; i \leq 9; i++$
3.  $Q [i] = matrix\_mult\_4\_4 (Q [i-1], TRANSITION\_MATRIX)$
4.  $PROBABILITIES = array (Q [8][0][0], Q [8][0][1], Q [8][0][2], Q [8][0][3])$

Set a price range of the probability to define the forecast more widely. Formula,

1.  $PRICE = STOCK\_PRICES\_OF\_A\_COMPANY$
2.  $LAST\_PRICE = reset (PRICE)$
3.  $for\ i = 0; i < 5; i++$
4.  $PROBABILITY\_PRICE\_RANGE [i] = LAST\_PRICE + RANGE [i]$

And lastly, sum up the all calculation for final prediction result.

1.  $for\ i = 0; i < 4; i++$
2.  $Probability = PROBABILITIES [i] * 100$
3.  $Price\ Prediction = PROBABILITY\_PRICE\_RANGE [i]\ to\ PROBABILITY\_PRICE\_RANGE [i + 1]$

### 3.5.3 Markov Model

- It is a random process characterized as memory less. The next position depends only on the recent state and not on the sequence of events that lead it.
- A sequence of states like,  
 $X_1, X_2, X_3, X_4$   
 $P( X_{k+1} = y | X_1 = x_1, \dots, X_k = x_k ) = P( X_{k+1} = y | X_k = x_k ) P( X_{k+1} = y | X_k = x ) = P( X_2 = y | X_1 = x )$
- The transition from  $X_{t-1}$  to  $X_t$  depends only on  $X_{t-1}$  (Markov Property).



P3 = 3<sup>rd</sup> Interval,

P4 = 4<sup>th</sup> Interval

After establishing the intervals for each data set of difference prices, prices are labelled considering the intervals means each individual difference price fell in to the label if it is in the range of that particular interval. Once it finished it's labeling for each difference price with its corresponding interval, then we counted the number of transitions for each individual difference price interval to the next difference price interval. We also recorded the number of points or count number belonging to each interval. After this, the transition matrix was developed through that process. Every entry of the transition matrix is considered to be the probability of the data points transitioning from one state to another state and with the corresponding to the appropriate rows and columns.

Our transition matrix is of the following form:

<i>P11</i>	<i>P12</i>	<i>P13</i>	<i>P14</i>
<i>P21</i>	<i>P22</i>	<i>P23</i>	<i>P24</i>
<i>P31</i>	<i>P32</i>	<i>P33</i>	<i>P34</i>
<i>P41</i>	<i>P42</i>	<i>P43</i>	<i>P44</i>

Figure -3.5: Transition Matrix

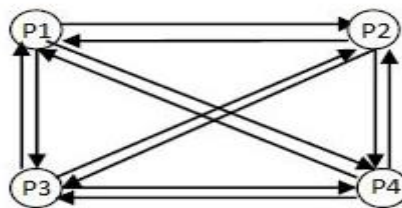


Figure -3.6: Transition Nodes

Let us consider that our transition matrix is represented by Q. We iteratively computed Q<sup>2</sup>, Q<sup>3</sup>, Q<sup>4</sup>, ....., Q<sup>8</sup>. Then we had considered 1 as being a particular state of the steady state probability. So that, we formed 4 steady state probability matrixes to predict or tell us the probability or possibility of the stock price lying in a particular range. The matrices are of the form:-

$$\mathbf{A} = [ 1 \ 0 \ 0 \ 0 ]$$

$$\mathbf{B} = [ 0 \ 1 \ 0 \ 0 ]$$

$$\mathbf{C} = [ 0 \ 0 \ 1 \ 0 ]$$

$$\mathbf{D} = [ 0 \ 0 \ 0 \ 1 ]$$

Matrix A when multiplied with Q2 , Q3 , ..... , Q8 iteratively, finally give us the probability of the stock price lying in interval P1. Similarly when matrix B is multiplied iteratively, it gives the probability of the stock price lying in interval P2, matrix C gives us the probability for interval P3 and D for interval P4 respectively.

### 3.5.5 Implementation

- Firstly, we have taken the historical data from Dhaka Stock Exchange (DSE) official website for five chosen stocks.
- We have calculated the moving average of the stock and the differences of it and then it was inserted into the MySQL database.
- We are using PHP for the calculation.
- The whole website has been designed using HTML, CSS, PHP, Ajax and JavaScript.
- To adapt to different screen sizes, we make the website responsive.

## Login Page

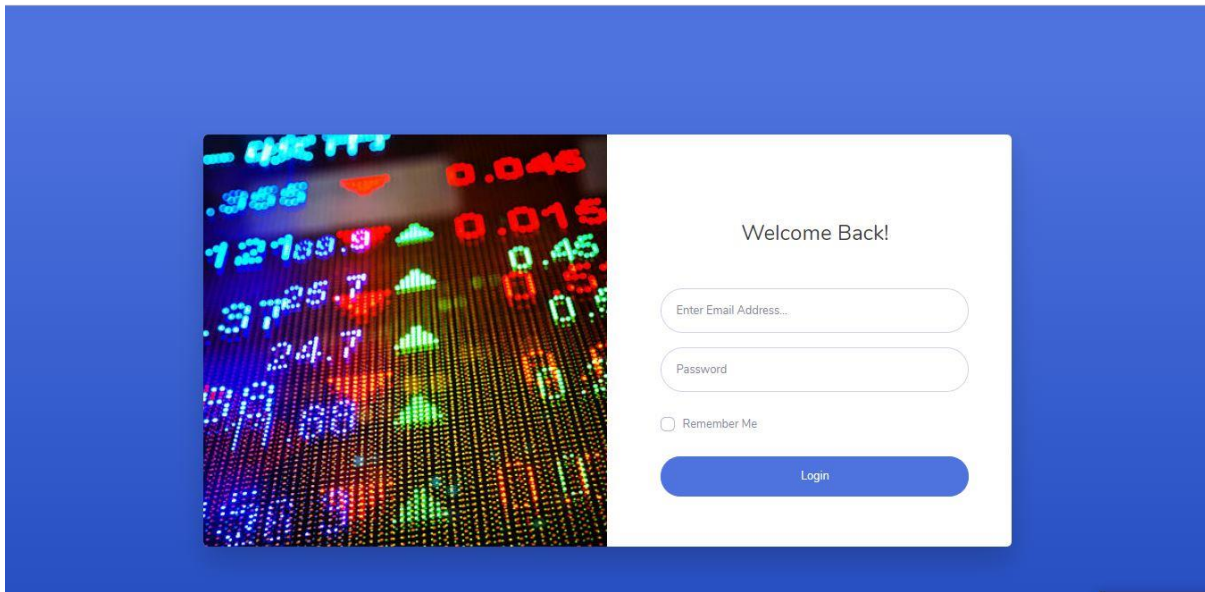


Figure -3.7: System Login Page

We are allowing a login validation of our system which help us to verify valid user and ensure security. There are two types of user available in our system who are allow to access, Admin and Normal User.

## Admin Panel

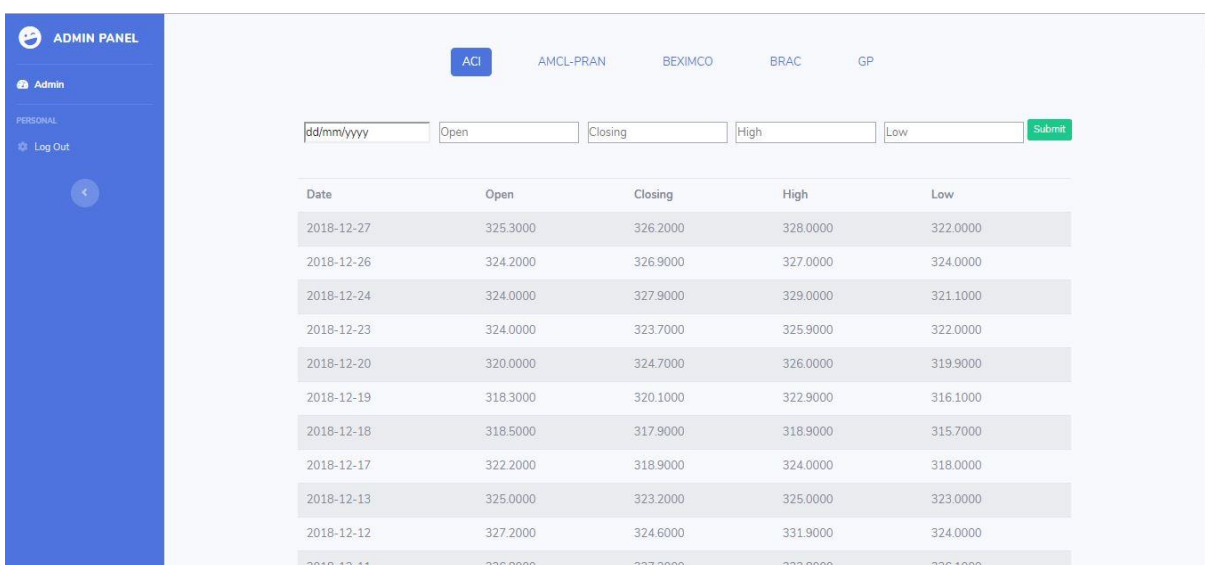


Figure -3.8: System Admin Panel

Admin is allowed to input stock data of different company and it will save into database.

## User Panel

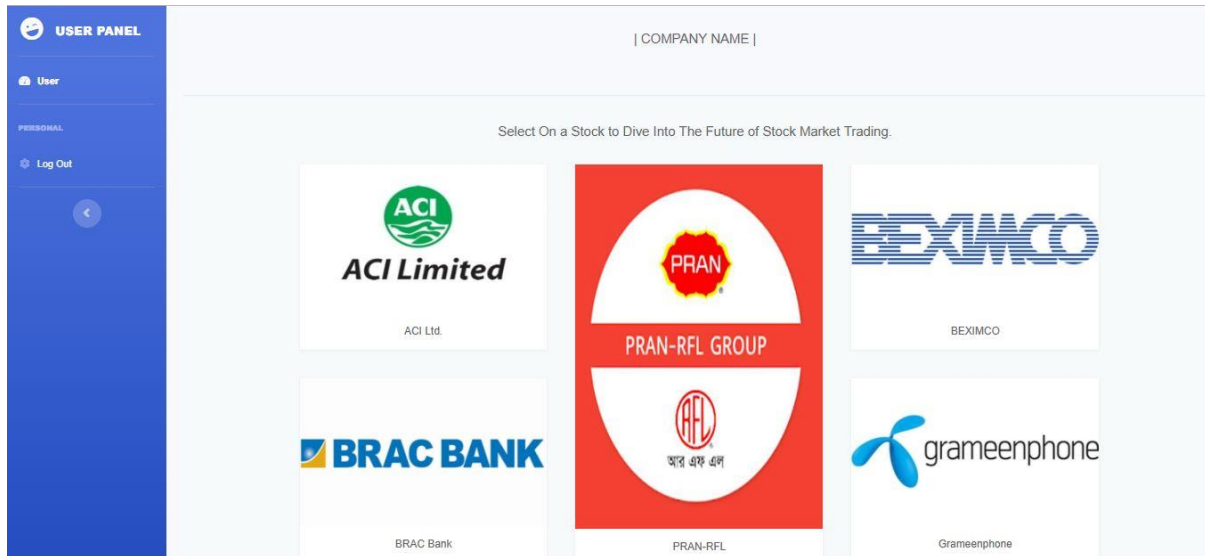


Figure -3.9: User View of Dashboard

After login to our site, user can see various company. If he or she click the image of the company, they can see the prediction prices.

## Prediction

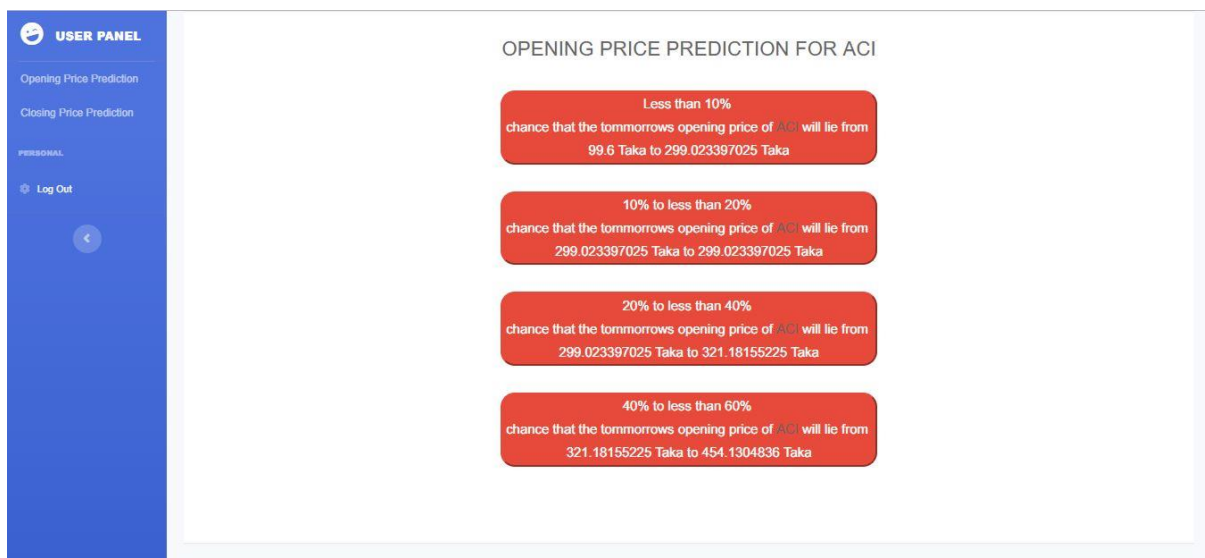


Figure -3.10: Opening Price of a company

After Click on a particular company, user can see the opening price prediction of that company.

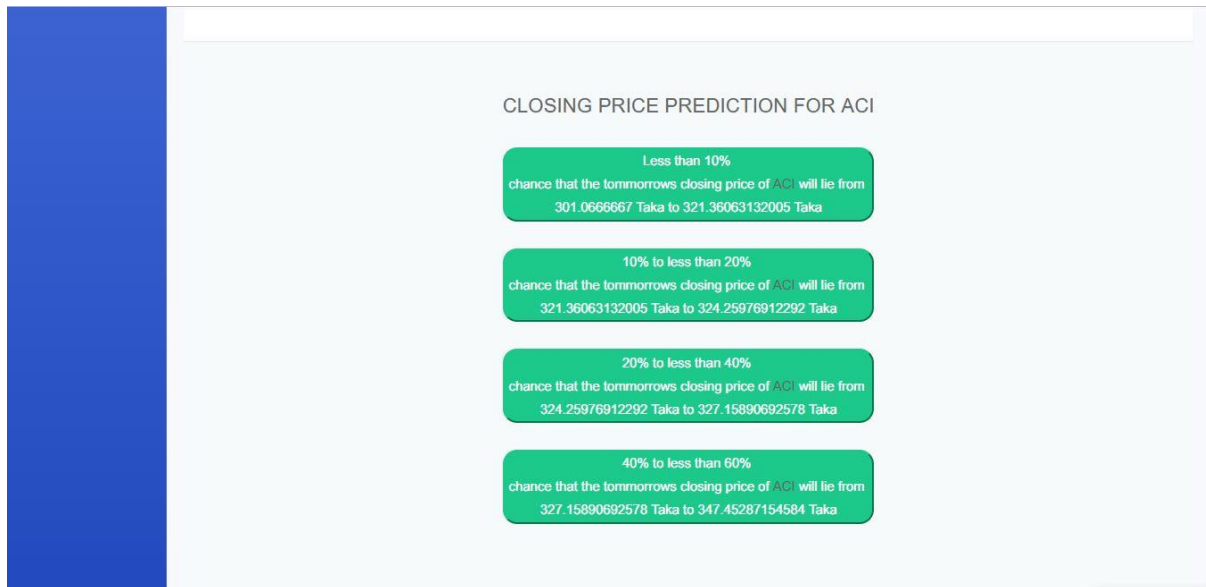


Figure -3.11: Closing Price of a company

Along with opening price, it also shows the closing price of that company.

### 3.5.6 Platform used for the system

- Platform : Web
- Language: HTML, CSS, PHP, JavaScript, Ajax, MySql
- Tools: Sublime, Notepad ++, Xampp



## **Sublime**

In our project, we have used sublime to program HTML, CSS, PHP, JavaScript and Ajax that is used on web.



Figure -3.12: Sublime

## **Xampp**

In our project, we have used xampp to create a local host server and create database for our project.



Figure -3.13: Xampp

## CHAPTER 4

### EXPERIMENTAL RESULTS AND DISCUSSION

#### 4.1 Introduction

In this section we will discuss what will be the result which is. We will also test our project and compare the result with actual result after implementation. We have to collect the data first to implement the project. Also we have to calculate the value in proper way. Lastly, we have to test all those features.

#### 4.2 Experimental Result

Table 4.1: Closest Predicted Results

Company Name	Closest Predicted Opening Price (TK)
ACI Ltd.	321.1815 – 454.1304
PRAN-RFL	233.4824 – 327.4119
BEXIMCO	22.7540 – 31.5864
BRAC Bank	70.9018 – 119.5143
Grameenphone	351.6953 – 545.8926

#### 4.3 System Testing

System testing is a very important part of a system. From a testing we can find out the problem which is called bug. If we can find the bug, we will able to fix that specific bug. To maintain the system properly, it should be bug free.

Table 4.2: System Testing

<b>Test Case</b>	<b>Expected Result</b>	<b>Actual Result</b>	<b>Test Result</b>
Login	User or admin should be able to login to the system successfully.	Can login successfully.	Pass
Logout	User or admin should be able to logout to the system successfully.	Can logout successfully.	Pass
Submit	Admin should be able to submit the input data to the system successfully.	Can submit successfully.	Pass
Opening Price Button	User should be able to see the predicted opening price from the system successfully.	Can see the opening price successfully.	Pass
Closing Price Button	User should be able to see the predicted closing price from the system successfully.	Can see the closing price successfully.	Pass

We did some test to our system to verify that our system is properly worked. We test few features of our system.

## 4.4 Discussion

Table 4.3: Result Comparing

<b>Company Name</b>	<b>Closest Predicted Opening Price (TK)</b>	<b>Actual Opening Price (TK)</b>	<b>Closest Predicted Closing Price (TK)</b>	<b>Actual Closing Price (TK)</b>
ACI Ltd.	321.1815 – 454.1304	329.9	327.1589 – 347.4528	327.6
PRAN-RFL	233.4824 – 327.4119	237.1	235.9208 – 238.3142	236.3
BEXIMCO	22.7540 – 31.5864	23.7	23.6627 – 24.4269	25.5
BRAC Bank	70.9018 – 119.5143	73.6	72.3311 – 86.5247	74.9
Grameenphone	351.6953 – 545.8926	370	368.3377 – 380.9670	368.3

Here we can see, ACI Ltd. predict four range of prediction for opening and closing but the closest prediction that match with the actual price is respectively 321.1815-454.1304 and 327.1589-347.4528. So we can consider that for ACI Ltd. our system is worked much properly and the prediction of closing price is on point.

For PRAN-RFL, it also shows four result and among this results the closest value which is matched with actual value is respectively 233.4824 – 327.4119 and 235.9208 – 238.3142. So also for PRAN-RFL we can say our system worked correctly.

In BEXIMCO, the scenario is different, the opening price is matched with actual price of it but the closing price doesn't match with actual price. So in here, we can say our system is not fully correct for BEXIMCO. It is partially correct.

Considering BRAC Bank, it also match with the actual opening and closing price of this company. So our system again worked properly for BRAC Bank.

Finally, Grameenphone shows our accuracy in new level. The opening price of it is matched with actual price but the closing price is fully matched accurately with the actual closing price of it. We can now say that our system may work so perfectly for many companies.

## **CHAPTER 5**

### **SUMMARY, CONCLUSION, LIMITATIONS AND IMPLICATION OF FUTURE RESEARCH**

#### **5.1 Summary**

Markov Chain model is a method of predicting future data. We use it for predicting the share market's stock price of Bangladesh. After implemented the model, we saw that the accuracy level is not that so high as other prediction using markov model has. Its because of the circumstances of Bangladesh. Bangladesh share market's stock price depend on many things. We can call it factors. By varying these factors, the stock price goes up and downs. If we don't consider those factors the accuracy level might not be high. Our model show almost accurate result but it will be more accurate if we consider that.

#### **5.2 Conclusion**

In this study, we propose and implemented a system which can predict future stock price of Bangladesh share market. System provide lots of information about share data. Our system can provide reliable information about future stock price of a share. It will earn the reliability of people and will encourage the people to invest more in share market. Their hope will be grown up again for Bangladesh share market.

#### **5.3 Limitations**

We have some limitation in our project. We can't provide much company to predict. Our accuracy is not that high. We can't include much feature in our web page, there should be more feature in our web page. Our web site is easily accessible which could be problematic in future.

#### **5.4 Implication of Future Research**

In our present work, we have tried to making the system reliable to people. Our plan for future is to make it more reliable. Add more companies to involve more people. We will try to increase our prediction accuracy. We will add more feature to our web page in near future. Which will be also beneficial for the people.

## REFERENCES

- [1] Learn about Stock Market, available at << <https://www.investopedia.com/articles/investing/082614/how-stock-market-works.asp>>>, last accessed on 25-03-2019 at 9:00pm.
- [2] Chiang W C, Urban T L and Baldrige G W (1996), A Neural Network Approach to Mutual Fund Net Asset Value Forecasting. *Omega*, Vol. 24 (2), pp. 205-215.
- [3] White H (1998), Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns, *Proceedings of the Second Annual IEEE Conference on Neural Networks*, Vol. 2, pp. 451-458.
- [4] Kim S H and Chun S H (1998), Graded forecasting using an array of bipolar predictions: application of probabilistic neural networks to a stock market index. *International Journal of Forecasting*, Vol. 14, pp. 323- 337.
- [5] Kimoto T, Asakawa K, Yoda M and Takeoka M (1990), Stock market prediction system with modular neural networks, *Proc. International Joint Conference on Neural Networks*, San Diego, Vol. 1, pp. 1-6.
- [5] Kim S H and Chun S H (1998), Graded forecasting using an array of bipolar predictions: application of probabilistic neural networks to a stock market index. *International Journal of Forecasting*, Vol. 14, pp. 323- 337.
- [6] Raposo R De C T and Cruz A J De O (2004), Stock Market prediction based on fundamentalist analysis with Fuzzy-Neural Networks. [http://www.labic.nce.ufrj.br/downloads/3wses\\_fsfs\\_2002 .pdf](http://www.labic.nce.ufrj.br/downloads/3wses_fsfs_2002.pdf)
- [7] Thammano A (1999), Neuro-fuzzy Model for Stock Market Prediction, *Proceedings of the Artificial Neural Networks in Engineering Conference*, ASME Press, New York, pp. 587-591.
- [8] Abraham A, Nath B and Mahanti P K (2001), Hybrid Intelligent Systems for Stock Market Analysis, *Proceedings of the International Conference on Computational Science*. Springer, pp. 337-345.
- [9] Romahi Y and Shen Q (2000), Dynamic Financial Forecasting with Automatically Induced Fuzzy Associations, *Proceedings of the 9th international conference on Fuzzy systems*, pp. 493-498.
- [10] Learn about Data Mining, Available at << <https://searchsqlserver.techtarget.com/definition/data-mining>>>, last accessed on 25-03-2019 at 10:00pm.

- [11] Vinciarelli A and Luettin J (2000), Off-line cursive script recognition based on continuous density HMM, Proceedings of the 7th International Workshop on Frontiers in Handwriting Recognition, Amsterdam, pp. 493-498.
- [12] Liebert M A (2004), Use of runs statistics for pattern recognition in genomic DNA sequences. Journal of Computational Biology, Vol. 11, pp. 107-124.
- [13] Jelinek F, Kaufmann M, Mateo C S (1990), Selforganized language modelling for speech recognition, in Readings in Speech Recognition (Eds. Alex Waibel and Kai-Fu Lee), Morgan Kaufmann, San Mateo, California, pp. 450-506.
- [14] Huang X, Ariki Y, Jack M (1990), Hidden Markov Models for speech recognition. Edinburgh University Press.
- [15] Xie H, Anrae P, Zhang M, Warren P (2004), Learning Models for English Speech Recognition, Proceedings of the 27th Conference on Australasian Computer Science, pp. 323-329.
- [16] R. Sasikumar, A. Sheik Abdullah (April 2016), Forecasting The Stock Market Values Using Hidden Markov Model, International Journal of Business Analytics and Intelligence, Vol. 4, Issue 1.
- [17] Li Z, Wu Z, He Y and Fulei C (2005), Hidden Markov model-based fault diagnostics method in speed-up and speed-down process for rotating machinery. Mechanical Systems and Signal Processing, Vol. 19(2), pp. 329-339.
- [18] Md. Rafiul Hassan, Baikunth Nath (2005), Stock Market Forecasting Using Hidden Markov Model: A New Approach, 5th International Conference on Intelligent Systems Design and Applications (ISDA'05).
- [19] Nguyet Nguyen (2018), Hidden Markov Model for Stock Trading, International Journal of Financial Studies 2018, 6, 36.

# APPENDIX

## Appendix A: Plagiarism Report



Figure -A1: Plagiarism Report