A ‘Financial Market’ is a mechanism that allows people to buy and sell or trade financial securities such as stocks and bonds. Indian financial market is one of the oldest markets in the world with Bombay Stock Exchange (BSE) starting in 1850s, when 4 Gujarati and 1 Parsi stockbroker would gather under banyan trees in front of Mumbai’s Town Hall and do business. The group eventually moved to Dalal Street in 1874 and in 1875 became an official organization known as 'The Native Share & Stock Brokers Association'. In 1956, the BSE became the first stock exchange to be recognized by the Indian Government under the Securities Contracts Regulation Act. The Bombay Stock Exchange developed the BSE Sensex in 1986, giving the BSE a means to measure overall performance of the exchange. The stock exchange in India is mainly centered around Bombay (presently Mumbia) with security brokers doing the business. The Bombay Stock Exchange switched to an electronic trading system in 1995. This automated, screen-based trading platform called BSE On-line trading (BOLT) currently has a capacity of 80 lakh orders per day. The BSE has also introduced the world's first centralized exchange-based internet trading system, BSEWEBx.co.in to enable investors anywhere in the world to trade on the BSE platform.

Around 1990s the participation of people in financial markets was poor, although India had a high household savings rate. Less than five percent of the population participated in the market directly or through mutual funds. There was little foreign investment as India’s market was perceived to be highly risky due to poor trading infrastructure and poor corporate governance. In April 1992 Indian government had created the Securities and Exchange Board of India (SEBI) to control and develop the financial market. In 1994, a team of specialist put together by Indian government, were asked to design and implement a series of programs to address the problems facing India’s financial markets. This regulatory project the ‘Financial Institutions Reforms and Expansion’ (FIRE) has initiated several market reforms.
Since then financial markets in India have become more organized and developed with Bombay Stock Exchange (BSE) and National Stock Exchange (NSE) overseeing the securities exchange. These structural changes have made BSE, one of the leading exchanges in the world. Even with all the required changes participation of the people in the market has not improved much. Added to it, understanding the market has become very difficult due to its complexity. An investor is interested to predict the market movement in order to take a right decision at the right time that yields profit. Unscrupulous market players have manipulated the market trends to suit their interest there by making an investor to look for other alternatives to invest.

The advent in Information Technologies made several business applications online. Hence, a large number of emerging applications such as Stock Markets, Communication Networks etc., have implemented Information Technology tools to better the business performance. In Stock Markets, the primary interest of the investor is to find the possible market movements, and predict the price of scrip in advance. It is a challenge to researchers and the academicians to propose innovative methods to predict the markets with a good degree of accuracy, as the market movements are highly volatile. However using the huge amount of data generated by Stock Market Systems several investor friendly applications were development. The data thus generated is useful in finding latest trends hidden using Data Mining techniques for business continuity. In particular, investors analyze the relevant trends favoring them, making such analysis as a continuous online process is of recent interest. Several Data Mining techniques have been proposed to extract patterns from data streams fall in a period of interest. Recently researchers are proposing computationally effective algorithms relevant to stock market data.

1.1 Stock Market Basics

In financial terminology, stock is the capital raised by a corporation, through the issuance and sale of shares. A share is a basic unit of account for various financial instruments including stocks, mutual funds, and limited partnerships. In British English, the usage of the word “share” alone to refer solely to stocks.
A share is one of a finite number of equal portions in the capital of a company, entitling the owner to a proportion of distributed but not reinvested profits (dividends) and to a portion of the value of the company in case of liquidation. Shares can be voting or non-voting, meaning they either do or do not carry the right to vote on the board of directors and corporate policy. A shareholder is any person or organization that owns one or more shares of a corporation's stock. The aggregate value of a corporation's issued shares is its market capitalization.

A stock market is an institution that facilitates the exchange of stocks between buyers and sellers. It is a general term used to refer to the organized trading of securities through various exchanges and through over-the-counter market. A “Stock Exchange” is a specific form of the stock market, and is a physical location where stocks and bonds are brought and sold, such as “National Stock Exchange of India”, “Bombay Stock Exchange”, New York Stock Exchange, NASDAQ, etc.,. The growth of electronic transactions facilitated through Information Technology, a large fraction of stock market transactions is not centrally located in a particular location. Any buyer can now purchase stocks through a brokerage or invest through Dividend Reinvestment Plans and Direct Investment Plans.

Stock trading has evolved tremendously. Co-production between technology and society has led the push for effective and efficient ways of trading. Most of the recent advancements with the trading have been due to the Internet. The Internet has allowed online trading. Within seconds of an order for a stock, the transaction can now take place. In contrast to the past where only those who could afford the expensive stockbrokers, anyone who wishes to be active in the stock market can now do so at a very low cost per transaction. Trading can even be done through Computer-Mediated Communication (CMC) use of mobile devices such as handheld computers and cellular phones. These advances in technology have made day trading possible. In the present day scenario anyone having an on-line account for brokerage, can easily enter into a digital stock exchange site and do the transaction on-line.
A stock market performs the following functions:

a. Connecting those who seek money (people who want public money to be invested in their ventures) with those who can provide it (investors who want buy the stocks).

b. Create an auction mechanism that decides the prices of stocks.

c. Distributing the future risk of investments across many millions of share holders.

d. Providing the claim tickets upon which future wealth can be staked.

e. Connecting financial institutions together to create money.

1.2 Motivation

The financial service industries, like stock markets throughout the world have become increasingly complex in recent times. Experts recognize the importance of finance and understanding basic finance relationships are the key to modern financial security. Hence the financial education and awareness in the functioning of the financial markets are essential to make good financial choices for investing in stock markets. Due to limited knowledge in financial markets and their functionality, a majority of the people depend on financial reports published in leading business papers and magazines or on agents in the financial markets who themselves depend on different market predicting software. These software though provides lot of features for trading have certain limitations. One of the main limitations being prediction of a stock price from time-to-time in a day.

The financial markets use automated computer programs developed with different technologies to predict the trades in the market. With the potential techniques in Data Mining and with the growth of the technologies to handle the huge databases, the predictive technologies have started growing tremendously. The academic research in Data Mining also contributed a lot to the predictive technologies. The use of Data Mining is well founded on the theory that the historic data holds essential hidden and previously unknown knowledge that can be used for
predicting the future direction and assist in decision making. The prediction of stock markets is regarded as a challenging task of financial time series prediction. Data analysis is one way of predicting increase or decrease of future stocks price.

In Stock Market Analysis, identifying the factors which are going to affect the prices in advance and there by predicting the trends of a specific stock correctly becomes very difficult. Though the physical factors are generally known in advance, the euphoria created by some people pulling prices away from their true trend levels are not known. Hence the only way to predict the prices is on the data that is generated from time to time and using it in identifying the interesting patterns from it. The recent research is mainly concentrated on identifying methodologies for analysis of stream data to identify these patterns.

The objective is to use these methodologies on stock market data to process it dynamically to match the current data with the existing stored patterns for identifying the trend, and there by predict the price changes for the current day and possibly for the days to come. In the financial market several experts use the technical analysis to predict the price trends for short term and long term based on, the business done by the companies, economy of the country, political decisions of the Governments, economic constraints, etc. and there by certain stocks are identified as hot or cold. An investor will look at a support price and resistance levels for a given stock to decide the transaction.

The above factors have motivated the work in this thesis on identifying the methodologies and come out with a generic algorithm and applying them to domains which have the time series data. To verify the applicability of the algorithm we use the stock market data which is essentially a time series data and predict the price movements as best as we can.

This thesis focuses on tailoring Data Mining techniques to stock market domain. As the data is continuous and time stamped, temporal data mining techniques were mainly looked at for finding interesting patterns to predict the price. Following of the temporal data mining techniques were explored in this research work:
i. Similarity Profile Association Mining.

ii. Frequent Episode Mining.

Similarity Profile Association Mining is to find the similar records in the huge data set and to find a mechanism to use them for predicting the price of the scrip on a given day.

Frequent Episode Mining on the other hand is to discover frequently occurring sequence of events. In the process identify the sequence of events which have some dependency on each other, and appear in the order of occurrence. These sequences of events can occur serially, parallelly or mixed depending on which episodes are categorized as serial episodes, parallel episodes or mixed episodes. As the stock market data is sequential in nature, we will be looking for serial episodes. The events here are gain, loss or neutral increase in price of the scrip. An even sequence is a combination of these transactions over a time interval. This period of time interval can be taken as the window for an episode. Hence an episode is defined here as a triplet \((S, T_s, T_e)\), where \(S\) is a Event Sequence a set of events occurring at different times between start time \(T_s\) and end time \(T_e\). Our task here is to discover significant episodes and then identify the frequency at which these episodes occur, and thereby identify some mechanism to predict the price of scrip.

1.3 Objectives

It is difficult to predict the exact price of stock at any given time, rather we can only predict nearer to the future price of stock. The main objective of this work is to find the methods which can best predict the dynamic movement of stock prices in the market. We have chosen some Data Mining techniques for this purpose. The idea is to identify existing patterns in the historical data, and maintain a database for it. Comparing the current price movement with the existing patterns, we predict the possible price movement in the future. In the process we identify any new patterns that we come across. For the purpose of identifying the interestingness of the pattern we look into the percentage of gain or loss we make, and the amount of volume transacted.
Hence our objectives in this work using time series data is to:

1. Identify interesting patterns from the historical data that is available and maintain a database for it.

Stock market data consists of a set of time-stamped transactions. Mining frequent patterns in transaction database has been studied extensively in data mining research. However, it is observed that, most of the existing frequent pattern mining algorithms like Apriori, FP-growth do not consider the time stamps associated with the transactions. Taking the time stamps in the transactions into account we define a transitional pattern of the set of transactions for the day, which captures the dynamic behavior of frequent patterns in a transaction database. These identified transitional patterns include both positive and negative transitional patterns indicating the upward and downward movement of price of scrip. Their frequencies increase/decrease dramatically at some point of time in a day, which are registered as significant milestones (time points of a transaction at which the frequency of the pattern changes most significantly). Using the Elliot wave principle, we identify the most significant frequency changes to be between three to five variations in the stock price. These interesting patterns are identified and recorded as interesting knowledge from the database.

2. Search for the matching pattern for the current data to best predict the future price of the scrip, for the immediate future or for the day and further future.

Once this knowledge base is created, for the current trend of price movement we need to find a matching pattern through which we can predict the future price of the scrip. Filtering this huge knowledge base for finding appropriate pattern dynamically is a challenge. We use opening and closing prices, gain/loss percentages, volume transacted for the scrip to filter the database. If the current trend does not find a matching pattern, then that trend is identified as a new pattern and is added to the existing knowledge base.
Predicting the price movement of a stock is done at two levels, predicting for the immediate movement and predicting for the day. For predicting the price of scrip for immediate future we use pattern matching for transitional frequencies with significant milestones, and for predicting the price for the day and for further we use similarity profiled association mining.

3. Discover frequently occurring sequence of high gain events in the data set and there by identify frequent episodes.

4. Model a system using the user's previous experience in buy/sell for further enhancing the prediction.

The aim of developing the system is to utilize the knowledge gained by the user along with the other mining methods for enhancing the reliability of the decision made by the user as he buys/sells to maximize/minimize his profit/loss. Many a times the user forgets his previous experience and may tend to do the same mistake again and again unintentionally. His decisions mainly depend on the market predictions and his own judgment of timing which may be purely instinctive. In this context, it is essential that user should have support of his own previous experience. The proposed system is developed using supportive knowledge base along with user's previous experience in transaction, which interacts with the user to enhance his decision making.

1.4 Organization of the Thesis

In Chapter 2 we present a generic algorithm for similarity profiles-based prediction which is applicable for different domains. The algorithm is applied on stock market data in the remaining part of the work to show the efficiency of the algorithm.

Chapter 3 makes survey of relevant literature such as working of Stock market domain and several market indicators which help in stock price prediction. The previous research on Stock price movements and market predictions were
discussed mainly the Different Fundamental Analysis methods and Technical Analysis methods were looked at. The role of Data Mining in Stock market predictions and specially the temporal data mining techniques were discussed. Market movement has motivated to study different association rule mining and classification and prediction methods. Since timing is important in stock market transactions, we have studied temporal data mining methods. Sequential patterns and frequent episodes, and their different use in applications are studies to find mechanism to predict stock market movement.

In Chapter 4 we have used similarity profile temporal association for indentifying patterns in the stock market data. A specific mechanism was identified looking at the support for different gains in different time intervals and there by predicting the price of the stock from time to time for a given scrip on a given day. Using the historical data we create a support tables for different gain ranges. We find a mechanism to link the similarity in current opening price of the day and the support table to predict the price of a stock from time to time for the day. Depending on the available historic data, the data is converted into 23 time slots time stamped data. For the prediction process, we have derived certain prediction equations and applied it on the current days opening price to continuously predict the price in future time slots.

Chapter 5 presents a mechanism to predict a stock movement on long term basis. The technique involved is to identify high gain event sequences and further identifying the effect of frequency of these sequences on the price. The prediction can be made for a week or projection can be done for a month depending on the requirements of the user. Predicting stock price across a time span of a week involves the discovery of repetitive patterns in the stock data. Identifying sequential events in the data and the relation between these sequential events represents episodes. Episodes in these patterns for the day are identified to study the effect of these episodes on the price movement. Our objective in this work was to identify these fluctuating patterns and the relations among them, and there by classify the data on these fluctuating patterns, and there by representing
classification rules. Data is further processed to create week wise test data for the purpose of classification. Remaining data is used for predicting.

In Chapter 6 we have presented a knowledge-based stock market prediction information system. Depending on the available products in the market, we have developed a product combining the two methods that are developed in Chapter 3 and Chapter 4. In this work we have also included the previous experience of a user, so that an investor can analyze the user’s previous experience under different circumstances and provide assistance to enhance his decision making.

Chapter 7 concludes the present work done under different constraints and provides possible improvements that can be done in future.