Chapter 7

Conclusions

7.1. Present Work

As we know the financial markets use automated computer programs developed with different technologies to predict the market movements. On observation we have found that all these products use different Artificial Intelligence mechanism in predictions. 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. With the intension of using Data Mining techniques we started this work to find an effective mechanism for prediction. 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.

We have specified the following objectives in this work:

1. Identifying 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. The available data that we have collected for this work was completely random in terms of number of records available for a day. Lot of pre-processing is done to standardize the data into 23 time slots. Taking the time stamps in the transactions into account we have defined a transitional pattern of the set of transactions for the day, which captured the dynamic behavior of frequent patterns in a transaction database. 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. Since the prediction was for the current day, the available value as the transaction starts will be the opening price. This value is further processed to calculate the gain on opening 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. Using all the knowledge acquired we would like to propose a generic algorithm which could be applicable to predict the a given transactions attribute value for all similar domains.

3. Discover frequently occurring sequence of high gain events in the data set and there by identify frequent episodes. The episodes width and the height has significance, the width indicates the

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

Our proposed generic algorithm for prediction using temporal data available in different application domain was successfully proved to be effective through the implementation over the Stock market domain.

Proposed approach is useful in predicting the market trends in their beginnings of the transactions. We proposed an approach to predict the intraday price of a stock using the historic data. Considerably large data is used for experimental verification of our approach to predict the future prices of the stocks. Given the time stamped transactions, the stock data is represented as
pattern database and similarity profiled temporal association mining is used to
discover all associated pattern records that are of relevance. Using the support
value for different price gain and the opening price of the stock for the day, we
extracted all the significant pattern records from the pattern database. Using the
current trend of the stock, we could project the future prices from time to time for
the day as efficiently as possible. Having created the pattern table, we are also
working on sequential patterns, and on classification of these sequential patterns
to come out with classifiers and there by predict the future patterns for a week or
for a specific day.

We could meet our objective of identifying fluctuating patterns and the
relations among these patterns, and there by classify the data depending on these
fluctuating patterns to represent classification rules. In the process of achieving
our objective, we have identified some recurring pattern of sequence of events
called “episodes”. The frequent episodes are the interesting patterns in the time-
series stock market data which has provided proof to our belief that these episodes
do have a significant effect on the price movement. This information is used to
formulate the problem of prediction as a classification problem, and identified
classification rules. Testing of these rules has proved that the classification is
efficient and hence prediction is accurate.

7.2. Future Enhancements

We had the limitation in choosing the number of time slots while
predicting the stock price, mainly due to the data that was available to us. If
historical data available is for smaller time intervals, the efficiency of the method
can be improved further.

In the process of classification we found the effect of episodes over the
prices to be very high, we wish to extend this work for predicting the stock price
for time-to-time in a day. Similarly support for different episodes can be studied
for prediction of the stocks.