A Review on Data Mining Applications to the Performance of Stock Marketing

D. Venugopal Setty,
Assistant Professor,
Department of Industrial Engineering and Management,
R.V. College of Engineering,
Bangalore – 560059, India

T.M.Rangaswamy
Professor,
Department of Industrial Engineering and Management,
R.V. College of Engineering,
Bangalore – 560059

K.N.Subramanya
Assistant Professor,
Department of Industrial Engineering and Management,
R.V. College of Engineering,
Bangalore – 560059, India

ABSTRACT
Data mining is the extraction of hidden predictive information from large data bases. Data mining is used to increase revenues and reduce costs. Data mining finds patterns and correlations or relationships in data by using sophisticated techniques. In today's society of high demanding customers, increasing competitiveness and in order to avoid the initial investment of gathering a data without being certain that it will be successful. A customer-focused approach will discover customer's preferences, needs and behavioral characteristics. The current approach is described for data development in share market to find the correlations between the customer feelings or perceptions and the physical characteristics of a data. Stock Market is an organized set-up with a regulatory body and the members who trade in shares are registered with the stock market and regulatory body SEBI. The stock market is also called the secondary market as it involves trading between two investors and itis often considered the primary indicator of a country's economic strength and development. The NSE is the largest stock exchange in India in terms of daily turnover and number of trades, for both equities and derivative trading.

Key words
Data mining, Stock Market, future trends, turnover, number of trades.

1. INTRODUCTION TO DATA MINING
Data are any facts, numbers, or text that can be processed by a computer. This includes: Operational or transactional data such as sales, cost, inventory, payroll & accounting; Non - operational data such as industry sales, forecast data & macro economic data; and Meta data such as logical data base design or data dictionary definitions. Data Mining is an analytic process designed to explore data (usually large amounts of data, typically business or market related) and in search of consistent patterns and /or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. Data Mining is often considered to be "a blend of statistics. Data mining helps marketing professionals improve their understanding of customer behavior. Data mining is primarily used today by companies with a strong consumer focus - retail, financial, communication, and marketing organizations. It enables these companies to determine relationships among "internal" factors such as price, product positioning & staff skills, and "external" factors such as economic indicators, competition & customer demographics.

1.1 Data mining process

The process of data mining consists of three stages: (1) initial exploration, (2) model building or pattern identification with validation / verification and (3) deployment (i.e., the application of the model to new data in order to generate predictions).

1.2 Data mining tools

The data mining tools are as follows:

1.2.1 Artificial neural networks

These are biological systems that detect patterns, make predictions and learn. The artificial ones are computer programs implementing sophisticated pattern detection and machine learning algorithms on a computer to build predictive models from large historical data bases. Neural networks grew out of the community of Artificial Intelligence rather than from the discipline of statistics. These are Non-linear predictive models that learn through training and resemble biological neural networks in structure.

1.2.2 Decision trees

A decision tree is a predictive model that can be viewed as a tree. Tree shaped structures represent a set of decisions. These decisions generate rules for the classification of a data set. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID). Linear regression is a statistical technique used to find the best-fitting linear relationship between a target (dependent) variable and its predictors (independent variables).

1.2.3 Genetic algorithms:

Optimization techniques that use process such as genetic combination, mutation and natural selection in a design based on the concepts of evolution.

1.2.4 Nearest neighbor method:

Nearest neighbor method is a prediction technique and its essence is that in order to predict what a prediction value is in one record look for records with similar predictor values in the historical data base and use the prediction value from the record that it “nearest” to the unclassified record.

1.2.5 Rule induction:

Rule induction is the form of data mining, most closely resembles the process that most people think about a vast data base. Rule induction on a data base can be a massive undertaking, where all possible patterns are systematically pulled out of the data and then an accuracy and significance are added to them that tell the user how strong the pattern is and how likely it is to occur again. The rule induction system retrieves all possible interesting patterns.

1.3 Data mining modeling:

This act of model building is something that people have been doing for a long time. Once the model is built it can then be used in similar situations where you don't know the answer. Table 1 illustrates the data used for building a model for new customer prospecting in a data warehouse and table 2 illustrates another common scenario for building models: predict what is going to happen in the future. Once the mining is complete, the results can be tested against the data held to confirm the model’s validity

Table 1 - Data Mining for Prospecting

Information	Customers	Prospects
General (e.g. demographic data)	Known	Known
Proprietary (e.g. customer transactions)	Known	Target

Table 2 - Data Mining for Predictions

Information	Yesterday	Today	Tomorrow
Static (e.g. demographic data, marketing plans)	Known	Known	Known
Dynamic (e.g. customer transactions)	Known	Known	Target

1.4 Applications of data mining:

A wide range of companies have deployed successful applications of data mining. The applications of data mining includes: banking industries, credit card company, financial industries, telecommunications industries, transportation company, turnover management, food-service menu analysis, fraud detection, package goods company, government policy setting, hiring profiles, medical management, process control, Quality control, store management, student recruiting and retention, warranty analysis, etc

2. INTRODUCTION TO STOCK MARKETING

There are many different kinds of customers with different kinds of needs and preferences. The market segmentation means; divide the total market, choose the best segments and design strategies for profitability serving the chosen segments better than the company's competitors do. In order to make superior value and customer satisfaction, the data needs information at every level. In fact, today's managers sometimes receive too much data, in which case the real problem is how to separate the useful and relevant nuggets of information from the mountains of data. The proposed system will have incorporated a method to collect data that will be relevant, accurate, current, and unbiased.

2.1 Stock exchange

A stock exchange, (formerly a securities exchange) is a corporation or mutual organization which provides "trading" facilities for stock brokers and traders, to trade stocks and other securities, thus providing a marketplace (virtual or real). Stock exchanges also provide facilities for the issue and redemption of securities as well as other financial instruments and capital events including the payment of income and dividends. The securities traded on a stock exchange include: shares issued by companies, unit trusts, derivatives, pooled investment products and bonds. To be able to trade a security on a certain stock exchange, it has to be listed there. Trade on an exchange is by members only. The initial offering of stocks and bonds to investors is by definition done in the primary market and subsequent trading is done in the secondary market. A stock exchange is often the most important component of a stock market. There is usually no compulsion to issue stock via the stock exchange itself, nor must stock be subsequently traded on the exchange. Such trading is said to be off exchange or over-the-counter. This is the usual way that derivatives and bonds are traded. Increasingly, stock exchanges are part of a global market for securities. There are 20 major Stock Exchanges in the world. Market capitalization & turnover at the end of May 2009 for few stock exchanges are as follows:

Region /Stock Exchange	Market Value (millions USD)	Total Share Turnover (millions USD)
Africa/ Johannesburg Securities Exchange	605,040	117,424
Americas/ NASDAQ	2773,684	12,256,704
Americas/ New York Stock Exchange	9574,066	7,986,835
Asia-Pacific/ Australian Securities Exchange	839,062	273,205
Asia-Pacific/ Bombay Stock Exchange	1032,589	83,906
Asia-Pacific/ National Stock Exchange of India	968,815	242,641
Europe/ London Stock Exchange	2204,320	1,483,263

2.2 The role of stock exchanges

Stock exchanges have multiple roles in the economy, this may includes: Raising capital for businesses, mobilizing savings for investment, Facilitating company growth, Profit sharing, corporate governance, creating investment opportunities for small investors, Government capital-raising for development projects, Barometer of the economy.

2.3 Listing requirements by stock exchange

Listing requirements are the set of conditions imposed by a given stock exchange on companies that want to be listed on that exchange. Conditions sometimes include minimum number of shares outstanding, minimum market capitalization and minimum annual income. Companies have to meet the requirements of the exchange in order to have their stocks and shares listed and traded there, but requirements vary by stock exchange.

1.2.6 Bombay Stock Exchange (Asia-Pacific Region)

Bombay Stock Exchange (BSE) has requirements for a minimum market capitalization of Rs.250 Million and minimum public float equivalent to Rs.100 Million.

1.2.7 NASDAQ Stock Exchange (America Region)

To be listed on the NASDAQ a company must have issued at least 1.25 million shares of stock worth at least $70 million and must have earned more than $11 million over the last three years.

2.4 Stock market

A stock market is a public market for the trading of company stock and derivatives at an agreed price; these are securities listed on a stock exchange as well as those only traded privately. It is an organized set-up with a regulatory body and the members who trade in shares are registered with the stock market and regulatory body SEBI. The stock market is also called the secondary market as it involves trading between two investors. Stock market gets investors together to buy and sell shares in companies. Share market sets prices according to supply and demand. A stock that is highly in demand will increase in price, whereas as a stock that is being heavily sold will decrease in price. Companies that are permitted to be traded in this market place are called "listed companies".

2.5 Importance of stock market

The stock market is one of the most important sources for companies to raise money. This allows businesses to be publicly traded, or raise additional capital for expansion by selling shares of ownership of the company in a public market. History has shown that the price of shares and other assets is an important part of the dynamics of economic activity, and can influence or be an indicator of social mood. In fact, the stock market is often considered the primary indicator of a country's economic strength and development. Rising share prices tends to be associated with increased business investment and vice versa. Share prices also affect the wealth of households and their consumption. Therefore, central banks tend to keep an eye on the control and behavior of the stock market. Exchanges also act as the clearing house for each transaction, meaning that they collect and deliver the shares, and guarantee payment to the seller of a security. This eliminates the risk to an individual buyer or seller that the counterparty could default on the transaction. The smooth functioning of all these activities facilitates economic growth, lower costs; promote the production of goods and services as well as employment. In this way the financial system contributes to increased prosperity.

1.2.8 Primary market

Primary market deals with the new issues of securities. In the primary markets, securities are bought by way of the public issue directly from the company. An official prospectus is published under the Corporations Law and contains all the information that is reasonably required to allow you to make an informed investment decision about the company.

1.2.9 Secondary market

The Secondary market is where existing securities are bought and sold. Secondary market deals with outstanding securities. In the secondary market shares are traded among investors. This market is made of organized exchanges and may have a trading floor, where orders are transmitted for execution. This is where all the trading of stocks are maintained and guided by the rules set down by the exchange.

2.5 Market index

An index is a statistical composite measure of the movement in the overall market or industry. Basically, indexes allow measuring the performance of a group of companies over a period of time. Companies are organized in an index according to two main methods or weighting as it is commonly termed. The movements of the prices in a market or section of a market are captured in price indices called stock market indices, e.g., the S&P, the FTSE and the Euro next indices. Such indices are usually market capitalization weighted, with the weights reflecting the contribution of the stock to the index. The constituents of the index are reviewed frequently to include/exclude stocks in order to reflect the changing business environment.There is two major classes of indexes in use:

1.2.10 Equally weighted price index

The index is calculated by taking the average of the prices of a set of companies.

> Equally weighted price Index = Sum (Prices of N companies) / divisor

1.2.11 Market capitalization weighted index

In this index, each of the N companie's prices is weighted by the market capitalization of the company.

> Market capitalization weighted index = Sum (Company market capitalization * Price) over N companies / Market capitalization for these N companies.

2.6 Market participants

A few decades ago, worldwide, buyers and sellers were individual investors, such as wealthy businessmen with long family histories and emotional ties to particular corporations. Over time, markets have become more "institutionalized" buyers and sellers. The market participants includes; investors, large institutions, issuers of securities, intermediaries and regulators.

2.8 Stock brokers

Stock broker is person who is licensed to trade in shares. They also have direct access to the share market and can act as agent in share transactions. For this service, they charge a fee. Stock brokers can also offer additional services such as portfolio management or advice. The type of broker will depend on own confidence in trading shares. Often investors, who know exactly what they want to buy, will go to a discount broker to enact the trade. Stock broker may be full service broker or discount broker.

2.10Full service broker

Full-service broker will provide you with advice on which stocks to trade. They can often operate as financial planners and help with other aspects of your investment portfolio. Because they offer advice, a full service broker usually charges between 2 and 2.5 per cent fees, depending on the size of the transaction.

2.11Discount broker

Discount broker will execute trades, but will not provide any advice. As a result brokerage charges are low. Discount brokers generally operate via the telephone, Internet or both.

2.9 Trader

In finance, a trader is some one who buys and sells financial instruments such as stocks, bonds and derivatives. Traders are either professionals working in a financial institution or a corporation, or individual investors. They buy and sell financial instruments traded in the stock markets, derivatives markets and commodity markets. Several categories and designations for diverse kinds of traders are found in finance, these includes: stock trader, day trader, pattern day trader, swing trader, floor trader and rogue trader**.**

2.12 Trading

Participants in the stock market range from small individual stock investors to large hedge fund traders, who can be based anywhere. Exchange are physical locations, where transactions are carried out on a trading floor, by a method known as open outcry. This type of auction is used in stock exchanges and commodity exchanges where traders may enter "verbal" bids and offers simultaneously. The other type of stock exchange (derivative exchanges) is a virtual kind, composed of a network of computers, where trades are made electronically via traders. Buying or selling at market means accepting any asked price or bid price for the stock, respectively. When the bid and ask prices match, a sale takes place on a first come first served basis.

2.11 Investment strategies

One of the many things people always want to know how to make money investing. There are two basic methods, namely fundamental analysis and technical analysis.

2.13Fundamental analysis

Fundamental analysis refers to analysing companies by their financial statements found in SEC filings, business trends, general economic conditions, etc Fundamental analysis uses to investigate the value of the company with regards to it potential growth in earnings. It starts with the broad analysis of the economy: economy growth, inflation, unemployment, and the level and direction of interest rate. By considering the indicators that affect the economy, financial analysts can then forecast future levels of GDP. The key concepts of fundamental analysis are: economic analysis, industry analysis, company analysis and ratio analysis

2.14Technical analysis

Technical analysis is based on an examination of the price and volume movements of individual stocks, sectors, or the market as a whole. It helps in searching for the direction of the market, sector or stock. Market technicians depend on technical analysis apart from price & volume data for identifying turning points in the market. Technical analysis includes tools like; pattern identification, moving averages, advance/decline charting styles, short selling, put/call ratio, relative strength indicators, etc.

2.11 Indian stock market

The history of Indian stock market is about 200 years old. The first organized stock market that was governed by the rules and regulations came into the existence in the form of The Native Share and Stock Brokers' Association in 1875. After gone through numerous changes this association is today better as Bombay Stock Exchange (BSE), which remains the premier stock exchange since its inception. Presently, there are 19 recognized stock exchanges out of which four are national level exchanges and the remaining are regional exchanges. National Stock Exchange and Bombay Stock Exchange are the leaders of Indian securities market in terms of listing, trading and volumes. The last 15 years of the Indian securities market can be considered as the most important part of the history where the market gone through the post liberalization era of Indian economy and witnessed the formation of Securities and Exchange Board of India (SEBI). SEBI brought substantial transparency in share market practices and managed to bring trust of domestic and international investors.

2.12 National stock exchange of India:

Capital market reforms in India and the launch of the Securities and Exchange Board of India (SEBI) accelerated the incorporation of the second Indian stock exchange called the National Stock Exchange (NSE) in 1992. After a few years of operations, the NSE has become the largest stock exchange in India. Three segments of the NSE trading platform were established one after another. The Wholesale Debt Market (WDM) commenced operations in June 1994 and the Capital Market (CM) segment was opened at the end of 1994. Finally, the Futures and Options segment began operating in 2000. Today the NSE takes the 14th position in the top 40 futures exchanges in the world. In 1996, the National Stock Exchange of India launched S&P CNX Nifty and CNX Junior Indices that make up 100 most liquid stocks in India. The Indices are owned and managed by India Index Services and Products Ltd (IISL) that has a consulting and licensing agreement with Standard & Poor's. In 1998, the National Stock Exchange of India launched its web-site and was the first exchange in India that started trading stock on the Internet in 2000. The National Stock Exchange of India Limited (NSE) is a Mumbai-based stock exchange. It is the largest stock exchange in India in terms of daily turnover and number of trades, for both equities and derivative trading. The NSE's key index is the S&P CNX Nifty, known as the Nifty, Nifty is a diversified index of 50 stocks from 25 different economy sectors weighted by market capitalization. NSE provides a fully automated screen-based trading system with the major market segments namely; Equity Market, Capital Markets, Futures & Options market, Derivatives Market, Wholesale Debt Market, Mutual Funds (MF) and Initial Public Offerings (IPO)

2.13 S&P CNX NIFTY

The S&P CNX Nifty (Standard & Poor CNX National Fifty) is the head line index on the National Stock Exchange of India Ltd. It tracks the behavior of a portfolio of blue chip companies, the largest and most liquid Indian securities. It includes 50 of the approximately 935 companies listed on the NSE. The index has been trading since April 1996 and is well suited for benchmarking. The S&P CNX Nifty is based on solid economic research and is created for those interested in investing and trading in Indian equities.

2.14 Market Representation

The S&P CNX Nifty stocks represent about 60% of the total market capitalization of the National Stock Exchange (NSE).

2.15 Diversification

The S&P CNX Nifty is a diversified index, accurately reflecting overall market conditions. The reward-to-risk ratio of S&P CNX Nifty is higher than other leading indices; is making more attractive portfolio and offers similar returns like others at lesser risk

2.16 Liquidity

Market impact cost is the best measure of the liquidity of a stock. It accurately reflects the costs faced when actually trading an index. For a stock to qualify for possible inclusion in the S&P CNX Nifty, it has to reliably have market impact cost below 0.75 %. The current impact cost of the S&P CNX Nifty for a portfolio size of Rs. 5 million is 0.07%.

2.17 Hedging Effectiveness

The basic risk of the S&P CNX Nifty futures will be lower compared to other index portfolios because of its liquidity and the S&P CNX Nifty has higher correlations with typical portfolios in India compared to other indices. These two factors imply that hedging using S&P CNX Nifty futures is superior.

2.18 Eligibility Criteria

Selection of the index set is based on following criteria:

2.19 Liquidity (Impact Cost)

For inclusion in the index, the security should have traded at an average impact cost of 0.75% or less during the last six months, for 90% of the observations. Impact cost is the cost of executing a transaction in a security in proportion to the weight of its market capitalization against the index market capitalization. This is the percentage mark up suffered while buying or selling the desired quantity of a security compared to its ideal price.

2.20 Market Capitalization

Companies eligible for inclusion in the S&P CNX Nifty must have a six-month average market capitalization of Rs 5 billion or more, during the latest six months.

2.21 Shares Outstanding

Companies eligible for inclusion in the S&P CNX Nifty should have at least 12% of its stock available to investors (float).

2.22 Domicile

The Company must be domiciled in India and trade on the NSE.

2.23 Eligible Securities

All common shares listed on the NSE are eligible for inclusion in the S&P CNX Nifty index. Convertible stock, bonds, warrants, rights and preferred stock that provide a guaranteed fixed return are not eligible.

2.24 Other Variables

A company which comes out with an IPO will be eligible for inclusion in the index. Such company must fulfills the normal eligibility criteria ie; impact cost, market capitalization and floating stock for a 3-onth period instead of a 6-month period.

2.25 Timing of changes

The index is reviewed every quarter and a six-week notice is given to the market before making any changes to the index constituents. The complete list of eligible securities is compiled based on the market capitalization criteria. The top stocks, in terms of size (market capitalization) are identified for inclusion in the index from the replacement pool. Stocks may be deleted due to mergers, acquisitions or spin-offs. Otherwise, every quarter a new eligible stock list is drawn up to review against the current constituents. If this new list warrants changes in the existing constituent list, then the smallest existing constituents are dropped in favor of the new additions.

2.26 Stock market basics

Stock market basics include shares and stocks. A Share or stock is a document issued by a company, which entitles its holder to be one of the owners of the company.

2.26.1 Share

A share is directly issued by a company through IPO or can be purchased from the stock market. By owning a share one can earn a portion of the company's profit called dividend. Also, by buying and selling the shares gets capital gain. So, return is the dividend plus the capital gain. However there is a risk of making a capital loss, if selling price of the share is below than the buying price.

2.27 Stock

A stock is nothing but a collection or a group of shares. The stock may be common stock or preferred stock.

2.28 Common Stock

Common stock represents the majority of stock. It represents ownership in a company and a claim on a portion of profits, or dividends. The dividend amount fluctuates and is not guaranteed. Shareholders are entitled to one vote per share to elect board members, who oversee the major decisions made by the company's management. In the long run, common stock yields higher returns than most other investments.

2.29 Preferred Stock

Preferred stock represents a degree of ownership in a company but usually doesn’t include voting rights. The stock holders of this type have the right to get a guaranteed fixed rate of dividend before the payment of dividend to the equity holders. They also have the right to get back their capital before the equity holders in case of winding up of the company.

2.30 Categories of stocks

2.31 Blue Chip stocks

Blue Chip stocks are considered to be the most prestigious, well-established stocks in the Stock Market. They pay dividends for a long time during good years and bad years. Examples: General Electric, IBM and Proctor & Gamble.

2.32 Growth stocks:

Growth stocks are those companies with strong growth potential. They put a great deal of effort into R&D and utilize the previous profits for future expansion. They usually pay small or no dividends at all. Investors are usually willing to wait for capital appreciation. The market price of growth stocks can be quite volatile.

2.33 Income stocks

Income stocks are purchased for income. Cyclical company is one whose earnings tend to fluctuate sharply with changes in the business cycle and such company makes of products for which the demand curve is flexible. A cyclical company includes; Caterpillar, General Motors, etc. Non-Cyclical company makes of products for which the demand curve is inflexible. A non-cyclical company includes; Coca-Cola, Proctor & Gamble (PG), etc.

2.34 Speculative stocks

Speculative stocks are those stocks, which may have Price Earnings Ratios which are at multiples of 50 to 100. New issue is a speculative stock.

2.35 Defensive stocks

Defensive stocks are those, which are stable and relatively safe in declining markets. Stocks which have this characteristic are food companies, drug manufacturers and utilities. They tend to decline less in recessions. And their products are necessary in any economic climate.

2.36 Dividends

The profit of the company distributed among the share holders is called dividend. Each share holder gets dividend proportionate to the face value of the shares held. Dividend is usually expressed as a percentage.

2.37 Financial market

Financial market is a mechanism that allows people to easily buy and sell financial securities (stocks and bonds), commodities (precious metals and agricultural goods) and other fungible items of value at low transaction costs. Financial markets can be domestic or international. The financial markets can be divided into different sub types namely; capital markets (stock markets, bond markets and commodity markets), money markets, derivatives markets, insurance markets, foreign exchange markets. Financial markets facilitates: raising of capital in the capital markets, transfer of risk in the derivatives markets, international trade in the currency markets, and match those who want capital to those who have it.

2.38 Financial Ratio analysis

Financial ratio analysis uses a company's financial information to predict its future projections of earnings. It is a tool that can assist the investor in the selection of stocks. From financial ratio analysis, an investor is able to assess the present and fast financial strength of a company. There are many ratios that can be used, namely; profitability ratios, liquidity ratios, activity ratios, debt ratios, market ratios and coverage ratios.

3. CONCLUSION

Competition requires timely and sophisticated analysis on an integrated view of the data. However, there is a growing gap between more powerful storage and retrieval systems A new technological leap is needed to structure and prioritize information for specific end-user problems. The data mining tools can make this leap. Quantifiable business benefits have been proven through the integration of data mining with current and future information systems.

4. REFERENCES

1. Calderon T.G., and Cheh J.J., (2002), 'A roadmap for future neural networks research in auditing and risk assessment', International Journal of Accounting Information Systems, Vol. 3, No. 4, pp. 203-236.
2. Das, S. and Chen, M. 2001. Yahoo for Amazon: Extracting Market Sentiment from Stock Message Boards, Proceedings of the 8th Asia Pacific Finance Association Annual Conference (APFA 2001), Bangkok, Thailand, July 22-25, 2001
3. Fanning K. and Cogger K., (1998), 'Neural Network Detection of Management Fraud Using Published Financial Data', International Journal of Intelligent Systems in Account-ing, Finance & Management, Vol. 7, No. 1, pp. 21-24.
4. R. Gencay, "Non-linear prediction of security returns with moving average rules," Journal of Forecasting, vol. 15, no.3, pp. 165-174, 1996
5. World Federation of Exchanges. "2007 Annual Report and Market Statistics". Retrieved on 2008-09-30.