Dr.G.Manoj Someswar, B. Satheesh, G.Vivekanand / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 4, June-July 2012, pp.717-723

Finance Mining – Analysis Of Stock Market Exchange For Foreign Using Classification Techniques.

Dr.G.Manoj Someswar, B. Satheesh, G.Vivekanand

Abstract –

Data Mining is a Business intelligence technique, which the finance model is considered to hardest way to make easy money. Mining certainly motivated by the prospect of discovering a financial issue such as buying/selling for stock/contract analysis. However the survey designs a successful model poses many challenges includes securing and cleaning data acquiring a sufficient amount of financial domain knowledge bounding the complexity. This work describes the modern finance finding efficient way to summarize and visualize the stock market data to give individuals useful information about the behavior for investment market decision and the data mining classification for stock markets. Our proposed analysis reveals progressive applications in addition to existing analysis trading technique using data mining or genetic algorithms for exchange of foreign rates.

Keywords – Finance, Data Mining, Classification, Stock Market, Trading.

INTRODUCTIONII

Knowledge mining is increasing synonymous to wealth creation and as a strategy plan for competing in the market place can be no better than the [1] information on which it is based, the importance of information in today's business can never be seen as an exogenous factor to the business. Organizations and individuals having access to the right information at the right moment, have greater chances of being successful in the epoch of globalization and cut-throat competition. Currently, huge electronic data repositories are being maintained by banks and other financial institutions across the globe. Valuable bits of information are embedded in these data repositories. The huge size of these data sources make it impossible for a human analyst to come up with interesting information that will help in the decision making process. A number of commercial enterprises have been quick to recognize the value of this concept, as a consequence of which the software [14] market itself for data mining is expected to be in excess of 10 billion USD. Business Intelligence focuses on discovering knowledge from various electronic data repositories, both internal and external, to support better decision making. Data mining techniques become important for this knowledge discovery from databases. In recent years, business intelligence systems have played pivotal roles in helping organizations to fine tune business goals such as improving customer retention, market penetration, profitability and efficiency. In most cases, these insights are driven by analyses of historical data.

Dynamic markets, and rapidly decreasing cycles of technological innovation provide important challenges for the banking and finance industry.

Worldwide just-in-time availability of information allows enterprises to improve their flexibility. In financial institutions considerable developments in information technology have led to huge demand for continuous analysis of resulting data. Data mining can contribute to [15] solving business problems in banking and finance by finding patterns, causalities, and correlations in business information and market prices that are not immediately apparent to managers because the volume data is too large or is generated too quickly to screen by experts.

Almost every computational method has been explored and used for financial modeling. We will name just a few recent studies. Monte-Carlo simulation of option pricing, finite-difference approach to interest rate derivatives and fast Fourier transform for derivative pricing New developments augment traditional technical analysis of stock market curves (Murphy, 1999) that has been used extensively by financial institutions. Such stock charting helps to identify buy/sell signals (timing "flags") using graphical patterns. Data mining as a process of discovering useful patterns, correlations has its own niche in financial modeling. Similarly to other computational methods almost every data mining method and technique has been used in financial modeling. An incomplete list includes a variety of linear and non-linear models, multi-layer neural

networks, k-means and hierarchical clustering; k-nearest neighbors, decision tree analysis, regression (logistic regression; general multiple regression), ARIMA, principal component analysis, and Bayesian learning. Less traditional methods used include rough sets, relational data mining methods (deterministic inductive logic programming and newer probabilistic methods, support vector machine, independent component analysis, Markov models and hidden Markov models.

SECTION II

2.1. Finance with Data Mining

The current efficient market theory hypothesis attempt to discover long-term trading rules regularities with significant profit, Greenstone and Oyer (2000) examine the month by month measures of return for the computer software and computer systems stock indexes to determine whether these indexes' price movements reflect genuine deviations from random chance using the standard t-test. They concluded that although Wall Street analysts recommended to use the "summer swoon" rule (sell computer stocks in May and buy them at the end of summer) this rule is not statistically significant. However they are able to confirm several previously known 'calendar effects" such as "January effect" noting meanwhile that they are not the first to warn of the dangers of easy data mining and unjustified claims of market inefficiency. The market efficiency theory does not exclude that hidden short-term local. Conditional regularities may exist. These regularities can not work "forever," they should be corrected frequently. It has been shown that the financial data are not random and that the efficient market hypothesis is merely a subset of a larger chaotic market hypothesis (Drake and Kim, 1997). This hypothesis does not exclude successful short term forecasting models for prediction of chaotic time series (Casdagli and Eubank, 1992). Data mining does not try to accept or reject the efficient market theory.

Data mining creates *tools* which can be useful for discovering subtle short-term conditional patterns and trends in wide range of financial data. This means that retraining should be a permanent part of data mining in finance and any claim that a silver bullet trading has been found should be treated similarly to claims that a perpetuum mobile has been discovered. The impact of market players on market regularities stimulated a surge of attempts to use ideas of statistical physics in finance (Bouchaud and Potters, 2000). If an observer is a large marketplace player then such observer can potentially change regularities of the marketplace dynamically. Attempts to forecast in such dynamic environment with thousands active agents leads to much more complex models than traditional data mining models designed for. This is one of the major reasons that such interactions are modeled using ideas from statistical physics rather than from statistical data mining.

2.2. Application of Finance data mining: Bank industry data mining is heavily used to model and predict credit fraud to evaluate risk to perform trend analysis profitability. In the financial markets neural networks have been used in stock price forecasting in option trading inbound rating in portfolio management in commodity price prediction mergers and acquisitions as well as in forecasting financial applications.

2.2.1 Forecasting the Stock: Several software applications on the market use data mining methods for stock analysis such an application used for stock prediction.

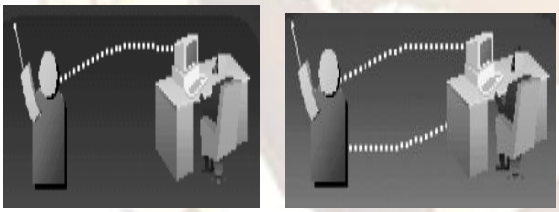
Screen 1 Shows the Stock Market.

NETPROPHET by neural networks corporation is a stock prediction application that makes use of neural networks, the most widespread use of data mining in banking is in the area of fraud detection. HNC is the credit fraud detection to monitor 160 million payment card accounts in a year. They also claim a healthy return on investment. While fraud is decreasing applications for payment card accounts are rising as much as 50 % a year. Widespread use of data mining in banking has not been unnoticed. In 1996 bank[3] systems & technology commented that most important data mining application is financial services. Finding banking companies who use data mining is not easy, given their proclivity for silence. The following list of financial companies that use data

mining required some digging into SEC reports from data mining vendors that are made available to the public. The list includes: Bank of America, First USA Bank, Headlands Mortgage Company, FCC National Bank, Federal Home Loan Mortgage Corporation, Wells Fargo Bank, Nations- Banc Services, Mellon Bank N.A., Advanta Mortgage Corporation, Chemical Bank, Chevy Chase Bank, U.S. Bancorp, and USAA Federal Savings Bank. Again it is reasonable to assume that most large banks are performing some sort of data mining.

2.2.2. Finance Halifax Bank in Real Time: Halifax PLC is second largest bank in London has chosen its right point real marketing suite as the foundation for a customer relationship initiative right point will enable Halifax customer service representatives to mobilize vital information about a customer and determine which campaigns product or services to offer at the point of customer.

Halifax direct customer service center receives more than 20 million customer calls per year and employs 800 customer service representatives. With the call center increasingly becoming the customer interaction center a customer decision to do business with a company is often on whether a company is aware of that customer's preferences and acts upon them accordingly. Using RightPoint, Halifax representatives will have a valuable tool for reliably predicting and delivering on the requirements of their customers in real time



Call with problem Agent Solution
Call with problem Agent Solution

Figure 1 describes before real time marketing call with 20M +Calls is equal to 20M+Missed Opportunities.

Figure 2 describes after real time marketing call with 20M+calls annually is equal to significant increase in customer wallet share.

Direct channel is playing an ever-increasing role in delivering customer contact, with call volumes predicted to grow to more than 50 million calls per year over the next three years," says Dick [4] Spelman, director of distribution at Halifax. "We need to ensure that we can harness customer data at the point of contact so that a customized service is delivered in real-time. This is what the RightPoint solution will deliver to our agents. The other parallel challenge that call centers face is generating sales income. Rather than add more agents and use questionable handover techniques, RightPoint offers us the potential to convert inbound service calls into profitable sales. If organizations don't tackle the revenue-generation aspect of their call-center activities, they will not be able to [5] afford the current unbridled growth in service traffic.

SECTION III

3. Problem Definition: Data mining is a business and knowledge sharing application, which we can apply for shopping, stock market, sales, compares the loss & gain results using operations and transactional data. Financial data produce huge data sets that build a foundation for approaching these enormously complex and dynamic problems with data mining tools. Potential significant benefits of solving these problems motivated extensive research for years. Our paper analyzes the overview on how to apply data mining on finance, its benefits more efficient.

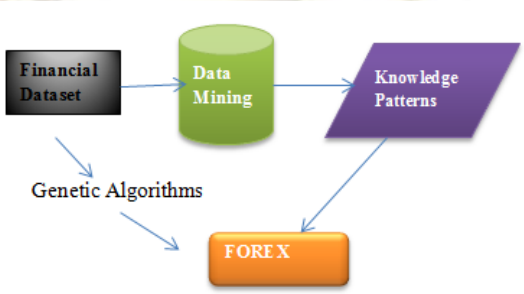


Figure 3 depicts the proposed analysis for Foreign Exchange using Data Mining Techniques.

3.1. Mining Stock Market Techniques: Forecasting stock market bank bankruptcies understanding and managing financial risk trading futures credit rating loan management bank customer profiling and money laundering analyses are core financial tasks for data mining. Stock market forecasting includes uncovering market trends planning investment strategies identifying the best time to purchase the stocks and what stocks to purchase.

3.1.1. Decision Tree on Stock Market: In a stock market, how to find right stocks and right timing to buy has been of great interest to investors. To achieve this objective, Muh-Cherng et al. (2006) present a stock trading method by combining the filter rule and the decision tree technique. The filter rule, having been widely used by investors, is used to generate candidate trading points. These points are

subsequently clustered and screened by the application of a decision tree algorithm. Chi-Lin Lu and Ta-Cheng Chen have employed decision tree-based mining techniques to explore the classification rules of information transparency levels of the listed firms in Taiwan's stock market. The main purpose of their study is to explore the hidden knowledge of information disclosure status among the listed companies in Taiwan's stock market. Moreover, the multilearner model constructed with decision tree algorithm has been applied. The numerical results have shown that the classification accuracy has been improved by using multi-leaner model in terms of less Type I and Type II errors. In particular, the extracted rules from the data mining approach can be developed as a computer model for the prediction or classification of good/poor information disclosure potential. By using the decision tree-based rule mining approach, the significant factors with the corresponding equality/ inequality and threshold values were decided simultaneously, so as to generate the decision rules. Unlike many mining approaches applying neural networks related approaches in the literature; the decision tree approach is able to provide the explicit Classification rules. Moreover, a multi-learner model constructed by boosting ensemble approach with decision tree algorithm has been used to enhance the accuracy rate in this work. Based on the extracted rules, a prediction model has then been built to discriminate good information disclosure data from the poor information disclosure data with great precision. Moreover, the results of the experiment have shown that the classification model obtained by the multi-learner method has higher accuracy than those by a single decision tree model. Also, the multi-learner model has less Type I and Type II errors. It indicates that the multilearner model is appropriate to elicit and represent experts' decision rules, and thus it has provided effective decision supports for judging the information disclosure problems in Taiwan's stock market. By using the rule based decision models, investors and the public can accurately evaluate the corporate governance status in time to earn more profits from their investment. It has a great meaning to the investors, because only prompt information can help investors in correct investment decisions (Jie and Hui, 2008).

3.1.2. Neural Networks on Stock Market: Advantage of neural networks is that they can approximate any nonlinear function to an arbitrary degree of accuracy with a suitable number of hidden units (Hornik et al., 1989). The development of powerful communication and trading facilities has enlarged the scope of selection for investors. David Enke and Suraphan Thawornwong introduced an information gain technique used in machine learning for data mining to evaluate the predictive relationships of numerous financial and economic variables. Neural network models for level estimation and classification are then examined for their ability to provide an effective forecast of future values. A cross validation technique is also employed to improve the generalization ability of several models. The results show that the trading strategies guided by the classification models generate higher risk-adjusted profits than the buy-and-hold strategy, as well as those guided by the level-estimation based forecasts of the neural network and linear regression models (David and Suraphan, 2005). Defu et al. (2007) dealt with the application of a well-known neural network technique, multilayer back propagation (BP) neural network, in financial data mining. A modified neural network forecasting model is presented, and an intelligent mining system is developed. The system can forecast the buying and selling signs according to the prediction of future trends to stock market, and provide decision-making for stock investors. The simulation result of seven years to Shanghai composite index shows that the return achieved by this mining system is about three times as large as that achieved by the buy-and-hold strategy, so it is advantageous to apply neural networks to forecast financial time series, so that the different investors could benefit from it (Defu et al., 2004). Accurate volatility forecasting is the core task in the risk management in which various portfolios' pricing, hedging, and option strategies are exercised. Tae (2007) proposes hybrid models with neural network and time series models for forecasting the volatility of stock price index in two viewpoints: deviation and direction. It demonstrates the utility of the hybrid model for volatility forecasting.

3.1.3. Clustering on Stock Market: Basaltoa et al. (2005) apply a pair wise clustering approach to the analysis of the Dow Jones index companies, in order to identify similar temporal behavior of the traded stock prices. The objective of this attention is to understand the underlying dynamics which rules the companies' stock prices. In particular, it would be useful to find, inside a given stock market index, groups of companies sharing a similar temporal behavior. To this purpose, a clustering approach to the problem may represent a good strategy. To this end, the chaotic map clustering algorithm is used, where a map is associated to each company and the correlation coefficients of the financial time series to the coupling strengths between maps. The simulation of a chaotic map dynamics gives rise to a natural partition of the data, as companies belonging to the same industrial

branch are often grouped together. The identification of clusters of companies of a given stock market index can be[7][8] exploited in the portfolio optimization strategies (Basaltoa et al., 2005). Graph representation of the stock market data and interpretation of the properties of this graph gives a new insight into the internal structure of the stock market. Vladimar et al. (2006) study different characteristics of the market graph and their evolution over time and came to several interesting conclusions based on their analysis. It turns out that the power-law structure of the market graph is quite stable over the considered time intervals; therefore one can say that the concept of self-organized networks, which was mentioned above, is applicable in finance, and in this sense the stock market can be considered as a "self-organized" system. Another important result is the fact that the edge density of the market graph, as well as the maximum clique size, steadily increases during the last several years, which supports the well-known idea about the globalization of economy
which has been widely discussed recently. They also indicate the natural way of dividing the set of financial instruments into groups of similar objects (clustering) by computing a clique partition of the market graph.

3.1.4. Association Rule on Stock Market: Agrawal et al. (1993) discovering association rules is an important data mining problem, and there has been considerable research on using association rules in the field of data mining problems. The associations' rules algorithm is used mainly to determine the relationships between items or features that [9] occur synchronously in the database. For instance, if people who buy item X also buy item Y, there is a relationship between item X and item Y, and this information is useful for decision makers. Therefore, the main purpose of implementing the association rules algorithm is to find synchronous relationships by analyzing the random data and to use these relationships as a reference during decision making (Agrawal et al., 1993). One of the most important problems in modern finance is finding efficient ways to summarize and visualize the stock market data to give individuals or institutions useful Hajizadeh et al. 115 information about the market behavior for investment decisions. The enormous amount of valuable data generated by the stock market has attracted researchers to explore this problem domain using different methodologies. Shu-Hsien et al. (2008) investigated stock market investment issues on Taiwan stock market using a two stage data mining approach. The first stage apriori algorithm is a methodology of association rules, which is implemented to mine knowledge and illustrate knowledge patterns and rules in order to propose stock category association and possible stock category investment collections. Then the K-means algorithm is a methodology of cluster analysis implemented to explore the stock cluster in order to mine stock category clusters for investment information. By doing so, they propose several possible Taiwan stock market portfolio alternatives under different circumstances (Shu-Hsien et al., 2008).

3.2. How to avoid Risk by Applying Data mining in Finance: Managing of risk is at the core of every financial organization major challenge in the banking and insurance world is therefore the implementation of risk systems in order to identify measure and control business exposure. Integrated measurement of different kinds of risk is moving into focus. These all are based on models representing single financial instruments or risk factors their behavior and their interaction.

3.2.1. Financial Risk: For single financial instruments that is stock indices interest rates or currencies market risk measurements is based on models [10] depending on a set of underlying risk factor such as interest rates stock indices or economic development. Today different market risk measurement approaches exist. All of them rely on models representing single instrument their behavior and interaction with overall market.

SECTION IV

4.1. Technical Application of Data Mining in Finance: Several ways for deciding when to buy/sell a stock/contract, the process of financial data mining synthesizes technical analysis with machine learning for constructing successful models. A model is defined as a series of trades, which identifies a buy and sell date and time, the stock/contract process upon entering a trade and a number of shares. Technical analysis involves constructing one or more mathematical models based upon the stock/contract movement or change in Volume. One of the keys to effective data mining is acquisition of specific domain knowledge. Possessing domain knowledge speeds up the data mining process because it allows the modeler to discriminate between multitudes of strategies that may be available. Furthermore, possessing domain knowledge helps in recognizing a "good" model. In this case, the technical indicators serve as the domain-specific knowledge. A technical indicator is an algorithm constructed using price or volume parameters. There are more than 100 technical indicators available. Common examples include moving averages, relative strength index, or commodity channel index. One of the challenges in the model formulation process is to focus on only a

relatively few technical indicators which appear to be most promising. Models may be constructed solely with technical indicators. However, students are encouraged to synthesize technical indicators with various machine learners.

4.2. Financial Trading: Trading techniques are going long means buying a stock contract at a particular price with the intent that the stock contract will increase in price. When short a stock contract is initially sold with the intent that the stock contract will decrease in price. Example a stock sold for $10 then later bought at $8 would result in a $2 profit thus it is possible to make money in either market direction. When an investor initially buys sells a stock contract the price may move in the opposite direction resulting in a losing position. An investor must decide how large a loss he or she is willing to tolerate or whether to risk the situation will turn in their favor amount of tolerance for loss is called drawdown for this percentage may be 10% for futures it may increase or fixed points.

The profit for a stock trade is equal to the purchase price P minus the selling price S times the number of shares N purchased.

Profit (Long Trade) = (P-S)*N

Futures contracts operates as follows for a long position an investor may buy one EMini S &P goes up the price of the contract for $2,000 for each point the S & P goes up the price of the contract increases by 50 dollars. If a contract is bought at $ 950 sold at $ 960 then the investor nets 10 points times $ 50 for a profit of 500 dollars on an initial investment of $ 2,000 or 25% increase in just one transaction.

CONCLUSION V

Increase of economic globalization information technology financial data are being generated and accumulated at an unprecedented pace. A critical need for automated approaches to effective and efficient utilization of massive amount of financial data to support organizations in strategic investment decisions. Data mining techniques have been used to uncover hidden patterns and predict future trends and behaviors in financial markets. The competitive advantages achieved by data mining include increased revenue, reduced cost, and much improved marketplace responsiveness and awareness. There has been a large body of research and practice focusing on exploring data mining techniques to solve financial problems. To be successful, a data mining project should be driven by the application needs and results should be tested quickly. Financial applications provide a unique environment where efficiency of the methods can be tested instantly, not only by using traditional training and testing data but making real stock forecast and testing it the same day. This process can be repeated daily for several months collecting quality estimates. This paper proposes a data mining in finance stock market and specific requirements for data mining methods including in making interpretations, incorporating relations and probabilistic learning for exchange.

Reference

[1] Agrawal R, Imilienski T, Swami A (1993). Mining association rules between sets of items in large databases, In Proceedings of the ACM SIGMOD international conference on management of data.

[2] Basaltoa N, Bellottib R, De Carlob F, Facchib P, Pascazio S (2005).

[3] Clustering stock market companies via chaotic map synchronization, Physica A. Berry MJA, Linoff GS (2000). Mastering data mining, New York: Wiley.

[4] Boris K, Evgenii V (2005). Data Mining for Financial Applications, the Data Mining and Knowledge Discovery Handbook.

[5] Chi-Lin L, Ta-Cheng C (2009). A study of applying data mining approach to the information disclosure for Taiwan's stock market investors, Expert Systems with Applications,. Cowan A (2002). Book review: Data Mining in Finance, Int. J. forecasting.

[6] David E, Suraphan T (2005). The use of data mining and neural networks for forecasting stock market returns, Expert Systems with Applications.

[7] Defu Z, Qingshan J, Xin L (2004). Application of Neural Networks in Financial Data Mining, Proceedings of world academy of science, Eng. Technol.

[8] Hornik K (1989). Stinchcombe M. and White H., "Multilayer feed forward networks are universal approximators", Neural Networks.

[9] Hsiao-Fan W, Ching-Yi K (2004). Factor Analysis in Data Mining, Computers and Mathematics with Applications. http://en.wikipedia.org/wiki/Stock_market http://www.resample.com/xlminer/help/Assocrules/associationrules_intr o.htm.

[10] Huarng K, Yu HK (2005). A type 2 fuzzy time series model for stock index forecasting, Physica.

[11] Jar-Long W, Shu-Hui C (2006). Stock market trading rule discovery using two-layer bias decision tree, Expert Systems with Applications.

[12] Jie S, Hui L (2008). Data mining method for listed companies' financial distress prediction, Knowledge-Based Systems.

[13] Teaching Financial Data Mining using Stocks and Futures Contracts **Gary D.** BOETTICHER Department of Computer Science, University of Houston – Clear Lake Houston, Texas 77058, USA

[14] Petra Hunziker, Andreas Maier, Alex Nippe, Markus Tresch, Douglas Weers, and Peter Zemp, Data Mining at a major bank: Lessons from a large marketing application retrieved 5th January, 2006 from http://homepage.sunrise.ch/homepage/pzemp/info/pkdd98.pdf

[15] Michal Meltzer, Using *Data Mining on the road to be successful part III*, published in October 2004, retrieved 2nd January, 2006 from http://www.dmreview.com/editorial/newsletter_article.cfm?nl=bireport&articleId=1011392&issue=20082

[16] Fuchs, Gabriel and Zwahlen, Martin, *What's so special about insurance anyway?*, published in DM Review Magazine, August 2003 issue, retrieved 5th January, 2006 from http://www.dmreview.com/article_sub.cfm?articleId=7157

Dr.G.Manoj Someswar, B.Tech., M.S.(USA), M.C.A., Ph.D. is having 20+ years of relevant work experience in Academics, Teaching, Industry, Research and Software Development. At present, he is working as an Associate Professor, Dept. of IT, MVSR Engineering College, Nadergul, Hyderabad, A.P. and utilizing his teaching skills, knowledge, experience and talent to achieve the goals and objectives of the Engineering College in the fullest perspective. He has attended more than 100 national and international conferences, seminars and workshops. He has more than 10 publications to his credit both in national and international journals. He is also having to his credit more than 50 research articles and paper presentations which are accepted in national and international conference proceedings both in India and Abroad.

B. Satheesh Ph.D M.Tech Current working Jyothismathi Inst of Tech & Scie Karimnagar as an Asst Prof in the Dept of CSE. His areas of Web Technologies Computer Networks, Attended a work shop on TESTING-TOOLS conducted by SREE Dattha Institutions, HYD in the year 2008, work shop on DESIGN PATTERNS conducted by SVAROG SOFTWARE SOL., HYD in the year 2011and Conferences : Attended-SUNTECH-DAYS FEB-2009 on JAVA-FX, JSF, HITEX, HYDERABAD

G.Vivekanand pursuing Ph.D M.Tech Current working Jyothismathi Inst of Tech & Scie Karimnagar as an Asst Prof in the Dept of CSE. His areas of Web Technologies, Data Mining, Attended a work shop on TESTING-TOOLS conducted by SREE Dattha Institutions, HYD in the year 2008, work shop on DESIGN PATTERNS conducted by SVAROG SOFTWARE SOL., HYD in the year 2011and Conferences: Attended-SUNTECH-DAYS FEB-2009 on JAVA-FX, JSF, HITEX, HYDERABAD