Vol 5. No. 4, June 2012 ISSN 2006-1781

African Journal of Computing & ICT

www.ajocict.net

Forecasting Portfolio Investment Using Data Mining

[1]Adebimpe, L.A
Dept of Computer Science
Emmanuel Alayande College of Education
Oyo, Nigeria
dradebimpela@yahoo.com

Adedara, Olusola
Department of Computer Science
The Federal Polytechnic
Ado Ekiti, Nigeria
fafvfk@yahoo.com

Longe, O.B (PhD)
Fulbright Fellow & Scholar-in-Residence
Southern University System
Baton Rouge, Louisiana, USA.
longeolumide@fulbrightmail.org

[1] *Corresponding Author*

ABSTRACT

Following the exponential and unprecedented rate at which commercial databases are growing, it becomes imperative to mine the data to provide guidance to investors for selecting portfolios which will afford them good income and enable them achieve appreciable capital gain. The ability to manage, manipulate, store and interpret enormous data available will be the focus of this paper. A good number of statistical tools will be applied on mined data to arrive at a robust and proper decision for investment in different portfolios.

Keywords: Portfolio; Finance; Investment; Securities.

1. INTRODUCTION

In the post war period, finance for development has overwhelmingly taken the form of loan capital. Non debt creating capital flows have been modest in size. Reliance on loan capital has resulted in an accumulation of debt which now constrains economic management in a wide range of developing countries - both low income as well as middle income developing countries. This debt has also become a major problem for lending institutions in the developed countries as well as a source of friction between North and South. Accordingly, there is renewed attention being given to non-debt creating capital flows - portfolio investment (WIDER, 1990). A portfolio investment is a passive investment in securities. Jean et al [5] defines investment as the sacrifice of current consumption in order to obtain increased consumption at a later date. From this perspective, an investment is undertaken with the expectation that it will lead, ultimately, to a preferred pattern of consumption for the investor.

African Journal of Computing & ICT Reference Format:

L.A. Adebimpe,, O. Adedara, & O.B. Longe (2012). Forecasting Portfolio Investment Using Data Mining. Afr J. of Comp & ICTs. Vol 5, No. 4. pp 101-106.

A fundamental step in portfolio investment is the purchase and sale of securities. Security according to Jean et al [5] can be defined as "a legal contract representing the right to receive future benefits under a stated set of conditions." The piece of paper (e.g. the share certificate or the bond) defining the property rights is the physical form of the security. The return on a security is the fundamental reason for wishing to hold it. The return is determined by the payments made during the lifetime of the security plus the increase in the security's value. The importance of risk comes from the fact that the return on most securities (if not all) is not known with certainty when the security is purchased. This is because the future value of security is unknown and its flow of payments may not be certain. It is a fundamental assumption of investment analysis that investors wish to have more return but do not like risk [1]

The risk inherent in holding a security has been described as a measure of the size of the variability, or the uncertainty, of its return. Several factors can be isolated as affecting the riskiness of a security and these include Maturity, Creditworthiness, Priority, Liquidity and Underlying economic activities of the issuer. Before investment can be undertaken, investors are often confronted with a number of issues that may put their wealth at risk if not properly addressed. These include the type of securities to be purchased, determination of the precise quantities, returns and risks of each of the chosen securities [2][3]

ISSN 2006-1781

In making these decisions, investors will be guided by examining available dataset on those securities. This is very essential for predictable trends. As observed by Larrain [7] not all of those decisions are taken by the person herself as they are delegated to asset managers. The partial privatization of social security has resulted in an increased delegation of portfolio choices in the shoulders of asset managers. These latter ones find their task more complex due to two newly appeared elements: the massive character of this process, which imply the management of significantly more specific investment strategies, and the fact that the financial literacy of those new clients is reduced, implying that there are more asymmetries of information and simpler informational requirements. In any event, understanding the choices made by investors will shed light on the important factors explaining the pricing of risk in financial markets. Focusing on the accuracy and timing of those decisions, this paper uses data mining approach in managing portfolio investment.

2. RELATED WORKS

Companies collect terabytes upon terabytes of information every day - anything from transactional data, to demographics, to product sales based on seasons. But what do they do with it all once it is neatly organized into a database? The concept of data mining is just as it sounds. This information can be more valuable than mining for gold, because the results are almost a guarantee. The data mining process used to be a highly technical process requiring mathematicians to build the analysis for companies. But today's data mining technology offers the tools they need to make sense of their customer data and apply it to business[11].

The term "data mining" refers to new methods for the intelligent analysis of large data sets. These methods have emerged from several historically disjoint fields, such as applied statistics, information systems, machine learning, data engineering, artificial intelligence, and knowledge discovery. One of the most enticing application areas of these emerging technologies is finance, becoming more amenable to data-driven modeling as large sets of financial data become available [1]. Financial institutions produce huge datasets that build a foundation for approaching these enormously complex and dynamic problems with data mining tools [4]. There has been a critical need for automated approaches to effective and efficient utilization of massive amount of financial data to support companies and individuals in strategic planning and investment for decision making [10]. Forecasting stock market, currency exchange rate, 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 [9].

There is currently a surge of interest in financial markets data mining. Large amount of historical data is available for this domain in machine readable form. Analyses of this data for the purpose of abstracting and understanding market behavior, and using the abstractions for making predictions about future market movements, is being seriously explored (AI on Wall Street 93).

Data mining in finance typically follows a set of general for any data mining task steps such as problem understanding, data collection and refining, building a model, model evaluation and deployment [12][9][6]. Data mining approach covers empirical models and regularities derived directly from data and almost only from data with little domain knowledge explicitly involved. Historically, in many domains, deep field-specific theories emerge after the field accumulates enough empirical regularities. We see that the future of data mining in finance would be to generate more empirical regularities and combine them with domain knowledge via generic analytical data mining approach [8].

Some firms have also deployed analytical data mining methods for actual investment portfolio management (Barr and Mani 93). As pointed out by Zafar et al [12][13], literature to provide some guideline of investment portfolio is very limited. Volumes of data are therefore being generated and stored in databases on regular basis. This data could be mined to provide guidance to investors for selecting optimum mix of portfolios. This activity is being performed at firm level and could be expanded to establish a national database for use by investors. Present research was focused on using data mining and decision analysis methodologies to extract meaningful patterns from business data. A number of statistical tools were applied on mined data to arrive at decisions for investment in various portfolios. The same methodology could be used to analyze the national database for providing useful information for investments in any country[11].

Rashid and Mohd [10] observed that many statistical and data mining techniques have been used to predict time series stock market. However, most statistical and data mining methods suffer from serious drawback due to requiring long training times, results are often hard to understand, and producing inaccurate prediction.

Boris [4] claimed that Data mining in finance has the same challenge as general data mining in data selection for building models. In finance, this question is tightly connected to the selection of the target variable. There are several options for target variable y: $y=T(k+1)$, $y=T(k+2)$,: : : ,$y=T(k+n)$, where $y=T(k+1)$ represents forecast for the next time moment, and $y=T(k+n)$ represents forecast for $n$moments ahead. Selection of dataset $\boldsymbol{T}$ and its size for a specific desired forecast horizon n is a significant challenge.

3. METHODOLOGY

Several parameters characterize data mining methodologies for financial forecasting. Data categories and mathematical algorithms are most important among them. This paper uses decision tree technique. As pointed out by Kovalerchuk and Vityaev [6], decision tree methods are very popular in data mining applications in general and in finance specifically. They provide a set of human readable, consistent rules, but discovering small trees for complex problems can be a significant challenge in finance. Rules can be extracted from decision trees.

ISSN 2006-1781
African Journal of Computing & ICT

www.ajocict.net

The data used for this paper was obtained from six organizations whose investing risk was assessed. This data was originally meant for business purposes. Thus researchers were required to transform the data for mining.

Table 1: Dataset analysis

Earning growth rate (EGR)	Positive
Earnings Per Share (EPS)	>80%
Return on Equity (ROE)	15% MINIMUM, PREFERRED 20%
Price Earning P/E ratio	10% MAXIMUM
Share Price	Undervalue

Table 2: Additional criteria for Ranking

Market Capitalization	Small cap
Inside Buy	At lest One
Institutional Ownership	Between 20% and 50%
Profitability	3 years above
Executive Performance Measure	Positive

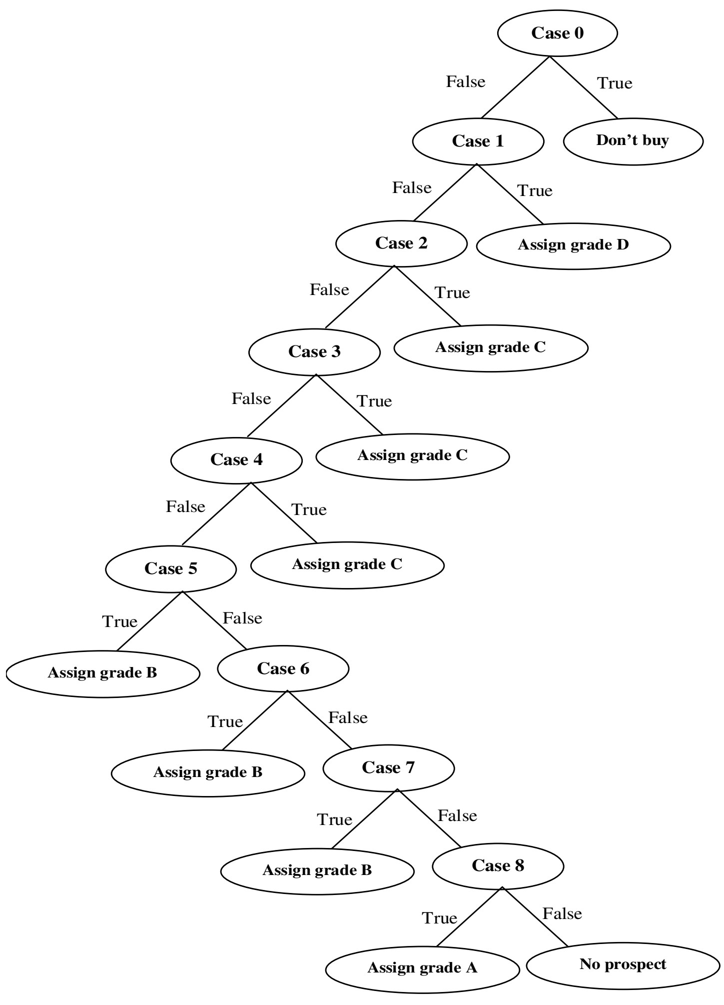

Fig. 1: Use Case Diagram

ISSN 2006-1781

Table 3: Decision Table for the Proposed System

	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
PER	0	1	0	1	1	1	1	1	1	1	1
ROE				0	0	0	1	0	1	1	1
EPS				0	0	1	0	1	0	1	1
EGR				0	1	0	0	1	1	0	1
CSP	0	0	1								
Investment strength = 60%				X							
Investment strength = 73.3%					X	X	X				
Investment strength = 86.7%								X	X	X	
Investment strength = 100%											X
Grade A											X
Grade B								X	X	X	
Grade C					X	X	X				
Grade D				X							
Comment A											X
Comment B1										X	
Comment B2									X		
Comment B3								X			
Comment C1							X				
Comment C2						X					
Comment C3					X						
Comment D				X							
No prospect	X	X	X								

Comment A – Best buy and good prospect

Comment B1 – Better investment, but risky due to company inefficiency which might lead to loss in investment.

Comment B2 – Better investment but risky due to lack of operational continuity

Comment B3 – Better investment but risky because stock price may drop in the nearest future.

Comment C1 – Good investment but risky because of company inefficiency and lack of operational continuity.

Comment C2 – Good investment but risky due to company inefficiency and possible drop in price.

Comment C3 – Good investment but risky due to lack of operational continuity and stock price may drop in the nearest future.

Comment D – Stock has prospect, but run the risk of drop in share price because of the company inefficiency.

ISSN 2006-1781

4. CONCLUDING AND FUTURE WORK

It is a truism that large amounts of business data is available for the purpose of abstracting and understanding market behavior and for making informed decisions. There is no doubt that this paper has provided a good approach for portfolio investment. It will no doubt provide a gateway to personal financial success. Finally, we believe the mechanism we suggest in this paper is more general, and will go a long way to guide investors. We expect that in coming years application of data mining in portfolio investment will be deepen.

REFERENCES

1. Andreas S. Weigend (1997)Data Mining in Finance: Intelligent Information Systems and Computer Intensive Methods for Financial Modeling and Data Analysis. Leonard N. Stern School of Business New York University

2. Artificial Intelligence Applications on Wall Street Proceedings. Software Engineering Press, 1993.

3. Barr D. and Mani G. (1993). Neural Nets in Investment Management: Multiple Uses. In Artificial Intelligence Applications on Wall Street Proceedings, pages 81-87, 1993.
4. Boris Kovalerchuk (2000). Data Mining for Financial Applications. Central Washington University, USA.

5. Jean Hindriks & Gareth D. Myles, 2003. "Strategic Inter-Regional Transfers," Journal of Public Economic Theory, Association for Public Economic Theory, vol. 5(2), pages 229-248, 04.Kl¨osgen W., Zytkow J. (2002) Handbook of data mining and knowledge discovery, Oxford Univ. Press, Oxford.

6. Kovalerchuk, B., Vityaev, E., Data Mining in Finance: Advances in Relational and Hybrid Methods, Kluwer, 2000.

7. Larrain Rios, G. (2007), "Portfolio Investment in an Intertemporal Setting: Assessment of the Literature and Policy Implications for Latin American Pension Systems", *OECD Working Papers on Insurance and Private Pensions*, No. 10, OECD Publishing.
8. Mitchell, T. (1997). Machine learning. McGraw Hill.

9. Nakhaeizadeh, G., Steurer, E., Bartmae, K., Banking and Finance, In: Kl¨osgen W., Zytkow J. (2002). Handbook of data mining and knowledge discovery, Oxford Univ. Press, Oxford.

10. Rashid H.K.and Mohd N.M.S., (2004). Classification with Degree of Importance of Attributes for Stock Market Data Mining. Jurnal Teknologi Maklumat Jilid 16, Bil. 2

11. Stephanie Roussel-Dupre (2001) Data Mining Provides Retail Understanding. Integrated Solutions for Retailers

12. World Institute for Development Economics Research of the United Nations University, March 1990.

13. Zafar S., Manarvi I., Ahmed J., Khan K. and Rehman K. (2009). Investment Portfolio Decisions for Shares through Data Mining. Computers and Industrial Engineering 2009 International Conference Pg 1458-1463.

Vol 5. No. 4, June 2012
African Journal of Computing & ICT

www.ajocict.net
ISSN 2006-1781

Authors' Briefs

Mr. ADEBIMPE, Lateef Adekunle is on faculty at the Department of Computer Science, Emmanuel Alayande College of Education, Oyo, Oyo State, Nigeria. His area of specialization is Network and Information Security and at present a doctoral candidate in Information Technology at AIT Accra Ghana and OUM, Malaysia to complement HND, B.Sc and M.Sc earlier obtained in Computer Science. He is an ardent believer of the maxim "there is never perfection in humanity". He can be reached by phone on +2348189521979 and through E-mail dradebimpela@yahoo.com

Mr. ADEDARA Olusola lectures at the Department of Computer Science, The Federal Polytechnic, Ado Ekiti, Nigeria. His holds the MSc Computer Science Degree from the University of Ibadan with specialization in data mining He can be reached through e-mail at fafvfk@yahoo.com. Phone on +2348088257765.

Dr. Longe Olumide is on Faculty at the Department of Computer Science, University of Ibadan, Nigeria. His research has focused on using social theories and computer security theories to explain causation and apprehension of cyber crimes. He is currently a Fulbright Fellow at the College of Business, Southern University, Baton Rouge, LA, USA. He can be reached by phone on +18572078409 and through E-mail longeolumide@fulbrightmail.org