New Approach of Bi-Clustering Learning for Trading Rules: An Overview

Yogesh Khambalkar, Prof. Rugraj

Dept. of Computer Engineering, Alard college of Engineering and Management, Pune, India

ABSTRACT: Historical financial entity prices are used to forecast the direction of future stock prices. Thedeveloped financial entity price prediction model uses a narrative two-layer reasoning approach thatemploys domain knowledge from scientific analysis in the first level of reasoning to direct a second layer of reasoning based on learning approaches. The proposed model is supplemented by a money managing strategy that uses the historical achievement of predictions made by thesculpt to take the action like buy/sell or no action using one's capital.

INTRODUCTION

Stock fundamentals can be measured bystock financial ratios. For example, the return on-equity ratio measures a stock's efficiencyin using its assets to generate profit,the debt–equityratio measures the amount ofthe stock's assets that are debts, and the priceearnings ratio measures the ratio of the stock'scurrent price to its current earnings.Therefore, scrutinizing financial ratios isimportant in finding undervalued stocks.However, there's no perfect rule that showswhich financial ratios and what values ofthese ratios are related to underpricedfinancial entities.For example, Benjamin Graham, the founderof value investment, prefers stocks with aprice–earnings ratio of no more than Using Graham's rules on picking stocks hasbeen demonstrated to createreturns for value seekers.An experiment conducted over aneight-year period from 1973 to 1980 showedthis strategy to be profitable.1 We propose using3D subspace clustering to generate rulesto pick potential undervalued stocks. The3D subspace-clustering method is effective inhandling high-dimensional financial data andis easily adaptable to new data set. Also, its resultsaren't influenced by human biases andemotions, and can be easily interpreted. Weconducted broadresearch in thestock market over a period of 28 years (from1980 to 2007), and we found that using guidelines generated by two 3D subspace-clustering algorithms (CAT Seeker and MIC)results in 60 percent more profits than using Graham's rules alone.

For an inexperienced investor, manuallysetting rules on the financial ratioscan be challenging, and even for the veteraninvestor, he or she might beprone to set irrational and prejudiced rules.The financier can stick to Graham's

rules, but the relevance of these rulesat present time remains to be seen.Hence, the belowissues needs tobe addressed: How do we find rules onfinancial ratios that are related to highstock price returns? We should notehere that we define the price returnof a stock as (sold price – boughtprice)/bought price.In financial domain we find some financialliterature that examinethe influence of single financialratios on financial entity prices. Different financial ratios compute differentcharacteristics of a stock, so to get thecomplete picture, it will be beneficial tostudy the combinedeffect of financialratios on the financial entity prices, and thisis a nontrivial issue.

II.BACKGROUND OF RESEARCH

Our goal with the Feature Generation module is to provide a knowledge-intensive andcomputationally efficient analysis of *past prices* which can be analyzedfurther in a second layer of reasoning. The field knowledge applied in the moduleis therefore limited to methods and techniques in technical analysis. The technical analysisliterature includes a wealth of different stock analysis techniques, some of which involvecomplicated and complex price patterns subjective in both discovery and analysis.These methods would be both computationally costly to identify and assess, and havesubsequently

been disregarded. We thus apply Occam's razor to the choice of proceduresin technical analysis, concentrating on the most prevalent indicators that can be efficientlyoperationalized and are intuitive in interpretation.

III.LITERATURE SURVEY

In this chapter we review overall literature of system. The [1] propose using 3D subspace clustering systems to mine guidelines that are related to high stock price returns. The 3D subspace-clustering approach groups stocks that have similar fundamentals (financial ratios) and high price returns across years. The highlighted region in Figure 1a is a 3D subspace cluster containing stocks $s2, s3, s4$ that have similar details reflected in financial entity ratios $R2, R3$ *and* $R4$ for years 1–3, 5–6, and 8–10. From Figure 1b, we can see that stocks $s2, s3, s4$ have high price returns.

Decision tree for stock trading is built based on technical analysis indicators. Indicators are set of mathematical formula that is calculated from stock prices data. These indicators are ordered to create certain trading rule, this rule become testing node for the decision tree. Based on the trading rule it can be calculated which price of stock would be suitable for a particular class. Those classes represent the decision that would be taken in the system The result will also be plotted to a chart for an easier trend analysis. Some systems are built based on financial market technical studysigns (Simple Moving Average, Moving Average Convergence Divergence, RSI, Money Flow Index and parabolic SAR) [2] method is arrangingindicators set into decision tree based on stock trading rules and classification which describe the rules and create buy, hold, and sell classes which represented conclusions ininvestment decisions. Final decision classes then are analyzed for their profitability,geometric mean yield, and accumulativecapital index.Moreover, sensitivity analysis is added into profitabilityanalysis to obtain more confident value trading/investing in judgement making.The study purpose is to enhance decision making in technical. stock trading.

System [3] describe an intelligent stock trading system by combining support vector machine (SVM) algorithm and box theory of financial entity. This theory believes a successful stock buying/selling generally occurs when the price effectively breaks out the previous oscillation box into one more new box. In the system, support vector machine algorithm is utilized to make predictions of the new top and bottom of the oscillation box. Then a trading strategy based on the box theory is built to make trading and investmentjudgements. The various stock movement patterns, i.e., bull, bear and fluctuant market are used to test the practicability of the method.

The resulting model is intended to be used as a decision support tool or as an autonomous artificial trader if extended with an interface to the stock exchange.Machine learning approach is also very effective for stock market analysis. System [4] define a stock price prediction model will be created using notions and methods in technical analysis and machine learning. The resulting prediction model should be engaged as simulatedagent that can be used to select stocks to trade on any given stock exchange.

In understanding the difficulties facing investors during the investment decision process; consider the case of common stocks in financial markets that produce on average significantly large return over the years than the saving account. However, a worthy range of investors avoid realizing these large returns due to the fact that "to pursue high returns investors must assume large risks. Model [5] formulated as a multi-criteria optimization model (maximizing the net profit and minimizing the maximum drawdown) to be solved for the contribution ratio of each trading decision model component in the trading decision pool. It has been demonstrated that the proposed strategy of combining different trading decision models results in noticeable increase of profit as well as considerable decrease in the maximum drawdown.

Previous research showed promising results on thepossibility of correctly predicting the price direction of astock or market index. We mention some of that workhere. [6] proposed 5-days-ahead and 10-days-ahead predictive models are built using the random forests algorithm. Theprototypes are constructed on the historic data of the CROBEX index and on a few companies listed at the Zagreb StockExchange from severalsegments. Numerous technical indicators, popular in quantitative analysis of stock markets, areselected as model inputs.

Some systems classify bi-clustering patterns into two, based on whetherthe pattern is defined on a single cluster or multiple clusters.If a bicluster pattern is defined on a single bicluster,we call the pattern a localized pattern. Else, we call thepattern a

globalized pattern. [7] describe address thisless studied, yet important problem and formally analyzeseveral bi-clustering algorithms in terms of the bicluster patternsthey attempt to discover. We methodicallyframethe requirements for familiar patterns and show the constraintsimposed by bi-clustering algorithms that definetheir capacity to identify such patterns. We also give experimentalresults from a carefully designed testbed to evaluatethe power of the employed search strategies.

The system [8] show the detail explanation of k-means classification. The [9] subsets of genes that have similar behavior under subsets of conditions, so we say that they express in tandem i.e. coexpress andactautonomously under other subsets of conditions. Discovering such coexpressions can be supportive to discover genomic facts such as gene networks or gene interactions. That is why, it is of utmost significance to make a concurrentgrouping of genes and conditions to identify clusters of genes that are coexpressedin clusters of situations. This type of grouping is called bi-clustering. It also focused on bi-clustering of genome expression dataset. The remaining of this article is structured as below - First, some definitions related to bi-clustering of microarray data. We then present in section 3 some evaluation functions and bi-clustering algorithms. Finally, we show how to validate biclusters via bi-clustering tools on microarrays datasets.

System [10] defineinnovativemethodology for customer division which is the elementary issue for an effective CRM (Customer Relationship Management). Initially, the chi-square statistical study is applied to choose set of attributes and K-means algorithm is employed to quantify the importance of each feature. Subsequently DBSCAN algorithm based on density is introduced to classify the customers into three clusters (First, Second and Third class). Finally,bi-clustering based on improved Apriori algorithm is used in these 3 groups to get more detailed information. Experimental results on the dataset of an airline company show that the bi-clustering wouldcategorise the customers more precisely and methodically.

IV.CONCLUSION

In the proposed survey we studied the different bi-clustering approaches as well clustering. We got the idea of proposed approach how to apply such scenarios on historical data like financial market or any benchmark data. All method outperforms four existing trading strategies, indicatingits sound performance for trading in financial markets.We expect that the novel idea of converting the bi-clusters intotrading rules will lead to a new research topic in finance theoryand applications.

V.FUTURE WORK

- Take the volume as one of the parameter as an input along with set of other technical indicators. Volume will act as important parameter to boost the confidence in taking overall decision.
- Calculate SMA-10 and SMA-30 (Simple Moving Average) and EMA-10, EMA-30(Exponential Moving Average) cross over status. This would be added as one more gating criteria to support the final decision.

REFERENCES

[1] Kelvin Sim and VivekanandGopalkrishnan proposed 3D Subspace Clustering for Value Investing by the IEEE Computer Society in 2014.

[2] F.X. Satriyo D. NugrohoTeguhBharataAdj, SilmiFauziat proposed the DECISION SUPPORT SYSTEM FOR STOCK TRADING USING MULTIPLE INDICATORS DECISION TREE in I st International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE) year 2014.

[3] Qinghua Wen, Zehong Yang, Yixu Song, PeifaJia proposed Intelligent Stock Trading Methodestablished on SVM Algorithm & Oscillation Box Prediction Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia, USA, June 14-19, 2009.

[4] Jan Ivar Larsen [4] proposed Predicting Stock Prices Using TechnicalAnalysis and Machine Learning Norwegian University of Science and TechnologyDepartment of Computer and Information Science year June 2010.

[5] Tarek M. Gadallah and M. NashatFors proposed A New Approach for Combining Multi-Criteria Trading Decision Models, Proceedings of the 2015 International Conference on Industrial Engineering and Operations ManagementDubai, United Arab Emirates (UAE), March 3 – 5, 2015.

[6] T. Manojlovic and I. Stajduhar proposed Predicting Stock Market Trends Using Random Forests Algorithm - A Example of the Zagreb Stock Exchange MIPRO 2015, 25-29 May 2015, Opatija, Croatia.

[7] DorukBozdag, Ashwin S. Kumar and Umit V. Catalyurek proposed Comparative Analysis of Bi-clustering Algorithms National Cancer Institute Grant R01CA141090; bythe U.S. DOE SciDAC Institute Grant DE-FC02-06ER2775;by the USA National Science Foundation inScholarships 1- CNS-064396 2-OCI-0904809, 3-OCI-0904802 and CNS-0403342.

[8] KardiTeknomoproposed K-Means Clustering methodologyTeknomo, Kardi. K-Means (k-NN) Clustering Lessons.URL - http:\\people.revoledu.com/kardi/ tutorial/kMean/July 2007.

[9] Haifa Ben Saber and MouradElloumi proposed A NEW SURVEY ON BI-CLUSTERING OFMICROARRAY DATA NatarajanMeghanathan et al. (Eds) : NeTCoM, CSIT, GRAPH-HOC, SPTM - 2014

[10] Xiaohui Hu and Haolan Zhang A New Customer Segmentation FrameworkBased on Bi-clustering AnalysisJOURNAL OF SOFTWARE, VOL. 9, NO. 6, JUNE 2014.