Using boosting for automated planning and trading systems

Germán Creamer

Submitted in partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2007

ABSTRACT

Using boosting for automated planning and trading systems

Germán Creamer

The problem: Much of finance theory is based on the efficient market hypothesis. According to this hypothesis, the prices of financial assets, such as stocks, incorporate all information that may affect their future performance. However, the translation of publicly available information into predictions of future performance is far from trivial. Making such predictions is the livelihood of stock traders, market analysts, and the like. Clearly, the efficient market hypothesis is only an approximation which ignores the cost of producing accurate predictions.

Markets are becoming more efficient and more accessible because of the use of ever faster methods for communicating and analyzing financial data. Algorithms developed in machine learning can be used to automate parts of this translation process. In other words, we can now use machine learning algorithms to analyze vast amounts of information and compile them to predict the performance of companies, stocks, or even market analysts. In financial terms, we would say that such algorithms discover inefficiencies in the current market. These discoveries can be used to make a profit and, in turn, reduce the market inefficiencies or support strategic planning processes.

Relevance: Currently, the major stock exchanges such as NYSE and NASDAQ are transforming their markets into electronic financial markets. Players in these markets must

process large amounts of information and make instantaneous investment decisions.

Machine learning techniques help investors and corporations recognize new business opportunities or potential corporate problems in these markets. With time, these techniques help the financial market become better regulated and more stable. Also, corporations could save significant amount of resources if they can automate certain corporate finance functions such as planning and trading.

Results: This dissertation offers a novel approach to using boosting as a predictive and interpretative tool for problems in finance. Even more, we demonstrate how boosting can support the automation of strategic planning and trading functions.

Many of the recent bankruptcy scandals in publicly held US companies such as Enron and WorldCom are inextricably linked to the conflict of interest between shareholders (principals) and managers (agents). We evaluate this conflict in the case of Latin American and US companies. In the first part of this dissertation, we use Adaboost to analyze the impact of corporate governance variables on performance. In this respect, we present an algorithm that calculates *alternating decision trees* (ADTs), ranks variables according to their level of importance, and generates representative ADTs. We develop a board *Balanced Scorecard* (BSC) based on these representative ADTs which is part of the process to automate the planning functions.

In the second part of this dissertation we present three main algorithms to improve forecasting and automated trading. First, we introduce a link mining algorithm using a mixture of economic and social network indicators to forecast earnings surprises, and cumulative abnormal return. Second, we propose a trading algorithm for short-term technical trading. The algorithm was tested in the context of the Penn-Lehman Automated Trading

Project (PLAT) competition using the Microsoft stock. The algorithm was profitable during the competition. Third, we present a multi-stock automated trading system that includes a machine learning algorithm that makes the prediction, a weighting algorithm that combines the experts, and a risk management layer that selects only the strongest prediction and avoids trading when there is a history of negative performance. This algorithm was tested with 100 randomly selected S&P 500 stocks. We find that even an efficient learning algorithm, such as boosting, still requires powerful control mechanisms in order to reduce unnecessary and unprofitable trades that increase transaction costs.

Contents

List of Figures

List of Tables

Acknowledgments

This thesis would not have been possible without the help of many people. I would like to thank the advisors of my dissertation, Yoav Freund and Salvatore J. Stolfo. Their advice has been very helpful in walking me through this complicated area of interdisciplinary research. Yoav's advice in the definition of algorithms and in the application of boosting to finance problems, was extremely useful in developing the appropriate forecasting algorithms. Sal's suggestions in the area of link mining, as well as the inclusion of my dissertation topic, as part of his overall project of intrusion detection, helped to give a broader perspective to my dissertation. Additionally, Sal has also been encouraging and offered me very valuable advice regarding my academic development. Special thanks to all the members of my committee, David Waltz, Tony Jebara and Vasant Dhar, who have helped me in the different stages of my dissertation, and have given me insightful thoughts about my research.

My deep appreciation specially to David Waltz who gave me very critical and detailed comments about my dissertation. His invitation to participate in the Center for Computational Learning Systems (CCLS) has helped me to enjoy my stay at Columbia where I have found an informal atmosphere of academic collegiality. Conversations with Hubert Delany, Christina Lesley, Marta Arias, Nizar Habash, Wei Chu, Raphael Pelossof, Ray Kwan, Aaron W. Harnly, and Andrew Howard helped to think about new perspectives of my research.

I appreciate the feedback and the invitation by Michael Kearns to participate in the Penn-Lehman automated trading competition (PLAT). His comments and suggestions about

To my mother

To María Consuelo, Mateo and Carolina

Chapter 1

Introduction

1.1 Quantitative analysis in finance

Quantitative evaluation of econometric models is usually done by evaluating the statistical significance of linear models. For example, previous studies on US securities (see the pioneering works of Altman [9], and Beaver [26], and also see [10, 11, 12, 21, 70, 61, 65, 114, 126, 153, 124, 182, 186, 195, 197, 199, 241]) have used linear discriminant analysis or logistic regression for the prediction of financial distress, bankruptcy, and credit risk. This analysis is based on estimating the parameters of an underlying stochastic system, usually assumed to be a linear system. One limitation of this methodology is that nonlinearities have to be incorporated manually. Another limitation is that the number of parameters that can be estimated reliably is limited by the amount of available data, and is often very small.

By contrast, machine learning methods such as decision trees [47], boosting [105] and support vector machines [183] avoid the question of estimating the parameters of the under-

lying distribution and focus instead on making accurate predictions for some variables given others variables. Breiman [46] contrasts these two approaches as the data modeling culture and the algorithmic modeling culture. According to Breiman [46], while most statisticians adhere to the data-modeling approach, people in other fields of science and engineering use algorithmic modeling to construct predictors with superior accuracy. The main drawback of algorithmic modeling, according to Breiman, is that although the models are easy to generate, they are hard to interpret.

In this research, we apply algorithmic modeling to predict and interpret the determinant factors of corporate performance, and to forecast stock prices, cumulative abnormal return, and earnings surprises. We use two variations of boosting, Adaboost and Logitboost, as the learning algorithm. Adaboost is a general discriminative learning algorithm invented by Freund and Schapire [105]. The basic idea of Adaboost is to repeatedly apply a simple learning algorithm, called the *weak* or *base* learner, to different weightings of the same training set. The simplest form of Adaboost is intended for binary prediction problems where the training set consists of pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$, x_i corresponds to the features of an example, and $y_i \in \{-1, +1\}$ is the binary label to be predicted. A *weighting* of the training examples is an assignment of a non-negative real value w_i to each example (x_i, y_i). Friedman et al. [106], followed by Collins, Schapire, and Singer [69], suggested a modification of Adaboost, called Logitboost. Logitboost can be interpreted as an algorithm for step-wise logistic regression.

We use boosting both to learn the decision rules constituting the tree, and to combine these rules through a weighted majority vote. The form of the generated decision rules is called an *alternating decision tree* (ADT) [104].

In the first part of this dissertation (chapter 3), we use Adaboost to define a board Balanced Scorecard (BSC), focusing on the conflict of interest between principal (shareholders) and agents (managers). We concentrate on Latin American ADRs, banks domiciled in Latin American countries, and in the S&P 500 companies.

In the second part of this dissertation (chapters 4, 5 and 6), we apply Logitboost to forecast company earnings, and cumulative abnormal return.

Finally, we use Logitboost to find profitable trading strategies and apply them in a financial trading agent competition using a single stock, and as part of a forecasting and multi-stock automated trading system. We decided to use boosting as our learning algorithm because of its feature selection capability, its error bound proofs [105], its interpretability, and its capacity to combine quantitative, and qualitative variables.

In the next section, we introduce the related work for these two major parts of the dissertation: corporate governance and automated planning using the BSC, and automated trading systems. The following section presents the main contributions, and the final section presents the organization of this dissertation.

1.2 Related work

1.2.1 Corporate governance and automated planning

Many of the recent bankruptcy scandals in publicly held US companies such as Enron and WorldCom are inextricably linked to a conflict of interest between shareholders (principals) and managers (agents). This conflict of interest is called the principal agent problem in finance literature. The principal agent problem stems from the tension between the interests

of the investors in increasing the value of the company and the personal interests of the managers.

One of the main areas where the agency conflict is expressed is in the compensation of the top executives of the firm. Before the 1970s, compensation was based mostly on salaries and bonuses that were linked to performance, but now most of the compensation is based on stock options. Jensen and Murphy [134] calculate that the average total remuneration of CEOs from S&P 500 firms has gone from $850,000 in 1970 to $9.4 million in 2002 (using 2002-constant dollars). The value of the options in the same period went from almost zero to $4.4 million. Jensen and Murphy [132, 133] as well as shareholder representatives suggested that executive compensation should include a larger options component. Compensations committees tend to grant stock options to CEOs and top managers since stock options are not debited as a cost to the firm. Jensen and Murphy [134] recognize the excess of these compensations committees and propose instead that the cost of granting options is the opportunity cost of not selling these options in the market. The Sarbanes-Oxley Act of 2002 introduced important provisions for executive compensation such as the prohibition of executive loans, and periods when insider trading is not allowed. The Financial Accounting Standards Board (FASB) and the Securities and Exchange Commission (SEC) complemented these rules requiring that if companies grant options to employees, those options should be registered in the financial statements as an expense for fiscal years beginning after June 15, 2005.

Several authors have studied the effect of altering the terms of executive stock options on performance [48] and as a reincentivization strategy [3]. Firms with agency problems and with insider-dominated boards are more to likely reprice executive stock options [58], while

companies that have more outsiders directors grant more compensation packages, such as equity-based compensation, to directors aligned with shareholders' interests. [202].

Jensen and Murphy [134] indicates that the fundamental motivation to grant executive options, which is to align corporate and managerial goals, is not fulfilled by the current executive compensation policy. On the other hand, current research shows that some of the policies are at least partially effective. An algorithm that would establish what are the appropriate thresholds, and under what conditions executive options should be granted can potentially be used to align corporate and managerial interests and reduce the principal agent conflict. Another major area where the principal-agent problem is evident is insider ownership. According to Jensen and Meckling [131], the separation of ownership and control is often seen as an opportunity for managers to accumulate wealth at the expense of shareholders [29, 210]. Ang, Rebel and Lin [14], using a sample of small US companies, show how agency costs increase with a decrease in managerial ownership as proposed by Jensen and Meckling [131]. Based on previous study by Weston [233] who indicates that beyond board ownership of 20-30% a hostile bid cannot succeed, Morck, Shleifer and Vishny [180] highlight the opposing effects of extensive insider ownership. On the one hand, a high proportion of insider ownership has a positive impact on performance because of insiders' incentive alignment with other shareholders (convergence-of-interests hypothesis) [130, 52, 74]. On the other hand, a high proportion of insider ownership has a negative impact on performance because of the insider's bargaining power that may lead managers to make self-interested decisions (entrenchment hypothesis) [131]. Stulz [217] finds–through a formal model–that the relationship between ownership and performance follows a roof-shaped curve where low levels of ownership improve performance while high levels of ownership affects performance.

McConnell and Servaes [170] and Fuerst and Kang [107] empirically confirm the implications of this model. Other studies show mixed results [174].

The structure and size of the board of directors also have an important effect on the principal-agent problem. The board of directors plays a high-level counsel and control role in any organization. However, it is necessary that the board of directors include outsiders (members who are not part of the management team) and maintain a minimal level of ownership to ensure their interest in the performance of the company. A board of directors may fail due to a strong emphasis on the CEO's personal agenda, low equity ownership among the board's members, an excessively large board of directors, and a culture that discourages dissent [130].

There is no a theoretical support to indicate the optimal values of organizational variables such as insider ownership and the structure and size of the board of directors. Moreover, these variables may change from industry to industry and country to country. Therefore a system that is able to recognize the optimal combination and the mechanism that connects these variables, would contribute significantly to an efficient planning process.

1.2.1.1 The Board Balanced Scorecard and automated planning

In response to the recent corporate scandals in the USA, several organizations and researchers have proposed corporate governance scorecards. Gompers et al. [112] use 24 different provisions related to takeover defense and shareholder rights to create a governance index. They show that a trading strategy based on this index outperforms the market. Standard & Poor's [207] have developed a method which combines macro and micro variables

and uses qualitative and quantitative analysis.[1] The German Society of Financial Analysts [212] has proposed a corporate governance scorecard for German corporations which is based on a "Code of Best Practice" following the German law. The German Society of Financial Analysts and partially Standard & Poor's use a qualitative framework based on "best practices" and require a lengthy due diligence process for each company under study, while the Gompers approach is purely quantitative.

In a different line of research, Kaplan and Norton [138] introduced the Balanced Scorecard (BSC) as a management system that helps organizations define their vision and strategy, and translate them into specific actions. The BSC provides feedback on internal business processes, performance, and market conditions in order to review the strategy and future plans [139, 141, 140, 142]. Large US companies, such as General Electric and Federal Express, and non-profit and public organizations have implemented the BSC approach [13, 226].

The BSC suggests that an organization should be evaluated from four perspectives:

1. **The financial perspective** emphasizes the long-term objectives of the company in terms of revenue growth and productivity improvement. The financial goals should be the final goals for the other perspectives.

2. **The customer perspective** emphasizes the lifetime relationship and service delivery with clients.

3. **The internal processes perspective** focuses on the use of clients' information to

---

[1]Even though the Standard & Poor's corporate governance scoring has been very successful in emerging markets, Standard & Poor's corporate governance services decided to pull out of the US market in September 2005.

sell new products and services according to the clients' needs.

4. **The learning and growth perspective** is the foundation of the BSC. This per-
   spective looks at the motivation, training, and capacity to innovate that employees
   need to have in order to implement the new strategies.

 The BSC is generally implemented at the corporate level, business unit level, and indi-
vidual level.

 At difference of the corporate governance scorecards presented at the beginning of this
section which emphasize corporate governance scoring, Kaplan and Nagel [137] proposed
the creation of a board Balanced Scorecard that includes corporate governance variables,
and is oriented to strategic planning at the board level. According to Kaplan and Nagel an
effective BSC program should include three parts:

1. **An enterprise BSC** that presents the company strategy, with detailed description
   of objectives, performance measures, targets, and initiatives to be implemented by
   the CEO and managers throughout the organization.

2. **A board BSC** which defines the strategic contribution of the board, includes the
   strategic data necessary for the board operation, and offers an instrument to monitor
   the structure and performance of the board and its committees.

3. **An executive BSC** allows the board of directors and the compensation committee
   to evaluate the performance of the top managers of the organization. Epstein and
   Roy [92, 93] explain the importance of the board BSC as an instrument to monitor
   and implement the best-practices of corporate governance, and also as a mechanism
   to evaluate the board of directors by the stakeholders.

The strategy of an organization, its main objectives, and its key business drivers define the indicators of the BSC. However, the choice of indicators is, in general, highly subjective and is often driven by the company management or industry practices. There are several proposals for more objective methods for quantifying board performance. YoungBlood and Collins [239] describe a method based on indicators using multi-attribute utility theory. Clinton et al. [68] base their method on Analytic Hierarchy Process. However, these methods still require a mix of quantitative measure with a qualitative evaluation by managers or experts. This dissertation proposes a method to automate the definition of the board BSC, thus making the process of company evaluation more objective and more transparent.

1.2.2 Forecasting and automated trading systems

1.2.2.1 Earnings prediction

The conflict of interest between principal and agents has also led to the so-called "earnings game". CEOs' compensation depends on their stock options. So, top managers concentrate on the management of earnings and surprises. Wall Street companies want to keep selling stocks. Thus, analysts try to maintain positive reviews of the companies. [2] Once a prediction is published, CEOs do whatever is necessary to reach that prediction or boost the results above analysts' prediction. CEOs play this game, even though a company may lose value in the long-term, because it boosts the potential value of their stock options.

The Institutional Brokers' Estimate System, IBES, has collected the analysts' earnings forecast and their revisions since the early seventies. Several other companies such as Zacks

---

[2] In the last years, this situation is changing because of the new separation between research and investment banking.

Investment Research and First Call Corporation[3] have also joined this effort and have extended the service to include other accounting indicators such as revenue announcements. These databases provide an estimation of markets expectations or market "consensus" about the future earnings announcement which is a simple average of the market analysts' predictions. Investors follow very closely these consensus indicators to forecast and take their investment decisions. Another important use of this information is screening and ranking the analysts' according to their previous performance. The company Starmine provides a new indicator called "Smart estimate". This indicator is an estimator of the earnings or revenue quarterly announcement based on the information provided by the most highly ranked analysts.

From the computer science perspective, the existence of these financial databases offers a great opportunity to evaluate the capacity of several machine learning methods to search for new finance time-series patterns that may improve the current forecasts. From the finance industry perspective, small investors or large institutional investors in the international finance market do not have the capacity to conduct detailed investment research in every market that they invest. Therefore, machine learning methods can help them to process a large amount of existent data to forecast earnings surprises, and cumulative abnormal return. Earnings surprise is the difference between actual quarterly earnings and consensus. Dhar and Chou [86] have already compared the predictive accuracy of tree-induction algorithms, neural networks, naive Bayesian learning, and genetic algorithms to classify the earnings surprise before announcement. Cumulative abnormal return is the return of a

---

[3]Recently, Thomson Financial acquired First Call and IBES and plans to integrate these two databases using First Call format.

specific asset less the average return of all assets in its risk-level portfolio for each trading date.

1.2.2.2 Financial electronic markets and automated trading strategies

The major stock exchanges such as NYSE and NASDAQ are transforming their markets into electronic financial markets. Many traders in these markets must rely on automated trading systems in order to process large amounts of information and make instantaneous investment decisions.

The early automated trading systems embedded classical artificial intelligence approaches such as expert systems, fuzzy logic, neural networks and genetic algorithms (see Trippi and Turban [223, 224], Trippi and Lee [222], Deboeck [78] and Chorafas [63] for a review of these systems. Goonatilake and Treleaven [113] survey an application of the above methods to automated trading and several other business problems such as credit risk, direct marketing, fraud detection, and price forecasting.).

Generally, the current automated trading systems include a backtest or simulation module. In this respect, the models developed by the agent-based perspective could be useful to explore new ideas without risking any money. The Santa Fe stock market model [4] has inspired many other agent-based financial market models such as Ehrentreich's [88], which is based on the Grossman and Stiglitz model [117]. In the Santa Fe stock market agents can classify and explore several forecasting rules that are built using genetic algorithms. Many of the models built in this perspective test the performance of agents or algorithms

---

[4]For a presentation of the Santa Fe stock market model see Arthur et al. [16], LeBaron et al. [157], and a later version at Lebaron [159]. Lebaron [156] has also a general review of papers in the area of agent-based finance.

that have unique characteristics. For example, Lettau [162] builds an agent-based financial market using simple agent benchmarks based on genetic algorithms; Gode and Sunder [110] develop a double auction market using random or zero intelligence traders; Arifovic [15] builds a model of the foreign exchange market using genetic algorithms; Routledge [201] extends the basic framework of Grossman and Stiglitz [117] with agents that can learn using genetic algorithms; Chan et al. [56] and Chan [57] use the artificial market framework to explore the behavior of different trading approaches and their microstructure impact. The application of the Ising model [127] to financial markets has led to several versions of the spin model where a sell position is a spin-up and a buy position is a spin-down. Prices are determined by the aggregation of traders' positions.[5]

These agent-based models are useful for simulation; however they do not automate the trading functions. On the contrary, Seo et al. [206] and Decker [79] describe a multi-agent portfolio management system that automatically classifies financial news. In this line of research, Thomas [219] combines news classification with technical analysis indicators in order to generate new trading rules. Lavrenko et al. [155] describe a system that recommends news stories that can affect market behavior. This is a special case of the activity monitoring task as suggested by Fawcett and Provost [99]. In a manner similar to fraud detection systems, activity monitoring generates alarms when an unusual event happens. These signals try to recognize when the trend of the market is positive, and therefore can generate new trading signals. Wuthrich et al. [237] and Cho, Wuthrich, and Zhang [62] weight keywords based on their occurrences to predict the direction of major stock indices. Text

---

[5][41] modified this model introducing an anti-ferromagnetic coupling between the global magnetization and each spin, as well as a ferromagnetic coupling between the local neighborhood and each spin.

classification and retrieval applied to finance is still an area under-explored in the literature. However, several investment banks and hedge funds are developing systems to automatically incorporate the impact of daily news into their trading systems. Our current work does not incorporate the news aspect, however the methods used to automate forecasting, trading or the extraction of new indicators are an important antecedent for our work in automated trading systems.

1.2.2.3 Trading strategies and technical analysis

Another important line of research in the trading algorithmic literature is the use of learning algorithms to generate trading rules using technical analysis indicators. Technical analysis or technical trading strategies try to exploit statistically measurable short-term market opportunities, such as trend spotting and momentum, in individual industrial sectors (e.g. financial, pharmaceutical etc.).

The presence of technical analysis has been very limited in finance literature because of its lack of a solid statistical or mathematical foundation, its highly subjective nature, and its visual nature. In the 60s and 70s researchers studied trading rules based on technical indicators and did not find them profitable [6] [96]. These findings led Fama [94] to dismiss technical analysis as a profitable technique and support the efficient market hypothesis. Part of the problem of the studies during the 60s was the *ad hoc* specifications of the trading rules that led to spurious patterns in the data. Specification of rules retroactively may have led to biased studies. Recently, Allen and Karjalainen [8] found profitable trading rules using genetic algorithms for the S&P 500 with daily prices from 1928 to 1995. However, these rules were not consistently better than a buy-and-hold strategy in the out-of-sample

test periods.

In the last years, there has been a growing interest on applying machine learning methods to formulate trading strategies using technical indicators such as the following: Lo, Mamaysky, and Wang [166], who used nonparametric kernel regression for technical pattern recognition of a large number of stocks for the period 1962 - 1996, found that technical indicators provide incremental information for investors comparing the unconditional empirical distribution of daily stock returns to the conditional distribution on specific technical indicators such as head and shoulders. Moody and Saffell [179] found that a trading system using direct reinforcement learning outperforms a Q-trader for the asset allocation problem between the S&P 500 and T-bill. Dempster et al. [82] compared four methods for foreign exchange trading (reinforcement learning, genetic algorithms, Markov chain linear programming, and simple heuristic) and concluded that a combination of technical indicators leads to better performance than using only individual indicators. Dempster and Leemans [80] reached a similar conclusion using adaptive reinforcement learning. Bates et al. [22] used Watkin's Q-learning algorithm to maximize profits; these authors compared order flow and order book data, and compared with technical trading rules. They concluded that using order flow and order book data was usually superior to trading on technical signal alone. LeBaron [158] applied bootstrapping to capture arbitrage opportunities in the foreign exchange market, and then used a neural network where its network architecture was determined through an evolutionary process. Finally, Towers and Burgess [221] used principal components to capture arbitrage opportunities.[6]

---

[6]See also [43] for an algorithmic approach to find trading strategies without including technical analysis indicators.

In this research we follow the tradition of the papers in this section that use machine learning algorithms to find profitable trading strategies and also build completely automated trading systems.

1.3   Main contributions of the dissertation

This dissertation studies the application of machine learning technology to very hard problems in computational finance such as financial time-series forecast and cross-stock analysis. So far, computational finance has been mostly associated with options valuation. We think this discipline can be enriched with the adoption of efficient computational methods to manage, optimize, and extract information from large financial datasets used in planning, trading, and risk management. Specifically, this dissertation compares boosting with other learning algorithms and offers a novel approach to using boosting as a predictive and an interpretative tool for problems in finance. Even more, we demonstrate how boosting can be part of a link mining algorithm that integrates accounting and organizational information. We are not aware of any other previous application of boosting in automated planning and trading. We decided to use boosting as our learning algorithm because of its feature selection capability, its error bound proofs [105], its interpretability, and its capacity to combine quantitative, and qualitative variables. This dissertation partially covers the existing gap between finance and machine learning literature especially in these areas. The application of boosting to forecasting of financial time-series requires an important effort of fine-tuning as well as selection of parameters. In this dissertation we present the details of the methodology and algorithms so that a finance or computer science researcher or practitioner could

apply these algorithms to obtain new predictions with Adaboost or Logitboost, interpret ADTs, automate planning and trading functions, or apply them to other finance problems.

Our main contributions are in two areas:

1.3.1 Automated planning systems:

1. **Representative alternating decision trees (ADTs) algorithm:** We develop an algorithm that ranks variables according to their level of importance in the ADTs, and generates representative ADTs with the most important variables.

   This research shows that Adaboost performed similarly to logistic regression, random forests, and bagging with stable datasets when we compared small and large samples from different countries and economic conditions. Additionally, we show how boosting and representative ADTs can be used as interpretative tools to evaluate the impact of corporate governance factors on performance and efficiency. Representative ADTs are particularly useful to understand the non-linear relationship between the variables that affects performance and efficiency.

2. **Performance management and automated planning system:** We demonstrate that the representative ADT is a useful tool to select and establish the relationship among the most important indicators of a board BSC. Additionally, the thresholds of the representative ADT define targets or ranges of values of the indicators that managers could follow to improve corporate performance. With this combined tool, managers can concentrate on the most important strategic issues and delegate the calculation of the targets to an automated planning system supported by Adaboost.

The main group of variables that we use are related to the principal agent problem because of their effect on company performance and efficiency. Boosting is the learning algorithm and interpretative tool of our board BSC. This model evaluates whether a company's performance or a bank's efficiency is above or below par as a function of the main corporate governance factors (executive compensation, insider ownership, and board of directors structure), and of selected accounting ratios that are known to be important in evaluating corporate governance. These features, selected and quantified by stumps-averaged classifier trained using boosting, become the main indicators and targets of the board BSC.

1.3.2 Forecasting and automated trading systems:

1. **Link mining algorithm and earnings forecast:** We propose a link mining algorithm, CorpInterlock, that selects the largest strongly connected component of a social network and ranks its vertices using several indicators of distance and centrality. These indicators are merged with other relevant indicators in order to forecast new variables using a boosting algorithm. We apply this link mining algorithm to integrate accounting variables of US companies with statistics of social networks of directors only (basic corporate interlock) and social networks of directors and analysts (extended corporate interlock) to forecast earnings surprises and cumulative abnormal returns. Link mining[7] is a set of techniques that uses different types of networks and their indicators to forecast or to model a linked domain.

   We implement CorpInterlock with Logitboost because it is a very flexible method

---

[7]For a recent survey see [109]

and can analyze a large and diverse group of quantitative and qualitative variables. The boosting approach also generates a score with each prediction which can be associated with the strength of the prediction. We establish that CorpInterlock implemented with Logitboost improves the prediction of earnings surprise in relation to the implementation of CorpInterlock with logistic regression.

2. **Small world and corporate interlock:** This research shows that the basic and extended corporate interlocks have the properties of a "small world" network. The "small world" model was formalized by Watts [229] and Watts et al. [230, 184, 185] based on the pioneering work of Milgram [175] who shows how apparently distant people are connected by a very short chain of acquaintances.

The statistics of an extended corporate interlock, directors and financial analysts, bring additional information to predict cumulative abnormal return, especially during a "bull" market.

3. **Constant rebalanced portfolio - technical analysis trading algorithm (CRP_TA):** We propose an algorithm for short-term technical trading. The algorithm was tested in the context of the Penn-Lehman Automated Trading Project (PLAT) competition, and is based on three main ideas. The first idea is to use a combination of technical indicators to predict the daily trend of the stock. The combination is optimized using a boosting algorithm. The second idea is to use the constant rebalanced portfolios (CRP) [7] within the day in order to take advantage of market volatility without increasing risk. The third idea is to use limit orders rather than market orders to minimize transaction costs.

The algorithm was profitable during the PLAT competition, and after the competition we enhanced it by including a market maker component. We show that the constantly rebalanced portfolio can improve if a classifier can anticipate the direction of the market. Additionally, transaction costs play a central role to raise performance. Instead of an automatic rebalance of the portfolio, the results of the PLAT competition indicate that if the CRP strategy is implemented only with limit orders, its results improve because of the rebates.

4. **Automated trading system:** We propose a multi-stock automated trading system. The system is designed to trade stocks, and relies on a layered structure consisting of ADT, which is implemented with Logitboost, as the machine learning algorithm; an online learning utility; and a risk management overlay. [8] The system generates its own trading rules, and weights the suggestion of the different ADTs or experts to propose a trading position. Finally, the risk management layer can validate the trading signal when it exceeds a specified non-zero threshold, and limit the use of a trading strategy when it is not profitable.

We test the expert weighting algorithm with data of 100 randomly selected companies of the S&P 500 index during the period 2003-2005. We find that this algorithm generates excess returns during the test period. Every component of the trading algorithm is important to obtain positive abnormal returns, and brings some functionality that is complemented by the rest of the layers. We observe that even an efficient learning algorithm, such as boosting, still requires powerful control mechanisms in order to

---

[8]See Dempster and Leemans [80] for a previous trading system using machine learning algorithms and a layered structure.

reduce unnecessary and unprofitable trades that increase transaction costs. Hence, the contribution of new predictive algorithms by the computer science or machine learning literature to finance still needs to be incorporated under a formal framework of risk management.

As part of the optimization of the trading system, we propose a method to simultaneously calculate the same features using different parameters, leaving the final feature selection to boosting. Many trader systems become very inefficient because they try all the parameters or are forced to select in advance parameters that are not adequate after a trading period. Our experiments show that the boosting approach is able to improve the predictive capacity when indicators are combined and aggregated as a single predictor. Even more, the combination of indicators of different stocks are demonstrated to be adequate in order to reduce the use of computational resources, and still maintain an adequate predictive capacity.

1.4   Organization of the dissertation

The remainder of the thesis is organized as follows:

Chapter 2 introduces the main methods (logistic regression, bagging, random forests, efficiency calculations, boosting, and ADTs) used in this research.

Chapter 3 presents the algorithm to calculate the representative ADT that ranks variables according to their level of importance in the ADTs. This algorithm supports the generation of a board BSC for S&P 500 companies. Instead of defining in advance a normative framework, we use a data-driven model in which the relevant features are selected

according to their positive impact on corporate efficiency. Additionally, the thresholds of the representative ADTs establish targets or ranges of values of the indicators that managers could follow to improve corporate performance.

Chapter 4 explores the effect of the "earnings game", and determines social networks indicators used for forecasting. We evaluate whether a social network that includes financial analysts and directors of US companies improves the forecast of earnings surprises and cumulative abnormal return in comparison to a social network that only includes directors. We combine accounting indicators with social network indicators to forecast earnings surprises and cumulative abnormal returns using a link mining algorithm. Appendix A includes the accounting and social network indicators used to forecast earnings surprises and cumulative abnormal return.

Chapter 5 introduces a trading algorithm for short-term technical trading used during the PLAT competition for the case of Microsoft. Appendix B includes the technical analysis indicators used for the trading algorithm during this competition.

Chapter 6 presents a complete multi-stock automated trading system with a boosting algorithm, online expert weighting, and a risk management layer. This chapter also includes the results of its implementation for a group of S&P 500 companies. Appendix C has the investment signals used for this trading system. This appendix includes most of the indicators included in appendix B plus an additional group of technical analysis indicators, and investment signals especially related with volatility such as generalized uutoregressive conditionally heteroskedastic (GARCH). Appendix C.1 explains in detail the calculation of GARCH.

Chapter 7 concludes the dissertation. This chapter discusses limitations of the methods

used, especially concentrating on the process of adjusting a machine learning algorithm to a

finance problem. We finally recommend areas of future research, in terms of how to expand:

a) the automated trading system and b) the automated planning system or the enterprise

BSC using boosting.

Chapter 2

Methods

2.1   Introduction

This chapter introduces the main methods that are used in this dissertation. The most important learning algorithms that we have explored in this research are Adaboost and Logitboost. We have also compared our results with other methods such as logistic regression, random forest, and bagging. In this research we have used the boosting approach as:

1. A predictive and interpretative tool to select and measure the main features used in a board Balanced Scorecard (chapter 3)

2. A predictive system for all the US stock market that combines accounting variables with social network and organizational variables (chapter 4).

3. A predictive tool that is able to integrate several technical analysis indicators to forecast a single stock price for an automated trading competition (chapter 5).

4. A predictive and integrative tool that is part of a multi-stock and long term automated trading system (chapter 6)

Every chapter brings a different perspective and introduces several algorithms that are appropriate to solve the main problems presented. However, the main algorithms that are used in all chapters, Adaboost or Logitboost, and additional algorithms that are used as benchmarks to compare the performance of boosting such as logistic regression, bagging, and random forests, are introduced in the next section. The final section introduces efficiency calculations using frontier analysis and data envelopment analysis that are used in chapter 3 for the evaluation of Latin American banks.

2.2 Learning methods

2.2.1 Logistic regression

The logistic regression models the posterior probabilities $Pr(C = l|X = x)$ of L classes C using linear regression in the observed values x of the input variable X. The model is a series of ordinary regressions:

$$\log \frac{Pr(C=1|X=x)}{Pr(C=L|X=x)} = \beta_{10} + \beta_1^T x$$

$$\log \frac{Pr(C=2|X=x)}{Pr(C=L|X=x)} = \beta_{20} + \beta_2^T x$$

...

$$\log \frac{Pr(C=L-1|X=x)}{Pr(C=L|X=x)} = \beta_{(L-1)0} + \beta_{L-1}^T x$$

where L-1 logit transformations or log-odds are the dependent variables with the logistic regression coefficients β_{l0}.

Taking the exponential of the log-odds, we can calculate the probabilities of each class

as follows:

$$Pr(C = l | X = x) = \frac{e^{\beta_{l0} + \beta_l^T x}}{1 + \sum_{r=1}^{L-1} e^{\beta_{r0} + \beta_r^T x}}, \; l = 1, \, ..., \, \text{L-1}$$

$$Pr(C = L | X = x) = \frac{1}{1 + \sum_{r=1}^{L-1} e^{\beta_{r0} + \beta_r^T x}}$$

The summation of these probabilities equal one. Logistic regression results are better interpreted using the odds ratios which can be computed by taking the exponential of the logistic regression coefficients [120].

2.2.2   Bagging and random forests

Bagging was proposed by Breiman [44] as a method that reduces the variance of a prediction function. If the training set Υ consists of pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$, x_i corresponds to the features of an example, and y_i is either a class label or a numerical response to be predicted. The predictor of y is $\psi(x, \Upsilon)$. Bagging or bootstrap aggregation generates uniform bootstrap samples with replacement of Υ. These samples and their predictors are $\Upsilon^{(B)}$ and $\psi(x, \Upsilon^{(B)})$ respectively.

When y_i is a numerical response, the final predictor is obtained by the average of the predictors of the bootstrap samples as

$$\psi_B(x) = av_B \psi(x, \Upsilon^{(B)}).$$

If y_i is a class label, $\psi_B(x)$ is obtained by the majority vote of $\psi(x, \Upsilon^{(B)})$.

Bagging has been shown to be particularly effective for reducing the variance of decision trees.

Random forests is a variant of bagging decision trees also proposed by Breiman [45], and for which free computer code is available. We chose this algorithm because it presents the best publicly available combination of decision trees and bagging.

This algorithm generates multiple trees (θ_i) from the training data, and from a random vector (x) sampled independently and with the same distribution for any tree that is part of the forest. As a result each tree generates a classifier $h(x, \theta_i)$. The majority vote of all the trees determine the predicted class. When the number of trees is very large, the generalization error for forests converges. Breiman [45] indicates that the accuracy of random forests is as good as Adaboost or better.

2.2.3 Boosting

Adaboost is a general discriminative learning algorithm invented by Freund and Schapire [105].

The basic idea of Adaboost is to repeatedly apply a simple learning algorithm, called the *weak* or *base* learner[1], to different weightings of the same training set. In its simplest form, Adaboost is intended for binary prediction problems where the training set consists of pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$, x_i corresponds to the features of an example, and $y_i \in \{-1, +1\}$ is the binary label to be predicted. A *weighting* of the training examples is an assignment of a non-negative real value w_i to each example (x_i, y_i).

On iteration t of the boosting process, the weak learner is applied to the training sample with a set of weights $w_1^t, \ldots, w_m^t$ and produces a prediction rule h_t that maps x to $\{0, 1\}$.[2] The requirement on the weak learner is for $h_t(x)$ to have a small but significant correlation with the example labels y when measured using the *current weighting of the examples*. After the rule h_t is generated, the example weights are changed so that the weak predictions $h_t(x)$ and the labels y are decorrelated. The weak learner is then called with the new weights

---

[1]Intuitively, a weak learner is an algorithm with a performance at least slightly better than random guessing

[2]Mapping x to $\{0, 1\}$ instead of $\{-1, +1\}$ increases the flexibility of the weak learner. Zero can be interpreted as "no prediction".

over the training examples, and the process repeats. Finally, all of the weak prediction rules are combined into a single *strong* rule using a weighted majority vote. One can prove that if the rules generated in the iterations are all slightly correlated with the label, then the strong rule will have a very high correlation with the label – in other words, it will predict the label very accurately.

The whole process can be seen as a variational method in which an approximation $F(x)$ is repeatedly changed by adding to it small corrections given by the weak prediction functions. In Figure 2.1, we describe Adaboost in these terms. We shall refer to $F(x)$ as the *prediction score* in the rest of the document. The strong prediction rule learned by Adaboost is $\mathrm{sign}(F(x))$.

A surprising phenomenon associated with Adaboost is that the test error of the strong rule (percentage of mistakes made on new examples) often continues to decrease even after the training error (fraction of mistakes made on the training set) reaches zero. This behavior has been related to the concept of a "margin", which is simply the value $yF(x)$ [204]. While $yF(x) > 0$ corresponds to a correct prediction, $yF(x) > a > 0$ corresponds to a *confident* correct prediction, and the confidence increases monotonically with a.

$$
\begin{aligned}
&F_0(x) \equiv 0 \\
&\text{for } t = 1 \ldots T \\
&\qquad w_i^t = e^{-y_i F_{t-1}(x_i)} \\
&\qquad \text{Get } h_t \text{ from } \textit{weak learner} \\
&\qquad \alpha_t = \tfrac{1}{2} \ln \left(\frac{\sum_{i:h_t(x_i)=1, y_i=1} w_i^t}{\sum_{i:h_t(x_i)=1, y_i=-1} w_i^t} \right) \\
&\qquad F_{t+1} = F_t + \alpha_t h_t
\end{aligned}
$$

Figure 2.1: The Adaboost algorithm [105]. y_i is the binary label to be predicted, x_i corresponds to the features of an instance i, w_i^t is the weight of instance i at time t, h_t and $F_t(x)$ are the prediction rule and the prediction score at time t respectively

$$F_0(x) \equiv 0$$
$$\text{for } t = 1 \ldots T$$
$$w_i^t = \frac{1}{1 + e^{y_i F_{t-1}(x_i)}}$$
$$\text{Get } h_t \text{ from } \textit{weak learner}$$
$$\alpha_t = \frac{1}{2} \ln \left(\frac{\sum_{i:h_t(x_i)=1, y_i=1} w_i^t}{\sum_{i:h_t(x_i)=1, y_i=-1} w_i^t} \right)$$
$$F_{t+1} = F_t + \alpha_t h_t$$

Figure 2.2: The Logitboost algorithm [106]. y_i is the binary label to be predicted, x_i corresponds to the features of an instance i, w_i^t is the weight of instance i at time t, h_t and $F_t(x)$ are the prediction rule and the prediction score at time t respectively

Friedman et al. [106], followed by Collins, Schapire, and Singer [69] suggested a modification of Adaboost, called Logitboost. Logitboost can be interpreted as an algorithm for step-wise logistic regression (see section 2.2.1). This modified version of Adaboost–known as Logitboost–assumes that the labels $y_i's$ were stochastically generated as a function of the $x_i's$. Then it includes $F_{t-1}(x_i)$ in the logistic function to calculate the probability of y_i, and the exponents of the logistic functions become the weights of the training examples (see Figure 2.2).

2.2.4 Alternating decision trees

One succesful and popular way of using boosting is to combine it with decision tree learning as the base learning algorithm [106]. We use boosting both to learn the decision rules constituting the tree and to combine these rules through a weighted majority vote. The form of the generated decision rules is called an *alternating decision tree* (ADT) [104]. In ADTs each node can be understood in isolation.

We explain the structure of ADTs using Figure 2.3. The problem domain is corporate performance prediction, and the goal is to separate stocks with high and low values based

on 17 different variables. The tree consists of alternating levels of ovals (*prediction nodes*) and rectangles (*splitter nodes*) (hence the word "alternating" in the name). The first number within the ovals defines contributions to the prediction score, and the second number (between parentheses) indicates the number of instances. In this example, positive contributions are evidence of high performance, while negative contributions are evidence of corporate financial problems. To evaluate the prediction for a particular company we start at the top oval (0.04) and follow the arrows down. We follow *all* of the dotted arrows that emanate from prediction nodes, but we follow *only one* of the solid-line arrows emanating from a splitter node, corresponding to the answer (yes or no) to the condition stated in rectangle. We sum the values in all the prediction nodes that we reach. This sum represents the prediction score $F(x)$ above, and its sign is the final, or strong, prediction. For example,
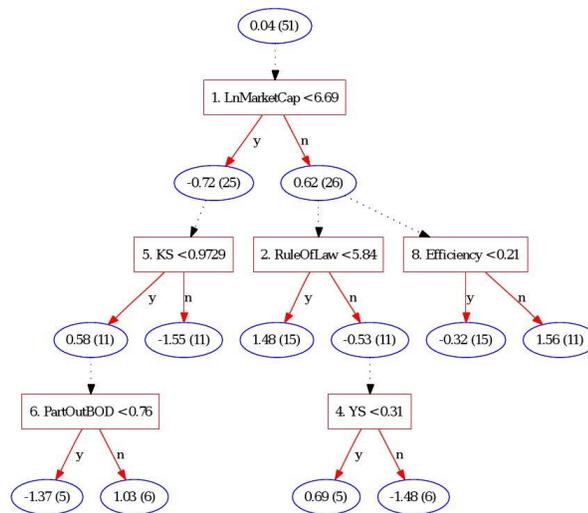
Figure 2.3: LAADR: representative ADT.

suppose we had a company for which LnMarketCap=6, KS=0.86, RuleOfLaw=7.02, and PartOutBOD=0.76. In this case, the prediction nodes that we reach in the tree are

0.042, −0.7181, 0.583, and 1.027. Summing gives a score of 0.9339, i.e., a very confident indicator that the company has a high market value.

This example demonstrates how alternating decision trees combine the contributions of many indicators to generate a prediction. The ADT in the figure was generated by Adaboost from training data. In Adaboost's terms, each prediction node represents a weak prediction rule, and at every boosting iteration, a new splitter node together with its two prediction nodes is added to the model. The splitter node can be attached to any previous prediction node, not only leaf nodes, unless it already has a splitter node attached. Each prediction node is associated with a weight α that contributes to the prediction score of every example reaching it. The weak hypothesis $h(x)$ is 1 for every example reaching the prediction node and 0 for all others. The number in front of the conditions in the splitter nodes of Figure 2.3 indicates the iteration number on which the node was added. In general, lower iteration numbers indicate that the decision rule is more important.

2.3 Measuring efficiency of Latin American banks

We conduct the evaluation of performance of Latin American banks using an efficiency measure because some of the banks under study are not public companies or participate in very illiquid markets. The present banking literature gives significant importance to efficiency evaluation of financial institutions, applying parametric and nonparametric frontier analysis techniques to a specific company as part of an industry, or to a firm's branches. Frontier analysis, based on optimization methodologies, selects the "best practice" firms or areas of a firm, obtains an efficiency score, and recognizes those areas where there is

overuse of inputs or underproduction of outputs within complex operations. Regulators use these techniques [23] to recognize the efficiency gain of a merger between two financial institutions. Frontier analysis can also be used to relate the level of risk that the firm is taking to its overall efficiency, and to establish "benchmarks" for financial institutions based on a "best-practice" frontier. These "benchmarks" can be established by regulators and also by managers who want to assure that the firms they run are competitive nationally or internationally in comparison with the rest of the industry [28].

From an economics point of view, the study of efficiency has been influenced by Leibenstein [161] and his concept of X-efficiency. The economic concept of efficiency includes technical efficiency and also implies allocative efficiency, where the firm must choose an optimal combination of input and output that minimizes costs or maximizes output based on the production technology as well as relative market prices. X-efficiency refers to technical efficiency. Examples of this approach appear in the early nonparametric frontier models [60], and in some of the early parametric frontier models such as in Aigner et al. [4].

The frontier approaches used to measure efficiency can be based on:

1. Nonparametric methods:

   (a) Data Envelopment Analysis (DEA): is a linear programming technique to measure X-efficiency where the set of best-practice (frontier) observations are those for which no other (combination of) firm(s) has as much of every output (given input) or as little of every input (given output). The institutions subject of study receive a score based on how efficient they are in relation to the best-practice institution. The drawback to this method is that it assumes that there is not

random error that leads to overestimating inefficiency.

2. Parametric methods:

   (a) Stochastic Frontier Approach (SFA) or the Econometric Frontier Approach: imposes a functional form such as a cost function and recognizes the random error.

   (b) Thick Frontier Approach (TFA): similar to SFA, but the estimations are based on the best performers in the data as estimators of the best-practice cost function for the whole group.

   (c) Distribution Free Approach (DFA): handles a cost function as the two previous techniques do, but assumes that there is an average efficiency and that the random error term tends to be eliminated.

These efficiency studies in the financial sector have been conducted mainly in the USA [172, 85, 40, 220], and on a smaller scale in Europe [18, 168, 116], Canada [203], Saudi Arabia [5], Tunisia [54], Turkey [240], and India [33].[3] In Latin America, efficiency studies in the banking sector have been scarce [220]. Pastor, Perez and Quesada [189], and Berger and Humphrey [28] have compared international studies on banking efficiency. We are not aware of previous studies that have addressed the relationship between efficiency, and corporate governance structure in Latin America.

DEA measures the performance of each producer relative to the best observed practice among k producers. The DEA frontier is a piecewise linear combination that connects the set of best-practice observations, creating a convex production possibilities set. The rest of the firms that are not in the frontier are ranked accordingly. DEA calculation implies the

---

[3]See Hall [118] for a collection of articles on bank efficiency from 1973 until 1998 for many countries.

minimization of a weighted sum of inputs in relation to a weighted sum of outputs:

$$\min_{u,v} \frac{v^T x_0}{u^T y_0}$$

subject to

$$\frac{v^T x_i}{u^T y_i} \geq 1$$

$$u, v \geq 0$$

where:

i= 1,...,0,...k

(x_0, y_0): input-output vector of the firm that is evaluated

(x_i, y_i): input-output vector of $ith.$ firm in the sample

u: vector of weights given to output

v: vector of weights given to input

This minimization problem can also be expressed as a linear programming problem:

$$\min_{u,v} v^T x_0$$

subject to

$$u^T y_0 = 1$$

$$v^T x_i \geq u^T y_i \quad \text{where} \quad i= 1,...,0,...k$$

$$u, v \geq 0$$

and then as the dual linear programming "envelopment" problem:

$$\max_{\theta,\gamma} \theta$$

subject to

$$X\gamma \leq x_o$$

$$\theta y_o \leq Y\gamma$$

$$\gamma \geq 0$$

X is an n by k input matrix, Y is an m by k output matrix, γ is a k by 1 intensity vector, and x_i and y_i are the columns of the input and output matrix respectively.

θ is a radial measure of technical efficiency. An optimal firm will have its efficiency measure (θ) equal to one. If it is more than one, it can still increase its output with the same unit of input. This version of DEA is output oriented, assumes constant returns to scale and was proposed by Charnes, Cooper, and Rhodes [60] (see also [167]).

We calculate efficiency for the Latin American banking sector using the DEA with output-oriented constant returns to scale as our measure of banking efficiency. As input, we use interest-paying deposits, and non-interest expenses which may include personnel, administrative costs, commissions, and other non-interest operating costs. As output, we use total income which includes interest, and non-interest income. Banks are ranked according to this measure country by country. If a bank shows a great level of inefficiency, a potential agency conflict might be present.

Chapter 3

Using boosting for a board

Balanced Scorecard

3.1   Introduction

The objective of this chapter is to demonstrate how the boosting approach can be used to

define a data-driven board Balanced Scorecard (BSC) with applications to Latin American

markets and S&P 500 companies. We compare our results using Adaboost with logistic

regression, bagging, and random forests. We conduct tenfold cross-validation experiments

on one sample of Latin American Depository Receipts (ADRs), on another sample of Latin

American banks, and on the S&P 500 companies. We find that if the dataset is uniform

(similar types of companies, and same source of information), as is the case with the Latin

American ADRs dataset, the results of Adaboost are similar to the results of bagging,

and random forests. Only when the dataset shows significant non-uniformity does bagging

improve the results. Additionally, the uniformity of the dataset affects the interpretability

of the results.

Using Adaboost, we can generate *alternating decision trees* (ADTs) that explain the relationship between corporate governance variables, performance and efficiency. We also propose an algorithm to build a representative ADT based on cross-validation experiments. The representative ADT selects the most important indicators for the board BSC. Additionally, the thresholds of the representative ADT establish targets or ranges of values to be used in the board BSC. As a final result, we propose an automated strategic planning system combining Adaboost with the BSC for board-level or investment decisions.

The rest of the chapter is organized as follows: section 3.2 presents the Representative ADT algorithm; section 3.3 introduces the data, and variables selected; section 3.4 explains in detail our experiments; section 3.5 presents the results of our forecast; sections 3.6 and 3.7 discuss the results from a methodological, and financial perspective, and section 3.8 presents the conclusions. Chapter 2 already introduced the main methods used in this chapter.

3.2 The Representative ADT algorithm and the Balanced Scorecard

A common complaint about boosting and the alternating decision tree algorithm (see section 2.2.3) is that the ADTs generated do not have a clear interpretation, especially if they are very different as it may happen after cross-validation. Considering these problems we propose an algorithm to calculate a representative ADT that extracts the common features among several trees.

The Representative ADT algorithm selects the most important features of a group of ADTs according to an internal ranking procedure.

Input:

Set of n cross-validation samples

1. For each feature i, select the node j that has the maximum number of cases over all the n cross-validation samples with the threshold k ($freq_{i,j,k}$) and calculate the rank[a] of the feature i:

$rank_{i,j,k} = \frac{avgIter_{i,j,k}}{freq_{i,j,k}}$

where $iter_{i,j,k,m}$ is the iteration when feature i is selected in node j with threshold k, and in sample m. The average iteration is:

$avgIter_{i,j,k} = \sum_{m=1}^{n}(iter_{i,j,k,m})/freq_{i,j,k}$

The thresholds values are simplified using the most significant two digits when written in scientific notation.

2. Select the first V nodes with the lowest $rank_{i,j,k}$ and with $avgIter_{i,j,k} \leq A$. If the node has a feature that already exists in another node, include it only if it is in at least 60% of the ADTs.[b]

3. Put a node under the root, if there is another node with higher priority in the same location.

Output:

Representative ADT

Figure 3.1: The Representative ADT algorithm.

[a]A low value of $rank_{i,j,k}$ shows that it is a more important node because in most cases the feature is included in early iterations.

[b]Based on previous tests, we chose to work with values of V and A equal to seven and ten respectively.

The Representative ADT algorithm looks for the most frequent nodes among a set of ADTs with same positions, features and thresholds. The selected nodes are ranked according to a rank coefficient obtained by the multiplication of the average iteration and the inverse of the frequency of nodes that share the same characteristics. A low value of this coefficient indicates that the node is more important because is present in many nodes of the original set of ADTs and/or appears in the first iterations. The algorithm selects the most important nodes. In case that the algorithm selects two nodes with the same position, the node with lower priority is put under the root (see Figure 3.1).

The features and their relationship of the representative ADT is used to identify key business drivers, strategic objectives and indicators of the BSC (see section 1.2.1.1). In a first step, the features and thresholds of the representative ADT become the indicators and targets of the BSC. If the representative ADT has several levels, then the relationship among the nodes also determine the relationship among the indicators of the BSC. In a second step, the indicators are transformed into objectives of the BSC and of the board strategy map. This second step requires a dialogue among managers where the results of the representative ADTs are reviewed according to the priorities of senior management.

3.3 Data and variable description

The data we used in our experiments are from 1) Latin American ADRs (LAADR), 2) Latin American banks (LABANKS), and 3) S&P 500 companies.

Our first dataset is called LAADR because it is a sample of 51 stocks domiciled in Latin America (LAADR) (Argentina, Brazil, Chile, Colombia, Peru, Mexico, and Venezuela) that have issued ADRs of level II, and III for the year 1998. Level I ADR are least restricted in their required compliance with US regulations, so we have not included them in our analysis. Level II ADRs correspond to foreign companies that list their shares on NASDAQ, AMEX, or NYSE. These companies must fully obey the registration requirements of the SEC, including complying with US GAAP. Level III ADRs refer to foreign companies that issue new stocks directly in the United States. This means that they have the same compliance requirements as a US public company, and are therefore the most regulated. We chose ADRs from countries on the list of emerging markets database (EMDB) of the

International Finance Corporation (IFC).[1]

We obtained the financial information from COMPUSTAT for the year 1998. The information on the value of market capitalization is from CRSP, and is compared with information from the NYSE. We extracted corporate governance information–such as list of directors, executives, and major shareholders–from the proxy statements published at Disclosure, Edgar, and companies' websites for the year 1998. In the case of LAADR, insider ownership is defined as ownership of a company by the CEO, managers, or relatives of the CEO, and members of the board of directors.

Our second dataset is called LABANKS because it is a list of 104 Latin American banks. LABANKS consists of banks headquartered in Argentina, Brazil, Chile, Colombia, Peru, Ecuador, and Bolivia representing about 80% of the total assets of the private sector in the major Latin American countries.[2] We obtained the banks' corporate information from Internet Securities Inc., central bank, regulator, and company websites. We collected financial as well as corporate information similar to that collected for ADRs. Our sample of banks is restricted by the availability of corporate finance information. Most of the financial information is from 2000. A few companies that were merged or disappeared in 1998 were included using the financial statements of 1997. The corporate information is gathered from the period 1998-2000. Considering that the information about ownership structure is relatively stable, we do not foresee any major consistency problem.

Our third dataset is called S&P 500 because it includes the companies that are part of

---

[1]Standard & Poor's acquired this database in January 2000, and it became the Standard & Poor's EMDB.

[2]We were not able to include Venezuela's banks because the President of the Venezuelan Banking Association declined to supply any information to our research team, and asked member banks not to supply any corporate information to us due to the increased risk of kidnapping that its members would be subject to if this information were distributed.

the S&P 500 index. The main sources of data for S&P 500 companies were ExecuComp for executive compensation information, and Compustat North America for accounting information. These two datasets are products of Standard & Poor's. We restricted our dataset to S&P 500 companies with available data from 1992 to 2004. We eliminated observations that did not have enough information to calculate Tobin's Q[3] or incomplete executive compensation information.

The main group of variables that we have selected from these datasets, and that we introduce in the next subsections are related to the principal agent problem because of their effect on company performance and efficiency. We use machine learning techniques to quantify this effect.

3.3.1 Independent variables or features: corporate governance factors and accounting indexes

In the experiments that we describe in the next sections we used the following as dependent variables or as features of the machine learning algorithms.

For the corporate governance variables, in the case of ADRs and banks we include the percentage of insider ownership (T_Insider) because the separation of ownership and control is seen as an opportunity for managers to accumulate wealth at the expense of the shareholders.

The next group of variables that we include for LAADR and LABANKS are those related to the structure of the board of directors (outsiders on the board of directors [PartOutBOD], natural logarithm of the size of the board of directors [LnDIR], and the double role of the

---

[3]Tobin's Q is the ratio of the market value of assets to the replacement cost of assets.

CEO as chairman of the board of directors and manager [ChairmanCEO]). Among these variables, outsiders on the board of directors seem the most important. Fama [95] and Fama and Jensen [98] explain how the separation between control and security ownership can be an efficient structure because professional outside directors may limit the power of managers to expropriate the residual claimants' interest. The size of the board of directors is also a relevant variable, according to Yermack [238] and Fuerst and Kang [107], because the size of the board of directors has an inverse association with firm value in the case of large US industrial corporations. Lipton and Lorsch [164] and Jensen [130] recommend that companies limit board membership to no more than seven or eight members. Additionally, Jensen [130] suggests that companies should separate the CEO role from the chairman role because of the need for independence. If the CEO is also chairman of the board, the dual role may have a negative impact on performance. Even more, Jensen recommends including active investors who hold a large equity or debt position in a company and take part in their strategic decisions. Institutional ownership (InstPart) is another variable that we include because large institutional shareholders act as monitors of managers' actions. Results might be ambiguous if there is insider ownership or hidden investment, because large shareholders may manage the firm for their own benefit only, and not for the benefit of the majority of small shareholders.

For LAADR and LABANKS, we also include corporate governance indicators at the country level according to La Porta et al. [151]: efficiency of the judicial system [EfficiencyJudicialSystem], rule of law [RuleOfLaw], risk of expropriation [RiskOfExpropriation], risk of contract repudiation [RiskOfContractRepudiation], corruption [Corruption], quality of accounting system [Accounting], and legal system [English/French]. Based on these in-

dicators, La Porta et al. [151] found that French-civil law countries have the weakest, and common-law (English) countries have the strongest legal protection of investors. We include these variables because we wanted to separate the effect of country variables from the effect of company variables.

For S&P 500 companies we include insider ownership, and variables related to executive compensation for the top five senior managers. The variables of executive compensation are total compensation for officers (TotalCompExec) and CEOs (totalCompCEO), value of options for officers (OptionAllValExec), CEOs (TotalValOptCEO), and directors (Options-Directors), value of stock options for officers (OptionAllValExec), fees paid for attendance to board of directors meeting (TotalMeetingPay), annual cash paid to each director (PayDirectors), indicator variables to specify if directors are paid additional fees for attending board committee meetings (DcommFee), and annual number of shares granted to non-employee directors (StockDirectors). The discussion about the link between executive compensation and performance is very extensive. Himmelberg et al. [123] and Palia [188] do not find any important association between Tobin's Q as a proxy for performance and equity incentives granted to managers. On the contrary, Hillegeist and Penalva [122] find that those firms with higher options incentives show better performance than the other US firms that were studied. The contradictory results of previous research as well as the importance of executive compensation in corporate governance policies led us to extend our analysis to the study of how total, and stock options compensation for the top five officers, CEOs and directors of a broad sample of US firms affect performance.

We have selected a group of accounting variables for all companies that are well-known for their predictive power, and also are indirect indicators of corporate governance variables.

These accounting variables are: the logarithm of market capitalization (LnMarketCap) for ADRs and S&P 500 companies, and an equity index per country as a proxy for size for Latin American banks[4]; long-term assets to sales ratio (KS) for ADRs and S&P 500 companies, and long-term assets to deposits (KD) for banks for their effect in the reduction of the agency conflict[5]; debt to total assets ratio (DebtRatio) as a capital structure indicator[6]; operating expenses to sales ratio (Efficiency) as an efficiency or agency cost indicator[7]; operating income to sales ratio (YS) as a market power proxy, and to indicate cash available from operations; and capital expenditures to long-term assets ratio (IK)[8] as a proxy for the relationship between growth and the possibility of investing in discretionary projects. A large IK ratio may indicate agency problems if managers are developing new projects that may increase their power, but do not add market value to the company. We use region and sector as indicators of the geographical area and industrial sector in which the company operates.[9] For S&P 500 companies sectors we also include Standard & Poors index membership (SPindex) (see Table 3.1).

---

[4]We used the equity index instead of equity value because efficiency is calculated country by country. We are interested in the effect of the relative size by country on efficiency instead of its absolute value.

[5]Assets can be monitored very easily and they can become collateral either for the development of new projects or to finance new acquisitions.

[6]Harvey et al. [119] find that in emerging market companies with extreme managerial agency costs shareholders benefit from intensively monitored debt.

[7]If operating costs are too high in relation to industry peers or previous years, it might be due to excessive perquisite consumption or other direct agency costs.

[8]Operating expenses to sales ratio and operating income to sales ratio are calculated only for ADRs and S&P 500 companies because these ratios are highly correlated with the efficiency indicator calculated for the banking sector. The capital expenditures to long-term assets ratio is also calculated only for the ADRs and S&P 500 companies.

[9]Sectors of activity for the S&P 500 companies is by the Global Industry Classification Standard, and for ADRs is by the North American Industrial Classification System (NAICS).

Table 3.1: Variables used for corporate governance experiments and for a board Balanced Scorecard. Third column indicates the type of company or dataset where each variable is used: LAADR for Latin American ADRs, LABANKS for Latin American banks, and S&P 500 for S&P 500 companies. Corporate governance variables at the country level are from La Porta et al. [151]. These variables are English, French, RuleOfLaw, Corruption, EfficiencyJudicialSystem, RiskOfExpropriation, RiskOfContractRepudiation, and Accounting.

Indicator	Definition	Type of companies
TobinQ	Tobin´s Q, which is the ratio of the market value to the replacement cost of assets. We use a proxy for Tobin´s Q as the ratio of book value of debt plus market value of common stocks and preferred stocks to total assets	LAADR,S&P 500
PartOutBOD	% outsiders on the board of directors	LAADR,S&P 500, LABANKS
LnDIR	Natural logarithm of board size	LAADR LABANKS
InstPart	% institutional ownership	LAADR LABANKS
T_Insider	% insiders' ownership. In the case of LAADR and the Latin American banks, insider ownership is defined as ownership of a company by the CEO, managers, or relatives of the CEO, and members of the board of directors.	LAADR, S&P 500, LABANKS
ChairmanCEO	1 if CEO is chairman, 0 otherwise	LAADR LABANKS
TotalCompCEO	Total compensation for CEOs. It includes the same items as TotalCompExec (thousands of dollars).	S&P 500
TotalValOptCEO	Value of options for CEOs (thousands of dollars).	S&P 500
TotalCompExec	Total compensation for officers. It includes the following items: salary, bonus, other annual, total value of restricted stock granted, total value of stock options granted (using Black-Scholes), long-term incentive payouts, and all other total (thousands of dollars).	S&P 500
OptionStockValueExec	Value of stock options granted to the executive during the year as valued using S&P's Black Scholes methodology (thousands of dollars).	S&P 500
OptionAllValExec	The aggregate value of all options granted to the executive during the year as valued by the company (thousands of dollars).	S&P 500
DexecDir	Dummy variable to indicate if officer was also a director for the reference year.	S&P 500
OptionsDirectors	Number of options and additional options granted to each non-employee director during the year (thousands).	S&P 500
StockDirectors	Stock shares (including restricted stock) granted to each non-employee director (thousands).	S&P 500
PayDirectors	Annual cash retained paid to each director (thousands of dollars).	S&P 500
TotalMeetingPay	Fees paid for attendance to board of directors meeting (thousands of dollars).	S&P 500
DcommFee	Dummy variable to indicate if directors are paid additional fees for attending board committee meetings.	S&P 500
SPindex	Standard and Poor's index membership. It indicates if companies are part of S&P500 (SP), S&P midcap index (MD), S&P smallcap index (SM), or is not part of a major US index (EX).	S&P 500
LnMarketCap	Natural logarithm of market capitalization, used to measure firm size	LAADR, S&P 500
KS or KD	Ratio of long term assets (property, plant and equipment) to sales (KS) for LAADRs and S&P 500 companies, and to deposits (KS) for LABANKS. This ratio is considered for its effect in the reduction of the agency conflict because these assets can be monitored very easily and they can become collateral for the development of new projects.	LAADR, S&P 500, LABANKS
YS	The ratio of operating income to sales	LAADR, S&P 500, LABANKS
DebtRatio	The ratio of debt to total assets, used as a capital structure variable. Emerging markets are much less liquid than those of developed countries. Hence, firms may give more importance to debt, rather than equity, as a source of capital.	LAADR, S&P 500, LABANKS
Equity index	Index of equity according to country of residence. This is a measure of size applied to LABANKS.	LABANKS
Efficiency	The ratio of operating expenses to sales. This is the efficiency ratio and works as a proxy for market power. It also indicates cash flow available for management use. Similarly, this efficiency ratio may also reveal agency costs or agency conflicts. (This is different from the DEA technical efficiency indicator).	LAADR, S&P 500, LABANKS
IK	The ratio of capital expenditures to long term assets (stocks of property, plant and equipment)	LAADR, S&P 500, LABANKS
AvgParticipation	Measure of ownership concentration. This is calculated as the average of the participation of the three largest shareholders per firm	LAADR,LABANKS
English	If the firm is domiciled in a country whose legal regime is part of the common law or English law legal family according to La Porta et al. (1998)	LAADR,LABANKS
French	If the firm is domiciled in a country that is part of the Napoleonic or French legal family according to La Porta et al.	LAADR,LABANKS
RuleOfLaw	Law and order tradition according to the agency International Country Risk (ICR). Scores are from 0 to 10. Lower values indicate that a country is characterized by less tradition of law and order.	LAADR,LABANKS
Corruption	Indicator of level of government corruption according to ICR. Low levels indicate higher corruption, such as solicitation of bribery by government officials	LAADR,LABANKS
EfficiencyJudicialSystem	Index about the level of efficiency of the legal system according to the agency Business International Corp. Scale is from zero to ten. Lower values correspond to lower efficiency levels.	LAADR,LABANKS
RiskOfExpropriation	Risk of confiscation or nationalization according to ICR. Scale is from zero to ten. Lower values imply higher risks.	LAADR,LABANKS
RiskOfContractRepudiation	Risk of modification of a contract by economic, social or political reasons as defined by ICR. Lower values correspond to higher risks.	LAADR,LABANKS
Accounting	Index based on 1990 annual reports according to their inclusion or omission of 90 items. These items are classified into the following categories: general information, income statements, balance sheets, fund flow statement, accounting standards, stock data and special items. For each country, a minimum of three companies were studied.	LAADR,LABANKS

3.3.2 Dependent variables or measures of company performance

We use Tobin's Q as the measure of performance for ADRs and S&P 500 companies.[10]

Tobin's Q, as a measure of the value of intangibles of a firm, is the ratio of the market value of assets to the replacement cost of assets. This is a measure of the real value created by management.[11] A higher value of Tobin's Q indicates that more value has been added or there is an expectation of greater future cash flow. Hence, the impact of management quality on performance is captured by Tobin's Q. Any difference of Tobin's Q from one indicates that the market perceives that the value of total assets is different from the value to replace their physical assets. The value of internal organization, management quality, or expected agency costs is assumed to explain the difference. Values of Tobin's Q above one indicate that the market perceives the firm's internal organization as effective in leveraging company assets, while a Tobin's Q below one shows that the market expects high agency costs. We use as a proxy for Tobin's Q the ratio of book value of debt plus market value of common stocks, and preferred stocks to total assets.[12] Tobin's Q is a measure of the value of intangibles of a firm.[13]

For Latin American banks, we use an efficiency measure based on DEA (see section 2.3) instead of Tobin's Q because some of the banks under study are not public companies or

---

[10]Tobin's Q is the preferred indicator of performance in corporate governance studies such as in La Porta et al. [152]

[11]The intangibles can also refer to other factors such as intellectual capital or the value of information technology. In this research we control for differences among countries, and economic sectors where companies may have similar technology. So we assume that Tobin's Q reflects management quality.

[12]Several papers [194, 64, 190] indicate that this proxy is empirically close to the well-known Lindenberg and Ross [163] proxy. For international stocks, the information to calculate the Lindenberg and Ross proxy is very limited.

[13]The discrimination between the contribution to performance of top management and other intangibles assets such as intellectual capital requires a more detailed analysis.

participate in very illiquid markets. Additionally, efficiency indicators calculate the agency costs to the firm.[14]

3.4 Experiments

We conducted a logistic regression using Tobin's Q as the dependent variable for LAADR and S&P 500 companies, and the DEA technical efficiency indicator as the dependent variable for LABANKS (see sections 2.2.1 and 2.3). As independent variables we used the variables that we introduced in section 3.3.1.

The logistic regression includes indicator variables for industrial sectors. We calculated the efficiency indicators for each country because of the differences between accounting systems in the countries under study. Hence, efficiency of banks is calculated in relation to their peers in their country.

For the logistic regression analysis and for all the learning algorithms, we eliminated variables that indicated multicollinearity. Multicollinearity is the presence of correlation among dependent variables. For LABANKS the variables eliminated were risk of contract repudiation, legal system, region, corruption, and debt ratio. For LAADR, we eliminated risk of expropriation, risk of contract repudiation, and region. For S&P 500 companies, we eliminated total compensation of officers, and CEOs.

We used Adaboost (see section 2.2.3) to classify stocks above and below the median. In the LAADR sample, the median of the Tobin's Q is very close to one. So, the results can be interpreted as the classification between those stocks with a market value of its

---

[14]Conflicts between managers and shareholders may arise when operating costs increase in relation to a fixed output.

assets above (Tobin's Q greater than one) or below (Tobin's Q smaller than one) its costs of replacement. For LABANKS, the classification is between more efficient and less efficient banks. The results of ADTs must be interpreted as companies with positive scores that have high Tobin's Q or in the case of banks as efficient institutions, while companies with negative scores have low Tobin's Q or are inefficient banks.

We performed tenfold cross-validation experiments to evaluate classification performance on held-out experiments using Adaboost. For LAADR and LABANKS we ran our experiments with 10 iterations. For S&P 500 companies, we run 300 iterations. We used the MLJAVA package, which implements the alternating decision tree algorithm described in Freund and Mason [104].[15] We ranked the variables as an average of the iteration when each variable is selected, weighted by their frequency.

To evaluate the difficulty of the classification task, we compared our method, Adaboost, with random forests (see section 2.2.3) using the software Random Forests V5.0.[16] We ran our experiments with 1,000 trees. We also used four variables for LAADR and LABANKS, and eight for S&P 500 companies randomly selected at each node in order to reduce the test error.

To check for the possibility that the Adaboost results could be improved because of the characteristic instability of Adaboost, we run bagging on top of Adaboost (bagged boosting). We created ten folds for testing and training. We obtained 100 bootstrap replicates of each testing fold. We averaged the score of the bootstraps of each fold to get the estimated class. Finally, we averaged the test error of the ten folds. We also compared ADTs with a single

---

[15]If interested in using MLJAVA, please contact `yfreund@cs.ucsd.edu`

[16]A working version of Random Forests V5.0 can be obtained from ⟨http://stat-www.berkeley.edu/users/breiman/RandomForests/⟩.

tree classifier and with a stumps-averaged classifier trained using boosting. We evaluated the differences between the average of the test error of Adaboost with the test errors of the rest of the learning algorithms using the t-test.

We applied the Representative ADT algorithm (see Figure 3.1) to the S&P 500 companies with stumps-averaged classifier trained using boosting. For LABANKS and LAADR, considering that we had very limited amount of information, we simplified the calculation of representative ADTs using the nodes that were present in at least 60% of the trees.

Finally, we use the main features and thresholds of the representative ADT as indicators and targets of the board BSC (see section 1.2.1.1).

In this research we restrict our analysis to the board BSC for the S&P 500 companies, although we could use a similar methodology to develop the enterprise BSC and the executive BSC. We mostly concentrate on the financial perspective and the internal process perspective because these are the perspectives mainly affected by the corporate governance variables.

3.5 Results

The evolution of the training and test errors are in Figure 3.2. The single tree boosting behaves similarly to Adaboost, and the stumps boosting shows a poorer performance for LAADR, while it shows a better performance for LABANKS during the first 10 iterations. In the case of S&P 500 companies, Adaboost is the dominant algorithm. The receiver operating characteristic (ROC) curve for LABANKS and S&P 500 companies generated using Adaboost shows a larger proportion of true positives versus false positives in comparison
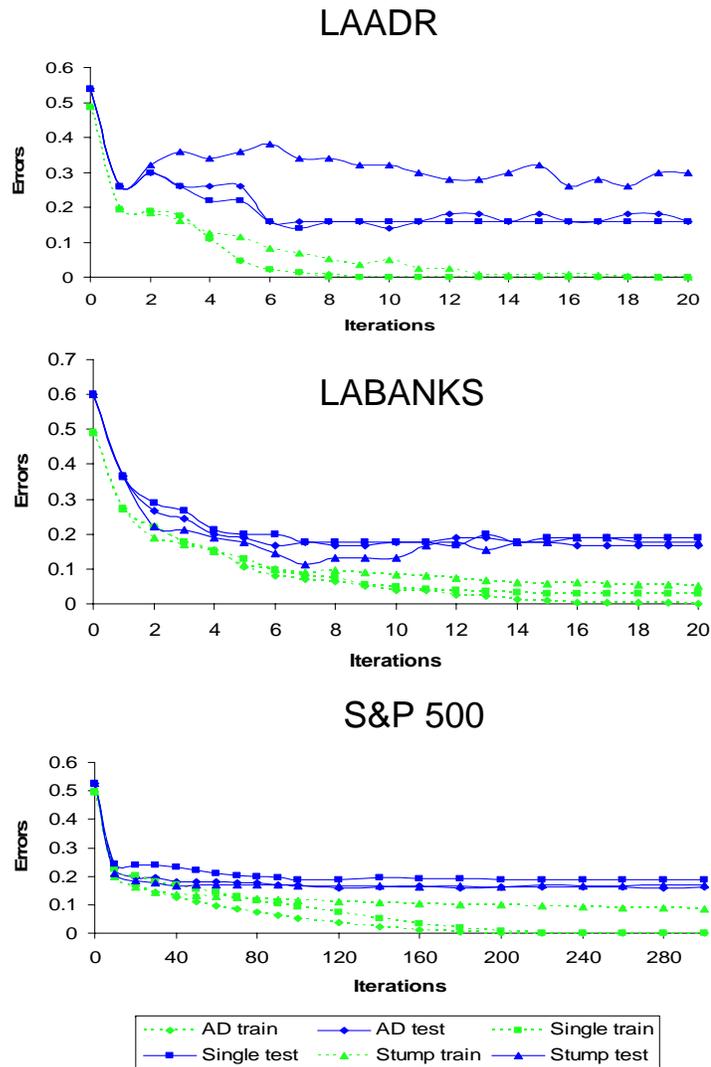
Figure 3.2: Training and test error across algorithms as described in section 3.4 by group of companies.

to the LAADR case (see Figures 3.3). The results of the test errors for the learning algorithms used are shown in Table 3.2. As both our LAADR and LABANKS datasets are very small (51 examples in LAADR and 104 examples in LABANKS), evaluating the statistical significance of the different models and the comparison of their test errors is difficult. Acknowledging these limitations, we present the results of the t-test: there is a significant
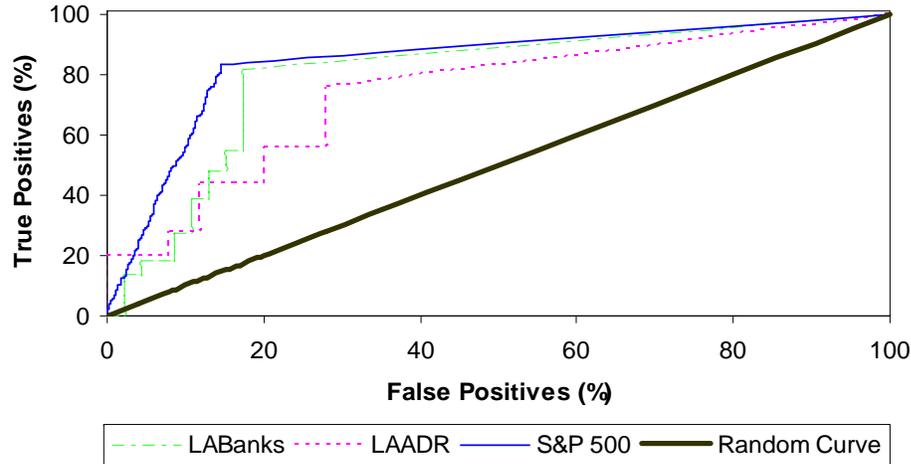
Figure 3.3: ROC curves by group of companies. y-axis presents the percentage of positive observations adequately classified (true positive) and x-axis presents the percentage of negative observations misclassified (false positives).

difference between the test errors of Adaboost and random forests for LAADR, while there are no differences of the test errors for the rest of the tests in both samples. In the case of S&P 500 companies, the test errors of single tree, bagged boosting, and random forests show a significant difference with Adaboost. Random forests presents the lowest test error for S&P 500 companies, and it is followed by bagged boosting.

Most of the S&P 500 subsets defined by the main accounting variables and economic sectors when only the corporate governance variables are included (Table 3.3) show a significant reduction of the test error. The most important reduction of the test error is observed in sectors 1 and 3, and when companies have both market capitalization and efficiency ratio above the median.

Tables 3.4 and 3.5 indicate the importance of each variable according to Adaboost and random forests. The results of both algorithms coincide in terms of what the four most

Table 3.2: Test errors and standard deviations of learning algorithms when all variables are included.

	LAADR		LABANKS		S&P 500	
	Test error	St. dev.	Test error	St. dev.	Test error	St. dev.
Adaboost	14.0%	16.5%	17.8%	9.4%	16.1%	2.1%
Single tree	16.0%	12.7%	17.8%	11.9%	18.7% **	1.8%
Stumps	32.0%	19.3%	13.3%	11.5%	16.8%	2.6%
Bagged boosting	22.0%	23.9%	13.33%	8.66%	14.01% **	0.50%
Random forests	32.0% *	16.87%	16.67%	17.57%	11.50% **	4.56%
Logistic regression	23.7%	18.5%	20.1%	13.2%	16.9%	2.9%
Number observations	51		104		2278	

* 5%, ** 1% significance level of t-test difference between test errors of algorithms and Adabooost

Table 3.3: Test errors and standard deviations of Adaboost for S&P 500 companies aggregated by variables (below and equal or above the median). Only corporate governance variables are used to make the prediction. Number of observations in parenthesis.

Segments	S&P 500	
	Test error	St. dev.
All variables, all sectors (2278)	16.11% **	2.02%
All variables, sectors 1 & 2 (1201)	16.24% **	4.53%
All variables, sectors 3, 4, 5 (1077)	16.53% **	2.76%
Only corporate governance variables (2278)	36.74%	3.50%
Only corporate governance variables, debt ratio < median (1139)	33.62%	1.77%
Only corporate governance variables, debt ratio ≥ median (1139)	33.74%	6.12%
Only corporate governance variables, efficiency < median (1139)	31.40%	4.88%
Only corporate governance variables, efficiency ≥ median (1139)	30.25% **	2.57%
Only corporate governance variables, cap.exp. / L.T.assets < median (1139)	32.23%	4.93%
Only corporate governance variables, cap.exp./ L.T.assets ≥ median (1139)	31.60%	4.31%
Only corporate governance variables, L.T. assets / sales < median (1139)	34.69%	4.38%
Only corporate governance variables, L.T. assets / sales ≥ median (1139)	30.80% *	3.80%
Only corporate governance variables, log market cap. < median (1139)	31.14% **	2.67%
Only corporate governance variables, log market cap. ≥ median (1139)	29.21% **	4.09%
Only corporate governance variables, oper.income / sales < median (1139)	29.91% *	5.50%
Only corporate governance variables, oper.income / sales ≥ median (1139)	30.00% **	5.60%

important variables are.[17] For the LAADR dataset the relevant variables are market capitalization, law and order tradition, % outsiders as directors, and operating expenses to sales ratio. For the LABANKS dataset the relevant variables are long-term assets to deposits ratio, equity index, risk of confiscation, and number of directors. For S&P 500 dataset the

[17]We accept a difference of two in the ranking between both algorithms.

Table 3.4: Results for LAADR and LABANKS. This table reports statistics and results of predicting Tobin's Q for LAADR and efficiency for LABANKS using logistic regression, Adaboost, and random forest. Country corporate governance variables are from [151]. RF: Random forests. z-score for random forests [46] is the raw importance score divided by standard deviation. Q25: 25th. percentile. Q75: 75th percentile. Logistic regression includes indicator variables to control for sector, although they are not included in the table. Variables that do not show any relevance are not included such as legal system, accounting, number of insiders in board of directors, and chairman as CEO. Corporate governance variables are in gray.

| | LAADR | | | | | | | | LABanks | | | | | | | |
| | Statistics | | | | Logit | Boost | RF | | Statistics | | | | Logit | Boost | RF | |
	Q25	Median	Q75	Mean	Odds ratios	Rank	z-score	Rank	Q25	Median	Q75	Mean	Odds ratios	Ranks	z-score	Rank
LnMarketCap (Nat. log market capitalization)	5.44	6.73	7.49	6.57	0.60	1	26	1	(Not used)							
Equity index	(Not used)								0.04	0.16	0.50	0.30	0.14	2	35	1
IK (Capital expenditures/ long-term assets)	0.05	0.08	0.13	0.10		10	6	3	(Not used)							
Efficiency (Operating expenses / sales)	0.10	0.16	0.23	0.16		6	5	4	(Not used)							
YS (Operating income / sales)	0.13	0.23	0.35	0.25	1.07	3			(Not used)							
Debtratio(Debt / total assets)	0.46	0.59	0.80	0.61			2	5	0.89	0.92	3.23	77.96				
KS (L.T. assets/sales)	0.73	1.44	2.20	1.82	1.25	4										
KD (LT ass./deposits)									0.04	0.06	0.10	0.11	46.85	1	33	2
TobinQ (Tobin's Q: performance)	0.91	1.04	1.44	1.38					(Not used)							
EfficiencyJudicialSystem (Effic. legal system)	6.00	6.00	7.25	6.50		9			6.00	6.25	6.75	6.43	0.63	9	12	4
RuleOfLaw (Law and order tradition)	5.35	5.35	7.02	5.82	0.93	2	6	2	2.50	6.32	6.67	5.27	0.94	7	8	6
Corruption (Level of government corruption)	4.77	5.30	5.30	5.21	0.92				5.00	5.18	6.02	5.45				
RiskOfExpropiation (Risk confiscation)	6.95	7.29	7.50	7.09					5.91	6.57	7.50	6.67	0.81	3	10	5
RiskOfContractRepudiation (Contract change)	6.30	6.55	6.80	6.33					4.91	5.18	6.30	5.68				
PartOutBOD (% outsiders as directors)	60.0%	77.0%	87.0%	68.1%	0.43	5	2	6	75.0%	94.4%	100.0%	84.8%	1.27	7	3	9
Avg Participation	(Not used)								0.50	0.93	1.00	0.75	0.57	6	4	8
LnDir (Natural log number directors)	1.95	2.20	2.30	2.08	1.03	7			1.79	2.20	2.40	2.10	0.61	5	12	3
InstPart (% institutional equity ownership)	15.0%	44.0%	71.0%	43.2%	0.77	8			(Not used)							
T_Insiders (% insider's equity ownership)	0.0%	0.0%	2.0%	10.4%			0.02	7	0.0%	0.0%	1.2%	8.8%	0.53	4	4	7

most important variables are operating expenses to sales ratio, operating income to sales ratio, debt to total assets, and long-term assets to sales ratio.

For some variables, there is an important discrepancy among boosting and random forests. In the case of the LAADR dataset, capital expenditures to sales ratio is considered the second and tenth most important variable according to random forests and boosting respectively, while for the LABANKS dataset the efficiency of the legal system is the variables that shows an important difference.

The results of bagged boosting cannot be interpreted in terms of the impact of each variable on performance and efficiency because of the large number of trees generated.

In the case of LABANKS, four variables chosen by Adaboost are ranked among the top

Table 3.5: Results for S&P 500 companies. This table reports statistics and results of predicting Tobin's Q for S&P 500 companies using logistic regression, Adaboost, and random forest. RF: Random forests. z-score for random forests [46] is the raw importance score divided by standard deviation. Q25: 25th. percentile. Q75: 75th percentile. Logistic regression includes indicator variables to control for sector, although they are not included in the table. Sector 3 (consumer staples and health care) according to Adaboost and sectors in general for random forests is the 7th. most important variable. Corporate governance variables are in gray.

	Statistics				Logit	Boost
	Q25	Median	Q75	Mean	Odds ratios	Ranks
LnMarketCap (Nat. log market capitalization)	7.97	8.66	9.48	8.81	0.36	5
IK (Capital expenditures/ long-term assets)	0.14	0.20	0.30	0.24	0.03	3
Efficiency (Operating expenses / sales)	0.12	0.22	0.34	0.24	0.00	1
YS (Operating income / sales)	0.11	0.17	0.25	0.19	0.00	2
Debtratio(Debt / total assets)	0.41	0.57	0.69	0.56	11.18	6
KS (L.T. assets/sales)	0.72	1.03	1.50	1.32	4.26	4
TobinQ (Tobin's Q: performance)	1.43	2.03	3.16	2.74	NA	
T_Insiders (% insider's equity ownership)	0.1%	0.3%	1.8%	3.27%	1.15	
totalMeetingPay (Total payment per meeting)	0	7	11	7.73	1.03	
TotalCompExec (Total compensation)	1572	2673	4782	4345.00		9
optionStockValueExec (Value stock option)	466.9	1174	2687	2677.00	1.00	
payDirectors (Annual cash pay to directors)	19	26	35	27.90	1.04	8
optionAllValExec (Total value options)	543.8	1436	3088	2972.00	1.00	
optionsDirectors (Number options directors)	0	3	10	9.42	1.00	
stockDirectors (Number of stocks directors)	0.00	0.00	0.84	0.71	1.15	7
totalCompCEO (Total compensation CEO)	2012.00	4215.00	8300.00	8166.00		
totalValOptCEO (Total value options CEO)	542.30	1919.00	4741.00	5340.00	1	

five variables according to random forests. Considering the similarity of the most important variables selected by random forests and Adaboost, we discuss the ADTs. ratio, and debt ratio for S&P 500 companies.

3.6 Methodological findings

The tenfold LAADR test errors do not show any significant difference between Adaboost and the other learning algorithms according to the t-test, with the exception of random forests, which shows a higher test error of 32%. For the tenfold LABANKS and S&P 500 cross-

validation, Adaboost has a 17.8% and 16.1% test error respectively. Bagged boosting and random forests reduces the Adaboost test error for LABANKS and for S&P 500 companies. The inverse situation happens in the case of LAADR companies.

It seems that the advantage of using bagging over Adaboost depends on the uniformity of the dataset. LAADR and S&P 500 companies are more uniform samples than LABANKS. LAADR only includes companies of large Latin American countries that fully obey the registration requirements of the SEC, including complying with US GAAP, while LABANKS includes banks of different size and following different accounting standards of Latin American countries. If the dataset is an agglomeration of several different datasets, such as in LABANKS or a combination of companies of diverse sectors such as in S&P 500, bagging can improve the results; however if the dataset is uniform such as in LAADR, bagging does not show any improvement over Adaboost. Therefore, stability is not a property that only depends on the learning algorithm; it also depends of the uniformity of the dataset.

The logistic regression analysis offered some insight about the relevance of the most important variables; however it was not possible to capture the interaction of these variables with the limited amount of data that we had. In contrast, Adaboost helped to rank the variables according to their importance, and also modeled their interaction.

In synthesis, in most of the cases Adaboost performed in a similar way to other learning algorithms such as bagging and random forests, and had the capacity to generate a score that evaluated the effect of corporate governance variables on performance (corporate governance score). Additionally, Adaboost also allowed us to interpret the results because of the limited number of trees that were generated in contrast to the requirements of the other methods

such as random forests.

3.7 Financial interpretation

Comparing the ADTs of LAADR and LABANKS (see Figures 2.3 and 3.4), the main distinctive variable is the size of the company measured by the logarithm of market capitalization *(1. LnMarketCap)*[18] for ADRs and equity index *(1. Equity)* for LABANKS. This result coincides with the classical study of Fama and French [97] in USA, which indicated that size is a key factor to explain the rate of return of stocks. ADRs with market capitalization around or above the median[19] perform better than the rest. Large companies in emerging markets are likely to be oligopolies or monopolies in their area of activity. The efficiency of smaller banks is also affected when there is a high country risk of expropriation *(4. RiskOf-Expropriation)*. However, the performance of LAADR improves in countries with a weak rule of law *(2. RuleOfLaw)* (see Figure 2.3). Large Latin American companies probably perform better in environments with a weak tradition of law and order because of the close family relationships that help them to influence government decisions in their favor. The benefits of these government private sector connections seem to be less important for small size companies *(1. LnMarketCap)*.

In countries with a strong rule of law and order, large companies may still have an important agency conflict that affects their performance if the cash available for operations is too high, as a large operating income to sales ratio *(4. YS)* indicates. An excessive

---

[18]We refer to a specific node of any ADT with its iteration number (first number of the node) and its variable name in italics

[19]We discuss the relative value of a threshold looking at Tables 3.4 and 3.5. For example, the threshold of *(1. LnMarketCap)* in Figure 2.3 is 6.69 and this is a value similar to the median of LnMarketCap (6.73) according to Table 3.4.

amount of cash may allow managers to spend it on projects that benefit them directly instead of increasing the value of their companies. A large operating expenses to sales ratio *(8. Efficiency)* may also indicate an agency conflict. Among the medium and large companies, 58% have an excessive efficiency ratio in relation to the threshold level found by Adaboost. The performance of small and medium size companies improves if the proportion
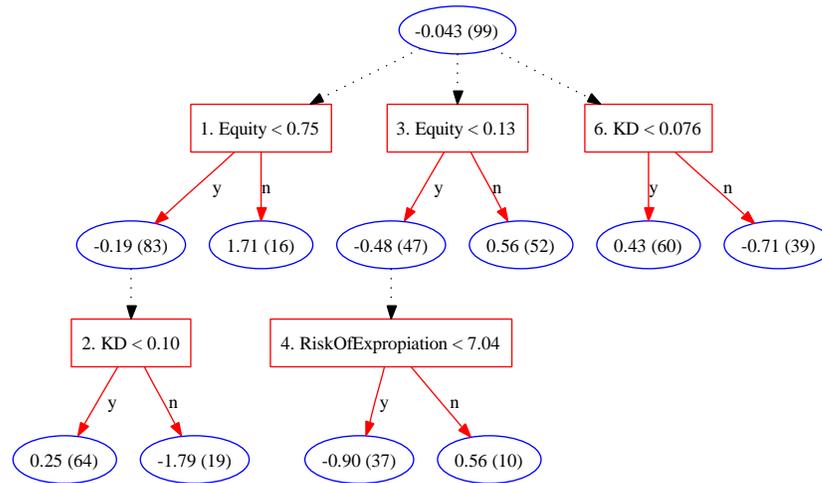
Figure 3.4: Representative ADT of LABANKS. The tree has ovals (*prediction nodes*) and rectangles (*splitter nodes*). The first number within the ovals defines contributions to the prediction score, and the second number (between parentheses) indicates the number of instances. The root node has the minimum prediction score for all cases. The sum of first values in all relevant prediction nodes, including the root, is the prediction score. The sign of prediction score is the prediction. The representative ADT is calculated selecting nodes present in at least 60% of the trees obtained from tenfold cross-validation.

of long-term assets to sales *(5. KS)* is below 0.97 (below the median) for LAADR companies. For Latin American banks, the efficiency improves when the long-term assets to deposits ratio *(6. KD)* is below 0.076 (close to the median). These indicators are important for revealing agency problems. The long-term assets are easy to monitor, and can become collateral to finance new projects. However, if the level of long-term assets is too high, it may indicate inefficiency and overspending.

According to the ADT for LAADR, the composition of the board of directors is important for smaller companies which have a capital sales ratio *(5. KS)* below 0.97. In these cases, the participation of outsiders on the board of directors *(6. PartOutBOD)* above a level of 76% is a relevant factor to improved performance. The finance literature indicates that outside directors supervise managers [231, 209]. Weisbach [231] finds that outsider-dominated boards are more likely to remove CEOs than firms with insider-dominated boards, especially when firms show poor performance.[20] Denis and Sarin [84] find that companies that increase the proportion of outsiders on the board of directors or reduce ownership concentration have above average returns in the previous year. However, Yermack [238], MacAvoy et al. [169], Hermalin and Weisbach [121], and Bhagat and Black [31, 32] find little correlation between composition of board of directors and performance. One possible explanation for these results is that the CEO hires outside directors; hence, directors do not dissent [74]. This hypothesis is reinforced by Core et al. [71], who find that CEO compensation is a decreasing function of the share of inside directors, and is an increasing function of the share of outside directors chosen by the CEO.

Inside directors also play an important role in the board of directors for strategic planning decisions, reviewing functional performance by areas and, in some cases, evaluating if there are important differences between the CEO's perspective and what is happening in the firm on a daily basis.[21] Baysinger and Butler [25] propose that an optimal board of directors should have a combination of inside, independent, and also affiliated directors.

---

[20]In the case of Italy this situation is different. Volpin [227] finds that the probability of turnover and its relationship to performance is lower for executives who are part of the family of the controlling shareholder. Rosenstein and Wyatt [200] find that announcements of outside directors are related to positive excess returns.

[21]Klein [148] found that inside director participation in investment committees correlates with better firm performance.

Bhagat and Black [32] suggest that boards should not be composed only of independent directors because of their findings that board independence does not improve performance, and because inside directors may bring the additional benefits explained above. This may explain why the ADT suggests that the maximum participation of outsiders in the board of directors of LAADR companies *(6. PartOutBOD)* should be 76%.

Insider ownership does not seem to affect the performance of LAADR companies. This can be understood in terms of the information published in the proxy statements of ADRs. These reports are not under the same strict control that the financial statements are. As a result, it is possible that many firms did not include relevant information about managers, ownership structure and board composition due to the need to protect shareholders against potential kidnapping or assault. Hence, only the major shareholders are registered. In the case of LABANKS, according to Adaboost and logistic regression, insider ownership is the third, and fourth most important variable in explaining efficiency. This is not the case for S&P 500 companies according to Adaboost and random forests. Furthermore, the mean of insider ownership for S&P 500 companies (3.27%) is much lower than the same value for LABANKS (8.8%) and LAADR (10.4%) (see Tables 3.4 and 3.5).

Management with a high level of ownership is likely to steer corporate decisions toward its own interests at the expense of corporate interests. This could be true in the case of strong family groups that control a company. These family groups may use their great bargaining power to make corporate decisions that benefit companies where they have a great interest. For example, banks may direct an important part of their loan portfolio to companies where managers or insiders have a significant interest. If the investment is successful, managers benefit. Otherwise, government and depositors assume the loss, as

occurred in the financial crisis of the Andean countries during the nineties. Jensen and Meckling [131] in their classic work described this behavior where large investors as equity holders will benefit when the firm takes an excessive risk because of the potential benefit on the upside, while the other stakeholders, such as the creditors, bear all the risk. Hermalin and Weisbach [121] had already proposed that agency costs increase with ownership, such as in the case of family firms. La Porta el al. [150] also mention that the agency problem in these companies is that the dominant family owner-manager may expropriate minority shareholders. Hence, there is a strong incentive to be a large shareholder in developing countries. However, expropriation is expensive; the cost of expropriation might be bigger than its potential benefit in the case of controlling shareholders, which explains why La Porta et al. [152] find that firms with higher cash-flow ownership by the controlling shareholder have higher valuation measured by Tobin's Q. La Porta et al. also find that firms in countries with better shareholder protections (common law countries) have higher valuation. Large management ownership may avoid the risks of takeovers and reduce the pressure of the board over managers [83].

In the case of LAADR and LABANKS, the limited impact of size of board of directors, the double role of CEO as manager and chairman of the board of directors, and composition of board of directors (percent of outsiders) on performance, and efficiency using logistic regression or Adaboost are findings similar to what previous studies have indicated in USA. Bhagat and Black [32] do not find that board independence leads to improved profitability after controlling for firm size, board size, industry effects, CEO stock ownership, ownership by outsiders, and size and number of outside 5% blockholders.

In our sample of S&P 500 companies, we present the representative ADT when we

include all variables (Figure 3.5), and only the corporate governance variables (Figure 3.6)

for all companies and aggregated by sectors of economic activity.

The representative ADT for all variables has selected mostly accounting ratios. If the

efficiency ratio (operating expenses / sales) *(6.38 Efficiency)* is below 0.17 in the top panel

of Figure 3.5, performance deteriorates. This counterintuitive result is explained because

sector 1 (energy and materials) and 2 (industrials and consumer discretionary)[22] are the

sectors with the largest presence (52.2%) among the S&P 500 companies, and a large

proportion of these companies (84.5%) with an efficiency ratio below 0.17 has a low Tobin's

Q or show poor performance. The representative ADT with all variables for these sectors

(medium panel of Figure 3.5) has an efficiency ratio *(2. Efficiency)* similar to the top panel

of the same figure, while the representative ADT of sectors 3, 4, and 5 (bottom panel of

Figure 3.5) has an efficiency ratio *(5.57 Efficiency)* with a threshold of 0.33 which is a much

higher value than what is observed in the previous two graphs. Considering that companies

of sectors 1 and 2 are mostly of an industrial type or capital intensive, they may have higher

fixed costs than the rest of the industries. So, if the operating expenses to sales ratio is too

low, it may indicate that the operating expenses are not enough to cover an efficient level

of operation, and performance deteriorates.

---

[22]Energy includes energy equipment and services. Materials includes chemical industries, construction materials, containers and packaging, metals and mining, and paper and forest products. Industrials include capital goods; commercial services and supplies; and transportation. Consumer discretionary includes automobiles and components; consumer durables and apparel; hotels, restaurants and leisure; media, and retailing.
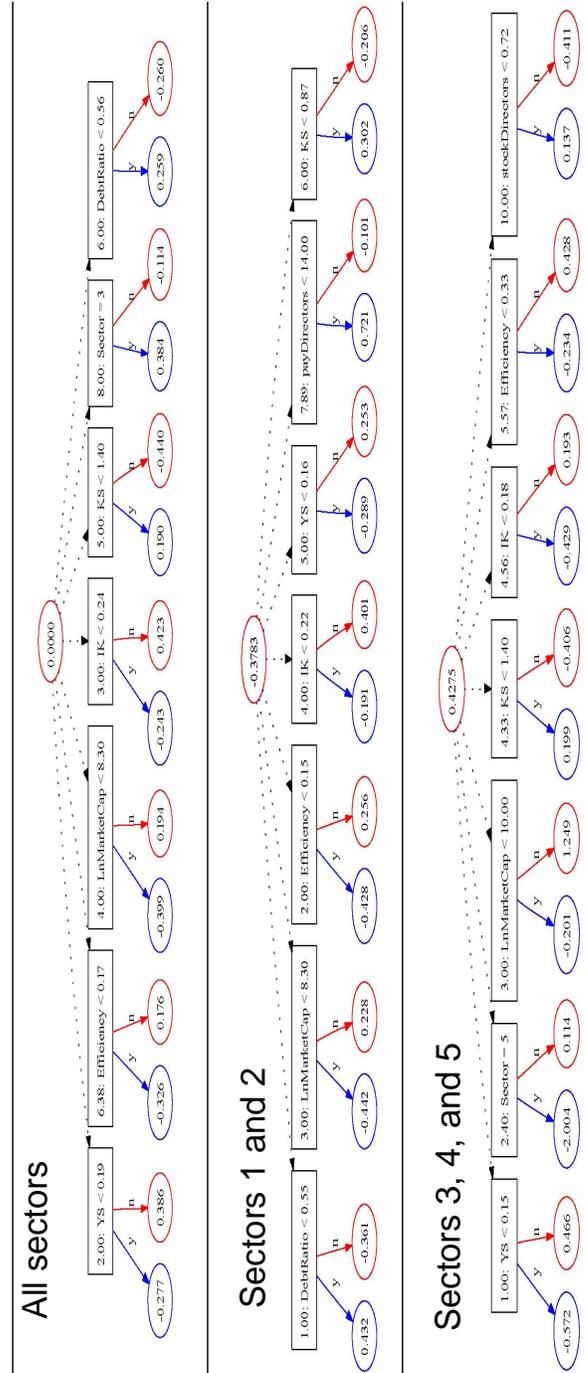
Figure 3.5: S&P 500: representative ADTs with all variables by sectors. This figure includes representative ADTs when all variables are considered (top panel), sectors 1 and 2 (medium panel), and sectors 3, 4, and 5 (bottom panel). First number in each rectangle is average iteration. Figure 3.1 describes the procedure to calculate representative ADTs.

3.7.1 Interpreting the S&P 500 representative ADTs with all variables

At difference of the initial assumptions, there is no indication that in S&P 500 companies (see top of Figure 3.5) a large operating income to sales ratio *(2. YS)* or capital expenditures to long-term assets ratio *(3. IK)* may lead to corporate governance problems, even more IK above the mean improves results. However, the representative ADT establishes a limit to the long-term assets to sales ratio *(5. KS)*. Companies that are in the top quartile according to the long-term assets to sales ratio show a lower performance than the rest of the companies.

The limitation of the ADTs of Figure 3.5 is that the accounting variables dominate, even when we separate our sample between sectors (medium and bottom panel of Figure 3.5). Companies of sectors 1 and 2 show only annual cash pay to directors *(7.89 payDirectors)*, and companies of sectors 3, 4, and 5 show only stock shares granted to directors *(10. stockDirectors)* as relevant corporate governance variables.

In order to capture the effect of corporate governance variables, in the next section we present a representative ADT that includes only these variables.

3.7.2 Interpreting the S&P 500 representative ADTs with only corporate governance variables

The representative ADT for all companies with only the corporate governance variables (top panel of Figure 3.6) captures most of the variables (payDirectors, optionAllValExec, stockDirectors, and TotalValOptCEO) or rules associated with high corporate performance in all companies. This ADT suggests that the compensation policy should have a larger variable component granting more options to top officers (companies in the top quartile)

(2.22 OptionAllValExec), with very broad limits for the value of the options to CEOs (companies in the fourth quartile) *(6.17 totalValOptCEO)*, and with a small cash payment to directors (companies in the first quartile) *(1.33 payDirectors)*. Additionally, this ADT recommends that insider ownership should at least be 10% *(9.33 T_Insiders)*. When we separate the representative ADTs by sectors of economic activity, the compensation policy varies. For sectors 1 and 2, the representative ADT (medium panel of Figure 3.6) suggests a policy that grants very limited cash compensations to CEOs (companies in the first quartile) *(7.75 totalCompCEO)*, and the rest of the compensation should largely be based on options *(3.5 totalValOptCEO)*. For the top officers, the value of the options granted should also be high (companies in the fourth quartile) *(4.8 OptionAllValExec)*. This ADT suggests that the compensation of directors should have a larger components of stocks (companies in the third quartile) *(2. stockDirectors)*, and a smaller annual cash payment (less than the median) *(1. and 4.5 payDirectors)*.

In the case of companies of sectors 3, 4, and 5, the representative ADT (bottom panel of Figure 3.6) indicates that the compensation policy should have a larger amount of stocks *(2. stockDirectors)*, and cash payment to directors *(4.4 payDirectors)*, and options for top executives including the CEO (companies in the fourth quartile) than in sectors 1 and 2 *(1. OptionAllValExec)*. Additionally, the representative ADT shows that insiders ownership *(5.14 T_Insiders)* improves performance.

The discrepancies between the two policy recommendations by sectors is explained by the main business processes of each economic sector. Companies of sectors 1 and 2 take major investment decisions that involve direct participation of the CEO as well as the strategic direction of the board of directors, such as the development of a new factory or
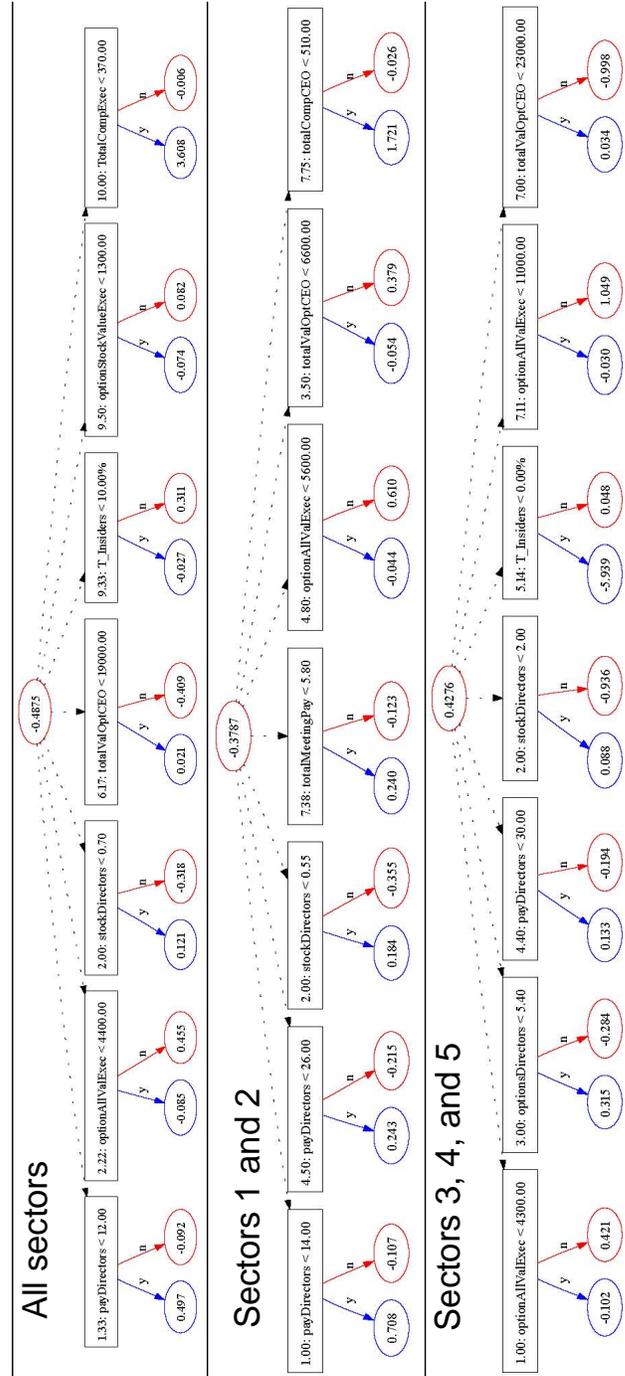
Figure 3.6: S&P 500: representative ADTs with only corporate governance variables by sectors. This figure includes representative ADTs when all variables are considered (top panel), sectors 1 and 2 (medium panel), and sectors 3, 4, and 5 (bottom panel). First number in each rectangle is average iteration. Figure 3.1 describes the procedure to calculate representative ADTs.

the exploration of a new oil region. Companies may motivate with large compensations, especially based on options, the involvement of CEOs and directors. However, once these investment decisions are taken the role of middle managers becomes more relevant, and CEO compensation can be restricted.[23] The profit of companies of sectors 3, 4 and 5, especially in the case of financial services, information technology and telecommunications, are driven by the quality of customer service and continuous technology update. Therefore, their success may significantly depend on the motivation of top officers and middle managers through flexible remuneration (options), while establishing a limit to the options granted to directors.

If the rules suggested by the top panel of Figure 3.6 are effective to improve company performance, the flexible part of executives' compensation might be reduced in some sectors, however it should not disappear as a result of the recent FASB rule which establishes that companies should register as expense any options granted to employees.

3.7.3 The Board Balanced Scorecard and automated planning

We include the variables that the representative ADTs selected for all companies (top panels of Figures 3.5 and 3.6) in the board strategy map and in the board BSC suggested by Kaplan and Nagel [137].

The board strategy map (Figure 3.7) shows the interrelationship between the objectives of each perspective. An important element of the board strategy map and the board BSC

---

[23]Jensen and Murphy [134] consider that there is a major misalignment between corporate performance and compensation paid to executives, especially CEOs. In recent years, there are well-known stories of CEOs who have been paid large compensations regardless of their performance. Michael Ovitz, former president of The Walt Disney Corp., received $140 million as his severance package when he was fired by unhappy shareholders after 14 months at the company. Core et al. [71] also find that CEOs have greater compensation in companies with greater agency problems.
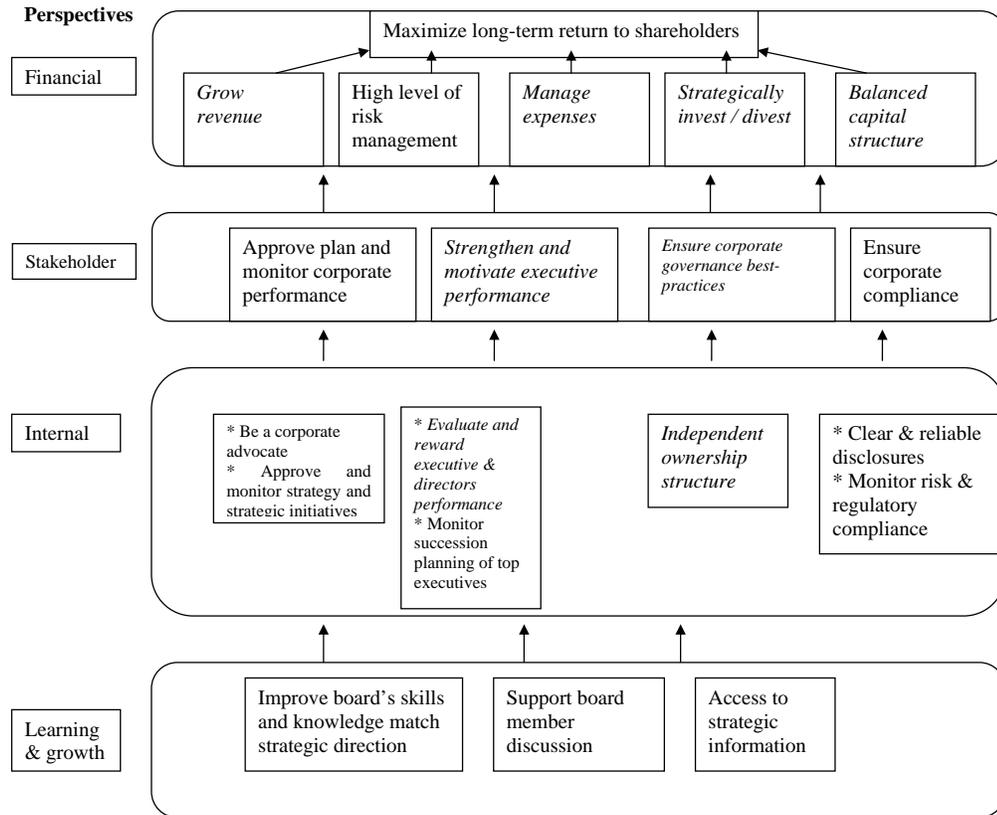
Figure 3.7: S&P 500: representative board strategy map. This figure shows the causal relationship among corporate variables. Adapted from Kaplan and Nagel [137]. Italics are the objectives selected or modified by representative ADTs

is the perspective of "stakeholder" instead of "consumer" as was proposed in the original

BSC. The reason to include the "stakeholder" perspective is that the stakeholders, such as

shareholders and financial analysts, are the consumers or clients of the board of directors.

We have expanded the board strategy map proposed by Kaplan and Nagel [137] to

incorporate new objectives that were consistent with the main variables selected by the

representative ADT. The new objectives that emerged are "Balanced capital structure"

in the financial perspective; "Independent ownership structure" in the internal perspec-

tive, and "Ensure corporate governance best-practices" in the stakeholder perspective. The

Board's strategic objectives		Indicators	Target(s)	Scores		Owners
High-level objectives	**Specific objectives**			**Yes**	**No**	
Financial:						
Maximize long-term	Grow revenues	Operating income / sales (YS)	> 0.19	-0.277	0.386	Executive
total return shareholders	Manage expenses	Operating expenses / sales (Efficiency ratio)	> 0.17	-0.326	0.176	management
	Strategically invest/divest	Long-term assets/sales (KS)	< 1.4	0.19	-0.44	
		Capital expenditures / Long-term assets (IK)	> 0.24	-0.243	0.423	
	Balanced capital structure	Debt ratio	< 0.56	0.259	-0.26	
Internal:						
Evaluate and reward directors'	Reduce fixed payment	Annual cash pay to directors (payDirectors)	< $12K	0.497	-0.092	Compensation
performance	Limit options payment	Number stocks granted to directors (stocksDirectors)	< 700	0.121	-0.318	committee
	to directors	Total value options CEO's (totalValOptCEO)	< $19M	0.021	-0.409	
Evaluate and reward	Increase options payments	Value of all options to officers (optionAllValExec)	> $4.4M	-0.085	0.455	
executive's performance	to top officers	Value stock options to officers (OptionStockValueExec)	> $1.3M	-0.074	0.082	
		Total compensation officers (TotalCompExec)	< $370K	3.608	-0.006	
Independent ownership struct.	Limit insiders' ownership	% insiders' ownership (T_Insiders)	> 10%	-0.027	0.311	Governance comm.

Figure 3.8: S&P 500: representative board BSC. Adapted from Kaplan and Nagel [137]. The board BSC assigns indicators to the objectives selected in the board strategy map. The indicators of the financial perspective are from the representative ADT that includes all variables (Figure 3.5) and the indicators of the internal perspective are from the representative ADT that includes only the corporate governance variables (Figure 3.6). The targets come from the rectangle and the scores from the ovals of the representative ADTs. K is thousands and M is millions.

board BSC (Figure 3.8) incorporates the new indicators and its targets according to the representative ADTs presented in Figures 3.5 and 3.6. The indicators are the most important variables selected by the representative ADTs and their targets are the threshold levels calculated for each variable. Finally, we can say that the representative ADT and BSC complement each other. The representative ADT selects what are the most important features or variables that should be used as indicators and therefore helps to choose the key drivers and objectives of the BSC. Additionally, the representative ADT is able to calculate the targets for every metric. The BSC puts in perspective the findings of the representative ADT. The BSC, as a strategic management system, integrates the four perspectives already described, and offers a framework that connects the variables recognized by the representative ADT in a logical order towards the maximization of shareholders' return.

3.8   Final comments and conclusions

In this research we proposed an algorithm that ranked variables according to their level of importance in the ADTs, and generated representative ADTs with the most important variables. This research showed that Adaboost performed similarly to logistic regression, random forests, and bagging with stable datasets. Additionally, we showed how representative ADTs can be used as interpretative tools to evaluate the impact of corporate governance factors on performance and efficiency. Representative ADTs were particularly useful to understand the non-linear relationship between the variables that affected performance and efficiency.

We demonstrated that the representative ADT is a useful tool to select and establish the relationship among the most important indicators of the BSC. Additionally, the thresholds of the representative ADTs established targets or ranges of values of the indicators that managers could follow to improve corporate performance. With this combined tool, managers can concentrate on the most important strategic issues and delegate the calculation of the targets to an automated planning system supported by Adaboost.

The use of ADTs in finance requires time-series or cross-sectional data in order to calculate meaningful nodes. Indicators that do not have enough information cannot be quantified using ADTs. So, the initial versions of a BSC still require an important participation of the board of directors, middle and senior management. However, as the planning team or the company creates its own database, then the representative ADT can select the relevant indicators and their targets. As the first part of this research showed, Adaboost also worked adequately with small datasets. However, the variance of the test error increased as the size

of the dataset decreased. So, we suggest that companies that use Adaboost to build BSCs use large datasets (industrial surveys or compensation surveys) or build their own internal dataset using the company's historical information.

Comparative regional studies always have a major problem in terms of how to integrate data coming from different sources, and generally with different standards. We saw that this problem was implicit in the LABANKS dataset. We think that the research of emerging markets can be improved by enlarging the dataset and running the learning algorithms in subsets aggregated by regions or corporate governance systems.

Chapter 4

Link analysis and boosting for earnings forecast

4.1 Introduction

The application of networks to social science has a long tradition since the seminal works of Moreno [181] and Milgram [175] about the representation of group dynamics in a sociogram and the "small world" problem. In Milgram's experiment letters are passed from acquaintance to acquaintance. As a result, he showed how apparently distant people are connected by a very short chain of acquaintances. Most of the current literature in social networks is oriented to classify networks, to identify their properties, or to develop new cluster algorithms. Less attention has been devoted to use social networks as a forecasting tool. Recently, link mining has emerged as a new area of research that partially fills this gap. Link mining[1] is a set of techniques that uses different types of networks and their indicators

---

[1]For a recent survey see [109]

to forecast or to model a linked domain. Link mining has had several applications [205] to different areas such as money laundering [147], telephone fraud detection [99], crime detection [211], and surveillance of the NASDAQ and other markets [147, 111]. However, very limited research has been done combining social network indicators with other relevant indicators. In this chapter we propose a link mining algorithm called CorpInterlock that merges social network indicators with any other relevant indicators to forecasting a variable that is mostly associated with the social network. We apply this algorithm for financial forecasting using social networks of corporate directors and financial analysts.

Several networks in the social and natural sciences have been identified to have the properties of a "small world" [230, 19]. We are particularly interested in those organizational studies about the corporate interlock or the social network of directors of major corporations. We refer to the social network among directors as the basic corporate interlock, and the social network among directors and analysts as the extended corporate interlock. In this respect, Davis et al. [76] have found that the basic corporate interlock of the major US corporations (those in the Fortune 500 list) between 1982 and 1999 has the characteristics of a "small world" as described in section **??**sec:smallWorld). A "small world" in the case of the corporate interlock implies that the average distance between firms, between directors, and (if applicable) between analysts is very short. Davis et al. [76] also find that the basic corporate interlock is highly stable, even after major changes in corporate governance. Mintz and Schwartz [177], following Mills' [176] thesis, study how commercial banks have a central position in the corporate interlock because of the participation of the major leaders of US nonfinancial corporations on the banks' boards. The original thesis of Mills is that a small group of business leaders, interconnected by being part of the same boards of directors, is

able to coordinate policies, share practices, and finally control the major corporations. One of the contributions of the "small world" literature in this area is to understand that this connection in the corporate elite is based on the direct link among different actors such as directors, and is not necessarily based on the banking sector or does not require a high level of ownership concentration.[2] Larcker et al. [154] have found that the distance between inside and outside directors, excluding the links when directors are part of the same board, affect CEO's compensation. The interesting aspect of this latter paper is that the authors control for standard economic determinants besides the organizational variables. Very few previous papers have studied the economic effects of corporate interlocks such as their effect on the decision process of: 1. making political contributions [178], 2. poison pills [75], and 3. switching from NASDAQ to NYSE [198].

We use the definition of cumulative abnormal return (CAR) as the return of a specific asset less the average return of all assets in its risk-level portfolio for each trading date, and earnings surprise or forecast error (FE) as the difference between the forecast of financial analysts and the actual earnings at the end of the period of evaluation. The implementation of our algorithm specifically forecast CAR and FE using indicators of the basic and extended corporate interlock and a group of well-known investment variables presented in appendix A. From our perspective, we do not know of any previous research that has used social network indicators combined with economic determinants to forecast CAR and FE. We think that if the corporate interlock plays such an important role in corporate governance, it may also have an impact to forecast CAR and FE.

---

[2]For a dynamic demonstration of the network of directors of the largest American companies see ⟨http://www.theyRule.net/⟩.

The reason that we study the extended social network of directors and analysts is because their relationship is part of what is called the principal agent problem, specifically to the so-called "earnings game" introduced in section 1.2.2.1. Hence, the extended corporate interlock could bring more information to forecast earnings surprise than a basic corporate interlock. Additionally, we expect that statistics of an extended corporate interlock could be able to predict return or earnings surprises better than cumulative abnormal return because of the relationship among directors and analysts that may explain earnings surprises. This methodology could also be applied to a larger class of measures as long as the social network used is relevant to the selected indicator. For instance, a labor economist may use a social network that includes board of directors members and workers leaders in order to evaluate labor productivity or quality of workers benefits. Considering the existence of the "earnings game", our objectives in this chapter are: a) evaluate whether the basic and extended corporate interlock (directors and analysts) of the US stock market has the properties of a "small world" network; b) evaluate the contribution of social network indicators of the basic and extended corporate interlock to predict the trend of FE and CAR, and c) present and test a link mining algorithm that selects the largest strongly connected component of a social network, ranks its vertices using several indicators of distance and centrality, and with other relevant indicators forecast new variables using a boosting algorithm.

The rest of the chapter is organized as follows: section 4.2 describes the "small world" model; section 4.3 introduces the finance literature on earnings surprise; section 4.4 presents a link mining algorithm to forecast the stock market; section 4.5 explains in detail our forecasting strategy; section 4.6 presents the results of our forecast; section 4.7 discusses the results, and section 4.8 presents the conclusions. Appendix A introduces the main

investment indicators used in this research.

4.2 Small World

Watts [229] and Watts et al. [230, 184, 185] have formalized and extended the "small world" model. The relevant aspect of the "small world" model is that it is possible to characterize an undirected graph $G(V, E)$ by its structural indicators where $V = v_1, v_2, ..., v_n$ is the set of vertices, E is the set of edges, and e_{ij} is the edge between vertices v_i and v_j:

- Clustering coefficient: $C \doteq \frac{1}{n} \sum_{i=1}^{n} CC_i$, where:

  - $CC_i \doteq \frac{2|\{e_{ij}\}|}{deg(v_i)(deg(v_i)-1)} : v_j \in N_i, e_{ij} \in E$. Each vertex v_i has a neighborhood N defined by its immediately connected neighbors: $N_i = \{v_j\} : e_{ij} \in E$.

  - $deg(v_i)$ is the degree centrality or degree of a vertex v_i: $deg(v_i) \doteq \sum_j a_{ij}$

  - a_{ij} is an element of the adjacent matrix A of G

  - k is the average degree of the vertices

  - n is the number of vertices in G

- Mean of characteristic path lengths between its vertices: $L \doteq \frac{1}{n} \sum_j d_{ij}$, where $d_{ij} \in D$ and D is the geodesic distance matrix (matrix of all shortest path between every pair of vertices) of G.

In the case of a random network, these structural indicators are $L_{random} \approx \frac{ln(n)}{ln(k)}$ and $C_{random} \approx \frac{k}{n}$.

Using the above indicators, the four properties that characterize a "small world" network are:

I. n is fixed and numerically large ($n \gg 1$).

II. k is fixed so that G is sparse ($k \ll n$), and with a minimum number of potential structures ($k \gg 1$).

III. G is decentralized. So, there is not a single dominant vertex: $k_{max} \ll n$ where k_{max} is the maximal degree.

IV. G must be strongly connected.

C works as a measure of order in G, where if $C >> k/n$, then G is considered locally ordered, while random graphs are not ordered and therefore C_{random} is very small as the above property 2 ($k \ll n$) implies. If a graph is locally ordered or highly clustered, then it should have long characteristic path lengths in order to communicate its different clusters. Obviously, a random graph is not ordered, therefore $C_{random} \ll C$, and $L \approx L_{random}$. As a result, a simple way to evaluate the "small world" properties of a network is if the "small world" ratio ($SW \doteq \frac{C}{L} \cdot \frac{L_{random}}{C_{random}}$) is much larger than one.

Other additional indicators of social networks that we have used in this study are:

1. Closeness centrality (normalized): $C_c(v_i) \doteq \frac{n-1}{\sum_j d_{ij}}$, where d_{ij} is an element of the geodesic distance matrix D [102, 39].

2. Betweenness centrality $B_c(v_i) \doteq \sum_i \sum_j \frac{g_{kij}}{g_{kj}}$. This is the proportion of all geodesic distances of all other vertices that include vertex v_i where g_{kij} is the number of geodesic paths between vertices k and j that include vertex i, and g_{kj} is the number of geodesic paths between k and j [102].

3. Normalized clustering coefficient: $CC'_i \doteq \frac{deg(v_i)}{MaxDeg} CC_i$, where MaxDeg is the maximum degree of vertex in a network [77].

4.3 Earnings surprise

A very well-known phenomenon studied in the accounting and behavioral finance literature is the earnings surprise effect. Earnings surprise or forecast error refers to the difference between financial analysts' predictions and the actual earnings reported by companies. The earnings surprise effect emphasizes how the market reacts more to negative surprises than to positive surprises. Therefore, investors and fund managers have developed many trading strategies around the earnings announcement period and invest significant resources trying to predict earnings surprises. An important source of information for investors are the predictions of more than 3,000 analysts collated in huge databases created by several companies such as IBES International Inc., Zacks Investment Research, and First Call Corporation. These provide investors with a "consensus", or simple average of the market analysts' predictions, which they use to estimate what the market will do.

Other researchers use analysts' predictions for such forecasts, allowing them to make early investment decisions before quarterly announcements. The method they use is linear regression analysis using variables such as the characteristics of companies, and analysts. These studies suggested that analysts' forecasts may have predictive value [187, 214, 30, 171, 2, 192, 191]. Brown et al. [51] standardized a method to calculate the earnings surprise with an indicator that they call "earnings surprise predictor". This indicator is calculated by taking the difference between actual quarterly earnings and the Institutional Brokers' Estimate System (IBES) consensus forecast (calculated just before the quarterly earnings report) and dividing that by the standard deviation of IBES analyst estimates. This "earnings surprise predictor" outperforms the market using a portfolio of S&P 500

companies during the period 1985-1994. [3] We believe that recent developments in the area of machine learning and link mining can contribute to this debate, and especially formalize the study of patterns of behavior for trading and financial forecasting as proposed by the behavioral finance approach. This approach sustains that markets are inefficients and move on individual biases or behavioral patterns [218]. In this chapter we propose a link mining algorithm that improves the earnings and return predictions combining well-known corporate variables with metrics of a social network of directors and analysts. The association among directors and financial analysts may allow companies to adjust earnings to the forecast of financial analysts. However, this relationship is not easily captured by linear regression analysis. Link mining algorithms may explain the relationship among organizational and economic variables, and therefore improve stock price prediction.

Earlier studies on analysts and earnings surprise show at least two types of major variables that are typical of these studies. First, researchers have quantified companies' characteristics or actions, since companies' changes have been shown to relate to analysts' recommendations [213] [4]. Secondly, there are variables which quantify analysts' predictions, such as the quality of their recommendation [236, 90, 20]; the accuracy of their past predictions [50]; the revisions they make [193]; the company variables they use [100, 213, 149]; the career moves of analysts [125]; the timing of analyst's predictions [5]; the herding behavior of analysts [67]; and the information content of analysts' reports [17].

---

[3]For a detailed list of references about the academic use of analysts' predictions see [49].

[4]Beckers et al. [27] find that after the European integration in 1992, country differences is not a relevant factor to explain earnings forecasts differences between analysts, however sector is still an important factor

[5]Ivkovic and Jegadeesh [128] find that the information content of upward earnings forecast revisions and recommendation upgrades increase near the earnings announcement date, while they are less informative in the week that follows this date. This situation is not observed for recommendation downgrades and downward revisions.

Several studies have evaluated investment strategies that follow consensus recommendations of analysts. A particularly sophisticated model was developed by the company Starmine, which ranks analysts and makes its predictions "Smart estimate" using the forecasts of the most highly ranked analysts. Barber et al. [20] find that after taking transaction costs into account, the high-trading level of strategies that follow consensus recommendations of analysts do not give a consistent return greater than zero. A similar result is obtained by Mikhail et al. [173] even after taking into account analysts' prior performance. They recommend that those investors that still want to follow analysts' recommendations may benefit if they use the forecasts of highly ranked analysts with at least five-years of superior performance in rankings surveys such as those collected by The Wall Street Journal. Jegadeesh et al. [129] reported that analysts from sell-side firms recommend mostly "glamour stocks" (characterized by positive momentum, high growth, high volume, and relatively high prices); however, investors that blindly follow a strategy that invests in these recommended stocks may not obtain positive returns because investment in these stocks also requires favorable quantitative indicators (i.e. high value and positive momentum).[6]

4.4 CorpInterlock: a link mining algorithm

In this chapter we propose a link mining algorithm called CorpInterlock (see Figure 4.1). This algorithm selects the largest strongly connected component of a social network and ranks its vertices using several indicators of distance and centrality. These indicators are merged with other relevant indicators in order to forecast new variables using a boosting

---

[6]Abarbanell [1] finds that analyst's forecasts do not completely integrate the information of past prices changes; additionally, Abarbanell et al. [2] find that the under-reaction of analysts to recent earnings is only a partial explanation for the under-reaction of stock prices to earnings.

algorithm.

Input:
Two disjoint nonempty sets V_{11} and V_{12}.

1. Build a bipartite graph $G_1(V_1, E_1)$ where its vertex set V_1 is partitioned into two disjoint sets V_{11} and V_{12} such that every edge in E_1 links a vertex in V_{11} and a vertex in V_{12}.

2. Build a one-mode graph $G_2(V_2, E_2)$ in which there exist an edge between v_i and v_j: v_i, $v_j \in V_2$ if and only v_i and v_j share at least a vertex $u_i \in V_{12}$. The value of the edge is equal to the total number of objects in V_{12} that they have in common.

3. Calculate the largest strongly connected component of G_2 and call it $G_3(V_3, E_3)$.

4. Calculate the adjacency matrix A and geodesic distance matrix D for G_3. a_{ij} and d_{ij} are the elements of A and D respectively. The mean of all the distances d_{ij} is L.

5. For each vertex $v_i \in V_3$ calculate the following social network indicators:

   - Degree centrality: $deg(v_i) = \sum_j a_{ij}$

   - Closeness centrality (normalized): $C_c(v_i) \doteq \frac{n-1}{\sum_j d_{ij}}$

   - Betweenness centrality: $B_c(v_i) = \sum_i \sum_j \frac{g_{kij}}{g_{kj}}$, where g_{kij} is the number of geodesic paths between vertices k and j that include vertex i, and g_{kj} is the number of geodesic paths between k and j.

   - Clustering coefficient: $CC_i = \frac{2|\{e_{ij}\}|}{deg(v_i)(deg(v_i)-1)} : v_j \in N_i,\ e_{ij} \in E$

   - Normalized clustering coefficient: $CC'_i = \frac{deg(v_i)}{MaxDeg} CC_i$, where MaxDeg is the maximum degree of vertex in a network

6. Merge social network indicators with any other relevant set of variables for the population under study and generate test and training sample.

7. Run machine learning algorithm with above test and training samples to predict Y variable.

Output:
Prediction of Y.

Figure 4.1: The CorpInterlock algorithm

4.4.1 Application to forecasting earnings surprise.

We used the CorpInterlock link mining algorithm (see Figure 4.1) to build a bipartite social network where the nodes of the partition V_{12} representing the directors and analysts are connected to nodes of the partition V_{11} representing companies that they direct or cover. This social network is converted into a one-mode network where the vertices are the companies and the edges are the number of directors and analysts that every pair of companies have in common. This is the extended corporate interlock. The basic corporate

interlock is calculated in the same way using only directors. The algorithm merges a group of investment variables presented in appendix A and a group of social network statistics obtained from the basic or extended corporate interlock. Finally, the algorithm predicts FE and CAR using a machine learning algorithm such as boosting. Additionally, for the basic and extended coporate interlock we calculate the "small world" ratio (see section 4.2) to evaluate if the extended corporate interlock is a "small world" network and compare with previous studies of corporate interlock; however, this ratio is not required to calculate the distance and centrality indicators.

We consider that this financial application of the CorpInterlock algorithm is appropriate because the increasing importance of organizational and corporate governance issues in the stock market requires the extraction of indicators from the extended and basic corporate interlock and merges them with more traditional economic indicators in order to forecast CAR and FE. The indicators calculated by the CorpInterlock algorithm captures the power relationship among directors and financial analysts as follows:

1. Degree centrality: directors and analysts of a company characterized by a high degree or degree centrality coefficient are connected among them through several companies.

2. Closeness centrality: directors and analysts of a company characterized by a high closeness centrality coefficient are connected among them through several companies that are linked through short paths.

3. Betweenness centrality: directors and analysts of a reference company characterized by a high betweenness centrality coefficient are connected among them through several companies. Additionally, the reference company mentioned above has a central role because it lies between several other companies, and no other company lies between

this reference company and the rest of the companies.

4. Clustering coefficient: directors and analysts of a company characterized by a high clustering coefficient are probably as connected among them as it is possible through several companies.

Each of the above measures show a different perspective of the connection between directors and analysts as described in the "earnings game" where the earnings forecast of analysts are aligned with management's expectations. Hence, we can include them in a decision system to forecast FE and CAR.

Additionally, we hypothesized that boosting will be able to detect a combination of economic and organizational variables to optimize the earnings surprise and cumulative abnormal return prediction.

Dhar and Chou [86] have already compared the predictive accuracy of tree-induction algorithms, neural networks, naive Bayesian learning, and genetic algorithms to classify the earnings surprise before announcement. They used a definition of earnings surprise or forecast error that we have also adopted in this research:

$$FE \doteq \frac{\text{CONSENSUS}_q - \text{EPS}_q}{|\text{CONSENSUS}_q| + |\text{EPS}_q|}$$

where CONSENSUS_q is the mean of earnings estimate by financial analysts for quarter q, and EPS_q is the actual earnings per share for quarter q. FE is a normalized variable with values between -1 and 1. Additionally, when CONSENSUS_q is close to zero and EPS_q is not, then the denominator will not be close to zero.

4.5 Experiments

We restricted our experiments to companies that are part of the US stock market. We obtained the price series from the Center for Research in Security Prices (CRSP), the accounting variables from COMPUSTAT[7], the list of financial analysts and earnings forecast or consensus from IBES, and the list of directors from the Investor Responsability Research Center. The list of directors exists only on an annual basis for the period 1996 - 2003. This restricts our analysis to this period. The number of companies under study changes every year. The minimum and maximum number of companies included in our study are 2,900 for 2002 and 4,018 for 1998.

We applied the CorpInterlock algorithm described in Figure 4.1 using the softwares EMT [215] and Pajek [77] to obtain the basic and extended corporate interlock. We computed the investment signals and a group of the social network statistics introduced in section 4.2 [average distance, betweenness centrality, closeness centrality, degree centralization, degree, and clustering coefficient (normalized and unnormalized)] of the basic and extended corporate interlock. We merged our accounting information, analysts' predictions (consensus) and social networks statistics using quarterly data, and selected the last quarter available for every year. Most of the fundamental and accounting variables used are well-known in finance literature and [129] demonstrated that these variables are good predictors of cross-sectional returns (see appendix A for an explanation of the variables used).

We forecasted two different trends: FE and CAR. In both cases, we labeled an instance as 1 if the trend was positive and -1 otherwise. We calculated the label of CAR using

---

[7]COMPUSTAT is an accounting database managed by Standard & Poor's.

the cumulative abnormal return of the month following the earnings announcement. We computed FE using the predictions of the analysts available 20 days before the earnings announcement as fund managers may suggest [86]. Fund managers take a position, short or long, a certain number of days before the earnings announcement and, according to their strategy, they will liquidate the position a given number of days after the earnings announcement. If fund managers know the trend of FE or CAR, they make take a position according to their expectations; however they do not need to know exactly what the future stock price is going to be. They profit when the market moves in the direction expected, and above a certain threshold, even though the market movement might not be in the exact amount forecasted. For this reason, the emphasis of this chapter is in the improvement of the prediction of the trend of FE and CAR–and not in their value–with the inclusion of the extended corporate interlock information.

We implemented the CorpInterlock algorithm using Logitboost with decision stumps. To evaluate the difficulty of the classification task, we compared our method with random forests [45], and logistic regression. The latter algorithm was our baseline method. We implemented ADTs and Logitboost with 50 iterations, and random forests with 100 trees and five features[8] using the Weka package [235]. We generated seven training models for each learning algorithm on an accumulative rolling basis, each one for every year from 1996 to 2002. The training data was accumulated from year to year. We tested our results with the information of the following year. We split our test sample in two sets per year. In total we had 14 test sets. The test errors that we obtained were the result of averaging our results over the 14 sets.

---

[8]We implemented random forests with five features in order to optimize its performance.

As we are including all the companies that are part of the US stock market for every year, if a company is listed during our period of evaluation it becomes part of our sample. Likewise, if a company is delisted during our period of evaluation, then this company is not anymore part of our sample. Therefore, we avoided the very common survivorship bias. We eliminated companies that did not have earnings or CAR information.

We ran linear regressions using FE and CAR as dependent variables, and evaluated the importance of the variables listed in appendix A for the model. We tested our model for heteroscedasticity and multicollinearity using the score test for non-constant error variance, and the variance inflation factor (VIF) respectively. We did not find heteroscedasticity or multicollinearity in our sample. In any case, if there was any multicollinearity, it was overcome by boosting's feature selection capability.[9]

We split the presentation of our results before and after 2001 because during this year there were a significant numbers of IPOs, mergers, and acquisitions that were affected by the presence of analysts; it was also the year when the market turned down after the internet "bubble", and also after this year the market became more regulated. In May 10, 2002 the Securities and Exchange Commission (SEC) approved the rule 2711 "Research Analysts and Research Reports" issued by the National Association of Securities Dealers (NASD), and the rule 472 "Communications with the Public" issued by the New York Stock Exchange (NYSE). These rules establish that no research analyst might be controlled by a firm's investment banking department. It also shows that the company that is subject of the report can review the report only for factual accuracy checks. In October 23, 2000,

---

[9]The regression is heteroscedastic if the variance of the residuals is not constant across observations. Multicollinearity is the presence of correlation among dependent variables.

the SEC issued Regulation Fair Disclosure (FD). This regulation requires that companies disseminate material information evenly, without giving any preferences to any investors or analysts. Critics of this regulation indicated that market volatility may increase and the volume of information disseminated in the market will be reduced. However, Lett et al. [160] neither find any significant increase in volatility, nor an increase in certain components of the bid-ask spread around new releases as a result of Regulation FD.

4.6 Results

The "small world" ratio for the basic and extended corporate interlock is much larger than one according to Table 4.1. Hence, both corporate interlocks are clearly considered to be of the "small world" type as Davis et al. [76] found for the Fortune 500 companies. The reason why Davis et al. chose to work with the Fortune 500 companies was to study the US corporate elite. To compare our results with the previous studies, we also present the "small world" ratios of the corporate interlocks of the S&P 500 companies[10] and we still find that these ratios are larger than one, and follow a similar path–even though about three times smaller–as in the case of the complete US stock market.

One of the most important facts that appears among the social network indicators in Table 4.1 is that while the indicators of the basic corporate interlock were very stable between 1996 and 2003, some of the indicators of the extended corporate interlock show a great variation. There is a very important increase in the degree during the year 2001 and then it drops significantly during the years 2002 and 2003. This also reduces the "small

---

[10]The companies in the S&P 500 index are not necessarily the same than the one selected by the Fortune 500. However, both represent the largest and richest US companies.

	N	N component	CC	Degree Centrality	Betweenness Centrality	Closeness Centrality	Degree	L (Distance)	SW (US stock market)	SW (S&P 500)
1996	3676	3573	0.601	0.010	0.001	0.324	51.855	3.129	27.404	2.117
1997	3953	3848	0.591	0.009	0.001	0.323	49.922	3.135	30.689	2.253
1998	4018	3928	0.566	0.009	0.001	0.319	46.700	3.180	32.240	2.457
1999	3960	3867	0.610	0.019	0.000	0.355	89.196	2.877	16.908	2.153
2000	3804	3692	0.582	0.015	0.001	0.342	69.563	2.978	20.076	2.112
2001	3504	3388	0.676	0.058	0.000	0.401	216.979	2.587	6.170	1.509
2002	2900	2677	0.506	0.008	0.001	0.275	24.167	3.739	37.129	5.247
2003	2932	2760	0.545	0.008	0.001	0.280	25.055	3.653	40.461	6.695

(a) Extended Corporate Interlock for US stock market

	N	N component	CC	Degree Centrality	Betweenness Centrality	Closeness Centrality	Degree	L (Distance)	SW (US stock market)	SW (S&P 500)
1996	3676	993	0.222	0.007	0.003	0.238	7.333	4.332	24.043	8.037
1997	3953	1044	0.222	0.006	0.003	0.231	7.011	4.450	26.552	6.925
1998	4018	1217	0.201	0.005	0.003	0.227	6.886	4.545	28.825	6.493
1999	3960	1272	0.199	0.005	0.003	0.227	6.920	4.526	29.859	6.839
2000	3804	1256	0.209	0.005	0.003	0.224	6.674	4.593	32.213	7.523
2001	3504	1349	0.216	0.004	0.003	0.218	6.506	4.730	36.445	7.292
2002	2900	1097	0.210	0.006	0.003	0.227	6.689	4.544	27.872	8.155
2003	2932	1129	0.190	0.005	0.003	0.225	6.608	4.559	26.565	8.770

(b) Basic Corporate Interlock for US stock market

Table 4.1: Social network indicators for the corporate interlock of total US stock market. CC is clustering coefficient. Last two columns are "small world" ratio for US stock market and S&P 500 companies respectively

world" ratio.

The implementation of CorpInterlock using Logitboost shows a significantly lower test error than its implementation using logistic regression, our baseline algorithm. Additionally, the Logitboost implementation shows similar test errors than the implementation of CorpInterlock using random forests (see Table 4.2). Based on these results, we decided to limit our analysis to the implementation of the CorpInterlock algorithm with Logitboost.

Table 4.3 shows that for the prediction of the trend of CAR for the period 1997-2001, the extended corporate interlock has a test error significantly lower than the test error with the basic corporate interlock. However, there is not a significant difference of the means of the test error to predict the trend of FE, even though overall the test error for the prediction of FE trend decreased to 20% from a 31% test error using logistic regression.

The regression analysis indicates that our model explains FE much better than CAR in the period 1997-2003. Table 4.4 has an adjusted R-square of 0.42-0.43 for FE while this

	Total			CAR		FE		
	Mean		St. Dev.	Mean	St Dev	Mean		St Dev
Logistic regression	40.15%	**	9.74%	48.99%	4.55%	31.32%	**	3.30%
Random forests	34.46%		14.51%	48.52%	3.86%	20.40%		2.02%
ADTs	35.24%		14.02%	48.86%	3.08%	21.63%	**	2.50%
Logitboost	34.71%		14.61%	48.90%	3.40%	20.51%		2.37%

Table 4.2: Mean of test errors for learning algorithms by CAR and FE. ** and * represent significance levels of 1% and 5% respectively for the paired t-test of the difference between test errors among each algorithm and Logitboost.

Indicators	Variables used	1997-2003		1997-2001		2002-2003
CAR	Extended corporate interlock	49.58%	*	50.38%	*	49.70%
	Basic corporate interlock	48.00%		46.67%		48.10%
FE	Extended corporate interlock	20.84%		19.68%		20.43%
	Basic corporate interlock	21.13%		20.00%		20.59%

Table 4.3: Mean of test errors using Logitboost for US stock market. * represents significance levels of 5% for the paired t-test of the test errors difference between the extended corporate interlock and the basic corporate interlock by CAR and FE.

value is about 0.022-0.035 for the prediction of CAR. However, in all cases when all the variables are used, the p-value of the F-statistics is highly significant indicating that the model has explanatory power. This result is also confirmed by the fact that the prediction of the trend of FE has a test error (20.5%) significantly lower than for the case of CAR (48.9%).

We could eliminate Table 4.4, and the basic results of this research will not be altered; however, we included this table because social scientists are used to analyzing the data using linear regressions. So we wanted to show the results using only linear regression, and the benefits of a link mining algorithm.

	Directors & analysts		Directors		Directors & analysts		Directors	
Variables	Economic var.	All	Economic var.	All	Economic var.	All	Economic var.	All
CAR1	0.031	0.032	0.008	0.009	-0.008	-0.008	0.005	0.005
	(3.815)***	(3.874)***	(0.750)	(0.875)	(-1.303)	(-1.355)	(0.776)	(0.750)
CAR2	0.000	0.000	0.015	0.017	0.000	0.000	0.005	0.005
	(0.060)	(0.037)	(1.529)	(1.692).	(0.047)	(-0.034)	(0.796)	(0.790)
SIZE	0.006	0.006	0.005	0.005	-0.002	-0.004	-0.001	-0.002
	(3.989)***	(3.425)***	(2.870)**	(2.163)*	(-1.991)*	(-3.419)***	(-0.827)	(-1.022)
FREV	-0.005	-0.004	-0.063	-0.064	-0.003	0.000	0.004	0.004
	(-0.182)	(-0.140)	(-1.173)	(-1.182)	(-0.151)	(0.020)	(0.094)	(0.103)
LTG	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
	(-1.595)	(-1.671).	(-0.140)	(-0.006)	(-1.344)	(-1.363)	(-0.769)	(-0.558)
SUE	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000
	(-0.190)	(-0.271)	(0.368)	(0.344)	(-0.434)	(-0.394)	(0.086)	(0.129)
SG	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(1.738).	(1.693).	(0.917)	(0.899)	(0.342)	(0.400)	(-1.109)	(-1.088)
TA	-0.008	-0.008	-0.001	-0.002	0.007	0.006	0.011	0.011
	(-0.486)	(-0.455)	(-0.067)	(-0.092)	(0.580)	(0.496)	(0.756)	(0.782)
CAPEX	-0.011	-0.012	-0.042	-0.041	0.002	0.001	0.047	0.049
	(-0.284)	(-0.298)	(-0.791)	(-0.768)	(0.061)	(0.035)	(1.261)	(1.304)
BP	0.000	0.000	-0.001	-0.001	0.000	0.000	-0.002	-0.002
	(-0.105)	(-0.142)	(-1.584)	(-1.536)	(-0.729)	(-0.738)	(-3.691)***	(-3.670)***
EP	0.004	0.004	0.002	0.002	0.000	0.000	-0.003	-0.003
	(2.045)*	(2.011)*	(0.727)	(0.716)	(-0.387)	(-0.371)	(-1.999)*	(-2.020)*
ANFOR	0.002	0.002	0.003	0.003	-0.002	-0.002	-0.002	-0.002
	(2.004)*	(1.787).	(2.730)**	(2.716)**	(-2.650)**	(-2.378)*	(-3.524)***	(-3.530)***
ANFORLAG	0.002	0.002	0.002	0.002	0.000	0.001	0.001	0.001
	(2.825)**	(2.958)**	(2.303)*	(2.333)*	(0.783)	(1.053)	(0.929)	(0.907)
FELAG	0.628	0.628	0.633	0.633	0.056	0.056	0.050	0.050
	(67.966)***	(67.966)***	(53.766)***	(53.741)***	(8.422)***	(8.519)***	(6.043)***	(6.057)***
CONSENSUS	-0.001	0.000	-0.001	-0.001	0.000	0.000	0.000	0.000
	(-0.625)	(-0.569)	(-0.885)	(-0.873)	(0.409)	(0.432)	(0.049)	(0.057)
B_c		0.042		0.042		-0.063		-0.011
		(1.199)		(1.560)		(-2.497)*		(-0.589)
C_c		0.206		-1.150		1.252		1.170
		(0.125)		(-0.917)		(1.065)		(1.338)
deg		-0.267		0.155		0.182		0.096
		(-3.440)***		(1.086)		(3.287)**		(0.965)
CC'		0.000		0.000		0.000		-0.001
		(1.814).		(-0.072)		(0.762)		(-1.191)
Adj.R square	0.43	0.431	0.421	0.421	0.02	0.022	0.035	0.035
p-vale (F-stat.)	0.00	0.000	0.000	0.000	0.00	0.000	0.000	0.000
	(a) Forecast error (FE)				(b) Cumulative abnormal return (CAR)			

Table 4.4: Regression results of two models using FE (panel a) or CAR (panel b) as the dependent variable and the following independent variables: 1. only economic variables, and 2. economic and social network variables. Models include intercept and dummy variables to control for economic sector of activity. Non relevant variables are not included. Economic variables are cumulative abnormal return for the preceding six months (CAR1) and for the second preceding six months (CAR2) since the earnings announcement day; natural logarithm of market capitalization (SIZE); analysts earnings forecast revisions to price (FREV); mean of analysts' long-term growth forecast (LTG); standardized unexpected earnings (SUE); sales growth (SG); total accruals to total assets (TA); rolling sum of capital expenditures to total assets(CAPEX); book to price ratio (BP); earnings to price ratio (EP); number of analysts predicting that earnings surprise increase (ANFOR) and its lagged value (ANFORLAG); and lagged forecast error (FELAG). Numbers in parentheses are t-statistics. ***, **, *, and . represent significance levels of 0.1%, 1%, 5%, and 10% respectively.

4.7 Discussion

Our results indicate that the CorpInterlock algorithm improves the prediction of the trend of FE during all the years under study. This finding can be explained if we consider that many fund managers or their representatives have influence or even have a seat or more in the board of the corporations where they invest. Hence, they can use their knowledge

about the financial health of the companies where they have some presence to optimize their portfolios. Additionally, institutional investors have access to their own research team and could maintain certain independence of the analysts' influence. They are able to deeply evaluate the companies in which they are interested in investing. Therefore, they have an understanding of the fundamental valuation of the companies where they invest regardless of the day to day market speculation. This fact explains that even though our algorithm is able to improve the forecast of the trend of FE in relation to logistic regression, our baseline algorithm, the inclusion of the social network information does not improve the prediction of FE.

The inclusion of analysts in the social network of directors improves the prediction of the trend of CAR only during the period 1997-2001. The main explanation is that the period 1997-2001 corresponds to the last part of the internet "bubble". During this period, stock prices increased very quickly and the valuation multiples such as price-to-earnings ratio of technology companies like YAHOO were much higher than what a fundamental analysis would indicate. Many individual investors were participating in the market, and even small investors left their regular jobs to become full-time day traders. An important source of information for these investors was the forecast of the analysts (consensus). Suddenly, technology analysts became stars and were interviewed in popular shows. Their opinions were able to influence the market and therefore the returns, while fundamental or value investors had less importance. Additionally, analysts were also hired by investment banks that were participating in new deals such as IPOs, mergers, and acquisitions. Analysts had a strong pressure from the investment bankers to favorably cover companies where they expected to have a new deal or already had one. Also, if an analyst was covering a company

that was merged or acquired another company, suddenly she expanded her coverage to a new company or even a new industry, if the company was trying to diversify itself. For example, Microsoft has grown through acquisitions and has significantly expanded its initial area of economic activity as "software developer". The analysts of Microsoft have to understand the new business operations. This latter idea also explains why in 2001 there is such an unusual increase in the degree of the extended corporate interlock of the US market as Table 4.1 shows.

The relationship between analysts and directors is part of the "earnings game" that we introduced in section 1.2.2.1. The value of the stock options of CEO's and senior managers depends on the earnings surprises. Managers try to reach or improve the analysts' predictions. At the same time, analysts need the investment banking business because their compensation might be based on it. As a result there are incentives on both sides to find a mutually satisfying prediction. If the same game is played in several companies with various common directors, then the inclusion of analysts in the social network of directors may increase the explanatory power of returns predictions as we found in our model during the period 1997-2001. However, the degree of the extended corporate interlock and the explanatory power of returns predictions are reduced in the period 2002-2003, probably because of the regulations introduced by the Sarbanes-Oxley Act.

The most important variables in the prediction of FE are lagged cumulative abnormal return for the last six months (CAR1), size, earnings price ratio (EP), the number of analysts predicting that earnings surprise increase and its lagged value (ANFOR and ANFORLAG), and the lagged value of FE (FELAG) while in the case of CAR only size, ANFOR and FELAG are relevant variables (see Table 4.4). The importance of lagged variables in the

prediction of FE may explain that the regression analysis show a better fit for FE than for CAR. These results are not surprising if we consider that the "earnings game" may explain that companies that have shown earnings surprises in the past or analysts that have predicted earnings surprise in the past may also have similar trends in the future.

4.8   Conclusions

The link mining algorithm, CorpInterlock, demonstrated to be a flexible mechanism to increase the explanatory power of social networks with the forecasting capability of machine learning algorithms, such as boosting. The capacity to improve the forecast of earnings surprises and abnormal return using a mixture of well-known economic indicators with organizational and behavioral variables also enriches the debate between the modern finance theory and behavioral finance to show how behavioral patterns can be recognized under a rigorous method of analysis and forecast.

The basic and extended corporate interlocks have the properties of a "small world" network. This research shows that the basic corporate interlock with only directors, following the spirit of Mills' [176] thesis, has a stable mechanism to influence economic events as this chapter shows. However, the expansion of the original corporate interlock to include new actors, such as financial analysts, bring additional information especially during a "bull" market.

The application of link mining algorithms to problems of finance or social sciences may enrich the discussion in two ways: on one hand, a link mining algorithm can contribute to the understanding of social phenomena with the integration of different domains and espe-

cially quantifying the network perspective. On the other hand, the complex social problems offer scenarios to tests the adequacy or the development of new algorithms to solve interdisciplinary problems. For example, the oil supply is controlled by rich-oil countries with authoritarian or autocratic governments. A link mining algorithm may help to integrate the different domains in play: political, social, economical and cultural, and to find links that may bring new solutions to old problems.

Chapter 5

A boosting approach for

automated trading

5.1 Introduction

The recent development of electronic communication networks (ECNs) or electronic financial markets has allowed a direct communication between investors, avoiding the additional cost of intermediaries such as the specialists of the New York Stock Exchange (NYSE). A very important aspect of the ECNs is the access and publication of the real-time limit order book. For many years such access was not available to most traders. For example, in the NYSE only specialists could observe the entries of the limit order book. Other investors could only see the price and number of shares of the executed orders.

Electronic markets maintain a centralized order book for each traded stock. This book maintains lists of all active limit orders and is used as the basis for matching buyers and sellers. By making the content of this book accessible to traders, electronic markets provide a

very detailed view of the state of the market and allow for new and profitable trading strategies. For example, Kakade, Kearns, Mansour, and Ortiz in [135] present a competitive algorithm using volume weighted average prices (VWAP).[1] Kavajecz and Odders-White [144] study how technical analysis indicators can capture changes in the state of the limit order book.

In this chapter we present an automated trading algorithm that was tested in the context of the Penn-Lehman Automated Trading Project (PLAT) competition. The algorithm is based on three main ideas. The first idea is to use a combination of technical indicators to predict the daily trend of the stock. The trading algorithm uses the stock price of the previous ninety days, and the open price of the current trading day to calculate a set of well-known technical analysis indicators. Based on this information, the trading algorithm anticipates the direction of the market using a boosting algorithm, and then takes a long or short position if expects that the market will go up or a down respectively. The second idea is to use constant rebalanced portfolios [7] within the day in order to take advantage of market volatility without increasing risk. This part of the trading algorithm puts limit orders to assure that there is a constant mix between the value of the stocks and of the portfolio. The third idea is to use limit orders rather than market orders in order to minimize transaction costs. The trader accesses the order book to put limit orders out of the bid-ask spread to capture the rebates that ECNs such as ISLAND pay to the trader whose submission was in the order books at the moment of execution. [2]

---

[1]VWAP is calculated using the volumes and prices present on the order book.

[2]A market order is an order to buy an asset at the current market price. A buy (sell) limit order is executed only at a price less (greater) or equal than the limit price. The ECNs register the orders in the order book which is continuously updated with new orders or when an order is executed.

The bid-ask spread refers to the difference between the bid price or the highest price that a trader is

The rest of the chapter is organized as follows: section 2 explains in detail our trading strategy introduced above; section 3 presents the results of the participation of our trading algorithm in the PLAT competition; section 4 introduces improvements to our algorithm such as the integration of the market maker strategy, and section 5 discusses futures lines of research.

5.2 Trading strategies and PLAT Competition

5.2.1 Automated trading – PLAT

Our trading algorithm has been tested with the Penn Lehman Automated Trading Project (see Kearns and Ortiz [145]).[3] This project, which is a partnership between the University of Pennsylvania and the quantitative trading division of Lehman Brothers, simulates ISLAND, one of the major ECNs, and has had trading competitions since the Fall of 2002.

The simulator that supports PLAT captures price and volume information of ISLAND about every 3 seconds, and provides an architecture where clients can connect and submit limit orders. During the competition for April-May 2004, Microsoft (MSFT) is the only stock traded. The simulator creates its own order book receiving the information of ISLAND and mixes it with the orders of its clients.

The simulator generates detailed information about the position of each trader: market and price simulator, outstanding shares, present value, and profit and loss position.

---

willing to pay for an asset, and the ask price or the lowest price that a trader is willing to sell an asset.

A long position is the result of buying a security expecting that the value of the underlying asset goes up. A short position is the result of selling a borrowed security expecting that the value of the underlying asset goes down.

[3]This description of PLAT refers to Spring 2004 when we participated in the competition.

PLAT is different from the well-known trading agent competition (TAC) run at the University of Michigan [232] because of PLAT's strict limitation to the financial market and because only one stock is traded: Microsoft. The classic TAC game is based on the travel industry market, and since 2003, it has also included a supply chain management game. Wellman et al. in [232] reports recent results of TAC. Both competitions, PLAT and TAC, are similar in terms of offering a platform and software for agents to develop their trading strategies.

5.2.1.1 PLAT Competition

We designed the trading algorithm "CRP_TA" that participated in the PLAT competition run in the period April 26 to May 7, 2004. The rules used during this competition were[4]:

1. The performance of each trader is measured by the Sharpe ratio calculated as the mean return and standard deviation of the 10-day profit and loss positions.

2. Traders do not have a limit in terms of number of shares that they can hold. However, positions must be liquidated at the end of the day. Any long position will completely lose its value, and any short position must pay a penalty of twice its market value.

3. Transaction costs follow ISLAND's fee/rebate policy: when a trade is executed, the party whose order was in the order books receives a rebate of $ 0.002, and the party that submitted the incoming order pays a transaction fee of $ 0.003.

During this competition, participants were split into two groups: red and blue. Our agent was team 1 in the red group.

---

[4]Further explanation of the PLAT project can be found at <http://www.cis.upenn.edu/ mkearns/projects/plat.html>

The competition also included an agent per team that bought and sold large number of shares each day following the volume weighted average price (VWAP).

5.3   Trading algorithm

Our basic approach is to separate our analysis of the market into two time scales. The long time scale is on the order of days or hours, the short time scale is on the order of seconds or minutes. When operating on the long time scale we use a variety of technical indicators (see appendix B) to predict price trends. In other words, we try to predict whether the stock price goes up or down in the next day or next hour. When operating on the short time scale we stick to the prediction given by the long time scale analysis and place orders in a way that would take maximal advantage of volatility, and minimize transaction costs.

In more detail, our long time scale analysis is based on an adaptive combination of technical analysis indicators. The combination is optimized using the boosting learning algorithm and past month as training data. The short time-scale trading is based on constant rebalanced portfolios with a time-based profile selected according to the long-term analysis. Finally, the actual market orders are generated in a way designed to take advantage of the transaction cost policy used in ISLAND.

We call our trading algorithm CRP_TA because it implements a hybrid strategy of a) forecasting the daily stock price with Logitboost using technical indicators (TA), and b) intra-day trading following a constant rebalanced portfolio (CRP) strategy.

5.3.1 Applying Logitboost to the selection of technical trading rules

The trading algorithm CRP_TA forecasts the direction of the stock price using ADTs which are implemented with Logitboost. We introduced this algorithm in section 2.2.3. CRP_TA trains ADTs using the following technical analysis indicators of the previous ninety days and described in appendix B: simple moving average, average directional movement index, directional movement index, Bollinger bands, moving average convergence divergence, relative strength index, stochastic indicators, and money flow index. We calculated these indicators using R and its financial engineering package called Rmetrics. [5]

The instances are labeled using the following rules:

Buy, if $P^{\mathrm{c}} \geq P^{\mathrm{o}} + \tau$

Sell, if $P^{\mathrm{c}} \leq P^{\mathrm{o}} - \tau$

Hold, otherwise

where τ is a constant that at least covers the transaction costs (\$0.003), and P^{o} and P^{c} are the close and open price respectively.

Logitboost generates a new set of trading rules. Hence, instead of using the rules that each technical analysis indicator suggests, Logitboost defines what are the appropriate rules based on the market conditions and the combination of a list of very well-known technical indicators.

---

[5] Information about R and Rmetrics can be found at <http://cran.r-project.org> and at <http://www.rmetrics.org> respectively.

5.3.2   Constant rebalanced portfolio

Constant rebalanced portfolio, known in the financial world as constant mix, is a well-known strategy in the investment community. Kelly [146] showed that when individuals invest the same proportion of their money on a specific asset–the constant rebalanced portfolio–their portfolio value will increase (or decrease) exponentially. Kelly introduced the log-optimal portfolio as the one which achieves the maximum exponential rate of growth. Algoet and Cover [7] showed that if the market is stationary ergodic, the maximum capital growth rate of a log-optimal portfolio is equivalent to the maximum expected log return. Cover [73] and later on many other researchers such as Vovk and Watkins [228], Cover and Ordentlich [72] Blum and Kalai [36], and Kalai and Vempala [136] extended CRP to the concept of universal constant rebalanced portfolio.

CRP simply requires that traders maintain a fixed proportion of stocks to portfolio value. If stock price increases (decreases), the stock to portfolio value ratio increases (decreases), then part of the stocks must be liquidated (bought). This strategy works better when the stock price is unstable, so the trader is able to sell when the price is high, and buy when the price is low.

We tested the trading algorithm CRP_TA in the PLAT competition run between April 26 to May 7, 2004. Every day of the competition CRP_TA trained an ADT with Logitboost using the information of the last ninety days and then using P^o took a long position (50% of the portfolio invested in MSFT), short position (25% of the portfolio) or did not trade. During the first half hour CRP_TA built its position, and during the half hour before the market closes, CRP_TA liquidated its position. There was an asymmetry between the long

---

Input:
Set of price series (open (P^{o}), close P^{c}, high (P^{h}), low (P^{l})), and volume
τ is a constant that at least covers the transaction costs ($\$0.003$)
q_g is goal mix of stocks and cash for MSFT

Forecast with machine learning algorithm (Logitboost) and technical indicators (TA):
1. At the beginning of the day, train an ADT with Logitboost using training set with technical analysis indicators, and labels (see appendix B) calculated with price and volume series of the last 90 days.
2. Forecast trend of P^{c} using P^{o} and technical analysis indicators for trading day, and take one of the following positions for single stock (MSFT) in first half hour of trading:
Long ($q_g = 50\%$), if $E(P^{\text{c}}) \geq P^{\text{o}} + \tau$
Short ($q_g = -25\%$), if $E(P^{\text{c}}) \leq P^{\text{o}} - \tau$
Hold, otherwise

Intra-day constant rebalanced portfolio (CRP):
3. Sends simultaneously buy and sell limit orders for δ according to:
Submit buy limit order for δ, if $q_t < q_g - \delta/W$
Submit sell limit order for δ, if $q_t > q_g - \delta/W$
Hold, otherwise
where W is net value portfolio, q_t is current mix of stocks and cash for MSFT, and δ is amount of dollars to buy or sell in order to reach q_g.

4. If ($q_t! = q_g$) after 60 ticks (about one minute), cancel limit orders, submit market orders to obtain q_t, and submit new limit orders.

5. Liquidate position in the last half hour before market closes.

Output:
Profit/loss of algorithm

Figure 5.1: The CRP_TA algorithm.

position (50%) and the short position (-25%) because of the higher penalty that a trader

with a short position would pay during the competition. The training of ADTs was done

using the MLJAVA package.[6]

The trading algorithm CRP_TA traded during the day balancing the portfolio according

to a goal mix as Figure 5.1 explains. CRP_TA intended to increase revenues sending limit

orders and expected that these orders arrived before than the counterparty's orders when

the orders were executed. In this case, the trader received rebates, and avoided paying fees.

---

[6]If interested in using MLJAVA, please contact yfreund@cs.ucsd.edu.

5.4 PLAT competition results

After ten trading days of participating in the PLAT competition, CRP_TA obtained a return of $27,686 and the Sharpe ratio was 0.83. Its performance was the second best in its group as Table 5.2 shows. CRP_TA forecasted correctly a short or long position eight out

	Profit and loss										
	Sharpe Ratio	26/4	27/4	28/4	29/4	30/4	3/5	4/5	5/5	6/5	7/5 Total
Team1	0.8334	2249	-151	7527	7198	6628	-2523	1567	2238	1885	1068 27687
Team2	-0.1619	27	-513	-3062	1219	3204	-153	327	15	61	-4601 -3476
Team3	1.1221	3574	7083	-127	-2832	2040	6691	4335	6108	5915	3061 35847
Team4	-0.4232	-44962	3147	-1185	-1832	-988	-88302	946	1129	1907	2316 -127825
Team5	-12.6200	-9.E+06	-8.E+06	-9.E+06	-8.E+06	-9.E+06	-7.E+06	-8.E+06	-8.E+06	-8.E+06	-7.E+06 -8.E+07
Team6	0.7213	1045	4729	243	-6694	12508	11065	-2377	5708	9271	11755 47252
Team7	2.4963	3433	1374	2508	2928	3717	3444	1322	3300	2199	966 25190
Team8	0.7559	271	538	-242	-248	13	636	386	452	461	121 2387
Team9	0.5778	1307	2891	-1563	-1349	-1339	3230	1850	2037	2465	1041 10569
Team10	0.0432	-4655	-1370	2178	2820	2766	2961	2665	-5746	2402	-2545 1475
Team11	-12.5931	-9.E+06	-8.E+06	-7.E+06	-8.E+06	-8.E+06	-7.E+06	-8.E+06	-8.E+06	-8.E+06	-8.E+06 -8.E+07

Figure 5.2: Profit and Loss of PLAT competition for all players. Competition was split in the first five teams (red group) and the next five teams (blue group). First column shows the Sharpe ratio for each team during the whole competition. Additional columns have daily profits or losses for each team expressed in US$. CRP_TA is team 1. Teams 5 and 11 are artificial traders who bought and sold large volume of shares following the VWAP.

of the ten days of the PLAT competition. These results were better than the results of a simulation for a sample of 840 days when the predictor was trained with information of the last 90 days. In this last case the test error was 48.8%. These differences could be explained because the optimization of the parameters used to calculate the technical indicators at the beginning of the competition might have not been adequate for other periods. We spent a significant amount of time fine tuning the parameters used for the forecast. Additionally, the trader did not get its position at the open price as the above simulation did it. It reached its position after the first half hour of trading.

To understand the intra-day dynamic, we present the results of a trading day when

the market is up and down (Figure 5.3). May 3rd was a very volatile day and the market

was up, while CRP_TA got a short position. The losses of a short position were partially

compensated by the benefits of intra-day trading thanks to the CRP strategy. On April

28th the market went down. CRP_TA assumed a short position that led to a profitable

position. This last result is evident in the top panel of Figure 5.3 that shows an important

difference between the portfolio value index and the index price or buy and hold (B&H)

position.

    (a) Market up   (b) Market down
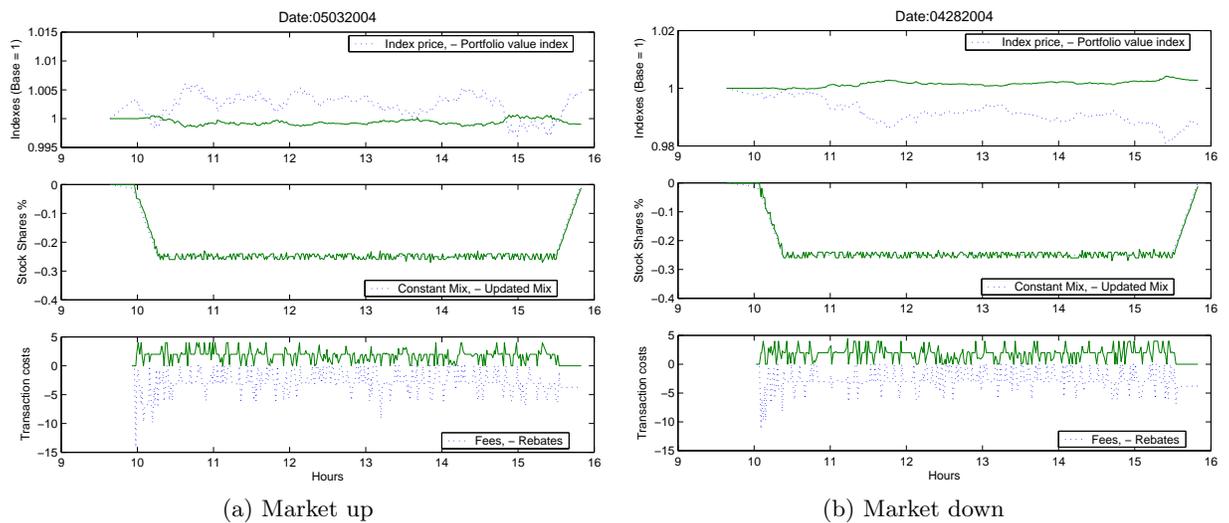
Figure 5.3: Representative intraday results of PLAT competition for CRP_TA when market
is up (a) and down (b). Top graphs compare portfolio value index with an index price or
a simple buy and hold (B&H) position. Middle graphs compare the goal or constant mix
of stocks and cash with the updated mix according to the trading algorithm. The steeper
curve at the beginning and at the end of the trading day is the period when CRP_TA builds
and liquidates its goal position. Bottom graphs include fees (> 0) and rebates (< 0). The
differences between rebates and fees are transaction costs.

During each trading day there were a large number of trading operations. However, the

process to adjust the portfolio to reach the goal mix affected the results because the trader

CRP_TA paid more fees than received rebates as the bottom of Figure 5.3 shows. The winner on CRP_TA's group during the PLAT competition, team 3, acted as a market maker placing limit orders outside the current spread. Hence, an important amount of CRP_TA's orders were plausibly traded with this team; however this trader did not pay fees, only received rebates because their orders were limit orders that most of the time arrived before CRP_TA's orders. If CRP_TA could incorporate this market marker strategy, probably its results may improve as we show in the next section.

5.5 Improved order strategy

After the PLAT competition, we integrated the market maker strategy into the CRP_TA, and we call the modified version of the algorithm as the "Market maker CRP_TA". The most important aspect of the revised version of the algorithm is that the orders should be executed as limit orders, and not as market orders as follows: Market maker CRP_TA starts with a balanced position according to the proportion of shares over portfolio value established as a goal (q_g). Then it sends simultaneously a buy limit order at a price slightly below ($0.005) the price at the top of the buy order book (P_{BuyB}), and a sell limit order at a price slightly above ($0.005) the price at the top of the sell order book (P_{SellB}). If the order is not completely filled within ten minutes of being issued, existent limit orders are canceled, and limit orders are reissued. In all cases, orders are reissued for the amount necessary to reach the goal mix of stocks and cash (see Figure 5.4).

Input:
Set of price series (open (P^o), close P^c, high (P^h), low (P^l), and volume
τ is a constant that at least covers the transaction costs ($\$0.003$)
q_g is goal mix of stocks and cash for MSFT
κ is minimum amount above or below top price of order books ($\$0.005$)

Forecast with machine learning algorithm (Logitboost) and technical indicators (TA):
1. At the beginning of the day, train an ADT with Logitboost using training set with technical analysis indicators, and labels (see appendix B) calculated with price and volume series of the last 90 days.
2. Forecast trend of P^c using P^o and technical analysis indicators for trading day, and take one of the following positions for single stock (MSFT) in first half hour of trading:
Long ($q_g = 50\%$), if $E(P^\text{c}) \geq P^\text{o} + \tau$
Short ($q_g = -25\%$), if $E(P^\text{c}) \leq P^\text{o} - \tau$
Hold, otherwise

Intra-day market maker constant rebalanced portfolio (CRP):
3. Sends simultaneously a buy and sell limit orders for δ according to:
Buy limit order for δ and $P_B = P_{BuyB} - \kappa$, if $q_t < q_g - \delta/W$
Sell limit order for δ and $P_S = P_{SellB} + \kappa$, if $q_t > q_g - \delta/W$
Hold, otherwise

where:
W is net portfolio value
q_t is current mix of stocks and cash for MSFT
δ is amount of dollars to buy or sell in order to reach q_g
P_B and P_S are prices of long and short limit orders
P_{BuyB} and P_{SellB} are prices at the top of the buy and sell order book respectively

4. If ($q_t\,! = q_g$) after 600 ticks (about 10 minutes), cancel and resubmit limit orders to obtain q_t.

5. Liquidate position in last half hour before market closes.

Output:
Profit/loss of algorithm

Figure 5.4: The money market CRP_TA algorithm.

We ran this new trading strategy and the original CRP_TA strategy during the period January 5-9, 2004. We present the results of January 8th for the market maker CRP_TA strategy and for the CRP_TA agent in Figure 5.5. During the week of January 5-9, the Sharpe ratio is 0.03 and -0.28 for the Market maker CRP_TA strategy and for the CRP_TA strategy respectively. The bottom of Figure 5.5 shows that Market maker CRP_TA received more in rebates than the amount it had to pay in fees. This difference helped to improve the financial result of the algorithm which is the major shortcoming of the CRP_TA strategy.

(a) Market maker CRP_TA   (b) CRP_TA
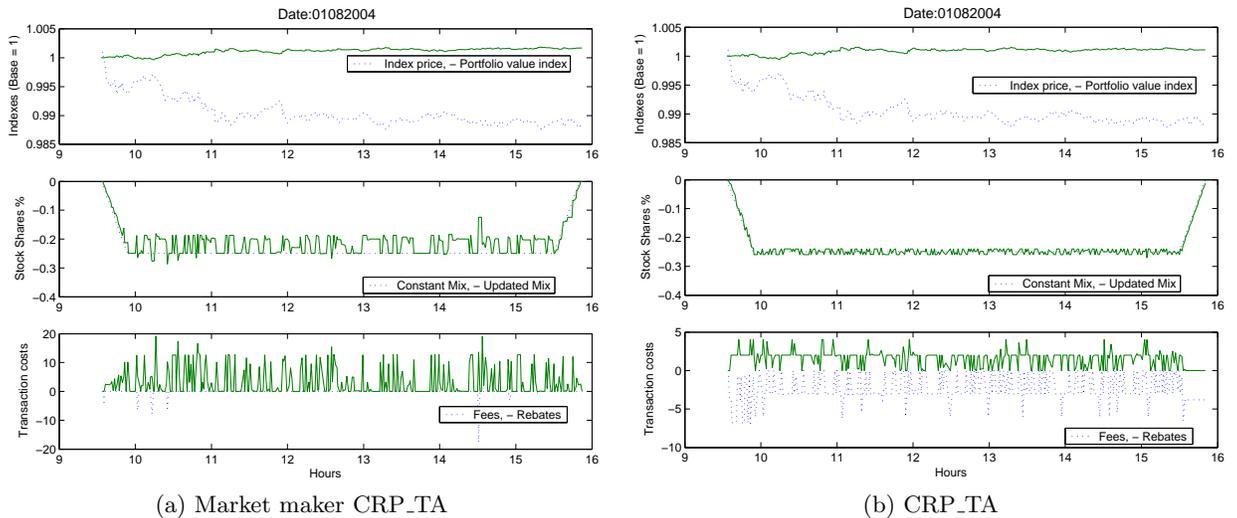
Figure 5.5: Representative intraday results for Market maker CRP_TA (a) and CRP_TA (b) in January 8th, 2004. Top graphs compare portfolio value index with an index price or a simple buy and hold position. Middle graphs compare the goal or constant mix of stocks and cash with the updated mix according to the trading algorithm. The steeper curve at the beginning and at the end of the trading day is the period when trading algorithms build and liquidate their goal position. Bottom graphs present fees (> 0) and rebates (< 0). The differences between fess and rebates are transaction costs.

Another shortcoming of the CRP_TA strategy is that this strategy takes a high risk when it keeps only a short or long position during the day. A variation of the CRP_TA strategy could be the creation of a portfolio that has a long and short position simultaneously. The scores obtained from Logitboost to forecast the stock price could be used to weight the long and short position. Hence, the position with higher score would have a higher weight. A market neutral portfolio could also be obtained using the same proportion of stocks to portfolio value for the short and long position. We also tried this final alternative for the week of January 5-9, 2004 and the Sharpe ratio deteriorates to -2.06. Obviously, this alternative misses the benefit of market forecasting using ADTs.

5.6   Final Comments and Conclusions

In this chapter we show that the constant rebalanced portfolio or constant mix strategy can improve if a classifier can anticipate the direction of the market: up, down or no change. Additionally, transaction costs play a central role to improve performance. Instead of an automatic rebalance of the portfolio, the results of the PLAT competition indicate that if the CRP strategy is implemented only with limit orders, its results improve because of the rebates.

We used very well known technical indicators such as moving averages or Bollinger bands. Therefore, the capacity to anticipate unexpected market movements is reduced because many other traders are expected to be trying to profit from the same indicators. In our case, this effect is reduced because we tried to discover new trading rules using Logitboost instead of following the trading rules suggested by each indicator. However, we are aware that our predictor may improve if we transform the technical indicators into more accurate ratios or select more informative indicators such as the effect of current news into stock prices.

Finally, our algorithm can be enriched by the introduction of risk management mechanisms in order to change strategy or liquidate its position if market behaves in unexpected ways.

Chapter 6

Automated trading with expert weighting

6.1 Introduction

This chapter introduces boosting and online expert weighting as the basis for an automated trading system that, in contrast to the algorithm presented in the last chapter, trades a large number of stocks for a long time.

The rest of the chapter is organized as follows: section 2 introduces the trading system, section 3 explains in detail the experiments, section 4 presents the results of our trading algorithm, and section 5 presents the conclusions and final comments. Appendix C introduces the main investment indicators used in this research.

6.2 The trading system

The trading system is designed to trade stocks and relies on a layered structure consisting

of a machine learning algorithm, an online learning utility, and a risk management overlay.

ADT, which is implemented with Logitboost, was chosen as the underlying algorithm. One

of the strengths of our approach is that the algorithm is able to select the best combination of

rules derived from well-known technical analysis indicators and is also able to select the best

parameters of the technical indicators. Additionally, the online learning layer combines the

output of several ADTs and suggests a short or long position. Finally, the risk management

layer can validate the trading signal when it exceeds a specified non-zero threshold and limit

the application of our trading strategy when it is not profitable (see Figure 6.1).
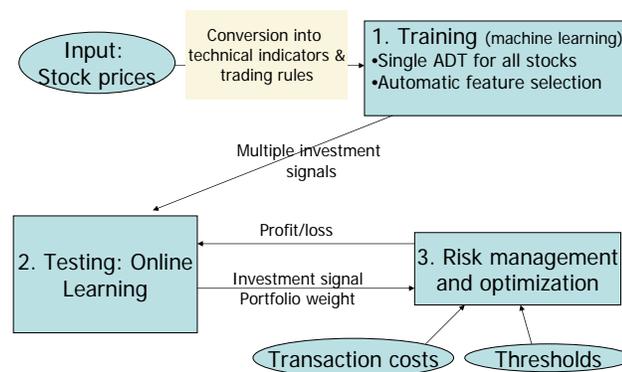
Figure 6.1: Trading system as a process flow. The inputs of the algorithm are price series that are transformed into technical indicators and trading rules. The first component trains several ADTs (experts) and each of them generates an investment signal. The second layer weights the investment signals of each ADT, and generates a weighted single investment signal for each stock using an online learning algorithm. The third layer filters the weak investment signals, and restricts non-profitable trading strategies.

6.2.1 Layer 1: The machine learning algorithm and its implementation

The basic algorithm that we used is ADT implemented with Logitboost. We introduced this algorithm in section 2.2.3, and we presented an initial application for automated stock trading in the context of the PLAT competition (see chapter 5). Several adaptations of the basic algorithm were necessary to improve its performance.

The inputs to the algorithm were a group of well-known technical indicators and investment signals introduced in section 1.2.2.3 and presented in appendix C. Additionally, we included the Sharpe ratio as a performance indicator and several measures of volatility such as GARCH (see appendix C.1 for a description of its calculation), Chaikin, and Garman-Klass volatility. The instances were weighted by the Sharpe ratio. We also included ratios and trading rules suggested by the practice of technical analysis. We calculated these indicators using R and its financial engineering package called Rmetrics. [1]

Our goal was to predict the trend of beta abnormal returns $(BXRET)$[2] using the above investment signals. $y_t \in [-1, +1]$ is the binary label to be predicted where 1 represents the expectation of a positive beta abnormal return, and -1 otherwise.

A major problem with the use of technical indicators is their calibration or the adequate selection of their best parameters, as with the optimal number of days to calculate the moving averages used by stochastic indicators or the number of standard deviations used to calculate the Bollinger bands. There are many versions of optimizers such as the "brute force" approach where all the alternatives are tested one-at-a-time, and simulated annealing

---

[1]Information about R and Rmetrics can be found at <http://cran.r-project.org> and at <http://www.rmetrics.org> respectively.

[2]Beta abnormal return is the return of a specific asset less the average return of all assets in its beta portfolio for each trading date.

or genetic optimizers as suggested by Katz and McCormick [143]. In our case, we initially tried one-at-a-time optimization where each parameter was tested with several values while keeping the others constant. We found that this approach was very inefficient because it required too much computer power and time.

The solution that we implemented was the simultaneous recalculation of the technical indicators with several values of their parameters. Then the boosting algorithm would select the best combination of parameters and technical indicators. Our initial parameters were the parameters recommended in the literature. We tested 16 different variations of the initial parameters and did not find a major difference when we used only three different variations of the parameters. So, we included three different versions of most of the technical indicators in our training, validation, and testing set. In total, each instance of the training, validation, and testing set had 112 inputs. We were able to use this approach because of the boosting's feature selection capability (see appendix C for a presentation of the parameters used).

In our experiments we observed that a small change in the sample may lead to different ADTs changing the investment recommendation. In order to prevent this problem, our algorithm generates different ADTs and, in the next layer, selects the output of the best ADTs.

6.2.2 Layer 2: Online learning and expert weighting

We improved the performance of the boosting algorithm by adding online learning capacity with expert weighting. Our algorithm is derived from a previous algorithm proposed by Freund, Manssour and Schapire [103] that predicts with an exponentially weighted aver-

age of the training error of all hypotheses. The "exponential weights" formula comes from the weighted majority algorithm introduced by Littlestone and Warmuth [165] and further studied by Cesa-Bianchi et al. [53]. The "exponential weights" is a different formula to compute the posterior distribution as it would be proposed by Bayesian analysis. An interesting feature of this algorithm is that it abstains from predicting on certain instances, so the predictions that it makes are very reliable. A final comment about the original algorithm proposed by Littlestone and Warmuth [165] is that it generates a fixed classification rule, so it works in the standard batch learning model. The algorithm that we present in the next subsection is an extension of this original algorithm applied to financial time series prediction, and therefore is an online learning model.[3]

6.2.2.1 The expert weighting algorithm

To simplify the presentation of our algorihtm, we introduce the case of one asset which can easily be extended to N assets. The final outcome sequence is the net weights of an asset $(W = W_1, W_2, .., W_t, ..., W_l)$ where $W_t \in [-1, 1]$, t refers to a time step and l represents the last step of the sequence.

In this research, every expert is an ADT calculated with the training set which is a sequence of pairs as introduced in section 2.2.3. We refer to the sequence of experts by $\psi = \psi_1, \psi_2, ..., \psi_E$ where E is the number of experts.

The outcome of expert ψ_i at time t is the prediction score S_t^i. In order to calculate the experts' weight, we need to calculate the cumulative abnormal return of each expert i (ψ_i)

---

[3]Borodin and El-Yaniv [42, 89] propose a different approach to use online learning for trading and portfolio selection. They measure the performance of their trading algorithm in relation to a "statistical adversary", and an optimal offline algorithm.

as:

$$car_t^i \doteq \sum_{s=t_{i+1}}^{t} \text{sign}(S_s^i) \cdot r_s^i$$

where $t_1 = 0$, t_i is the time step at which ψ_i was added to the pool when $i > 1$, and r_s^i is the abnormal return for expert i at time s.

car_t^i is a sum of random variables that are close to $\mathcal{N}(\epsilon, 1)$ where ϵ is the slight advantage that the expert has over random guessing. Our goal is to give higher weight to experts that have higher ϵ. On the one hand, ϵ is masked by the noise, but on the other hand, the noise increases only as $\sqrt{t}$ while the drift increases like ϵt.

The weight of the first expert is:

$$w_t^1 \doteq \exp\left(\frac{C \cdot car_t^1}{\sqrt{t}}\right)$$

where C is an exogenous parameter.

The weight of expert ψ_i at time t where $t > t_i$ is:

$$w_t^i \doteq I_i \cdot \text{ramp}(t - t_i) \cdot \exp\left(\frac{C \cdot car_t^i}{\sqrt{t - t_i}}\right)$$

where $I_i \doteq \frac{\sum_{j=1}^{i-1} w_{t_i}^j}{i-1}$ is the initial weight assigned to the new expert i (ψ_i) and is the average weight of the previous experts at the time that the new expert is added (t_i), $\text{ramp}(t - t_i) \doteq \min\left(\frac{t-t_i}{t_{t+1}-t_i}, 1\right)$ is a function that brings in the new expert gradually, and t_{i+1} is the time that the next expert is added.

The experts' weight (W_t) is the result of weighting the answers of all the experts:

$$W_t = L_t - S_t$$

where the fraction of experts that suggest a long or short position is $L_t = \frac{\sum_{i:S_t^i>0} w_t^i}{\sum_i w_t^i}$ or $S_t = 1 - L_t^i$, respectively. Therefore, $W_t \in [-1, +1]$ is also a trading suggestion to take either a long $(W_t > 0)$, hold $(W_t = 0)$ or a short position $(W_t < 0)$.

The expert weighted cumulative abnormal return is:

$$CAR \doteq \sum_t (W_t \cdot r_t - (W_t - W_{t-1}) \cdot tc)$$

where tc are transaction costs.

We call this version of the algorithm as the "Base" version. Additionally, we include two variations of the algorithm where we modify the denominator of the weight functions as:

1. "Simple time adjustment" version where:

$$w_t^i \doteq I_i \cdot \text{ramp}(t - t_i) \cdot \exp\left(\frac{C \cdot car_t^i}{t - t_i}\right)$$

$$w_t^1 \doteq \exp\left(\frac{C \cdot car_t^1}{t}\right)$$

2. "No time adjustment" version where:

$$w_t^i \doteq I_i \; \cdot \; \text{ramp}(t - t_i) \; \cdot \; \exp\left(C \; \cdot \; car_t^i\right)$$

$$w_t^1 \doteq \exp\left(C \; \cdot \; car_t^1\right)$$

6.2.3   Layer 3: Risk management and optimization

An important aspect of the layered structure of the trading system is that the decision to trade is separated from the trade recommendation made by layers 1 and 2. While layer 1 and 2 suggest a preferred long or short position, layer 3 evaluates this position by considering market factors before taking an investment decision. Any trading system requires risk management rules. However they are difficult to include in the trading module itself, so it was easier to include them in an independent layer.

The risk management layer evaluates the strength of the trading signal W_t given by layer 2. As W_t is normalized between a completely short (-1) and completely long (1) position, the risk management system eliminates the trading signals that are not above (for long positions) or below (for short positions) a non-zero threshold γ_0. This parameter is fixed based on the information generated during the training and optimization stage.

A common complaint about automated trading systems is that they are not always profitable because market conditions change, and therefore what used to be adequate in certain period of time is not in another moment. In this respect, an indicator to evaluate the performance of an algorithm used by traders is the maximum drawdown $(D_{i,t})$ [81, 179]. This indicator defined for stock i over a period $t - t_0$ is:

$$D_{i,t} \doteq max(R_{i,t_x} - R_{i,t_y}|t_0 \leq t_x \leq t_y \leq t)$$

R_{i,t_x} and R_{i,t_y} are the accumulated return from time t_0 until time t_x and t_y respectively for stock i. We denote the vector of the maximum drawdown for all stocks of our portfolio as D_t. If $D_{i,t} < \gamma_1$ for a certain period of time (30 trading days), then the system holds its current position, and if $D_{i,t-1} == min(D_{t-1})$, the system liquidates the current stock position. If these conditions improve, the stock can be traded again. After the first thirty trading days (which is still even part of the validation period) these rules are continuously applied. The rationality for the first risk management rule is very obvious: the trading system should not invest more in a strategy that has been demonstrated to be unprofitable for a specific stock. Considering that the market conditions may change, the trading system only holds the position but does not liquidate it. If the situation is even worse, and if the maximum drawdown of a stock i is the maximum drawdown among all the selected stocks, the trading system liquidates the position. We think that with a large porfolio a relative indicator is useful because it incorporates market conditions. However, an additional rule, common among trading desks, that the system stops trading or liquidates its current position when its return is below a certain nominal amount, could also coexist with the above rules.

We also used the Sharpe ratio (SR_t) and Sterling ratio (St_t) to evaluate the risk-adjusted return of our experts. The Sharpe ratio is calculated as the mean of return divided by its standard deviation:

$$SR_t \doteq \frac{\mu(R_t)}{\sigma(R_t)}$$

The Sharpe ratio can be very unstable for small return variances, and in a long period it cannot distinguish between shorter periods where there are large profits or losses. This is the reason that traders also incorporate the maximum drawdown in their risk analysis. The maximum drawdown helps to recognize those clusters of profits and losses that are undetected by the Sharpe ratio. In this respect, the Sterling ratio brings additional information that is not captured by the Sharpe ratio when it includes the maximum drawdown as its denominator and the mean of the return as the numerator:

$$St_t \doteq \frac{\mu(R_t)}{D_t}$$

We calculated the Sharpe ratio and Sterling ratio on a monthly basis (20 trading days).

6.2.3.1 Optimization and training

We divided our time series between a training, validation, and test set. The training set was used to generate the first expert or ADT. The objective of the validation set is to choose the appropriate version of the expert weighting algorithms and the optimal parameters C, γ_0, and γ_1. In section 6.4, we present the result of all the versions of the model for comparative reasons. Based on the initial expert, optimal parameters, and test set, we conducted the experiments that we describe in the next section.

The complete algorithm that includes the three layers of the trading system is presented in Figure 6.2.

Input:

Set of price series (close, open, high, and low), volume and beta abnormal return ($BXRET$)

r is number of different values of parameters to calculate investment signals

N is number of stocks to be selected from market

d is number of days between experts' training

γ_0 and γ_1 are thresholds to filter experts' weight.

C is an exogenous parameter for expert weighting.

Train with machine learning algorithm:

1. Select a representative sample of N number of stocks from targeted market.

2. Calculate investment signals, and labels [a] with basic parameters for a selected group of stocks.

3. Recalculate investment signals with r variations of basic parameters, and include all of the investment signals as features in the training and test sets where the binary label is $y_t = sign(BXRET)$.

4. Integrate all the instances of the N stocks in a single training and single test set.

5. Train an initial expert (ψ_1) with Logitboost. Every d days train a new expert i ψ_i. Call the sequence of experts as $\psi = \psi_1, \psi_2, ..., \psi_E$ where E is the number of experts.

6. Every day recalculate test set and weight experts as in next steps.

Expert weighting algorithm (this part is simplified for one asset, even though can be extended to N assets):

7. Calculate the weight of the first expert at time t as $w_t^1 \doteq \exp\left(\frac{C \cdot car_t^1}{\sqrt{t}}\right)$ where:

   a. $car_t^i \doteq \sum_{s=t_{i+1}}^{t} \text{sign}(S_s^i) \cdot r_s^i$

   b. $t_1 = 0$, t_i is the time step at which ψ_i is calculated when $i > 1$

   c. r_s^i is the abnormal return for expert i at time s.

8. Calculate the weight of expert ψ_i at time $t > t_i$ as $w_t^i \doteq I_i \cdot \text{ramp}(t - t_i) \cdot \exp\left(\frac{C \cdot car_t^i}{\sqrt{t-t_i}}\right)$ where:

   a. $I_i \doteq \frac{\sum_{j=1}^{i-1} w_{t_i}^j}{i-1}$ is the initial weight assigned to i (ψ_i)

   b. $\text{ramp}(t - t_i) \doteq \min\left(\frac{t-t_i}{t_{i+1}-t_i}, 1\right)$

   c. t_{i+1} is the time that the next expert is added.

9. Calculate the experts' weight as $W_t = L_t - S_t$ where $L_t = \frac{\sum_{i:S_t^i>0} w_t^i}{\sum_i w_t^i}$ and $S_t = 1 - L_t^i$.

Risk management:

10. if $|W_t| < \gamma_0$, then $W_t = 0$

11. If for stock i, $D_{i,t-1} == min(D_{t-1})$ then $W_{i,t} = 0$, else if $D_{i,t-1} < \gamma_1$ then $W_{i,t} = W_{i,t-1}$ where:

   a. $D_{i,t} \doteq max(R_{i,t_x} - R_{i,t_y} | t_0 \leq t_x \leq t_y \leq t)$ is the maximum drawdown for stock i

   b. R_{i,t_x}, R_{i,t_y}, and $W_{i,t}$ are the accumulated return from time t_0 until time t_x and t_y, and experts' weights for stock i respectively

   c. D_t is the matrix of maximum drawdowns for selected stocks

Output:

Expert weighted cumulative abnormal return for stock i is:

$CAR_i \doteq \sum_t \left(W_{i,t} \cdot r_{i,t} - (W_{i,t} - W_{i,t-1}) \cdot tc_i\right)$ where tc_i are transaction costs for stock i.

Figure 6.2: Forecasting and trading algorithm.

---

[a]In this research we used all the investment signals described in appendix C

6.3   Experiments

We selected a random sample of 100 stocks from the S&P 500 index with daily data from January 2001 through December 2004. For every stock we had a complete series of close, open, high and low prices, volume, and $BXRET$. Companies that did not have the variable $BXRET$ were eliminated. We obtained these price series from CRSP. The time series were distributed in the following way:

- Training set included two years (500 trading days). This calculation required about 540 trading days because of about 40 lagged values that were discarded. This dataset includes data from October 2000 to December 2002. We aggregated in one set the observations of all the stocks. So, our training set had 50,000 observations. In previous tests, we tried the generation of individual ADTs for each stock, however this solution used a significant amount of computer power and time, and the results were still very similar.

- Validation (in-sample) set is based on 100 trading days (10,000 observations). This dataset includes data from December 2002 to May 2003.

- Test (out-of-sample) set had the remaining observations (411 trading days or 41,100 observations). This dataset includes data from May 2003 to December 2005.

We ran our algorithm using a moving window of two years for training followed by about two months out-of-sample testing (50 days). We conducted our tests on a daily rollover basis. Even though we trained a new expert (ADT) only every 50 trading days, we tested the existent experts every day with the additional price information. The algorithm presented

in Figure 6.2 reweighted the participation of each expert according to individual performance and time elapsed since the initial training. Additionally, the risk management module would hold or liquidate positions when they were not profitable or became too risky.

The selection of 50 trading days between training periods was partially justified by the fact that using longer periods (100 days) or shorter periods (25 days) did not generate significant differences in the results obtained. However, the computational overhead was very important when the days between training were reduced. The training of each expert required about 45 minutes and the testing of every day lasted about two to three minutes with one expert and about eight to ten minutes with the eleven experts that we had in total. [4]

We aggregated the daily results in 21 test sets (2,000 observations per set equivalent to 20 trading days) that we used to calculate the performance and risk indicators. The averages of these results are presented in the next section. We ran the expert weighting algorithm using its three versions and the following values of the parameter C: 0, 0.05, 0.5, 1, 5, 10, 20, 30, 40, and 50. C is 0 is equivalent to a simple average of all the experts. The threshold to eliminate very weak expert weights (γ_0) is set to 0.20, and the threshold to restrict trading (γ_1) is set to 0.

We tested our results with transaction costs of $0, $0.001, $0.002, and $0.003 per stock. These values are realistic if we consider that ISLAND has the policy to pay a rebate of $0.002 per stock to the trader whose order was in the order books, and charges a rebate of $0.003 per stock to the trader that submitted the incoming order. As we showed in

---

[4]During the PLAT competition (see chapter 5), we trained our expert every day. However, this competition was only for ten days and we were not combining new experts.

chapter 5, traders can use this policy in their favour and do not have to pay the full fee of $0.003. On the contrary, they can capture the rebate using only limit orders as the market maker strategy suggests. Even more, large brokerage firms have much lower direct transaction costs and the initial investment to trade becomes a sunk cost that does not have a major impact in individual trades.

6.4 Results

Our results do not show major differences when the different values of C are used, with the exception of cases when $C > 10$ where results often deteriorate. Hence, we present only the cases where $C = 1$ and $C = 0$ (simple average of experts). We also compare our results with a baseline alternative of buy and hold (B&H).

Figure 6.3 shows the annualised in-sample and out-of-sample CAR for all stocks by transaction costs for the "Base" version and with $C = 1$. These figures indicate that all the alternatives performed better than a simple B&H strategy. The validation or in-sample set was not profitable during the first two thirds of the trading period when transaction costs were $0.002 and $0.003. All of the alternatives, with the exception of B&H, recovered in the last period with the introduction of a second expert, and also when the algorithm had accumulated some additional information. At the end of this period, all of the alternatives showed positive CAR (see Table 6.1).

Figure 6.3 also shows that all the alternatives of the out-of-sample group have positive CAR while the B&H alternative has negative CAR. Also all models show improvements until about halfway through the trading period or the end of the year 2004. Then, there

is a decline and finally a smoother recovery. An explanation for these results is that the

trading system was able to choose an adequate combination of experts that was efficient

during a certain period of time. These experts became less efficient, and the incorporation

of new experts and probably, new market conditions, led to the final improvement of the

results. The above figures and Tables 6.1, and 6.2 indicate that an increase of \$0.001 in

(a) In-sample

(b) Out-of-sample
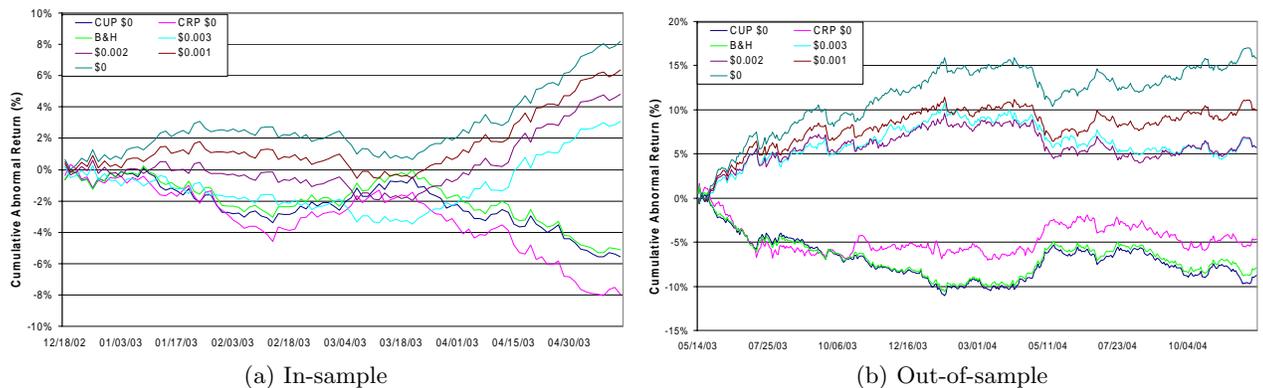
Figure 6.3: Cumulative abnormal return (CAR) by transaction costs with $C = 1$. CAR is calculated as weighted sum of profits of portfolio.

transaction costs when transaction costs are \$0 and \$0.001 represents about 3-6 percent

points of differences in total and average CAR over the whole trading period. These results

are even more evident with Sharpe ratios (see Table 6.3). The immediate explanation for

these results is the high number of transactions required by the expert weighting algorithm.

In this respect, one of the roles played by the risk management and optimization layer is to

reduce the number of transactions to only those that seem to be profitable.

The Sterling ratio [5] in Table 6.4 shows sharper differences among the different versions

of the model. While the Sharpe ratio is useful for comparing the performance of each alter-

---

[5]The in-sample Sterling ratio table is not included because an important part of its results are voided in order to calculate the maximum drawdown.

| Model	In-sample				Out-of-sample			
	$0.003	$0.002	$0.001	$0	$0.003	$0.002	$0.001	$0
Avg.weight	3.15	4.80	6.34	8.19	7.41	7.62	11.98	18.01
Base	3.10	4.81	6.35	8.18	6.87	7.11	11.51	17.52
Simple	3.15	4.80	6.36	8.19	6.93	6.96	11.54	18.07
No adj.	3.10	4.65	6.40	8.02	7.04	8.26	11.68	19.46
B&H	−5.42	−5.32	−5.22	−5.12	−9.24	−9.24	−9.24	−9.24

Table 6.1: Cumulative abnormal return (%) by transaction costs with $C = 1$. CAR is calculated as weighted sum of profits of portfolio.

Model	In-sample				Out-of-sample			
	$0.003	$0.002	$0.001	$0	$0.003	$0.002	$0.001	$0
Avg.weight	12.02	16.03	20.17	26.63	6.69	6.99	9.88	14.29
Base	11.82	16.09	20.20	26.61	6.36	6.63	9.56	13.97
Simple	12.03	16.04	20.23	26.63	6.39	6.55	9.58	14.34
No adj.	11.85	15.49	20.39	25.93	6.63	7.56	9.57	15.14
B&H	−10.82	−10.59	−10.37	−10.14	−4.61	−4.61	−4.61	−4.61

Table 6.2: Average annualised abnormal return (%) by transaction costs with $C = 1$. Average is calculated over the monthly in-sample (December 2002 to May 2003) and out-of-sample (May 2003 to December 2005) sets.

native over the complete trading period, the Sterling ratio, and specifically the maximum drawdown, seem to be more useful for capturing large variations of risk-adjusted return inside each test set. For instance, the "No time adjustment" version of the model has a slightly higher return than the "Base" and "Simple" versions. However, this situation is mostly reversed with the Sterling ratio.

6.5   Final comments and conclusions

The trading system introduced in this chapter generated positive abnormal returns for a large group of stocks. This trading system was able to obtain these results based on a ma-

	In-sample				Out-of-sample			
Model	$0.003	$0.002	$0.001	$0	$0.003	$0.002	$0.001	$0
Avg.weight	0.09	0.14	0.19	0.23	0.09	0.11	0.15	0.20
Base	0.09	0.14	0.19	0.23	0.09	0.10	0.14	0.19
Simple	0.09	0.14	0.19	0.23	0.09	0.10	0.15	0.20
No adj.	0.09	0.13	0.18	0.22	0.10	0.11	0.15	0.21
B&H	−0.15	−0.15	−0.15	−0.15	−0.11	−0.11	−0.11	−0.11

Table 6.3: Average Sharpe ratios by transaction costs with $C = 1$. Average is calculated over the monthly in-sample (December 2002 to May 2003) and out-of-sample (May 2003 to December 2005) sets. Sharpe ratio is a risk-adjusted return indicator calculated as the mean of return divided by the standard deviation.

Model	$0.003	$0.002	$0.001	$0
Avg.weight	0.07	0.08	−2.62	0.07
Base	0.34	−0.02	1.33	0.08
Simple	0.51	0.03	−4.41	0.08
No adj.	0.16	0.10	−0.15	0.06
B&H	0.02	0.02	0.02	0.02

Table 6.4: Average out-of-sample Sterling ratios by transaction costs with $C = 1$. Average is calculated over the monthly out-of-sample (May 2003 to December 2005) sets. Sterling ratio is a risk-adjusted return indicator calculated as the mean of return divided by the maximum drawdown. Higher values imply higher return. At difference of the Sharpe ratio, the Sterling ratio captures large variations of profits or losses as can be observed in this table.

chine learning algorithm that makes the prediction, a weighting algorithm that combines the experts, and a risk management layer that selects only the strongest prediction and avoids trading when there is a history of negative performance. Every component of the trading system is important to obtain positive abnormal returns, and brings some functionality that is complemented by the rest of the layers. We find that even a very efficient learning algorithm, such as boosting, still requires powerful control mechanisms in order to reduce unnecessary and unprofitable trades that increase transaction costs. Hence, the contribu-

tion of new predictive algorithms by the computer science or machine learning literature to finance still needs to be incorporated under a formal framework of risk management.

As part of the optimization of the trading system, we propose a method to simultaneously calculate the same features using different parameters, leaving the final feature selection to boosting. Many trader systems become very inefficient because they try all the parameters or are forced to select in advance parameters that are not adequate after a trading period. Our experiments show that the boosting approach is able to improve the predictive capacity when indicators are combined and aggregated as a single predictor. Even more, the combination of indicators of different stocks was demonstrated to be adequate in reducing the use of computational resources, while still maintaining an adequate predictive capacity.

Chapter 7

Final comments and conclusions

The main contribution of this dissertation is the definition of a methodology to apply several learning algorithms–mainly Adaboost, Logitboost, and link mining–to automate two main corporate finance functions: strategic planning and trading. Both of these activities require a significant amount of expensive corporate resources (highly trained financial analysts, traders and computer equipment) that could be reduced if at least certain stages of the planning and trading processes are automated using machine learning algorithms.

From the planning point of view, the complexity of large organizations and the significant amount of data implies that those organizations that implement the most efficient methods to manage this information would have a competitive advantage. From the trading point of view, the development of electronic financial markets requires the development of efficient algorithms that can process significant amounts of diverse information, and make investment decisions in seconds. This is a field where computer science can contribute to solve finance's problems. Financial economists design economic models with leading indicators to predict prices. However, these indicators may not be profitable if they are not included in an

efficient algorithm that is able to transform several indicators or a mixture of them into instantaneous trading signals. The main findings of the preceding chapters have shown that learning algorithms can be applied to automated planning to 1) develop a representative ADT algorithm, and 2) automate the generation of BSC using boosting. Also, learning algorithms can be applied to forecasting and automated trading in the following areas: 3) developing a link mining algorithm to integrate accounting and social network variables, 4) identifying new predictive indicators using social networks of directors and analysts, 5) exploring a trading strategy in a controlled competition, and 6) developing a multi-stock automated trading system. In the following section I comment on each of these factors, and in the last section I discuss limitations, recommendations, and future work.

7.1 Contributions

7.1.1 Automated planning system

1. **Representative ADT algorithm:** We developed an algorithm that ranks variables according to their level of importance in the ADTs, and generates representative ADTs with the most important variables.

   This research showed that Adaboost performed similarly to logistic regression, random forests, and bagging with stable datasets when we compared small and large samples from different countries and economic conditions. Additionally, we showed how boosting and representative ADTs can be used as interpretative tools to evaluate the impact of corporate governance factors on performance and efficiency. Representative ADTs are particularly useful to understand the non-linear relationship between

the variables that affects performance and efficiency.

2. **Performance management and automated planning system:** We demonstrated that the representative ADT is a useful tool for selecting and establishing the relationship among the most important indicators of a board BSC. Additionally, the thresholds of the representative ADT establish targets or ranges of values of the indicators that managers could follow to improve corporate performance. With this combined tool, managers can concentrate on the most important strategic issues and delegate the calculation of the targets to an automated planning system supported by Adaboost.

We think that this is an important contribution for the machine learning and performance management literature because while the BSC has been widely applied to S&P 500 companies, no other machine learning algorithm has been used for this purpose. There are very few studies defining methods to quantify a BSC and all methods require significant participation of human experts. Additionally, the methodology described in chapter 3 contributes to automating the function of strategic planning which is one of the major responsibilities of the board of directors. Most organizations have a method to control the performance of intermediate and senior managers. However, the supervision of the board of directors is the responsibility of shareholders and regulatory organizations. Besides the legal periodic reports, there are very limited mechanisms that shareholders, regulators or directors themselves can use to automatically establish standards to evaluate the performance of the board. The generation of a board BSC using boosting covers this gap.

7.1.2 Forecasting and automated planning system

1. **Link mining algorithm and earnings forecast:** We integrated boosting with a link mining algorithm, CorpInterlock, to demonstrate that the relationship between analysts and directors improves cumulative abnormal return predictions during "bull" markets or in periods characterized by a great number of IPOs, mergers and acquisitions. This role is less important in a "bear" market, especially after the Sarbanes-Oxley Act. We established that CorpInterlock implemented with Logitboost improves the prediction of earnings surprise in relation to the implementation of CorpInterlock with logistic regression.

   We found that CorpInterlock is a flexible mechanism for increasing the explanatory power of social networks with the forecasting capability of machine learning algorithms, such as boosting. The capacity to improve the forecast of earnings surprises and abnormal return using a mixture of well-known economic indicators with social network variables also enriches the debate between the modern finance theory and behavioral finance to show how behavioral patterns can be recognized under a rigorous method of analysis and forecast.

2. **Small world and corporate interlock:** This research showed that the basic and extended corporate interlocks have the properties of a "small world" network. The statistics of an extended corporate interlock, directors and financial analysts, bring additional information to predict cumulative abnormal return, especially during a "bull" market.

3. **Constant rebalanced portfolio - technical analysis trading algorithm (CRP_TA):**

We propose the algorithm CRP_TA that combines intraday trading based on constant rebalanced portfolio with daily price forecast based on technical analysis indicators. This algorithm was profitable during the PLAT competition, and after the competition we enhanced it by including a market maker component. We showed that the constant rebalanced portfolio can improve if a classifier can anticipate the direction of the market. Additionally, transaction costs play a central role in raising performance. Instead of an automatic rebalance of the portfolio, the results of the PLAT competition indicated that if the CRP strategy is implemented only with limit orders, its results improve because of the rebates.

4. **Automated trading system:** We proposed a multi-stock automated trading system. The system is designed to trade stocks, and relies on a layered structure consisting of ADT, which is implemented with Logitboost as the machine learning algorithm; an online learning utility; and a risk management overlay. The system generates its own trading rules, and weights the answers of the different ADTs or experts to suggest a trading position. Finally, the risk management layer can validate the trading signal when it exceeds a specified non-zero threshold, and limit the use of a trading strategy when it is not profitable.

We tested the expert weighting algorithm with data of 100 randomly selected companies of the S&P 500 index during the period 2003-2005. We found that this algorithm generated excess returns during the test period, and every component of the trading algorithm is important to obtain positive abnormal returns, and brings some functionality that is complemented by the rest of the layers. We found that even an

efficient learning algorithm, such as boosting, still requires powerful control mechanisms in order to reduce unnecessary and unprofitable trades that increase transaction costs. Hence, the contribution of new predictive algorithms by the computer science or machine learning literature to finance still needs to be incorporated under a formal framework of risk management.

As part of the optimization of the trading system, we proposed a method to simultaneously calculate the same features using different parameters, leaving the final feature selection to boosting. Many trader systems become very inefficient because they try all the parameters or are forced to select in advance parameters that are not adequate after a trading period. Our experiments show that the boosting approach is able to improve the predictive capacity when indicators are combined and aggregated as a single predictor. Even more, the combination of indicators of different stocks was shown to reduce the use of computational resources, and still maintain a satisfactory predictive capacity.

7.2   Limitations, recommendations, and future work

Economists are generally very suspicious of the so-called "data mining" methods because of the risk of overfitting, and the lack of parameters' estimation and interpretability of results. This perceptual barrier, and the limited knowledge of learning algorithms by economists have slowed down the incorporation of machine learning methods into mainstream econometrics. In relation to the first objection of economists, we have indicated how to prevent overfitting in the case of boosting. For instance, we used a validation set to select a number

of iterations, and the value of other parameters that are associated with a reduction of the test error without overfitting. In all the experiments conducted with the test set, we used the same number of iterations and parameters that we selected when we used the validation set. In relation to the second objection, lack of parameter estimation and interpretability, we have also shown how boosting can be used as an interpretative method instead of a "black box" that simply forecasts without a clear understanding of the underlying rules. The use of boosting as an interpretative tool is possible because of the generation of a representative ADT where main variables are selected together with an ordered relationship.

We think that economists' questions about algorithmic modeling are reasonable because of previous naive or straightforward applications of learning methods to the financial markets. Our experience in adapting boosting to finance problems is that a simple and straightforward application of boosting to finance does not bring a significant improvement in forecasting as we showed in chapter 3. There are other well-known methods used for finance problems, such as logistic regression, that have a similar performance to boosting. However, boosting can work with a mixture of quantitative and qualitative indicators, and also with non-linear time series. Furthermore, boosting can be used to understand the non-linear relationship between the variables as we mentioned above, and can automatically select the best features. Based on these factors, we strongly recommend the inclusion of boosting as one of the methods used by econometricians or applied economists for prediction or for finding causal relationships among variables of large datasets.

An important aspect that we observed in our experiments is that the calibration of machine learning methods requires a significant amount of time. In chapter 3 we proposed a method that leads to an automatic selection of features and parameters by boosting.

The application of this mechanism improved our results, and significantly reduced the time of adaptation of boosting to new problems. Boosting's predictions also depended of the quality and value of the indicators used. Hence, a direct application of boosting without the adequate calibration of the model or search for the best indicators may generate a very poor prediction. Based on this, we also included new indicators to predict earnings surprises, such as social network statistics.

Additionally, we recognize that boosting or another learning algorithms used to forecast time series may have a predictive ability for only a certain period of time. However, the randomness and continuous change of the financial market may cause a trading strategy based on boosting or another predictor to become ineffective. Hence, we are also extremely suspicious of trading systems that are based on learning algorithms and show spectacular results in certain periods of time, without incorporating risk management mechanisms. Even though several hedge funds and investment banks are currently using learning algorithms to find trading strategies, they do not rely only in the learning algorithmic component. These funds or banks have strict risk management systems based on their practice and finance theory. Hence, we recommend the implementation of trading strategies using learning algorithms only as part of a trading system based on sound finance principles.

The automated trading system proposed in this dissertation could easily be adapted to other domains such as foreign exchange, fixed income, or the international equity market. Future research could be directed to evaluating the specific risk management mechanisms or alternative methods for selecting the most profitable trading strategies. Additionally, the comparison of boosting with other machine learning methods in different domains may bring new light about the strengths and weaknesses of each method. For instance, the

international market is much more volatile and exposed to many more risk factors than the US market. A trading system for the international equity market should be able to deal with these heterogeneous conditions.

A potential extension of this research is the application of boosting for portfolio selection using the Black Litterman model [34, 35]. This model includes the subjective expectations of investors in a risk variance optimization model. An alternative line of research is to use the scores of boosting instead of the subjective expectations of the investors. This approach would combine the optimal predictive capability of boosting with a risk return optimization model.

Finally, this research can also be extended using boosting for the design of the enterprise BSC, and including other perspectives of those reviewed in this study. Initially, the corporate governance variables did not seem to be very relevant to predicting corporate performance. However, when the results of these variables were interpreted together with the accounting variables using representative ADTs, the effect of corporate governance on performance became evident as the BSC demonstrated. A similar situation may happen with the variables of the other perspectives of the BSC. The recent cases of US bankruptcies have demonstrated that when companies are doing very well, corporate governance variables do not seem to be relevant. However, in moments of financial distress, corporate governance variables play a very important role in improving performance and efficiency. In this respect, another future direction for this research line is the evaluation of the abnormal return of two portfolios with top and bottom tier companies based on the suggestions of the representative ADTs and board BSC. Additionally, the combination of Adaboost with the BSC can be used as a automated strategic planning system that continuously updates

itself for board-level decisions of directors or for investment decisions of portfolio managers.

Bibliography

[1] J. Abarbanell. Do analysts earnings forecasts incorporate information in prior stock price changes? *Journal of Accounting and Economics*, 14:1471–65, 1991.

[2] J. Abarnabell and V. Bernard. Tests of analysts' overreaction/ underreaction to earnings information as an explanation for anomalous stock price behavior. *Journal of Finance*, 47(3):1181–1207, 1992.

[3] V. V. Acharya, K. John, and R. K. Sundaram. On the optimality of resetting executive stock options. *Journal of Financial Economics*, 57(1):65–101, 2000.

[4] D. Aigner, C. Lovell, and P. Schmidt. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6:21–37, 1977.

[5] T. Al-Farj, A. Alidi, and K. Bu-Bshait. Evaluation of bank branches by means of data envelopment analysis. *International Journal of Operations and Production Management*, 15, 1993.

[6] S. Alexander. Price movements in speculative markets: trends or random walks. *Industrial Management Review*, 2:7–26, 1961.

[7] P. H. Algoet and T. M. Cover. Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. *Annals of Probability*, 16:876–898, 1988.

[8] F. Allen and R. Karjalainen. Using genetic algorithms to find technical trading rules. *Journal of Financial Economics*, 51(2):245–271, 1999.

[9] E. I. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4):589–609, Sep 1968.

[10] E. I. Altman. Evaluation of a company as a going concern. *Journal of Accountacy*, pages 50–57, 1974.

[11] E. I. Altman. Measuring corporate bond mortality and performance. *Journal of Finance*, 54, 1989.

[12] E. I. Altman, J. B. Caouette, and P. Narayanan. Credit-risk measurement and management: The ironic challenge in the next decade. *Financial Analysis Journal*, pages 7–11, 1998.

[13] M. Ampuero, J. Goransson, and J. Scott. Solving the measurement puzzle. how EVA and the Balanced Scorecard fit together. *Perspectives on Business Innovation*, 1998.

[14] J. Ang, R. Cole, and J. W. Lin. Agency costs and ownership structure. *Journal of Finance*, 40:81–106, 2000.

[15] J. Arifovic. The behavior of the exchange rate in the genetic algorithm and experimental economies. *Journal of Political Economy*, 104(3):510–541, 1996.

[16] W. B. Arthur, J. H. Holland, B. LeBaron, R. Palmer, and P. Talyer. Asset Pricing Under Endogenous Expectations in an Artificial Stock Market. In W. Arthur, S. Durlauf, and D. Lane, editors, *The Economy as an Evolving Complex System II*, pages 15–44. "Addison-Wesley", "Reading, MA", 1997.

[17] P. Asquith, M. B. Mikhail, and A. S. Au. Information content of equity analyst reports. *Journal of Financial Economics*, 75(2):245–282, 2005.

[18] A. Athanassopoulos. Service quality and operating efficiency synergies for management control in the provision of financial services: Evidence from Greek bank branches. *European Journal of Operational Research*, 98(2):300–313, 1997.

[19] A. Barabasi. *Linked: The New Science of Networks*. Perseus, Cambridge, MA, 2002.

[20] B. Barber, R. Lehavy, M. McNichols, and B. Trueman. Can investors profit from the prophets? security analysts recommendations and stock returns. *Journal of Finance*, 56(2):531–563, 2001.

[21] R. Barr, L. M. Seiford, and F.Thomas. Forecasting bank failure: A non-parametric frontier estimation approach. *Recherches Economiques de Louvain*, 60(4), 1994.

[22] R. Bates, M. Dempster, and Y. Romahi. Evolutionary reinforcement learning in FX order book and order flow analysis. In *Proceedings of the IEEE International Conference on Computational Intelligence for Financial Engineering*, pages 355–362, 2003.

[23] P. Bauer, A. Berger, G. Ferrier, and D. Humphrey. Consistency conditions for regulatory analysis of financial institutions: A comparison of frontier efficiency methods. *Journal of Economics and Business*, 50(2):85–114, 1998.

[24] L. Bauwens and P. Giot. *Econometric Modelling of Stock Market Intraday Activity.* Kluwer Academic Publishers, Boston, 2001.

[25] B. Baysinger and H. Butler. Corporate governance and the board of directors: performance effects of changes in board composition. *Journal of Law, Economics and Organization*, 1:101–204, 1985.

[26] W. Beaver. Financial ratios as predictors of failure. *Journal of Accounting Research*, 4:71–111, 1966.

[27] S. Beckers, M. Steliaros, and A. Thomson. Bias in European analysts' earnings forecasts. *Financial Analysts Journal*, 60(2):74–85, 2004.

[28] A. Berger and D. Humphrey. Efficiency of financial institutions: International survey and directions for future research. *European Journal of Operational Research*, 98(2):175–212, 1997.

[29] A. Berle and G. Means. *The Modern Corporation and Private Property.* Harcourt, New York, 1932.

[30] V. L. Bernard and J. K. Thomas. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, 13(4), 1990.

[31] S. Bhagat and B. Black. The uncertain relationship between board composition and firm performance. *Business Lawyer*, 54:921–963, 1999.

[32] S. Bhagat and B. Black. The non-correlation between board independence and long-term firm performance. *Journal of Corporation Law*, 27:231–273, 2002.

[33] A. Bhattacharyya, C. Lovell, and P. Sahay. The impact of liberalization on the productive efficiency of Indian commercial banks. *European Journal of Operational Research*, 98(2):332–345, 1997.

[34] F. Black and R. Litterman. Asset allocation: Combining investor views with market equilibrium. Fixed income research, Goldman Sachs & Co., September 1990.

[35] F. Black and R. Litterman. Global asset allocation with equities, bonds, and currencies. Fixed income research, Goldman Sachs & Co., October 1991.

[36] A. Blum and A. Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999. Special Issue for COLT '97.

[37] T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[38] J. A. Bollinger. *Bollinger on Bollinger Bands*. McGraw-Hill, New York, 2001.

[39] S. P. Borgatti and M. Everett. A graph-theoretic perspective on centrality. *Social Networks*, 2006. In press.

[40] B. D. Borger, G. D. Ferrier, and K. Kerstens. The choice of a technical efficiency measure on the free disposal Hull reference technology: a comparison using US banking data. *European Journal of Operational Research*, 105(3):427–446, 1998.

[41] S. Bornholdt. Expectation bubbles in a spin model of markets: Intermittency from frustration across scales. *International Journal of Modern Physics C*, 12:667, 2001.

[42] A. Borodin and R. El-Yaniv. *Online computation and competitive analysis*. Cambridge University Press, Cambridge, 1998.

[43] V. Boyarshinov. *Machine learning in computational finance*. PhD thesis, Rensselaer Polytechnic Institute, 2005.

[44] L. Breiman. Bagging predictors. *Machine Learning*, 24:123–140, 1996.

[45] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[46] L. Breiman. Statistical modeling: The two cultures. *Statistical Science*, 16:199–231, 2001.

[47] L. Breiman, J. H. Friedman, R. Olshen, and C. Stone. *Classification and Regression Trees*. Wadsworth and Brooks, Belmont, 1984.

[48] M. Brenner, R. K. Sundaram, and D. Yermack. Altering the terms of executive stock options. *Journal of Financial Economics*, 57(1):103–128, 2000.

[49] L. D. Brown. *I/B/E/S Research Bibliography*. I/B/E/S International Incorporated, 6 edition, 2000.

[50] L. D. Brown. How important is past analyst forecast accuracy? *Financial Analysts Journal*, 57(6):44–49, 2001.

[51] L. D. Brown, J. C. Y. Han, E. K. Jr., and W. H. Quinn. Predicting analysts' earnings surprise. *Journal of Investing*, 5(1):17–23, 1996.

[52] J. A. Byrne. The best and worst boards. *Business Week*, 1996.

[53] N. Cesa-Bianchi, Y. Freund, D. Haussler, D. P. Helmbold, R. E. Schapire, and M. K. Warmuth. How to use expert advice. *Journal of the ACM*, 44(3"):427–485, may 1997.

[54] M. Chafai. Estimating input-specific technical inefficiency: The case of the Tunisian banking industry. *European Journal of Operational Research*, 98(2):314–331, 1997.

[55] M. Chaikin. How to find big winners using persistency of money flow. CD, NA.

[56] N. Chan, B. LeBaron, A. Lo, and T. Poggio. Agent-based models of financial markets: A comparison with experimental markets. Technical Report 124, MIT Artificial Markets Project, 1999.

[57] T. Chan. *Artificial markets and intelligent agents*. PhD thesis, Massachusetts Institute of Technology, 2001.

[58] D. M. Chance, R. Kumar, and R. B. Todd. The 'repricing' of executive stock options. *Journal of Financial Economics*, 57(1):129–154, 2000.

[59] T. S. Chande and S. Kroll. *The New Technical Trader: Boost your profit by plugging into the latest indicators*. John Wiley & Sons, Inc., New York, 1994.

[60] A. Charnes, W. Cooper, and E. Rhodes. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2:429–44, 1978.

[61] K. Chen and J. Lee. Accounting measures of business performance and Tobin's Q theory. *Journal of Accounting, Auditing and Finance*, 10(3):587–609, 1995.

[62] V. Cho, B. Wuthrich, and J. Zhang. Text processing for classification. *Journal of Computational Intelligence in Finance*, 1998.

[63] D. N. Chorafas. *Chaos Theory in the Financial Markets*. Probus publishing, Chicago, 1994.

[64] K. Chung and S. W. Pruitt. A simple approximation to Tobin's Q. *Financial Management*, 23(3):70–74, 1994.

[65] D. G. Clarke and J. B. McDonald. Generalized bankruptcy models applied to predicting consumer credit behavior. *Journal of Economics and Business*, 44(1):47–62, 1992.

[66] J. F. Clayburg. *Four Steps to Trading Success: Using everyday indicators to achieve extraordinary profits*. John Wiley & Sons, Inc., New York, 2001.

[67] M. Clement and S. Tse. Financial analyst characteristics and herding behavior in forecasting. *Journal of Finance*, 40(1):307–341, 2005.

[68] D. Clinton, S. Webber, and J. Hassell. Implementing the Balanced Scorecard using the analytic hierarchic process. *Management Accounting Quarterly*, 3(3):1–11, 2002.

[69] M. Collins, R. E. Schapire, and Y. Singer. Logistic regression, adaboost and Bregman distances. *Machine Learning*, 48(1-3):253–285, 2004.

[70] R. Collins and R. D. Green. Statistical methods for bankruptcy forecasting. *Journal of Economics and Business*, 34:349–354, 1982.

[71] J. E. Core, R. W. Holthausen, and D. F. Larcker. Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics*, 51(3):371–406, 1999.

[72] T. Cover and E. Ordentlich. Universal portfolios with side information. *IEEE Transactions on Information Theory*, 42(2), March 1996.

[73] T. M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

[74] G. Crystal. *In Search of Excess: The overcompensation of American executives*. W. W. Norton Company, New York, 1991.

[75] G. Davis. Agents without principles? the spread of the poison pill through the intercorporate network. *Administrative Science Quarterly*, 36:586–613, 1991.

[76] G. Davis, M. Yoo, and W. Baker. The small world of the american corporate elite, 1982-2001. *Strategic Organization*, 1(3):301–326, 2003.

[77] W. de Nooy, A. Mrvar, and V. Batagelj. *Exploratory social network analysis with Pajek*. Cambridge University Press, New York, 2005.

[78] G. J. Deboeck, editor. *Trading on the Edge: Neural, genetic, and fuzzy systems for chaotic financial markets*. John Wiley & Sons, Inc., New York, 1994.

[79] K. Decker, K. Sycara, and D. Zeng. Designing a multi-agent portfolio management system, 1996.

[80] M. Dempster and V. Leemans. An automated FX trading system using adaptive reinforcement learning. *Expert Systems with Applications: Special issue on financial engineering*, 30:534–552, 2006.

[81] M. Dempster, T. W. Payne, Y. Romahi, and G. Thompson. Computational learning techniques for intraday FX trading using popular technical indicators. *IEEE Transactions on neural networks*, 12:744–754, 2001.

[82] M. Dempster and Y. Romahi. Intraday FX trading: An evolutionary reinforcement learning approach. In *Third International Conference on Intelligent Data Engineering and Automated Learning IDEAL 02, Manchester, UK, August 12-14, 2002. Proceedings*, volume 2412 of *Lecture Notes in Computer Science*, pages 347–358. Springer, 2002.

[83] H. Demsetz. The structure of ownership and the theory of the firm. *Journal of Law and Economics*, pages 375–393, 1983.

[84] D. Denis and A. Sarin. Ownership and board structures in publicly traded corporations. *Journal of Financial Economics*, 52:187–223, 1999.

[85] R. DeYoung. A diagnostic test for the distribution-free efficiency estimator: An example using U.S. commercial bank data. *European Journal of Operational Research*, 98(2):243–249, 1997.

[86] V. Dhar and D. Chou. A comparison of nonlinear methods for predicting earnings surprises and returns. *IEEE Transactions on Neural Networks*, 12(4):907–921, 2001.

[87] J. F. Ehlers. *Rocket Science for Traders: Digital signal processing applications*. John Wiley & Sons, Inc., New York, 2001.

[88] N. Ehrentreich. The Santa Fe artificial stock market re-examined - suggested corrections. Computational Economics 45/02, Martin Luther Universitat Halle Wittenberg, Sept. 2002.

[89] R. El-Yaniv. Competitive solutions for online financial problems. *ACM Computing Surveys*, 30(1):28–69, 1998.

[90] J. E. Elton, M. J. Gruber, and S. Grossman. Discrete expectational data and portfolio performance. *Journal of Finance*, 41:699–714, 1986.

[91] R. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1006, 1982.

[92] M. J. Epstein and M.-J. Roy. Measuring and improving the performance of corporate boards. Technical report, The Society of Management Accountants of Canada, 2002.

[93] M. J. Epstein and M.-J. Roy. How does your board rate? *Strategic Finance*, pages 25–31, February 2004.

[94] E. Fama. Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25:383–417, 1970.

[95] E. Fama. Agency problems and the theory of the firm. *Journal of Political Economy*, 88:288–307, 1980.

[96] E. Fama and M. Blume. Filter rules and stock market trading. security prices: a supplement. *Journal of Business*, 39:226–241, 1970.

[97] E. Fama and K. French. The cross-section of expected stock returns. *Journal of Finance*, 47:427–465, 1992.

[98] E. Fama and M. Jensen. Separation of ownership and control. *Journal of Law and Economics*, 26:301–325, 1983.

[99] T. Fawcett and F. Provost. Activity monitoring: Noticing interesting changes in behavior. In *Proceedings of the Fifth ACM SIGKDD International conference on knowledge discovery and data mining (KDD-99)*, pages 53–62, 1999.

[100] C. A. Finger and W. R. Landsman. What do analysts' stock recommendations really mean? Working paper, University of Illinois and U.N.C. Chapel Hill, 1999.

[101] N. Fosback. *Stock Market Logic: A sophisticated approach to profits on Wall Street.* Dearborn Trade Publishing, Chicago, 1991.

[102] L. Freeman. Centrality in networks: I. conceptual clarification. *Social Networks*, 1:215–39, 1979.

[103] Y. Freund, Y. Mansour, and R. Schapire. Generalization bounds for averaged classifiers. *The Annals of Statistics*, 32(4), 2004.

[104] Y. Freund and L. Mason. The alternating decision tree learning algorithm. In *Machine Learning: Proceedings of the Sixteenth International Conference*, pages 124–133, 1999.

[105] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

[106] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: A statistical view of boosting. *The Annals of Statistics*, 38(2):337–374, apr 2000.

[107] O. Fuerst and S. Kang. Corporate governance, expected operating performance and pricing. Corporate Ownership and Control, 2004.

[108] M. Garman and M. Klass. On the estimation of security price volatilities from historical data. *Journal of Business*, 53:67–78, 1980.

[109] L. Getoor and C. P. Diehl. Link mining: A survey. *SIGKDD Explorations*, 7(2):3–12, 2005.

[110] D. K. Gode and S. Sunder. Allocative efficiency of markets with zero intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119–137, 1993.

[111] H. G. Goldberg, J. D. Kirkland, D. Lee, P. Shyr, and D. Thakker. The NASD securities observation, news analysis and regulation system (sonar). In *IAAI 2003*, 2003.

[112] P. A. Gompers, J. L. Ishii, and A. Metrick. Corporate governance and equity prices. *Quarterly Journal of Economics*, 118(1):107–155, 2003.

[113] S. Goonatilake and P. Treleaven, editors. *Intelligence Systems for Finance and Business.* John Wiley & Sons, Inc., New York, 1995.

[114] A. Goudie and G. Meeks. Forecasting and explaining corporate failure: A study of contingent financial statements. Technical report, University of Cambridge, Department of Economics, 1992.

[115] J. Granville. *Granville's New Key to Stock Market Profits.* Prentice Hall, Upper Saddle River, 2000.

[116] E. Grifell-Tatje and C. Lovell. The sources of productivity change in Spanish banking. *European Journal of Operational Research*, 98(2):364–380, 1997.

[117] S. Grossman and J. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70:393–408, 1980.

[118] M. J. Hall, editor. *The Regulation And Supervision of Banks: Regulation and efficiency in banking*, volume 4. Edward Elgar Pub, Cheltenham, 2001.

[119] C. R. Harvey, K. V. Linsc, and A. H. Roper. The effect of capital structure when expected agency costs are extreme. *Journal of Financial Economics*, 74(1):3–30, 2004.

[120] T. Hastie, R. Tibishirani, and J. Friedman. *The Elements of Statistical Learning.* Springer, New York, 2003.

[121] B. Hermalin and M. Weisbach. The effects of board composition and direct incentives on firm performance. *Financial Management*, 20:101–112, 1991.

[122] S. A. Hillegeist and F. Pealva. Stock option incentives and firm performance. IESE Research Papers D/535, IESE Business School, Jan. 2004. available at http://ideas.repec.org/p/ebg/iesewp/d-0535.html.

[123] C. Himmelberg, G. Hubbard, and D. Palia. Understanding the determinants of managerial ownership and the link between ownership and performance. *Journal of Financial Economics*, 53:353–384, 1999.

[124] A. L. Hing-Ling. A five-state financial distress prediction model. *Journal of Accounting Research*, 25:127–138, 1987.

[125] H. G. Hong and J. D. Kubik. Analyzing the analysts: Career concerns and biased earnings forecasts. *Journal of Finance*, 58(1):313–351, 2003.

[126] J. Hudson. Company bankruptcies and births matter. *Applied Economics*, 29:647–654, 1997.

[127] E. Ising. Beitrag zur theorie des ferromagnetismus. *Z. Physik*, 31:253–258, 1925.

[128] Z. Ivkovic and N. Jegadeesh. The timing and value of forecast and recommendation revisions. *Journal of Financial Economics*, 73(3):433–463, 2004.

[129] N. Jegadeesh, J. Kim, S. D. Krische, and C. M. C. Lee. Analyzing the analysts: When do recommendations add value? *Journal of Finance*, 59(3):1083–1124, 2004.

[130] M. Jensen. The modern industrial revolution, exit, and the failure of internal control systems. *Journal of Finance*, 48:831–880, 1993.

[131] M. Jensen and W. H. Meckling. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3:305–360, 1976.

[132] M. Jensen and K. J. Murphy. Performance pay and top management incentives. *Journal of Political Economy*, 98(2):225–265, 1990.

[133] M. Jensen and K. J. Murphy. CEO incentives: It's not how much you pay, but how. *Harvard Business Review*, 68(3):138–153, 1990.

[134] M. Jensen and K. J. Murphy. Remuneration: Where we ve been,how we got to here,what are the problems,and how to fix them. Finance Working Paper 44/2004, European Corporate Governance Institute, July 2004.

[135] S. M. Kakade, M. Kearns, Y. Mansour, and L. E. Ortiz. Competitive algorithms for VWAP and limit order trading. In *Proceedings of the 5th ACM conference on Electronic commerce*, pages 189–198. ACM Press, 2004.

[136] A. Kalai and S. Vempala. Efficient algorithms for universal portfolios. *Journal of Machine Learning Research*, 3:423–440, 2003.

[137] R. S. Kaplan and M. E. Nagel. Improving corporate governance with the Balanced Scorecard. *Directors Monthly*, pages 16–20, March 2004.

[138] R. S. Kaplan and D. P. Norton. The Balanced Scorecard: measures that drive performance. *Harvard Business Review*, 70(1):72–79, 1992.

[139] R. S. Kaplan and D. P. Norton. Putting the Balanced Scorecard to work. *Harvard Business Review*, 71(5), 1993.

[140] R. S. Kaplan and D. P. Norton. Linking the Balanced Scorecard to strategy. *California Management Review*, 49(1):53–79, 1996.

[141] R. S. Kaplan and D. P. Norton, editors. *The Balanced Scorecard*. Harvard Business School Press, Boston, 1996.

[142] R. S. Kaplan and D. P. Norton. Using the Balanced Scorecard as a strategic management system. *Harvard Business Review*, 74(1):75–85, 1996.

[143] J. Katz and D. McCormick. *The Encyclopedia of Trading Strategies*. McGraw-Hill, New York, 2000.

[144] K. A. Kavajecz and E. R. Odders-White. Technical analysis and liquidity provision. *Review of Financial Studies*, 2004.

[145] M. Kearns and L. Ortiz. The Penn-Lehman Automated Trading Project. *IEEE Intelligent Systems*, 2003.

[146] J. L. Kelly. A new interpretation of information rate. *Bell System Technical Journal*, 35:917–926, 1956.

[147] J. D. Kirkland, T. E. Senator, J. J. Hayden, T. Dybala, H. G.Goldberg, and P. Shyr. The nasd regulation advanced detection system (ads). *AI Magazine*, 20(1):55–67, 1999.

[148] A. Klein. Firm performance and board committee structure. *Journal of Law and Economics*, 41:275–303, 1998.

[149] S. D. Krische and C. M. C. Lee. The information content of analyst stock recommendations. Working paper, Cornell University, 2000.

[150] R. La-Porta, F. L. de Silanes, and A. Shleifer. Corporate ownership around the world. *Journal of Finance*, 54:471–517, 1999.

[151] R. La-Porta, F. L. de Silanes, A. Shleifer, and R. Vishny. Law and finance. *Journal of Political Economy*, 196:1113–1155, 1998.

[152] R. La-Porta, F. L. de Silanes, A. Shleifer, and R. Vishny. Investor protection and corporate valuation. *Journal of Finance*, 57:1147–1170, 2002.

[153] W. Lane, S. W. Looney, and J. W. Wansley. An application of the Cox proportional hazards model to bank failure. *Journal of Banking and Finance*, 10:511–531, 1986.

[154] D. F. Larcker, S. A. Richardson, A. J. Seary, and I. Tuna. Back door links between directors and executive compensation. Working paper, February 2005.

[155] V. Lavrenko, M. Schmill, D. Lawrie, P. Oglivie, D. Jensen, and J. Allan. Language models for financial news recommendation. In *Proceedings of the International Conference on Information and Knowledge Management*, pages 389–396, 2000.

[156] B. Le Baron. Agent based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control*, 24:679–702, 2000.

[157] B. Le Baron, W. B. Arthur, and R. Palmer. The time series properties of an artificial stock market. *Journal of Economic Dynamics and Control*, 21:1487–1516, 1998.

[158] B. LeBaron. An evolutionary bootstrap method for selecting dynamic trading strategies. In *Refenes, A.-P. N., Burgess, A. and Moody, J. eds., Decision Technologies for Computational Finance, Proceedings of the Fifth International Conference Computational Finance*, pages 141–160. Kluwer Academic Publishers, 1998.

[159] B. LeBaron. Empirical regularities from interacting long and short memory investors in an agent-based financial market. *IEEE Transactions on Evolutionary Computation*, 5:442–455, 2001.

[160] C. I. Lee, L. Rosenthal, and K. C. Gleason. Effect of Regulation FD on asymmetric information. *Financial Analysts Journal*, 60(3):79–89, 2004.

[161] H. Leibenstein. Allocative efficiency vs "x-efficiency". *American Economic Review*, 56:392–415, June 1966.

[162] M. Lettau. Explaining the facts with adaptive agents: The case of mutual funds flows. *Journal of Economic Dynamics and Control*, 21:1117–1148, 1997.

[163] E. Lindenberg and S. A. Ross. Tobin's Q ratio and industrial organization. *Journal of Business*, 54:1–32, 1981.

[164] M. Lipton and J. Lorsch. A modest proposal for improved corporate governance. *Business Lawyer*, 48:59–77, 1992.

[165] N. Littlestone and M. Warmuth. The weighted majority algorithm. *Information and Computation*, 108:212–261, 1994.

[166] A. Lo, H. Mamaysky, and J. Wang. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *Journal of Finance*, 4:1705–1765, 2000.

[167] C. Lovell. Production frontiers and productive efficiency. In *The Measurement of Productive Efficiency*. Oxford University Press, New York, 1993.

[168] A. Lozano. Profit efficiency for Spanish savings banks. *European Journal of Operational Research*, 98(2):381–394, 1997.

[169] P. MacAvoy, S. Cantor, J. Dana, and S. Peck. ALI proposals for increased control of the corporations by the board of directors: An economics analysis, 1983.

[170] J. McConnell and H. Servaes. Additional evidence on equity ownership and corporate value. *Journal of Financial Economics*, 27:595–612, 1990.

[171] R. R. Mendenhall. Evidence on the possible underweighting of earnings information. *Journal of Accounting Research*, 29(1):170–179, 1991.

[172] L. Mester. Measuring efficiency at U.S. banks. *European Journal of Operational Research*, 98(2):230–242, 1997.

[173] M. B. Mikhail, B. Walther, and R. Willis. Do security analysts exhibit persistent differences in stock picking ability? *Journal of Financial Economics*, 74(1):67–91, 2002.

[174] W. H. Mikkelson and R. Ruback. An empirical analysis of the interim equity investment process. *Journal of Financial Economics*, 14:523–553, 1985.

[175] S. Milgram. The small world problem. *Psychology Today*, 2:60–7, 1967.

[176] C. Mills. *The Power Elite.* Oxford Press, New York, 1956.

[177] B. Mintz and M. Schwartz. *The Power Structure of American Business.* University of Chicago Press, Chicago, 1985.

[178] M. Mizruchi. *The Structure of Corporate Political Action: Interfirm relations and their consequences.* Harvard University Press, Cambridge, MA, 1992.

[179] J. Moody and M. Saffell. Learning to trade via direct reinforcement. *IEEE-NN*, 12:875–889, jul 2001.

[180] R. Morck, A. Shleifer, and R. Vishny. Management ownership and corporate performance: An empirical analysis. *Journal of Financial Economics*, 20:293–316, 1988.

[181] J. Moreno. *Application of the Group Method to Classification.* National Committee on Prisons and Prison Labor, New York, 1932.

[182] C. R. Moyer. Forecasting financial failure: a re-examination. *Financial Management*, 1977.

[183] K. R. Müller, A. J. Smola, G. Rätsch, B. Schölkopf, J. Kohlmorgen, and V. Vapnik. Predicting time series with support vector machines. In *Proceedings of ICANN*, volume 1327, pages 909–1004. Springer, 1997.

[184] M. Newman, S. Strogatz, and D. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(026118), 2001.

[185] M. Newman, D. Watts, and S. Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(2566-2572), 2002.

[186] J. Ohlson. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18:1, 1980.

[187] J. A. Ou and S. H. Penman. Accounting measurement, price-earnings ratios, and the information content of security prices. *Journal of Accounting Research*, 27, 1989.

[188] D. Palia. The endogeneity of managerial compensation in firm valuation: A solution. *Review of Financial Studies*, 14:735–764, 2001.

[189] J. M. Pastor, F. Prez, and J. Quesada. Efficiency analysis in banking firms: An international comparison. *European Journal of Operational Research*, 98(2):395–407, 1997.

[190] S. B. Perfect and K. W. Wiles. Alternative construction of Tobin's Q: An empirical comparison. *Journal of Empirical Finance*, 1:313–41, 1994.

[191] D. Peters. Are earnings surprises predictable? *Journal of Investing*, 2(2):47–51, 1993.

[192] D. Peters. The influences of size on earnings surprise predictability. *Journal of Investing*, 2(4):54–59, 1993.

[193] D. Peterson and P. Peterson. Abnormal returns and analysts earnings forecast revisions associated with the publication of 'stock highlights' by value line investment survey. *Journal of Financial Research*, 18(4):465–477, 1995.

[194] P. Peterson and D. Peterson. *Company Performance and Measures of Value Added*. The Research Foundation of the Institute of Chartered Financial Analysts, Charlottesville, VA, 1996.

[195] G. Pinches and K. Mingo. A multivariate analysis of industrial bond ratings. *Journal of Finance*, 28:1, 1973.

[196] M. Pring. *Technical Analysis Explained*. McGraw-Hill, New York, 4 edition, 2002.

[197] M. Queen and R. Roll. Firm mortality: using market indicators to predict survival. *Financial Analysts Journal*, 43(3):9–26, 1987.

[198] H. Rao, G. Davis, and A. Ward. Embeddedness, social identity and mobility: Why firms leave the NASDAQ and join the New York Stock Exchange. *Administrative Science*, 45:268–92, 2000.

[199] P. S. Rose and G. A. Giroux. Predicting corporate bankruptcy: an analytical and empirical evaluation. *Review of Business and Economics Research*, 19(2), 1984.

[200] S. Rosenstein and J. Wyatt. Outside directors, board independence, and shareholder wealth. *Journal of Financial Economics*, 26:175–191, 1990.

[201] B. R. Routledge. Genetic algorithm learning to choose and use information. *Macroeconomic dynamics*, 5:303–325, 2001.

[202] H. J. Ryan and R. I. Wiggins. Who is in whose pocket? director compensation, board independence, and barriers to effective monitoring. *Journal of Financial Economics*, 73(3):497–524, 2004.

[203] C. Schaffnit, D. Rosen, and J. C. Paradi. Best practice analysis of bank branches: an application of DEA in a large Canadian bank. *European Journal of Operational Research*, 98(2):269–289, 1997.

[204] R. E. Schapire, Y. Freund, P. Bartlett, and W. S. Lee. Boosting the margin: A new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26(5):1651–1686, October 1998.

[205] T. E. Senator. Link mining applications: Progress and challenges. *SIGKDD Explorations*, 7(2):76–83, 2005.

[206] Y.-W. Seo, J. Giampapa, and K. Sycara. Financial news analysis for intelligent portfolio management. Technical Report CMU-RI-TR-04-04, Robotics Institute,Carnegie Mellon University, 2004.

[207] S. . P. G. Services. Standard & poor's corporate governance scores and evaluations, 2004.

[208] C. J. Sherry. *The New Science of Technical Analysis*. Probus publishing, Chicago, 1994.

[209] A. Shivdasani. Board composition, ownership structure and hostile takeovers. *Journal of Accounting and Economics*, 16:167–98, 1993.

[210] A. Shleifer and R. Vishny. A survey of corporate governance. *Journal of Finance*, 2:737–783, 1997.

[211] M. Sparrow. The application of network analysis to criminal intelligence: An assessment of the prospects. *Social Networks*, 13:251–274, 1991.

[212] C. Stenger. The corporate governance scorecard. *Corporate Governance: An international review*, 12(1):11–15, 2004.

[213] S. E. Stickel. The anatomy of the performance of buy and sell recommendations. *Financial Analysts Journal*, 51:25–39, 1995.

[214] T. Stober. Summary financial statements measures and analysts' forecasts of earnings. *Journal of Accounting and Economics*, 15:347–372, 1992.

[215] S. Stolfo, G. Creamer, and S. Hershkop. A temporal based forensic discovery of electronic communication. In *Proceedings of the National Conference on Digital Government Research*, San Diego, California, 2006.

[216] T. Stridsman. *Trading Systems and Money Management*. McGraw-Hill, New York, 2003.

[217] R. Stulz. Managerial control of voting rights. *Journal of Financial Economics*, 20:25–59, 1988.

[218] R. Thaler. *Advances in Behavioral Finance II*. Princeton University Press, Princeton, NJ, 2005.

[219] J. D. Thomas. *News and Trading Rules*. PhD thesis, Carnegie Mellon University, 2003.

[220] R. Thompson, E. J. Brinkmann, P. Dharmapala, M. Gonzalez-Lima, and R. M. Thrall. DEA/AR profit ratios and sensitivity of 100 large U.S. banks. *European Journal of Operational Research*, 105(3):213–229, 1998.

[221] N. Towers and A. N. Burgess. Implementing trading strategies for forecasting models. In *Y. Abu-Mostafa et al. (eds.) Computational finance. Proceedings of the Sixth International Conference on Computational Finance*, pages 313–325. The MIT Press, 1999.

[222] R. Trippi and J. Lee, editors. *Foundations of Investment System Using Artificial Intelligence and Web*. McGraw-Hill, New York, 2000.

[223] R. Trippi and E. Turban, editors. *Investment management: Decision support and expert systems*. Van Nostrand Reinhold, New York, 1990.

[224] R. Trippi and E. Turban, editors. *Neural Networks in Finance and Investing*. Probus publishing, Chicago, 1993.

[225] T. Tsay. *Analysis of Financial Time Series*. Wiley, New York, 2002.

[226] S. Voelpel, M. Leibold, R. Eckhoff, and T. Davenport. The tyranny of the Balanced Scorecard in the innovation economy. In *4th. International Critical Management Studies Conference*, 2005.

[227] P. Volpin. Governance with poor investor protection: Evidence from top executive turnover in Italy. *Journal of Financial Economics*, 64:61–90, 2002.

[228] V. Vovk and C. Watkins. Universal portfolio selection. In *Proceedings of the 11th Annual Conference on Computational Learning Theory (COLT-98)*, pages 12–23, New York, jul 1998. ACM Press.

[229] D. Watts. Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, 105:493–527, 1999.

[230] D. Watts and S. Strogatz. Collective dynamics of small world networks. *Nature*, 393:440–2, 1998.

[231] M. Weisbach. Outside directors and CEO turnover. *Journal of Financial Economics*, 20:431–460, 1988.

[232] M. P. Wellman, A. Greenwald, P. Stone, and P. R. Wurman. The 2001 trading agent competition. *Electronic Markets*, 13(1), 2002.

[233] F. Weston. The tender takeover. *Mergers and Acquisitions*, pages 74–82, 1979.

[234] W. Wilder. *New Concepts in Technical Trading Systems*. Trend Research, 1978.

[235] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, 2nd. edition, 2005.

[236] K. Womack. Do brokerage analysts recommendations have investment value? *Journal of Finance*, 51:137–167, 1996.

[237] B. Wuthrich, D. Permunetilleke, S. Leung, V. Cho, J. Zhang, and W. Lam. Daily prediction of major stock indices from textual www data. In *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining*, pages 364–368, 1998.

[238] D. Yermack. Higher market valuation of companies with a small board of directors. *Journal of Financial Economics*, 40:185–211, 1996.

[239] A. Youngblood and T. Collins. Addressing Balanced Scorecard trade-off issues between performance metrics using multi-attribute utility theory. *Engineering Management Journal*, 15(1):11–17, 2003.

[240] O. Zaire. The effect of financial liberalization on the efficiency of Turkish commercial banks. *Applied Financial Economics*, 5, 1995.

[241] C. Zavgren. The prediction of corporate failure: the state of the art. *Journal of Accounting Literature*, 2, 1983.

[242] E. Zivot and J. Wang. *Modeling Financial Time Series with S-Plus*. Springer, New York, 2003.

Appendices

A   Investment signals for earnings surprise prediction

The fundamental variables are calculated using the information of the previous quarter (SUE,SG,TA,and CAPEX) and our notation is similar to the notation used by Jegadeesh et al. [129]. We do not include firm-specific subscripts in order to clarify the presentation. Subscript q refers to the most recent quarter for which an earnings announcement was made.

Variable	Description	Calculation detail
Company description:		
SECTOR	Two-digit sector classification according to the Global Industrial Classification Standards (GICS) code.	Energy 10, Materials 15, Industrials 20, Consumer Discretionary 25, Consumer Staples 30, Health Care 35, Financials 40, Information Technology 45 Telecommunication Services 50, Utilities 55
Price momentum:		
CAR1	Cumulative abnormal return for the preceding six months since the earnings announcement day	$[\Pi_{t=m-6}^{m-1}(1+R_t)-1]-[_{t=m-6}^{m-1}(1+R_{tw})-1]$, where R_t is return in month t, R_{tw} is value weighted market return in month t, and m is last month of quarter
CAR2	Cumulative abnormal return for the second preceding six months since the earnings announcement day	$[\Pi_{t=m-12}^{m-7}(1+R_t)-1]-[\Pi_{t=m-6}^{m-1}(1+R_{tw})-1]$
Analysts variables:		
ANFOR (ANFOR-LAG)	Number of analysts predicting that earnings surprise increase (lagged value)	
CONSENSUS	Mean of earnings estimate by financial analysts	
FELAG	Lagged forecast error	$\frac{\text{CONSENSUS}_q - \text{EPS}_q}{|\text{CONSENSUS}_q| + |\text{EPS}_q|}$ [86]
Earnings momentum:		
FREV	Analysts earnings forecast revisions to price	$\sum_{i=0}^{5} \frac{\text{CONSENSUS}_{m-i} - \text{CONSENSUS}_{m-i-1}}{P_{m-i-1}}$ where P_{m-1} is price at end of month $m-1$, and i refers to the previous earnings revisions
SUE	Standardized unexpected earnings	$\frac{(EPS_q - EPS_{q-4})}{\sigma_t}$ where EPS is earnings per share, and σ_t is standard deviation of EPS for previous seven quarters
Growth indicators:		
LTG	Mean of analysts' long-term growth forecast	
SG	Sales growth	$\frac{\sum_{t=0}^{3} \text{Sales}_{q-t}}{\sum_{t=0}^{3} \text{Sales}_{q-4-t}}$
Firm size:		
SIZE	Market cap (natural log)	$ln(P_q\ shares_q)$ where $shares_q$ are outstanding shares at end of quarter q
Fundamentals:		
TA	Total accruals to total assets	$\frac{\triangle\text{C.As.}_q - \triangle Cash_q - (\triangle\text{C.Lb.}_q - \triangle\text{C.Lb.D}_q) - \triangle\text{T}_q - \text{D\&A}_q}{\frac{(\text{T.As.}_q - \text{T.As.}_{q-4})}{2}}$

where $\triangle X_q = X_q - X_{q-1}$ and C.As., C.Lb.,C.Lb.D.,T,D&A,and T.As. stands for current assets, current liabilities, debt in current liabilities, deferred taxes, depreciation and amortization, and total assets respectively.

CAPEX	Rolling sum of capital expenditures to total assets	$\frac{\sum_{t=0}^{3} \text{capital expenditures}_{q-t}}{(T.As._q - T.As._{q-4})/2}$
Valuation multiples:		
BP	Book to price ratio	$\frac{\text{book value of common equity}_q}{\text{market cap}_q}$, where market cap$_q = P_q$ shares$_q$
EP	Earnings to price ratio (rolling sum of EPS of the previous four quarters deflated by prices)	$\frac{\sum_{t=0}^{3} EPS_{q-t}}{P_q}$
Social networks:		
$deg(v_i)$	Degree centrality or degree: number of edges incidents in vertex v_i	$\sum_j a_{ij}$, where a_{ij} is an element of the adjacent matrix A
$C_c(v_i)$	Closeness centrality (normalized): inverse of the average geodesic distance from vertex v_i to all other vertices	$\frac{n-1}{\sum_j d_{ij}}$, where d_{ij} is an element of the geodesic distance matrix D [102, 39]
$B_c(v_i)$	Betweenness centrality: proportion of all geodesic distances of all other vertices that include vertex v_i	$\sum_i \sum_j \frac{g_{kij}}{g_{kj}}$, where g_{kij} is the number of geodesic paths between vertices k and j that include vertex i, and g_{kj} is the number of geodesic paths between k and j [102]
CC_i	Clustering coefficient: cliquishness of a particular neighborhood or the proportion of edges between vertices in the neighborhood of v_i divided by the number of edges that could exist between them [230]	$\frac{2|\{e_{ij}\}|}{deg(v_i)(deg(v_i)-1)} : v_j \in N_i,\ e_{ij} \in E$, where each vertex v_i has a neighborhood N defined by its immediately connected neighbors: $N_i = \{v_j\} : e_{ij} \in E$.
CC_i'	Normalized clustering coefficient	$\frac{deg(v_i)}{MaxDeg} CC_i$, where MaxDeg is the maximum degree of vertex in a network [77]
C (not used for forecasting)	Mean of all the clustering coefficients	$\frac{1}{n} \sum_{i=1}^{n} CC_i$
SW (not used for forecasting)	"Small world" ratio [230].	$\frac{C}{L} \frac{L_{random}}{C_{random}}$, where $L_{random} \approx \frac{ln(n)}{ln(k)}$ and $C_{random} \approx \frac{k}{n}$
Labels:		
LABELFE	Label of forecast error (FE)	1 if $CONSENSUS \geq EPS$ (current quarter) , -1 otherwise
LABELCAR	Label of cumulative abnormal return (CAR)	1 if $CAR_{m+1} \geq 0$, -1 otherwise, where CAR_{m+1} refers to the CAR of the month that follows the earnings announcement

B Technical analysis indicators used during PLAT competition

Technical indicators are statistics of the market that quantify market trends. Most technical indicators have been developed by professional traders using trial and error. It is common practice to use rules based on technical indicators to choose the timing of buy and sell orders. These rules are called buy and sell "signals". In this work we use a combination of market indicators and trading signals. We define these indicators in this appendix and provide the basic intuition that motivates them. Throughout this section we assume a single fixed stock.

We start with some basic mathematical notation. We index the trading days by $t = 1, 2, \ldots$. We denote by $P_t^o, P_t^c, P_t^{uc}, P_t^h$, and P_t^l, the open, adjusted close, unadjusted close[1], high, and low price of the tth trading day. We eliminate the lower index when we wish to refer to the whole sequence, i.e. P^c refers to the whole sequence $P_1^c, P_2^c, \ldots$. Using this notation we define the median price $P^{med} = (P^h + P^l)/2$, the typical or average price $P^{typ} = (P^h + P^l + P^{uc})/3$, and the weighted close price $P^{wc} = (P^h + P^l + 2P^{uc})/4$.

Many of the technical indicators incorporate time averages of prices or of other indicators. We use two types of time averages, the simple moving average and the exponentially weighted moving average.[2] Let $\mathbf{X}$ denote a time sequence $X_1, X_2, \ldots$. The **simple moving average** is defined as

$$\mathbf{SMA}_t(\mathbf{X}, n) = \frac{1}{n} \sum_{s=0}^{n-1} X_{t-s} \ ,$$

---

[1]Unadjusted close prices are the actual published prices at the end of the trading day. The adjusted stock price removes the effect of stock splits and dividend payments. Our goal is to predict P_t^c, the *adjusted* close price.

[2]We follow Zivot and Wang [242] in describing the technical analysis indicators. Additional useful references about technical analysis and trading are [143, 196, 59, 208, 216, 87, 66].

and the **exponentially weighted moving average** is defined as

$$\mathbf{EMA}_t(\mathbf{X}, n) = \lambda \sum_{s=0}^{\infty} (1 - \lambda)^s X_{t-s}; \quad \lambda = \frac{2}{n+1} \ .$$

A useful property of $\mathbf{EMA}_t(\mathbf{X}, n)$ is that it can be calculated using a simple update rule:

$$\mathbf{EMA}_t(\mathbf{X}, n) = \lambda X_t + (1 - \lambda)\mathbf{EMA}_{t-1}(\mathbf{X}, n) \ .$$

In the following table we describe the technical indicators. The parameters of each indicator are in parentheses. Most of the parameters used refer to the length of the period (n) selected to calculate the indicator. In case of exponential moving average, the parameter used is λ which also depends of n. We have assigned parameters which are typically used in the industry for each indicator.

Variable	Description	Calculation detail [Source]
Price indicators:		
$SMA_t^c(n)$	Simple moving average of the last n observations of a time series P^c.	$\mathbf{SMA}_t(P^c, n)$ *where $n = 3$, and 6*
Bollinger bands:	Using the moving average or the median band $(Boll_t^m(n))$ as the reference point, the upper and lower Bollinger [38] bands $(Boll_t^u(n)$ and $Boll_t^d(n)$ respectively) are calculated in function of s standard deviations. When price crosses above (below) the upper (lower) Bollinger band, it is a sign that the market is overbought (oversold). Technical analysts typically calculate Bollinger bands using 20 days for the moving average and 2 standard deviations.	$Boll_t^m(n) = SMA_t^c(n)$ where n=6
$Boll_t^u(n)$	Upper Bollinger band	$Boll_t^m(n) + s \cdot \sigma_t^2(n)$ where s=2.6 [Katz [143]]
$Boll_t^d(n)$	Lower Bollinger band	$Boll_t^m(n) - s \cdot \sigma_t^2(n)$ where s=2.6 [Katz [143]]

| $ADX_t(n)$ | Average directional movement index: indicates if there is a trend and the overall strength of the market [234]. Range of values from 0 to 100. A high number is a strong trend, and a low number is a weak trend. The directional movement index (DX_t) is the percentage of the true range $(TRange_n)$ that is up $(+DI_t(n))$ or down $(-DI_t(n))$. The true range determines the trading range of an asset. | $ADX_{t-1}(n) \cdot (n-1) + DX_t)/n$
 where:
 $DX_t \doteq \frac{(+DI_t(n))-(-DI_t(n))}{(+DI_t(n))+(-DI_t(n))}$
 $TRange_n = max(\mathbf{P_n^h}) - min(\mathbf{P_n^l})$
 $n = 5$
 $\mathbf{P_n^h} = (P_{t-n}^h, P_{t-n+1}^h, P_{t-n+2}^h, \dots, P_t^h)$
 $\mathbf{P_n^l} = (P_{t-n}^l, P_{t-n+1}^l, P_{t-n+2}^l, \dots, P_t^l)$ |

Momentum and oscillation indicators:

| $MACD_t(s,f)$ | Moving average convergence divergence: difference between two moving averages of different periods (s,f) where s stands for a slow period and f for a fast period. $MACD_t(s,f)$ is regularly calculated using 26 (s) and 12 (f) periods. | $\mathbf{EMA}_t(P^c, s) - \mathbf{EMA}_t(P^c, f)$
 where s=26, and f=12. |

| $MACDS_t(s,f,n)$ | MACD signal line: moving average of $MACD_t(s,f)$ of past n periods. A buy (sell) signal is generated when the $MACD_t(s,f)$ crosses above (below) the signal line or a threshold. | $\mathbf{EMA}_t(MACD_t(s,f), n)$
 where f=12, n=9, and s=26. |

| $MACDH_t(n,l)$ | MACD histogram: difference between the fast MACD line and the MACD signal line. | $\mathbf{EMA}_t(c, l, \lambda) - MACDS_t(n)$
 $where \ f = 26$ |

| $RSI_t(n)$ | Relative strength index: compares the days that stock prices finish up against those periods that stock prices finish down. Technical analysts calculate this indicator using 9, 14 or 25 periods. A buy signal is when $RSI_t(n)$ crosses below a lower band of 30 (oversold) and a sell signal when $RSI_t(n)$ crosses above an upper band of 70 (overbought). | $100 - \dfrac{100}{1 + \dfrac{\mathbf{SMA}_t(\mathbf{P_n^{up}}, n)}{\mathbf{SMA}_t(\mathbf{P_n^{dn}}, n)}}$
 where n = 5, and n is the length of the time series
 $P_t^{up} = \begin{cases} P_t^c & \text{if } P_t^c > P_{t-1}^c \\ \text{empty} & Otherwise \end{cases}$
 $P_t^{dn} = \begin{cases} P_t^c & \text{if } P_t^c < P_{t-1}^c \\ \text{empty} & Otherwise \end{cases}$
 $\mathbf{P_n^{up}} = (P_{t-n}^{up}, P_{t-n+1}^{up}, P_{t-n+2}^{up}, \dots, P_t^{up})$
 $\mathbf{P_n^{dn}} = (P_{t-n}^{dn}, P_{t-n+1}^{dn}, P_{t-n+2}^{dn}, \dots, P_t^{dn})$ |

| Stochastic oscillator: | Compares close price to a price range in a given period to establish if market is moving to higher or lower levels or is just in the middle. The oscillator indicators are: | |

$FAST\%K_t(n)$	Percent measure of the last close price in relation to the highest high and lowest low of the last n periods (true range). Typically a period (n) of 5 is used for $FAST\%K_t(n)$ and 3 for the rest of stochastic indicators. We follow this convention.	$\frac{P_t^{\mathrm{uc}}-min(\mathbf{P_n^l})}{max(\mathbf{P_n^h})-min(\mathbf{P_n^l})}$
	Vector with low prices of last n periods	$\mathbf{P_n^l} = (P_{t-n}^l,\ P_{t-n+1}^l,\ P_{t-n+2}^l,\ \ldots,\ P_t^l)$
	Vector with high prices of last n periods	$\mathbf{P_n^h} = (P_{t-n}^h,\ P_{t-n+1}^h,\ P_{t-n+2}^h,\ \ldots,\ P_t^h)$
$FAST\%D_t(n)$	Moving average of $FAST\%K_t(n)$.	$\mathbf{SMA}_t(FAST\%K_t(n), 3)$
$SLOW\%K_t(n)$	Identically calculated to $FAST\%D_t(n)$ using a 3-period moving average of $FAST\%K_t(n)$.	$\mathbf{SMA}_t(FAST\%K_t(n), 3)$
$SLOW\%D_t(n)$	Moving average of $SLOW\%K_t(n)$. Typically a period of 3 is used. A buy (sell) signal is generated when any oscillator (either $\%K$ or $\%D$) crosses below (above) a threshold and then crosses above (below) the same threshold. Typically a threshold of 80 is used for the above threshold, and 20 for the below threshold. Buy and sell signal are also generated when $FAST\%K_t(n)$ or $SLOW\%K_t(n)$ crosses above or below $FAST\%D_t(n)$ or $SLOW\%D_t(n)$ respectively.	$\mathbf{SMA}_t(SLOW\%K_t(n), 3)$
$MFI_t(n)$	Money flow index: measures the strength of money flow (MF_t) in and out of a stock. At difference of the $RSI_t(n)$ which is calculated using stock prices, $MFI_t(n)$ is calculated using volume. When $MFI_t(n)$ crosses above (below) 70 (30), this is a sign that the market is overbought (oversold).	$100 - \dfrac{100}{1 + \dfrac{PMF_t(n)}{NMF_t(n)}}$ $where\ n = 15$ $MF_t = P_t^{\mathrm{typ}} \cdot VOL_t$ $PMF_t(n) = \mathbf{SMA}_t(MF_t, n)\text{when } MF_t > 0$ $NMF_t(n) = \mathbf{SMA}_t(MF_t, n) \text{ when } MF_t < 0$ VOL_t is volume of day t $PMF_t(n)$ is positive money flow $NMF_t(n)$ is negative money flow

C   Investment signals for automated trading system

This appendix lists the indicators and the rules or signals that we have used for our predictions in the model presented in chapter 6. We use the same conventions and most of the indicators included in appendix B. We name as "rule" and follow by an identification number the most common rules associated with each indicator. Most of the buy and sell signals are generated when the value of an indicator crosses some threshold or the value of another indicator. The input to our learning system includes both signals and indicators. We use *normalized* indicators, by which we mean indicators whose value does not change if all the prices in the sequence are multiplied by a constant factor. This is important when working with adjusted stock prices.

Additionally, we recalculate a selected group of indicators and their rules with three different values of the main parameters that are close to the industry practice. So, our learning system should be able to select the optimal combination of indicators and parameters. Additionally, we include ratios of the indicators which generally are calculated as the indicator divided by its moving average. Most of these ratios are part of the trading rules. However, we include the ratios by themselves so that our learning system finds its own rules. Finally, there are indicators that do not have a specific trading rule such as the volatility and return indicators. We include several measures of volatility, so that our model is able to discover its own rules of risk management.

Variable	Description	Calculation detail
Price indicators:		
$EMA_t^c(\lambda)$	Exponential moving average of a time series P^c.	$\mathbf{EMA}_t(P^c, \lambda)$ $where\ \lambda = 0.9,\ 0.84,\ and\ 0.78$
$rule1_t - rule4_t$	Exponential moving average to price $(P_t^c, P_t^{med}, P_t^{typ}, and\ P_t^{wc})$.	$\frac{EMA_t^c(\lambda)}{P^c}, \frac{EMA_t^{med}(\lambda)}{P^{med}}, \frac{EMA_t^{typ}(\lambda)}{P^{typ}}, and\ \frac{EMA_t^{wc}(\lambda)}{P^{wc}}$

where $n=$ 0.9, 0.84, and 0.78

$SMA_t^c(n)$	Simple moving average of the last n observations of a time series P^c.	$\mathbf{SMA}_t(P^c, n)$ *where* n= 10, 16 *and* 22
$rule5_t$	Simple moving average to P_t^c	$\frac{SMA_t^c(n)}{P^c}$ where n=10, 16 and 22
Bollinger bands:	Using the moving average or the median band $(Boll_t^m(n))$ as the reference point, the upper and lower Bollinger [38] bands $(Boll_t^u(n)$ and $Boll_t^d(n)$ respectively) are calculated in function of s standard deviations. When price crosses above (below) the upper (lower) Bollinger band, it is a sign that the market is overbought (oversold). Technical analysts typically calculate Bollinger bands using 20 days for the moving average and 2 standard deviations.	$Boll_t^m(n) = SMA_t^c(n)$
$Boll_t^u(n)$	Upper Bollinger band	$Boll_t^m(n) + s\sigma_t^2(n)$ where s=2, n=20, 26 and 32
$Boll_t^d(n)$	Lower Bollinger band	$Boll_t^m(n) - s\sigma_t^2(n)$ where s=2, n=20, 26 and 32
$PBoll_t^u(n)$	Price to upper Bollinger band	$\frac{P_t^c}{Boll_t^u(n)}$
$PBoll_t^d(n)$	Price to lower Bollinger band	$\frac{P_t^c}{Boll_t^d(n)}$
$rule6_t$	Bollinger trading rule	$\begin{cases} Buy & \text{if } P_{t-1}^c \geq Boll_t^d(n) \text{ and } P_t^c < Boll_t^u(n) \\ Sell & \text{if } P_{t-1}^c \leq Boll_t^d(n) \text{ and } P_t^c > Boll_t^u(n) \\ Hold & \text{Otherwise} \end{cases}$
Momentum and oscillation indicators:		
$MOM_t(n)$	Momentum: price (P_t^c) change in the last n periods. When it crosses above (below) zero, it indicates that trend is up (down). The default value of n is 12.	$P_t^c - P_{t-n}^c$ *where* $n = 12, 18,$ *and* 24
$MomEMA_t(n, \lambda)$	Momentum to $\mathbf{EMA}_t(MOM_t(n), \lambda)$	$\frac{MOM_t(n)}{\mathbf{EMA}_t(MOM_t(n)\lambda)}$ where n=12, 18, and 24 and $\lambda = 0.75$
$rule7_t$	Momentum trading rule	$\begin{cases} Buy & \text{if } MOM_{t-1}(n) \leq \mathbf{EMA}_t(MOM_t(n), \lambda) \\ & \text{and } MOM_t(n) > \mathbf{EMA}_t(MOM_t(n), \lambda) \\ Sell & \text{if } MOM_{t-1}(n) \geq \mathbf{EMA}_t(MOM_t(n), \lambda) \\ & \text{and } MOM_t(n) < \mathbf{EMA}_t(MOM_t(n), \lambda) \\ Hold & \text{Otherwise} \end{cases}$

$ACCEL_t(n)$	Acceleration: difference of price change. The default value of n is 12.	$MOM_t(n) - MOM_{t-1}(n)$ $where\ n = 12,\ 18,\ and\ 24$
$rule8_t$	Acceleration trading rule	$\begin{cases} Buy & \text{if } ACCEL_{t-1}(n) + 1 \leq 0 \\ & \text{and } ACCEL_t(n) + 1 > 0 \\ Sell & \text{if } ACCEL_{t-1}(n) + 1 \geq 0 \\ & \text{and } ACCEL_t(n) + 1 < 0 \\ Hold & \text{Otherwise} \end{cases}$
$ROC_t(n)$	Rate of change: rate of change of $P_t^{\mathbf{c}}$. Technical analysts recommend using 10 periods to calculate this indicator.	$\frac{P_t^{\mathbf{c}} - P_{t-n}^{\mathbf{c}}}{P_{t-n}^{\mathbf{c}}} \cdot 100$ $where\ n = 10,\ 16,\ and\ 22$
$rule9_t$	ROC trading rule	$\begin{cases} Buy & \text{if } ROC_{t-1}(n) \leq 0 \text{ and } ROC_t(n) > 0 \\ Sell & \text{if } ROC_{t-1}(n) \geq 0 \text{ and } ROC_t(n) < 0 \\ Hold & \text{Otherwise} \end{cases}$
$MACD_t(s, f)$	Moving average convergence divergence: difference between two moving averages of slow and fast periods (s, f). $MACD_t(s, f)$ is regularly calculated using 26 (s) and 12 (f) periods.	$\mathbf{EMA}_t(P^{\mathbf{c}}, s) - \mathbf{EMA}_t(P^{\mathbf{c}}, f)$ $where\ f = 12,\ and\ s = 18,\ 24,\ and\ 30$
$MACDS_t(s, f, n)$	MACD signal line: moving average of $MACD_t(s, f)$ of past n periods. A buy (sell) signal is generated when the $MACD_t(s, f)$ crosses above (below) the signal line or a threshold.	$\mathbf{EMA}_t(MACD_t(s, f), n)$ $where\ f = 12,\ n = 9,\ and\ s = 18,\ 24,\ and\ 30$
$rule10_t$	MACD trading rule	$\begin{cases} Buy & \text{if } MACD_{t-1}(s, f) \leq MACDS_t(s, f, n) \\ & \text{and } MACD_t(s, f) > MACDS_t(s, f, n) \\ Sell & \text{if } MACD_{t-1}(s, f) \geq MACDS_t(s, f, n) \\ & \text{and } MACD_t(s, f) < MACDS_t(s, f, n) \\ Hold & \text{Otherwise} \end{cases}$
$MACDR_t(s, f, n)$	$MACD_t(s, f)$ to $MACDS_t(s, f, n)$	$\frac{MACD_t(s, f)}{MACDS_t(s, f, n)}$
$RSI_t(n)$	Relative strength index: compares the days that stock prices finish up against those periods that stock prices finish down. Technical analysts calculate this indicator using 9, 14 or 25 periods. A buy signal is when $RSI_t(n)$ crosses below a lower band of 30 (oversold), and a sell signal when $RSI_t(n)$ crosses above an upper band of 70 (overbought)	$100 - \dfrac{100}{1 + \dfrac{\mathbf{SMA}_t(\mathbf{P_n^{up}}, n_1)}{\mathbf{SMA}_t(\mathbf{P_n^{dn}}, n_1)}}$ $where\ n_1 = 8,\ 14,\ and\ 20$ and n is the length of the time series $P_t^{up} = \begin{cases} P_t^{\mathbf{c}} & \text{if } P_t^{\mathbf{c}} > P_{t-1}^{\mathbf{c}} \\ empty & Otherwise \end{cases}$ $P_t^{dn} = \begin{cases} P_t^{\mathbf{c}} & \text{if } P_t^{\mathbf{c}} < P_{t-1}^{\mathbf{c}} \\ empty & Otherwise \end{cases}$ $\mathbf{P_n^{up}} = (P_{t-n}^{up},\ P_{t-n+1}^{up},\ P_{t-n+2}^{up},\ \cdots,\ P_t^{up})$ $\mathbf{P_n^{dn}} = (P_{t-n}^{dn},\ P_{t-n+1}^{dn},\ P_{t-n+2}^{dn},\ \cdots,\ P_t^{dn})$
$rule11_t$	RSI trading rule	$\begin{cases} Buy & \text{if } RSI_{t-1}(n) \geq 30 \text{ and } RSI_t(n) < 70 \\ Sell & \text{if } RSI_{t-1}(n) \leq 30 \text{ and } RSI_t(n) > 70 \\ Hold & \text{Otherwise} \end{cases}$

Stochastic oscillator:	Compares close price to a price range in a given period to establish if market is moving to higher or lower levels or is just in the middle. The oscillator indicators are:	
$FAST\%K_t(n)$	Percent measure of the last close price in relation to the highest high and lowest low of the last n periods (true range).	$\frac{P_t^{\mathrm{uc}} - min(\mathbf{P_n^l})}{max(\mathbf{P_n^h}) - min(\mathbf{P_n^l})}$ $where\ n = 12,\ 18,\ and\ 24$
	Vector with low prices of last n periods	$\mathbf{P_n^l} = (P_{t-n}^l,\ P_{t-n+1}^l,\ P_{t-n+2}^l,\ \ldots,\ P_t^l)$
	Vector with high prices of last n periods	$\mathbf{P_n^h} = (P_{t-n}^h,\ P_{t-n+1}^h,\ P_{t-n+2}^h,\ \ldots,\ P_t^h)$
$FAST\%D_t(n)$	Moving average of $FAST\%K_t(n)$.	$\mathbf{SMA}_t(FAST\%K_t(n), 3)$
$SLOW\%K_t(n)$	Identically calculated to $FAST\%D_t(n)$ using a 3-period moving average of $FAST\%K_t(n)$.	$\mathbf{SMA}_t(FAST\%K_t(n), 3)$
$SLOW\%D_t(n)$	Moving average of $SLOW\%K_t(n)$. Typically a period of 3 is used.	$\mathbf{SMA}_t(SLOW\%K_t(n), 3)$
$rule12_t$	Fast stochastic trading rule	$\begin{cases} Buy & \text{if } FAST\%K_{t-1}(n) \leq FAST\%D_t(n) \\ & \text{and } FAST\%K_t(n) > FAST\%D_t(n) \\ Sell & \text{if } FAST\%K_{t-1}(n) \geq FAST\%D_t(n) \\ & \text{and } FAST\%K_t(n) < FAST\%D_t(n) \\ Hold & \text{Otherwise} \end{cases}$
$rule13_t$	Slow stochastic trading rule	$\begin{cases} Buy & \text{if } SLOW\%K_{t-1}(n) \leq SLOW\%D_t(3) \\ & \text{and } SLOW\%K_t(n) > SLOW\%D_t(3) \\ Sell & \text{if } SLOW\%K_{t-1}(n) \geq SLOW\%D_t(3) \\ & \text{and } SLOW\%K_t(n) < SLOW\%D_t(3) \\ Hold & \text{Otherwise} \end{cases}$
$slowKslowD_t(n)$	$SLOW\%K_t(n)$ to $SLOW\%D_t(3)$	$\frac{SLOW\%K_t(n)}{SLOW\%D_t(3)}$
$fastKfastD_t(n)$	$FAST\%K_t(n)$ to $FAST\%D_t(n)$	$\frac{FAST\%K_t(n)}{FAST\%D_t(n)}$
$WILL_t(n)$	Williams indicator: the calculation is similar to the stochastic oscillator with a scale from 0 to -100. It tries to capture moments when the market is overbought (0 - -20) or oversold (-80 - -100).	$\frac{max(\mathbf{P_n^h}) - P_t^{\mathrm{uc}}}{max(\mathbf{P_n^h}) - min(\mathbf{P_n^l})}(-100)$ $where\ n = 14$
$rule14_t$	Williams trading rule	$\begin{cases} Buy & \text{if } WILL_{t-1}(n) \geq -20 \text{ and } WILL_t(n) < -80 \\ Sell & \text{if } WILL_{t-1}(n) \leq -20 \text{ and } WILL_t(n) > -80 \\ Hold & \text{Otherwise} \end{cases}$

$MFI_t(n)$	Money flow index: measures the strength of money flow (MF_t) in and out of a stock. At difference of the $RSI_t(n)$ which is calculated using stock prices, $MFI_t(n)$ is calculated using volume.	$100 - \dfrac{100}{1 + \dfrac{PMF_t(n)}{NMF_t(n)}}$ where: $n = 14$ $MF_t = P_t^{\text{typ}} \cdot VOL_t$ $PMF_t(n) = \mathbf{SMA}_t(MF_t, n)$ when $MF_t > 0$ $NMF_t(n) = \mathbf{SMA}_t(MF_t, n)$ when $MF_t < 0$ VOL_t is volume of day t $PMF_t(n)$ is positive money flow $NMF_t(n)$ is negative money flow
$rule15_t$	Money flow index trading rule. When $MFI_t(n)$ crosses above (below) 70 (30), this is a sign that the market is overbought (oversold).	$\begin{cases} Buy & \text{if } MFI_{t-1}(n) \geq 30 \text{ and } MFI_t(n) < 70 \\ Sell & \text{if } MFI_{t-1}(n) \leq 30 \text{ and } MFI_t(n) > 70 \\ Hold & \text{Otherwise} \end{cases}$
Volatility and return indicators: [3] $CHV_t(n, n_1)$	Chaikin volatility: evaluates the widening of the range between high and low prices. This indicator also calculates the rate of change of the moving average of the difference between high and low prices. Chaikin [55] suggests using 10 periods (n_1) to calculate this indicator. Chaikin also considers that a very fast increase (decrease) of the Chaikin volatility is a signal that the bottom (top) of the market is near.	$\mathbf{EMA}_t(P^{\text{h}} - P^{\text{l}}, n)/\mathbf{EMA}_{t-n_1}(P^{\text{h}} - P^{\text{l}}, n) - 1$ $where n = n_1 = 10$
GK_t	Garman-Klass volatility: this is an extreme-value indicator proposed by Garman and Klass [108] that takes into account intraday price range to calculate the variation of the stock price. According to Garman and Klass variance is minimized when $\alpha = 0.12$. f is the fraction of the day that trading is closed. We use a value suggested by Rmetrics (0.19) which is about 4.5 hours.	$\alpha \dfrac{(P_t^{\text{o}} - P_{t-1}^{\text{uc}})}{f} + (1 - \alpha)\dfrac{\hat{\sigma_t^2}(n)}{1-f}$ $where\ f < 1$ $u = (P_t^{\text{h}} - P_t^{\text{o}})$ $d = (P_t^{\text{l}} - P_t^{\text{o}})$ $\hat{\sigma_t^2}(n) = 0.511(u - d)^2 - 0.019P_t^{\text{c}} \cdot (u + d) - 2ud - 0.383(P_t^{\text{uc}})^2$
$\hat{r}_{t+1}, \sigma_t^2$	Next period return and volatility calculated using the $GARCH(1,1)$ model (see section C.1)	
Sharpe ratio	Risk adjusted return (see section C.1).	$\hat{r}_t/\sigma_t^2$

Volume indicators:

OBV_t	On balance volume: this indicator was developed by Granville [115] to evaluate the impact of positive and negative volume flows. OBV_t adds the volume when the close price has increased and substracts it when the close price has decreased. A sign of market reversal is when OBV_t diverges with the price movement.	if $P_t^{\text{c}} > P_{t-1}^{\text{c}}$ $OBV_t = OBV_{t-1} + VOL_t$ if $P_t^{\text{c}} < P_{t-1}^{\text{c}}$ $OBV_t = OBV_{t-1} - VOL_t$
ADL_t	Accumulation/distribution line: this indicator was developed by Chaikin [55] to evaluate the effect of accumulative flow of money of a particular security. ADL_t is calculated using the close location value (CLV_t). This indicator compares the unadjusted close price with the range of prices for the same period without comparing with the previous period as the OBV_t does. ADL_t range is from -1 to +1, and zero is the central point. A positive value indicates buying pressure and a negative value indicates selling pressure. If there is an important positive (negative) divergence between the accumulation distribution line and the price,we have a bullish (bearish) signal.	$\sum_{t=1}^{n} CLV_t \cdot VOL_t$ $where\ CLV_t = \frac{2P_t^{\text{uc}} - P_t^{\text{l}} - P_t^{\text{h}}}{P_t^{\text{h}} - P_t^{\text{l}}}$ and n refers to the length of the time series.
CHO_t	Chaikin oscillator: this indicator was also developed by Chaikin [55]. The Chaikin oscillator is the $MACD$ of the ADL. This oscillator is the difference between a short and a long $\mathbf{EMA}_t(ADL, n)$ of the ADL. The interpretation of this indicator is similar to the $MACD$.	$\mathbf{EMA}_t(ADL, n_1) - \mathbf{EMA}_t(ADL, n_2)$ $where\ n_1 = 3,\ n_2 = 10$
$rule16_t$	Chaikin volatility trading rule	$\text{sign}\,(CHO_t)$

NVI_t and PVI_t	Negative and positive volume index: these indicators were introduced by Fosback [101] as signals of bull markets. NVI_t (PVI_t) concentrates on days when volume decreases (increases). The rationality is that "informed" investors take positions on days when volume decreases, while the "uninformed" investors take position on days when the volume increases. NVI_t (PVI_t) is calculated as the cumulative sum of $ROC_t(n)$ when volume decreases (increases). Fosback maintains that there is 95% probability that a bull market is going to develop when NVI_t crosses above its one year moving average, and 67% probability of a bear market when PVI_t crosses below its one year moving average.	if $VOL_t < VOL_{t-1}$ $NVI_t = NVI_{t-1} + ROC_t(n)NVI_{t-1}$ $PVI_t = PVI_{t-1}$ if $VOL_t \geq VOL_{t-1}$ $PVI_t = PVI_{t-1} + ROC_t(n)PVI_{t-1}$ $NVI_t = NVI_{t-1}$ $where$ n=1
$rule17_t$	Negative volume index trading rule.	$\begin{cases} Buy & \text{if } NVI_{t-1} \leq \mathbf{SMA}_t(NVI,l) \\ & \text{and } NVI_t > \mathbf{SMA}_t(NVI,l) \\ Hold & \text{Otherwise} \end{cases}$
$rule18_t$	Positive volume index trading rule.	$\begin{cases} Buy & \text{if } PVI_{t-1} \leq \mathbf{SMA}_t(PVI,l) \\ & \text{and } PVI_t > \mathbf{SMA}_t(PVI,l) \\ Sell & \text{if } PVI_{t-1} \geq \mathbf{SMA}_t(PVI,l) \\ & \text{and } PVI_t < \mathbf{SMA}_t(PVI,l) \\ Hold & \text{Otherwise} \end{cases}$ where $l = 10$
$nviSMA_t(l)$	NVI_t to $\mathbf{SMA}_t(NVI,l)$	$\dfrac{NVI_t}{\mathbf{SMA}_t(NVI,l)}$, where $l = 10$
$pviSMA_t(l)$	PVI_t to $\mathbf{SMA}_t(PVI,l)$	$\dfrac{PVI_t}{\mathbf{SMA}_t(PVI,l)}$, where $l = 10$
$PV_t(n)$	Price-volume trend: this indicator is similar to OBV_t. It calculates a cumulative total of volume where the portion of volume added/substracted is given by the increase or decrease of close prices in relation to the previous period.	$\sum_{t=1}^{n} VOL_t \cdot ROC_t(n_1)$ $where\ n_1 = 1$ and n is the length of the time series.

---

[3]GARCH is also another indicator of volatility and it is explained in detail in appendix C.1

C.1 GARCH and Value at Risk

ARCH and GARCH models have been widely used in finance literature to forecast volatility and assets' return, especially since the volatility of assets return seems to be serially correlated. Engle [91] introduced the Autoregressive Conditionally Heteroskedastic (ARCH) model as a way to simulate the serial correlation of volatility. We follow Tsay [225] in describing the ARCH/GARCH models.

In essence, in ARCH models the price changes by a normal distribution with constant mean and time-varying variance. We denote by r_t log-return on day t:

$$r_t = \log(p_{t+1}/p_t)$$

where p_t is the price (at a specific time, usually the close price) on day t. At this moment, we are not considering transaction costs.

The ARCH(m) model defines a stochastic process for generating the sequence of log-returns $r_1, r_2, \ldots$. The process is defined over the *mean adjusted returns* which are defined as $a_t \doteq r_t - \mu$ where μ is the fixed *mean return*. The source of randomness is a sequence of "noise" random variables $\epsilon_1, \epsilon_2, \ldots$ which are chosen independently at random according to the distribution $N(0, 1)$. For $t = 1, 2, \ldots$ we compute the variance $\sigma_t^2(n)$ according to the formula

$$\sigma_t^2(n) = \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + \alpha_m a_{t-m}^2 \ .$$

where $\alpha_0 > 0$ and $\alpha_i \geq 0$.

Given σ_t we set $a_t = \sigma_t \epsilon_t$. To initialize the process we set $\sigma_0^2(n) = 0$.

Bollerslev [37] extended the ARCH model and proposed the Generalized Autoregressive Conditionally Heteroskedastic (GARCH) model to simulate volatility without having to calculate a large number of coefficients for polynomials of high-order. The GARCH models assume that r_t can be simulated with an autoregressive moving-average (ARMA) model. [4]

The $GARCH(m,s)$ model adds a distributed lag structure to simulate the conditional variance:

$$\sigma_t^2(n) = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2(n)$$

assuming that $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{s} \alpha_i + \sum_{i=1}^{s} \beta_i < 1$. This last condition assures that the unconditional variance of a_t is not infinite or a_t is stationary, and its conditional variance changes over time ($\sigma_t^2(n)$).

GARCH has also been used to calculate the value at risk (VAR). Value at risk is the maximum amount of a portfolio that can be lost in the worst case scenario with a certain confidence level.

VAR can be calculated using the one step ahead forecast of the log return $\hat{r}_{t+1}$ and the volatility calculation $\hat{\sigma}_{t+1}^2$ using the $GARCH(m,s)$ model. Then the VAR calculation at time t+1 is:

$$VAR = \hat{r}_{t+1} - \frac{t_v(p)\hat{\sigma}_{t+1}^2}{\sqrt{v/(v-2)}}$$

where $\frac{t_v(p)}{\sqrt{v/(v-2)}}$ is the pth quantile of a Student-t distribution with v degrees of freedom. In practice, this value is 1.65 when VAR is calculated assuming that the probability of the

---

[4] An ARMA(p,q) model to calculate r_t has the following form:

$$\hat{r}_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i}^2$$

occurrence of a major loss is 5%.

We calculate the one step ahead forecast of the log return $\hat{r}_{t+1}$, the volatility $\hat{\sigma}_t^2$, and the Sharpe ratio $(\hat{r}_{t+}/\hat{\sigma}_t^2)$ using the $GARCH(1,1)$ model (See Bauwens and Giot [24] for further details in the calculation of VAR using GARCH.).