Decision Support Systems 57 (2014) 212–223

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Empirical evaluation of an automated intraday stock recommendation system incorporating both market data and textual news

Tomer Geva *, Jacob Zahavi

Recanati Business School, Tel Aviv University, Israel

ARTICLE INFO

Article history:
Received 15 December 2011
Received in revised form 7 September 2013
Accepted 9 September 2013
Available online 5 October 2013

Keywords:
Stock returns
Investment decisions
Data-mining
Prediction
Recommendation

ABSTRACT

In this study we evaluate the effectiveness of augmenting numerical market data with textual-news data, using data mining methods, for forecasting stock returns in intraday trading. Integrating these two sources of data not only enriches the information available for the forecasting model, but it can potentially capture joint patterns that may not otherwise be identified when each data source is employed separately. We start with market data and then gradually add various textual data representations, going from simple representations, such as word counts, to more advanced representations involving sentiment analysis. To find the incremental value of each data representation, we build an end-to-end recommendation process including data preprocessing, modeling, validation, trade recommendations and economic evaluation. Each component of the modeling process is optimized to remove human bias and to allow us to impartially compare the results of the various models. Additionally, we experiment with several forecasting algorithms to find the one that yields the "best" results according to a variety of performance criteria. We employ data representation procedures and modeling improvements beyond those used in previous related studies. The economic evaluation of the results is conducted using a simulation procedure that inherently accounts for transaction costs and eliminates biases that have potentially affected previous related data-mining studies. This research is one of the largest-scale data-mining studies for evaluating the effectiveness of integrating market data with textual news data for the purpose of stock investment recommendations. The results of our study are promising in that they show that augmenting market data with advanced textual data representation significantly improves stock purchase decisions. Best results are achieved when the approach is implemented with a nonlinear neural network forecasting algorithm.

1. Introduction

1.1. Background

Successful strategies for stock trading decisions have clear and obvious benefits. Consequently, this topic had received much attention in data-mining studies, which report on a myriad of methods and algorithms for automated stock investment decisions.

Yet, successful construction of such models still remains a challenging undertaking. Key financial theories such as the Random Walk model of stock prices [18] and the Efficient Market Hypothesis (EMH) [19] indicate that it is difficult to generate excessive, risk-adjusted trading profits, based on currently available public information. Data-mining literature has also reported on a host of difficulties in predicting stock behavior. For instance, Dhar and Chou [11] mention that financial markets are inherently "noisy" with several types of nonlinearities. Likewise, Dhar et al. [12] refer to (a.) weak theory that results in a large number of variables, thus increasing the dimensionality of the problem; (b.) relationships between variables being weak and nonlinear; and (c.) the potential significance of variable interactions.

Most studies that predict future stock returns rely primarily on historical market data such as stock price and trading volume (e.g., [25,33,38,56]). While historical market information is relatively simple to obtain and process, the downside is that it practically ignores news concerning real-world events such as: mergers and acquisitions, dividends, interim results, analyst upgrades or downgrades, changes in management, and others—which have been documented to have an impact on stock returns [30,48].

News reports about companies abound these days and are commonly available in textual format on websites, through financial news providers such as Reuters and Bloomberg, and through other delivery systems. These data sources provide significant information about real-world events that historical market data can capture, at best, only partially and indirectly. Indeed, a host of methods, collectively called text mining or text analytics, have been developed in recent years for capturing and representing textual data. The output of such methods can potentially be used as additional explanatory information in financial prediction models. Yet, in contrast to the myriad of studies using

* Corresponding author.
E-mail address: tgeva@tau.ac.il (T. Geva).

0167-9236/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dss.2013.09.013

market-based explanatory data, relatively few studies have explored financial recommendation based only on explanatory variables obtained from textual data (e.g., [22,23,32,39,41,42,46]).

Using only textual data in prediction models also has its drawbacks, as it misses out on joint patterns of stock price changes and news publication. Examples of such joint patterns have been described in [7] and [45]. Clearly, only a trading recommendation system that is based on both market data and news data can benefit from the synergy between these two different sources of information and has the potential to detect their combined effect on stock prices.

Despite the potential benefits of integrating both types of data for predicting stock returns, this approach has been addressed in only a very limited number of data-mining studies. Among these studies we mention [50,53,54].[1] Additionally, previous data-mining studies have not conducted a systematic evaluation of the informativeness of augmenting market data with textual-news data.

In addition to being investigated in the data-mining literature, algorithmic trading has also been discussed in the financial literature. Hasbrouck and Saar [27], Hendershott et al. [28], Hendershott and Riordan [29] as well as Domowitz and Yegerman [15] examined how the stock market is influenced by algorithmic trading. These studies found that algorithmic trading generally improves various aspects of market quality such as liquidity and volatility. Other papers from the financial literature [1,9,51,52,55] have examined the informativeness of different data sources such as news and message boards which appear in textual format. For example, the objective of Antweiler and Frank [1] was to use message board postings to predict returns, and Tetlock [52] looked at whether media pessimism could be used to predict downward pressure on market price and trading volume. Our study contributes to the evaluation of the informativeness of textual news data by comparing it to the predictive capabilities of market data, and by measuring the incremental value of adding more sophisticated textual data representations. For this purpose we have constructed an "end-to-end" decision support system that first brings the different data representations to a "common ground" for comparison purposes, and then derives and evaluates actionable trading recommendations.

1.2. Research goal and methodology

Our objective was to conduct a systematic evaluation of the effectiveness of augmenting market-based data with textual news-based data for intraday stock recommendation decisions. For this purpose, we successively augmented market data with five different sets of textual data representations at different levels of complexity and sophistication, and measured the effect of each additional representation on predictive performance. We began with simple data representations, including a news item count and a "Bag of Words" representation, and gradually moved to more elaborate methods such as categorization into business events, news item sentiment scores and, finally, "calibrated sentiment" scores—a method proposed for the first time in this study.

We conducted this analysis while taking into account the *joint* effect of data representation and forecasting algorithm on intraday stock recommendation decisions. Our contention is that, due to the strong interrelationship that commonly exists between the data representation and the forecasting algorithm, this joint effect has considerable influence on the results of recommendation models. It is well known that a data representation that produces superior results with a certain data-mining algorithm might underperform when a different type of data-mining algorithm is employed. This contention is also supported by a previous pilot study [24] conducted as part of this research.

Specifically, to explore the influence of the forecasting algorithm, we ran each data setup using each of the following three data-mining algorithms, selected from three different "families" of algorithms: (a.) non-parametric feed-forward neural network; (b.) decision tree involving a genetic algorithm; and (c.) parametric linear logistic regression model.

To implement a systematic evaluation methodology, we have built an end-to-end modeling process that encompasses an entire array of tasks, including data pre-processing, feature creation, feature selection, forecasting algorithm, recommendations and economic evaluation. The various components were optimized for each data and algorithm combination in order to bring all modeling setups to a "common ground" for comparison purposes. The main modules included in this process are presented schematically in Fig. 1.

The output of the process consists of a series of trade recommendations over time. The economic value of the recommendations was evaluated by means of an elaborate simulation process that inherently accounts for the transaction cost of buying and purchasing stocks, thus avoiding biases that might have affected previous related data-mining studies.

Finally, to address the dynamics of this process over time, we used a "sliding window" approach, in which we used three months' worth of data to train a model and then validated the model's performance on data from the succeeding month. We used 11.5 months of available data in our research, resulting in 9 successive modeling setups and 8.5 months of independent data for validating the modeling performance. Fig. 2 illustrates the sliding window approach.

1.3. Intraday prediction challenge

Our study focuses on the short-term, intraday level, where stocks are purchased and sold during the same trading day. We chose to work at the intraday level since it provides a good test-bed for studies dealing with stock recommendation decisions while using news-based data. This is because utilization of intraday data is challenging due to their inherent "noisiness", and intraday trading is generally considered a risky practice.[2] On the other hand, intraday trading provides significant opportunities for automated, news-based, trading systems. These opportunities arise from the short time intervals during which stock prices may remain inefficient after news reaches the market. It is during this time that automated trading systems, reacting rapidly to new information, can potentially render profitable trades. The length of the interval in which the stock price remains inefficient has been researched in various financial studies. Studies such as [7] and [8] observed stock price inefficiency during intraday time intervals ranging from less than 1 min to 30 min.[3]

1.4. Contribution

Our main contribution is evaluating the informativeness of augmenting market data with textual-news data by means of an "end-to-end" recommendation system. This recommendation system is designed to bring different data source representations to a common ground for comparison purposes. Specific contributions include the following items, not previously reported in related data-mining literature: (a.) Using a multi-stage automated process for feature selection and transformations, designated to maximize the informativeness of market

[1] We note that [58] and [49] might have also utilized both market and textual data as explanatory variables within their forecasting models. However, the exact way in which market data are appended to textual news data is not explicitly mentioned. We also note that a few studies mentioning that they concurrently use or integrate textual news data with market data effectively utilize the market data only in relation to defining the dependent variable (e.g., [22,32]).

[2] Intraday trading usually involves substantial commissions and execution costs due to voluminous trading activity. Intraday trading usually also requires expensive data feeds as well as processing, analysis and testing software. Due to the common practice of "margin trading" in conjunction with intraday trading, risk substantially increases.

[3] [7] went further to demonstrate the feasibility of exploiting short-term market inefficiencies in the case of a financial news television program.