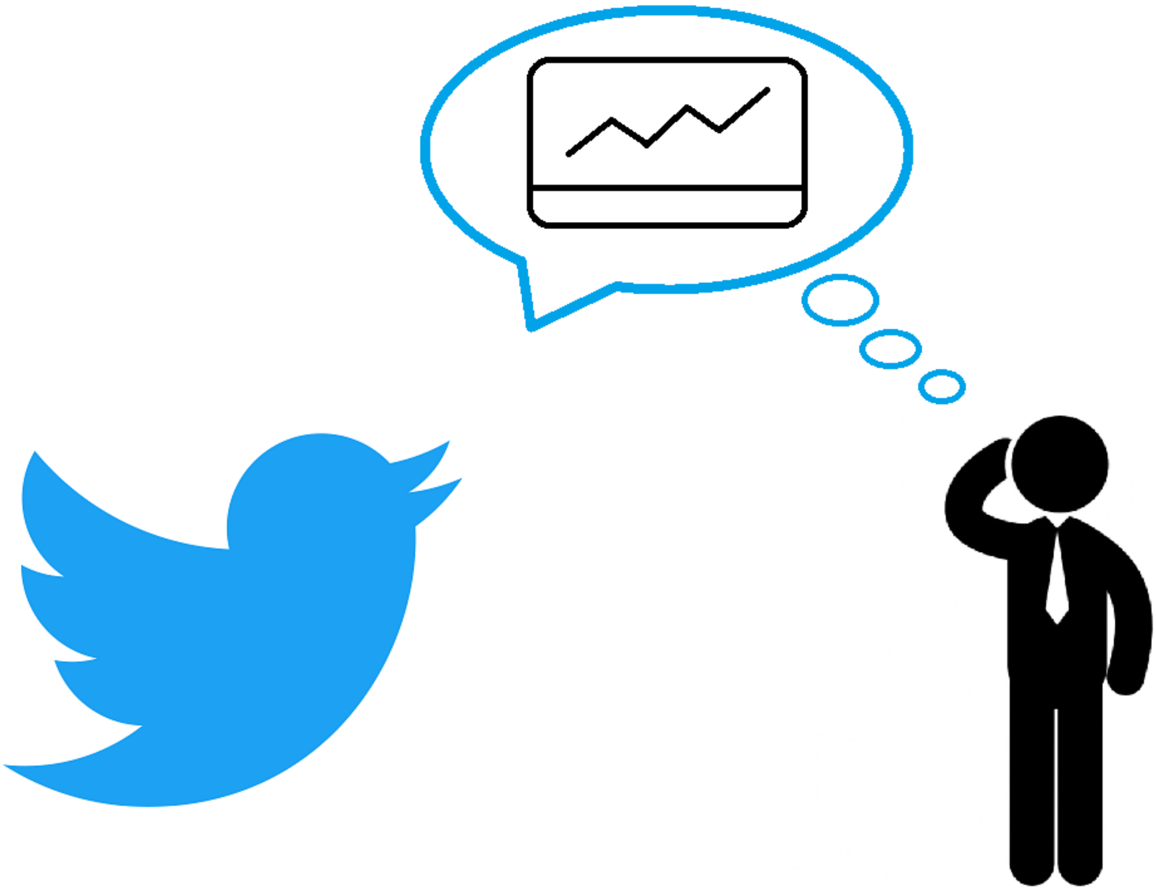
Breaking News: The Influence of the Twitter Community on Investor Behaviour



Bachelorarbeit

zur Erlangung des akademischen Grades „Bachelor of Science (B. Sc.)“ im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik, Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Matthias Hörnschemeyer

[redacted]

Prüfer: Prof. Dr. H.-J. von Mettenheim

Hannover, den 15.11.2016

Contents

List of Figures

List of Tables

List of abbrevations

EMH	Efficient Market Hypothesis
EPU	Index of economic policy unvertainty
GI	General Inquirer
SVM	Support Vector Machine
SVR	Support Vector Regression
S&P500	Standard & Poor’s 500
VIX	CBOE Volatility Index

1. Introduction

The financial market is the place where people trade with securities. According to the 'Efficient Market Hypothesis' ('EMH') (Fama, 1965) prices follow a 'random walk' that is unpredictible by past information whereas markets are ‘informationally efficient‘. This means that since relevant news spread very quickly amongst investors all available information are at any time included in price building and therefore do not allow returns in excess of market average over the long run. This theory is based on two basic assumptions:

a) All pieces of information are available and assessable for every possible investor at any time.
b) All market participants act rationally on the basis of these information.

on a)

On real markets this is rarely the case. Ekanshigupta, Preetibedi and Poonamlakra (2014) state in their work that different lifestyles influence the access to information, whereas the vast variety of informative channels reduces the chance to keep up with the amount of available news. Nevertheless the existance and unbroken success of financial institutes and famous investors like Wall Street Billionaire Warren Buffet suggest that above average returns can indeed be achieved. In contradiction to the EMH before buying or selling securities investors usually use the so-called 'Fundamental Analysis' (Taylor and Allen (1992), Oberlechner, T. (2001)). In the course of this they create an image of the company or the security itself and gain pieces of information that are not included in the price and allow a more precise forecast.

on b)

There are indications, that investors do not act fully rational: In 1990 De Long et al. found, that irrational market behaviour of noise traders causes mispricing, higher volatility and further noise trading in markets (De Long et al., 1990). Further, researchers like Baker and Wurgler (2006) and Bathia and Bredin (2013) have proved evidence for a strong correlation between sentiment of investors and stock returns. Since creators of all trades are human beings the theory of Behavioral Finance puts forward the influence of emotions on peoples' choices to address such findings. It uses psychological and social aspects to explain human decision making in financial markets involving irrational behavior like loss aversion and the concepts of mental accounting, prospect theory and self-control (Shefrin and Statman, 1985). Recent researches from this field show that the way traders react towards information varies depending on gender (Lee et al. 2013) and cognitive biases (Friesen and Weller, 2006). Effects of social groups are also addressed: Hong, Kubik and Stein (2004) prove that a decision for or against market participation is affected by social relations. Moreover, Banerjee (1992) describes 'Herding' as the process, in which individuals follow group behaviour instead of using their own information. Transporting this into the market behavior one can assume that social structures affect decisions of investors. Indeed, in his work ‘social mood and financial economics’ John R Nofsinger (2005) states that ‘the social mood influences financial decision makers' and leads to irrational, 'noisy' market behaviour. Consequently a crucial point for researchers as well as for investors is to detect social mood and investor sentiment to evaluate and predict price changes properly.

These two points lead to the following conclusion: Investors are to be interested in achieving unique, reliable information on stocks, companies and business trends. Their behaviour is shaped by social and psychological factors. Therefore it is crucial for researchers to understand their relation to sources of information on the one hand and the community which they are part of on the other hand. It is the aim of this paper to understand the impact on the average investor more precisely.

Chapter 2 deals with the development of a Behavioural Economic framework converning the influence on an average investor. In chapters 4 and 5 a Twitter data set is analysed regarding two research questions. In either case, the results are discussed based on former results. Chapter 3 gives an overview of the methodology and the observed data. In chapter 6 the results are summarised and an outlooks and suggestions for further research are given.

2. Literature Review and development of the behavioural model 'The Average Investor'

2.1: Literature review: Text mining for market prediction

Over the last two decades the internet has developed into the major medium of spreading information in which investors can obtain financial news on websites or blogs. With the growing field of Data Mining various approaches have been made in order to gain valuable information and detect sentiment of investors. These divide into two basic categories regarding their assumptions on the investor and the market theory they propose:

Category A: Some researchers predict stock market movements by extracting information directly from news articles, blogs, company announcements and other public sources. Jhai and Cohen (2011) analyse newspaper articles to detect sentiment in these and successfully predict the stock market as they concentrate on finance-related columns. Hagenau, Liebmann and Neumann (2013) for their part focus on company announcements. Their model explains stock movements following the release. Other authors use publications from internet platforms where statements on the market are released more frequently than via newspaper articles. Jin et al. (2013) for example make use of this to filter and analyse articles from a common financial news platform and forecast trends in the Foreign Exchange Market. In a similar manner Azou (2009) develops a model with relevant news on individual stocks from an online platform to predict market returns.

Category B: Authors belonging to this category analyse text content, which is neither published as a piece of financial news or announcement by a news site nor does it necessarily relate to a particular stock or brand. Instead, influential mood and sentiment revealed by random users is employed for market prediction. In most cases microblogging platforms such as Twitter are used. The advantage of micro blogs in this context is the high number of users and userposts every

6. Conclusion and suggestions for future work

> ***'Man is so intelligent that he feels impelled to invent theories to account for what happens in the world. Unfortunately, he is not quite intelligent enough, in most cases, to find correct explanations. So that when he acts on his theories, he behaves very often like a lunatic."***
>
> **Aldous Huxley (1932)**

In chapter 4 of this paper a data set of Twitter news has been used to develop a model for stock market prediction. Approaches on the basis of many regular sites without further specialisation on Finance in particular and a linear regression have not been successful. By contrast, the reduction of the observed authorset in combination with a Support Vector Regression has helped in achieving a significant prediction accuracy for the Standard&Poors 500 and two other indices. Even though the model is not a state-of-the-art engine, its success in predicting various indices proves that stock prediction based on Twitter news channels is indeed possible. At the same time it leaves space for further development with regard to noise reduction and strategies for market turbulences and the choice of observed channels.

Chapter 5 addresses the impact of different sets of authors on the prediction accuracy of various sectors. In order to create an efficient sentiement analysis tool for market prediction it is crucial to observe data that contains valuable information. Results of the sector-specific investigation show that in the case of Twitter news channels the validity regarding companies from the 'Information Technology'-sector exceeds those of other sectors. This aspect is useful for prediction in this case and it is useful for researchers' future prediction engines. The question whether the impact of the frequency of mentions and news releases on prediction accuracy differs among various media channels remains for further research.

The anterior part of the paper addresses the development of the behavioural model 'The Average Investor'. EkanshiGupta, Preetibedi and Poonamlakra (2014) state in their work: 'To conclude, the new paradigm of Behavioural Finance emerged as a model that successfully attempted to challenge and refute the traditional financial theory.' Despite this, 'The Average Investor' is not designed to 'refute' classical models of Finance. Instead, its purpose is to supplement them and to establish more sufficient models by including further aspects into conventional methods of market prediction.

Transferring this idea to the present case it would be compelling to investigate, by what kind of news public mood is changed and whether it is predictive of the stock market (Concept III and II from 'The Average Investor'). Moreover in order to reveal the influence of news releases on stock movements it would be important to find out whether determinants of the former investigation contain fundamental information on stocks (Concept I). This would also address questions concerning the existence and the impact of a dual causality for stock price movements - fundamental information and investor sentiment- which is suggested by the results. Finally, by this one could deconstruct the mechanisms of price building and point out, wether they relate to the doctrine of efficient markets (EMH) or actually originate from Behavioural Finance.