RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS No. 104
Business Informatics 16 2010

Aleksander Fafuła
Wrocław University of Economics, Wrocław, Poland
aleksander.fafula@ue.wroc.pl

A PROTOTYPE OF A PLATFORM FOR DATA-DRIVEN APPROACH TO DETECTION OF COGNITIVE BIASES

Abstract: Cognitive biases often influence decision processes related to investment on stock markets. These include complex problems with perception and understanding of surrounding reality. This research aims to detect cognitive biases in the data-driven way. The chosen are: gambler's fallacy, base rate fallacy, clustering illusion, bandwagon effect, overconfidence effect. Specially constructed trading platform was used to perform the experiment. Along with the standard functionalities of a stock market simulator a few tools from cognitive science were applied. Inquisitive results show some patterns in behaviour of traders. Results are promising and hopefully deserve further investigation.

Keywords: cognitive biases, trading simulator, patterns in behaviour.

1. Introduction

"Some things in life are bad / They can really make you mad / Other things just make you swear and curse" – like in Monty Python's famous song we can encounter bad moments in our lives. Sometimes they are not our fault but sometimes they are solely our achievement. As human beings we all make mistakes and this is actually natural. "Behaving" wrong on the stock market can cause serious damage to our future and even health. Sometimes we are under the influence of a very common source of problems – cognitive biases. Is it possible that we can detect them automatically?

The science of Investment Psychology points out various cognitive biases. They have been widely tested for many decades. The best known publications in this area are: *Judgment Under Uncertainty: Heuristics and Biases* by Kahneman, Slovic, Tversky (1982), *A Mind of Its Own: How Your Brain Distorts and Deceives* by Fine (2006), *Cognitive Biases in Market Forecasts* by Fisher and Statman. Polish best known seems to be *Psychologiczne pułapki oceniania i podejmowania decyzji* by Tadeusz Tyszka (2000). Due to the nature of this matter some models of behaviour are known better than the others. A few of them have been chosen as they appear to be more common and easy to detect.

1. Gambler's fallacy – the tendency to think that future probabilities are altered by past events, when in reality they are unchanged. Results from an erroneous con-

ceptualization of the law of large numbers. For example, if an abundance of heads comes up on a coin, observers may be heard to assert that a tail is due; that it is more likely to come up than another head [O'Neill, Puza 2004].

2. Base rate fallacy – this is people's tendency to ignore base rates in favour of, e.g., individuating information (when such is available), rather than integrate the two. This tendency has important implications for understanding judgment phenomena in many clinical, legal, and social-psychological settings [Bar-Hillel 1980].

3. Clustering illusion – the tendency to see patterns where actually none exist. For example Thomas Gilovich found that most people thought that the sequence "OXXXOXXXOXXOOOXOOXXOO" looked non-random, whereas, in fact, it has several characteristics maximally probable for a "random" stream, such as an equal number of each result and an equal number of adjacent results with the same outcome for both possible outcomes [Gilovich 1991].

4. Bandwagon effect – the tendency to do (or believe) things because many other people do (or believe) the same. Related to groupthink and herd behaviour [Leibenstein 1950].

5. Overconfidence effect – excessive confidence in one's own answers to questions. For example, for certain types of question, answers that people rate as "99% certain" turn out to be wrong 40% of the time [Pallier et al. 2002].[1]

Full list and extensive descriptions of experiments can be found in the literature. However, cognitive science has not been introduced to data-mining at scale provided by computing power and network availability. It is unclear and thus needs further investigation if data-driven approach is able to give additional point of view for cognitive biases detection.

This paper presents the results of a preparation to an experiment introducing the data-driven approach to detection of cognitive biases. To make results more clear it focuses on newbie investors.

2. Methods

The experiment was conducted in a form of investment game. Special stock market simulator was written to provide extra functionalities. These include abilities for: storing additional information in the database, collecting data using questionnaire, displaying additional views for administration. The codename of this web application is French *Étoile* which means *Star* in English.

Traders were able to deal only with one hard coded product at this early stage. Due to the data availability France Telecom (FTE.PA) was chosen. It is registered on the Euronext Paris, therefore is operates in convenient stock market opening hours. Every user was given $1,000,000 and time from December 2009 to January 2010.

[1] The numbers above correspond with the numbers of questions provided in the next chapter. Probably only one cognitive bias should be tested at once. This will be taken under consideration in the future work.

Trading was available only during the stock market opening hours (excluding holidays etc.).

Étoile was designed as an interface for acquiring data from traders and bots using various methods. It is constructed on a well-known client-server architecture. Further, processing is done by external applications and imported back to Étoile. The application is used for:

1) data collecting (from brokers and bots/agents),
2) data cleaning (to filter, remove noise and inconsistencies),
3) data integration (to merge data from stocks and various databases),
4) data selection (export from database to data mining applications),
5) data transformation (data is being transformed to suitable formats).

As mentioned earlier, the main front-end is a simple Website. It can be accessed by any modern Web browser, palm device, tablet or a smart phone. User is able to check here some basic information, such as current server time (UTC), quotes registered by the server, their capital and other. It also has buttons which one presses when buying or selling. Information displayed depends on the mode in which the program is operating, that is, anonymous, competition, or developer. Moreover, it is possible to change some basic settings.

First of all, the user is obliged to create an account, and then log in to the platform. Next, he or she should choose desired quotes and start ordinary trading by using own system and pressing buy/sell Étoile's buttons. History of trading operations was also implemented. It displays information like: type of operation (buy/sell), time/date, quantity, price, volume, commission, total turnover. Current close is updated with AJAX so there is no need to reload Web page.

On the other hand, the server basically collects buy/sell signals from brokers. Then they are prepared to be parsed. This was done by filtering abnormal signals and checking for erroneous values. It is possible to attach any number of trading bots using various decisions and collect their signals, too. This is done by http GET method and it is simple to implement. This gives the system an opportunity to compare signals from brokers and bots and detect which algorithm has been used by the broker. Users can adjust the system's parameters so the given probability is achieved. Moreover, it is possible to search for similarities and patterns not only among brokers (real and bots) but also in a single session registered for user. Decisions can be later evaluated and explained. This gives one a chance to find patterns in trading behaviour and offer some advice in the future. This is going to be implemented in the next release of Étoile that is going to have more of its features displayed in the real time.

Technology used by Étoile is based mainly on Open Source. The interface uses AJAX, and displays SVG quotes diagrams. On the server side, there is FreeBSD web server with Apache http daemon, PHP as cgi and MySQL. AJAX has been chosen to give the user a sense of reliability and to save net traffic. SVG is very fast and uses small amount of memory to operate and transfer. In addition, it is light for the server and can thus be used to display real time quotes.

Étoile is written in a very simple way and it does not make use of any frameworks, so it saves CPU time. To shorten the transaction time there is also native and direct communication with a database, and so there is no ORB. Queries are simple and there also appears strong tables optimization to improve timings.

For this amount of the data collected, there are a few improvements that can be done. Databases can be easily arranged into clusters. Redundancy can be built on the level of network routing (dns/Round-Robin, etc.).

Welcome to Étoile ;-)

Your decision has been registered. Please fill the form below to continue.

All questions must be answered.

Q 1. Look at the current situation of the stock market on the diagram. How do you think –– what is the most probable trend in close future?

Rise

Fall

No changes

Q 2. List your indicators and their generated signals (for instance: RSI→Buy = True), as well your final decision:

[Type your answer here]

Q 3. Have you seen any similar pattern in the diagram before you have made this decision?

yes

no

uncertain

Q 4. Have you consulted your decision with the following sources:

Internet / press

friend

intuition

irrational rules

fundamental

random/none

Q 5. How do you evaluate your decision (1 = risky, 5 = sure):

1 2 3 4 5

[Submit button]

Listing 1. Étoile – experimental platform for behavioural pattern recognition – questionnaire for the user after investment decision

To build a matrix of stock data and human behaviour, a bit of cognitive science was implemented. It was done with the use of simple questionnaire. Questions use various input methods, for example: yes/no/don't know, text input, multiple choice, scoring (1-5). After every operation (i.e. buy/sell) user is being forced to answer five simple questions. This is a work in its very beginning and needs further investigation.

3. Results

Sample collected with the usage of Étoile shows some encouraging information. Thus it is clear that more wide experiment should be carried out. Results show that the decisions of the users, supported and related to answers to simple survey, can indeed create patterns. Questions are asked every time a user makes a decision. Survey should contain more sophisticated questions, this is why this area can also be improved to make results more meaningful.

During the experiment time FTA has not experienced significant changes (Figure 1). December was in falling trend until 21st. Then a small rise followed. Users operated in mixed conditions.

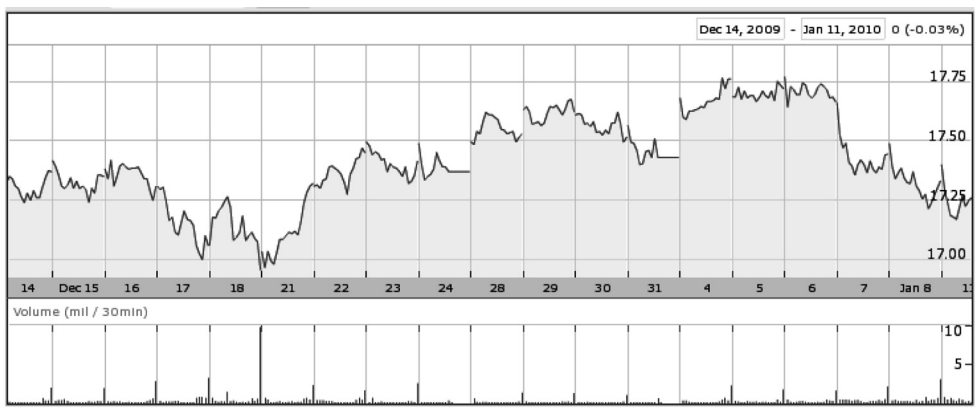

Figure 1. Chart of France Telecom during the experiment time

Source: www.google.com/finance.

Trading students when asked question #3 "Have you seen any similar pattern in the diagram before you have made this decision?" answered 68% for Yes, 15% No and 17% were Uncertain. This should signal the clustering illusion being in effect as they commonly see non-random-looking patterns. Especially when comparing to charting skills of non-experienced traders (we can assume they are not that familiar with charts). There is also a tendency, where users convert from other decisions to Yes over time. This happens even despite the fact the data set they worked on was small. Interesting is that the average of transaction amount is more than twice higher when the answer is No, then Yes.

Figure 2 shows that traders are more sure of their decisions when they are under the clustering illusion. This is of course a false signal of unjustified confidence. Many psychologists consider gambler's fallacy the main cause of clustering illusion. So perhaps detecting clustering illusion gives a right to indicate gambler's fallacy.

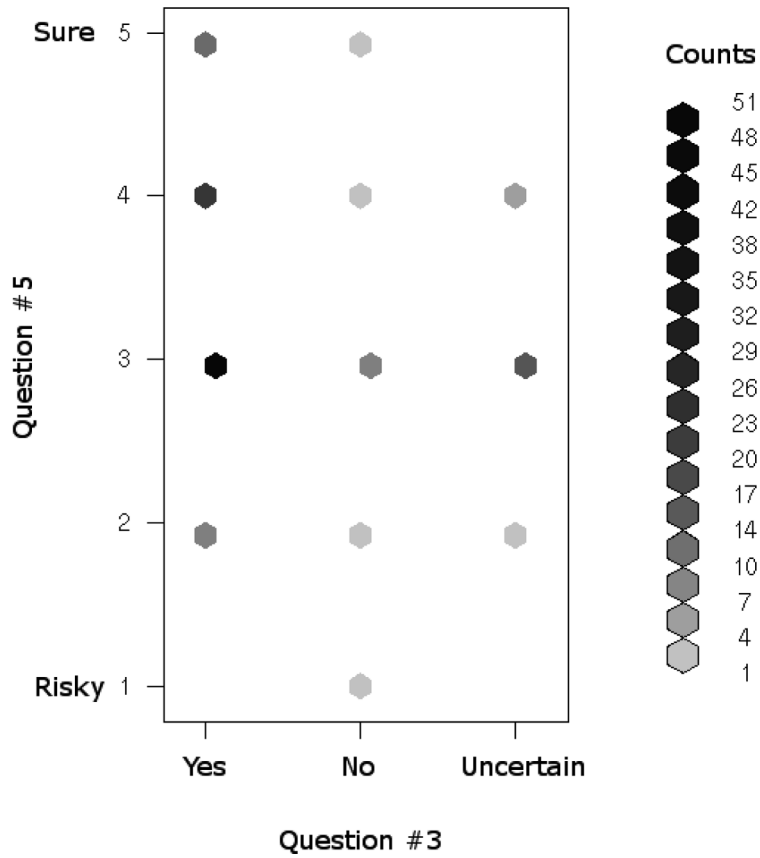

Figure 2. Scatterplot with binning

It is also worth mentioning that patterns appeared in results of the data collection process itself can be used for various purposes. Samples can be stored in databases, reused, discussed, validated and improved. However, the most significant achievement is to use it for searching for the patterns in behaviour.

Conducting this experiment with newbie users also brings a few interesting ideas. They appear to be easier to "tag". This is because of the fact they do not operate on the investment markets professionally. Moreover they tend to use the same tools like web resources and applications. It is also likely they use similar technical analysis tools. This is also confirmed by the results of the survey.

It is clear that the group of beginning investors is the largest and needs more decision-supporting tools. The experiment shows they are influenced by similar and simple cognitive biases. Finally it is possible that perhaps sharing this knowledge with them is going to improve their behaviour and increase the value of portfolio. This makes justification for students and newbie investors to give this simulator a try.

Despite the facts that the sample collected with this experiment is too small and the survey does not fulfil the standards of cognitive science, the results are encouraging for making further efforts to improve it.

4. Discussion

Étoile is a work in progress and still under heavy development. It is necessary to discuss this experiment with the wider audience, involving especially: psychologists, financial specialists and traders. This will enable the system to be extended and improved. Technical plans for the future direct project towards its usability. It may be advisable to create Google Gadget (http://desktop.google.com/plugins/) to present Étoile to a large number of users. There are many popular Google gadgets that are connected to stocks and trading, such as these listed at http://desktop.google.com/plugins/c/sidebar/finance.html. Another idea is to make the system available not only during the controlled sessions but worldwide at any time. This would involve a complete automation of data mining processes and some additional features like RSS with results, making notes, and additional expert panel.

Finally Étoile is also meant to be used as an educational platform that is able to point and thus eliminate many behavioural problems automatically.

Acknowledgements: I would like to express the deepest appreciation to professor Jerzy Korczak, whose encouragement, guidance and support helped me to write the paper. His knowledge and experience in the research area, artificial intelligence and stock markets were invaluable.

References

Bar-Hillel M. (1980), The base-rate fallacy in probability judgments, *Acta Psychologica*, Vol. 44, pp. 211-233.

Gilovich T. (1991), *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*, The Free Press, New York.

Leibenstein H. (1950), Bandwagon, snob, and Veblen effects in the theory of consumers' demand, *The Quarterly Journal of Economics*, Vol. 64, No. 2

O'Neill B., Puza B.D. (2004), *Dice have no memories but I do: A defence of the reverse gambler's belief*, reprinted in abridged form as: O'Neill B., Puza B.D., In defence of the reverse gambler's belief, *The Mathematical Scientist*, Vol. 30, No. 1, pp. 13-16.

Pallier G. et al. (2002), The role of individual differences in the accuracy of confidence judgments, *The Journal of General Psychology*, Vol. 129, No. 3.

WYKRYWANIE BŁĘDÓW POZNAWCZYCH POPRZEZ PODEJŚCIE ZORIENTOWANE NA DANE – PROTOTYP PLATFORMY

Streszczenie: błędy poznawcze wpływają często na procesy związane z podejmowaniem decyzji inwestycyjnych na giełdach papierów wartościowych. Do błędów tych zalicza się złożone problemy postrzegania oraz rozumienia rzeczywistości. Artykuł opisuje konstrukcję prototypu systemu, którego przeznaczeniem jest detekcja błędów poznawczych. Wykrywanie tychże zostało oparte na metodach drążenia danych. Wybrano następujące błędy poznawcze: paradoks hazardzisty, iluzję grupowania, zaniedbywanie miarodajności, efekt polaryzacji i zasadę podczepienia. Do prowadzenia badań skonstruowano symulator giełdowy, który oprócz standardowych funkcji został rozbudowany o mechanizm prowadzenia ankiet psychologicznych. Rezultaty badania są obiecujące i zasługują na pogłębienie studiów na ten temat.