Predicting Product Performance with Social Media

Liviu LICĂ, Mihaela TUTĂ
Academy of Economic Studies, Bucharest, Romania
lica.liviu@yahoo.com, mihaela.tuta@gmail.com

Last 20 years brought massive growth in IT&C world. Mobile solutions such as netbooks, laptops, mobile phones, tablets enable the wireless connection to the Internet. Anyone can access it anytime and anywhere. In this context, a part of the activities from the real world have a correspondence in the online discussions. Social media in general and social networks in particular have turned into marketing tools for organizations and a place where people can express their opinions and attitudes about products.The paper shows how social media can be used for predicting the success of a product or service. To showcase this, two case studies are presented; a test to prove that the conversations that take place in social media are a good indicator of success and the second is an exercise to predict the winner of the Oscar for best picture in 2011.
Keywords: *Social Media, Social Networks, Prediction, Movie, Internet*

1 Introduction

The current paper begins with a presentation of social networks in the current IT environment. The third section describes the premise of using social networks for predictions. The fourth part contains a short overview on the previous attempts made by Johan Bolen and Hal Varian to predict the evolution of the stock markets and to forecast the level of activity for a number of different economic time series.

Section five and section sixpresent two case studies. The first case study describes a method for predicting movie ratings by using discussions about these movies on Twitter. This topic of movies was chosen because it is popular and talked about, with enough variety to allow calibrating and validating the predictions against ratings from websites like IMDb or against the values of the box-office. The base of this first case study is the analysis of sentiments and the main tools used were an custom web applications that makes use of a third party service called Tweet Sentiments and the commercial tool called UberVU.

Two weeks before the Oscars, data from Twitter was collected and based on it, authors tried to make predictions about the winners. The result was accurate, compared with the results obtained by Google Trends and Yahoo who used a quantitative approach.

Section seven presents the conclusions, limits of the current study and further investigations that can be developed for future research.

2 Social networks are on the rise

Social networks are a manifestation of the concept of social media; they represent web applications that use the Internet as the backbone for communication and the idea of a social graph in order to allow easy communications between its users. A social graph is mapping of the relations between a group of people. What is important to take away from this is that social networks are based on relations. The nature of the relation can vary greatly, from friendship or kin, as it's mostly the case of Facebook, to professional, for LinkedIN and common interests in a topic for Twitter.

The last few years have brought impressive growth to the field of IT&C. Mobile devices have evolved with great speed, smartphones are becoming a commodity and their price keeps going down, together with the data plans. Wireless access to the Internet is becoming omnipresent. Mobile means also netbooks and laptops and tablets, which have been growing strongly as well. Some predict that by the end of 2015 the mobile web will surpass the desktop Internet connections [1]. Inseparably linked to the growth of mobile is the growth of social networks. And if the

growth of the mobile web is impressive, the growth of social networks is astounding. Facebook is the superstar of the moment, with more than 600 million users [2], it became the subject of an Oscar winning movie, and Mark Zuckerberg was named person of the year by Times. In addition more than 60% of Fortune 500 companies are now using social media as marketing tools [3] [4] and the number is growing. Time spent on social networks has increased in 2010 with 82% over the previous year [5] [6].

With all this it is hard to see social networks as just websites used by teenagers. They have turned into tools used by people of all ages to communicate and it is not unconceivable that they will turn into something as common as e-mail. One of the most impressive changes of heart towards social media has taken place inside the U.S. Army. If in 2010 the Army was banning the use of social networking sites for its members [7], in 2011 it published an excellent handbook for the use of social media [8] and embraced the useful parts of social media while trying to minimize the negative effects its use can have. In the handbook there is the following line „…*Social media is not a fad, if the Army ignores it, it will not go away*". This is also the opinion of the authors of this paper too. All the signs point that social media will continue to grow. It will find a way into our daily lives, just as e-mail did. Understanding and using correctly these tools is a must for any organization

3 How social networks can be used for predictions

Human behavior is notoriously difficult to predict. What happens with a product or service is based on that human behavior. Studying it has always been a struggle. A good salesman who owns a small shop can figure out, more or less, what his customers want and how good his products and services are received. But scale this to a giant organization and thousands of people and it becomes impossible to tell. Feedback from the market is crucial for understanding customer behavior.

Until not so long ago the only way to get feedback was by going out and doing a market research. Social networks bring to the table a new tool. In social networks, like Twitter, the data is publicly available and its many users constantly inform their followers about their opinions and beliefs. Users create, consume and share information and all these interactions are a snapshot of what happens in the world at every single moment. It is a stream of thoughts and opinions that can be harnessed to identify very accurately what people think about a product or service. Market research is still important. It is not possible to extrapolate from an 140 characters message on Twitter all the complex data a well built questionnaire can obtain, yet the possibility itself is amazing.

From this it is not hard to imagine that these thoughts and opinions and sentiments, expressed constantly, can be used to predict how well a product, service or organization is perceived. This is the hypothesis that the paper sets out to test. Before doing our own studies, a quick look at similar attempts is taken.

4 Previous attempts at predicting

Social networks and search engines that crawl and constantly take snapshots of everything that is online are invaluable sources of information. Besides the historical value, the information available online gives a good picture of how the society, at large, feels about the world around it or about specific topics. This is at the basis of some of the most interesting researches that use social networks and also at the base of the case studies presented in this paper.

One of the papers that generated a lot of interest is authored by Johan Bollen, Mao and Zeng. Their paper called "Twitter mood predicts the stock market" [9] uses a large data set of historical messages from Twitter, which is processed and filtered for messages that contain words like "I feel" and that express a state of mind to generate an index of the mood of the society at a certain moment in time. This information is fed into existing algorithms for predicting the stock market.

The result is that adding the mood of the society improves considerably the accuracy of the predictions. They managed to predict with an accuracy of 86.7% the evolution of the stock market in the analyzed period.

Another fascinating research was done by Hal Varian, as the Chief Economist at Google and Hyunyoung Choi [10]. They studied if Google queries help predict economic activity. They found that data from Google Trends can help improve forecasts of the current level of activity for a number of different economic time series, including automobile sales, home sales, retail sales, and travel behavior. The approach is different as it neglects sentiment analysis which, in this paper, is considered of great importance.

A takeaway from both papers is that in order to do a good analysis and to make predictions a field is needed that is narrow enough to allow focusing on specific topics, and wide enough to have the necessary data. It is also important to have accepted metrics, economic or of other nature, which can be the base of the calibration and testing of the predictions models. For the case studies in this paper movies were chosen as the topic of the prediction. There are other researches that make use of the same topic. SitaramAsur and Bernardo Huberman [11] use data from Twitter to predict the incomes from the box-office for a series of movies. Other researches [12], [13] take different approaches but Asur's is the most interesting. Google also has an interest in the subject, by using the same approach as in Varian's paper; they use data from Google Trends to try to predict the winners [14].

All these studies are a good example of the possibilities that the data available on social networks and on the Internet, as a whole, has.

5 Predicting movie ratings

In order to test if online conversations can truly be used for making predictions, a case study was built. It was chosen the topic of movies, as it is a popular and talked about subject, with enough variety to allow calibrating and validating the predictions against ratings from websites like IMDb or against the values of the box-office.

The study is described by the following statements:

- the subject of the case study are movies;
- the social networked used for gathering data about the movies is Twitter;
- the gathered messages are analyzed using a sentiment analysis tool;
- the sentiments are used as the primary input in the predictions;
- for sentiment analysis were used two services: Tweet Sentiments and uberVU;
- the calculated metrics were correlated against the IMDb ratings and the box-office.

From Twitter a total number of 856.196 messages were gathered and processed. The application used for gathering this data was purposely built and it ran for a total of 65 days. The messages came from 428.530 Twitter users. A number of 30 movies were chosen.

The main criterion for choosing the movies was the name of the movies. Movies with names that can cause confusion, like "Frozen", "Cash" or "Robin Hood" were avoided. The titles of such movies can be used in phrases that have nothing to do with the movie itself. On the other hand the words "How to train your dragon" are unlikely to be used in a conversation that is not about the movie with the same name. It was also a priority to group the movies by release dates, in batches of five or more, in order to have as much a homogenous selections as possible. The selected movies are listed in Table 1.

Table 1. The selected movies

Name of movie	Launch Date	Name of movie	Launch Date	Name of movie	Launch Date
Daybreakers	12/3/2010	Welcome to the Rileys	10/11/2010	Casino Jack	28/10/2010
Youth in Revolt	8/1/2010	Megamind	5/11/2010	Rabbit Hole	3/11/2010

The Spy Next Door	5/1/2010	Saw 3D	29/10/2010	Little Fockers	22/12/2010
The Book of Eli	29/1/2010	Morning Glory	10/11/2010	TRON: Legacy	17/12/2010
Tooth Fairy	22/1/2010	Paranormal Activity 2	22/11/2010	True Grit	25/2/2011
Edge of Darkness	19/2/2010	Hot Tub Time Machine	26/3/2010	Country Strong	7/1/2011
Falling Awake	29/1/2010	How to Train Your Dragon	19/3/2010	Season of The Witch	7/1/2011
44 Inch Chest	15/1/2010	Clash of the Titans	9/4/2010	Green Hornet	28/1/2011
From Paris with Love	5/2/2010	The Runaways	8/4/2010	The Dilemma	14/1/2011
Percy Jackson Lightning Thief	12/2/2010	Repo Men	19/3/2010	The Rite	28/1/2011

With all these precautions there were still a few problems with the chosen movies. Some had a very big budget, like "From Paris with love" and some a very low budget, like "Falling Awake". Also, the sentiment analysis tools are not smart enough yet not to be biased by words like "Revolt" in "Youth in Revolt", which influenced too much the re-

sults for this movie, being overly negative and rending it unusable.

After gathering all the information and calculating and index that in the paper was sentiment index, correlations were calculated between this index and the IMDb rating and the box office values. The index is calculated by using the following formula.

$$I_{sent} = 100 * \left(\frac{\left(Nr_{positiv} - Nr_{negativ} \right) / TotalMes}{2} + 0.5 \right)$$

where:

I_{sent} – Sentiments index

$Nr_{positiv}$ – Number of positive messages

$Nr_{negativ}$ – Number of negative messages

$TotalMes$ – Total number of messages, the sum of the positive, negative and neutral ones

Table 2 contains the data collected through the web application and also trough a third party service called uberVU. Below are given details about the meaning of the fields of the table.

Table 2. Data obtained through the web application and through uberVU

Query	Total nr of Twitter messages	Mesagges analyzed with TS	TS Neutral	TS Positive	TS Negative	TS sentiment index	Index Calculated using TS data	IMDb rating	Box Office	uberVU Sentiment Index
Daybreakers	4608	4596	2760	1327	509	58.89904	49	6.5	51,416,464	58.6
Youth in Revolt	6015	6005	312	1672	4021	**30.4413**	33	6.7	19,623,544	64.5
The Spy Next Door	1459	1458	884	422	152	**59.25926**	53.5	5.2	43,580,395	**94.8**
The Book of Eli	27514	27510	13718	8701	5091	56.56125	62.5	6.8	157,091,718	65.96
Tooth Fairy	37123	37080	12665	15709	8706	59.4431	97.5	4.8	112,361,476	44.7
Edge of Darkness	7259	7252	4663	1648	941	**54.87452**	57	6.7	81,124,129	64.1
Falling Awake	7420	7419	1572	2636	3211	**46.12481**	30.5	6.7	n/a	64.65
44 Inch Chest	421	420	231	106	83	52.7381	46.875	5.9	294,245	41
From Paris with Love	5952	5938	493	5174	271	**91.28494**	96.5	6.4	52,795,309	**96.4**
Percy Jackson Lightning Thief	4526	4515	2562	979	974	50.05537	50.5	5.8	226,497,209	58.1
Welcome to the Rileys	1559	1556	230	1288	38	90.1671	93	7.2	317,382	**97.4**
Megamind	17522	16876	7895	7150	1831	65.75907	64.5	7.3	317,415,120	76.15
Saw 3D	12809	12793	6914	4183	1696	59.72016	66	5.6	130,910,178	49.6

Morning Glory	10604	10595	196	10039	360	**95.67721**	95.5	6.8	46,686,565	62.75
Paranormal Activity 2	18219	18218	1830	4925	11463	**32.05621**	28.5	6	171,637,907	42.75
Hot Tub Time Machine	18100	18098	9438	6647	2013	62.80252	59	6.7	64,416,046	56.75
How to Train Your Dragon	37018	36958	17134	15951	3873	66.34017	67	8.2	494,878,759	88.05
Clash of the Titans	28438	28411	293	3226	24892	**11.8704**	9	5.9	493,214,993	55.85
The Runaways	56860	56760	15030	12094	29636	34.54722	31.5	6.6	4,681,651	65.1
Repo Men	8998	8996	5291	2293	1412	54.89662	46.5	6.3	18,409,891	51.5
Casino Jack	11093	11036	6875	2923	1238	57.63411	59.5	6.4	894,540	60.5
Rabbit Hole	38892	38762	21701	10848	6213	55.97879	49	7.5	1,789,719	49.05
Little Fockers	70038	69999	27055	37386	5558	**72.73461**	62	5.3	304,052,615	64.4
True Grit	64342	64172	20957	36003	7212	72.43268	70.5	8.1	161,868,013	60.5
Country Strong	52342	52296	23671	18473	10152	57.95568	65	5.4	20,051,573	51.6
Season of The Witch	39668	39582	24085	10196	5301	56.18337	52.5	5.5	64,750,774	57.15
Green Hornet	70932	70805	27075	37194	6536	71.6496	65	6.5	171,922,547	65.45
The Dilemma	54875	54802	1522	17200	36080	32.77435	32.5	4.9	60,592,125	55.75
The Rite	66365	66324	20155	32398	13771	64.04243	59	6.2	25,605,937	52.6
TRON: Legacy	0	0	0	0	0	**-1**	61.5	7.3	372,899,356	68.75

In order to be able to calculate the I sent index, the formula of which was given above, every collected message had to be analyzed semantically and its sentiment identified. For simplicity each message could have only one of the three states:

- positive;
- negative;
- neutral.

In order to do the actual sentiment analysis a third party service was used, called Tweet Sentiments. This service is abbreviated as TS in Table 2.

When all the data was collected it became obvious that the sentiment index for some of the movies contained inaccuracies. For example "Form Paris with Love" had a very high 96.5 sentiment index. After some further testing it became obvious that the problem was caused by the word "Love", which gives a very positive connotation to the message.

In these conditions, in order to double test the result obtain through the custom-built application, a commercial application called uberVU, who's features allowed to collect similar data, was also used. The advantage of using the commercial application is that it allowed for a very quick data collection, compared to the lengthy process required by the custom application. This advantage was used in a later stage of the research, when data about the Oscar nominated movies was collected.

With the data collected, it was possible to look at different correlations, between the values of the sentiment index and the IMDb rating, between the value of the box office and the same index and so on. The strongest correlation that came out of the case study was between the sentiment index calculated based on the data collected via the app and analyzed through the Tweet Sentiments (column Calculated using TS data in Table 2 and Isent_TSC in Figure 1) and the IMDb rating (Iimbd in Figure 1), the value of which is 0.614. The graphical representation of the data used for calculating the correlation is given in Figure 1.
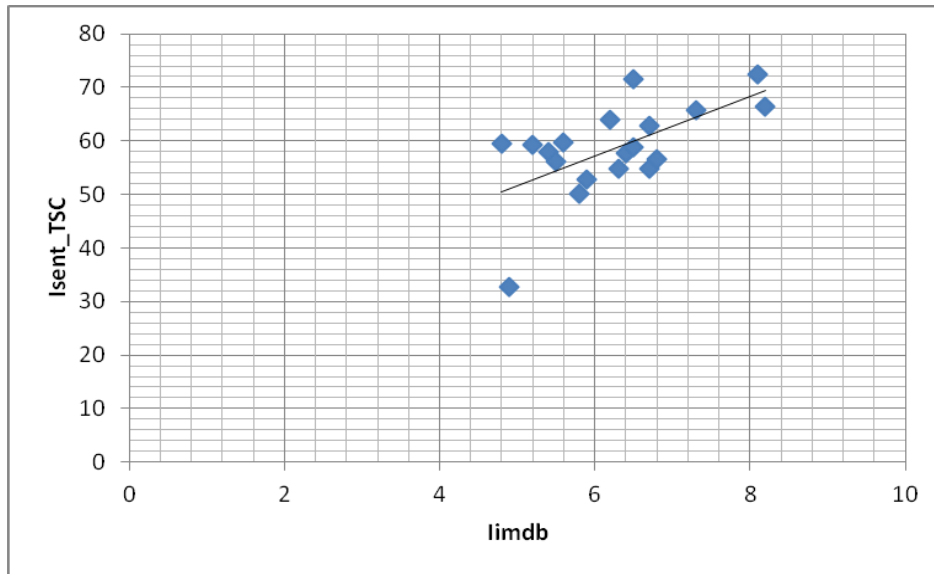
Fig. 1. Values of the IMDb rating and Isent index used for calculating the correlation

The strongest correlation between the sentiment index and the box office value was obtained when using the sentiment index computed based on the uberVU (Isent_UC in Figure 2) and the box office value (ValBO in Figure 2). The correlation gives a value of 0.607.
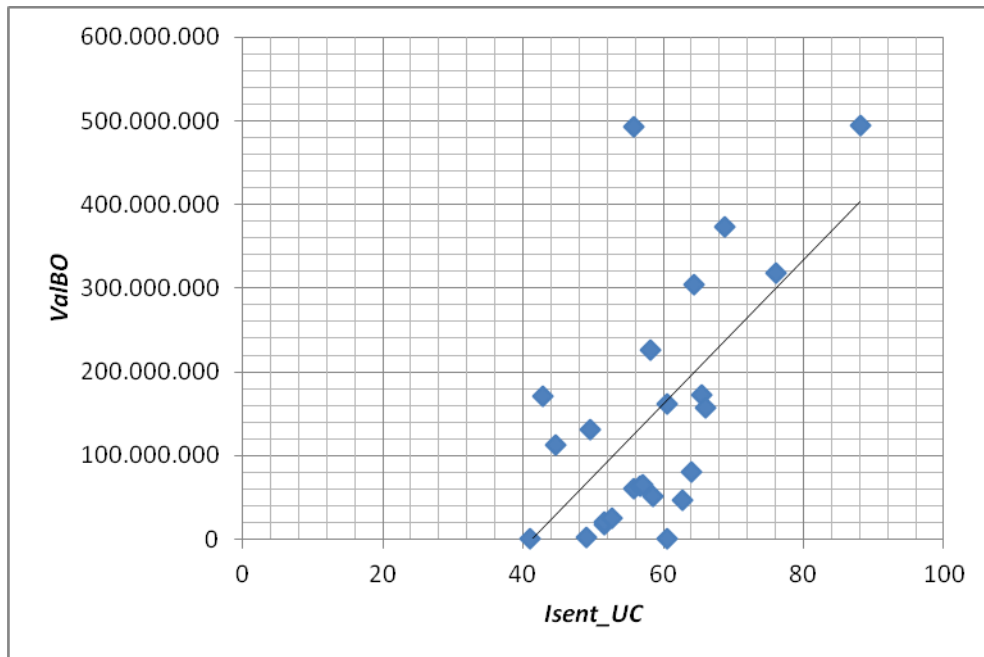
Fig. 2. Values of the Box office and Isent index, based on uberVU data, used for calculating the correlation between the two

After calculating this correlation and other possible combinations, the conclusion is that there is a definite correlation between the sentiment index and the IMDb rate and also the box-office values. Based on the values it was possible to identify a linear equation that defines how these values relate one to the other. Many forms of relation were computed, including exponential, logarithmic, polynomial and from all the linear relation seem to describe best the data.
When starting from these equations, it was attempted to predict the IMDb rate for movies outside the study the result were not very

accurate, showing that while there is a definite link, as expected, it is overly simplistic to use just the sentiment index in making predictions of exact values like the IMDb rate. The case study only scratches the surface and it points in a very interesting research direction.

6 Predicting the Oscars

While doing the research a new and exciting opportunity for testing the power of social networks for predicting success showed up. On February the 27th 2011 the 83rd Academy Awards ceremony took place. It occurred to

the authors that it was a perfect situation to use the experience gathered while doing the previously presented case study and build and do an actual prediction of an event that is widely covered.

The authors attempt at predicting the winners is not singular. Google attempted to predict the winners by looking at the Google Trends data. They claim that this data accurately predicted the winners in the previous 3 years. In 2011, however, the story was different. The most searched movie was "Black Swan", yet the winner was "The King's Speech". This is shown in Figure 3.
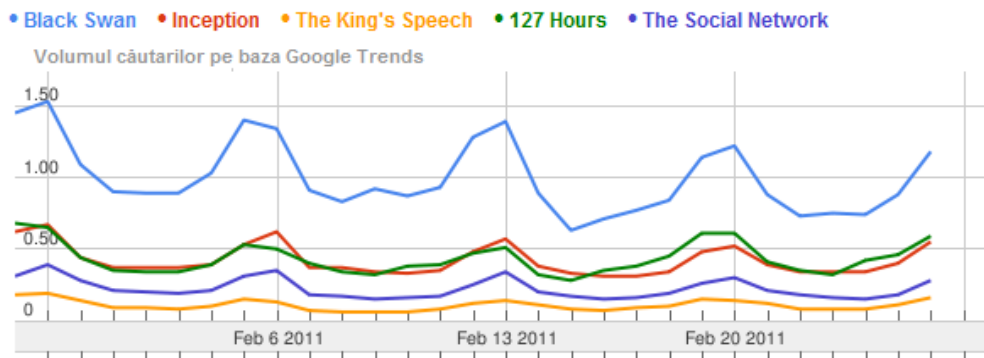

Fig. 3. Number of mentions of the top nominated movies, according to Google Trends

For best actor, Google put James Franco in the leading spot, while Colin Firth was the one that took home the trophy. Google took a quantitative approach. They looked only at the amount of searches.

Same approach was taken by Yahoo. A company called Webtrends did a poll and so did Yahoo. And another company called Elifemaps used the chatter on Twitter for their predictions. From all these attempts only the Yahoo poll, whose inputs were the opinions of people, managed to have some accurate

predictions. All methods that were based solely on the quantity of messages or searches about a certain movie were unsuccessful.

The approach used for the case study from this paper had as its foundation the assumptions that sentiments are the most important, the same used as in the previous example. In the same time, the authors collected their own quantitative data, using the same uber-VU tool. The values obtain for best movie nominees are summarized in Table 3.

Table 3. Number of mentions, data collected with uberVU, for the best movie nominees

Movie	...	16-Feb	17-Feb	18-Feb	19-Feb	20-Feb	21-Feb	22-Feb	23-Feb	24-Feb	25-Feb	26-Feb	27-Feb
The King's Speech		3422	4137	6676	7453	6802	7303	7292	6839	6908	7815	9693	17917
Back Swan		2517	8145	18748	22692	22811	23025	20951	18966	17545	18869	21250	36016
The Social Network		6035	5460	5706	6780	7920	8257	7267	6216	5960	7426	9038	15769
127 Hours		4104	4580	5517	6502	6486	5517	4751	4641	5866	8140	10832	12579
Inception					53	94	5207	10362	10230	10600	11831	12543	19085
TrueGrit		6772	5464	2215	606	3451	6480	6191	5701	6608	8191	8057	10120

Plotted on a graph, the data is similar to the
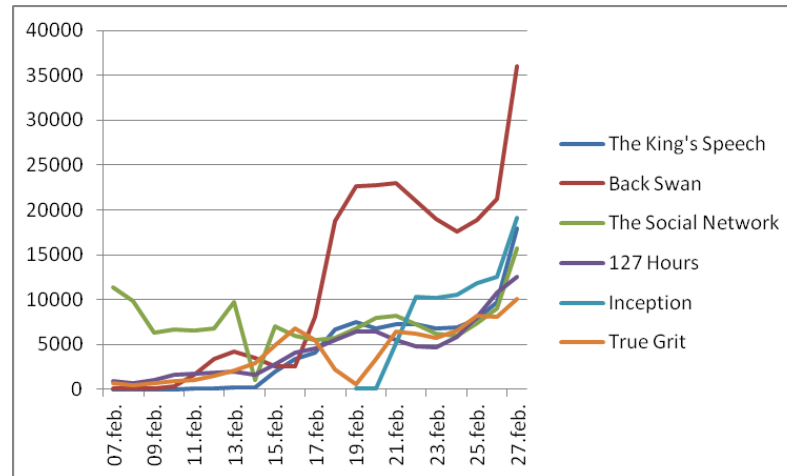Google Trends collected data (Figure 4).

Fig. 4. Number of mentions of the top nominated movies, according to uberVU data

Data for the 2 weeks before the Oscars was gathered about the movies with most chances at winning an Oscar and for the actors and actresses nominated for a leading role. Based on the data sentiment indexes were calculated by using the same methods as in the previous case study. Only these three categories were chosen because many of the movies and actors were nominated in other categories too, for example for best sound, makeup, film editing. Because of the difficulties and inaccuracies of differentiating on all these categories, only the overall and most general

were chosen. From the best movie category only 6 of the ten were actually included in the study, the ones that were considered to have most chances of winning. In the same time the number of mentions was also collected, for comparison with the results of the sentiment analysis and with the result provided by Google trends. The results of the sentiment index was accurate in two of the three cases, the indexes predicting correctly the winner for best picture (Table 4) and for best male actor in a leading role (Table 5).

Table 4. Sentiment indexes for movie nominated for best picture

Movie \ Date	20 Feb	21 Feb	22 Feb	23 Feb	24 Feb	25 Feb	26 Feb	27 Feb	Average (excl. 27)
The King's Speech	76.50	77.55	78.10	77.05	76.65	75.35	77.15	91.95	**76.91**
Back Swan	67.45	68.40	68.40	69.15	69.05	71.90	72.10	81.10	69.49
The Social Network	75.80	76.45	77.40	77.30	76.45	75.95	75.80	90.55	76.45
127 Hours	64.70	66.75	66.70	66.95	63.85	56.45	55.90	70.00	63.04
Inception	79.10	66.85	67.25	68.45	67.70	67.30	66.20	78.55	68.98
True Grit	76.25	76.60	75.35	73.35	76.30	77.90	76.45	88.15	76.03

Table 5. Sentiment indexes for best actors in leading role nominees

Name	14-26 February 2011				27 February 2011			
	Positive	**Neutral**	**Negative**	**Isent**	**Positive**	**Neutral**	**Negative**	**Isent**
Javier Bardem	94	1.8	4.2	94.9	88.3	4.4	7.3	90.5
Jeff Bridges	85.3	10.8	3.9	90.7	84.7	5.8	9.5	87.6

Jesse Eisenberg	88.1	5.3	6.6	90.75	87.2	3.8	9	89.1
Colin Firth	94.7	3.8	1.5	**96.6**	94.1	4.3	1.6	**96.25**
James Franco	81	17.3	1.7	89.65	82.1	15.9	2	90.05

But this approach is not without its problems. In case of best actress in a leading role nominees, the award went to Natalie Portman, the winner, was actually last, based on the sentiment index (Table 6), presumably because of the difficult and unconventional role she played.

Table 6. Sentiment indexes for best actress in leading role nominees

	14-26 February 2011				27 February 2011			
Name	**Positive**	**Neutral**	**Negative**	**Isent**	**Positive**	**Neutral**	**Negative**	**Isent**
Annette Bening	92.8	0	7.2	92.8	98.4	0	1.6	98.4
Nicole Kidman	99.5	0	0.5	99.5	91.2	0	8.8	91.2
Jennifer Lawrence	96.4	0.7	2.9	96.75	94.7	0	5.3	94.7
Natalie Portman	79.1	19.2	1.7	**88.7**	88.5	10.8	0.7	93.9
Michelle Williams	99.1	0.1	0.8	99.15	99.3	0.2	0.5	99.4

In this case the quantitative approach was more accurate. This is seen in Figure 5, in which it is clear that the most talked about person, before the awards took place, was Natalie Portman.
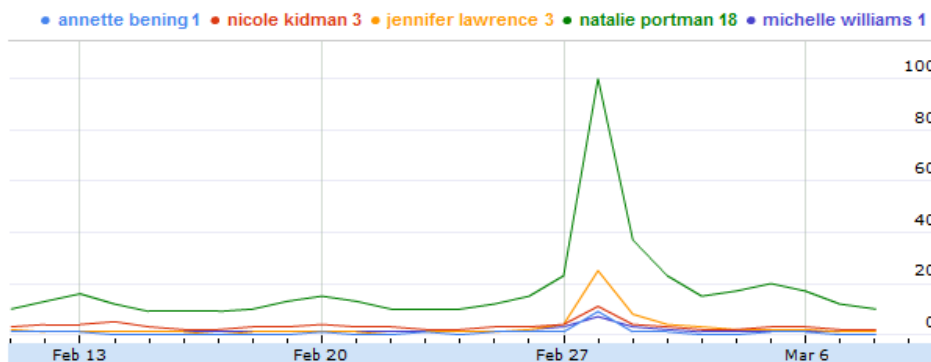
Fig. 5. Number of mentions of the nominated actress in leading role, according to Google Trends

From Table 5 and 6 and from Figure 5 a very interesting phenomenon can be seen. These tables and figure show data from before the event and immediately after the event, when the winners were known. In all cases, both the sentiment and the number of mentions grew very much for the winner. This is interesting to note as a proof of how people react and how the attitude and interest changes and shifts after a clear winner is nominated.

Overall the sentiment indexes calculated based on the discussion from social networks have been good and far better predictors of the winners of the Oscars in the main categories than the quantitative approaches.

7 Conclusions

Social media in general and social networks in particular, combined with the rest of the information available online through search engines and other services, represent a crawlable, indexable, searchable snapshot of the attitudes of society toward specific topics. This information can be used to calculate the mood of a group of people toward a specific subject and this, combined with quantitative

data and traditional techniques, can be used to improve prediction methods or to create new ones.

The studies reached the objective of testing the possibility of predicting how successful a service pr product is, using social media. We saw there is a correlation between the online discussions and some performance indicators. In the same time, it is difficult to establish a mathematical relation that can be used to build predictions.

There are some frontiers that can be further investigated and developed:

- **Semantic analysis**. At this moment this is a sensitive topic. The applications that were used for those two studies can be improved. At this stage they cannot treat the title of a movie as a single instance. For example, the title "From Paris with Love" can generate a very high sentiment index, having generally a positive connotation. During the analysis, such titles had to be omitted to avoid introducing errors;

- **Spam and unrelated messages**. The collected data contained enough irrelevant messages to make us think about the errors they can introduce. For the two case studies we chose films with titles containing words combinations that cannot be easily used in other contexts but this is not a fail proof method. A better filtering and spam removing solution would improve greatly the accuracy of the collected data. As at the previous point, advanced semantic analysis techniques can be very useful.

Also the use of metadata or sources that provided metadata would be a solution;

- **Narrow niches**. The movies and that were analyzed were picked from a very broad range of genres and budget. This proved not to be ideal, as some pictures with very high production and marketing budgets are bound to have a wider reach and a better penetration then low-budget indie films. Future research, no matter if it uses movies or any other form of goods and services need to focus on a narrow niche that with products that are as homogenous as possible.

- **Multiple sources of data**. We have used only one source of data and that is Twitter. Yet there are numerous other services, some of which have a wider reach, as is the case of Facebook or blogs. Using data from multiple sources will improve the quality of the results.

The case studies presented in the paper are only the tip of the iceberg. The available information represents a gold mine for marketers and economists and will lead to a deeper understanding of consumer behavior.

References
[1] J. O'Dell, *New Study Shows the Mobile Web Will Rule by 2015*, mashable.com, April 2010.
[2] Pingdom*, Internet 2010 in numbers*, published at royal.pingdom.com, 12 January 2011.
[3] N. Ganim Barnes, *The Fortune 500 and Social Media: A Longitudinal Study of Blogging, Twitter and Facebook Usage by America's Largest Companies*, Center for Marketing Research at the University of Massachusetts, 2011.
[4] C. Delcea, M. Dascălu, C. Ciurea, "A model for Improving Enterprise's Performance based on Collaborative E-learning,"*proceeding of ICEIS (International Conference on Enterprise Information Systems)*, pp. 5-12, Funchal, Madeira, Portugal, 2010.
[5] Nielsen Report, *Led by Facebook, Twitter, Global Time Spent on Social Media Sites up 82% Year over Year,* Nielsen.com, published 22 January 2010.

[6] Experian Simmons Report, *2010 Social Networking Report*, available at Experian.com, published 10 June 2010.

[7] U. S. Department of Defense (DOD), *Directive-Type Memorandum (DTM) 09-026 - Responsible and Effective Use of Internet-based Capabilities*, http://www.defense.gov/NEWS/DTM%2009-026.pdf, 25 February 2010.

[8] U. S. Army, Office Of The Chief Of Public Affairs, Online and Social Media Division, *U.S. Army Social Media Handbook*, dodlive.mil, January 2011.

[9] J. Bollen, H. Mao, X. J. Zeng, *Twitter mood predicts the stock market,* arXiv, 10101.3003v1, to be published in Journal of Computational Science, 14 October 2010.

[10] H. Varian, H. Choi*, Predicting the Present with Google Trends,* Google Research Paper, 2010.

[11] S. Asur, B .A. Huberman*,* "Predicting the Future with Social Media", *WI-IAT 2010 IEEE/WIC/ACM International Conference*, ISBN 978-1-4244-8482-9, published by IEEE, 1 November 2010, Toronto, Canada.

[12] R. Sharda, D. Delen, "Predicting box-office success of motion pictures with neural networks,"*Expert Systems with Applications*, vol. 30, no. 2, ISSN:0957-4174, published Pergamon Press, Inc. Tarrytown, NY, USA, February 2006.

[13] W. Zhang, S. Skiena, "Improving Movie Gross Prediction Through News Analysis," WI-IAT '09 Proceedings, Volume 01, published by IEEE Computer Society Washington, DC, ISBN: 978-0-7695-3801-3, 2009.

[14] Google, Search trends: a clue to 2011 Oscar winners?, The Google Blog, 15 February 2011.

Liviu LICĂ is a web developer with the company BSC. He is a PhD candidate since October 2008 at the Economic Informatics Department within the Academy of Economic Studies (ASE), Bucharest, Romania. He holds a Master in Science in the field of Project Management, from ASE, Bucharest. He graduated the Faculty of Economic Cybernetics, Statistics and Informatics, the Economic Informatics specialty, in March 2006 and the Faculty of Marketing in August 2006 at the Academy of Economic Studies, Bucharest. His main research areas are: the Internet;the influence of the Internet on society; web development; marketing; social media; social networks; promotion mechanisms.

Mihaela TUTĂ is PhD candidate since October 2010, at Marketing Department within Academy of Economic Studies, Bucharest. She graduated the Faculty of Cybernetics, Statistics and Informatics (being specialized in Economic Informatics) in March 2006 and the Faculty of Marketing (Marketing Policies specialty) in August 2006. She obtained a Master Degree in "Online Marketing", in March 2008, at ASE Bucharest and is affiliate member of Chartered Management Institute, United Kingdom, since March 2009.Since July 2006 she has been working for Oracle, within Advanced Customer Support (ACS), having the chance to combine her technical knowledge with her business knowledge. Main research areas for her are: online marketing, the Internet, CRM applications, Business Intelligence solutions and their economic impact.