NewsCATS: A News Categorization And Trading System

Marc-André Mittermayer
Glenfield Fund, c/o Swiss Capital Group
Talacker 41, CH-8039 Zurich, Switzerland
mittermayer@swisscap.com

Gerhard F. Knolmayer
Institute of Information Systems, Univ. of Bern
Engehaldenstr. 8, CH-3012 Bern, Switzerland
gerhard.knolmayer@iwi.unibe.ch

Abstract

NewsCATS is an Automated Text Categorization (ATC) prototype using a hand-made thesaurus to forecast intraday stock price trends from information contained in press releases. Due to a unique labeling approach and by carefully selecting the appropriate training data NewsCATS achieves a performance which is clearly superior to other ATC prototypes used for stock price trend forecasting. In this paper we describe the architecture, training, and testing of NewsCATS as well as the results of an extensive robustness analysis.

1. Introduction

Several prototypes for predicting short-term market reactions to news based on Automated Text Categorization (ATC) techniques have been developed. Wüthrich et al. were pioneers in building such a prototype almost ten years ago [1][2][23]. In the meantime, several other prototypes were developed and described in Lavrenko et al. [8][9][14], Thomas et al. [18][19][21][22], Elkan/Gidófalvi [5][6], Peramunetilleke/Wong [15], Fung/Lam/Yu [3] [4], Schulz/Spiliopoulou/Winkler [16] [20], and Mittermayer [11][12]. A detailed comparison of these eight prototypes and their performance data is given in [13].

Many of these prototypes use 3-category models (distinguishing between good, bad, and neutral news) and Naïve Bayes approaches as classifiers. In some of the prototypes the features are selected manually, which marks a substantial difference to classical ATC research where the feature selection is usually completely automated. On the other hand, labeling of training data is done automatically in almost all prototypes which again differs from classical ATC where training data is typically pre-categorized manually by domain experts.

Despite the truly interesting concepts of these prototypes some deficiencies exist. Most of them use daily closing data to measure the impact of news on the price of the security. However, it is highly questionable whether one is able to capture market reactions following the publication of news with an hourly or even daily resolution since financial markets are assumed to react very fast. Only three of the eight prototypes use a temporal granularity lower than one hour and it is interesting to observe that this does not necessarily result in better performance [13].

None of the prototypes attempting to forecast stock price or equity index trends distinguishes between the first publication of news and related follow-up statements. Also no effort is made to differentiate between news types that may be capable of moving market prices and others. As a consequence, the training data provided to the learning algorithms show a high degree of noise and it is very likely that this results in a sub-optimal pattern detection process.

Like most of the previously developed prototypes NewsCATS (News Categorization And Trading System) aims to forecast stock price trends. A preliminary version of NewsCATS was sketched in [11]. Based on the comparison described in [13] we significantly redesigned the system and tested it under more realistic and broader assumptions. When using the term NewsCATS in this paper we refer to the redesigned system.

2. Architecture of NewsCATS

NewsCATS is organized in three engines: The *Document Preprocessing Engine* automatically preprocesses incoming press releases. The *Categorization Engine* sorts the press releases into different categories and, finally, the *Trading Engine* triggers trading recommendations for the corresponding security.

The first two engines are trained in the learning phase by analyzing a set of press releases. The Document Preprocessing Engine creates the feature list with the bag-of-words method using either trivial functions like Collection Term Frequency (CTF), Inverse Document Frequency (IDF), and CTFxIDF or ATC-specific functions like Chi-squared (CHI), Information Gain (IG), and Odd's Ratio (OR) [17].

In addition to this automated procedure we created a handcrafted thesaurus containing features that were assumed to drive stock prices. Among the features in the thesaurus are single words (e.g., up, down), phrases (like 'formal investigation', 'sales climb'), and tuples of words/phrases (e.g., reduction NEAR 'financial guidance', approve NEAR 'share buyback'). Features from the thesaurus are forced into the final feature set, partially over-

Proceedings of the Sixth International Conference on Data Mining (ICDM'06)
0-7695-2701-9/06 $20.00 © 2006 IEEE

riding the results of the bag-of-words method. This enforcement happens if texts of the training data contain a feature occurring in the thesaurus but not in the current feature list. The thesaurus used is available online at http://www.ie.iwi.unibe.ch/NewsCATS_Thesaurus.html.

Finally, the Document Preprocessing Engine maps the press releases used for training into vectors; their elements represent the values obtained by the functions Within-Document Frequency (WDF), IDF, WDFxIDF, or Boolification [10]. Still in the training phase, the Categorization Engine receives the vectors and learns classifiers by applying algorithms like Rocchio, k Nearest Neighbors (kNN), linear SVM (lSVM) or non-linear SVM (nlSVM) with Gauss, sigmoid, or polynomial kernel [17].

The classifiers are trained with respect to a 3-category model which distinguishes good, bad, and neutral press releases. Such an approach is also used by most other prototypes. A press release is labeled as good (bad) if the stock price of the issuing company increases (decreases) significantly in the minutes after the news is published; press releases not leading to significant price moves are labeled as neutral. The labeling process is described in more detail in Section 3.3.

In the operational phase the Document Preprocessing Engine converts newly incoming press releases (or test documents) into a structured format using the feature list determined during training. Depending on the results provided by the Categorization Engine (i.e. the category assigned to the press release), the Trading Engine produces trading signals that may be executed via a Direct Market Access system.

3. The Data and Its Preparation

3.1. Price Data

Price data used during training and testing of NewsCATS comprise of transaction prices in 15-second intervals for S&P 500 stocks from April 1 to December 31, 2002. These stocks were chosen because correctly predicted price trends can only be converted into money if the underlying security shows enough trading volume. Transaction prices were extracted from the NYSE Trade and Quote Database. The dataset includes all transactions performed during the market hours at all U.S. marketplaces. In addition, for NASDAQ stocks also pre-market and after-hours transactions are provided.

As mentioned above, four transactions per minute (every 15 seconds) are allowed in the performance simulations. The price of a certain stock at time t was approximated by using the price of the last transaction recorded with time stamp t. If no transaction happened at this time, the last transaction before t was taken as proxy.

3.2. Press Releases

NewsCATS only considers company-specific news articles that have a high probability of being distributed simultaneously to market participants (and, thus, are new to the market). These publications are enforced by regulations like the Securities and Exchange Act of 1934 and the Regulation FD (Fair Disclosure), both requiring the simultaneous disclosure of certain material nonpublic information. Tests show that such news articles (also known as "press releases") have far more impact on stock prices than other news that often only repeat or comment the basic news contained in press releases ("editorial news" from, e.g., Reuters or Dow Jones) [12].

Usually companies do not publish press releases directly but via media partners to ensure compliance with legal requirements. In the U.S., PRNewswire and Businesswire are the dominating intermediaries in distributing such press releases. Together they control about 99% of the market, each managing roughly half of the press releases.

Generously, PRNewswire provided us with its archive of press releases covering the time period from April 1, 2002 to December 31, 2002. The roughly 18,000 press releases issued by S&P 500 companies were delivered in news industry text format (NITF) containing meta information about the companies mentioned in the text (ticker symbols), the topic, publication date and time, etc. PRNewswire distinguished 46 topics like sales reports, earnings, venture capital but also, for instance, religion or women-related news.

We excluded press releases that

- contained more than one ticker symbol or
- were published before 9:30 am EST (8:30 am EST in the case of NASDAQ stocks because of pre-market trading) or after 4 pm EST (5 pm EST in the case of NASDAQ stocks because of after-hours trading) or
- were published on weekends or exchange holidays.

After this filtering process 9,128 press releases remained in the dataset. We then implemented a heuristic to separate press releases with topics that typically affect market prices from other press releases. The goal of this separation is to provide the learning algorithm with a dataset which is as noise-free as possible. The seven topic classes with the largest impact on prices were ‘Dividends’, ‘Earnings projections or forecasts’, ‘Financing agreements’, ‘Legal issues’, ‘Licensing/marketing agreements’, ‘Offerings’, and 'Sales reports' (for more details cf. [12]). These seven topic classes appear in the metadata section of 989 press releases; these were used to train and test NewsCATS.

3.3. Labeling

Most of the prototypes mentioned above apply a 3-category model for training and performance testing, respectively. In NewsCATS we follow [3][4] and allow for a fourth category 'UNCLEAR' in the labeling process. Press releases are assigned to this category if their influence on stock prices is unclear. A press release has an unclear impact if, after its publication, the stock is trading at a remarkably higher or lower price than before but reverses completely within 15 minutes. Such press releases are excluded from the training process. The rationale behind this design decision is to support the learning algorithm by eliminating ambiguous training examples.

The 989 selected press releases were labeled as follows: For each press release we calculated a series of 2-minute moving price averages after the publication of the corresponding press release. The first moving average (referred to as MA_1) ranged from 1 minute after the publication to 3 minutes after the publication, the second (MA_2) from 1 minute and 15 seconds to 3 minutes and 15 seconds, the third (MA_3) from 1 minute and 30 seconds to 3 minutes and 30 seconds, and so forth. The last moving average (MA_{49}) covered the period from 13 minutes to 15 minutes after publication. These 49 moving averages were put in relation to the average from 1 minute before to 1 minute after the publication of the press release, referred to as MA_0. We determined the smallest and the largest value out of the 49 moving averages (MA_{min} and MA_{max}) and calculated the continuously compounded stock price return between MA_0 and MA_{min}, referred to as CCR_{min}, as well as between MA_0 and MA_{max}, referred to as CCR_{max}, as $CCR_{min/max} = \ln (MA_{min/max} / MA_0)$.

A press release is labeled as good news if it contains at least one feature of the thesaurus and either

- the maximum gain during 15 minutes after publication CCR_{max} is large (> 3%) and the maximum loss CCR_{min} is small (< 3%) or
- both the maximum gain CCR_{max} and the maximum loss CCR_{min} exceed 3% and the maximum gain is more than twice the maximum loss.

Bad news is defined analogously. If neither the gain nor the loss exceeds 3%, the press release is labeled as neutral. All other cases appear as unclear. After this labeling, 83 press releases belonged to category 'GOOD', 42 to 'BAD', 504 to 'NEUTRAL', and 360 to 'UNCLEAR'.

As already mentioned, press releases assigned to the category 'UNCLEAR' are not used during training in order to avoid ambiguous signals to the learning algorithms. But, of course, these news articles were used during performance testing.

4. Performance Simulation

4.1. Initial Training of NewsCATS

As described in Section 2, both the Document Preprocessing Engine and the Categorization Engine are able to apply different techniques to fulfill their preprocessing and categorization tasks. In the default setting the feature selection was performed by using IDF as selection function. We limited the number of features to 15% of the number of documents used in the training phase; this is close to the average of different recommendations outlined in [10]. The documents were represented with WDFxIDF, the most popular function for this task. The default classifier used was the highly accepted lSVM for which the software package SVM*light* provided by Thorsten Joachims at http://svmlight.joachims.org/ was used.

The default setting may look somehow arbitrary. Modification of this setting and the resulting effects are discussed in Section 4.3.

4.2. Testing of NewsCATS

A 10-fold cross validation was performed to test the performance of NewsCATS. As a consequence, all documents were available for testing. Due to integral requirements, 3 documents from 'GOOD', 2 from 'BAD', and 4 from 'NEUTRAL' were excluded randomly, leaving a total of 980 documents. In each of the 10 runs 90% of the 620 press releases assigned to the categories 'BAD', 'GOOD', and 'NEUTRAL' were used for training and 10% for testing. To each of the 10 test sets we added 10% of the 360 press releases from the category 'UNCLEAR' in order to use all press releases for testing. In each test run the 98 press releases available for testing were processed in a batch job. For each press release falling in category 'GOOD' a long roundtrip (RT) was simulated with the underlying stock (i.e. the stock was bought and sold later) and for each press release in the category 'BAD' a short roundtrip was performed (i.e. the stock was sold short and bought back later).

The positions were opened at the effective price 30 seconds after publication of the press release (in order to leave NewsCATS enough time to do its analysis). Some of the papers mentioned in the introduction ([5][8][9]) allow an early exit: Profits are taken immediately if the value of the investment rises 1% or more. However, in these simulations no stop loss rules as components of early exit strategies were applied. For this study we defined a two-sided early exit strategy (w/ EE) using a more conservative threshold of 0.5% for gains and triggering stop loss transactions if the position falls to or below -2%. If we do not accept an early exit (w/o EE), the positions were closed 15 minutes after opening them.

Several statistics can be used to describe the performance of a certain parameter setting. In the following we use

- the harmonic mean F_1 [in %] of macro-averaged precision and recall,
- the overall accuracy α [in %], i.e. the percentage of correct predictions,
- the average return *per* roundtrip [in %] achieved with and without the early exit option, and
- the total return [in %] resulting from *all* roundtrips achieved with and without the early exit option.

Table 1 shows the averages of the 10 runs of the performance simulation obtained with the default setting and compares them to the performance one would have achieved by simply guessing the short-term stock price trend (top row).

Table 1. Average performance achieved with News-CATS' default setting vs. a random trader

	F_1	α	Profit / RT		Total Profit		#RT
			w/o EE	w/ EE	w/o EE	w/ EE	
Random	33	33	0.00	0.00	0	0	653
Default	66	82	0.22	0.27	72	89	329

NewsCATS performs significantly better than a trader who applies a 3-category model and assigns press releases randomly according to a uniform distribution. With an F_1 of 66% the accuracy of the categorization performed by the Categorization Engine looks, at the first glance, worse than the one achieved in simple categorizations of document collections like Reuters-22173/21578, Reuters RCV1 or OHSUMED. However, the forecasting prototypes referenced above do not give F_1-values but report α values. NewsCATS categorizes α = 82% of the press releases correctly whereas a random trader would achieve only 33%. None of the prototypes obtains an overall accuracy value α > 50%. From this point of view, α = 82% appears as an exceptionally good result.

Without early exit NewsCATS achieves a profit per roundtrip of 0.22%. This is similar to the best result reported in earlier studies (0.23% in [9]). However, the result in [9] is achieved by allowing early profit taking. If early exits are allowed in NewsCATS, the result is boosted to 0.27%. We assume that the careful selection of the input data, the application of the noise-reducing heuristic described in Section 3.2, and the unique labeling approach have contributed to this good result. The rationale behind this assumption is that without these steps the performance was slightly lower than the one reported in [9] (for more details cf. [12]).

4.3. Robustness Analysis

The engines implemented in NewsCATS can be parameterized in a number of ways. We conducted a robustness analysis for each of the adjustable parameters of the default setting (π_1: feature selection function, π_2: size of feature set, π_3: document representation function, and π_4: classifier), leaving the other parameters unchanged. Varying the parameter setting may give insights how the performance values change (compared to a certain benchmark). In the analysis of parameter π_i (i>1) we used the best setting for $\pi_1, \ldots, \pi_{i-1}$ as benchmark. Using again the seven statistics suggested in Section 4.2 to describe the performance of a certain setting, we had to decide which statistic (or which combination of statistics) determines whether or not a setting delivers better results than another. While having an eye on all statistics we decided to leave the final call to the F_1-values.

4.3.1. Varying the Feature Selection Function (π_1). The default setting for feature selection was IDF. Table 2 shows the performance results obtained with other feature selection functions. The top row contains the result obtained with the default setting, as given in Table 1.

Table 2. Average performance achieved with different feature selection functions

	F_1	α	Profit / RT		Total Profit		#RT
			w/o EE	w/ EE	w/o EE	w/ EE	
IDF	66	82	0.22	0.27	72	89	329
CTFxIDF	66	82	0.23	0.27	76	89	331
CTF	69	83	0.21	0.28	70	94	335
CHI	64	78	0.11	0.10	42	38	383
IG	62	67	0.13	0.08	73	45	559
OR	67	82	0.25	0.28	83	93	332

Table 2 shows that the function yielding the best result for F_1 is CTF (gray shaded), which also increases profit per roundtrip and total profit when an early exit is allowed. It is somewhat puzzling that the ATC-specific functions, often regarded as being superior to the trivial functions CTF, IDF, and CTFxIDF, end up well below the results achieved with CTF.

4.3.2. Varying the Size of the Feature Set (π_2). The size of the feature set has been set by default to 15% of the number of documents. Table 3 depicts the performance of NewsCATS if the size of the feature set is varied between 5 and 25%. The top row repeats the best result for π_1.

Table 3. Average performance achieved with different sizes of the feature set

	F_1	α	Profit / RT		Total Profit		#RT
			w/o EE	w/ EE	w/o EE	w/ EE	
15%	69	83	0.21	0.28	70	94	335
5%	68	83	0.24	0.33	64	88	267
10%	68	83	0.27	0.29	87	94	324
20%	68	82	0.19	0.27	64	91	336
25%	68	83	0.20	0.28	66	92	329

Changing the size of the feature set does not improve the performance of NewsCATS: Both accuracy measures (F_1 and α) remain unchanged or deteriorate. With the early exit option the profit per roundtrip is largest when the size of the feature set is reduced to 5%. However, the increase in the profit per roundtrip comes along with a reduction in the number of roundtrips, leaving the user with a total return of 88%, which is below the 94% in the top row.

4.3.3. Varying the Document Representation (π_3). By default the press releases have been represented by using WDFxIDF. Table 4 gives an overview of the performances achieved with NewsCATS if the document vectors are determined by Boolification, WDF or IDF. The vectors have been cosine normalized in order to comply with the input requirements.

Table 4. Average performance achieved with different document representation techniques

	F_1	α	Profit / RT		Total Profit		#RT
			w/o EE	w/ EE	w/o EE	w/ EE	
WDFxIDF	69	83	0.21	0.28	70	94	335
Boolean	61	84	-0.01	-0.05	-3	-13	255
WDF	61	84	0.03	-0.03	8	-8	259
IDF	67	82	0.19	0.23	66	80	346

Varying the document representation technique also does not improve the performance of NewsCATS. In contrast, with WDF and the Boolean representation significant performance deteriorations in terms of F_1 occur. This deterioration even leads to the somehow puzzling fact that the profit per roundtrip falls below zero.

4.3.4. Varying the Classifier (π_4). The last parameter varied is the learning/categorization algorithm. By default the Categorization Engine is trained with a lSVM. Alternatively, NewsCATS may be trained with Rocchio, kNN, and three different nlSVM (g: Gauss, s: sigmoid, and p: polynomial kernel). The kNN was applied with k=10 and the nlSVM with a polynomial kernel was trained with a third-order polynom. Table 5 shows the resulting performances.

Table 5. Average performance achieved with different classifiers

	F_1	α	Profit / RT		Total Profit		#RT
			w/o EE	W/ EE	w/o EE	w/ EE	
lSVM	69	83	0.21	0.28	70	94	335
Rocchio	62	80	0.09	0.14	31	48	344
kNN	64	79	0.16	0.02	62	8	389
nlSVM (g)	63	82	0.13	0.17	41	54	315
nlSVM (p)	69	83	0.27	0.29	91	98	337
nlSVM (s)	68	82	0.27	0.29	89	95	329

These results are in line with others reported in the ATC literature. Typically, Rocchio performs worst and so it does in NewsCATS. It is somewhat surprising that the nlSVM with a Gauss kernel (in terms of F_1) performed worse than the kNN which is usually inferior to the nlSVM (cf. [7][24]). The best result is achieved by applying the nlSVM with a polynomial kernel (gray shaded row). The difference between the F_1 values obtained with lSVM is very small (68.9% vs. 68.7%). However, the total profit values rise quite remarkably.

5. Discussion

NewsCATS is an ATC prototype using a hand-made thesaurus to forecast intraday stock price trends from information contained in press releases. By carefully selecting the appropriate training data and due to a unique approach to label them, NewsCATS achieves a performance which is clearly superior to other ATC prototypes used for stock price trend forecasting. The best result achieved so far was 0.23% per roundtrip [9]. NewsCATS outperforms this result by more than 25%: The best performance achieved with the "final" setting is 0.29% (Table 5). Moreover, the robustness analysis described in Section 4.3 shows that the performance results are fairly robust.

However, also the 0.29% must be regarded as a rather moderate financial result since neither the prototypes mentioned in this paper nor NewsCATS consider costs in their performance simulations. But costs can be quite significant because transaction costs, the costs of immediate execution (the bid/ask spread), and costs related with market illiquidity (e.g., the limited volume at the bid/ask prices) may exist. Since most of the previous prototypes achieve a gross profit between 0.1% and 0.2% per round-trip it is rather likely that net of all costs these prototypes yield about a zero return, except for institutional investors with favorable fee agreements.

In this paper we did not take these types of costs into account because we wanted to compare previous results with the performance obtained by NewsCATS. However, the cost issue will be addressed in a forthcoming paper describing in-depth performance simulations with NewsCATS under more realistic trading assumptions.

6. References

[1] V. Cho, *Knowledge Discovery from Distributed and Textual Data*, Ph.D. Thesis, Hong Kong Univ. of Science and Technology, Hong Kong, 1999.

[2] V. Cho, B. Wüthrich, and J. Zhang, "Text Processing for Classification", *Journal of Computational Intelligence in Finance*, Vol. 7, No. 2, 1999, pp. 6-22.

[3] G. P. C. Fung, J. X. Yu, and W. Lam, "News Sensitive Stock Trend Prediction", *Proceedings of the 6th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Taipei, May 6-8, 2002, pp. 481-493.

[4] G. P. C. Fung, J. X. Yu, W. Lam, "Stock Prediction: Integrating Text Mining Approach Using Real-time News", *Proceedings of the IEEE International Conference on Computational Intelligence for Financial Engineering*, Hong Kong, Mar. 20-23, 2003, pp. 395-402.

[5] G. Gidófalvi, *Using News Articles to Predict Stock Price Movements*, Project Report, Univ. of California San Diego, San Diego, 2001.

[6] G. Gidófalvi and C. Elkan, *Using News Articles to Predict Stock Price Movements*, Technical Report, Univ. of California San Diego, San Diego, 2003.

[7] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features", *Proceedings of the 10th European Conference on Machine Learning (ECML 1998)*, Chemnitz, Apr. 21-24, 1998, pp. 137-142.

[8] V. Lavrenko, M. Schmill, D. Lawrie, P. Ogilvie, D. Jensen, and J. Allan, "Mining of Concurrent Text and Time Series", *Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Boston, Aug. 20-23, 2000, pp. 37-44.

[9] V. Lavrenko, M. Schmill, D. Lawrie, P. Ogilvie, D. Jensen, and J. Allan, "Language Models for Financial News Recommendation", *Proceedings of the 9th International Conference on Information and Knowledge Management (CIKM 2000)*, Washington, Nov. 6-11, 2000, pp. 389-396.

[10] D. D. Lewis, *Representation and Learning in Information Retrieval*, Ph.D. Thesis, Univ. of Massachusetts Amherst, Amherst, 1992.

[11] M.-A. Mittermayer, "Forecasting Intraday Stock Price Trends with Text Mining Techniques", *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, Big Island, January 5-8, 2004, p. 64.

[12] M.-A. Mittermayer, *Einsatz von Text Mining zur Prognose kurzfristiger Trends von Aktienkursen nach der Publikation von Unternehmensnachrichten*, Ph.D. Thesis, Univ. of Bern, Bern, 2005.

[13] M.-A. Mittermayer and G. Knolmayer, "Text Mining Systems for Predicting the Market Response to News: A Survey", Working Paper No. 184, Institute of Information Systems, Univ. of Bern, Bern, 2006.

[14] P. Ogilvie and M. Schmill, *Ænalyst - Electronic Analyst of Stock Behavior*, Project Proposal 791m, Univ. of Massachusetts Amherst, Amherst, 1999.

[15] D. Peramunetilleke and R. K. Wong, "Currency Exchange Rate Forecasting from News Headlines", *Proceedings of the 13th Australasian Database Conference*, Melbourne, Jan. 28 - Feb. 2, 2002, pp. 131-139.

[16] A. Schulz, M. Spiliopoulou, and K. Winkler, "Kursrelevanzprognose von Ad-hoc-Meldungen: Text Mining wider die Informationsüberlastung im Mobile Banking", in: W. Uhr, W. Esswein, and E. Schoop (eds.), *Wirtschaftsinformatik 2003*, Vol. 2, Physica, Heidelberg, 2003, pp. 181-200.

[17] F. Sebastiani, "Machine Learning in Automated Text Categorization", *ACM Comp. Surveys*, Vol. 34, No. 1, 2002, pp. 1-47.

[18] Y. Seo, J. A. Giampapa, and K. Sycara, *Text Classification for Intelligent Portfolio Management*, Technical Report, Carnegie Mellon Univ., Pittsburgh, 2002.

[19] Y. Seo, J. A. Giampapa, and K. Sycara, *Financial News Analysis for Intelligent Portfolio Management*, Technical Report, Carnegie Mellon Univ., Pittsburgh, 2004.

[20] M. Spiliopoulou, A. Schulz, and K. Winkler, "Text Mining an der Börse: Einfluss von Ad-hoc-Mitteilungen auf die Kursentwicklung", in: C. Becker and H. Redlich (eds.), *Data Mining und Statistik in Hochschule und Wirtschaft*, Shaker, Aachen, 2003, pp. 215-228.

[21] J. D. Thomas, *News and Trading Rules*, Ph.D. Thesis, Carnegie Mellon Univ., Pittsburgh, 2003.

[22] J. D. Thomas and K. Sycara, "Integrating Genetic Algorithms and Text Learning for Financial Prediction", *Proceedings of the GECCO-2000 Workshop on Data Mining with Evolutionary Algorithms*, Las Vegas, Jul. 8, 2000, pp. 72-75.

[23] B. Wüthrich, V. Cho, S. Leung, D. Permunetilleke, K. Sankaran, J. Zhang, and W. Lam, "Daily Prediction of Major Stock Indices from Textual WWW Data", *Proceedings 4th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, Aug. 27-31, 1998, pp. 364-368.

[24] Y. Yang and X. Liu, "A Re-examination of Text Categorization Methods", *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in: Information Retrieval (SIGIR 1999)*, Berkeley, Aug. 15-19, 1999, pp. 42-49.