Reading the Markets:
Forecasting Public Opinion of Political Candidates by News Analysis

Kevin Lerman
Dept. of Computer Science
Columbia University
New York, NY USA
klerman@cs.columbia.edu

Ari Gilder and **Mark Dredze**
Dept. of CIS
University of Pennsylvania
Philadelphia, PA USA
agilder@alumni.upenn.edu
mdredze@cis.upenn.edu

Fernando Pereira
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA USA
pereira@google.com

Abstract

Media reporting shapes public opinion which can in turn influence events, particularly in political elections, in which candidates both respond to and shape public perception of their campaigns. We use computational linguistics to automatically predict the impact of news on public perception of political candidates. Our system uses daily newspaper articles to predict shifts in public opinion as reflected in prediction markets. We discuss various types of features designed for this problem. The news system improves market prediction over baseline market systems.

1   Introduction

The mass media can affect world events by swaying public opinion, officials and decision makers. Financial investors who evaluate the economic performance of a company can be swayed by positive and negative perceptions about the company in the media, directly impacting its economic position. The same is true of politics, where a candidate's performance is impacted by media influenced public perception. Computational linguistics can discover such signals in the news. For example, Devitt and Ahmad (2007) gave a computable metric of polarity in financial news text consistent with human judgments. Koppel and Shtrimberg (2004) used a daily news analysis to predict financial market performance, though predictions could not be used for future investment decisions. Recently, a study conducted of the 2007 French presidential election showed a correlation between the frequency of a candidate's name in the news and electoral success (Véronis, 2007).

This work forecasts day-to-day changes in public perception of political candidates from daily news. Measuring daily public perception with polls is problematic since they are conducted by a variety of organizations at different intervals and are not easily comparable. Instead, we rely on daily measurements from prediction markets.

We present a computational system that uses both external linguistic information and internal market indicators to forecast public opinion measured by prediction markets. We use features from syntactic dependency parses of the news and a user-defined set of market entities. Successive news days are compared to determine the novel component of each day's news resulting in features for a machine learning system. A combination system uses this information as well as predictions from internal market forces to model prediction markets better than several baselines. Results show that news articles can be mined to predict changes in public opinion.

Opinion *forecasting* differs from that of opinion *analysis*, such as extracting opinions, evaluating sentiment, and extracting predictions (Kim and Hovy, 2007). Contrary to these tasks, our system receives objective news, not subjective opinions, and learns what events will impact public opinion. For example, "oil prices rose" is a fact but will likely shape opinions. This work analyzes news (cause) to predict future opinions (effect). This affects the structure of our task: we consider a time-series setting since we must use past data to predict future opinions, rather than analyzing opinions in batch across the whole dataset.

We begin with an introduction to prediction markets. Several methods for new feature extraction are explored as well as market history baselines. Systems are evaluated on prediction markets from the 2004 US Presidential election. We close with a discussion of related and future work.

2 Prediction Markets

Prediction markets, such as TradeSports and the Iowa Electronic Markets[1], provide a setting similar to financial markets wherein shares represent not companies or commodities, but an outcome of a sporting, financial or political event. For example, during the 2004 US Presidential election, one could purchase a share of "George W. Bush to win the 2004 US Presidential election" or "John Kerry to win the 2004 US Presidential election." A pay-out of $1 is awarded to winning shareholders at market's end, e.g. Bush wins the election. In the interim, price fluctuations driven by supply and demand indicate the perception of the event's likelihood, which indicates public opinion of an event. Several studies show the accuracy of prediction markets in predicting future events (Wolfers and Zitzewitz, 2004; Servan-Schreiber et al., 2004; Pennock et al., 2000), such as the success of upcoming movies (Jank and Foutz, 2007), political stock markets (Forsythe et al., 1999) and sports betting markets (Williams, 1999).

Market investors rely on daily news reports to dictate investment actions. If something positive happens for Bush (e.g. Saddam Hussein is captured), Bush will appear more likely to win, so demand increases for "Bush to win" shares, and the price rises. Likewise, if something negative for Bush occurs (e.g. casualties in Iraq increase), people will think he is less likely to win, sell their shares, and the price drops. Therefore, prediction markets can be seen as rapid response indicators of public mood concerning political candidates. Market-internal factors, such as general investor mood and market history, also affect price. For instance, a positive news story for a candidate may have less impact if investors dislike the candidate. Explaining market behavior requires modeling news information external to the market and internal trends to the market.

This work uses the 2004 US Presidential election markets from Iowa Electronic Markets. Each

market provides a daily average price, which indicates the overall market sentiment for a candidate on a given day. The goal of the prediction system is to predict the price direction for the next day (up or down) given all available information up to the current day: previous days' market pricing/volume information and the morning news. Market history represents information internal to the market: if an investor has no knowledge of external events, what is the most likely direction for the market? This information can capture general trends and volatility of the market. The daily news is the external information that influences the market. This provides information, independent of any internal market effects to which investors will respond. A learning system for each information source is developed and combined to explain market behavior. The following sections describe these systems.

3 External Information: News

Changes in market price are likely responses to current events reported in the news. Investors read the morning paper and act based on perceptions of events. Can a system with access to this same information make good investment decisions?

Our system operates in an iterative (online) fashion. On each day (round) the news for that day is used to construct a new instance. A logistic regression classifier is trained on all previous days and the resulting classifier predicts the price movement of the new instance. The system either profits or loses money according to this prediction. It then receives the actual price movement and labels the instance accordingly (up or down). This setting is straightforward; the difficulty is in choosing a good feature representation for the classifier. We now explore several representation techniques.

3.1 Bag-of-Words Features

The prediction task can be treated as a document classification problem, where the document is the day's news and the label is the direction of the market. Document classification systems typically rely on bag-of-words features, where each feature indicates the number of occurrences of a word in the document. The news for a given day is represented by a normalized unit length vector of counts, excluding common stop words and features that occur fewer than 20 times in our corpus.

3.2 News Focus Features

Simple bag-of-words features may not capture relevant news information. Public opinion is influenced by new events – a change in focus. The day after a debate, most papers may declare Bush the winner, yielding a rise in the price of a "Bush to win" share. However, while the debate may be discussed for several days after the event, public opinion of Bush will probably not continue to rise on old news. Changes in public opinion should reflect changes in daily news coverage. Instead of constructing features for a single day, they can represent differences between two days of news coverage, i.e. the novelty of the coverage. Given the counts of feature i on day t as c_i^t, where feature i may be the unigram "scandal," and the set of features on day t as $\mathbf{C}^t$, the fraction of news focus for each feature is $f_i^t = \frac{c_i^t}{|\mathbf{C}^t|}$. The news focus change (Δ) for feature i on day t is defined as,

$$\Delta f_i^t = \log\left(\frac{f_i^t}{\frac{1}{3}(f_i^{t-1} + f_i^{t-2} + f_i^{t-3})}\right) , \quad (1)$$

where the numerator is the focus of news on feature i today and the denominator is the average focus over the previous three days. The resulting value captures the change in focus on day t, where a value greater than 0 means increased focus and a value less than 0 decreased focus. Feature counts were smoothed by adding a constant (10).

3.3 Entity Features

As shown by Wiebe et al. (2005), it is important to know not only what is being said but about whom it is said. The term "victorious" by itself is meaningless when discussing an election – meaning comes from the subject. Similarly, the word "scandal" is bad for a candidate but good for the opponent. Subjects can often be determined by proximity. If the word "scandal" and Bush are mentioned in the same sentence, this is likely to be bad for Bush. A small set of entities relevant to a market can be defined a priori to give context to features. For example, the entities "Bush," "Kerry" and "Iraq" were known to be relevant before the general election. Kim and Hovy (2007) make a similar assumption.

News is filtered for sentences that mention exactly one of these entities. Such sentences are likely about that entity, and the extracted features are conjunctions of the word and the entity. For example, the sentence "Bush is facing another scandal" produces the feature "bush-scandal" instead of just "scandal." [2] Context disambiguation comes at a high cost: about 70% of all sentences do not contain any predefined entities and about 7% contain more than one entity. These likely relevant sentences are unfortunately discarded, although future work could reduce the number of discarded sentences using coreference resolution.

3.4 Dependency Features

While entity features are helpful they cannot process multiple entity sentences, nearly a quarter of the entity sentences. These sentences may be the most helpful since they indicate entity interactions. Consider the following three example sentences:

- Bush defeated Kerry in the debate.

- Kerry defeated Bush in the debate.

- Kerry, a senator from Massachusetts, defeated President Bush in last night's debate.

Obviously, the first two sentences have very different meanings for each candidate's campaign. However, representations considered so far do not differentiate between these sentences, nor would any heuristic using proximity to an entity. [3] Effective features rely on the proper identification of the subject and object of "defeated." Longer n-grams, which would be very sparse, would succeed for the first two sentences but not the third.

To capture these interactions, features were extracted from dependency parses of the news articles. Sentences were part of speech tagged (Toutanova et al., 2003), parsed with a dependency parser and labeled with grammatical function labels (McDonald et al., 2006). The resulting parses encode dependencies for each sentence, where word relationships are expressed as parent-child links. The parse for the third sentence above indicates that "Kerry" is the subject of "defeated," and "Bush" is the object. Features are extracted from parse trees containing the pre-defined entities (section 3.3), using the parent, grandparent, aunts,

---

[2] Other methods can identify the subject of sentiment expressions, but our text is objective news. Therefore, we employ this approximate method.

[3] Several failed heuristics were tried, such as associating each word to an entity within a fixed window in the sentence or the closer entity if two were in the window.

Feature	Good For
Kerry ← plan → the	Kerry
poll → showed → Bush	Bush
won → Kerry[4]	Kerry
agenda → 's → Bush	Kerry
Kerry ← spokesperson → campaign	Bush

Table 1: Simplified examples of features from the general election market. Arrows point from parent to child. Features also include the word's dependency relation labels and parts of speech.

nieces, children, and siblings of any instances of the pre-defined entities we observe. Features are conjoined indicators of the node's lexical entry, part of speech tag and dependency relation label. For aunts, nieces, and children, the common ancestor is used, and in the case of grandparent, the intervening parent is included. Each of these conjunctions includes the discovered entity and back-off features are included by removing some of the other information. Note that besides extracting more precise information from the news text, this handles sentences with multiple entities, since it associates parts of a sentence with different entities. In practice, we use this in conjunction with News Focus. Useful features from the general election market are in table 1. Note that they capture events and not opinions. For example, the last feature indicates that a statement by the Kerry campaign was good for Bush, possibly because Kerry was reacting to criticism.

4   Internal Information: Market History

News cannot explain all market trends. Momentum in the market, market inefficiencies, and slow news days can affect share price. A candidate who does well will likely continue to do well unless new events occur. Learning general market behavior can help explain these price movements.

For each day t, we create an instance using features for the price and volume at day $t - 1$ and the price and volume change between days $t - 1$ and $t - 2$. We train using a ridge regression [5] on

all previous days (labeled with their actual price movements) to forecast the movement for day t, which we convert into a binary value: up or down.

5   Combined System

Since both news and internal market information are important for modeling market behavior, each one cannot be evaluated in isolation. For example, a successful news system may learn to spot important events for a candidate, but cannot explain the price movements of a slow news day. A combination of the market history system and news features is needed to model the markets.

Expert algorithms for combining prediction systems have been well studied. However, experiments with the popular weighted majority algorithm (Littlestone and Warmuth, 1989) yielded poor performance since it attempts to learn the optimal balance between systems while our setting has rapidly shifting quality between few experts with little data for learning. Instead, a simple heuristic was used to select the best performing predictor on each day. We compare the 3-day prediction accuracy (measured in total earnings) for each system (news and market history) to determine the current best system. The use of a small window allows rapid change in systems. When neither system has a better 3-day accuracy the combined system will only predict if the two systems agree and abstain otherwise. This strategy measures how accurately a news system can account for price movements when non-news movements are accounted for by market history. The combined system improved over individual evaluations of each system on every market [6].

6   Evaluation

Daily pricing information was obtained from the Iowa Electronic Markets for the 2004 US Presidential election for six Democratic primary contenders (Clark, Clinton, Dean, Gephardt, Kerry and Lieberman) and two general election candidates (Bush and Kerry). Market length varied as some candidates entered the race later than others: the DNC markets for Clinton, Gephardt, Kerry, and Lieberman were each 332 days long, while Dean's was 130 days and Clark's 106. The general

---

[4]This feature matches phrases like "Kerry won [the debate]" and "[something] won Kerry [support]"

[5]This outperformed more sophisticated algorithms, including the logistic regression used earlier. This may be due to the fact that many market history features (e.g. previous price movements) are very similar in nature to the future price movements being predicted.

[6]This outperformed a single model built over all features, perhaps due to the differing natures of the feature types we used.

election market for Bush was 153 days long, while Kerry's was 142. [7] The price delta for each day was taken as the difference between the average price between the previous and current day. Market data also included the daily volume that was used as a market history feature. Entities selected for each market were the names of all candidates involved in the election and "Iraq."

News articles covering the election were obtained from Factiva[8], an online news archive run by Dow Jones. Since the system must make a prediction at the beginning of each day, only articles from daily newspapers released early in the morning were included. The corpus contained approximately 50 articles per day over a span of 3 months to almost a year, depending on the market. [9]

While most classification systems are evaluated by measuring their accuracy on cross-validation experiments, both the method and the metric are unsuitable to our task. A decision for a given day must be made with knowledge of only the previous days, ruling out cross validation. In fact, we observed improved results when the system was allowed access to future articles through cross-validation. Further, raw prediction accuracy is not a suitable metric for evaluation because it ignores the magnitude in price shifts each day. A system should be rewarded proportional to the significance of the day's market change.

To address these issues we used a chronological evaluation where systems were rewarded for correct predictions in proportion to the magnitude of that day's shift, i.e. the ability to profit from the market. This metric is analogous to weighted accuracy. On each day, the system is provided with all available morning news and market history from which an instance is created using one of the feature schemes described above. We then predict whether the market price will rise or fall and the system either earns or loses the price change for that day if it was right or wrong respectively. The system then learns the correct price movement and the process is repeated for the next day.[10] Sys-

Market		History	Baseline
DNC	Clark	20	13
	Clinton	38	-8
	Dean	23	24
	Gephardt	8	1
	Kerry	-6	6
	Lieberman	3	2
General	Kerry	2	15
	Bush	21	20
Average (% omniscience)		13.6	9.1

Table 2: Results using history features for prediction compared with a baseline system that invests according to the previous day's result.

tems that correctly forecast public opinions from the news will make more money. In economic terms, this is equivalent to buying or short-selling a single share of the market and then selling or covering the short at the end of the day.[11] Scores were normalized for comparison across markets using the maximum profit obtainable by an omniscient system that always predicts correctly.

Baseline systems for both news and market history are included. The news baseline follows the spirit of a study of the French presidential election (Véronis, 2007), which showed that candidate mentions correlate to electoral success. Attempts to follow this method directly – predicting market movement based on raw candidate mentions – did very poorly. Instead, we trained our learning system with features representing daily mention counts of each entity. For a market history baseline, we make a simple assumption about market behavior: the current market trend will continue, predict today's behavior for tomorrow.

There were too many features to learn in the short duration of the markets so only features that appeared at least 20 times were included, reducing bag-of-words features from 88.8k to 28.3k and parsing features from 1150k to 15.9k. A real world system could use online feature selection.

6.1 Results

First, we establish performance without news information by testing the market history system alone. Table 2 shows the profit of the history pre-

---

[7]The first 11 days of the Kerry general election market were removed due to strange price fluctuations in the data.

[8]http://www.factiva.com/

[9]While 50 articles may not seem like much, humans read far less text before making investment decisions.

[10]This scheme is called "online learning" for which a whole class of algorithms apply. We used batch algorithms since training happens only once per day.

[11]More complex investment schemes are possible than what has been described here. We choose a simple scheme to make the evaluation more transparent.
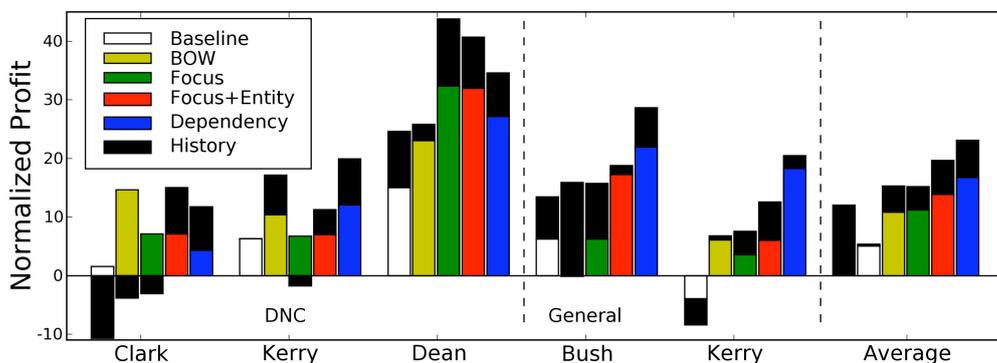
Figure 1: Results for the different news features and combined system across five markets. Bottom bars can be compared to evaluate news components and combined with the stacked black bars (history system) give combined performance. The average performance (far right) shows improved performance from each news system over the market history system.

diction and baseline systems. While learning beats the rule based system on average, both earn impressive profits considering that random trading would break even. These results corroborate the inefficient market observation of Pennock et al. (2000). Additionally, the general election markets sometimes both increased or decreased, an impossible result in an efficient zero-sum market.

During initial news evaluations with the combined system, the primary election markets did either very poorly or quite well. The news prediction component lost money for Clinton, Gephardt, and Lieberman while Clark, Dean and Kerry all made money. Readers familiar with the 2004 election will immediately see the difference between the groups. The first three candidates were minor contenders for the nomination and were not newsmakers. Hillary Clinton never even declared her candidacy. The average number of mentions per day for these candidates in our data was 20. In contrast, the second group were all major contenders for the nomination and an average mention of 94 in our data. Clearly, the news system can only do well when it observes news that effects the market. The system does well on both general election markets where the average candidate mention per day was 503. Since the Clinton, Gephardt and Lieberman campaigns were not newsworthy, they are omitted from the results.

Results for news based prediction systems are shown in figure 1. The figure shows the profit made from both news features (bottom bars) and

market history (top black bars) when evaluated as a combined system. Bottom bars can be compared to evaluate news systems and each is combined with its top bar to indicate total performance. Negative bars indicate negative earnings (i.e. weighted accuracy below 50%). Averages across all markets for the news systems and the market history system are shown on the right. In each market, the baseline news system makes a small profit, but the overall performance of the combined system is worse than the market history system alone, showing that the news baseline is ineffective. However, all news features improve over the market history system; news information helps to explain market behaviors. Additionally, each more advanced set of news features improves, with dependency features yielding the best system in a majority of markets. The dependency system was able to learn more complex interactions between words in news articles. As an example, the system learns that when Kerry is the subject of "accused" his price increases but decreased when he is the object. Similarly, when "Bush" is the subject of "plans" (i.e. Bush is making plans), his price increased. But when he appears as a modifier of the plural noun "plans" (comments about Bush policies), his price falls. Earning profit indicates that our systems were able to correctly forecast changes in public opinion from objective news text.

The combined system proved an effective way of modeling the market with both information sources. Figure 2 shows the profits of the depen-

dency news system, the market history system, and the combined system's profits and decision on two segments from the Kerry DNC market. In the first segment, the history system predicts a downward trend in the market (increasing profit) and the second segment shows the final days of the market, where Kerry was winning primaries and the news system correctly predicted a market increase.

Véronis (2007) observed a connection between electoral success and candidate mentions in news media. The average daily mentions in the general election was 520 for Bush (election winner) and 485 for Kerry. However, for the three major DNC candidates, Dean had 183, Clark 56 and Kerry (election winner) had the least at 43. Most Kerry articles occurred towards the end of the race when it was clear he would win, while early articles focused on the early leader Dean. Also, news activity did not indicate market movement direction; median candidate mentions for a positive market day was 210 and 192 for a negative day.

Dependency news system accuracy was correlated with news activity. On days when the news component was correct – although not always chosen – there were 226 median candidate mentions compared to 156 for incorrect days. Additionally, the system was more successful at predicting negative days. While days for which it was incorrect the market moved up or down equally, when it was correct and selected it predicted buy 42% of the time and sell 58%, indicating that the system better tracked negative news impacts.

7 Related Work

Many studies have examined the effects of news on financial markets. Koppel and Shtrimberg (2004) found a low correlation between news and the stock market, likely because of the extreme efficiency of the stock market (Gidófalvi, 2001). Two studies reported success but worked with a very small time granularity (10 minutes) (Lavrenko et al., 2000; Mittermayer and Knolmayer, 2006). It appears that neither system accounts for the time-series nature of news during learning, instead using cross-validation experiments which is unsuitable for evaluation of time-series data. Our own preliminary cross-validation experiments yielded much better results than chronological evaluation since the system trains using future information, and with much more training data than is actu-
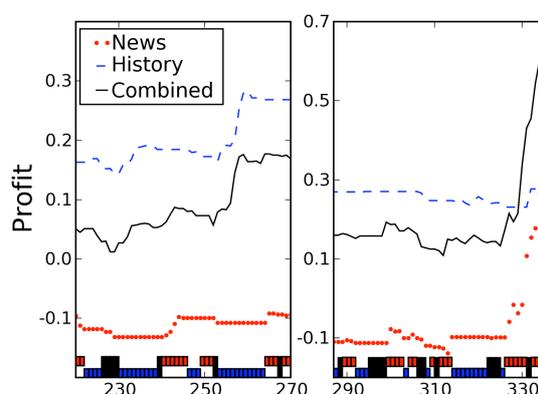

Figure 2: Two selections from the Kerry DNC market showing profits over time (days) for dependency news, history and combined systems. Each day's chosen system is indicated by the bottom stripe as red (upper) for news, blue (lower) for history, and black for ties.

ally available for most days. Recent work has examined prediction market behavior and underlying principles (Serrano-Padial, 2007). [12] Pennock et al. (2000) found that prediction markets are somewhat efficient and some have theorized that news could predict these markets, which we have confirmed (Debnath et al., 2003; Pennock et al., 2001; Servan-Schreiber et al., 2004).

Others have explored the concurrent modeling of text corpora and time series, such as using stock market data and language modeling to identify influential news stories (Lavrenko et al., 2000). Hurst and Nigam (2004) combined syntactic and semantic information for text polarity extraction.

Our task is related to but distinct from sentiment analysis, which focuses on judgments in opinions and, recently, predictions given by opinions. Specifically, Kim and Hovy (2007) identify which political candidate is predicted to win by an opinion posted on a message board and aggregate opinions to correctly predict an election result. While the domain and some techniques are similar to our own, we deal with fundamentally different problems. We do not consider opinions but instead analyze objective news to learn events that will impact opinions. Opinions express subjective statements about elections whereas news reports events. We

---

[12] For a sample of the literature on prediction markets, see the proceedings of the recent Prediction Market workshops (http://betforgood.com/events/pm2007/index.html).

use public opinion as a measure of an events impact. Additionally, they use generalized features similar to our own identification of entities by replacing (a larger set of) known entities with generalized terms. In contrast, we use syntactic structures to create generalized ngram features. Note that our features (table 1) do not indicate opinions in contrast to the Kim and Hovy features. Finally, Kim and Hovy had a batch setting to predict election winners while we have a time-series setting that tracked daily public opinion of candidates.

8 Conclusion and Future Work

We have presented a system for forecasting public opinion about political candidates using news media. Our results indicate that computational systems can process media reports and learn which events impact political candidates. Additionally, the system does better when the candidate appears more frequently and for negative events. A news source analysis could reveal which outlets most influence public opinion. A feature analysis could reveal which events trigger public reactions. While these results and analyses have significance for political analysis they could extend to other genres, such as financial markets. We have shown that feature extraction using syntactic parses can generalize typical bag-of-word features and improve performance, a non-trivial result as dependency parses contain significant errors and can limit the selection of words. Also, combining the internal market baseline with a news system improved performance, suggesting that forecasting future public opinions requires a combination of new information and continuing trends, neither of which can be captured by the other.

References

Debnath, S., D. M. Pennock, C. L. Giles, and S. Lawrence. 2003. Information incorporation in online in-game sports betting markets. In *Electronic Commerce*.

Devitt, Ann and Khurshid Ahmad. 2007. Sentiment polarity identification in financial news: A cohesion-based approach. In *Association for Computational Linguistics (ACL)*.

Forsythe, R., T.A. Rietz, , and T.W. Ross. 1999. Wishes, expectations, and actions: A survey on price formation in election stock markets. *Journal of Economic Behavior and Organization*, 39:83–110.

Gidófalvi, G. 2001. Using news articles to predict stock price movements. Technical report, Univ. of California San Diego, San Diego.

Hurst, Matthew and Kamal Nigam. 2004. Retrieving topical sentiments from online document collections. In *Document Recognition and Retrieval XI*.

Jank, Wolfgang and Natasha Foutz. 2007. Using virtual stock exchanges to forecast box-office revenue via functional shape analysis. In *The Prediction Markets Workshop at Electronic Commerce*.

Kim, Soo-Min and Eduard Hovy. 2007. Crystal: Analyzing predictive opinions on the web. In *Empirical Methods in Natural Language Processing (EMNLP)*.

Koppel, M. and I. Shtrimberg. 2004. Good news or bad news? let the market decide. In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.

Lavrenko, V., M. Schmill, D. Lawrie, P. Ogilvie, D. Jensen, and J. Allan. 2000. Mining of concurrent text and time series. In *KDD*.

Littlestone, Nick and Manfred K. Warmuth. 1989. The weighted majority algorithm. In *IEEE Symposium on Foundations of Computer Science*.

McDonald, R., K. Lerman, and F. Pereira. 2006. Multilingual dependency parsing with a two-stage discriminative parser. In *Conference on Natural Language Learning (CoNLL)*.

Mittermayer, M. and G. Knolmayer. 2006. NewsCATS: A news categorization and trading system. In *International Conference in Data Mining*.

Pennock, D. M., S. Lawrence, C. L. Giles, and F. A. Nielsen. 2000. The power of play: Efficiency and forecast accuracy in web market games. Technical Report 2000-168, NEC Research Institute.

Pennock, D. M., S. Lawrence, F. A. Nielsen, and C. L. Giles. 2001. Extracting collective probabilistic forecasts from web games. In *KDD*.

Serrano-Padial, Ricardo. 2007. Strategic foundations of prediction markets and the efficient markets hypothesis. In *The Prediction Markets Workshop at Electronic Commerce*.

Servan-Schreiber, E., J. Wolfers, D. M. Pennock, and B. Galebach. 2004. Prediction markets: Does money matter? *Electronic Markets*, 14.

Toutanova, K., D. Klein, C. Manning, and Y. Singer. 2003. Feature-rich part-of-speech tagging with a cyclic dependency network. In *HLT-NAACL*.

Véronis, Jean. 2007. La presse a fait mieux que les sondeurs. http://aixtal.blogspot.com/2007/04/2007-la-presse-fait-mieux-que-les.html.

Wiebe, Janyce, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *LREC*, 39:165–210.

Williams, L.V. 1999. Information efficiency in betting markets: A survey. *Bulletin of Economic Research*, 51:1–30.

Wolfers, J. and E. Zitzewitz. 2004. Prediction markets. *Journal of Economic Perspectives*, 18(2):107–126.