

Time-Sensitive Opinion Mining for Prediction

Wenting Tu, David Cheung, and Nikos Mamoulis

Department of Computer Science, The University of Hong Kong, Pokfulam, Hong Kong

E-mail: {wtu, dcheung, nikos}@cs.hku.hk

Complementary-material webpage: <http://sites.google.com/site/wenting4aaai/>

Abstract

Users commonly use Web 2.0 platforms to post their opinions and their predictions about future events (e.g., the movement of a stock). Therefore, opinion mining can be used as a tool for predicting future events. Previous work on opinion mining extracts from the text only the polarity of opinions as sentiment indicators. We observe that a typical opinion post also contains temporal references which can improve prediction. This short paper presents our preliminary work on extracting reference time tags and integrating them into an opinion mining model, in order to improve the accuracy of future event prediction. We conduct an experimental evaluation using a collection of microblogs posted by investors to demonstrate the effectiveness of our approach.

1 Introduction

By analyzing the *prediction opinions* of users, we can aggregate personal wisdom and different viewpoints into an objective prediction. This paper aims at extending the previous opinion mining framework (Dave et al. 2003) to derive a more accurate prediction model. Our work is motivated by the observation that within a post item by an individual, prediction opinions may refer to different times. An opinion may not necessarily refer to the future and therefore may not constitute a prediction. For example, in the sentence “Although the market was bullish today, tomorrow it may be changed to bearish”, the opinion “bullish” refers to “today” (past) and the opinion “bearish” refers to “tomorrow” (future). Obviously, the author actually predicts a bearish trend here. If we disregard temporal references in sentiment analysis of this sentence, there is a “bullish” and a “bearish” component in the content, and the result is amphibolous. On the other hand, by considering the reference time of the opinions, we can correctly identify the author’s meant prediction. In this work, we propose a model that extracts the *reference time* (RT) of prediction opinions posted by users. When we aggregate prediction opinions about a subject in order to derive a prediction, we use the RT information to weigh the influence of the opinions, by favoring those that refer to the future. This way, we achieve a better prediction accuracy, as we demonstrate experimentally, using a real-world dataset.

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

2 Our methodology

Our work extends the basic opinion mining framework to use reference time (RT) information in the formation of future predictions. For the extraction of RT, we design four features that discriminate prediction opinions with different RT tags. Opinion text around each sentiment indicator is translated into a four-dimensional vector. Then, using learning or rule-based methods, we can identify the RT value of the opinioned text segment. Finally, we construct a prediction model that uses the RT information.

Extraction of Reference Time

We design four features that can be used to characterize future and past RT semantics (Several real-world examples of opinioned text items and their corresponding RT features can be seen in Figure 1):

(i) **Temporal Expression (TE)**. Expressions in the text may refer to specific time. For an opinion, we find all terms it contains and can be characterized as temporal expressions. Then, we map each term into an absolute time (AT) and compare it to the current time (CT). If AT is after CT, the value of feature TE of the opinion is defined as 1; if AT is before CT, TE is -1 . If there is no temporal expression in the opinion, TE is set to 0. Moreover, if we cannot decide which one of AT and CT comes first (e.g., when AT is a time interval containing CT), the corresponding term does not contribute to TE.

(ii) **Tense Indicator (TI)**. Certain words or verb endings are tense indicators. Tense indicators have strong discriminative power for separating future and past semantics. We keep a dictionary where the value of each TI word indicating future tense (e.g., “will”, “to be”, etc.) is $+1$ and the values of those indicating past tense (e.g., “was”, past forms of verbs, etc.) are -1 . Then, for analyzing the reference time of an opinion, we find all TI words contained in its context. Finally, we use the sum of values of these TI words in the dictionary as the value of TI feature. If no TI word appears in the opinion’s text, the value of TI is set to 0.

(iii) **Prediction/Summary Indicator (PSI)**. Certain keywords have prediction or summary semantics such as “expect”, “look back”, etc. Prediction semantics typically refer to the future, while summary semantics refer to the past. We follow the same approach as in the case of the TI feature,

i.e., building a dictionary for PSI terms and aggregate positive (predictive) and negative (summative) occurrences of them to derive the value of PSI for the opinion text item.

(iv) **Specificity Level (SL)**. This feature is motivated by the observation that when people talk about things that already happened, they tend to use more details. For example, when people talk about the current or past price of a stock, they provide exact numbers. However, when people post predictions, they often give relative or abstract values (e.g., predict that the value will increase or provide value range). We define specificity clues for descriptions. For example, if a description refers to a market index in single digits (e.g., “2373 points”), we regard this as a clue for the specificity. As another example, if a description of an earthquake’s magnitude has one or two digits after the decimal point, we regard this as a specificity clue. The value of the SL feature is binary: -1 means that the opinion contains some specificity clues and 0 means that it does not contain any specific clues.

In order to determine the RT tag (past or future) for a given opinioned text item, We aggregate (sum) the values of the RT feature vector. If the sum is positive, the corresponding RT tag is future; otherwise, we regard the opinion as non-future. This simple model does not rely on human-labeled examples; we employ in our experiments and demonstrate its effectiveness in practice.

Text content	RT features [TE, TI, PSI, SL]	RT tag (F-future, P-past)
2012年4月25日大盘涨到了全年的最高	[1(2012年4月25日), 0, 0, 0]	P
预计必涨无疑	[0, 0, 1(预计), 0]	F
将会长时间内保持上行	[0, 0, 1(将会), 0]	F
大盘指数经过大跌最终收于 2269.13 点	[0, 0, 0, -1(2269.13 点)]	P

Figure 1: Opinioned text items and their RT features and tags.

Prediction Signal with RT Tags

After extracting the RT tags of opinions, we can use them as weights of opinions’ sentiment indicators to perform RT-aware prediction: when we aggregate the sentiment indicators, we assign different weights to opinions with different RT tags (i.e., higher weights are given to opinions whose RT tag is future). For example, the weight of the sentiment of an opinion in the aggregation can be $1 + \alpha$ if the RT tag of the opinion is future, and 1 otherwise. Here, parameter α controls the impact of RT information in the final prediction.

3 Experimental Evaluation

We evaluate the effectiveness of identifying RT tags by using our RT features and the performance of utilizing RT tags on improving predictions about the stock market.

Table 1 shows the accuracy and coverage (the corresponding feature value is not zero) of using all features and each individual feature on identifying 1,000 manually labeled (future or non-future) opinions. We can see that all of the features can identify the RT tags to some extent. Our model, which combines all them achieves the best performance. Thus, it is important to consider all these features if we want to achieve the best performance.

We now showcase the effectiveness of the identified RT tags on predicting the movement of the Shanghai

Table 1: Performances on identifying RT tags.

Feature(s)	ALL	TE	TI	PSI	SL
Accuracy	95%	82%	73%	96%	86%
Coverage	98%	48%	27%	30%	7%

market Index (SHI) between 2013-01-02 and 2013-06-01 based on 50,169 microblogs of investors posted on SinaWeibo. Figure 2 shows the effectiveness of our model (RT tags+Sentiments+Baseline) and its competitors (*Traditional*: Sentiments+Baseline and *Baseline*) (Oh and Sheng 2011). More details about the experimental setup and the competitor approaches are on our complementary-material webpage. Observe that our *Time-sensitive* model, which is based on RT tags performs better than *Baseline* and *Traditional* on most α values, indicating that the consideration of RT can help in improving the prediction accuracy.

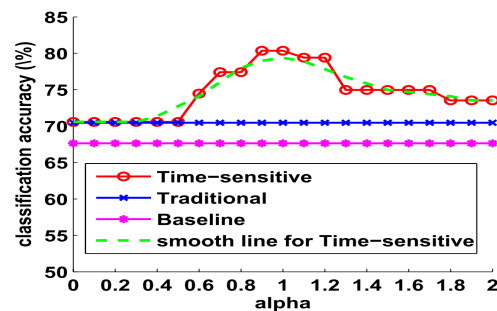


Figure 2: Performances on prediction.

4 Related Work

Our work is closely related to temporal information retrieval (Chambers et al. 2007). Due to space limitations, we refer the reader to our complementary-material webpage for a more detailed coverage of related work.

5 Conclusion and Future Work

In this paper, we propose a framework for discovering temporal references (future and non-feature) from prediction opinions and for using the extracted RT tags to build accurate prediction models. In the future, we plan to refine the RT tags to refer to a more detailed time granularity (e.g., past, present, near-future, far-future). This way, we could achieve more accurate prediction signals and further improve the prediction accuracy of our framework.

References

Chambers, N.; Wang, S.; and Jurafsky, D. 2007. Classifying temporal relations between events. In *ACL*.

Dave, K.; Lawrence, S.; and Pennock, D.M. 2003. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *WWW*.

Oh, C.; and Sheng, O. 2011. Investigating predictive power of stock micro blog sentiment in forecasting future stock price directional movement. In *ICIS*.