

Emotion Modeling from Writer/Reader Perspectives Using a Microblog Dataset

Yi-jie Tang and Hsin-Hsi Chen

Department of Computer Science and Information Engineering
National Taiwan University, Taipei, Taiwan

tangyj@nlg.csie.ntu.edu.tw; hhchen@csie.ntu.edu.tw

Abstract

Most recent studies on emotion analysis and detection focus on how writers express their emotions through textual information. In this paper, we model emotion generation on the Plurk microblogging platform from both writer and reader perspectives. Support Vector Machine (SVM)-based classifiers are used for emotion prediction. To better model emotion generation on such a social network, three types of non-linguistic features are used: social relation, user behavior, and relevance degree, along with textual features. We found that each of the non-linguistic features can be combined with linguistic features to achieve higher performance. In fact, the combination of linguistic, social, and behavioral features performs the best.

1 Introduction

Emotions express humans' feeling and experiences on some subject matters. They are typically recognized in text, speech, body gestures, and some visual information. Emotion mining is crucial for many applications, including customer care (Gupta, Gilbert, and Fabrizio 2010), sale prediction (Liu, et al. 2007), game animation (Bernhaupt et al. 2007), and robot simulation (Becker, Kopp, and Wachsmuth 2004). Capturing people's feelings, predicting their reactions to events, and generating suitable emotions are typical tasks in emotion mining.

Emotion-tagged corpora are indispensable for emotion modeling. Recently, the social media known as weblogs, or blogs, encourage users to share their emotions through writing. For example, bloggers regularly use emoticons to express personal feelings in their written posts. To encourage increased reader interaction, some news media, e.g., Yahoo! Kimo News, provide a vot-

ing mechanism for news readers to express their feelings about news articles they've just read. The collection of blogger posts and news reader responses forms writer and reader emotion-tagged corpora, respectively, facilitating writer emotion and reader emotion mining.

Previous studies (e.g., Yang, Lin, and Chen 2007a; Yang, Lin, and Chen 2007b; Yang, Lin, and Chen 2008) have used an emotion-tagged weblog corpus to investigate the ways in which people express their emotions, trying to detect writers' affective status with textual contents they have written. While these studies aimed to perform emotion analysis and detection from the writer's perspective, a few papers have studied reader emotion generation (Lin, Yang, and Chen 2007; Lin and Chen 2008; Lin, Yang, and Chen, 2008) using an emotion-tagged news corpus, modeling how readers react to articles on news websites.

To study how writer emotion affects readers' feelings, Yang, Lin and Chen (2009) used the Yahoo! Kimo Blog and Yahoo! Kimo News to produce a dataset annotated with both writer and reader emotions. They constructed a document-level reader-emotion classifier using the Yahoo! Kimo News corpus, and applied the resulting classifier on the Yahoo! Kimo Blog corpus. In this way, a new blog corpus labeled with both writer and reader emotions was obtained.

The major problem with the above approach is that the reader emotion tagging on the writer corpus depends on classification performance. Plurk, a unique social network and microblogging platform, provides a new opportunity in which a dialogue consists of posts and corresponding replies. A poster begins by publishing a post along with an emotion, then a replier responds to the post and labels it with an emotion symbol. The replier serves as a reader, and also as a writer when the reply has been attached.

Therefore, the original poster has only a writer emotion, but the replier has both a writer emotion and a reader emotion.

In this paper, we model emotion mining from the writer perspective, reader perspective, and the combined writer and reader perspective. To collect data including both writers' and readers' emotional information, we extracted messages from Plurk, ending up with 50,000 conversations in the dataset.

Support Vector Machine (SVM) was chosen as classifiers to predict repliers' emotion. Like other related studies, this experiment included textual features for training and testing. Since the conversations collected present communication and interaction between social network users, some non-linguistic features were taken into account. As a result, 4 types of features are used, including linguistic features, social relation, user behavior, and relevance degree.

The rest of this paper is organized as follows. Section 2 discusses previous work related to emotion studies. Section 3 introduces the Plurk social network and describes the extraction of the dataset. Section 4 discusses how emotions from reader and writer perspectives are analyzed. Section 5 describes the SVM classifier, along with the feature set. Section 6 details the performance of the prediction tasks, and discusses and compares the usefulness of different types of features. The final section concludes the paper.

2 Related Work

Mishne (2005) adopts mood taggings in LiveJournal articles to train a mood classifier on document-level with SVM. Mishne and Rijke (2006) use a blog corpus to identify the intensity of community mood during some given time intervals. Jung, Choi, and Myaeng (2007) also focus on the mood classification problem in LiveJournal.

Yang, Lin, and Chen (2007a) use Yahoo! Kimo Blog as corpora to build emotion lexicons. A collocation model is proposed to learn emotion lexicons from weblog articles. Emotion classification at sentence level is experimented by using the mined lexicons to demonstrate their usefulness. Yang, Lin, and Chen (2008) further investigate the emotion classification of weblog corpora using SVM and conditional random field (CRF) machine learning techniques. The emotion classifiers are trained at the sentence level and applied to the document level. Their experiments

show that CRF classifiers outperform SVM classifiers.

Lin, Yang and Chen (2007) pioneer reader emotion analysis with an emotion-tagged Yahoo! Kimo news corpus. They classify documents into reader emotion categories with SVM and Naïve Bayes classifiers (Lin, Yang and Chen, 2008). Besides classification, Lin and Chen (2008) propose pairwise loss minimization (PLM) and emotional distribution regression (EDR) to rank reader emotions. They show that EDR is better at predicting the most popular emotion, but PLM produces ranked lists that have higher correlation with the correct lists. Yang, Lin, and Chen (2009) further introduce the application of emotion analysis from both the writer's and reader's perspectives. The relationships between writer and reader emotions are discussed in their works.

Besides long articles, some studies also deal with emotion detection of short messages from microblogs and news headlines. Strapparava and Mihalcea (2007) focus on the emotion classification of news headlines. Go, Huang, and Bhayani (2009) use distant supervision for sentiment classification of Twitter messages. In their study, SVM outperforms Naïve Bayes and Maximum Entropy, and has the accuracy of 82.2%. Sun et al. (2010) focus on the Plurk microblogging platform, using text content and the NTU Sentiment Dictionary to build their feature set. These studies all focus on writer's emotions rather than reader's emotions.

Our contributions are different from the others. We employ the emotion tagging of both posters and repliers in Plurk and investigate reader and writer emotion analysis with both linguistic and non-linguistic features using a machine learning approach.

3 The Plurk Dataset

3.1 The Plurk Social Network

Plurk is a web-based social network that allows users to post short messages limited to 140 characters. From this viewpoint, Plurk is similar to Twitter and other microblogging platforms. Unlike Twitter, however, Plurk also acts like an instant messaging system because a user can see replies as soon as they are sent by another user. A post and its replies are grouped within a box on the screen, indicating that they are messages from the same conversation. Every post can be given an optional "qualifier," which is a one-word verb indicating the poster's action or feeling. There are 18 qualifiers, including Loves,

Likes, Shares, Gives, Hates, Wants, Wishes, Needs, Will, Hopes, Asks, Has, Was, Wonders, Feels, Thinks, Says, and Is. Figure 1 shows a typical conversation on Plurk. In this conversation, the first line was entered by a poster. He chose "loves" as the qualifier, stating that he "loves the iPod." The other messages were en-

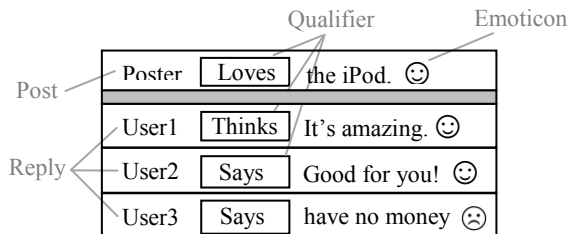


Figure 1. A conversation on Plurk

tered by other users as replies to the poster. Their messages are followed by graphic emoticons that express their emotions.

Plurk provides 78 basic graphic emoticons, and these emoticons are commonly used in users' messages. We choose 35 of the emoticons and categorize them into the positive and negative group. The other 43 are either neutral or cannot be clearly categorized, so we exclude them to minimize uncertainty. Figure 2 lists the Plurk emoticons used in this study.

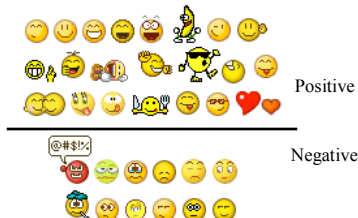


Figure 2. Emoticons as positive and negative labels

Plurk is very popular in Taiwan and some other Asian countries. Figure 3 shows the number of unique daily visitors from Taiwan. As of August 2009, it has about four hundred thousand unique daily visitors, and the number keeps increasing. Thus, we can easily obtain an enough amount of data suitable for training and testing.



Figure 3: Number of unique daily visitors of Plurk according to Google Trends.

3.2 Dataset

We prepare our dataset from the Plurk platform. In this dataset, there are 50,000 conversations dating from Jun 21, 2008 to Nov 7, 2009, and each of them consists of a post with or without emoticon and a corresponding reply with an emoticon. All the replies have to be the first reply to a post, because this can help us make sure the reply is a response to the original post rather than to other responses. All messages are in Traditional Chinese.

We filter out some messages by their qualifiers. For example, we filter out the messages with the "share" qualifier, because most "shares" are website links or images rather than general text messages. If a message contains an emoticon that is not shown in Figure 2, it will also be filtered out. Such an emoticon does not present obvious positive or negative emotion, and will not be used in our study.

In the dataset, there are 42,115 conversations with a positive reply and 7,885 conversations with a negative reply. These conversations are obtained randomly from the Plurk website, and we think this should reflect their actual distribution on Plurk. For this reason we use this dataset without adjusting the proportion of the two emotion types. The proportion of positive conversations (84.23%) is used as baseline.

4 Reader/Writer Perspective

Most related studies focus on the analysis and detection of writer's emotion, since a writer's content has a more direct link to his emotion, and corpora containing writer's emotion are easier to find on the Web. In this paper, we try to model the generation of reader's emotion, and this kind of emotion can be related to the content written by poster, replier, or both. Depending on different perspectives, we have 3 types of models: reader model, writer model, and reader + writer model. Figure 4 shows important components in the modeling: a poster pt and the text $T(pt)$ that pt posts; a replier rp , the text $T(rp)$ used by rp to reply to pt and rp 's emotion $E(rp)$; $S(pt, rp)$ denotes the social relationship between pt and rp ; $B(rp)$ denotes the behavior of rp ; and $R(T(pt), T(rp))$ denotes the relevance between post $T(pt)$ and reply $T(rp)$. The uses of the components will be discussed in detail in the following.

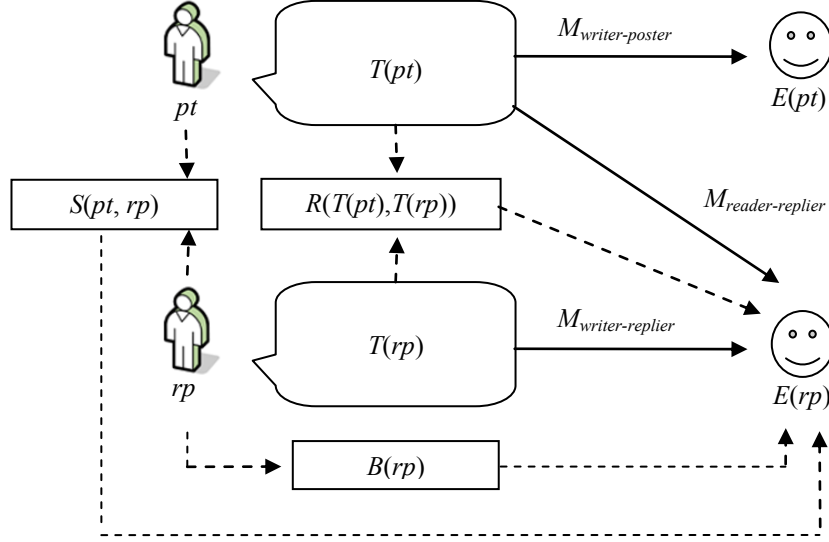


Figure 4. Different emotion generation models on Plurk

4.1 Reader Perspective

By looking at a replier’s emotion from reader perspective, we can build a reader model. In this model, we assume a replier’s emotion is directly generated by reading the poster’s message, and then the replier expresses his emotion by using an emoticon in his reply. It is indicated by the model $M_{reader-replier}$ in Figure 4. That is, $E(rp) = M_{reader-replier}(T(pt))$, where $M_{reader-replier}$ is a function that maps $T(pt)$ into an emotion. Besides $T(pt)$, we can consider social relationship between rp and pt , and the behavior of rp such that $E(rp) = M_{reader-replier}(T(pt), S(pt, rp), B(rp))$.

4.2 Writer Perspective

In a conversation, both the poster and the replier produce textual contents. To model emotion generation from writer’s perspective, we assume users’ emotions are related to their own contents. Thus, we have two types of writer model: poster’s writer model and replier’s writer model. In our study, we mostly deal with replier’s writer model, while poster’s writer model is listed for comparison. For replier’s writer model, a replier’s content is used to predict his own emotion. The model $M_{writer-replier}$ in Figure 4 indicates the generation of a replier’s emotion from writer’s perspective. That is, $E(rp) = M_{writer-replier}(T(rp))$, where $M_{writer-replier}$ is a function that maps $T(rp)$ into an emotion. Besides $T(rp)$, we can consider social relationship between rp and pt , and the behavior of rp such that $E(rp) = M_{writer-replier}(T(rp), S(pt, rp), B(rp))$. For poster’s writer model, a

post’s content is used to predict his emotion. That is, $E(pt) = M_{writer-poster}(T(pt))$.

4.3 Reader and Writer Perspective

We combine both reader and writer perspectives, and assume a replier’s emotion is related to both poster’s content and the replier’s own content. Thus, a replier’s emotion is predicted using poster’s and replier’s texts. In this case, $E(rp) = M_{reader-writer}(T(pt), T(rp), R(T(pt), T(rp)))$, where $M_{reader-writer}$ is a function maps $T(pt)$, $T(rp)$, $R(T(pt), T(rp))$ into an emotion. Besides textual information, we can also introduce social relationship between rp and pt , and the behavior of rp into this function.

5 Emotion Modeling

SVM is adopted as classifiers to predict emotion from reader and/or writer perspectives. Besides textual features, we also incorporate non-textual features such as social relation, user behavior, and relevance degree.

5.1 Text Features (T)

Since about 70% of Chinese words are disyllabic, and new words and slangs are commonly used in social media, we use bigrams instead of words as features. Chinese character bigrams in all poster’s and/or replier’s messages are extracted. We model the relationship between a bigram w and an emotion e as probability $P(w|e)$.

A training set is composed of conversations between posters and repliers. A conversation scenario between a poster and a replier is as fol-

lows. A poster pt writes down a post $T(pt)$ with emotion $E(pt)$. After a replier rp reads the post $T(pt)$, rp writes down a reply $T(rp)$ with emotion $E(rp)$. Note poster pt writes and replier rp reads the same message $T(pt)$, and express emotions $E(pt)$ and $E(rp)$, respectively. In contrast, replier rp reads and writes different messages, i.e., $T(pt)$ and $T(rp)$, with the same emotion $E(rp)$.

In this way, we have three data sets $D_{writer-poster}$, $D_{reader-replier}$, and $D_{writer-replier}$. Here, $D_{writer-poster}$ is composed of all the messages of posters along with their emotions. $D_{reader-replier}$ consists of all the messages which repliers read and emotions they express. $D_{writer-replier}$ denotes a set of messages and emotions that repliers make. These three data sets are used to train $P_{writer-poster}$, $P_{reader-replier}$, and $P_{writer-repliers}$ respectively.

To apply SVM in the experiments, libSVM is used as the classification tool (Chang and Lin 2001). The libSVM parameter selection tool found that $C=3$ and $\gamma=0.13$ yielded the best results.

5.2 Social Relation (S)

The text-based emotion model does not consider the personalization issue. Intuitively, each replier has his own preference. Social relationship between a poster and a replier is the first cue. We measure the social relationship between two users with their interaction degree. The following three features are proposed.

S_1 defines an interaction degree between users u_1 and u_2 as their total number of interactions.

$$S_1(u_1, u_2) = \sum_{(u_1, u_2) \in D} 1 \quad (1)$$

where D is a multiset of conversations (u_1, u_2) , and u_1 and u_2 are poster and replier in the conversation.

Feature S_2 considers how often user u_1 posts messages.

$$S_2(u_1, u_2) = \frac{S_1(u_1, u_2)}{(end-start) \sum_{(u_1, replier) \in D} 1} \quad (2)$$

Where $start$ and end denote the starting day and the ending day of the interaction between user u_1 and u_2 . Here S_2 equals to S_1 divided by the frequency of posts by poster u_1 .

We also consider how often a replier posts a reply. S_3 defined as follows captures this idea.

$$S_3(u_1, u_2) = \frac{S_1(u_1, u_2)}{(end-start) \sum_{(poster, u_2) \in D} 1} \quad (3)$$

5.3 User Behavior (B)

Individual user behavior is another feature. It models the subjective tendency of a user. The history of a specific replier shows which emotions he tends to express often. B_{-int} defines the negative tendency of user u .

$$B_{-int}(u) = \frac{C(E(u)=0)}{C(E(u)=0)+C(E(u)=1)} \quad (4)$$

where u is a replier, $E(u)$ is the replier's emotion with a value 0 (negative) or 1 (positive). C is the frequency of $E(u)$. This indicator does not take the interaction with posters into account.

We also consider how often he expresses his positive emotion to a specific poster. This feature is called interactive behavior (B_{+int}) and is defined as follows.

$$B_{+int}(u) = \frac{C(E(u)=1)}{\sum_{(poster, u) \in D} 1} \quad (5)$$

In some cases, replier's behavior history is not available. We use back-off smoothing to deal with this issue. Interactive user behavior after smoothing (B_s) is defined as:

$$B_s(E(rp) = e | T(pt)) = \begin{cases} P(E(rp) = e | pt = u) & \text{if } \sum_{e \in EM} C(E(rp) = e, pt = u) > K_1 \\ P(E(rp) = e) & \text{if } \sum_{e \in EM} C(E(rp) = e) > K_2 \\ P(E(RP) = e) & \text{otherwise} \end{cases} \quad (6)$$

where rp is a replier, pt is a poster, and RP is a set of all repliers. We set K_1 and K_2 to 1 in the experiments.

5.4 Relevance Degree (R)

Although a post and its reply are in the same conversation, they are not necessarily on the same topic or fully related to each other. This may affect the use of emoticons, so we also deal with relevance degree. $R(T(pt), T(rp))$ is defined as follows:

$$R(T(pt), T(rp)) = \begin{cases} 1 & \text{if there exists an anaphoric element in } T(rp) \\ 0.5 + \frac{\text{total overlapped bigrams in } T(pt) \text{ and } T(rp)}{\text{total bigrams in } T(pt)} & \text{otherwise} \end{cases} \quad (7)$$

If there exists an anaphoric element or a conjunction in replier's message, then we say the conversation is related and assign relevance degree to 1. Nine anaphoric elements and 43 conjunctions are adopted. Otherwise, we check if the post and the reply overlap. More overlapped words mean more related. We assume the post and the reply have some basic relationship, so

that the default relevance degree is set to 0.5. In current design, although the relevance degree is measured based on some linguistic markers, we still call it a non-linguistic feature for comparison with Text feature.

5.5 Normalization

The size of the linguistic feature set is much larger than the three non-linguistic feature sets, so we apply the following vector normalization method to deal with the issue:

$$F = (f_1, f_2, f_3, \dots, f_n) \quad (8)$$

$$NF = (nf_1, nf_2, nf_3, \dots, nf_n) \quad (9)$$

$$nf_i = \frac{f_i}{\sqrt{f_1^2 + f_2^2 + f_3^2 + \dots + f_n^2}} \quad (10)$$

F is a vector representing a feature set with n features $f_1, f_2, f_3, \dots, f_n$. To get the normalized F , each f value is divided by the length of F . Thus, we have the normalized F , which is defined as NF above, with features $nf_1, nf_2, nf_3, \dots, nf_n$.

6 Results and Discussion

Classifiers were trained and tested with 10-fold cross-validation. In this section, the results of the models from the three perspectives are shown and discussed.

6.1 Text Features (T)

T	Reader model	80.67%
	Writer model	88.75%
	Reader+Writer model	88.71%
S		82.78%
B_{-int}		84.14%
B_{+int}		86.25%
B_s		86.93%
R		81.53%

Table 1. Accuracy of different feature sets

First, we use an individual feature set at a time to compare their performance. The linguistic feature set (T) is used to model replier’s emotion generation from three different perspectives. When performing the prediction task with the reader model and the writer model, 3,000 bigrams from either poster’s or replier’s messages were used, respectively. For performing the task with the reader + writer model, all the bigrams from both the reader and writer models were used, for a total of 6,000 features.

Table 1 shows that the writer model and the reader+writer model achieved much higher performance than the reader model. The performance of the writer model is slightly higher than that of the reader+writer model, but the t-test shows that the difference is insignificant. The performance of the writer model and the reader+writer model is higher than the baseline (84.23%), while the performance of the reader model is lower than that of the baseline.

Interactive user behavior (B_{+int}) outperformed non-interactive user behavior (B_{-int}), and achieved performance (86.25%) higher than the baseline. After applying back-off smoothing, the interactive user behavior (B_s) proved to achieve even higher performance (86.93%), which is the best among all non-linguistic feature sets.

Social relation (S) and relevance degree (R) performed lower than the baseline, with relevance degree (R) performing the worst. Most replies should be related to their posts since they are in a conversation, and because participants are usually friends. However, 85.27% of conversations have a relevance degree of 0.5, the lowest value, which means there were not anaphoric elements, conjunctions, and overlaps. Relevance is not easy to be measured accurately between two short messages. In summary, when each of the non-linguistic feature sets is used individually, the following results are seen: $B_s > B_{+int} > B_{-int} > S > R$. For the behavior feature set, back-off smoothing is useful. In addition, the behavior pattern in response to a specific poster is more useful than to all posters, suggesting that the affective interaction between two given users may be based on a certain pattern.

6.2 Combination of Feature Sets

Experimentation with some combinations of different feature sets was also conducted. Table 2 shows the results of these combinations, from reader, writer, or reader and writer perspectives. Writer models still outperformed reader models, and are slightly better than reader+writer models for all feature combinations except for the SVM model with the $T + B_s + S$ combination.

When combined with textual features, the behavioral feature set was still more powerful than social relation and relevance degree. However, all these 3 feature sets are helpful since paired t-tests show that the differences between T and $T + B_s$, T and $T + S$, and T and $T + R$ to be significant ($p < 0.05$).

	Reader Models	Writer Models	Reader + Writer Models
T	80.67%	88.75%	88.71%
$T + S$	83.42%	89.60%	89.26%
$T + B_s$	88.02%	91.42%	91.16%
$T + R$	82.73%	89.14%	88.93%
$T + B_s + R$	88.14%	91.48%	91.27%
$T + B_s + S$	88.42%	91.60%	91.61%
$T + B_s + S + R$	88.37%	91.53%	91.30%

Table 2. Accuracy of models with different feature set combinations

Because B_s is most useful when used with textual features, $T + B_s$ with $T + B_s + S$ and $T + B_s + R$ were compared to see how S and R can improve performance. For the reader models with SVM, the difference between $T + B_s$ and $T + B_s + S$ was significant ($p < 0.05$), but the difference between $T + B_s$ and $T + B_s + R$ was insignificant. This shows that $T + B_s + S$ is a more useful combination than $T + B_s + R$. For writer and reader + writer models, $T + B_s + S$ still outperformed $T + B_s + R$.

Although each of the 3 non-linguistic features can improve performance, combining all of them ($T + B_s + S + R$) does not achieve the highest performance. The highest performance is achieved by the combination of $T + B_s + S$, regardless of which perspective is used. According to results from the paired t-test, the difference between $T + B_s + S + R$ and $T + B_s + S$ is insignificant for the reader model and the writer model. This shows that although adding R to the combination does not decrease the performance significantly, it is also not helpful. The reasons for this may be the following: both social relation and interactive behavior are related to interaction between two specific users, so their effects may overlap; only 14.73% of conversations have a relevance value higher than 0.5.

6.3 Different Perspectives

For all feature set combinations, the writer models and the reader+writer models achieve better performance than the reader models. These differences are significant according to the paired t-tests, which suggests that for predicting a replier’s emotion, the message generated by the replier him- or herself contains more useful infor-

mation than the message generated by the poster and then read by the replier.

When using the textual feature set only, the reader model’s SVM performance (80.67%) was much lower than the writer model’s (88.75%) and that of the reader+writer model (88.71%). When T is used with B_s and S , in contrast, the SVM performance of the reader model is 88.42%, only slightly lower than the performance of the writer model (91.60%) and the reader+writer model (91.61%). This indicates that when modeling emotion generation on a social network, non-linguistic features play more important roles.

The performance of textual feature set for the writer model with SVM is 88.75%, slightly higher than that for the reader+writer model (88.71%). According to results of the paired t-test, the difference between them is insignificant. For the $T + B_s + S$ combination, the performance of the reader+writer model (91.61%) is slightly higher than the performance of the writer model (91.60%), though the difference is also insignificant. Thus, it makes little difference in performance whether emotion generation is modeled from writer perspective or both reader and writer perspectives. In this series of experiments, 91.61% was the highest accuracy achieved.

6.4 Writer Model

As mentioned in the Section 4, another kind of writer model exists, for which the content is written by the poster, of which was also included as experiment with poster’s writer model. In this case, only the linguistic feature set can be used. Results seen included an accuracy of 89.19%. Results of t-test for the posters’ and repliers’ writer model showed the difference as insignificant ($p < 0.082$). However, it is important to note that the dataset used for the posters’ writer model differs from the one used for the repliers’ writer model, so this comparison is for reference only.

7 Conclusion

To better model emotion generation on a micro-blogging platform with social network characteristics, different models from the reader and/or writer perspectives were included in the experiment, and showed their differences. Discoveries included that predicting emotion from the reader perspective is more challenging than from the writer perspective. In addition, using non-linguistic features with linguistic features for emotion prediction resulted in discovering that each of the non-linguistic feature sets is useful.

In this study, the combination of all feature sets did not achieve the best performance. In future work, the weights and combination methods of different feature sets will need to be further studied. Additional efforts will also be needed to precisely represent the characteristics of user interaction and message contents. The relevance degree used in this study, for example, deals with only anaphoric elements, conjunctions, and overlapped bigrams in this study. Other factors and resources also will be needed to more effectively determine the relevance of two messages.

As this paper suggests, a writer model is different from a reader model. The same bigrams or words can have different effects on writers' and readers' emotional expression. For example, greetings can cause a positive reader response even if the writer uses a negative emoticon and shows some negative feelings. Thus, these findings suggest that reader emotion be further explored in future studies.

The models presented in this paper are useful for a wide range of applications, especially those related to conversation and interaction between humans and machines. They can also help improve the performance of automated customer service and writing assistance systems, in which readers' emotional responses are important. Different types of features can be used for different application domains. The behavioral feature, for example, can be used when a user's conversation history is available.

References

- C. Becker; S. Kopp; and I. Wachsmuth. 2004. Simulating the Emotion Dynamics of a Multimodal Conversational Agent. In *Proceedings of Tutorial and Research Workshop on Affective Dialogue Systems*, 154-165.
- R. Bernhaupt; A. Boldt; T. Mirlacher; D. Wilfinger; and M. Tscheligi. 2007. Using Emotion in Games: Emotional Flowers. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, 41-48.
- C.C. Chang and C.J. Lin. 2001. LIBSVM: a Library for Support Vector Machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- N. Gupta; M. Gilbert; and G.D. Fabrizio. 2010. Emotion Detection in Email Customer Care. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 10-16.
- A. Go; L. Huang; and R. Bhayani. 2009. *Twitter Sentiment Classification Using Distant Supervision*. CS224N Project Report, Stanford University, Stanford, CA.
- Y. Jung; Y. Choi; and S.H. Myaeng. 2007. Determining Mood for a Blog by Combining Multiple Sources of Evidence. In *Proceedings of International Conference on Web Intelligence*, 271-274.
- H.Y. Lin and H.H. Chen. 2008. Ranking Reader Emotions Using Pairwise Loss Minimization and Emotional Distribution Regression. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 136-144.
- H.Y. Lin; C.H. Yang; and H.H. Chen. 2007. What Emotions Do News Articles Trigger in Their Readers? In *Proceedings of 30th Annual International ACM SIGIR Conference*, 733-734.
- H.Y. Lin; C.H. Yang; and H.H. Chen. 2008. Emotion Classification of Online News Articles from the Reader's Perspective. In *Proceedings of International Conference on Web Intelligence*, 220-226.
- Y. Liu; X. Huang; A. An; and X. Yu. 2007. ARSA: A Sentiment-Aware Model for Predicting Sales Performance Using Blogs. In *Proceedings of the 30th Annual International ACM SIGIR Conference*, 607-614.
- G. Mishne. 2005. Experiments with Mood Classification in Blog Posts. In *Proceedings of the 1st Workshop on Stylistic Analysis of Text for Information Access*, Salvador, Brazil.
- G. Mishne and M. De Rijke. 2006. Capturing Global Mood Levels Using Blog Posts. In *Proceedings of AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs*, 145-152.
- Y.T. Sun; C.L. Chen; C.C. Liu; C.L. Liu; and V.W. Soo. 2010. Sentiment Classification of Short Chinese Sentences. In *Proceedings of 22nd Conference on Computational Linguistics and Speech Processing*, Nantou, Taiwan, 184-198.
- C. Strapparava and R. Mihalcea. 2007. Affective Text. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, 70-74.
- C.H. Yang; H.Y. Lin; and H.H. Chen. 2007a. Building Emotion Lexicon from Weblog Corpora. In *Proceedings of 45th Annual Meeting of Association for Computational Linguistics*, 133-136.

- C.H. Yang; H.Y. Lin; and H.H. Chen. 2007b. Emotion Classification Using Web Blog Corpora. In *Proceedings of 2007 IEEE/WIC/ACM International Conference on Web Intelligence*, 275-278.
- C.H. Yang; H.Y. Lin; and H.H. Chen. 2008. Sentiment Analysis in Weblog Using Contextual Information: A Machine Learning Approach. *International Journal of Computer Processing of Languages* 21(4): 331-345.
- C.H. Yang; H.Y. Lin; and H.H. Chen. 2009. Writer Meets Reader: Emotion Analysis of Social Media from both the Writer's and Reader's Perspectives. In *Proceedings of International Conference on Web Intelligence*, 287-290.