Sentiment Prediction using Collaborative Filtering

Jihie Kim, Jaebong Yoo, Ho Lim, Huida Qiu, Zornitsa Kozareva, Aram Galstyan
Information Sciences Institute, University of Southern California
{jihie, jaebong, kozareva, galstyan}@isi.edu, huidaqiu@usc.edu

Abstract

Learning sentiment models from short texts such as tweets is a notoriously challenging problem due to very strong noise and data sparsity. This paper presents a novel, collaborative filtering-based approach for sentiment prediction in twitter conversation threads. Given a set of sentiment holders and sentiment targets, we assume we know the true sentiments for a small fraction of holder-target pairs. This information is then used to predict the sentiment of a previously unknown user towards another user or an entity using collaborative filtering algorithms. We validate our model on two Twitter datasets using different collaborative filtering techniques. Our preliminary results demonstrate that the proposed approach can be effectively used in twitter sentiment prediction, thus mitigating the data sparsity problem.

Introduction

Social media has become an important tool for information dissemination, search and marketing. Many campaigns (marketing, non-commercial, political) use Facebook and Twitter to communicate their messages and opinions. Social media participants often respond to those messages (and to each other) with expressions towards or against particular subjects or targets. Detecting and characterizing these "sentiment" expressions can be very important for capturing attitudes, positions or traits of individual participants or groups.

While the majority of the research focuses on detecting and analyzing sentiments expressed in individual twitter message (Kouloumpis, Wilson, and Moore 2011; Calais Guerra et al. 2011), there is still a necessity for developing sentiment analyzers in interactive discussions that are formed through reply-to chains. Unlike retweets where the user's retweeting activity expresses an agreement with what the original user has posted, in reply-to messages any sentiment (positive, negative, neutral) can be expressed towards the previous message or the discussion topics. Participation in discussions also involves more effort from the poster than a simple retweet, which can indicate higher engagement of the participants. Through various forms of interactions, users express their sentiment toward the previous message (agree/disagree) or certain discussion topics (positive/negative).

The goal of our research is to use twitter discussion threads and identify the sentiments of users towards other posters and topics. There are multiple challenges associated with the solution of this problem. One of them concerns the fact that users do not often use sentiment bearing words to express their opinion, hence a more complex sentiment prediction model is necessary. Another significant issue is the data sparsity, as users do not always express their opinion towards all topics or users, which is making it hard to build a prediction model.

In this paper, we address these challenges by introducing a novel sentiment prediction framework, which relies on *collaborative filtering* techniques. Collaborative filtering has been successfully used to solve various preference prediction problems. For instance, product purchase patterns can be generalized to predict preferences of similar users. The main idea behind our framework is that the same reasoning can be applied to sentiments: *A sentiment of a user toward a specific target (or another user) can be predicted based on the sentiments of similar users.* We explore several alternative approaches for representing and predicting the sentiment based on content features in the messages and message reply-to relations in discussions. Our experimental studies on two different twitter political discussion datasets indicate that collaborative filtering techniques can effectively predict user's sentiment towards discussion topics or relations among discussants.

The main contributions of the paper are two-fold: (1) application of collaborative filtering techniques in the context of sentiment classification and (2) demonstration of the effectiveness of the proposed approach using two different twitter political discussion datasets.

Sentiment Prediction in Twitter

In Figure 1 we show an example of a discussion thread that forms within Twitter. In this example, the discussion has a sequence of messages: M1, M2, . . . , M7 for which the reply-to relation forms a tree-like structure among the messages. The users (with user-prefix) represent the discussants. The first user initiates the thread by expressing his/her sentiment towards the election candidates (i.e. *Tony Blair who the former prime minister in UK and campaigned for Labour in UK election 2010*). Then the other users respond to the message with further emotional expressions or sympathy.

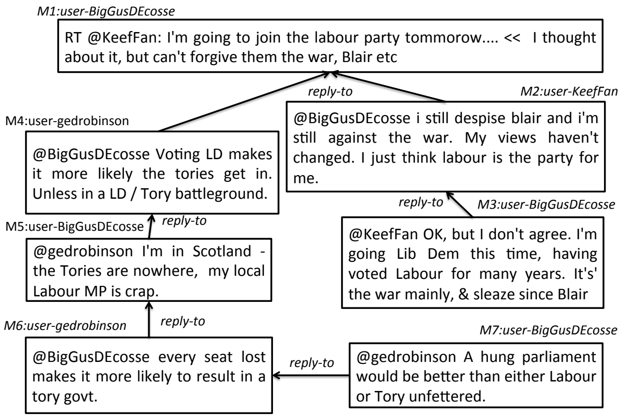

Figure 1: An example discussion in Twitter.

The main goal of our research is to use twitter discussion threads such as the one shown above for identifying the sentiment of a given user towards other posters and topics. This is a very hard and challenging research problem because often sentiment words like "*I don't agree*" or "*despise*" are not explicitly used in the twitter messages and standard word-based sentiment analysis techniques would fail to capture the sentiment. In addition, the limited amount of discussion data for each user is making it hard to observe and infer the sentiment of a user towards other participants and topics.

Given the nature of our task and the twitter discussion threads, we propose a sentiment prediction framework that is based on collaborative filtering, shown pictorially in Figure 2. The rows represent *Sentiment Holders*, whereas the columns represent *Sentiments Targets*.

Target (users + entity mentioned)

Holder (users)

	u2	u3	u5	u6	u7	u8	e1	e2	e3
u1	-	+	-	?	-	?	+	?	+
u2	?	?	?	?	?	+	?	+	?
u3	-	?	-	?	+	+	+	?	+
u4	?	+	?	+	-	+	+	+	+
u5	+	+	-	?	?	?	-	-	?
u6	-	-	-	-	-	?	?	?	?
u7	-	+	+	-	?	+	+	+	+

Figure 2: Sentiment holder - sentiment target matrix

Sentiment Holder: Given a message, we assume that the sentiment holder is the person who posted the message. There are cases where the poster quotes other people's sentiment but such cases are relatively fewer.

Sentiment Target: The target of the sentiment can be the message that the current message replies to or the person who posted the prior message. That is, the sentiment depends on agree/disagree relations between the messages.

As a concrete example, consider the first row in the matrix in Figure 2, which represented the user u_1. This user holds a negative sentiments towards the users u_2, u_5, u_7, and positive sentiment towards the user u_3 as well as the entities e_1 and e_3. The other rows can be interpreted similarly. The users express sentiment toward (+) or against (-) certain targets, as indicated in the matrix. Note that only some of the attitude information can be available or the given information can be ambiguous. The cells containing "?" represent unknown sentiments, e.g., due to lack of sufficient data. Our goal is to use collaborative filtering to predict the unknown sentiments which are not explicitly present in the data.

In our settings, the sentiment target can be either another user ("*@u2 Where is* **your** *evidence for this fake news?*) or another entity ("*Unable to stand on their own against the* **Mujahideen**, **KDF** *now serve submissively under* **Ugandan command**.").

Sentiment Classification

In this section we describe two methods for sentiment identification in twitter discussion thread messages. This information is later used by the collaborative filtering approaches to predict the unknown sentiment toward another user/entity.

LIWC We characterize sentiment expressions in twitter messages using Linguistic Inquiry and Word Count (Tausczik and Pennebaker 2010), or LIWC. LIWC is widely used to evaluate emotional and psychological features in diverse texts including short texts, blogs and tweets (Quercia et al. 2011). To represent the overall sentiment of a message towards the previous post, we use the *posemo* (positive emotion) and *negemo* (negative emotion) scores, which reflect the ratio of positive/negative words in the message out of all words in the message. Unlike previous approaches, we also detect the entities mentioned in the discussions and the *posemo* and *negemo* values of their neighboring words are used to capture the sentiment towards the target entity. For entity extraction we use state-of-the-art entity recognizer (Finkel, Grenager, and Manning 2005).

Sentiment Classification with Deep Learning We use deep learning to automatically learn phrase representations, hierarchical structure, and sentiment distributions (Socher et al. 2011). The model uses neural word representations that represent words as continuous vectors of parameters. These word vectors ($x \in R^n$) are stacked into a word embedding matrix $L \in R^{n \times |V|}$, where $|V|$ is the size of the vocabulary. We then can represent a sentence as an ordered list of these word vectors $(x_1, x_2, \ldots, x_{m-1}, x_m)$ by looking up the matrix L. Given a sentence with m words, the model tries to construct a tree of the sentence, a hierarchical and distributed vector representation, by minimizing the reconstruction errors. To build our model, we used $5,000$ positive and $5,000$ negative tweets. We used distant supervision technique to generate the training labeled data, since labeled twitter messages were limited. For it, we collected around 50 words that are indicative of a positive/negative emotions. For instance, *love*, *smile*, *happy* are positive words and we assume that tweets containing these words or their hashtags would express positive sentiment. Similarly words like *sad*, *cry*, *war* express negative emotions. The result of 10-fold cross validation on this dataset is .81 f-score. Once the model was built and validated, we used it to classify the unlabeled twitter discussion threads with positive and negative sentiment labels.

Prediction with Collaborative Filtering

For our evaluation, we use the following collaborative filtering algorithms that are used often for user preference prediction.

Average Baseline (User Average: UAVG): The user average (UAVG) value predicts unknown ratings by the average value of all available ratings from the users.

Weighted combination of sentiment values (UBasedCF): User-based collaborative filtering is adapted to predict the value. That is, the user's (sentiment holder's) sentiment is predicted based on a weighted sum of similar users. Given the cosine similarity values among sentiment holders s_{ik}, and their sentiment toward target $j(r_{kj})$, the predicted sentiment value is $p_{ij} = \sum_k s_{ik} r_{kj}$.

Regularized SVD (Singular Value Decomposition): We used the correlation matrix $C = \frac{1}{n-1} A^T A$, where A is a matrix of sentiment holder and target. Each cell in the matrix represents the holder's sentiment toward the target. We select the optimal principle components to project data onto the eigen plane. The similarity is calculated based on the projected data. Parameters are estimated by minimizing the sum of squared residuals, one feature at a time, using gradient descent with regularization and early stopping. Finally, prediction of sentiment of a given user is calculated based on the estimation.

PMF (Probabilistic matrix factorization): Recently, probabilistic matrix factorization became widely used for collaborative filtering (Salakhutdinov and Mnih 2007). The basic idea is that the emotion of a user towards another user or entity is determined by a few latent factors. Thus, the $N \times M$ emotion matrix can be factorized into the product of two low-rank factor matrices: $M = U^T V$, where U is a $D \times N$ user coefficient matrix and V is a $D \times M$ factor matrix. There is a lot of work on probabilistic matrix factorization in collaborative filtering, we employ the one from (Salakhutdinov and Mnih 2007). After factorizing the emotion matrix M, we multiply the two factor matrices and look-up the (i, j) entry as the emotion from user i to user/entity j.

Results

This section describes prediction accuracies with different settings: collaborative filtering approaches, different sentiment targets (user or entity) and varying data density.

Datasets For our study, we use two different political discussion datasets collected from Twitter. The first set has twitter messages on UK political election, posted between Mar/25/2010 and May/11/2010. The second set is from discussions among neighbors of a political entity, HSM-Press that is known to represent a militant organization Al-Shabaab in Africa. The data is collected between December 2011 and March 2012. The first (second) dataset has approximately 802K (172K) tweets from 173K (48K) users.

For each dataset (HSM and UK), the experiment data was formed by the following approach. First, the sentiment holder and target pair frequency is calculated. Then the most popular sentiment holders and targets are selected by identifying a threshold where the frequency drops significantly. This approach ensures that our selected data has enough common sentiment targets among the different sentiment holders for the evaluation. For evaluating prediction performance with varying density values, we randomly remove data according to the density rate, from 0.01 to 0.09.

Feature Representation We use the PREA (Lee, Sun, and Lebanon 2012) collaborative filtering toolkit, and map the sentiment information into rating values:

- LIWC: 1 if *negemo* is greater than *posemo*, 3 if *posemo* dominates, and 2 for neutral.
- Deep learning: 1 if negative emotion is greater, 3 if positive emotion dominates, and 2 for neutral.

For evaluating the sentiment classification results, we randomly hide 20% of the known values in the matrix using the sentiment classification described above. We then check whether the dominance of positive or negative sentiment is consistent with the original values.

Evaluation Configuration To evaluate the performance of our framework, we conduct the following test cases:

- Case1: predict individual user sentiments towards other users
- Case2: predict individual user sentiments towards other users and named entity targets

Discussion of Results

Given the two different cases described above, we predict the sentiments toward other users or named entities using different algorithms: naive CF (weighted sum), User-based CF, SVD, and PMF. The cell values of each matrix are the rating values that correspond to the predictions of the LIWC or deep learning sentiment classifiers.

Figure 3 shows predictive performances of these models on the different sets over different densities: the label of each plot follows this notation: $[Case]\_[SentimentInfoUsed]\_[TwitterData]$. For example, Case1_DEEP_HSM on the first (top left) plot of Figure 3 means Case1 evaluation using with sentiment values created by deep learning using the HSM twitter data. Each line represents each method (PMF, SVD, User Average, and User-based CF). For the given combination of cases and sentiment features, we ran the test for 100 times for each density level so that we can see the distribution of the RMSE (root mean squared error) as shown in Figure 3. We also used majority sentiment as a baseline, and its RMSE value for Case1_LIWC_UK is 1.143.

Several interesting points can be noted: 1) as expected RMSE tends to decrease as the density increases because we have more information (observations) to predict unknown sentiments. 2) Overall, PMF model works very

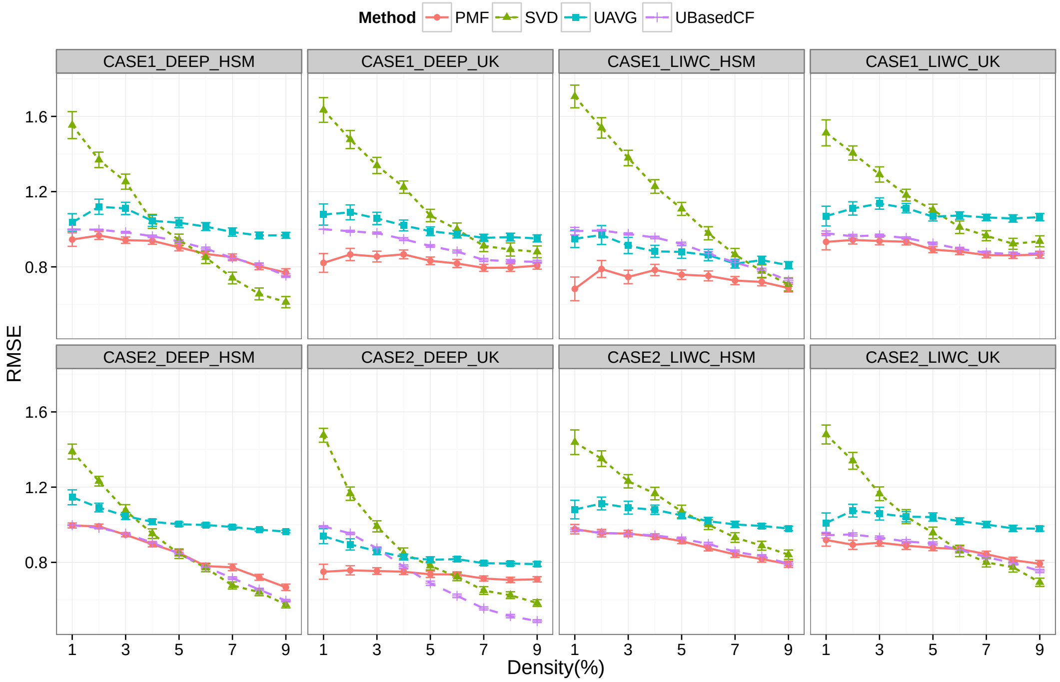

Figure 3: Sentiment prediction error for varying density of known sentiment holder/target pairs.

well compared to other methods over the different density in entire dataset. However, regularized SVD tends to be the most sensitive to the density level: the RMSE values decrease more as the density level changes. In several cases (Case1_DEEP_HSM, Case2_DEEP_HSM, and Case2_LIWC_UK), SVD is the best performing algorithm when the number of users and entities is sufficiently large. Although both SVD and PMF use matrix factorization technique to capture hidden semantic features, SVD may be more sensitive to noisy data than PMF. 3) We observed that the prediction results with deep learning data are slightly better (overall average RMSE = 0.929) than those of LIWC (overall average RMSE = 0.950). The deep learning approach provides a more stable sentiment model.

In summary, the above results show that PMF reaches the best performance for all case combinations and deep learning produces a robust sentiment model. The overall trend shows that RMSE decreases as the density increases. To conclude, the collaborative filtering techniques are useful tools for predicting sentiment in twitter conversation.

Conclusion and Future Work

Our experimental study and results showed that collaborative filtering techniques can be a useful and powerful tool for generating profiles of user sentiments. We expect that similar approaches can be used in other conversational data such as online message boards that discuss political issues. As retweet often indicates endorsement of the message, we plan to incorporate it into the collaborative filtering framework and generate a stronger model.

Acknowledgments This research was supported in part by DARPA grant No. W911NF–12–1–0034. We thank Chao Wang and Sam Shuster for helping to prepare the data.

References

Calais Guerra, P.; Veloso, A.; Meira, Jr., W.; and Almeida, V. 2011. From bias to opinion: a transfer-learning approach to real-time sentiment analysis. In *Procs. of KDD*, 150–158.

Finkel, J. R.; Grenager, T.; and Manning, C. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of ACL*.

Kouloumpis, E.; Wilson, T.; and Moore, J. 2011. Twitter Sentiment Analysis: The Good the Bad and the OMG! In *5th AAAI Conference on Weblogs and Social Media*.

Lee, J.; Sun, M.; and Lebanon, G. 2012. A Comparative Study of Collaborative Filtering Algorithms. *ArXiv e-prints*.

Quercia, D.; Ellis, J.; Capra, L.; ; and Crowcroft, J. 2011. In the mood for being influential on twitter. In *3rd IEEE International Conference on Social Computing*.

Salakhutdinov, R., and Mnih, A. 2007. Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems*.

Socher, R.; Pennington, J.; Huang, E. H.; Ng, A. Y.; and Manning, C. D. 2011. Semi-supervised recursive autoencoders for predicting sentiment distributions. In *Proceedings of EMNLP*, 151–161.

Tausczik, Y., and Pennebaker, J. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology* 29(1):24–54.