Content-based Influence Modeling for Opinion Behavior Prediction

Chengyao Chen, Zhitao Wang, Yu Lei and Wenjie Li
Department of Computing,The Hong Kong Polytechnic University
{cscchen, csztwang, csylei, cswjli}@comp.polyu.edu.hk

Abstract

Nowadays, social media has become a popular platform for companies to understand their customers. It provides valuable opportunities to gain new insights into how a person's opinion about a product is influenced by his friends. Though various approaches have been proposed to study the opinion formation problem, they all formulate opinions as the derived sentiment values either discrete or continuous without considering the semantic information. In this paper, we propose a Content-based Social Influence Model to study the implicit mechanism underlying the change of opinions. We then apply the learned model to predict users' future opinions. The advantage of the proposed model is the ability to handle the semantic information and to learn two influence components including the opinion influence of the content information and the social relation factors. In the experiments conducted on Twitter datasets, our model significantly outperforms other popular opinion formation models.

1 Introduction

Social media services, such as Twitter, Facebook, etc. provide fast and effective platforms for people to receive messages from their neighbors/friends and express their own opinions. The online communication can gradually influence one's opinions (Anagnostopoulos et al., 2008). In fact, according to a marketing survey[1], 71% of the consumers said they are more likely to make a purchase based on social media referrals. Naturally, social media offers a great chance for companies to conduct marketing by influencing the opinions of their potential customers. In order to achieve that, exploring and understanding the intrinsic mechanism of opinion formation is of great importance.

Informational influence is a primary process for forming opinions on products in social media (Das et al., 2014). It describes the following scenario: when users lack the necessary information, they will seek for the opinions of their neighbors to update their beliefs. Taking the informational influence as premise, several models are proposed with different assumptions of how a person updates her/his own opinions according to the neighbors' opinions (Clifford and Sudbury, 1973; DeGroot, 1974; Hegselmann and Krause, 2002). In these models, opinions are pre-defined as statuses through discrete categories including positive, negative and neutral opinion (Hegselmann and Krause, 2002; Galam, 2002; De et al., 2014) or continuous scales of opinion strengths (Clifford and Sudbury, 1973; DeGroot, 1974; Yildiz et al., 2011; Chazelle, 2012). However, on most social media platforms, people exchange their views by posting and replying through textual messages. The summarized opinion status simplifies the opinion formation process, and ignores the effects of semantic information hidden in the exchanged content information. Even if two messages have the same opinion category, different semantic information of the contents may result in different effects on others' opinions. We take two postings of the product "Samsung Galaxy" as examples:

(1): *"I can't post gifs on this stupid Galaxy S6."*

(2): *"Just lost my new Galaxy S6 and very sad."*

(1) expresses complaints about a problem in the usage of the product, which may lead other users to an unfavorable impression on the product. However, (2) expresses the personal sad mood for the loss and

---

[1]http://www.socialmediatoday.com/content/30-statistics-how-social-media-influence-purchasing-decisions-infographic

no bad effect on "Samsung Galaxy" is transmitted through this message. The two examples show that the summarized opinion representations are not able to differentiate the opinion influences of different expressions on other users. Therefore, it is necessary to deeply explore the user-generated content in social communication, especially to understand the actual opinion influence derived from the messages during the communication.

The problem becomes discovering the underlying relevance between a person's opinion and the received content information. The intuitive solution is to employ the co-occurrence patterns of one's opinion and her/his neighboring messages. However, as the data grows, the patterns of co-occurrences can be sparse and ineffective for prediction. Vector representations of words and phrases have been successfully applied in many Natural Language Processing tasks (Bengio et al., 2003; Le and Mikolov, 2014). Through encoding the semantic information, word embedding makes it possible to overcome the curse of dimensionality.

Therefore, we propose a Content-based Social Influence Model (CIM) based on the neural network framework which encodes the content information with word embeddings. We represent each opinion word as a dense vector in the continuous space. We then compose the opinion word vectors of one's previous message and her/his neighboring messages to form the social opinion context vector and feed the vector to a softmax layer for opinion prediction. To construct the social opinion context vector, we incorporate two social relation factors, stubbornness and interpersonal influence. Stubbornness represents the degree a user insists on her/his previous opinion and interpersonal influence represents the influence one receives from neighbors. Also, the social relation factors are polarity-related which can be either positive or negative.

Different from previous opinion formation models which only learn the opinion influence of social relationships, our proposed model learns two opinion influence components, i.e., the opinion word embeddings and the social relation factors. The learned word vectors reflect the opinion influence of different opinion words during the discussion on a specific issue, and the social relation factors including stubbornness and interpersonal influence. Integrating these two components together, our model has the capability to describe the opinion formation process more accurately. In the experiments conducted on three Twitter datasets, our proposed model performs better than other state-of-art opinion influence models. Besides, we also study the expression of users with different influence powers. The analysis could be as a reference for companies to understand the different effects of different wordings and furthermore manage their social accounts better.

The rest of paper is organized as follows. We first review the related work in Section 2. Section 3 formulates the problem and describes the framework of our proposed model. Then, the experiments and evaluation are given in Section 4. Finally, we conclude and mention potential future works in Section 5.

2 Related Work

2.1 Opinion Influence Modeling

Opinion formation is a problem firstly studied by the researchers in the sociology and statistics areas. One notable work is proposed by DeGroot (DeGroot, 1974), which takes opinions as continuous values and assumes that one updates her/his opinions by averaging neighboring opinions. Hegselmann et al., (Hegselmann and Krause, 2002) propose the Flocking model with another assumption. They assume that people are influenced by others depending on how close their opinions are. Different from these two studies which represent opinions with continuous values, voter model represents opinion as discrete category (Clifford and Sudbury, 1973). In this model, a person selects only one of her/his neighbors uniformly at random, and takes the current opinion of the neighbor as her/his own opinion. A modification is termed as the Majority voter model (Krapivsky and Redner, 2003), where the user adopts the majority opinion in his/her neighborhood. Apart from the neighboring influence, another social relation factor stubbornness is considered in opinion prediction models. It represents the degree that one insists on her/his own opinion. The DeGroot model, Flocking model and the Voter model are extended with the idea of stubbornness (Acemoglu and Ozdaglar, 2011; Yildiz et al., 2013). Recently, De et al., (De et al., 2014) propose an asynchronous linear model (AsLM) based on the DeGroot model, which first

introduces the negative influence and proves the effectiveness of the proposed model on the social media dataset. However, existing models fail to consider the effects of content information on the opinion formation problem. Our work is the first try to integrate semantic information into opinion behavior modeling.

2.2 Neural Network in NLP Tasks

Recently, neural network has received great achievements in Natural Language Processing tasks, such as language modeling (Bengio et al., 2003), machine translation (Cho et al., 2014) and sentiment classification (Tang et al., 2014). One of the most useful neural network techniques for NLP is the word embedding, which learns vector representations of words (Bengio et al., 2003; Collobert and Weston, 2008; Mikolov et al., 2013). The neural language model proposed by Bengio et al., (Bengio et al., 2003) uses the concatenation of several previous words (context) as the input of the feed-forward neural network, and then the encoded context vector is used to predict the next word (target word). Following the word embedding techniques, several models are extended to achieve the phrase-level and sentence-level representations by composing all vectors of words in the phrase or sentence together. The basic composition method is using weighted average of all word vectors (Zanzotto et al., 2010; Mikolov et al., 2013). In (Mikolov et al., 2013), they use a simple data-driven approach, where phrases are formed based on the unigram and bigram counts of the words. Furthermore, considering the syntactic structure of the phrases or sentences, a method combining the words by their orders in the syntactic tree is proposed (Socher et al., 2011).

The proposed content-based social influence model bears similarities with the neural language model. In the opinion formation tasks, we regard the neighboring opinions and one's previous opinion as the "contexts", and the "target" is one's future opinion category. The model has a more complex framework since the social relation factors including stubbornness and interpersonal influence are considered with the word embeddings.

3 Approach

3.1 Problem Definition

Formally, we denote the network of users who are interested in the same issue as $G = (V, E)$, where each vertex $u \in V$ represents a user, and each edge $e \in E$ represents a following friendship between two users. The number of users is N. The neighbor set for each user $u \in V$ is denoted by $F_u = \{v|(u, v) \in E\}$, whose size is $n(u)$.

Additionally, the opinion expression behavior of a user is formulated as a triple $< p, o, t >$ which represents that a user u posts a tweet p with the opinion category o at the timestamp t. There are three values of +1, 0, -1 for opinion category o indicating the "positive", "negative" and "neutral" sentiment respectively. Given a user u, his opinion behaviors are represented as a sequence of triples: $\{< p_u(1), o_u(1), t_u(1) >, \cdots, < p_u(i), o_u(i), t_u(i)) >, \cdots, < p_u(m(u)), o_u(m(u)), t_u(m(u)) >\}$

Furthermore, given the above definitions, we define the neighboring opinion set for each user u at each timestamp $t_u(i)$ as $C_u(i) = \{p_{F_u^1}(t_1), \cdots, p_{F_u^v}(t_v), \cdots, p_{F_u^{n(u)}}(t_{n(u)})\}$, where $t_u(i-1) < t_v < t_u(i)$ for each neighbor $v \in F_u$. It includes all the information u receives from his neighbors in F_u since previous posting time $t_u(i-1)$. Considering opinion words are the most representative parts to reflect one's opinion, we only keep the opinion words within each tweet. For brevity, we rewrite the neighboring opinion set as $C_u(i) = \{C_u^1(i), \cdots, C_u^v(i), \cdots, C_u^{n(u)}(i)\}$, where $C_u^v(i) = \{C_{u,1}^v(i), \cdots, C_{u,|C_u^v(i)|}^v(i)\}$ contains all opinion words in the tweet $p_{F_u^v}(t_v)$. If there does not exist a posting from a neighbor v during the time period, $C_u^v(i)$ is an empty set. Also, we represent the tweet $p_u(i)$ with the opinion words set $S_u(i) = \{S_{u,1}(i), \cdots, S_{u,|S_u(i)|}(i)\}$.

The problem can be defined as: given the neighboring opinion information received in previous timestamp $C_u(i)$ and previous personal opinion $S_u(i-1)$, our objective is to predict the future opinion category $o_u(i)$ at the timestamp $t_u(i)$.
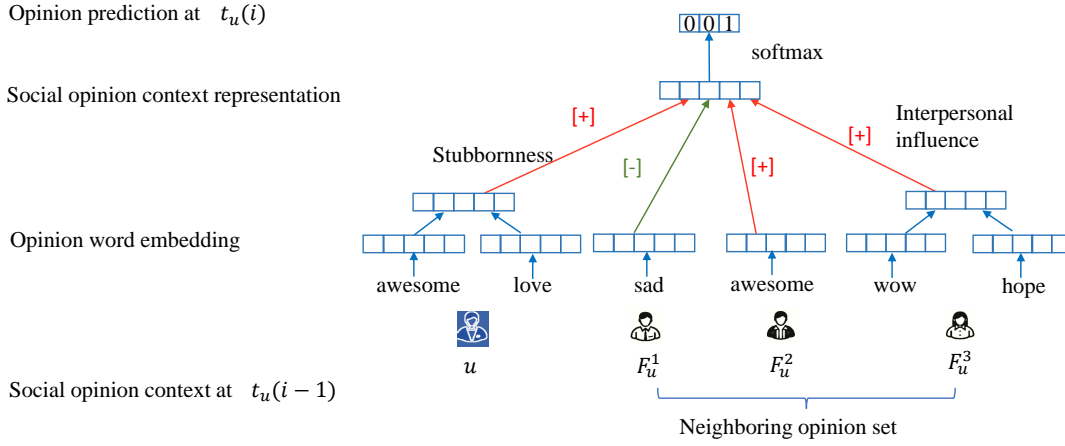
Figure 1: The graphical representation of the CIM on the opinion prediction

3.2 Framework

In this paper, we propose a novel influence model based on representation learning to solve the opinion prediction problem. Different from the existing models which learn the social relation factors including stubbornness and interpersonal influence for each user individually, our model proposes an unified framework by learning the opinion influence of the content information and the influence among social relationships together. We represent each opinion word as a dense vector, and present the composition method for the formation of social opinion context vector by concerning the polarity-related social relation factors (Section 3.2.1). Afterwards, the social opinion context vector is then used to predict one's opinion category (Section 3.2.2). Finally, we present how to learn the proposed model (Section 3.2.3). The graphical description for our proposed model is in Figure 1.

3.2.1 Social Opinion Context Composition with Polarity-related Influence

In this work, we represent each opinion word w as a low-dimensional continuous and real-valued vector $\Phi(w)$, with the dimension d. To obtain the representation of the opinions from u's vth friend, we sum the vectors of all opinion words in the set $C_u^v(i)$, and represent it as $\Phi(C_u^v(i))$. Given all neighboring opinion representations, the social opinion context vector $c_u(i)$ could be obtained by combining them together. Traditional composition methods form the phrase vector by combining word vectors with the weights obtained from the data, or applying the matrix transformation to the concatenation of word vectors (Le and Mikolov, 2014). In this work, we propose a composition method utilizing two social relation factors that have been commonly considered in previous influence models (Das et al., 2014; De et al., 2014). The social relation factors are used to describe the influence among users on the network. The stubbornness factor describes how much a person insists on her/his previous opinion, and the interpersonal influence represents the strength a neighbor has to change one's opinion. Because the interpersonal influence has the linear property (De et al., 2014), our method averages all the word vectors in the neighboring opinion set $C_u(i)$ and one's own previous opinion set $S_u(i-1)$ with the social relation factors. Formally, it is denoted as follows:

$$c_u(i) = \sum_{v=1}^{n(u)} tanh(\alpha_{uv})\Phi(C_u^v(i)) + tanh(\alpha_{u0})\Phi(S_u(i-1)) \tag{1}$$

where $\Phi(S_u(i-1)) = \sum_{k}^{|S_u(i-1)|} \Phi(S_{u,k}(i-1))$. α_{uv} represents the interpersonal influence on user u's opinion from the vth neighbor, and α_{u0} represents u's stubbornness. The two social relation factors are limited between -1 and 1 by using tangent function in Eq (2), which allows both positive and negative influence.

$$tanh(\alpha_{uv}) = \frac{e^{\alpha_{uv}} - e^{-\alpha_{uv}}}{e^{\alpha_{uv}} + e^{-\alpha_{uv}}} \tag{2}$$

The idea of polarity-related influence was firstly proposed by (De et al., 2014), and was proved quite effectively for opinion prediction on social network. The positive influence happens when a user trusts her/his friend, s/he will accept the opinion of her/his friend and express the same one. The negative influence implies that a user gets influenced by her/his friend, but to the opposite direction.

3.2.2 Opinion Prediction

Finally, social opinion context vector could be taken as the features to predict the future opinion category in the output layer. The output layer of our approach is expressed by the following equations.

$$P(o_u(i)|c_u(i)) = softmax(Vc_u(i) + b) \tag{3}$$

The softmax function represents the probability of current vector belonging to the jth class.

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \tag{4}$$

where $V \in \mathbb{R}^{K \times d}$, and $b \in \mathbb{R}^K$. K is the number of opinion categories, and it is set 3 in our model.

3.2.3 Learning

The model is parameterized by the social relation factors α, the word representation $\Phi(w)$ for each opinion word, and the output parameters V, b. The objective function we need to maximize is the log-likelihood of all opinion behavior sequences defined in Eq (5).

$$\mathcal{L}(O) = \sum_{u=1}^{N} \sum_{i=1}^{m(u)} \log P(o_u(i)|C_u(i), S_u(i-1)) \tag{5}$$

We learn the model using the stochastic gradient decent (SGD) algorithm. The dimensionality of the word embedding d is set as 30. During the training phrase, we normalize the gradients if the norm exceeds 1 (Pascanu et al., 2013). The training phrase stops when the training error has a decrease less than 1 or reaches the maximum iteration length of 100. The model is implemented by Theano library (Bastien et al., 2012).

4 Experiment

4.1 Data Collection

We select three well-known electronic products widely discussed on Twitter for the purpose of performance evaluation. They are "Samsung Galaxy", "Xbox" and "PlayStation". For each product, we collect all the tweets containing the product name, such as "Samsung Galaxy", published from 1st March, 2014 to 30th November, 2014 by using the Twitter streaming API[2]. We remove the inactive users with less than 30 tweets and the over active users with more than 1000 tweets. We also collect the following relationships among the users, and further construct the user network for each individual product.

Table 1 summarizes the statistics of the datasets. The "# of users" and the "# of avg friends" describe the size of the network. During each communication round, not all of a user's friends provide the suggestions, and the friends who actually post tweets and influence the user's future opinion are the active friends. Each communication round starts after a user posts a tweet, and ends when the user updates her/his opinion with a new tweet. Therefore, we define the average number of active friends by "# of avg active friends". The active level is denoted as $("\#of\,avg\,active\,friends")/("\#of\,avg\,friends")$. It implies the interests of the users on the discussion of a product. From the statistics, we observe that the products "Samsung Galaxy" and "Xbox" are actively discussed by the users with the 39%, 37% active level respectively. However, the communication on the topic "PlayStation" is not as frequent as the communication on the other two topics.

---

[2]https://dev.twitter.com/streaming/overview

Table 1: Network statistics.

Topic	Samsung Galaxy	Xbox	PlayStation
# of users	8921	4358	5158
# of avg friends	14.42	9.58	11.83
# of avg active friends	5.65	3.58	3.33
active level	**0.39**	**0.37**	**0.28**

4.2 Opinion Processing

Many approaches has been proposed to analyze the sentiment from the text (Hu and Liu, 2004; Pang and Lee, 2008; Mukherjee et al., 2012). However, all these methods fail to explore the reason why people express or change their opinions. In our work, we take the sentiment of tweets as premise and discover the social influence during the communication. The Vader method recently proposed by (Hutto and Gilbert, 2014) has been proved better than typical state-of-art benchmarks on analyzing the sentiment of tweets with 96% accuracy on the Twitter dataset. With the constructed twitter-specific sentiment lexicon, Vader method considers the grammatical and syntactical rules to access the sentiment scores of tweets. We utilize the Vader method to score the sentiment of each tweet and to tag the sentiment category. The tweet with positive sentiment score is tagged as positive, the one with negative score is tagged as negative, and the one with zero score is tagged as neutral.

Additionally, we obtain all the opinion words with the following rules. For each tweet, all the opinions words included in the twitter-specific sentiment lexicon (Hutto and Gilbert, 2014) are extracted. If an opinion word follows a negation word, we retain the phrase "not"+"opinion word" instead of the original opinion word. For example, the opinion word extracted from the tweet "I don't like the Samsung Galaxy S6." is the phrase "not like". For the tweets only stating the facts without expressing an opinion, we use the word symbol "NeuW" to represent them. To alleviate the word sparsity, we only keep the opinion words that occur more than 50 times in the whole dataset and replace the infrequent opinion words with the corresponding symbols. The positive opinion words are replaced with the symbol "PosW", and the negative opinion words are replaced with the symbol "NegW". Finally, the numbers of the remaining opinion words for the topic "Samsung Galaxy", "Xbox", and "PlayStation" are 880, 1146 and 505, respectively.

4.3 Experimental Set-up

We compare the proposed model CIM with four baseline models, i.e., the DeGroot model, the Flocking model, the Voter model and the AsLM model. These models have different assumptions for the opinion formation process. To be fair, all baseline models incorporate the factor of personal stubbornness. It means that all models take the influence from one's previous opinion into account. For the DeGroot model (Acemoglu and Ozdaglar, 2011), the Flocking model (Hegselmann and Krause, 2002) and the AsLM model (De et al., 2014), each tweet is represented as a continuous sentiment score. For the Voter model with the assumption of the majority adoption (Krapivsky and Redner, 2003), each tweet is summarized by its opinion category. To further verify the effectiveness of the content information, we develop another influence model Content_SVM which is implemented with LIBSVM (Chang and Lin, 2011). The model trains SVM classifiers individually for each user by taking all the neighboring opinion words and the opinion words in one's previous tweet as features. To be consistent with the linear influence assumption, the linear kernel is used in SVM training process. The parameters of each model are set for their best performances experimentally.

We split the data into the training dataset and test dataset according to the posting time. The training dataset is constructed by using the data before the $m(u)-1$ timestamp for each user u. With the influence model learned from the training set, we predict the last opinion for each user.
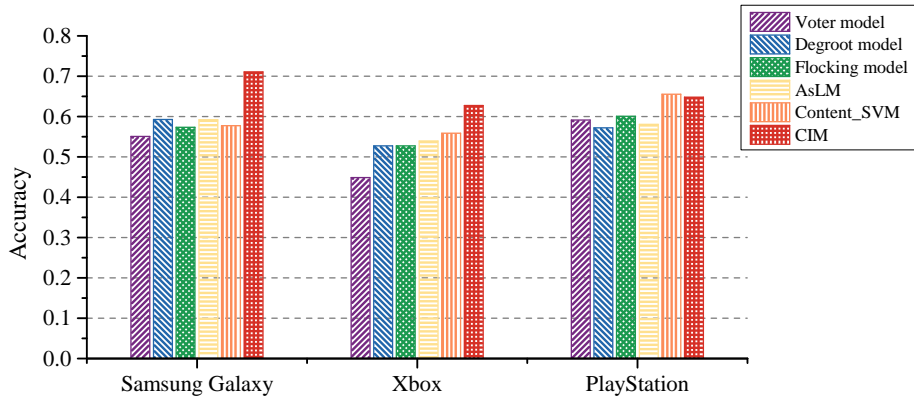
Figure 2: Performances on opinion prediction

4.4 Opinion Prediction Performances

We first evaluate the prediction accuracy for all the models. The results are displayed in Figure 2.

$$Accuracy = \frac{the\ number\ of\ correctly\ predicted\ users}{the\ number\ of\ all\ users}$$

The content-based models (Content_SVM and CIM) almost outperform the baseline methods in all three topics, which verifies that employing the detailed content information is more effective than only using the opinion statuses for opinion behavior prediction.

Meanwhile, CIM performs consistently much better than all baseline methods on the topics of "Samsung Galaxy", and "Xbox". Compared with other methods which only learns the social relation factors from opinion behaviors for each user individually, CIM encodes the semantic information into the dense vectors of the opinion words through learning from the opinion behaviors of all users. The good performance of CIM demonstrates its better ability to capture two types of opinion influence components including opinion influence of the opinion words and social relation factors together. However, CIM has a slightly lower accuracy compared with the best competitor on the topic "PlayStation". It can be attributed to the lower active level of users on the PlayStation than those on the other two topics. The insufficient communication histories over the network make it difficult to learn the actual influence of opinion words for opinion prediction, and may even harm the results.

4.5 The Effect on Opinion Category

For a more detailed analysis, we further evaluate the ability of CIM on predicting different opinion categories. We present the distributions of three opinion categories in both the training dataset and the test dataset in Table 2. The F1 score which considers both precision and recall, is used as the measurement on each opinion category. The experimental results are included in Table 3.

On the topics of "Samsung Galaxy" and "Xbox" with the active communication environment, CIM still has a significant improvement concerned with the evaluation metrics on all the three opinion categories. Specifically, the improvements compared with the best competitors on the positive opinion prediction and the negative opinion prediction are 17.7%, 21.5% for the topic "Samsung Galaxy" and 11.5%, 20.3% for the topic "Xbox" respectively. Compared with predicting the neutral opinions, forecasting the positive and negative opinions is more useful for companies to understand the customer needs and the brand reflection.

On the topic "PlayStation" with the relatively inactive communication, best performances of different evaluation metrics are obtained by different models. CIM performs well on the prediction of the positive and neutral opinions but poorly on the prediction of negative opinions. It reveals that the weakness of CIM is mainly on learning the negative opinion formation process when the communication is insufficient. We also note that the Voter model which performs poorly on the other two topics has better results on the "PlayStation" topic. Different from influence models based on the interpersonal influence,

Table 2: Opinion category statistics.

Topic	Samsung Galaxy		Xbox		PlayStation	
	Training set	Test set	Training set	Test set	Training set	Test set
% of negative opinion	11.05	14.61	16.33	6.81	11.99	19.73
% of positive opinion	19.96	19.95	41.56	26.88	25.03	19.28
% of neutral opinion	65.43	65.41	42.11	66.31	62.98	60.99

Table 3: Performances on three opinion categories.

Topic	Samsung Galaxy			Xbox			PlayStation		
	F1_Pos	F1_Neg	F1_Neu	F1_Pos	F1_Neg	F1_Neu	F1_Pos	F1_Neg	F1_Neu
Degroot	0.4950	0.1932	0.6913	0.5185	0.1935	0.6035	0.2531	0.1405	0.7064
Flocking	0.4449	0.2677	0.6780	0.4469	0.2069	0.6240	0.3513	**0.3711**	0.7125
AsLM	0.5812	0.2139	0.7028	0.5597	0.2298	0.6293	0.3210	0.2025	0.7338
Voter	0.4826	0.1762	0.6246	0.4637	0.1694	0.4709	**0.5655**	0.2688	0.6782
Content_SVM	0.4918	0.1436	0.6732	0.5972	0.2004	0.6106	0.5616	0.3410	0.7458
CIM	**0.6842**	**0.3253**	**0.7787**	**0.6658**	**0.2765**	**0.6677**	0.5568	0.1521	**0.7518**

the Voter model assumes that one will accept the mainstream view of her/his neighbors as the future opinion. The results indicate that when neighboring messages are insufficient, the group influence of all neighbors dominates. It motivates us to utilize the group influence with the interpersonal influence together for benefiting the opinion behavior prediction in the insufficient communication situation.

4.6 Analysis of Wording for Influential Users

With the learned model, the companies could get the insights into how to become an influential voice on the social media by improving their wordings. We analyze different expressions used by users with different social opinion influence degrees in the network. Based on the learned interpersonal influences, we calculate the influence strengths of Twitter users by averaging their outgoing influence strengths on their followers. Based on the influence strengths, we divide users into three groups. The users with influence strengths more than 0.5 are categorized as the positively influential users. The users with influence strengths less than -0.5 represent the negatively influential users. The remaining are regarded as the ordinary users with little influence.

We then extract the high frequent words from the users in different influence groups. The results show that the positively influential users more likely utilize the words describing the facts, e.g., "security", "special" and impress". However, the tweets posted by strong negative influential users are more emotional with the words like "Woo", "Wow" or the emoticons "o_o". The analysis indicates that the detailed information about the products tends to make positive effects, while heavily emotional expressions may annoy people and influence them in the opposite direction.

5 Conclusions

In this paper, we propose to characterize the users' tweets with detailed opinion content instead of discrete opinion categories or continuous scores. To the best of our knowledge, this is the first attempt to incorporate the content information into opinion behavior modeling. Existing models only learn the social relation factors from the pre-defined opinion sequences. Differently, our proposed model based on the feed-forward neural network framework is capable of learning the opinion word representations which encodes the actual influence of the opinions words, and learn the two social relation factors from the opinion behaviors of all users. The experiments conducted on the Twitter dataset demonstrate the effectiveness of our proposed model on the opinion prediction. We also examine the expressions of users with different influence degrees, which could provide useful information for companies to manage their accounts. Based on the current work, we will further combine more influencing factors including the personal interests and group influence in the future model.

Acknowledgments

References

Acemoglu, D. and Ozdaglar, A. (2011). Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1(1):3–49.

Anagnostopoulos, A., Kumar, R., and Mahdian, M. (2008). Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 7–15. ACM.

Bastien, F., Lamblin, P., Pascanu, R., Bergstra, J., Goodfellow, I., Bergeron, A., Bouchard, N., Warde-Farley, D., and Bengio, Y. (2012). Theano: new features and speed improvements. *arXiv preprint arXiv:1211.5590*.

Bengio, Y., Ducharme, R., Vincent, P., and Janvin, C. (2003). A neural probabilistic language model. *The Journal of Machine Learning Research*, 3:1137–1155.

Chang, C.-C. and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

Chazelle, B. (2012). Natural algorithms and influence systems. *Communications of the ACM*, 55(12):101–110.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Clifford, P. and Sudbury, A. (1973). A model for spatial conflict. *Biometrika*, 60(3):581–588.

Collobert, R. and Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, pages 160–167. ACM.

Das, A., Gollapudi, S., and Munagala, K. (2014). Modeling opinion dynamics in social networks. In *Proceedings of the 7th ACM international conference on Web search and data mining*, pages 403–412. ACM.

De, A., Bhattacharya, S., Bhattacharya, P., Ganguly, N., and Chakrabarti, S. (2014). Learning a linear influence model from transient opinion dynamics. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 401–410. ACM.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121.

Galam, S. (2002). Minority opinion spreading in random geometry. *The European Physical Journal B-Condensed Matter and Complex Systems*, 25(4):403–406.

Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3).

Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM.

Hutto, C. and Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International AAAI Conference on Weblogs and Social Media*.

Krapivsky, P. and Redner, S. (2003). Dynamics of majority rule in two-state interacting spin systems. *Physical Review Letters*, 90(23):238701.

Le, Q. V. and Mikolov, T. (2014). Distributed representations of sentences and documents. *arXiv preprint arXiv:1405.4053*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

Mukherjee, S., Bhattacharyya, P., et al. (2012). Sentiment analysis in twitter with lightweight discourse analysis. In *COLING*, pages 1847–1864.

Pang, B. and Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2):1–135.

Pascanu, R., Gulcehre, C., Cho, K., and Bengio, Y. (2013). How to construct deep recurrent neural networks. *arXiv preprint arXiv:1312.6026*.

Socher, R., Lin, C. C., Manning, C., and Ng, A. Y. (2011). Parsing natural scenes and natural language with recursive neural networks. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 129–136.

Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., and Qin, B. (2014). Learning sentiment-specific word embedding for twitter sentiment classification. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 1555–1565.

Yildiz, E., Acemoglu, D., Ozdaglar, A. E., Saberi, A., and Scaglione, A. (2011). Discrete opinion dynamics with stubborn agents. *Available at SSRN 1744113*.

Yildiz, E., Ozdaglar, A., Acemoglu, D., Saberi, A., and Scaglione, A. (2013). Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation*, 1(4):19.

Zanzotto, F. M., Korkontzelos, I., Fallucchi, F., and Manandhar, S. (2010). Estimating linear models for compositional distributional semantics. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 1263–1271. Association for Computational Linguistics.