Intersubjectivity and Sentiment: From Language to Knowledge

Lin Gui,[1] Ruifeng Xu,[1][†] Yulan He,[2] Qin Lu,[3] Zhongyu Wei[4]
[1]Laboratory of Network Oriented Intelligent Computation,
Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China
[2]School of Engineering and Applied Science, Aston University, United Kingdom
[3]Department of Computing, the Hong Kong Polytechnic University, Hong Kong
[4]Computer Science Department, The University of Texas at Dallas, Texas 75080, USA

Abstract

Intersubjectivity is an important concept in psychology and sociology. It refers to sharing conceptualizations through social interactions in a community and using such shared conceptualization as a resource to interpret things that happen in everyday life. In this work, we make use of intersubjectivity as the basis to model shared stance and subjectivity for sentiment analysis. We construct an intersubjectivity network which links review writers, terms they used, as well as the polarities of the terms. Based on this network model, we propose a method to learn writer embeddings which are subsequently incorporated into a convolutional neural network for sentiment analysis. Evaluations on the IMDB, Yelp 2013 and Yelp 2014 datasets show that the proposed approach has achieved the state-of-the-art performance.

1 Introduction

Sentiment analysis [Pang *et al.*, 2002] becomes a hot topic in natural language processing research. It aims to classify the polarity of a given text either at the sentence level or at the document level. In this paper, we focus on sentiment classification at the document-level. Traditionally, document-level sentiment classification methods trained supervised classifiers from documents labeled either positive or negative by using the bag-of-words assumption [Gui *et al.*, 2014]. Other methods explored the use of latent topics for sentiment classification [He *et al.*, 2013]. Most recently, there have been growing interests in using deep learning for sentiment classification. With the help of deep neural networks, such as Convolutional Neural Network [Kim, 2014; Hu *et al.*, 2014], Recursive Neural Network [Socher *et al.*, 2013] or Recurrent Neural Network [Tang *et al.*, 2015a], sentences and documents are better represented such that performance in sentiment classification are improved.

Generally speaking, these methods learn a function that maps a given text into certain class. That is, they aim to map the word sequences or latent representations to some sentiment classes. Such mapping is done mostly at the surface

level of the text using lexical information of the corresponding language.

However, text is intended to convey meaning that is latent in the script beyond the surface form of the language. The study on sociology theory showed that there is a gap between the surface form of a language and the corresponding abstract concepts, named to as intersubjectivity [Dunbar and Dunbar, 1998]. Intersubjectivity refers to a shared perception of reality among members in their social network. Looking through the history of language development, it is not hard to see that common knowledge is not built from language construction alone. Rather, it is evolved from intersubjective understanding. Intersubjectivity suggests that the meaning of a word or a phrase is not encoded in the surface form of that language as a mapping from a term to an object or a subject. Rather, it is a commonly accepted conceptualization by a society sharing the same language. In other words, the meaning of a term is assigned by its writers.

In sentiment analysis, intersubjectivity plays a very important role. In product reviews, review writers(author for short) always describe a product with different expressions and sentiments. The words chosen often reflect an author's point of view. For example, the fans of Sony camera may ridicule Zeiss, which tend to weigh more but of higher in picture quality, using words like a "hammer". Users of Zeiss, on the other hand, describe Sony cameras as "toys", which implies that Sony cameras may look good but unprofessional. Inspired by intersubjectivity, in this study, we make use of two kinds of information for sentiment analysis. The first one is the relationship between author and the words they use. The second one is the relationship between words and the polarities they are associated with.

More specifically, we propose an intersubjectivity based network embedding method to map authors into a continuous vector space. The basic idea is to use a naïve Bayes based method to link authors to the words they use. We also link words with polarities, which stands for subjectivity. Then, we utilize a network embedding method to identify similar stances through shared use of terms by mapping authors and the review words into a continuous vector space. The continuous vector of an author can reflect the degree of subjectivity and the polarities are naturally embedded. We then incorporate the learned author vectors as features into a convolutional neural network(CNN) for sentiment classification. Evalua-

---

[†]corresponding author: xuruifeng@hitsz.edu.cn

tions show that the proposed method achieves the state-of-the-art results on IMDB, Yelp2013 and Yelp2014 datasets. The main contributions of this paper include:

1. The first work to introduce the concept of intersubjectivity from sociology to sentiment analysis.

2. A proposed naïve Bayes based method to construct a heterogeneous network with reviwers, phrasal expressions and polarities, which aims to capture intersubjectivity in subjective terms.

3.A proposed sentiment analysis method based on the intersubjectivity network, which achieves the state-of-the-art performance on the IMDB, Yelp2013 and 2014 datasets.

The rest of this paper is organized as follow: Section 2 briefly introduces related works in sentiment analysis and review modeling. Section 3 presents our approach including the construction of intersubjectivity network and sentiment classification built upon the constructed network. Section 4 gives experimental setup and performance evaluation. Section 5 concludes the paper.

2 Related works

Sentiment analysis or opinion mining [Pang *et al.*, 2002; J *et al.*, 2015; He *et al.*, 2012] has become a hot research topic in recent years. It aims to determine the polarity of a given piece of text. Other works also aim to detect the key components from text such as opinion holders and targets [Xu *et al.*, 2015].

Simple approaches in sentiment classification relied on rules and lexicons [Taboada *et al.*, 2011]. However, these methods usually required heavy manual processing. Machine learning based methods use supervised classification or regression [Pang and Lee, 2005] to train models from polarity labeled text. Besides unigram word features, other features such as word n-grams, part-of-speech tags, negation words, modifiers, affective attributes, etc., are also used in sentiment classifiers [Abbasi *et al.*, 2011; Xia *et al.*, 2013].

More recently, deep learning based methods have been shown effective in many text classification tasks including sentiment analysis [Tang *et al.*, 2014a; Xu *et al.*, 2014]. Most deep learning based methods aim to learn the continuous representation of text, such as words, phrases, sentences and even documents. Representation learning is typically carried out at two levels, namely the basic word level [Mikolov *et al.*, 2013; Tang *et al.*, 2014b] or the compositional sequence level [Glorot *et al.*, 2011; Kalchbrenner *et al.*, 2014].

Apart from text, author information can also be used in sentiment classification. [Gao *et al.*, 2013] designed author-specific features to capture author leniency. [Dong *et al.*, 2014] incorporated textual topics and author-word factors into supervised topic modeling. [Hovy, 2015] utilized the demographic information in sentiment analysis. [Tan *et al.*, 2011] and [Hu *et al.*, 2013] utilized author-author relationship for Twitter sentiment analysis. In summary, existing methods used two types of author information, namely (1) shared personal profiles such as ages, gender etc. as shared background and (2) text written by writers as lexical features for statistical similarity measures as topic words. This similarity measure is not intersujectivity based as there is no mu-

tual reinforcement of writers and their written text.

3 Our Approach

In this work, we propose an author modeling method which learns reviewer embeddings from an intersubjectivity network built from review documents which links authors, terms and polarities in a unified network. As will be shown in our experiments, author embeddings learned offer a better explain ability and incorporating reviewer embeddings into a CNN gives the state-of-the-art results on three product review datasets. Our approach is inspired by the theory of intersubjectivity on shared conceptualization through their social interactions. In this section, we first discuss how to construct an intersubjectivity network from text and author information, and then propose a network embedding method to model authors, words and sentiments in the same embedding space, and finally incorporate the learned representations into a CNN for sentiment classification at the document level.

3.1 Intersubjective Network

The key problem here is how to construct an intersubjectivity network. We solve the problem by two subtasks. The first subtask is to construct a subjectivity network, which aims to capture subjective terms in text. The second one is to construct an author network, which aims to capture the relationship among different individuals (authors of review documents).

Subjective terms here refer to words or phrases carrying positive or negative sentiments. Since the extraction of subjective terms itself is a challenging research problem, we simplify the task by extracting word bigrams from review documents as candidate subjective terms. We then use naïve Bayes based log-count ratio, which has been shown helpful in sentiment analysis [Wang and Manning, 2012], to assign each candidate bigram a weight corresponding to the associated polarity.

Assuming that there is a training set D with n labeled samples, $D = \{s_1, s_2, \ldots, s_n\}$, the label is $y_i = \{1, -1\}, i = 1, 2, \ldots, n$, the vocabulary of bigrams is V, and $V = \{v_1, v_2, \ldots, v_k\}$. If we use c_j^i to represent the number of occurrences of v_j in s_i, we can define the probability of v_j assigned with different labels as:

$$p_j(1) = \alpha + \sum_{i:y_i=1}^{n} c_j^i \tag{1}$$

$$p_j(-1) = \alpha + \sum_{i:y_i=-1}^{n} c_j^i \tag{2}$$

Here, α is a smoothing parameter with a small positive number to avoid zero probabilities. Then the log-count ratio is defined as:

$$r_j = log\left(\frac{p_j(1)}{\sum_{m=1}^{k} p_m(1)}\right) / log\left(\frac{p_j(-1)}{\sum_{m=1}^{k} p_m(-1)}\right) \tag{3}$$

Note that Equation (3) is only suitable for binary classification. We need to extend Equation (3) to handle more general

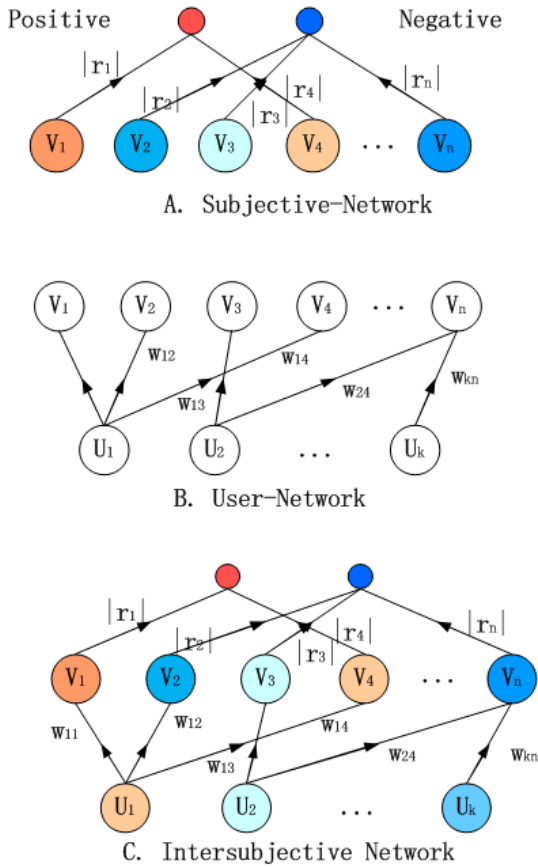Figure 1: Intersubjectivity network

cases with l different polarity ratings. The first step is to map l different polarity ratings to [-1,1]. Since different tasks in sentiment analysis require different polar intensity, we do not give a specific mapping function here. Instead, we define the properties of a mapping function f:

1.f is monotonous;

2.f maps the most positive label to 1 and the most negative label to -1;

Equation (3) can now be extended to a more flexible form:

$$r_j = \sum_{y_i} f(y_i) log \left(\frac{p_j(y_i)}{\sum_{m=1}^{k} p_m(y_i)} \right) \quad (4)$$

We use r_j to indicate the relationship between a bigram v_j and a polarity. If v_j occurs in positive reviews more often, then r_j will be positive. Otherwise, r_j will be negative. If a bigram is more closely related to a certain polarity label, the absolute value of r_j will be larger.

We then extract the top k positive and the top k negative candidate bigrams based on the values of $|r_j|$ as the most relevant subjective terms to build the subjective network as shown in Fig.1 A. Here, v_j is linked with its corresponding polarity label, and the weight of the edge between v_j and the label is $|r_j|$.

The next step is to add the author information into the subjective network. We create a link between an author and a

subjective term if the author used the term in his reviews. The corresponding link weight is set as the occurrence count of the term in his reviews. For instance, for an author u_i and a subjective term v_j, the weight between them is w_i^j, the number of times v_j appears in the reviews written by u_j. Since the values of $|r_j|$ and w_i^j are not in the same scale, we deploy a Gaussian based standardization. Assuming the mean value and variance of $|r_j|$ are μ_r and σ_r^2, the mean value and variance of w_i^j are μ_w and σ_w^2. The standardization of $|r_j|$ and w_i^j are:

$$\overline{|r_j|} = \frac{|r_j| - \mu_r}{\sigma_r} \quad (5)$$

$$\overline{w_{ij}} = \frac{w_{ij} - \mu_w}{\sigma_w} \quad (6)$$

Here, $\overline{|r_j|}$ and $\overline{w_{ij}}$ should follow Gaussian distribution, and we use the corresponding Gaussian probability as the smoothed weight of the edge in the intersubjectivity network.

3.2 Network Embedding and Author Representation Learning

The next problem is how to represent the authors in the intersubjectivity network. In recent years, representation learning has become a hot topic in natural language processing research. The basic idea is to use a continuous vector to represent words, sentences or documents. Here, we deploy a representation learning method to learn a continuous representation for authors. Given a large network $G = (V, E)$, where V is a set of vertexes which contains authors, terms, and polarities, and E is a set of edges, which stands for the relationship between vertices. Intersubjectivity Network Embedding aims to represent each vertex $v \in V$ in a lower-dimensional space R^d by learning a function $f : V \rightarrow R^d$, where $d \ll |V|$.

We first need to decide what to capture in the representations of the vertices. Since there are three different types of vertices, the intersubjectivity network is actually a heterogeneous network. As such, in the represented vector space, vertices of different types should have dissimilar distributions. This is because if the linked vertices come from different types, they should not be similar to each other in the representation vector space. In other words, this requires a more suitable measure: authors sharing similar subjective terms should be similar to each other. So, we propose a probability based on conditional representation as:

For any vertex v_i and its corresponding representation in the vector space, denoted as α_i, $\alpha_i \in R^d$, let the neighbor linked to v_i as v_i' with representation α_i'. The conditional probability $p(v_j|v_i)$ can be defined by softmax as:

$$p(v_j|v_i) = \frac{exp(\alpha_j'^{T} u_i)}{\sum_{k=1}^{|V|} exp(\alpha_k'^{T} u_i)} \quad (7)$$

Obviously, given the definition above, two author vertices share similar subjective terms should have high conditional probability. This is further defined as the second-order proximity. In order to preserve the second-order proximity, we

need to make the conditional distribution of $p(v_j|v_i)$ be close to its empirical distribution. The empirical distributions can be derived from weights in the intersubjectivity network. We define the objective function of our representation learning algorithm as:

$$O = - \sum_{(i,j) \in E} w_{ij} log\, p\,(v_j|v_i) \quad (8)$$

By minimizing the objective function O, we can represent every v_i with a d-dimensional vector α_i'. It should be noted that O includes all vertices in the network, not just the authors in Equation 7. In this work, we hypothesize that the terms with the same polarity will be similar to each other in the low dimensional representation space, and authors sharing similar terms will be similar to each other too.

In this study, we use a network embedding method to find people having similar sentiment through their shared use of words using a network model. Our method, similar to word embedding [Mikolov *et al.*, 2013], defines the context of an author, by other authors who share similar terms. Our algorithm makes sure that, in the embedding result, authors and terms are separated modeled as networks. Authors sharing similar terms are identified through mutual interaction of two networks dynamically for shared content of different authors. This makes our method original.

The next problem is how to utilize the author representation learned from the intersubjectivity network to improve the performance of sentiment analysis.

3.3 Incorporating Author Representations into CNN for Sentiment Classification

Theoretically speaking, the representation of authors can be embedded into any sentiment classification methods. Since CNN have shown good performance in previous works, we simply take CNN as the learning algorithm for sentiment analysis.

The architecture of the neural network is shown in Fig. 2. The input layer is the continuous word representations learned by word2vec [Mikolov *et al.*, 2013]. Let a document containing l words in the training data with label y as: $s = \{w_1, w_2, w_3, \ldots, w_l; y\}$.

Here, w_i is the word representation of the i-th word in the sentence sequence. The label y denotes the polarity, which can take the values of -1 or 1 as binary class labels. Assuming that the dimension of the word representation is d and can be represented as: $(v_{i1}, v_{i2}, v_{i3}, \ldots, v_{id})^T$. Let the concatenation of word sequence from i-th to j-th words be denoted as $w_{i:j}$, a convolution operation involving a filter $m \in R^{hd}$ which have the window of size h, can extract a feature on $w_{i:i+h-1}$ in the convolutional layer by:

$$c_i = f\,(xw_{i:i+h-1} + b) \quad (9)$$

Here,b is the bias and f is a non-linear mapping function such as the logistic function. We then use Equation (9) to extract features from a sentence s as: $(c_1, c_2, c_3, \ldots, c_{l-h+1})$.

A max pooling operation is used to extract the most relevant feature, the one with the highest value in the max pooling layer. Finally, all of the features are connected with a label by
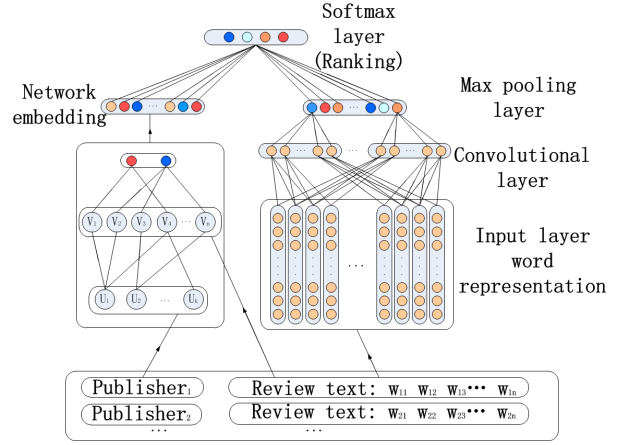

Figure 2: Architecture of Intersubjectivity Embedded CNN.

a softmax layer to train a classifier for sentiment classification.

In order to compose with the author representations learned from intersubjectivity embedding results, we join the author embeddings with the features in the max pooling layer. Specifically, the training of intersubjectivity embeddings uses the top 20k terms and the negative sampling method [Tang *et al.*, 2015c] and produces a representation for each author. Then for each training sample, we use the author representation as an additional feature to be combined with other features in the max pooling layer. In the training process, weights associated with both intersubjectivity embedding features and the convolutional features are updated simultaneously.

4 Evaluations and Discussions

4.1 Experimental Setup

We evaluate our algorithm on three product review datasets, including IMDB [Diao *et al.*, 2014] and the Yelp Dataset Challenge in 2013 and 2014. Statistics with respect to authors, products and reviews are given in Table 1:

Dataset	#Authors	#Products	#Reviews
IMDB	1,310	1,635	84,919
Yelp2013	16,31	1,633	78,966
Yelp2014	4,818	4,194	231,163

Table 1: The statistics of three datasets

IMDB is labeled with 10 different levels of sentiment, score 1 for the most negative and score 10 for the most positive. Yelp 2013 and Yelp 2014 are labeled with 5 different levels of sentiment, score 1 for the most negative and score 5 for the most positive. All the datasets are pre-processed by following [Tang *et al.*, 2015b].

We use accuracy (ACC) and root mean square error (RMSE) as the evaluation metrics. Let the predicted label of i-th testing sample be $predicted_i$, the actual label of i-th testing sample be $actual_i$, and the size of the test set be N. The metrics can be calculated as below:

method	IMDB		Yelp2013		Yelp2014	
	ACC	RMSE	ACC	RMSE	ACC	RMSE
Paragraph vector	0.341	1.814	0.554	0.832	0.564	0.802
RNTN+recurrent	0.400	1.764	0.574	0.804	0.582	0.821
CNN	0.440	1.464	0.596	0.770	0.610	0.762
JMARS*	—	1.773	—	0.985	—	0.999
UPNN	0.441	1.602	0.596	0.784	0.608	0.764
Our Approach	**0.476**	**1.396**	**0.623**	**0.714**	**0.635**	**0.690**

Table 2: Comparison with existing methods in terms of ACC and RMSE.

$$ACC = \frac{\sum_{predicted_i=actual_i} 1}{N} \qquad (10)$$

$$RMSE = \sqrt{\frac{\sum_i \left(predicted_i - actual_i\right)^2}{N}} \qquad (11)$$

4.2 Comparison to other Methods

We compare our proposed method with a number of existing methods as listed below:

1. Paragraph vector for document modeling [Le and Mikolov, 2014], used for sentiment classification in [Tang *et al.*, 2015b];

2. Recursive Neural Tensor Network (RNTN) for sentence modeling [Socher *et al.*, 2013], incorporated into recurrent neural networks for document modeling in [Tang *et al.*, 2015b];

3. Convolutional Neural Network (CNN) for sentiment classification [Kim, 2014];

4. Jointly Modeling Aspects, Ratings and Sentiments (JMARS) [Diao *et al.*, 2014], a topic modeling method which leverages on authors, product aspects and sentiments for sentiment classification;

5. User Product Neural Network (UPNN) [Tang *et al.*, 2015b], which incorporates user and product information using CNN.

Note that other than JMARS, which used topic modeling approach, all others are neural network based method and UPNN which also used both author and content information through CNN method was the state-of-the-art method for all three datasets. Table 2 shows the performance evaluation. Although, JMARS leverages information on authors, product aspects and sentiments in a unified topic model, it is outperformed by neural network based methods. Our method also outperforms the state-of-the-art system UPNN by 0.030 for average ACC and 0.117 for RMSE with p-value less than 0.01 indicating significant improvement.

4.3 Further Analysis on Author Modeling

Both UPNN and our approach are built upon CNNs. In order to understand the benefit of modeling authors using intersubjectivity networks, we conduct a set of experiments to compare CNN with different author modeling methods. Results are shown in Table 3. The basic CNN without any author modeling, labeled as CNN, is used as the baseline for comparison. We experiment with author modeling following UPNN [Tang *et al.*, 2015b], where each author is distributed into a low dimensional space by polarity. The method is labeled as DL (Distributed Label). Our own author modeling based on the intersubjectivity networks is labeled as ISN.

Dataset	Method	ACC	RMSE
	CNN	0.440	1.464
IMDB	CNN+DL	0.464	1.451
	CNN+ISN	0.476	1.396
	CNN	0.596	0.770
Yelp2013	CNN+DL	0.607	0.747
	CNN+ISN	0.623	0.714
	CNN	0.610	0.762
Yelp2014	CNN+DL	0.625	0.744
	CNN+ISN	0.635	0.690

Table 3: The comparison with or without author modeling

From Table 3 we can see that the performance improvement of CNN+DL compared to the baseline CNN is only modest for all three sets of data in terms of both measures of ACC and RMSE. However, the performance improvement of CNN+ISN is much more significant compared to the CNN baseline. This shows that author modeling using DL can only capture author characteristics in a coarse manner. On the other hand, intersubjectivity networks model authors and their shared subjective terms in a more unified manner. Both authors and subject terms are mapped into the same embedding space. As has been shown previously, under the optimal value of the objective function defined in Equation (8), authors sharing similar subjective terms should be similar to each other, and terms with similar polarities should be mapped to nearby locations as well.

We also study in details the top *k* most positive and negative authors identified by our intersubjectivity network. Due to space limit, only results from the Yelp 2013 dataset is presented here. Recall that we have three types of nodes in the intersubjectivity network: the author nodes, the term nodes, and the polarity nodes. We observe that the top 100 most similar nodes identified from the heterogeneous intersubjectivity network are all author nodes. We show the statistics of the top 5 most positive and most negative authors in Table 4 and Table 5, respectively. Note that authors in the same polarity group share very similar review patterns. For example, authors in the top 5 most positive group tend to give very positive ratings (either score 4 or 5) in most of their reviews, while authors in the top 5 negative group tend to spread out their ratings across all reviews.

author	Polarity Similarity	Review History(Times)					Average
		1	2	3	4	5	
1	0.676	1	2	12	34	25	4.08
2	0.675	0	3	2	35	17	4.16
3	0.673	6	11	12	13	5	3.00
4	0.672	0	3	8	2	15	4.04
5	0.664	1	1	2	21	11	4.11

Table 4: The top 5 most positive authors identified in the intersubjectivity network.

author	Polarity Similarity	Review History(Times)					Average
		1	2	3	4	5	
1	0.709	2	3	4	4	5	3.38
2	0.631	6	7	5	3	0	2.24
3	0.624	1	2	5	7	4	3.68
4	0.603	6	5	16	6	0	2.66
5	0.595	4	5	3	6	3	2.95

Table 5: The top 5 most negative authors identified in the intersubjectivity network.

Since the top 100 most similar nodes identified from the intersubjectivity network are all author nodes, we speculate that our proposed author representation learning method has the capability to distinguish between terms and authors. Fig. 3a shows the learned author and term embeddings in our intersubjectivity network. We also plot the representation by DL in Fig. 3b. Here, we use the PCA (Principal Component Analysis) to obtain the top two dimensions of the intersubjectivity network embedding results. Since the mean value of review ratings is 3.81, we mark authors or terms as positive if their average scores are over 3.81 and negative otherwise for visualization. It is shown that the authors and terms are separated well, and positive terms also differ from negative terms. However, we notice that positive and negative authors seem to overlap more in ISN than in DL as if DL has better distinguishing power. The fact is, however, our ability to distinguish terms are useful in the sentiment classification task. In other words, we also model authors based on their review behaviors (authors sharing similar subjective terms are close to each other), not based on their ratings. Using a simple thresholding method based on ratings to mark authors with different color in Fig. 3a may be misinterpreted as the ISN method has less distinguishing power as Fig. 3b seems to separate positive and negative authors better. However, the DL method models authors purely based on their review ratings. Authors who give positive review ratings most of time can also wrote negative reviews occasionally. As such, classifying authors as positive or negative based on their review ratings solely and ignoring the content generated from them can feed wrong information to document-level sentiment classifiers. Since ISN models authors and their shared subjective terms in a unified manner, it indeed gives better sentiment classification results compared to DL. This shows that it is more appropriate to model authors based on their content rather than their ratings.
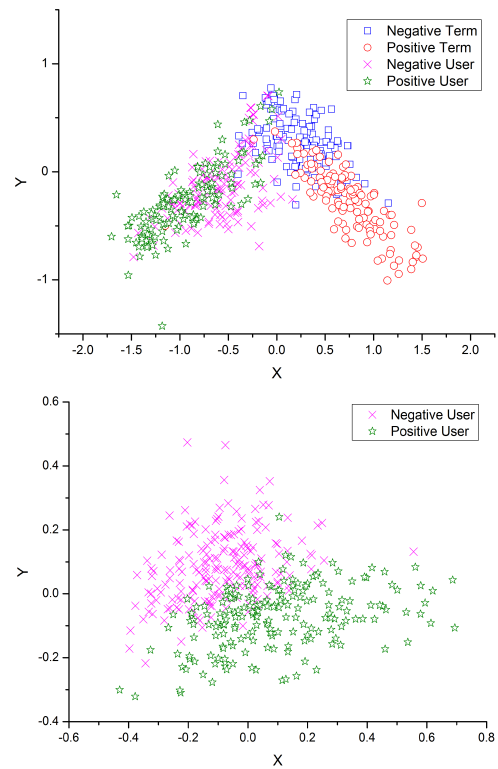

Figure 3: The distribution of terms and authors in the embedding space(a. Top-by ISN;b. Bottom-by DL)

5 Conclusion

In this paper, we present a new author modeling method for sentiment classification based on the notion of intersubjectivity. More specifically, we include authors, terms and polarities in a unified network. Both terms and authors are mapped into the same embedding space. The learned author representation is then incorporated into a CNN-based neural network for sentiment classification. Experimental results show that our proposed author embedding learning method not only offers a better semantic interpretation incorporating intersubjectivty but also improvement to sentiment classification to achieve the state-of-the-art results on the relevant datasets.

Acknowledgments

References

[Abbasi et al., 2011] Ahmed Abbasi, Stephen France, Zhu Zhang, and et al. Selecting attributes for sentiment classification using feature relation networks. *Knowledge and Data Engineering, IEEE Trans. on*, 23(3):447–462, 2011.

[Diao *et al.*, 2014] Qiming Diao, Minghui Qiu, Chao-Yuan Wu, and et al. Jointly modeling aspects, ratings and sentiments for movie recommendation (jmars). In *SIGKDD*, pages 193–202, 2014.

[Dong *et al.*, 2014] Li Dong, Furu Wei, Ming Zhou, and et al. Adaptive multi-compositionality for recursive neural models with applications to sentiment analysis. In *AAAI*, 2014.

[Dunbar and Dunbar, 1998] Robin Dunbar and Robin Ian MacDonald Dunbar. *Grooming, gossip, and the evolution of language*. Harvard University Press, 1998.

[Gao *et al.*, 2013] Wenliang Gao, Naoki Yoshinaga, Nobuhiro Kaji, and et al. Modeling user leniency and product popularity for sentiment classification. *IJCNLP*, 2013.

[Glorot *et al.*, 2011] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Domain adaptation for large-scale sentiment classification: A deep learning approach. In *ICML*, pages 513–520, 2011.

[Gui *et al.*, 2014] Lin Gui, Ruifeng Xu, Qin Lu, Jun Xu, and et al. Cross-lingual opinion analysis via negative transfer detection. In *ACL*, 2014.

[He *et al.*, 2012] Yulan He, Hassan Saif, Zhongyu Wei, and Kam-Fai Wong. Quantising opinions for political tweets analysis. 2012.

[He *et al.*, 2013] Yulan He, Chenghua Lin, Wei Gao, and et al. Dynamic joint sentiment-topic model. *ACM Trans. on Intelligent Systems and Technology*, 5(1):6, 2013.

[Hovy, 2015] Dirk Hovy. Demographic factors improve classification performance. In *Proceedings of ACL*, 2015.

[Hu *et al.*, 2013] Xia Hu, Lei Tang, Jiliang Tang, and et al. Exploiting social relations for sentiment analysis in microblogging. In *WSDM*, pages 537–546, 2013.

[Hu *et al.*, 2014] Baotian Hu, Zhengdong Lu, Hang Li, and Qingcai Chen. Convolutional neural network architectures for matching natural language sentences. In *NIPS*, pages 2042–2050, 2014.

[J *et al.*, 2015] Xu J, Zhang Y, Wu Y, Wang J, Dong X, and Xu H. Citation sentiment analysis in clinical trial papers. *AMIA*, 2015.

[Kalchbrenner *et al.*, 2014] Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom. A sentence model based on convolutional neural networks. In *ACL*, 2014.

[Kim, 2014] Yoon Kim. Convolutional neural networks for sentence classification. *EMNLP*, 2014.

[Le and Mikolov, 2014] Quoc V Le and Tomas Mikolov. Distributed representations of sentences and documents. pages 1188–1196, 2014.

[Mikolov *et al.*, 2013] Tomas Mikolov, Ilya Sutskever, Kai Chen, and et al. Distributed representations of words and phrases and their compositionality. In *NIPS*, pages 3111–3119, 2013.

[Pang and Lee, 2005] Bo Pang and Lillian Lee. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *ACL*, pages 115–124, 2005.

[Pang *et al.*, 2002] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *ACL*, pages 79–86, 2002.

[Socher *et al.*, 2013] Richard Socher, Alex Perelygin, Jean Y Wu, and et al. Recursive deep models for semantic compositionality over a sentiment treebank. In *EMNLP*, volume 1631, page 1642, 2013.

[Taboada *et al.*, 2011] Maite Taboada, Julian Brooke, Milan Tofiloski, and et al. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307, 2011.

[Tan *et al.*, 2011] Chenhao Tan, Lillian Lee, Jie Tang, and et al. User-level sentiment analysis incorporating social networks. In *SIGKDD*, pages 1397–1405, 2011.

[Tang *et al.*, 2014a] Duyu Tang, Furu Wei, Bing Qin, and et al. Building large-scale twitter-specific sentiment lexicon: A representation learning approach. In *COLING*, pages 172–182, 2014.

[Tang *et al.*, 2014b] Duyu Tang, Furu Wei, Nan Yang, and et al. Learning sentiment-specific word embedding for twitter sentiment classification. In *ACL*, volume 1, pages 1555–1565, 2014.

[Tang *et al.*, 2015a] Duyu Tang, Bing Qin, and Ting Liu. Document modeling with gated recurrent neural network for sentiment classification. In *EMNLP*, pages 1422–1432, 2015.

[Tang *et al.*, 2015b] Duyu Tang, Bing Qin, and Ting Liu. Learning semantic representations of users and products for document level sentiment classification. In *ACL*, 2015.

[Tang *et al.*, 2015c] Jian Tang, Meng Qu, Mingzhe Wang, and et al. Line: Large-scale information network embedding. In *WWW*, pages 1067–1077, 2015.

[Wang and Manning, 2012] Sida Wang and Christopher D Manning. Baselines and bigrams: Simple, good sentiment and topic classification. In *ACL*, pages 90–94, 2012.

[Xia *et al.*, 2013] Rui Xia, Chengqing Zong, Xuelei Hu, and et al. Feature ensemble plus sample selection: domain adaptation for sentiment classification. *Intelligent Systems, IEEE*, 28(3):10–18, 2013.

[Xu *et al.*, 2014] Liheng Xu, Kang Liu, and Jun Zhao. Joint opinion relation detection using one-class deep neural network. pages 677–687, 2014.

[Xu *et al.*, 2015] Ruifeng Xu, Lin Gui, Jun Xu, and et al. Cross lingual opinion holder extraction based on multi-kernel svms and transfer learning. *World Wide Web*, 18(2):299–316, 2015.