Aspect-based Sentiment Analysis using Vector Space Restricted Boltzmann Machines

No Author Given

No Institute Given

Abstract. Recent years, many studies have addressed problems in sentiment analysis in different levels, and building aspect-based methods has become a central issue for deep opinion mining. However, previous studies need to use two-separated modules in order to extract aspect-sentiment word pairs, then predict the sentiment polarity. In this paper, we use Restricted Boltzmann Machines in combination with Vector Space Model to build the joined model which not only extracts aspect terms appeared and classify them into respective categories, but also completes the sentiment polarity prediction task. The experiments show that the method we use in categories classification and polarity prediction tasks better than other state-of-the-art approaches.

Keywords: Aspect-based, Sentiment Analysis, Opinion Mining, Restricted Boltzmann Machine, Supervised Learning, Vector Space Model

1 Introduction

Sentiment Analysis (also known as opinion mining) is the process of determining whether a piece of writing is positive, negative. It aims to derive the opinion or attitude of a speaker or a writer with respect to some topic or the overall contextual polarity of a document. With the development of opinionated user-generated review sites, many customers can write reviews and express their opinions about the products, services, people or places on the websites. Each review contains many aspects/features about the products. For example, consider reviews in restaurant domain, we may have features such as view (of the restaurant), cost, dishes, etc. Therefore, new customers may find it difficult to explore the large number of reviews in order to make right decision, since many people have different purposes, which base on different aspects of the product. Thus, Sentiment Analysis can play an important role, not only address this issue but also help companies improve their products based on these reviews.

Aspect-based Sentiment Analysis (ABSA) has received much attention in recent years. ABSA consists of two parts, aspect extraction and sentiment classification. These two parts are analyzing opinionated texts, such as opinions, sentiments by extracting the aspect-term appeared in the sentences, and classifying them into two classes (positive or negative). It helps people make the right decisions thanks to the fine-grained analysis on each aspect of products. In previous works, Liu and Hu proposed combining two or more modules to complete the

ABSA model [1–3]. For instance, we can use one for extracting aspect-sentiment word pairs, one for sentiment prediction, another optional module is to summarize the opinions.

Some work has been investigated to extract the candidate product feature opinion pairs. Kumar and Raghuveer use the rules on the typed dependency tree, then they use lexicons to classify and generate a summary of the product [3]. In Toh and Wang [4], they build a Conditional Random Field (CRF) based classifier for Aspect Term Extraction and a linear classifier for Aspect Term Polarity Classification, but there is no method to extract the opinion words in the sentences. Unsupervised method such as Latent Dirichlet Allocation (Blei et al., 2003) [5] is also used to extract and group corresponding representative words into categories. Such approaches, however, must use two-separated modules to perform both aspect extraction and sentiment classification, or just only have ability to solve one of these tasks.

Hence, Wang et al. adapted the Sentiment-Aspect Extraction based on Restricted Boltzmann Machines (SERBM) to overcome this problem [6]. They use unsupervised version of RBM to group reviews in categories and also classify opinion. Three different types of hidden units are used to represent aspects, sentiments, and background words in this model, respectively. Furthermore, they blend background knowledge into this model using priors and regularization to help it acquire more accurate feature representations. The visible layer $\mathbf{v}$ of SERBM is represented as a $K \times D$ matrix, where K is the dictionary size and D is the document length, or the number of sentences in reviews. However, much uncertainty still exists about the defined hidden units in hidden layer. First, unsupervised method can just cluster reviews into categories and can not fix the unit for desired category. No information given to determine which position of the hidden units to represent aspects, sentiments or background words during the training process. Second, visible layer would be a matrix combined by high-dimensional vectors if training data set had varied vocabulary, which result in insufficient computational resources.

In this paper, we implement Vector Space model combined with Restricted Boltzmann Machine (VS-RBM) to extract aspect term appeared and classify sentiment polarity of the sentences. The reasons we propose VS-RBM include the ability to jointly model aspect and sentiment information together and the capability in reducing dimensionality of the input vectors. In this two-layer structure model, we do not use hidden layer as output layer but the dependencies between the components of observations in visible layer. Each input unit in the visible layer is a component of the feature vector, created by the Vector Space Model instead of using the Bag Of Words model as previous approach. This helps reduce the dimensionality of the input while keeping documents' semantics. The results obtained based on the dataset of reviews in restaurant domain, which widely adopted by previous work [7–9] (Ganu et al., 2009; Brody and Elhadad, 2010; Zhao et al., 2010).

The rest of this paper is organized as follows. Section 2 presents related work. Section 3 overviews the background information of models, describes our approach to classify reviews into aspect categories and predict sentiment polarity of the reviews. Experimental results are presented in Section 4. Finally, Section 5 concludes the paper and discusses future work.

2 Related work

For the supervised learning approach, to determine the dependency and the other relevant information in the sentence, Wei and Gulla (2010) propose a hierarchical classification model. Jiang et al. (2011), Boiy and Moens (2009) use dependency parser to generate a set of aspect dependent features for classification, which weighs each feature based on the position of the feature relative to the target aspect in the parse tree. Other supervised learning models have been widely used as Hidden Markov Models, SVMs, Conditional Random Fields to extract aspects which appear in sentences (Jin et al., 2009; Choi and Cardie, 2010; Jakob and Gurevych, 2010; Kobayashi et al., 2007).

Supervised learning is dependent on the training data and has difficulty to scale up to a large number of application domains. A model or classifier trained from labeled data in one domain often performs poorly in another domain [**?**]. Hence, unsupervised methods are often adopted to avoid this issue. Several recent studies investigate statistical topic models which are unsupervised learning methods that assumes each document consists of a mixture of topics. Specifically, Latent Dirichlet Allocation (LDA), Multi-Grain LDA model (Titov and McDonald, 2008) are used to model and extract topics from document collections. The two-step approach to detect aspect-specific opinion words (Brody and Elhadad, 2010) includes identifying aspects using topic models and identifying aspect-specific sentiment words by considering adjectives only. The joint sentiment/topic model (JST) is also used to separate aspect words and sentiment words (Lin and He, 2009). The topic-sentiment mixture model (Mei et al., 2007) based on three separated models with the help of some external training data is proposed to extract aspect and sentiment words. The SERBM model (Wang et al., 2015) also jointly address these two tasks in an unsupervised setting.

Other methods also use lexicon-based approach have been proposed (Hiu and Liu, 2004; Ding et al., 2008) which has been shown to perform quite well in a large number of domains. They use a sentiment lexicon, expressions, rules of opinions, and the sentence parse tree to help classify the sentiment orientation on each aspect appeared in a review. They also consider sentiment shifters (i.e. not, none, nobody, etc). However, these rule-based methods have a shortcoming in processing complex documents where the aspect is hidden in sentence, and failing to group extracted aspect terms into categories.

3 Proposed method

3.1 Background

Restricted Boltzmann Machine Energy-based models associate a scalar energy to each configuration of the variables of interest. Learning phrase modifies that energy function so that it will has desirable properties. Energy-based probabilistic models define a probability distribution through an energy function, as follows:

$$p(x) = \frac{e^{-E(x)}}{Z}. \tag{1}$$

where $E(x)$ is the energy function and Z is called the partition function defined by

$$Z = \sum_x e^{-E(x)} \tag{2}$$

Stochastic gradient descent can be used to perform learning phrase for energy-based model.

Boltzmann Machines (BMs) are a particular form of log-linear Markov Random Field (MRF), i.e., for which the energy function is linear in its free parameters. RBM further restrict BMs to those without visible-visible and hidden-hidden connections, or there is no connection between neurons in the same layer. A graphical depiction of an RBM is shown below.

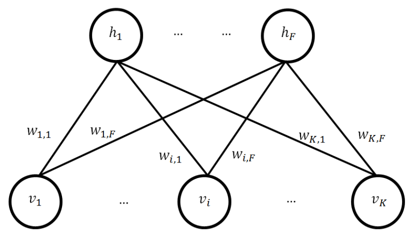

Fig. 1: The network graph of an RBM model with K visible and F hidden units

As shown in Figure 1, RBM model is a two-layer neural network which contains one visible layer and one hidden layer. Visible layer is constituted by visible units correspond to the components of an observation (e.g., one visible unit for each word in a input document). The hidden layer composed of hidden units which model dependencies between the components of observations (e.g., dependencies between the words in the document). To make them powerful enough to represent complicated distributions, we can increase the numbers of hidden variables (also called hidden units) in order to increase the modeling capacity of the RBM.

In this context, the visible layer $\mathbf{v}$ is represented as a $D \times K$ matrix, where D is the document length, or the number of sentences in reviews, and K is the dictionary size. More precisely, RBM model takes input as sentences in review, each visible unit correspond to each word in sentence. If a sentence i-th has word k-th in dictionary, we set $\mathbf{v}_i^k = 1$. The hidden layer can be expressed as a vector $h \in 0, 1^F$, where F is the dimension of the vector, or number of hidden units. Hidden units of RBM model corresponds to the relation between visible units.

The energy function of the model can be defined as

$$E(\mathbf{v}, h) = -\sum_{i=1}^{D}\sum_{j=1}^{F}\sum_{k=1}^{K} W_{ij}^{k} h_j v_i^k - \sum_{i=1}^{D}\sum_{k=1}^{K} v_i^k b_i^k - \sum_{j=1}^{F} h_j a_j. \tag{3}$$

where W_{ij}^k is the connection weight between i-th visible unit of value k to the j-th hidden node, b_i^k specifies a bias of visible unit v_i^k, and a_j specifies to a bias of h_j. The conditional probabilities from hidden to visible layer and from visible to the hidden layer we can obtain from (1) and (3) are probabilistic version of the usual neuron activation function.

$$P(v_i^k = 1|h) = sigm(b_i^k + \sum_{j=1}^{F} h_j W_{ij}^k) \tag{4}$$

$$P(h_j = 1|\mathbf{v}) = sigm(a_j + \sum_{i=1}^{D}\sum_{k=1}^{K} v_i^k W_{ij}^k) \tag{5}$$

where $sigm(x) = \frac{1}{(1+e^{-x})}$ is the logistic function.

Vector Space Model The Vector Space Model (VSM) is a proven and powerful paradigm in Natural Language Processing, in which documents are represented as vectors in a high-dimensional space [10]. The idea of this model is representing text documents as vectors. So that text documents can be calculated in form of vectors where semantics is conserved.

In VSM, each word i are represented as t-dimensional vector vector $w_i = (v_i, v_2, ..., v_t)$ The similarity of two words i and j can be calculated based on the cosine of the angle between the vectors.

$$\cos\theta = \frac{w_i.w_j}{||w_i||||w_j||} \tag{6}$$

where $w_i.w_j$ is the dot product of these two words. $||w_i||$ and $||w_j||$ is the norm of vector w_i and w_j, respectively.

3.2 Our Aspect-based Sentiment Analysis model

RBM model as a generative stochastic artificial neural network is treated like a model in the field of deep learning. RBM have found applications in dimensionality reduction, classification, collaborative filtering, feature learning, especially applications in the field of image processing. Recently, RBM is developed and applied in the field of Natural Language Processing, including topic modeling, sentiment analysis. It can be trained in supervised or unsupervised ways, which depends on the task.

For the customer restaurant reviews analyzing task, every word in a review may mention a certain aspect (e.g. “delicious” corresponds to *food* aspect), or a certain opinion (e.g. “good” is about positive sentiment, while “bad” is about negative sentiment). In addition, there are other words do not mention about aspect or sentiment, they can be removed during preprocessing phrase. The latent topics in a review are considered as the factors which generate aspect and sentiment words within that review. To technically illustrate this, hidden layer which includes hidden units in RBM would contain information of the latent topics. In the generation process, they produce the values of visible units, which are also the information of the aspect and sentiment words in reviews.

3.3 Structure

In previous approach, they have used unsupervised learning to build model of RBM [6]. Although this model can help to overcome the limitations of handlabeled training data, but there still exist certain shortcomings. First, this unsupervised RBM model fixes the units to serve a function in the hidden layer. In particular, they fix hidden units 0-6 to represent the target aspects *Food*, *Staff*, *Ambience*, *Price*, *Ambience*, *Miscellaneous*, and *Other Aspects*, respectively. This may reduce the accuracy of the classification, since unsupervised learning model does not fix the categories that it will cluster into the output units before the classification phrase is accomplished. There is no information in the visible layers for the connected hidden units to get which unit for which aspect. Second, each document which put into learning model is transformed into a $K \times D$ matrix $\mathbf{v}$, where K is the dictionary size, and D is the document length. If visible unit i in $\mathbf{v}$ takes the k-th value, v_i^k is set to 1. This makes the dictionary size will be very large if training reviews become abundant. From that, the feature vectors which are fed into the model will be sparse, leading to the decrease of the model's accuracy .

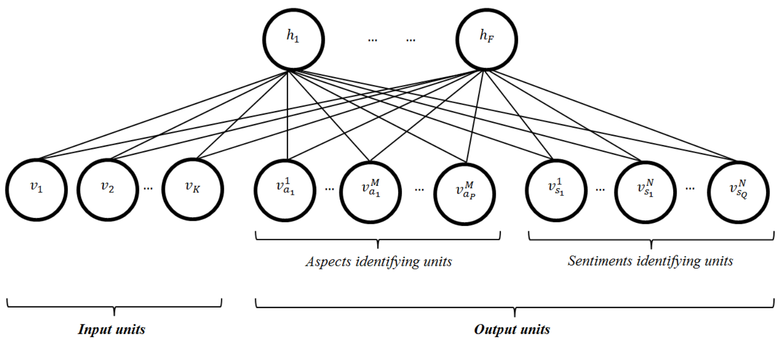

Fig. 2: The network graph of our Supervised Sentiment-Aspect Extraction RBM model

To overcome these problems, we propose using supervised RBM model which is illustrated in Figure 2. The ***Output units*** including the units represent the aspects and sentiment orientations of the reviews, which called *Aspects identifying units* and *Sentiments identifying units* respectively. These units will be put together with ***Input units*** in the visible layers, instead of being placed in hidden layers. Meanwhile, the units in the hidden layer will take the role of the relationship between the units in the visible layers. Each input unit in the visible layer is a component of the feature vector, created by the word representation model instead of using the bag of words model as previous approach. This helps reduce the dimensionality of the input matrix while keeping documents' semantics.

Compared to standard RBMs, the first difference of this model is the visible layer. Apart from the input units which are the feature vector made from reviews, there are also aspect and sentiment identifying units that represent the output component of the model. Suppose in the training set, there are reviews talking about P aspects and have Q sentiment orientations in total. For detail, set of aspects is $A = \{a_1, a_2, ..., a_P\}$, set of sentiment orientations is $S = \{s_1, s_2, ..., s_Q\}$.

For each aspect, the model will set M units to express the nature of the aspect. For example, if the review mentioned aspect a_i of A, the units from $v^1_{a_i}$ to $v^M_{a_i}$ would be set to 1, and 0 otherwise. The sentiment orientation of the reviews bears comparison with aspect. The model will set N units to express the nature of the sentiment. If the sentiment polarity is s_j included in S, the units from $v^1_{s_j}$ to $v^N_{s_j}$ would be set to 1, and 0 otherwise.

This structure which builds up the strength of the model can solve two tasks simultaneously, including aspect identification and sentiment classification. They are reflected in the aspect and sentiment identifying units which play important roles in the sampling process of RBM, while weight values of connecting edges between the units contain semantic information of the review. In addition, the fixed dimension of the feature vectors which do not depend on the number of vocabulary not only helps the model prevent the occurrence of decreasing speed and accuracy, but also solves the semantic problem in the review.

3.4 Training

Contrastive Divergence (CD), also called Approximate Gradient Descent, is the way that help RBM learns the connection weights in the network. CD has two phrases, which are Positive phrase and Negative phrase, in each phrase we compute the positive and negative value by the multiplication of visible and hidden units, then update the connection weight based on the subtraction from positive value of negative value. For detail, each epoch of CD can be expressed in several steps below:

Step 1. Update the states of the hidden units using the logistic activation rule described in (5). For the j-th hidden unit, compute its activation energy and set its state to 1 with corresponding probability.

Step 2. For each connection edge e_{ij}, get the value from positive phrase by computing $pos(e_{ij}) = v_i h_j$

Step 3. Reconstruct the visible units in a similar manner by using the logistic activation rule described in (4). For the i-th visible unit, compute its activation energy and set its state to 1 with corresponding probability. Then do ***Step 1*** to update the hidden units again.

Step 4. For each connection edge e_{ij}, get the value from negative phrase by computing $neg(e_{ij}) = v_i h_j$

Step 5. Update the connection weights W_{ij} by the equation

$$W_{ij} = W_{ij} + \mathbf{lr}(pos(e_{ij}) - neg(e_{ij})) \tag{7}$$

where **lr** is a learning rate.

After m steps of transfer between visible and hidden layers in a CD-m run of the above steps, values of the hidden units will reflect the relationship between the visible units in the model. The connection weight matrix between the two layers helps hidden layer generate visible units which include input units, aspect identifying units and also sentiment identifying units.

4 Experiments

Experimental results

đừng dùng mấy chữ không xác định, thay vào đó dùng two / three ...

In this section, we present sequence of experiments to evaluate the performance of our model on the aspect identification and sentiment classification tasks.

4.1 Data

restaurant review

To evaluate our model performance, we used review restaurant dataset which is used widely in previous work [7–9]. Data contain reviews about the restaurant with 1,644,923 tokens and 52.574 documents in total. Documents in this dataset are annotated with one or more labels from a gold standard label set $S = \{$*Food, Staff, Ambience, Price, Anecdote, Miscellaneous*$\}$.

Em ghi ra tên dataset là gì, do công trình nào đề ra đầu tiên
Vd: for evaluation, we use ABC dataset~\cite{xyz}. This data was also used in ...

4.2 Aspect Extraction

Experimental Setup Following the previous studies (Brody and Elhadad (2010) and Zhao et al. (2010)), reviews with less than 50 sentences are chosen. From that, we only use sentences with a single label for evaluation to avoid ambiguity, these sentences are selected from reviews with three major aspects chosen from the gold standard labels are $S' = \{$*Food, Staff, Ambience*$\}$. Then, we put all words in sentence as lowercase and remove stop words.

cite theo template

tách câu

To convert each word in the sentence into feature vector, we use word-to-vector technique [10] with pre-trained vectors trained on part of Google News dataset (about 100 billion words)[1]. The model contains 300-dimensional vectors for 3 million words and phrases. Each sentence's representation is a vector which is a summary of vectors that represent for words appeared in that sentence.

cách sum này cần cite

We use 300 visible units in our Vector Space RBM (VS-RBM) as aspects identifying units, where units 1-100 capture *Food* aspect, units 101-200 capture *Staff* aspect and units 201-300 capture *Ambience* aspect. For sentiment classification, we also use 200 visible units playing a role as sentiments identifying units, each 100 units capture positive and negative information similarly. At first, these units are set to 0. After Gibbs sampling process, we sum up each 100 of these units to determine the aspect appeared in document.

mấy chỗ bôi vàng cần coi lại tiếng Anh

[1] `https://code.google.com/archive/p/word2vec/`

thêm trường short-title

canh độ rộng bảng lại cho đẹp

Aspect	Method	Precision	Recall	F_1
food	Prior knowledge only	88.09	7.60	87.84
	SVM	**94.10**	75.74	83.93
	RBM + Word2Vec	74.58	14.79	24.68
	VS-RBM	84.04	**94.43**	**88.93**
staff	Prior knowledge only	**95.96**	45.39	61.63
	SVM	90.25	**78.14**	**83.76**
	RBM + Word2Vec	27.73	53.32	36.48
	VS-RBM	88.27	65.28	75.06
ambience	Prior knowledge only	16.02	**86.93**	27.06
	SVM	29.15	82.70	43.11
	RBM + Word2Vec	14.85	60.12	23.81
	VS-RBM	**69.22**	59.64	**64.07**

Table 1: Aspect identification results in terms of precision, recall, and F_1 scores on the restaurant reviews dataset

Evaluation To evaluate the model's performance, we use Precision, Recall and F_1 scores for each aspect identification on restaurant review dataset. As a baseline, we implement *Prior knowledge only*, method which uses only cosine similarity of the document vector and aspect vector to identify aspect. For detail, we generate the vectors of "food", "staff" and "ambience" words using Word2Vec model. And then compute the cosine similarity between the vector of document d_i and each of those 3 aspect vectors. The document will be put into aspect a_i category if the cosine similarity between $vector(d_i)$ and $vector(a_i)$ is the highest one.

tách thành 2 đoạn
Đoạn 1: evaluation metric
Đoạn 2: mô tả baseline

SVM and standard RBM with Word2Vec are also re-implemented to compare with our VS-RBM model, which process this same restaurant review dataset and identify aspects for every document in this dataset under the same experimental conditions. Evaluation results for aspect identification are given in Table 1.

giải thích lí do tại sao lại so sánh với các pp này?

feature đưa vào SVM ntn? tham số cho SVM ntn? dùng tool gì?

Discussion Considering the results from Table 1, we find that VS-RBM performs better than other methods. Specifically, it is evident that our VS-RBM model outperforms previous methods' F_1 scores on *Food* and *Ambience* aspects. Compared with Prior knowledge only, the F_1 scores improve by 1.09%, 13.43% and 37.01% respectively, for the *Food*, *Staff*, and *Ambience* aspects. This result proves that our VS-RBM model is not entirely based on Prior knowledge obtained from VSM to have better classification. Inheriting RBM's ability in modeling latent topics and VSM's capability in identifying aspects, VS-RBM model im? can achieve higher Precision and Recall scores for the unbalanced dataset. Compared with RBM + Word2Vec model, the F_1 scores yield relative improvements by 63.25%, 38.58% and 40.26% respectively, on the same aspects. This result reveals that RBM model performs badly without the prior knowledge from VSM when using Word2Vec technique.

Comparing with SVM performance, precision scores in food and staff domains of SVM are higher than VS-RBM's, but SVM can not address the unbalance data problem which results in reduced precision in ambience domain. Moreover, VS-RBM is a joint model which has ability to identify aspects and classify sentiment polarities simultaneously, while SVM is a classification model and we have to train two separated models to solve these two tasks.

In our model's performance evaluation process, we can not re-implement SERBM due to insufficient computational resources. However, we compare VS-RBM's result with SERBM's result which was presented by Wang et al. [6]. With the same settings in aspect identification task on restaurant review dataset, VS-RBM outperforms SERBM with improvement in F_1 scores by 1.73% and 7.06% on *Food* and *Staff* aspects, respectively.

4.3 Sentiment Classification

Sentiment classification task is modified similar to aspect identification task in our VS-RBM model. We assign a sentiment score to every document in the restaurant review dataset based on the output of VS-RBM's sentiment identifying units in visble layer. Then we use SentiWordnet[2], a famous lexical resource for opinion mining [11], and adapt SVM[3], a well-know machine learning technique [12], to compare the result with our VS-RBM model.

Following the previous studies, we consult SentiWordNet to obtain a sentiment label for every word and aggregate these to judge the sentiment information of an entire review document in terms of the sum of word-specific scores [6]. For SVM, we use linear kernel to train the model and the other setting is the same as VS-RBM's. Table 2 shows the comparison between SentiWordNet, SVM and VS-RBM with Accuracy as the evaluation metric.

canh lại table

Method	SentiWordNet	SVM	VS-RBM
Accuracy	73.36	78.26	**79.79**

Table 2: Accuracy for SentiWordNet, SVM and VS-RBM on sentiment classification task

As we can observe in Table 2, the best sentiment classification accuracy result on the restaurant review dataset is 79.79% achieved by VS-RBM. Compared with two baselines, VS-RBM yields a relative improvement in the overall accuracy by 6.43% over SentiWordNet and by 1.53% over SVM.

Comparing the result of VS-RBM with SERBM's which was presented by Wang et al. [6], VS-RBM increase the classification accuracy by 0.99%.

giải thích thêm lí do (có thể) tại sao VS-RBM lại tốt hơn các pp khác

[2] `http://sentiwordnet.isti.cnr.it`
[3] `https://www.csie.ntu.edu.tw/~cjlin/libsvm/`

5 Conclusion and future work

In this paper, we have proposed Vector Space Restricted Boltzmann Machine (VS-RBM) model to jointly identify aspect categories and classify sentiment polarities in supervised setting. Our approach modifies standard RBM model and combine it with VSM, which not only helps reduce dimension of the input vectors but also gives the RBM model prior knowledge for more accurate classification. Our experimental results show that this model can outperform state-of-the-art models.

In the future, we plan to collect Vietnamese review dataset with diverse domains and combine our VS-RBM with stacked RBMs to form Deep Belief Networks. We hope that it will solve the aspect identification and sentiment classification tasks as well as RBM family methods.

References

1. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of KDD. (2004) 168–177
2. Blair-goldensohn, S., Neylon, T., Hannan, K., Reis, G.A., Mcdonald, R., Reynar, J.: Building a sentiment summarizer for local service reviews. In: WWW Workshop on NLP in the Information Explosion Era. (2008)
3. kumar V, R., Raghuveer, K.: Web user opinion analysis for product features extraction and opinion summarization. International Journal of Web and Semantic Technology (2012) 69–82
4. Toh, Z., Wang, W.: Dlirec: Aspect term extraction and term polarity classification system. In: Proceedings of the 8th International Workshop on Semantic Evaluation. (2014) 235–240
5. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation. Journal of Machine Learning Research 3 (2003) 993–1022
6. Wang, L., Liu, K., Cao, Z., Zhao, J., de Melo, G.: Sentiment-aspect extraction based on Restricted Boltzmann Machines. In: Proceedings of the 53rd Annual Meeting of the ACL. (2015) 616–625
7. Ganu, G., Elhadad, N., Marian, A.: Beyond the stars: Improving rating predictions using review text content. In: Proceedings of WebDB 2009. (2009) 1–6
8. Brody, S., Elhadad, N.: An unsupervised aspect-sentiment model for online reviews. In: Proceedings of NAACL-HLT 2010. (2010) 804–812
9. Zhao, W.X., Jiang, J., Yan, H., Li, X.: Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. In: Proceedings of EMNLP 2010. (2010) 56–65
10. Řehůřek, R., Sojka, P.: Software Framework for Topic Modelling with Large Corpora. In: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, ELRA (2010) 45–50
11. Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In: Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10). (2010) 2200–2204
12. Chih-Chung, C., Chih-Jen, L.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology **2** (2011) 27:1–27:27 Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.