A Cross-Lingual Joint Aspect/Sentiment Model for Sentiment Analysis

Zheng Lin
Institute of Information
Engineering, CAS
No.91(A), Minzhuang Road
Beijing, China
linzheng@iie.ac.cn

Xiaolong Jin
Key Lab of Network Data
Science & Technology
ICT, CAS
No.6, Kexueyuan South Road
Beijing, China
jinxiaolong@ict.ac.cn

Xueke Xu
Key Lab of Network Data
Science & Technology
ICT, CAS
No.6, Kexueyuan South Road
Beijing, China
xuxueke@ict.ac.cn

Yuanzhuo Wang
Key Lab of Network Data
Science & Technology
ICT, CAS
No.6, Kexueyuan South Road
Beijing, China
wangyuanzhuo@ict.ac.cn

Weiping Wang
Institute of Information
Engineering, CAS
No.91(A), Minzhuang Road
Beijing, China
wangweiping@iie.ac.cn

Xueqi Cheng
Key Lab of Network Data
Science & Technology
ICT, CAS
No.6, Kexueyuan South Road
Beijing, China
cxq@ict.ac.cn

ABSTRACT

Sentiment analysis in various languages has been a research hotspot with many applications. However, sentiment resources (e.g., labeled corpora, sentiment lexicons) of different languages are unbalanced in terms of quality and quantity, which arouses interests in cross-lingual sentiment analysis aiming at using the resources in a source language to improve sentiment analysis in a target language. Nevertheless, many existing cross-lingual related works rely on a certain machine translation system to directly adapt the labeled data from the source language to the target language, which usually suffers from inaccurate results generated by the machine translation system. On the other hand, most sentiment analysis studies focus on document-level sentiment classification that cannot solve the aspect dependency problem of sentiment words. For instance, in the reviews on a cell phone, *long* is positive for the lifespan of its battery, but negative for the response time of its operating system. To solve these problems, this paper develops a novel Cross-Lingual Joint Aspect/Sentiment (CLJAS) model to carry out aspect-specific sentiment analysis in a target language using the knowledge learned from a source language. Specifically, the CLJAS model jointly detects aspects and sentiments of two languages simultaneously by incorporating sentiments into a cross-lingual topic model framework. Extensive experiments on different domains and different languages demonstrate that the proposed model can significantly improve the accuracy of sentiment classification in the target language.

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing

General Terms

Algorithms, Experimentation, Performance

Keywords

Sentiment Analysis, Cross-Lingual Sentiment Analysis, Topic Model, Aspect Discovery

1. INTRODUCTION

With the flourish of Web2.0 based applications, the Web has been full of various user-generated contents in different languages, which contain rich sentiment information (e.g., comments, reviews, blogs, etc.). It is no doubt that such information is of great value. For instance, based on the reviews or comments left by previous customers, product or service providers may improve the quality of their services or products, while incoming customers may choose well-content services or products. For this reason, sentiment analysis aiming at identifying the sentiment polarity (i.e., positive or negative) of given texts has rapidly emerged into a research hotspot with broad application prospects, including product recommendation and service follow-up etc., and has thus attracted much attention from both academia and industries in recent years [13, 19, 22]. A couple of sentiment analysis related issues have been extensively studied for online customer reviews, ranging from coarse-grained document-level sentiment classification [16, 28] to fine-grained extraction of aspects and opinion expressions [6, 20].

Currently, for some languages (e.g., English), there have been lots of sentiment resources available for conducting sentiment analysis, including labeled corpora and sentiment lexicons [19]. However, for other languages, sentiment resources are often insufficient or of low quality. Therefore, it is desirable to develop cross-lingual mechanisms so as to utilize the knowledge learning from those resource-rich languages to improve sentiment analysis in other languages. In doing so, a direct approach is to translate the resources in the resource-rich language to the target language using a machine

translation tool [1, 23]. However, it is known that machine translation is a very challenging task and the existing machine translation system is thus not powerful enough to generate very accurate results. Particularly, a machine translation system usually generates only one result, which may not be suitable for the situation therein.

In reviews, customers often pay more attention to specific aspects rather than the whole entity. Here, an aspect means a part or a feature of the entity, e.g., the rooms or the location of a hotel. A customer tends to be more interested in some particular aspects when making a purchase or booking decision. Therefore, it is necessary and appealing to simultaneously consider aspects and sentiments in sentiment analysis. However, although there have been some studies on cross-lingual sentiment analysis [1, 23], they are designed on document level and consequently cannot handle the aspect dependency problem of sentiment words, which means that a sentiment word may have different sentiment polarities regarding different aspects of an entity. For instance, in the reviews on a cell phone, *long* is positive for its battery's lifespan, but negative for the response time of its operating system. On the other hand, probabilistic topic models have been widely used on aspect-level sentiment analysis in recent years, because they provide an unsupervised and thus flexible way in discovering aspects from reviews [11, 12]. However, most of them perform well in a specific language and cannot be readily transferred to other languages. The main reason is that the resources are unbalanced among different languages. Directly applying these methods on poor-quality corpora often leads to poor results. Therefore, it is a potential solution to use a cross-lingual topic model to help sentiment analysis in a target language by exploiting the resources in a source language. However, the existing cross-lingual topic models are not suitable for sentiment analysis, as they do not take sentiment into account.

To address the above problems, in this paper we propose a novel Cross-Lingual Joint Aspect/Sentiment (CLJAS) model that can jointly extract aligned aspects and aspect-dependent sentiments in two different languages via a unified framework. To the best of our knowledge, this is the first study in the open literature on developing an unsupervised cross-lingual topic model for sentiment analysis. The proposed CLJAS model can help sentiment analysis in a resource-poor language by exploiting its correspondences with a resource-rich language, which, however, does not require parallel corpora, a machine translation system, and any labeled sentiment reviews. Extensive experiments on real datasets in different domains and different languages demonstrate that the proposed model can be successfully used in practical applications and significantly improve the accuracy of sentiment classification in target languages.

The rest of this paper is organized as follows. Section 2 reviews the related work of sentiment analysis on sentence and aspect levels and cross-lingual topic models. Section 3 elaborates the proposed CLJAS model. In Section 4, through comparison with existing representative models, we validate the effectiveness and investigate the performance of the CLJAS model. Finally, Section 5 concludes the paper.

2. RELATED WORK

In this section, we first give an overview of sentiment analysis on sentence level and aspect level, respectively, and then present related studies on cross-lingual topic models.

2.1 Sentence-level sentiment analysis

Sentence-level sentiment analysis has two tasks, i.e., subjectivity classification and sentiment classification. Subjectivity classification aims at separating objective sentences from subjective ones. It

has been earnestly explored in the field of natural language processing [24, 25]. Sentence-level sentiment classification considers each sentence as a separate unit and assumes that a sentence contains only one opinion. Existing studies on sentence-level sentiment classification can roughly be divided into two categories.

The first category is based on various statistical techniques. For instance, Pang and Lee [18] first showed that sentence-level extraction can improve the performance of sentiment analysis on document level. They developed a cascaded approach that first filters out objective sentences using a global min-cut inference. Afterward, the extracted subjective sentences are adopted as the input of a document-level sentiment classifier. Hu and Liu [10] proposed a bootstrapping technique to determine the semantic orientation of each opinion sentence using WordNet.

The second category makes use of the advanced natural language processing tools, such as, syntactic parsers. For example, Meena and Prabhakar [15] showed how atomic sentiments of individual phrases can be combined together in the presence of conjuncts to decide the overall sentiment of a sentence. Yao and Li [27] proposed a kernel-based approach to classify sentiments of sentences by incorporating multiple features from lexical and syntactic levels. Tan et al. [21] invented a linguistic approach that combines typed dependencies and subjective phrase analysis to detect sentence-level sentiment polarity.

In general, all of the above studies have demonstrated that regarding each sentence as a separate unit is reasonable in sentiment analysis and sentence-level sentiment analysis is helpful for document-level sentiment classification in some cases.

2.2 Aspect-level sentiment analysis

As compared to sentiment analysis on sentence level, sentiment analysis on aspect level is more delicate. There have been a few approaches proposed to detect aspects and aspect-specific sentiment, but not in a unified way. The approach presented in [5] first detects aspects using Local Latent Dirichlet Allocation (LDA) [2] and then identifies aspect-sensitive polarities of adjectives using polarity propagation based on an aspect-specific polarity graph. However, it simply selects adjectives as opinion words, which cannot cover all opinion words. In [14], Lu et al. proposed an optimization framework to combine different signals for determining aspect-aware sentiment polarities. However, the aspects were predefined with manually selected keywords, and the opinion words were extracted beforehand. Unlike these studies, our model presented in this paper can detect aspects and aspect-specific sentiment in different languages within a unified framework.

There have also existed unified models developed by incorporating sentiment factors into classic topic models to jointly detect aspects and sentiments. The Joint Sentiment/Topic (JST) model [12] is the first LDA based model considering topics and sentiments simultaneously. The Aspect and Sentiment Unification Model (ASUM) [11] follows a similar generative process to JST. But, in ASUM a sentiment-topic pair is selected for a sentence rather than a word as in JST. In essence, both ASUM and JST aim at detecting sentiment-coupled aspects/topics rather than explicitly detecting sentiments specific to aspects/topics. MaxEnt-LDA [30] jointly discovers both aspects and aspect-specific opinion words by integrating a supervised maximum entropy component to separate opinion words from objective ones. However, it does not further identify aspect-aware sentiment polarities of opinion words, which are very important for sentiment analysis. The Joint Aspect/Sentiment model proposed in [26] detects aspect-specific opinion words and identifies the sentiment polarities of opinion words.
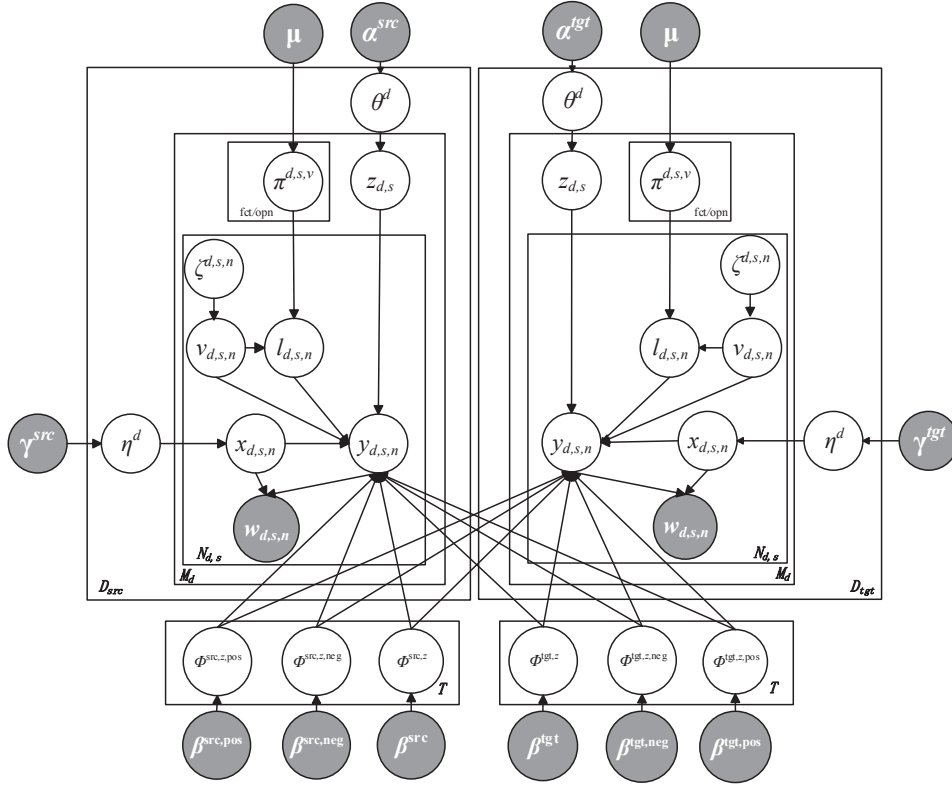
Figure 1: The graphical representation of the CLJAS model.

However, it is designed for monolingual scenarios instead of multilingual ones.

2.3 Cross-lingual topic models

For a cross-lingual topic model, the key is to find a bridge for connecting different languages. Existing models usually combine different languages via parallel or comparable corpora. For example, CI-LDA [7] preserves connections between languages by sharing distributions over latent topics of bilingual documents. Switch LDA [17] extends CI-LDA such that it can control the proportions of languages in each multilingual topic by introducing a per-word switch variable. Symmetric Correspondence LDA proposed in [8] is a topic model that incorporates a hidden variable to control a pivot language. Since parallel or comparable corpora are relatively scarce resources, this kind of models are confined in their applications.

Cross-lingual topic models on unaligned texts may be applied to a broader class of corpora. For instance, Zhang, Mei, and Zhai [29] incorporated soft bilingual dictionary-based constraints into Probabilistic Latent Semantic Analysis (PLSA) so that it can extract shared latent topics in text data of different languages. Boyd-Graber and Blei [3] developed the MUltilingual TOpic (MUTO) model to exploit matching across languages on term level to detect multilingual latent topics from unaligned texts. However, these models do not consider sentiment factors and thus cannot help cross-lingual sentiment analysis.

Moreover, Boyd-Graber and Resnik [4] proposed an LDA-based holistic model for multilingual sentiment analysis on word level. But, this model is supervised, while the model proposed in this paper is unsupervised. Besides, albeit our model is also LDA-based, it is on sentence level, which makes it more appropriate for aspect and sentiment detection from reviews.

3. THE PROPOSED MODEL

In this section, we elaborate the proposed Cross-Lingual Joint Aspect/Sentiment (CLJAS) model, which can jointly detect aspects and aspect-specific sentiments of two languages simultaneously. The key idea behind CLJAS is two-fold. First, for all reviews from the same domain, we assume that they share the same topic distribution despite they are in different languages. In this way, the words in the reviews from the same domain but in different languages tend to be assigned with the same topic. Second, a bilingual dictionary is adopted to connect the topics in two languages. Through dictionary-based translation, we can exploit correspondences between two different languages to learn semantically aligned topic distributions. Note that in this study we cope with the language transferring problem using a bilingual dictionary for two reasons. First, a bilingual dictionary can usually provide multiple translation words to a source word, which enables the proposed model to select the most suitable one for the situation at hand; Second, a bilingual dictionary based model does not require large-scale parallel corpora and is thus more flexible. Figure 1 depicts the graphical representation of the CLJAS model and the notation used in it is explained in Table 1. In what follows we describe the model in more detail. Since CLJAS is basically a left-right symmetrical model, as shown in Figure 1, where the left and right sides correspond to the source and target languages, respectively, in order to make it easy to understand we first describe it from one side and then explain its cross-lingual mechanism.

Assume that we have a corpus of D customer reviews from a specific domain and in different language. Each review d is a list of sentences, each sentence s in review d is a list of words, and the nth word $w_{d,s,n}$ in sentence s of review d is an entry from a vocabulary with V distinct words. Each sentence s in any review

d is associated with an aspect variable $z_{d,s}$. Similar to ASUM [11] and JAS [26], the CLJAS model also assumes that all words in a sentence are generated from the same topic. Therefore, the aspect variable, $z_{d,s}$, is shared by all words in sentence s.

Table 1: The notation used in the CLJAS model.

$D/M/N/T$	number of reviews/sentences/words/aspects
$V_{\text{src/tgt}}$	vocabulary size of the source/target language
$d/s/w/z$	a review/sentence/word/aspect
x	language label, src (source) vs. tgt (target)
v	subjectivity label, fct (fact) vs. opn (opinion)
l	sentiment label, pos (positive) vs. neg (negative)
y	intermediate variable indicating a word in either the source or the target language
ϕ	multinomial distribution over words
θ	multinomial distribution over aspects
π	binomial distribution over sentiment labels
η	binomial distribution over language labels
ζ	binomial distribution over subjectivity labels
$\tau^{\text{src},w}_{\text{tgt},y} / \tau^{\text{tgt},w}_{\text{src},y}$	probability of the translation from a target (source) word y to the given source (target) word w
$\boldsymbol{w}$	list of words in the entire corpora in both source and target languages
$\mathbf{z}/\mathbf{z}_{\neg(d,s)}$	assignments of aspects to all sentences (excluding sentence s in review d)
$\mathbf{v}/\mathbf{v}_{\neg(d,s,n)}$	assignments of subjectivity labels to all words (excluding the nth word in sentence s of review d)
$\mathbf{l}/\mathbf{l}_{\neg(d,s,n)}$	assignments of sentiment labels to all words (excluding the nth word in sentence s of review d)
$\mathbf{x}/\mathbf{x}_{\neg(d,s,n)}$	assignments of language labels to all words (excluding the nth word in sentence s of review d)
$\mathbf{y}/\mathbf{y}_{\neg(d,s,n)}$	assignments of intermediate variables to all words (excluding the nth word in sentence s of review d)
c^d/c^d_t	in review d, number of sentences (is assigned aspect t)
$c^d_{\text{src/tgt}}$	in review d, number of times any word is assigned language label src/tgt
$f^{d,s,\text{src/tgt,fct}}$	in review d, number of times any word of sentence s is assigned subjectivity label fct and language label src/tgt
$f^{d,s,\text{src/tgt,opn,pos/neg}}$	in review d, number of times any word of sentence s is assigned language label src/tgt, subjectivity label opn and sentiment label pos/neg
$c^{\text{src/tgt},t,\text{fct}}_{(y)}$	number of times any word is assigned language label src/tgt, subjectivity label fct (assigned to intermediate word y) under aspect t
$c^{\text{src/tgt},t,\text{opn,pos/neg}}_{(y)}$	number of times any word is assigned language label src/tgt, subjectivity label opn and sentiment label pos/neg (assigned to intermediate word y) under aspect t
$c^{d,s,\text{opn}}_{\text{pos/neg}}$	number of total words in sentence s is assigned subjectivity label opn and sentiment label pos/neg

Essentially, CLJAS is an LDA-based model. In the traditional LDA model, topics are associated with documents and words are associated with topics. CLJAS extends LDA by first adding two variables, i.e., sentiment label and subjectivity label. Specifically, in the CLJAS model, topics are associated with documents, sentiment labels are associated with subjectivity labels, and words are associated with topics, sentiment labels and subjectivity label-

Algorithm 1 The generation process of the CLJAS model on the target language side.

1. For each aspect z:
   a. For the source language:
      (1) Draw a multinomial distribution over words, $\phi^{\text{src},z} \sim Dir(\beta^{\text{src}})$;
      (2) Draw a multinomial distribution over positive and negative words, respectively, $\phi^{\text{src},z,\text{pos}} \sim Dir(\beta^{\text{src,pos}}), \phi^{\text{src},z,\text{neg}} \sim Dir(\beta^{\text{src,neg}})$;
   b. For the target language:
      (1) Draw a multinomial distribution over words, $\phi^{\text{tgt},z} \sim Dir(\beta^{\text{tgt}})$;
      (2) Draw a multinomial distribution over positive and negative words, respectively, $\phi^{\text{tgt},z,\text{pos}} \sim Dir(\beta^{\text{tgt,pos}}), \phi^{\text{tgt},z,\text{neg}} \sim Dir(\beta^{\text{tgt,neg}})$;
2. For each document d in the corpus of the target language:
   a. Draw a multinomial distribution, $\theta^d \sim Dir(\alpha^{\text{tgt}})$;
   b. Draw a binomial distribution, $\eta^d \sim Beta(\gamma^{\text{tgt}})$;
   c. For each sentence s in review d:
      (1) Draw an aspect, $z_{d,s} \sim \theta^d$;
      (2) For each subjectivity label v, draw a binomial distribution, $\pi^{d,s,v} \sim Beta(\mu)$;
      (3) For each word $w_{d,s,n}$ in the sentence s:
         (a) Draw a language label, $x_{d,s,n} \sim \eta^d$;
         (b) Choose a subjectivity label, $v_{d,s,n} \sim \zeta^{d,s,n}$;
         (c) If $v_{d,s,n} = $ opn:
            i. Choose a sentiment label, $l_{d,s,n} \sim \pi^{d,s,\text{opn}}$
            ii. If $x_{d,s,n} = $ tgt, draw a word in target language $y_{d,s,n} \sim \phi^{\text{tgt},z_{d,s,n},l_{d,s,n}}$; Generate a word directly, $w_{d,s,n} = y_{d,s,n}$;
            iii. If $x_{d,s,n} = $ src, draw a word in source language $y_{d,s,n} \sim \phi^{\text{src},z_{d,s,n},l_{d,s,n}}$; Generate a word by translating from $y_{d,s,n}$ according to the translation probability, $\tau^{\text{src},w_{d,s,n}}_{\text{tgt},y_{d,s,n}}$;
         (d) If $v_{d,s,n} = $ fct:
            i. Choose a sentiment label, $l_{d,s,n} \sim \pi^{d,s,\text{fct}}$;
            ii. If $x_{d,s,n} = $ tgt, draw a word in the target language $y_{d,s,n} \sim \phi^{\text{tgt},z_{d,s,n}}$; Generate a word directly, $w_{d,s,n} = y_{d,s,n}$;
            iii. If $x_{d,s,n} = $ src, draw a word in the source language $y_{d,s,n} \sim \phi^{\text{src},z_{d,s,n}}$; Generate a word by translating from $y_{d,s,n}$ according to the translation probability, $\tau^{\text{src},w_{d,s,n}}_{\text{tgt},y_{d,s,n}}$.

s. For each aspect z, CLJAS learns three multinomial distributions over words, which model its objective semantics, ϕ^z, and its positive and negative sentiments, $\phi^{z,\text{pos}}$ and $\phi^{z,\text{neg}}$, respectively. In the CLJAS model, one draws an aspect $z_{d,s}$ for each sentence s of any review d from a document-specific distribution over aspects θ^d. Then, for each word $w_{d,s,n}$ in sentence s, one draws a subjectivity label, $v_{d,s,n}$, and a sentiment label, $l_{d,s,n}$, where $v_{d,s,n}$ indicates that $w_{d,s,n}$ is a sentiment-bearing word or an objective one, whilst $l_{d,s,n}$ further suggests that $w_{d,s,n}$, if it is sentiment-bearing, conveys a positive or a negative sentiment. Finally, one draws the word, $w_{d,s,n}$, from a word distribution corresponding to objective semantics ϕ^z, or the positive $\phi^{z,\text{pos}}$ or negative $\phi^{z,\text{neg}}$ sentiment of the topic z.

Note that the CLJAS model differs from the state-of-the-art JST and ASUM models in that it first explicitly separates sentiment-bearing words from objective ones and then assigns their sentiment polarities. In addition, since CLJAS is a cross-lingual model, an extra language label $x_{d,s,n}$ is introduced to word $w_{d,s,n}$. Its role will be described later on. Therefore, for each word $w_{d,s,n}$, there are three indicator variables associated with it, namely, language label $x_{d,s,n}$, subjectivity label $v_{d,s,n}$ and sentiment label $l_{d,s,n}$. Algo-

rithm 1 presents the generation process of the CLJAS model on the target language side. The generation process on the source language side is not presented, but it can be deduced likewise.

In order to leverage the resources in the source language to improve sentiment analysis in the target language, CLJAS jointly detects aspects and sentiments of the two languages simultaneously. Upon the above description on one side, next we explain its cross-lingual mechanism. In CLJAS, each aspect $z_{d,s}$ corresponds to two clusters of word distributions, namely, $(\phi^{\mathrm{src},z}, \phi^{\mathrm{src},z,\mathrm{pos}}, \phi^{\mathrm{src},z,\mathrm{neg}})$ for the source language and $(\phi^{\mathrm{tgt},z}, \phi^{\mathrm{tgt},z,\mathrm{pos}}, \phi^{\mathrm{tgt},z,\mathrm{neg}})$ for the target language. The binary switch variable, $x_{d,s,n}$, is used to indicate that the intermediate variable $y_{d,s,n}$ is generated from the word distributions of the source language or the target language. In more detail, let's take a word, $w_{d,s,n}$, of review d in the target language as an example. If $x_{d,s,n} = \mathrm{tgt}$, suggesting that $y_{d,s,n}$ is directly generated from the word distributions of the target language, then we choose a specific word distribution from the target language cluster $(\phi^{\mathrm{tgt},z}, \phi^{\mathrm{tgt},z,\mathrm{pos}}, \phi^{\mathrm{tgt},z,\mathrm{neg}})$ and $w_{d,s,n}$ is equal to $y_{d,s,n}$; If $x_{d,s,n} = \mathrm{src}$, suggesting $y_{d,s,n}$ is generated from the word distributions for the source language, then we choose a specific word distribution from the source language cluster $(\phi^{\mathrm{src},z}, \phi^{\mathrm{src},z,\mathrm{pos}}, \phi^{\mathrm{src},z,\mathrm{neg}})$, and $w_{d,s,n}$ is equal to the word translated from $y_{d,s,n}$.

In the model, asymmetric hyperparameters $\gamma^{\mathrm{src/tgt}}_{\mathrm{src/tgt}}$ indicate the prior knowledge on choosing between the source and target languages. They enforce knowledge transfer from the auxiliary reviews in the source language to sentiment analysis in the target language. In particular, for reviews in the source language, we set $\gamma^{\mathrm{src}}_{\mathrm{tgt}} > \gamma^{\mathrm{src}}_{\mathrm{src}}$ such that they can contribute more to topic modeling on reviews in the target language; for reviews in the target language, we set $\gamma^{\mathrm{tgt}}_{\mathrm{src}} > \gamma^{\mathrm{tgt}}_{\mathrm{tgt}}$ such that they can adopt more knowledge from the source language.

Finally, the CLJAS model employs Gibbs sampling [9] for inference. Specifically, for the reviews in the target language, we first draw $z_{d,s}$ by the conditional probability presented in Equation (1). Then, we can jointly draw values for $v_{d,s,n}, l_{d,s,n}, x_{d,s,n}$, and $y_{d,s,n}$ as the conditional probabilities presented in Equations (2)-(5), respectively. For the sake of space limitation, we will not go through the corresponding details.

4. EXPERIMENTAL VALIDATION

In this section, we validate the proposed CLJAS model via extensive experiments. Specifically, we first examine its validity on discovering aspects, validate the effectiveness of its cross-lingual mechanism, and study the effect of the bilingual dictionary on its behavior. Next, we investigate its effectiveness and performance by applying it to practical sentiment classification tasks and quantitatively comparing with four baselines.

4.1 Experimental Settings

4.1.1 Datasets and Auxiliary Resources

In order to validate the proposed CLJAS model, we conducted experiments on real datasets that contain hotel reviews and product reviews in different languages (including English, Chinese, French, German, Spanish, Dutch, and Italian), collected from a few well-known websites. In the experiments, English is taken as the unique source language and the others are target ones. The product reviews cover four domains, i.e., electronics, kitchen, network, and health. In more detail, for English, the hotel dataset contains 12000 reviews collected from Booking.com[1], and the product dataset in

each domain has 2000 reviews collected from Amazon[2][3]. For Chinese, the hotel dataset contains 12000 reviews, downloaded from an open source website[4], and the product dataset contains 2000 reviews for each domain, collected from Jingdong[5], one of the most popular e-shopping websites in China. For other languages (including French, German, Spanish, Dutch, and Italian), the datasets are collected from Booking.com, each of which contains 4000 hotel reviews. All of the datasets were preprocessed with sentence segmentation. For Chinese, we further used ICTCLAS[6] to conduct word segmentation over sentences.

Besides datasets, other three resources are needed in the CLJAS model, including a bilingual dictionary between the source and target language, a few seed words and an opinion lexicon of each language. The bilingual dictionary is used to establish correspondences between two different languages. As in a bilingual dictionary a source word usually corresponds to multiple target words (translation words), some of which may not be suitable for the review scenarios, we employed a simple filtering strategy to overcome the interference of those irrelevant translation words. Specifically, given a source word, we first computed the Term Frequency (TF) for each of its corresponding translation words in the corpora in the target language; Next, we chose at most K terms with the highest TF values as candidate translation words; Finally, a translation probability was assigned to each candidate translation word w_i according to $TF(w_i)/\sum_{i=1}^{K} TF(w_i)$. In our experiments, K was set to be 5.

The seed words of each language used in the CLJAS model are subjective words with clear sentiment polarity. They are adopted to identify sentiment polarities of other subjective words. For English, we select 10 positive seed words and 10 negative seed words with highest TF in the corpus. For other languages, the seed words are obtained by translating the English ones. Table 2 presents the full lists of seed words of all target languages, where for each seed word an English translation word is provided in subsequent brackets, which was obtained by *Google Translate*.

In the proposed model, an opinion lexicon is used to separate subjective words from objective ones[7]. The English and Chinese opinion lexicons used in the experiments are acquired from the commonly-used knowledge base, HowNet[8]. The English one contains 3687 terms, while the Chinese one contains 5997 terms. For other languages, the opinion lexicons are obtained by translating the English one using *Google Translate*. If an opinion lexicon of a certain language is not available, all adjectives in the corpus in this language can be considered as subjective words.

4.1.2 Parameters Settings

In the experiments, we ran 100 iterations of Gibbs sampling, which were experimentally proven adequate for obtaining stable results. We set $\alpha_z = 50/T$ for each aspect z and $\beta_w = 0.1$ for each word w in the vocabulary, as in [9]. The default values for β_w^{pos} and β_w^{neg} were set to 0.1. For each positive seed word w, β_w^{neg} was set to 0; for each negative seed word w, β_w^{pos} was set to 0. For each positive non-seed word w, β_w^{neg} was set to 0.01; For each negative non-seed word w, β_w^{pos} was set to 0.01. The value of μ was set to 0.001 for each sentiment label. Without loss of generality,

---

[1]http://www.booking.com

[2]http://www.seas.upenn.edu/m̃dredze/datasets/sentiment/

[3]http://llt.cbs.polyu.edu.hk/l̃ss/ACL2010_Data_SSLi.zip

[4]http://nlp.csai.tsinghua.edu.cn/

[5]http://www.360buy.com

[6]http://ictclas.cn/index.html

[7]Note that the CLJAS model does not use the polarity information of words in the lexicon.

[8]http://www.keenage.com/html/e_index.html

$$P(z_{d,s} = t|\mathbf{z}_{\neg(d,s)}, \mathbf{v}, l, \mathbf{x}, \mathbf{y}, \mathbf{w}) \propto \frac{c_t^d + \alpha_t^{\text{tgt}}}{\sum_{t'=1}^{T}(c_{t'}^d + \alpha_{t'}^{\text{tgt}})} \left(\begin{array}{c} (\frac{\Gamma(\sum_{w=1}^{V_{\text{tgt}}}(c_w^{\text{src},t,\text{fct}} + \beta_w^{\text{tgt}}))}{\Gamma(\sum_{w=1}^{V_{\text{tgt}}}(c_w^{\text{src},t,\text{fct}} + \beta_w^{\text{tgt}}) + f^{d,s,\text{src},\text{fct}})}) \\ \cdot \prod_{w=1}^{V_{\text{src}}}(\frac{\Gamma(c_w^{\text{src},t,\text{fct}} + \beta_w^{\text{src}} + f_w^{d,s,\text{src},\text{fct}})}{\Gamma(c_w^{\text{src},t,\text{fct}} + \beta_w^{\text{src}})}) \end{array} \right) \times$$

$$\prod_{s\in\{\text{pos},\text{neg}\}} \left(\begin{array}{c} (\frac{\Gamma(\sum_{w=1}^{V_{\text{src}}}(c_w^{\text{src},t,\text{opn},l} + \beta_w^{\text{src},l}))}{\Gamma(\sum_{w=1}^{V_{\text{src}}}(c_w^{\text{src},t,\text{opn},l} + \beta_w^{\text{src},l}) + f^{d,s,\text{src},\text{opn},l})}) \\ \cdot \prod_{w=1}^{V_{\text{src}}}(\frac{\Gamma(c_w^{\text{src},t,\text{opn},l} + \beta_w^{\text{src},l} + f_w^{d,s,\text{src},\text{opn},l})}{\Gamma(c_w^{\text{src},t,\text{opn},l} + \beta_w^{\text{src},l})}) \end{array} \right) \times \left(\begin{array}{c} (\frac{\Gamma(\sum_{w=1}^{V_{\text{tgt}}}(c_w^{\text{tgt},t,\text{fct}} + \beta_w^{\text{tgt}}))}{\Gamma(\sum_{w=1}^{V_{\text{tgt}}}(c_w^{\text{tgt},t,\text{fct}} + \beta_w^{\text{tgt}}) + f^{d,s,\text{tgt},\text{fct}})}) \\ \cdot \prod_{w=1}^{V_{\text{tgt}}}(\frac{\Gamma(c_w^{\text{tgt},t,\text{fct}} + \beta_w^{\text{tgt}} + f_w^{d,s,\text{tgt},\text{fct}})}{\Gamma(c_w^{\text{tgt},t,\text{fct}} + \beta_w^{\text{tgt}})}) \end{array} \right) \times$$

$$\prod_{s\in\{\text{pos},\text{neg}\}} \left(\begin{array}{c} (\frac{\Gamma(\sum_{w=1}^{V_{\text{tgt}}}(c_w^{\text{tgt},t,\text{opn},l} + \beta_w^{\text{tgt},l}))}{\Gamma(\sum_{w=1}^{V_{\text{tgt}}}(c_w^{\text{tgt},t,\text{opn},l} + \beta_w^{\text{tgt},l}) + f^{d,s,\text{tgt},\text{opn},l})}) \\ \cdot \prod_{w=1}^{V_{\text{tgt}}}(\frac{\Gamma(c_w^{\text{tgt},t,\text{opn},l} + \beta_w^{\text{tgt},l} + f_w^{d,s,\text{tgt},\text{opn},l})}{\Gamma(c_w^{\text{tgt},t,\text{opn},l} + \beta_w^{\text{tgt},l})}) \end{array} \right) \tag{1}$$

$$P(v_{d,s,n} = \text{fct}, y_{d,s,n} = w_{d,s,n}, x_{d,s,n} = \text{tgt}|\mathbf{z}, \mathbf{v}_{\neg(d,s,n)}, \mathbf{l}_{\neg(d,s,n)}, \mathbf{x}_{\neg(d,s,n)}, \mathbf{y}_{\neg(d,s,n)}, \mathbf{w})$$
$$\propto P(v_{d,s,n} = \text{fct}|w_{d,s,n}) \cdot \frac{c_{\text{tgt}}^d + \gamma_{\text{tgt}}^{\text{tgt}}}{c^d + \gamma_{\text{tgt}}^{\text{tgt}} + \gamma_{\text{src}}^{\text{tgt}}} \cdot \frac{c_{w_{d,s,n}}^{\text{tgt},t,\text{fct}} + \beta_{w_{d,s,n}}^{\text{tgt}}}{\sum_{w=1}^{V_{\text{tgt}}}(c_w^{\text{tgt},t,\text{fct}} + \beta_w^{\text{tgt}})} \tag{2}$$

$$P(v_{d,s,n} = \text{opn}, l_{d,s,n} = l, y_{d,s,n} = w_{d,s,n}, x_{d,s,n} = \text{tgt}|\mathbf{z}, \mathbf{v}_{\neg(d,s,n)}, \mathbf{l}_{\neg(d,s,n)}, \mathbf{x}_{\neg(d,s,n)}, \mathbf{y}_{\neg(d,s,n)}, \mathbf{w})$$
$$\propto P(v_{d,s,n} = \text{opn}|w_{d,s,n}) \cdot \frac{c_l^{d,s,\text{opn}} + \mu_l}{\sum_{l'\in\{\text{pos},\text{neg}\}}(c_{l'}^{d,s,\text{opn}} + \mu_{l'})} \cdot \frac{c_{\text{tgt}}^d + \gamma_{\text{tgt}}^{\text{tgt}}}{c^d + \gamma_{\text{tgt}}^{\text{tgt}} + \gamma_{\text{src}}^{\text{tgt}}} \cdot \frac{c_{w_{d,s,n}}^{\text{tgt},t,\text{opn},l} + \beta_{w_{d,s,n}}^{\text{tgt},l}}{\sum_{w=1}^{V_{\text{tgt}}}(c_w^{\text{tgt},t,\text{opn},l} + \beta_w^{\text{tgt},l})} \tag{3}$$

$$P(v_{d,s,n} = \text{fct}, y_{d,s,n} = y, x_{d,s,n} = \text{src}|\mathbf{z}, \mathbf{v}_{\neg(d,s,n)}, \mathbf{l}_{\neg(d,s,n)}, \mathbf{x}_{\neg(d,s,n)}, \mathbf{y}_{\neg(d,s,n)}, \mathbf{w})$$
$$\propto P(v_{d,s,n} = \text{fct}|w_{d,s,n}) \cdot \frac{c_{\text{tgt}}^d + \gamma_{\text{src}}^{\text{tgt}}}{c^d + \gamma_{\text{tgt}}^{\text{tgt}} + \gamma_{\text{src}}^{\text{tgt}}} \cdot \tau_{\text{src},y}^{\text{tgt},w_{d,s,n}} \cdot \frac{c_y^{\text{src},t,\text{fct}} + \beta_y^{\text{src}}}{\sum_{w=1}^{V_{\text{src}}}(c_w^{\text{src},t,\text{fct}} + \beta_w^{\text{src}})} \tag{4}$$

$$P(v_{d,s,n} = \text{opn}, l_{d,s,n} = l, y_{d,s,n} = y, x_{d,s,n} = \text{src}|\mathbf{z}, \mathbf{v}_{\neg(d,s,n)}, \mathbf{l}_{\neg(d,s,n)}, \mathbf{x}_{\neg(d,s,n)}, \mathbf{y}_{\neg(d,s,n)}, \mathbf{w})$$
$$\propto P(v_{d,s,n} = \text{opn}|w_{d,s,n}) \cdot \frac{c_l^{d,s,\text{opn}} + \mu_l}{\sum_{l'\in\{\text{pos},\text{neg}\}}(c_{l'}^{d,s,\text{opn}} + \mu_{l'})} \cdot \frac{c_{\text{tgt}}^d + \gamma_{\text{src}}^{\text{tgt}}}{c^d + \gamma_{\text{tgt}}^{\text{tgt}} + \gamma_{\text{src}}^{\text{tgt}}} \cdot \tau_{\text{src},y}^{\text{tgt},w_{d,s,n}} \cdot \frac{c_y^{\text{src},t,\text{opn},l} + \beta_y^{\text{src},l}}{\sum_{w=1}^{V_{\text{src}}}(c_w^{\text{src},t,\text{opn},l} + \beta_w^{\text{src},l})} \tag{5}$$

Table 2: The list of seed words for different target languages.

	Positive	Negatvie
Chinese	不错 (not bad), 方便 (convenient), 干净 (clean), 热情 (friendly), 好 (good), 满意 (satisfied), 高兴 (happy), 漂亮 (beautiful), 喜欢 (like), 赞 (great)	不好 (not good), 惨 (miserable), 差 (bad), 陈旧 (old), 丑 (ugly), 恶劣 (wicked), 慢 (slow), 失望 (disappointed), 脏 (dirty), 糟(terrible)
French	bien (good), propreté (clean), calme (quiet), bon (good), propre (clean), aimable (friendly), confort (comfort), bonne (good), agréable (pleasant), excellent (excellent)	cher (expensive), bruyant (nosiy), bruyante (noisy), panne (breakdown), mal (evil), mauvaise (bad), sale (dirty), manque (lack), difficile (difficult), désagréable (unpleasant)
German	gut (good), gute (good), ruhig (quite), freundlich (friendly), freundliche (friendly), sauber (clean), hilfsbereit (helpful), schön (beautiful), guter (good), saubere (clean)	teuer (expensive), lärm (noise), schlecht (bad), schmutzig (dirty), enttäuscht (disappointed), hässlich (ugly), schlechte (poor), schmutzige (dirty), unzufrieden (unhappy), schlechtes (bad)
Spanish	buena (good), limpieza (clean), limpia (clean), buen (good), amable (friendly), excelente (excellent), limpio (clean), bueno (good), perfecto (perfect), contentos (happy)	caro (expensive), pobre (poor), malo (bad), peor (worse), sucio (dirty), sucios (dirty), costoso (expensive), ruidoso (noisy), decepcionado (disappointed), grosero (rude)
Italian	ottima (good), pulizia (clean), buona (good), gentile (kind), comodo (comfortable), ottimo (excellent), bene (good), bella (nice), pulito (clean), cortesia (courtesy)	rumore (noise), scarsa (poor), sporco (dirty), male (bad), scarso (poor), cattivo (bad), rumoroso (noisy), scortese (rude), deluso (disappointed), sgradevole (unpleasant)
Dutch	goed (good), vriendelijk (friendly), prima (fine), goede (well), schoon (clean), mooi (beautiful), uitstekend (excellent), behulpzaam (helpful), netjes (neatly), lekker (nice),	lawaai (noise), jammer (pity), slecht (bad), gehorig (noisy), duur (expensive), vies (dirty), pover (poor), slechter (worse), teleurstellend (disappointing), lelijke (ugly)

Table 3: The example aspects discovered by the CLJAS model.

Topic 2	Food	breakfast, hotel, coffee, room, food, restaurant, buffet, fruit, eggs, service, morning, pastries, wine, served, continental, staff, cheese, dinner, good, great, nice, bar, excellent, fresh, lovely, best, cold, hot, wonderful, delicious
	食物	品种 (variety), 不错 (good), 酒店 (hotel), 服务员 (waiter), 一般 (just so-so), 没什么 (nothing), 服务 (service), 送(deliver), 差 (bad), 丰盛 (rich), 份 (dish), 菜 (vegetable), 鸡蛋 (egg), 饱 (full), 早餐 (breakfast), 自助 (buffet), 味道 (taste),丰富 (abundant), 很好 (nice), 简直 (simply), 每天 (everyday), 粥 (porridge), 东西 (stuff), 早饭 (breakfast), 西式 (western), 单调 (tedious), 恭维 (compliment), 人气(popularity), 咸菜 (pickles), 美味 (yummy)
Topic 6	Location	hotel, walk, location, station, Rome, metro, walking, minutes, bus, steps, Vatican, distance, termini, Spanish, train, city, restaurants, minute, area, away, great, close, good, easy, right, quiet, attractions, nice, major, convenient
	位置	方便 (convenient), 酒店 (hotel), 离 (distance), 近 (close), 位置 (location), 打车 (taxi), 火车站 (railway station), 购物(shopping), 机场 (airport), 出门 (go out), 车 (car), 不错 (good), 远 (far), 地铁站 (subway station), 坐 (sit), 唯一 (sole), 交通 (traffic), 地理(location), 分钟 (minute), 市中心 (downtown), 很好 (nice), 距离 (distance), 宾馆 (hotel), 周围 (nearby), 左右 (approximately), 海边 (seaside), 地处(situated), 不到 (less than), 到达 (arrive), 边 (side)
Topic 7	Service	staff, hotel, location, English, service, desk, breakfast, room, reception, stay, spoke, restaurant, Rome, Italian, restaurants, concierge, people, helpful, friendly, great, good, clean, excellent, nice, pleasant, wonderful, help, polite, courteous, comfortable
	服务	服务 (service), 差 (bad), 态度 (attitude), 前台 (receptionist), 酒店 (hotel), 服务员 (waiter), 员 (staff), 客人 (customer), 大堂 (hall manager), 感觉 (feeling), 员工 (staff), 服务生 (waiter), 理 (respond), 小姐 (waitress), 不错 (good), 很好 (nice), 热情 (friendly), 慢(slow), 微笑 (smile), 素质 (quality), 行李 (luggage), 高 (high), 礼貌 (polite), 候 (wait), 笑容 (smile), 到位 (proper), 打招呼 (greet), 效率(efficiency), 早餐 (breakfast), 结帐 (check out)

the above parameters were set the same as in [26]. For the target language, the values of γ_{src}^{tgt} and γ_{tgt}^{tgt} are asymmetrical. We fixed $\gamma_{tgt}^{tgt} = 0.001$, but changed γ_{src}^{tgt} from 0.001 to 1 to investigate the effect of language transfer on model performance.

4.1.3 Baselines

For the purpose of comparison, we adopted four representative models as the baselines. We chose USL (Universal Sentiment Lexicon) and SVM as baselines, because USL is a typical unsupervised approach to sentiment classification, where a universal opinion lexicon collected from HowNet is adopted to determine the polarity of a review, and SVM is a typical supervised model for sentiment classification. We also employed the state-of-the-art ASUM and JAS models as baselines, as they both consider aspect in sentiment analysis. Note that both ASUM and JAS are monolingual models, whilst CLJAS is a cross-lingual model. The purpose of comparing CLJAS with these two monolingual aspect-sentiment models is mainly to examine the effectiveness of the cross-lingual mechanism implemented in CLJAS.

4.2 Experimental Results

4.2.1 Aspect Discovery

This experiment is to investigate the performance of the developed CLJAS model on discovering aspects on given datasets. For the purpose of illustration, Table 3 presents three major aspects discovered by CLJAS on hotel reviews, where, for each aspect, top 30 words are listed and sorted by its word probability distribution in a descending order.

It can be seen in Table 3 that the CLJAS model can effectively detect major aspects from the reviews in both the source and target languages. The extracted words under each aspect are quite coherent and meaningful. Moreover, the CLJAS model can obtain bilingually aligned aspects due to the mechanism of sharing the same topic distribution. This mechanism enables us to improve the

aspect modeling on the data in the target language by leveraging the rich resources in the source language.

4.2.2 Cross-Lingual Mechanism

This experiment aims to evaluate the proposed cross-lingual mechanism. For this purpose, we calculated the perplexity [2] of the test set under different γ_{src}^{tgt}, which controls the degree of knowledge transfer between the source and target languages, i.e., the extent that the knowledge learned from the dataset in the source language is used to learn knowledge for the dataset in the target language. It is known that the lower the perplexity, the better the performance of the topic model [2].

In the experiment, for each domain, 2000 reviews were used for model training and the rest reviews for inference. Table 4 presents the perplexity values of the test sets in Chinese under different domains. As seen in Table 4, with the increase of γ_{src}^{tgt}, the perplexity under different domains basically decreases in a monotonic manner. Therefore, the performance of the model was improved by introducing the cross-lingual mechanism. The improvement mainly benefits from the knowledge transferred from the data in the source language. According to the experimental results, we observe that within a certain range, the larger the γ_{src}^{tgt}, the more improvement that can be achieved. However, γ_{src}^{tgt} cannot be too large; Otherwise, it leads to the model to learn knowledge almost completely from the source language and neglect the characteristics of the target language.

We further evaluated the proposed cross-lingual framework on other languages. Taking English as the source language, Table 5 presents the perplexity values of test sets in different languages under the hotel domain. It can be observed from Table 5 that, with the increase of γ_{src}^{tgt}, the perplexity monotonically decreases under different languages. Therefore, we can claim that the proposed cross-lingual mechanism is effective for different domains and for different languages.

4.2.3 Effect of The Bilingual Dictionary

Table 4: The perplexity of test sets in Chinese under different domains with exponentially increasing γ_{src}^{tgt}.

$\gamma_{src}^{tgt} \rightarrow$	0.001	0.01	0.1	1
Electronics	891.577	886.291	875.624	829.403
Kitchen	806.246	804.469	788.822	751.710
Network	585.608	582.498	577.552	562.346
Health	696.527	690.537	680.129	654.075
Hotel	133.922	129.507	130.738	125.999
Average	534.835	529.835	524.656	508.012

Table 5: The perplexity of test sets in different languages under the hotel domain with exponentially increasing γ_{src}^{tgt}.

$\gamma_{src}^{tgt} \rightarrow$	0.001	0.01	0.1	1
French	664.795	661.59	659.498	658.047
German	1090.84	1067.38	1062.16	1044.94
Spanish	657.153	651.546	649.757	629.508
Dutch	868.732	866.921	855.115	831.888
Italian	1038.45	1031.68	1035.33	1024.94
Average	863.994	855.823	852.372	837.865

As aforesaid, in the CLJAS model a bilingual dictionary is needed to bridge the two languages. In order to investigate the dependency of the performance of the CLJAS model on the quality and scale of the bilingual dictionary, we compared the perplexity of the same test datasets of different domains in Chinese with two different bilingual dictionaries, which contain 224,385 and 41,814 pairs of translation words, respectively. Figure 2 presents the experimental results. We can observe that there is no significant difference between the results corresponding to the two dictionaries of different size. This observation suggests that the proposed CLJAS model is basically stable regarding the quality and scale of the bilingual dictionary.
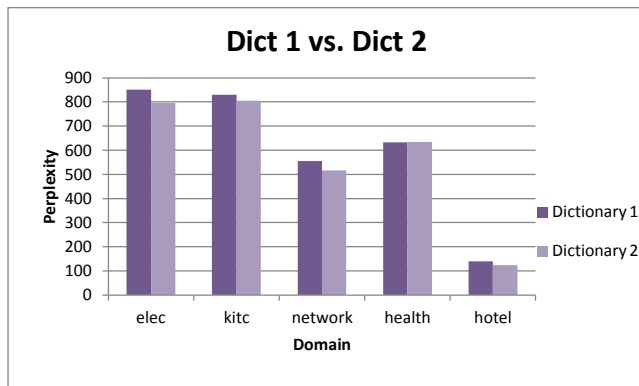

Figure 2: The perplexity of Chinese test sets in different domains with different bilingual dictionaries.

4.2.4 Sentiment Classification

In this experiment, we investigate the validity and performance of the CLJAS model by evaluating the quality of the aspect-dependent sentiment lexicons extracted using CLJAS. Specifically, we first adopted CLJAS to extract aspect-dependent sentiment lexicons and further applied them to practical sentiment classification tasks. Nex-

t, we evaluated the quality of the aspect-dependent sentiment lexicons according to how much they improve the performance of practical sentiment classification, as compared to the aforesaid four baselines. Table 6 (a) presents the results of both CLJAS and the baselines in sentiment classification tasks in different domains. We also compared CLJAS with the baselines on hotel review datasets in different languages. The multilingual experimental results are presented in Table 6 (b). In these two tables, the accuracy of sentiment classification on both positive and negative reviews as well as their combination is listed for all methods. The bottom row presents the accuracy of sentiment classification averaged on different domains/languages.

As CLJAS is an unsupervised model, we first compared it with the three unsupervised baselines, i.e., USL, ASUM, and JAS. It can be seen in Table 6 (a) that for datasets in different domains CLJAS consistently outperforms USL, as it increases the accuracy of sentiment classification by 0.087. This is because the sentiment lexicon extracted by CLJAS for determining the sentiment polarities of reviews is aspect-dependent, while the sentiment lexicon used in USL is aspect-independent. For datasets in other languages, the performance superiority of CLJAS is more obvious. Specifically, as observed in Table 6 (b), CLJAS performs better than USL by a performance improvement of 0.173 on average. In USL, since the multilingual opinion lexicons are obtained by translating the English one, they suffer from the loss of accuracy caused by the translation. Therefore, the performance of the unsupervised method based on translation for multilingual sentiment analysis is unsatisfactory. As for ASUM and JAS, we can find in Table 6 (a) that CLJAS improves the accuracy of sentiment classification by 0.061 and 0.063, respectively, while in Table 6 (b) the corresponding accuracy improvement is 0.126 and 0.141. That is to say, CLJAS performs better than ASUM and JAS as well. This is because CLJAS combine the corpora of two languages together and can thus capture more statistics characteristics than ASUM and JAS that depends only on monolingual corpora.

We also compared CLJAS with the typical supervised learning approach, SVM. Specifically, for each domain, 50% data was used for training and the rest 50% data was used for testing. Particularly, we first employed LibSVM[9] to train a sentiment classifier and then predicted sentiment polarities on test data. It should be mentioned that we tried different kernel functions to refine the performance of SVM and finally adopted a linear one as the kernel function, with which LibSVM gets the best results. From Table 6 (a), we can observe that CLJAS outperforms SVM on average, as it increases the accuracy of sentiment classification by 0.03. This counterintuitive phenomenon may be because the training data of product reviews in different domains used in SVM is of poor quality, whilst the unlabeled data used in CLJAS is of relatively high quality. From Table 6 (b), we can observe that the performance of CLJAS is inferior to that of SVM. This is intuitive, as CLJAS is an unsupervised method while SVM is supervised.

In general, CLJAS performs better than USL, ASUM, and JAS in different domains and different languages. Therefore, it can be concluded that no matter in what domains and in what languages, the CLJAS model is effective and exhibits better performance than the state-of-the-art unsupervised models, which highlights its value in practical applications.

5. CONCLUSIONS

In this paper, we have proposed for the first time an unsupervised Cross-Lingual Aspect/Sentiment (CLJAS) model for senti-

---

[9]http://www.csie.ntu.edu.tw/c̃jlin/libsvm/

Table 6: The accuracy of sentiment classification with (a) different domains and (b) different languages.

(a)

	USL			SVM			ASUM			JAS			CLJAS		
	pos	neg	all	pos	neg	all	pos	neg	all	pos	neg	all	pos	neg	all
Electronics	0.718	0.730	0.724	0.796	0.800	0.798	0.797	0.737	0.764	0.749	0.715	0.731	0.772	0.770	0.771
Kitchen	0.698	0.808	0.741	0.793	0.793	0.793	0.687	0.813	0.734	0.732	0.708	0.719	0.743	0.825	0.778
Network	0.696	0.691	0.694	0.791	0.750	0.769	0.769	0.660	0.701	0.770	0.704	0.732	0.757	0.792	0.774
Health	0.691	0.700	0.696	0.644	0.668	0.655	0.723	0.666	0.691	0.748	0.694	0.718	0.763	0.819	0.788
Hotel	0.715	0.686	0.700	0.800	0.857	0.826	0.737	0.86	0.793	0.833	0.743	0.775	0.841	0.915	0.878
Average	0.704	0.723	0.711	0.765	0.774	0.768	0.743	0.747	0.737	0.766	0.713	0.735	0.775	0.824	0.798

(b)

	USL			SVM			ASUM			JAS			CLJAS		
	pos	neg	all	pos	neg	all	pos	neg	all	pos	neg	all	pos	neg	all
French	0.72	0.696	0.707	0.895	0.884	0.889	0.747	0.785	0.764	0.729	0.739	0.734	0.887	0.867	0.877
German	0.777	0.641	0.687	0.926	0.900	0.913	0.876	0.707	0.768	0.738	0.739	0.738	0.863	0.860	0.862
Spanish	0.646	0.640	0.643	0.875	0.878	0.877	0.824	0.666	0.720	0.693	0.718	0.705	0.906	0.848	0.875
Italian	0.689	0.677	0.683	0.897	0.858	0.877	0.845	0.639	0.698	0.716	0.701	0.709	0.841	0.849	0.845
Dutch	0.683	0.752	0.712	0.918	0.896	0.907	0.862	0.652	0.714	0.698	0.717	0.707	0.801	0.883	0.837
Average	0.703	0.681	0.686	0.902	0.883	0.892	0.831	0.690	0.733	0.715	0.723	0.718	0.859	0.861	0.859

ment analysis. The CLJAS model requires neither parallel corpora nor labeled sentiment reviews. Through the cross-lingual mechanism, the model can improve sentiment analysis in a target language by leveraging data and knowledge available in a source language. Particularly, CLJAS can obtain bilingually aligned aspects from the reviews in the same domain due to the mechanism of sharing the same topic distribution. CLJAS can also extract aspect-dependent sentiment lexicons that are essential for sentiment analysis. We have further compared the CLJAS model with both supervised and unsupervised representative baselines by applying them to practical sentiment classification tasks. Experimental results have demonstrated that as compared to the unsupervised baselines, CLJAS can significantly improve the accuracy of sentiment classification, while comparing to the supervised baseline CLJAS also exhibits comparable performance, which have convincingly validated its effectiveness and highlighted its value.

Acknowledgement

6. REFERENCES

[1] C. Banea, R. Mihalcea, J. Wiebe, and S. Hassan. Multilingual subjectivity analysis using machine translation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 127–135, 2008.

[2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Machine Learning Research*, 3:993–1022, January 2003.

[3] J. Boyd-Graber and D. M. Blei. Multilingual topic models for unaligned text. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*, pages 75–82, 2009.

[4] J. Boyd-Graber and P. Resnik. Holistic sentiment analysis across languages: Multilingual supervised latent dirichlet allocation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 45–55, 2010.

[5] S. Brody and N. Elhadad. An unsupervised aspect-sentiment model for online reviews. In *Proceedings of Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 804–812, 2010.

[6] Z. C ̈acilia, s. N. Mathia, S. Heiner, and S. Michael. Fine-grained sentiment analysis with structural features. In *Proceedings of the 5th International Joint Conference on Natural Language Processing*, page 336–344, 2011.

[7] E. Erosheva, S. Fienberg, and J. Lafferty. Mixed-membership models of scientific publications. In *Proceedings of the National Academy of Sciences of the United States of America*, pages 5220–5227, 2004.

[8] K. Fukumasu, K. Eguchi, and E. Xing. Symmetric correspondence topic models for multilingual text analysis. In *Proceedings of Advances in Neural Information Processing Systems*, 2012.

[9] T. L. Griffiths and M. Steyvers. Finding scientific topics. In *Proceedings of the National Academy of Sciences*, pages 5228–5335, 2004.

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

[11] Y. Jo and A. H. Oh. Aspect and sentiment unification model for online review analysis. In *Proceeding of WSDM '11*, pages 815–824, 2011.

[12] C. Lin and Y. He. Joint sentiment/topic model for sentiment analysis. In *Proceeding of the 18th ACM conference on Infor-mation and knowledge management*, pages 375–384, 2009.

[13] B. Liu. Sentiment analysis and subjectivity. *Handbook of natural language processing*, 2:568, 2010.

[14] Y. Lu, M. Castellanos, U. Dayal, and et al. Automatic construction of a context-aware sentiment lexicon: an optimization approach. In *Proceedings of the 20th international conference on World Wide Web*, pages 347–356, 2011.

[15] A. Meena and T. Prabhakar. Sentence level sentiment analysis in the presence of conjuncts using linguistic analysis. In *Advances in Information Retrieval*, pages 573–580. Springer, 2007.

[16] R. Moraes, J. F. Valiati, and W. P. G. Neto. Document-level sentiment classification: An empirical comparison between svm and ann. *Expert Systems with Applications*, 40:621–633, February 2013.

[17] D. Newman, C. Chemudugunta, P. Smyth, and M. Steyvers. Statistical entity-topic models. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 680–686, 2006.

[18] B. Pang and L. Lee. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd annual meeting on Association for Computational Linguistics*, page 271. Association for Computational Linguistics, 2004.

[19] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, pages 79–86, 2002.

[20] K. Suin, Z. Jianwen, C. Zheng, H. O. Alice, and L. Shixia. A hierarchical aspect-sentiment model for online reviews. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*. AAAI, 2013.

[21] L. K.-W. Tan, J.-C. Na, Y.-L. Theng, and K. Chang. Sentence-level sentiment polarity classification using a linguistic approach. In *Digital Libraries: For Cultural Heritage, Knowledge Dissemination, and Future Creation*, pages 77–87. Springer, 2011.

[22] P. D. Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 417–424, 2002.

[23] X. Wan. Co-training for cross-lingual sentiment classification. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1*, pages 235–243, 2009.

[24] J. Wiebe and E. Riloff. Creating subjective and objective sentence classifiers from unannotated texts. In *Computational Linguistics and Intelligent Text Processing*, pages 486–497. Springer, 2005.

[25] J. Wiebe, T. Wilson, R. Bruce, M. Bell, and M. Martin. Learning subjective language. *Computational linguistics*, 30(3):277–308, 2004.

[26] X. Xu, S. Tan, Y. Liu, X. Cheng, and Z. Lin. Towards jointly extracting aspects and aspect-specific sentiment knowledge. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 1895–1899. ACM, 2012.

[27] T. Yao and L. Li. A kernel-based sentiment classification approach for chinese sentences. In *Computer Science and Information Engineering, 2009 WRI World Congress on*, volume 5, pages 513–518. IEEE, 2009.

[28] A. Yessenalina, Y. Yue, and C. Cardie. Multi-level structured models for document-level sentiment classification. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1046–1056. Association for Computational Linguistics, 2010.

[29] D. Zhang, Q. Mei, and C. Zhai. Cross-lingual latent topic extraction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1128–1137, 2010.

[30] W. X. Zhao, J. Jiang, H. Yan, and X. Li. Jointly modeling aspects and opinions with a maxent-lda hybrid. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 56–65, 2010.