No.1

Aspect Level Sentiment Classification with Deep Memory Network

深度记忆网络的方面等级情绪分类

Abstract

We introduce a deep memory network for aspect level sentiment classiﬁcation. Unlike feature-based SVM and sequential neural models such as LSTM, this approach explicitly captures the importance of each context word when inferring the sentiment polarity of an aspect. Such importance degree and text representation are calculated with multiple computational layers, each of which is a neural attention model over an external memory. Experiments on laptop and restaurant datasets demonstrate that our approach performs comparable to state-of-art feature based SVM system, and substantially better than LSTM and attention-based LSTM architectures. On both datasets we show that multiple computational layers could improve the performance. Moreover, our approach is also fast. The deep memory network with 9 layers is 15 times faster than LSTM with a CPU implementation.

我们引入了一个用于方面级情绪分类的深存储器网络。 不像基于特征的SVM和顺序神经模型，如LSTM，这种方法明确地捕获每个上下文词在推断一个方面的情感极性时的重要性。 使用多个计算层来计算这样的重要度和文本表示，每个计算层是在外部存储器上的神经注意模型。 笔记本电脑和餐厅数据集的实验表明，我们的方法可以与基于先进特性的SVM系统相媲美，并且比LSTM和基于注意的LSTM体系结构好得多。 在两个数据集上，我们显示多个计算层可以提高性能。 此外，我们的方法也很快。 具有9层的深存储器网络比具有CPU实现的LSTM快15倍。

1 Introduction

Aspect level sentiment classiﬁcation is a fundamental task in the ﬁeld of sentiment analysis (Pang and Lee, 2008; Liu, 2012; Pontiki et al., 2014). Given a sentence and an aspect occurring in the sentence, this task aims at inferring the sentiment polarity (e.g. positive, negative, neutral) of the aspect. For example, in sentence “great food but the service was dreadful!”, the sentiment polarity of aspect “food” is positive while the polarity of aspect "service" is negative.Researchers typically use machine learning algorithms and build sentiment classiﬁer in a supervised manner. Representative approaches in literature include feature based Support Vector Machine (Kiritchenko et al., 2014; Wagneret al., 2014) and neural network models (Dong et al., 2014; Lakkaraju et al., 2014; Vo and Zhang, 2015; Nguyen and Shirai, 2015; Tang et al., 2015a). Neural models are of growing interest for their capacity to learn text representation from data without careful engineering of features, and to capture semantic relations between aspect and context words in a more scalable way than feature based SVM.

方面层面的情绪分类是情绪分析的一个基本任务（Pang和Lee，2008; Liu，2012; Pontiki et al，2014）。给定句子和在句子中出现的方面，该任务旨在推断该方面的情感极性（例如正，负，中性）。例如，在句子“伟大的食物，但服务是可怕的”，方面“食物”的情感极性是积极的，而方面“服务”的极性是负的。研究人员通常使用机器学习算法和建立情感分类在监督方式。文献中的代表性方法包括基于特征的支持向量机（Kiritchenko等，2014; Wagneret等，2014）和神经网络模型（Dong等，2014; Lakkaraju等，2014; Vo和Zhang，2015; Nguyen和Shirai，2015; Tang等，2015a）。神经模型对于他们从数据学习文本表示而不仔细设计特征以及以比基于特征的SVM更可扩展的方式捕获方面和上下文词之间的语义关系的能力越来越感兴趣。

Despite these advantages, conventional neural models like long short-term memory (LSTM) (Tang et al., 2015a) capture context information in an implicit way, and are incapable of explicitly exhibiting important context clues of an aspect. We believe that only some subset of context words are needed to infer the sentiment towards an aspect. For example, in sentence “great food but the service was dreadful!”, “dreadful” is an important clue for the aspect “service” but “great” is not needed. Standard LSTM works in a sequential way and manipulates each context word with the same operation, so that it cannot explicitly reveal the importance of each context word. A desirable solution should be capable of explicitly capturing the importance of context words and using that information to build up features for the sentence after given an aspect word. Furthermore, a human asked to do this task will selectively focus on parts of the contexts, and acquire information where it is needed to build up an internal representation towards an aspect in his/her mind.

尽管有这些优点，常规神经模型如长期短期记忆（LSTM）（Tang等人，2015a）以隐式方式捕获上下文信息，并且不能明确地展示一方面的重要上下文线索。我们认为只需要上下文词的一些子集来推断一个方面的情感。例如，在句子“伟大的食物，但服务是可怕的”，“可怕”是方面“服务”的重要线索，但“伟大”不需要。标准LSTM以顺序方式工作，并且以相同的操作来操纵每个上下文词，使得它不能显式地揭示每个上下文词的重要性。理想的解决方案应该能够明确地捕获上下文单词的重要性并且使用该信息来在给定方面字之后为该句子建立特征。此外，被要求完成这项任务的人将选择性地集中于部分上下文，并且获取在其中需要建立对他/她的头脑中的一个方面的内部表示的信息。

In pursuit of this goal, we develop deep memory network for aspect level sentiment classiﬁcation, which is inspired by the recent success of computational models with attention mechanism and explicit memory (Graves et al., 2014; Bahdanau et al., 2015; Sukhbaatar et al., 2015). Our approach is data-driven, computationally efﬁcient and does not rely on syntactic parser or sentiment lexicon. The approach consists of multiple computational layers with shared parameters. Each layer is a content-and location- based attention model, which ﬁrst learns the importance/weight of each context word and then utilizes this information to calculate continuous text representation. The text representation in the last layer is regarded as the feature for sentiment classiﬁcation. As every component is differentiable, the entire model could be efﬁciently trained end-to-end with gradient descent, where the loss functionis the cross-entropy error of sentiment classiﬁcation.

We apply the proposed approach to laptop and restaurant datasets from SemEval 2014 (Pontiki et al., 2014). Experimental results show that our approach performs comparable to a top system using feature-based SVM (Kiritchenko et al., 2014). On both datasets, our approach outperforms both LSTM and attention-based LSTM models (Tang et al., 2015a) in terms of classiﬁcation accuracy and running speed. Lastly, we show that using multiple computational layers over external memory could achieve improved performance.

为了实现这一目标，我们开发了用于方面级情绪分类的深层存储网络，这是受到最近关注机制和显式记忆的计算模型的成功的启发（Graves等人，2014; Bahdanau等人，2015; Sukhbaatar等人al。，2015）。我们的方法是数据驱动的，计算效率高，不依赖于句法解析器或情感词典。该方法由具有共享参数的多个计算层组成。每个层是基于内容和位置的注意模型，其首先学习每个上下文词的重要性/权重，然后利用该信息来计算连续文本表示。最后一层中的文本表示被认为是情感分类的特征。由于每个组件都是可区分的，整个模型可以有效地用端到端的梯度下降训练，其中损失函数是情感分类的交叉熵误差。

我们将所建议的方法应用于SemEval 2014的笔记本电脑和餐厅数据集（Pontiki等，2014）。实验结果表明，我们的方法执行相当于使用基于特征的SVM的顶端系统（Kiritchenko等人，2014）。在这两个数据集上，我们的方法在分类精度和运行速度方面优于LSTM和基于注意的LSTM模型（Tang et al。，2015a）。最后，我们表明使用多个计算层在外部存储器可以实现提高性能。

No.2

A Joint Segmentation and Classification Framework for Sentiment Analysis

情感分析的联合分割和分类框架

Abstract

In this paper, we propose a joint segmentation and classiﬁcation framework for sentiment analysis. Existing sentiment classiﬁcation algorithms typically split a sentence as a word sequence, which does not effectively handle the inconsistent sentiment polarity between a phrase and the words it contains, such as “not bad” and “a great deal of”. We address this issue by developing a joint segmentation and classiﬁcation framework (JSC), which simultaneously conducts sentence segmentation and sentence-level sentiment classiﬁcation. Speciﬁcally, we use a log-linear model to score each segmentation candidate, and exploit the phrasal information of top-ranked segmentations as features to build the sentiment classiﬁer. A marginal log-likelihood objective function is devised for the segmentation model, which is optimized for enhancing the sentiment classiﬁcation performance. The joint model is trained only based on the annotated sentiment polarity of sentences, without any segmentation annotations. Experiments on a benchmark Twitter sentiment classiﬁcation dataset in SemEval 2013 show that, our joint model performs comparably with the state-of-the-art methods.

在本文中，我们提出一个联合分割和分类框架的情感分析。现有的情绪分类算法通常将句子分割为单词序列，其不能有效地处理短语与其包含的单词之间不一致的情绪极性，例如“不坏”和“大量”。我们通过开发一个联合分割和分类框架（JSC）来解决这个问题，它同时进行句子分割和句子级情感分类。具体来说，我们使用对数线性模型对每个分割候选进行评分，并利用排名最高的分割的短语信息作为特征构建情感分类器。为分割模型设计了边际对数似然目标函数，其被优化用于增强情感分类性能。仅基于句子的注释情感极性来训练联合模型，而没有任何分割注释。在SemEval 2013中的基准Twitter情绪分类数据集上的实验表明，我们的联合模型与最先进的方法具有可比性。

1 Introduction

Sentiment classiﬁcation, which classiﬁes the sentiment polarity of a sentence (or document) as positive or negative, is a major research direction in the ﬁeld of sentiment analysis (Pang and Lee, 2008; Liu, 2012; Feldman, 2013). Majority of existing approaches follow Pang et al. (2002) and treat sentiment classification as a special case of text categorization task. Under this perspective, previous studies typically use pipelined methods with two steps. They ﬁrst produce sentence segmentations with separate text analyzers (Choi and Cardie, 2008;Nakagawaetal.,2010;Socheretal.,2013b) or bag-of-words (Paltoglou and Thelwall, 2010; Maasetal.,2011). Then, feature learning and sentiment classiﬁcation algorithms take the segmentation results as inputs to build the sentiment classiﬁer (Socheretal.,2011;Kalchbrenneretal.,2014; Dong et al., 2014).

1简介

情感分类将句子（或文档）的情感极性分为正或负，是情感分析领域的主要研究方向（Pang和Lee，2008; Liu，2012; Feldman，2013）。 大多数现有的方法遵循Pang et al。 （2002），将情绪分类作为文本分类任务的特例。 在这个观点下，以前的研究通常使用具有两个步骤的流水线方法。 他们首先使用单独的文本分析器产生句子分割（Choi and Cardie，2008; Nakagawaetal。，2010; Socheretal。，2013b）或词袋（Paltoglou和Thelwall，2010; Maasetal。 然后，特征学习和情绪分类算法将分割结果作为输入来构建情感分类器（Socheretal。，2011; Kalchbrenneretal。，2014; Dong et al。，2014）。

The major disadvantage of a pipelined method is the problem of error propagation, since sentence segmentation errors cannot be corrected by the sentiment classiﬁcation model. A typical kind of error is caused by the polarity inconsistency between a phrase and the words it contains, such as (not bad, bad) and (a great deal of, great). The segmentations based on bag-of-words or syntactic chunkers are not effective enough to handle the polarity inconsistency phenomenons. The reason lies in that bag-of-words segmentations regard each word as a separate unit, which losses the word order and does not capture the phrasal information. The segmentations based on syntactic chunkers typically aim to identify noun groups, verb groups or named entities from a sentence. However, many sentiment indicators are phrases constituted of adjectives, negations, adverbs or idioms (Liu,2012; Mohammadetal.,2013a), which are splitted by syntactic chunkers. Besides, a better approach would be to utilize the sentiment information to improve the segmentor. Accordingly, the sentiment-speciﬁc segmentor will enhance the performance of sentiment classiﬁcation in turn. In this paper, we propose a joint segmentation and classiﬁcation framework (JSC) for sentiment analysis, which simultaneous conducts sentence segmentation and sentence-level sentiment classiﬁcation. The framework is illustrated in Figure 1.

流水线方法的主要缺点是错误传播的问题，因为句子分割错误不能由情绪分类模型校正。一个典型的错误是由一个短语和它包含的词之间的极性不一致引起的，例如（不坏，坏）和（大量，大）。基于词袋或句法块的分段不足以有效地处理极性不一致现象。原因在于词袋分割将每个词视为单独的单元，这丢失了词序并且不捕获短语信息。基于句法块的分段通常旨在从句子中识别名词组，动词组或命名实体。然而，许多情绪指标是由形容词，否定，副词或成语组成的短语（Liu，2012; Mohammadetal。，2013a），它们由句法块分割。此外，更好的方法是利用情绪信息来改进分段。因此，情感特定分割器将提高情绪分类的性能。在本文中，我们提出一个联合分割和分类框架（JSC）情感分析，同时进行句子分割和句子情绪分类。框架如图1所示。

We develop (1) a candidate generation model to generate the segmentation candidates of a sentence, (2) a segmentation ranking model to score each segmentation candidate of a given sentence, and (3) a classiﬁcation model to predict the sentiment polarity of each segmentation. The phrasal information of top-ranked candidates from the segmentation model are utilized as features to build the sentiment classiﬁer. In turn, the predicted sentiment polarity of segmentation candidates from classiﬁcation model are leveraged to update the segmentor. We score each segmentation candidate with a log-linear model, and optimize the segmentor with a marginal log-likelihood objective. We train the joint model from sentences annotated only with sentiment polarity, without any segmentation annotations.

We evaluate the effectiveness of our joint model on a benchmark Twitter sentiment classiﬁcation dataset in SemEval 2013. Results show that the joint model performs comparably with state-of-the-art methods, and consistently outperforms pipeline methods in various experiment settings. The main contributions of the work presented in this paper are as follows.

• To our knowledge, this is the ﬁrst work that automatically produces sentence segmentationforsentimentclassiﬁcationwithinajoint framework.

• We show that the joint model yields comparable performance with the state-of-the-art methods on the benchmark Twitter sentiment classiﬁcation datasets in SemEval 2013

我们开发了（1）候选生成模型以生成句子的分割候选，（2）分割排序模型以给出给定句子的每个分割候选，以及（3）分类模型以预测每个分割的情感极性。来自分割模型的排名最高的候选的短语信息被用作构建情感分类器的特征。反过来，来自分类模型的分割候选的预测情感极性被用于更新分割。我们用对数线性模型对每个分割候选进行评分，并使用边际对数似然目标优化分割。我们训练联合模型从仅用情感极性注释的句子，没有任何分割注释。

我们在SemEval 2013中评估我们的联合模型在基准Twitter情绪分类数据集上的有效性。结果表明，联合模型与最先进的方法具有可比性，并且在各种实验设置中始终优于管道方法。本文提出的工作的主要贡献如下。

•据我们所知，这是第一个工作，自动生成句子分割与相关框架的分类。

•我们证明联合模型在2013年SemEval基准Twitter情绪分类数据集中产生可比性能与最先进的方法

No.3

Sentence-level Emotion Classification with Label and Context Dependence

具有标签和上下文依赖性的句子级情感分类

Abstract

Predicting emotion categories, such as anger, joy, and anxiety, expressed by a sentence is challenging due to its inherent multi-label classification difficulty and data sparseness. In this paper, we address above two challenges by incorporating the label dependence among the emotion labels and the context dependence among the contextual instances into a factor graph model. Specifically, we recast sentence-level emotion classification as a factor graph inferring problem in which the label and context dependence are modeled as various factor functions. Empirical evaluation demonstrates the great potential and effectiveness of our proposed approach to sentence level emotion classification.

预测由句子表达的情绪类别（例如愤怒，快乐和焦虑）由于其固有的多标签分类难度和数据稀疏而具有挑战性。 在本文中，我们通过将情绪标签之间的标签依赖性和情境实例之间的上下文依赖性结合到因子图模型中来解决上述两个挑战。 具体来说，我们将句子级情感分类重写为因素图推断问题，其中标签和上下文相关性被建模为各种因子函数。 实证评价表明我们提出的方法对句子情感分类的巨大潜力和有效性。

1 Introduction

Predicting emotion categories, such as anger, joy, and anxiety, expressed by a piece of text encompasses a variety of applications, such as online chatting (Galik et al., 2012), news classification (Liu et al., 2013) and stock marketing (Bollen et al., 2011). Over the past decade, there has been a substantial body of research on emotion classification, where a considerable amount of work has focused on document-level emotion classification. Recently, the research community has become increasingly aware of the need on sentence-level emotion classification due to its wide potential applications, e.g. the massively growing importance of analyzing short text in social media (Kiritchenko et al., 2014; Wen and Wan, 2014). In general, sentence-level emotion classification exhibits two challenges.

预测情绪类别，如愤怒，喜悦和焦虑，由一段文字表示包括各种应用程序，如在线聊天（Galik等人，2012），新闻分类（Liu等人，2013）和股票 营销（Bollen等，2011）。 在过去十年中，已经有大量的关于情绪分类的研究，其中大量的工作集中在文档级情感分类。 最近，由于其广泛的潜在应用，研究界已经越来越意识到对句子级情感分类的需要。 在社交媒体中分析短文本的重要性越来越大（Kiritchenko等，2014; Wen和Wan，2014）。 一般来说，句子级情感分类表现出两个挑战。

On one hand, like document-level emotion classification, sentence-level emotion classification is naturally a multi-label classification problem. That is, each sentence might involve more than one emotion category. For example, as shown in Figure 1, in one paragraph, two sentences, i.e., S1 and S3, have two and three emotion categories respectively. Automatically classifying instances with multiple possible categories is sometimes much more difficult than classifying instances with a single label. On the other hand, unlike document-level emotion classification, sentence-level emotion classification is prone to the data sparseness problem because a sentence normally contains much less content. Given the short text of a sentence, it is often difficult to predict its emotion due to the limited information therein. For example, in S2, only one phrase “ 如愿以偿 (that is all I want)” expresses the joy emotion. Once this phrase fails to appear in the training data, it will be hard for the classifier to give a correct prediction according to the limited content in this sentence.

一方面，像文档级情感分类，句子级情感分类自然是一个多标签分类问题。也就是说，每个句子可能涉及多于一个情感类别。例如，如图1所示，在一个段落中，两个句子，即S1和S3，分别具有两个和三个情感类别。使用多个可能的类别自动对实例进行分类有时比使用单个标签对实例进行分类要困难得多。另一方面，与文档级情感分类不同，句级情感分类易于出现数据稀疏问题，因为一个句子通常包含少得多的内容。给定句子的短文本，由于其中的有限信息，常常难以预测其情绪。例如，在S2中，只有一个短语“如愿以偿”（即我所想要的）表达喜悦情绪。一旦该短语未出现在训练数据中，则分类器将难以根据该句中的有限内容给出正确的预测。

In this paper, we address above two challenges in sentence-level emotion classification by modeling both the label and context dependence. Here, the label dependence indicates that multiple emotion labels of an instance are highly correlated to each other. For instance, the two positive emotions, joy and love, are more likely to appear at the same time than the two counterpart emotions, joy and hate. The context dependence indicates that two neighboring sentences or two sentences in the same paragraph (or document) might share the same emotion categories. For instance, in Figure 1, S1, S2, and S3, from the same paragraph, all share the emotion category joy. Specifically, we propose a factor graph, namely Dependence Factor Graph (DFG), to model the label and context dependence in sentence-level emotion classification. In our DFG approach, both the label and context dependence are modeled as various factor functions and the learning task aims to maximize the joint probability of all these factor functions. Empirical evaluation demonstrates the effectiveness of our DFG approach to capturing the inherent label and context dependence. To the best of our knowledge, this work is the first attempt to incorporate both the label and context dependence of sentence-level emotion classification into a unified framework. The remainder of this paper is organized as follows. Section 2 overviews related work on emotion analysis. Section 3 presents our observations on label and context dependence in the corpus. Section 4 proposes our DFG approach to sentence-level emotion classification. Section 5 evaluates the proposed approach. Finally, Section 6 gives the conclusion and future work

本文中，我们通过建模标签和上下文相关性来解决句子情感分类中的上述两个挑战。这里，标签相关性指示实例的多个情感标签彼此高度相关。例如，两个积极的情绪，喜悦和爱，更可能出现在同一时间比两个对应的情绪，喜悦和恨。上下文相关性指示相同段落（或文档）中的两个相邻句子或两个句子可以共享相同的情感类别。例如，在图1中，来自同一段落的S1，S2和S3都共享情感类别喜悦。具体来说，我们提出了一个因子图，即依赖因子图（DFG），用于模拟句子情感分类中的标签和上下文相关性。在我们的DFG方法中，标签和上下文相关性被建模为各种因子函数，并且学习任务旨在最大化所有这些因子函数的联合概率。经验评估证明了我们的DFG方法捕获固有标记和上下文依赖性的有效性。据我们所知，这项工作是第一次尝试将句子级情感分类的标签和上下文依赖结合到一个统一的框架中。本文的其余部分安排如下。第2节概述了情感分析的相关工作。第3节介绍我们对语料库中标签和上下文依赖性的观察。第4节提出了我们的DFG方法对句子情感分类。第5节评估了拟议的方法。最后，第6节给出了结论和未来的工作

2 Related Work

Over the last decade, there has been an explosion of work exploring various aspects of emotion analysis, such as emotion resource creation (Wiebe et al., 2005; Quan and Ren, 2009; Xu et al., 2010), writer’s emotion vs. reader’s emotion analysis (Lin et al., 2008; Liu et al., 2013), emotion cause event analysis (Chen et al., 2010), document-level emotion classification (Alm et al., 2005; Li et al., 2014) and sentence-level or short text-level emotion classification (Tokushisa et al., 2008; Bhowmick et al., 2009; Xu et al., 2012). This work focuses on sentence-level emotion classification. Among the studies on sentence-level emotion classification, Tokushisa et al. (2008) propose a data-oriented method for inferring the emotion of an utterance sentence in a dialog system. They leverage a huge collection of emotion-provoking event instances from the Web to deal with the data sparseness problem in sentence-level emotion classification. Bhowmick et al. (2009) and Bhowmick et al. (2010) apply KNN-based classification algorithms to classify news sentences into multiple reader emotion categories. Although the multi-label classification difficulty has been noticed in their study, the label dependence is not exploited. More recently, Xu et al. (2012) proposes a coarse-to-fine strategy for sentence-level emotion classification. They deal with the data sparseness problem by incorporating the transfer probabilities from the neighboring sentences to refine the emotion categories. To some extent, this can be seen a specific kind of context information. However, they ignore the label dependence by directly applying Binary Relevance to overcome the multi-label classification difficulty. Unlike all above studies, this paper emphasizes the importance of the label dependence and exploits it in sentence-level emotion classification via a factor graph model. Moreover, besides the label dependence, our factor graph-based approach incorporates the context dependence in a unified framework to further improve the performance of sentence-level emotion classification.

2相关工作

在过去十年中，探索情感分析的各个方面的工作爆发了，例如情感资源的创造（Wiebe et al。，2005; Quan and Ren，2009; Xu et al。，2010）阅读者的情感分析（Lin et al。，2008; Liu et al。，2013），情绪事件分析（Chen et al。，2010），文档级情感分类（Alm et al。，2005; Li et al。 2014）和句子级或短文本级情感分类（Tokushisa et al。，2008; Bhowmick et al。，2009; Xu et al。，2012）。这项工作侧重于句子的情感分类。在句子级情感分类的研究中，Tokushisa et al。 （2008）提出了一种用于在对话系统中推断说话句子的情感的面向数据的方法。他们利用来自Web的大量的情感启发事件实例来处理句子级情感分类中的数据稀疏问题。 Bhowmick et al。 （2009）和Bhowmick等人（2010）应用基于KNN的分类算法将新闻句子分类为多个读者情感类别。尽管在他们的研究中已经注意到多标签分类难度，但是没有利用标签依赖性。最近，Xu et al。 （2012）提出了一种从句到情绪分类的粗到精的策略。他们通过并入来自相邻句子的转移概率来处理数据稀疏问题，以改进情感类别。在某种程度上，这可以看作是一种特定类型的上下文信息。然而，他们通过直接应用二进制相关性来克服标签依赖性，以克服多标签分类难度。与所有上述研究不同，本文强调标签依赖的重要性，并通过因子图模型在句子级情感分类中利用它。此外，除了标签依赖之外，我们的基于因子图的方法在统一的框架中并入上下文相关性，以进一步提高句子情感分类的性能。

No.4

Cross-Lingual Sentiment Classification with Bilingual Document Representation Learning

跨语言情感分类与双语文档表示学习

Abstract

Cross-lingual sentiment classiﬁcation aims to adapt the sentiment resource in a resource-rich language to a resource-poor language. In this study, we propose a representation learning approach which simultaneously learns vector representations for the texts in both the source and the target languages. Different from previous research which only gets bilingual word embedding, our Bilingual Document Representation Learning model BiDRL directly learns document representations. Both semantic and sentiment correlations are utilized to map the bilingual texts into the same embedding space. The experiments are based on the multilingual multi-domain Amazon review dataset. We use English as the source language and use Japanese, German and French as the target languages. The experimental results show that BiDRL outperforms the state-of-the-art methods for all the target languages.

跨语言情感分类旨在使资源丰富的语言中的情感资源适应资源贫乏的语言。 在本研究中，我们提出一种表示学习方法，同时学习源和目标语言中的文本的向量表示。 与以前只有双语词汇嵌入的研究不同，我们的双语文档表示学习模型BiDRL直接学习文档表示。 使用语义和情感相关性将双语文本映射到相同的嵌入空间中。 实验基于多语言多域Amazon评论数据集。 我们使用英语作为源语言，并使用日语，德语和法语作为目标语言。 实验结果表明，BiDRL优于所有目标语言的最先进的方法。

1 Introduction

Sentiment analysis for online user-generated contents has become a hot research topic during the last decades. Among all the sentiment analysis tasks, polarity classiﬁcation is the most widely studied topic. It has been proved to be invaluable in many applications, such as opinion polling (Tang etal.,2012), customer feedback tracking (Gamon, 2004), election prediction (Tumasjanetal., 2010), stock market prediction (Bollen et al., 2011) and so on.

Most of the current sentiment classiﬁcation systems are built on supervised machine learning algorithms which require manually labelled data. However, sentiment resources are usually unbalanced in different languages. Cross-lingual sentiment classiﬁcation aims to leverage the resources in a resource-rich language (such as English) to classify the sentiment polarity of texts in a resource-poor language (such as Japanese). The biggest challenge for cross-lingual sentiment classiﬁcation is the vocabulary gap between the source language and the target language. This problem is addressed with different strategies in different approaches. Wan (2009) use machine translation tools to translate the training data directly into the target language. Meng et al. (2012) and Lu et al. (2011) exploit parallel unlabeled data to bridge the language barrier. Prettenhofer and Stein (2010) use correspondence learning algorithm to learn a map between the source language and the target language. Recently, representation learning methods has been proposed to solve the cross-lingual classiﬁcation problem (XiaoandGuo,2013;Zhouetal.,2015). These methods aim to learn common feature representations for different languages. However, most of the current researches only focus on bilingual word embedding. In addition,these models only use the semantic correlations between aligned words or sentences in different languages while the sentiment correlations are ignored.

1简介

在线用户生成的内容的情绪分析已经成为过去几十年中的热研究主题。在所有情感分析任务中，极性分类是最广泛研究的主题。它被证明在许多应用中是无价的，例如观察投票（Tang etal。，2012），客户反馈跟踪（Gamon，2004），选举预测（Tumasjanetal。，2010），股票市场预测（Bollen等人， 2011）等。

大多数当前的情感分类系统建立在需要手动标记数据的监督机器学习算法上。然而，情感资源通常在不同语言中是不平衡的。跨语言情绪分类旨在利用资源丰富的语言（例如英语）中的资源来分类资源贫乏的语言（例如日语）中的文本的情绪极性。跨语言情感分类的最大挑战是源语言和目标语言之间的词汇差距。这个问题在不同的方法中用不同的策略来解决。 Wan（2009）使用机器翻译工具将训练数据直接翻译成目标语言。 Meng等人（2012）和Lu et al。 （2011）利用并行的未标记数据桥接语言障碍。 Prettenhofer和Stein（2010）使用对应学习算法来学习源语言和目标语言之间的映射。最近，提出了表示学习方法来解决跨语言分类问题（XiaoandGuo，2013; Zhouetal。，2015）。这些方法旨在学习不同语言的共同特征表示。然而，目前的大多数研究仅关注双语词嵌入。此外，这些模型仅使用不同语言中的对齐单词或句子之间的语义相关性，而忽略情感相关性。

In this study, we propose a cross-lingual representation learning model BiDRL which simultaneously learns both the word and document representations in both languages. We propose a joint learning algorithm which exploits both monolingual and bilingual constraints. The monolingual constraints help to model words and documents in each individual language while the bilingual constraints help to build a consistent embedding space across languages.

For each individual language, we extend the paragraph vector model (Le and Mikolov, 2014)to obtain word and document embeddings. The traditional paragraph vector model is fully unsupervised without using the valuable sentiment labels. We extend it into a semi-supervised manner by forcing the positive and negative documents to fall into different sides of a classiﬁcation hyperplane. Learning task-speciﬁc embedding has been proved to be effective in previous research. To address the cross-language problem, different strategies are proposed to obtain a consistentem bedding space across different languages. Both sentiment and semantic relatedness are exploited while previous studies only use the semantic connection between parallel sentences or documents.

The performance of BiDRL is evaluated on a multilingual multi-domain Amazon review dataset (Prettenhofer and Stein, 2010). By selecting English as the source language, a total of nine tasks are evaluated with different combinations of three different target languages and three different domains. The proposed method achieves the state-of-the-art performance on all the tasks.

The main contributions of this study are summarized as follows:

1) We propose a novel representation learning method BiDRL which directly learns bilingual document representations for cross-lingual sentiment classiﬁcation. Different from previous studies which only obtain word embeddings, our model can learn vector representations for both words and documents in bilingual texts.

2) Our model leverages both the semantic and sentiment correlations between bilingual documents. Not only the parallel documents but also the documents with the same sentiment are required to get similar representations.

3) Our model achieves the state-of-the-art performances on nine benchmark cross-lingual sentiment classiﬁcation tasks and it consistently outperforms the existing methods by a large margin.

在这项研究中，我们提出一个跨语言表示学习模型BiDRL，同时学习两种语言的单词和文档表示。我们提出一个联合学习算法，利用单语和双语约束。单语约束有助于模拟每种语言的单词和文档，而双语约束有助于跨语言建立一致的嵌入空间。

对于每种语言，我们扩展段落向量模型（Le和Mikolov，2014），以获得文字和文档嵌入。传统的段落向量模型是完全无监督的，而不使用有价值的情感标签。我们通过强迫正文档和负文档落入分类超平面的不同侧来将其扩展为半监督的方式。学习任务特定嵌入已被证明在以前的研究中是有效的。为了解决跨语言问题，提出了不同的策略来获得跨不同语言的一致性床铺空间。情感和语义相关性被利用，而先前的研究仅使用并行句子或文档之间的语义连接。

BiDRL的性能在多语言多域Amazon评论数据集（Prettenhofer和Stein，2010）上进行评估。通过选择英语作为源语言，使用三种不同的目标语言和三个不同域的不同组合来评估总共九个任务。所提出的方法实现了对所有任务的最先进的性能。

本研究的主要贡献总结如下：

1）我们提出一种新的表示学习方法BiDRL，其直接学习双语文献表示以用于交叉语言情感分类。与以前的研究只有获得单词嵌入，我们的模型可以学习双语文本中的单词和文档的向量表示。 2）我们的模型利用了双语文档之间的语义和情感相关性。不仅并行文档而且具有相同情感的文档都需要获得类似的表示。 3）我们的模型实现了九个基准跨语言情感分类任务的最先进的表现，并且它大大优于现有方法。

2 Related Work

Sentiment analysis is the ﬁeld of studying and analyzing people’s opinions, sentiments, evaluations, appraisals, attitudes, and emotions (Liu, 2012). Most of the previous sentiment analysis researches focus on customer reviews and classifying the sentiment polarity is the most widely studied task (Pang et al., 2002). Cross-lingual sentiment classiﬁcation is a popular topic in the sentiment analysis community which aims to solve the sentiment classiﬁcation task from a cross-language view. It is of great importance for the area since it can exploit the existing labeled information in a source language to build a sentiment classiﬁcation system in any other target language. It saves us from manually labeling data for all the languages in the world which is expensive and time-consuming. Cross-lingual sentiment classiﬁcation has been extensively studied in the very recent years. Mihalceaetal.(2007) translate English subjectivity words and phrases into the target language to build a lexicon-based classiﬁer. Baneaetal.(2010) also use the machine translation service to obtain parallel corpus. It investigates several questions based on the parallel corpus including both the monolingual sentiment classiﬁcation and cross-lingual sentiment classiﬁcation. Wan(2009) translates both the training data (English to Chinese) and the test data (Chinese to English) to train different models in both the source and target languages. The co-training algorithm (Blum and Mitchell, 1998) is used to combine the bilingual models together and improve the performance. In addition to the translation-based methods, several studies utilize parallel corpus or existing resources to bridge the language barrier. Balamurali(2012) use WordNet senses as features for supervised sentiment classiﬁcation. They use the linked WordNets of two languages to bridge the language gap. Lu et al. (2011) consider the multilingual scenario where small amount of labeled data is available in the target language. They attempted to jointly classify the sentiment for both source language and target language. Meng et al. (2012) propose a generative cross-lingual mixture model to leverage unlabeled bilingual parallel data. Prettenhofer and Stein(2010) use the structural correspondence learning algorithm to learn a map between the source language and the target language. Xiao and Guo (2014) treat the bilingual feature learning problem as a matrix completion task.

2相关工作

情绪分析是研究和分析人们的意见，情绪，评价，评价，态度和情绪的领域（Liu，2012）。大多数以前的情感分析研究集中于客户评论和分类情感极性是最广泛研究的任务（Pang等人，2002）。跨语言情感分类是情感分析社区中的一个热门话题，旨在从跨语言视角解决情感分类任务。这对于该领域是非常重要的，因为它可以利用源语言中现有的标记信息来构建任何其他目标语言的情感分类系统。它使我们免于手动标记世界上所有语言的数据，这是昂贵和耗时的。近年来，跨语言情绪分类已被广泛研究。 Mihalceaetal。（2007）将英语主观性词和短语翻译成目标语言，以构建基于词典的分类器。 Baneaetal。（2010）也使用机器翻译服务获得平行语料库。它研究了基于平行语料库的几个问题，包括单语言情感分类和跨语言情感分类。 Wan（2009）将训练数据（英语到中文）和测试数据（中英对照）翻译成源语言和目标语言的不同模型。共训练算法（Blum和Mitchell，1998）用于将双语模型组合在一起并提高性能。除了基于翻译的方法，几个研究利用平行语料库或现有资源来弥合语言障碍。 Balamurali（2012）使用WorthNetsenses特征进行监督情绪分类。他们使用两种语言的链接WordNets来弥合语言差距。 Lu et al。 （2011）考虑了多语言场景，其中少量的标签数据可用的目标语言。他们试图将两种源语言和目标语言的内容分类。 Meng等人（2012）提出了一个生成性的跨语言混合模型，以利用未标记的双语并行数据。 PrettenhoferandStein（2010）使用结构对应学习算法来学习源语言和目标语言之间的映射。 Xiao和Guo（2014）将双语特征学习问题作为矩阵完成任务。

This work is also related to bilingual representation learning. Zou et al. (2013) propose to use word alignment as the constraints in bilingual word embedding. Each word in one language should be similar to the aligned words in another language. Gouws et al. (2015) propose a similar algorithm but only use sentence-level alignment. It tries to minimize a sampled L2-loss between the bag-of-words sentence vectors of the parallel corpus. Xiao and Guo (2013) learn different representations for words in different languages. Part of the word vector is shared among different languages and the rest is language-dependent. Klementiev et al. (2012) treat the task as a multi-task learning problem where each task corresponds to a single word, and task relatedness is derived from co-occurrence statistics in bilingual parallel data. Hermann and Blunsom (2015) propose the bilingual CVM model which directly minimizes the representation of a pair of parallel documents. The document representation is calculated with a composition function based on words. Chandar A P et al. (2014) and Zhou et al. (2015) use the autoencoder to model the connections between bilingual sentences. It aims to minimize the reconstruction error between the bag-of-words representations of two parallel sentences. Luong et al. (2015) propose the bilingual skip-gram model which leverages the word alignment between parallel sentences. Pham et al. (2015) extend the paragraph vector model to force bilingual sentences to share the same sentence vector.

This study differs with the existing works in the following three aspects, 1) we exploit both the semantic and sentiment correlations of the bilingual texts. Existing bilingual embedding algorithms only use the semantic connection between parallel sentences or documents. 2) Our algorithm learns both the word and document representations. Most of the previous studies simply compute the average of the word vectors in a document. 3) Sentiment labels are used in our embedding algorithm by introducing a classiﬁcation hyperplane. It not only helps to achieve better embedding performance in each individual language but also helps to bridge the language barrier.

这项工作也与双语表示学习有关。 Zou et al。 （2013）提出使用词对齐作为双语词嵌入的约束。一种语言中的每个单词应与另一种语言中的对齐单词相似。 Gouws et al。 （2015）提出了一个类似的算法，但只使用句子级对齐。它试图最小化平行语料库的词袋语句向量之间的采样L2损失。 Xiao和Guo（2013）学习不同语言的单词的不同表示。词向量的一部分在不同语言之间共享，其余部分是语言相关的。 Klementiev et al。 （2012）将任务视为多任务学习问题，其中每个任务对应于单个词，并且任务相关性源自双语并行数据中的同现统计。 Hermann和Blunsom（2015）提出了双语CVM模型，其直接最小化一对平行文档的表示。使用基于单词的合成函数来计算文档表示。 Chandar A P et al。 （2014）和Zhou et al。 （2015）使用自动编码器来建模双语句子之间的连接。其目的是最小化两个并行句子的词袋表示之间的重建误差。 Luong等人（2015）提出了利用并行句子之间的词对齐的双语skip-gram模型。 Pham等人（2015）扩展段落向量模型，强制双语句子共享同一句子向量。

这项研究与现有作品在以下三个方面不同：1）我们利用双语文本的语义和情感相关性。现有的双语嵌入算法仅使用并行句子或文档之间的语义连接。 2）我们的算法学习词和文档表示。大多数以前的研究只是计算文档中单词向量的平均值。 3）通过引入分类超平面，在我们的嵌入算法中使用情感标签。它不仅有助于在每种语言中实现更好的嵌入性能，而且有助于弥合语言障碍。

No.5

Modeling Social Norms Evolution for Personalized Sentiment Classification

建模社会规范个性化情感分类的进化

Abstract

Motivated by the ﬁndings in social science that people’s opinions are diverse and variable while together they are shaped by evolving social norms, we perform personalized sentiment classiﬁcation via shared model adaptation over time. In our proposed solution, a global sentiment model is constantly updated to capture the homogeneity in which users express opinions, while personalized models are simultaneously adapted from the global model to recognize the heterogeneity of opinions from individuals. Global model sharing alleviates data sparsity issue, and individualized model adaptation enables efﬁcient online model learning. Extensive experimentations are performed on two large review collections from Amazon and Yelp, and encouraging performance gain is achieved against several state-of-the-art transfer learning and multi-task learning based sentiment classiﬁcation solutions

受社会科学中的发现，人们的意见是多样化的和可变的，同时他们由不断变化的社会规范塑造，我们通过共享的模型适应随着时间执行个性化情感分类。 在我们提出的解决方案中，全球情绪模型不断更新，以捕捉用户表达意见的同质性，而个性化模型同时适应全球模型，以识别个人意见的异质性。 全球模型共享减轻了数据稀疏问题，个性化模型适应使有效的在线模型学习。 对来自亚马逊和Yelp的两个大型评论集合进行广泛的实验，并且基于几个最先进的转移学习和基于多任务学习的情感分类解决方案实现了令人鼓舞的性能增益