Towards a Sentiment Analysis Model Based on Semantic Relation Analysis

Thien Khai Tran, Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology - VNU-HCM, Hồ Chí Minh, Vietnam & Faculty of Information Technology, Ho Chi Minh City University of Foreign Languages and Information Technology, Hồ Chí Minh, Vietnam

Tuoi Thi Phan, Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology - VNU-HCM, Hồ Chí Minh, Vietnam

ABSTRACT

Sentiment analysis is an important new field of research that has attracted the attention not only of researchers, but also businesses and organizations. In this article, the authors propose an effective model for aspect-based sentiment analysis for Vietnamese. First, sentiment dictionaries and syntactic dependency rules were combined to extract reliable word pairs (sentiment - aspect). They then relied on ontology to group these aspects and determine the sentiment polarity of each. They introduce two novel approaches in this work: 1) in order to “smooth” the sentiment scaling (rather than using discrete categories of 1, 0, and -1) for fined-grained classification, then extract multi-word sentiment phrases instead of sentiment words, and 2) the focus is not only on adjectives but also nouns and verbs. Initial evaluations of the system using real reviews show promising results.

KEYWORDS

Dependency Rules, Opinion Mining, Semantic Relations, Sentiment Analysis Model, Sentiment Analysis

1. INTRODUCTION

The task of sentiment analysis was first formulated in the early 2000s, particularly in the work of Dave, Lawrence, and Pennock (2003) as well as that of Nasukawa and Yi (2003). In the past fifteen years, a great deal of research has sought to analyse and evaluate opinions on products and services in media sources (Singh, Sharma, Dey, 2015). As elaborated in Sentiment Analysis and Subjectivity. Handbook of Natural Language Processing, Second Edition (2010), there are three levels of analysis worth considering: (i) document-level, (ii) sentence-level, and (iii) aspect-level. In reference to document level analysis, the research of Sharma, Nigam, and Jain (2014), Sharma, Hoque, and Chandra (2016), Tang, Qin, Liu, and Yang (2015), and Xia, Xu, Yu, and Qi, (2016) need to be taken into consideration. Significant works which discuss sentence-based sentiment analysis include Marcheggiani, Tackstrom, Esuli, and Sebastiani (2014) and Yang and Cardie, (2014). Our core research interest is in aspect-based sentiment analysis, the focus of several important studies, including those of Chinska and Joseph (2015) and Wang, Jiang, Lan, and Wu, (2017). The challenge of sentiment analysis is that the opinions which users provide on various topics are often complex and multifaceted. The sentiment valence of a word can change depending on its specific context, a phenomenon which is referred to as contextual valence shifting (Polanyi and Zaenen, 2006). Contextual valence shifting has made traditional methods, such as machine learning with bag of words or n-gram models, ineffective, as the latter tend to tackle single words, assigning them a positive or negative polarity based on pre-defined sentiment lexicons. In contrast, the most widely used techniques in sentiment analysis tend to account

DOI: 10.4018/IJSE.2018070104

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
for contextual valence shifting by paying considerable attention to semantic elements, namely words or phrases which are factors giving rise to polarity shifting (Carrillo de Albornoz, Plaza, & Gervas, 2010; Jia, Yu, & Meng, 2009; Tran & Phan, 2018).

In the present paper we propose an aspect-based sentiment analysis model for Vietnamese reviews, which combines an opinion dictionary and syntactic dependency rules to extract reliable word pairs (sentiment - aspect). Using domain-specific ontologies we grouped these aspects and determined the sentiment polarity of each.

The key innovations presented here are as follows:

- We have provided an effective model for solving the sentiment analysis problem by combining an opinion dictionary, domain-specific ontologies, and syntactic dependency rules.
- We have extracted multi-word sentiment phrases (instead of individual words) for fine-grained (multi-class) classification, including extracting nouns and verbs which bear emotional factors, rather than adjectives alone.

The remainder of this paper is organized as follows: In Section 2 related work is summarised. In Section 3 the proposed sentiment analysis model is outlined. In Section 4 the experiments used to evaluate the model are described. Finally, the results of the paper are summarized and avenues for future work are discussed.

2. RELATED WORK

In this section, we will present the various types of sentiment analysis problem: i) document-level sentiment analysis; ii) sentence-level sentiment analysis; iii) aspect-based sentiment analysis. Document-level sentiment analysis is the simplest form of opinion mining. This technique identifies users’ reviews as positive, negative or neutral using the assumption that each review is written regarding one object by only one holder. Opinion classification uses each opinion as a document, reducing the opinion classification process to document classification. Opinions are divided into three categories: positive, negative or neutral. Using this model allows a wide variety of document classification algorithms to be effectively applied to opinion classification problems. Machine learning represents one class of widely used document classification algorithms. Pang et al., (2002) applied naïve Bayes, maximum entropy and support vector machines to classify film reviews as either positive or negative. Their results showed that a SVM using unigram gave the highest accuracy (82.9%). In the case of Vietnamese text, Tran et al., (2016) have made use of a similar method accompanied by a multilabel classification approach.

The more fine-grained sentiment-level analysis is based on the understanding that a document can contain many different opinions, even regarding the same object/entity. To fully analyse the opinions in a review document, researchers need to perform sentiment analysis at the sentence level. This involves two main tasks: 1) classification of the sentence as a subjective sentence or an objective sentence, and 2) identification of the sentiment polarity of the subjective sentence. Yu and Hatzipassilisoglou, (2003) adopted the naïve Bayes and multiple naïve Bayesian methods to tackle the first task. For the second task, Matsumoto et al., (2005) extracted subtrees from the dependency parsing trees for classification of the features. The authors showed that subtrees yielded richer semantic information than n-grams and word sequences.

Aspect-based sentiment analysis involves two core tasks: extracting sentiment aspects and identifying their sentiments. In Liu, Hu, and Cheng, (2005) the authors proposed a method of extracting infrequent aspects based on the observation that “sentiments tend to appear together with aspects.” They identified infrequent aspects using the following procedure: if a sentence contains no frequent aspects but features sentiment words, then the nearest noun or noun phrase to the sentiment...
20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/article/towards-a-sentiment-analysis-model-based-on-semantic-relation-analysis/214876?camid=4v1


www.igi-global.com/e-resources/library-recommendation/?id=166

Related Content

Design and Control of a Hand-Assist Robot with Multiple Degrees of Freedom for Rehabilitation Therapy

www.igi-global.com/chapter/design-and-control-of-a-hand-assist-robot-with-multiple-degrees-of-freedom-for-rehabilitation-therapy/244040?camid=4v1a

Robust Integral of NN and Error Sign Control for Nanomanipulation Using AFM

www.igi-global.com/article/robust-integral-error-sign-control/68865?camid=4v1a
On the Forces Between Micro and Nano Objects and a Gripper
www.igi-global.com/article/forces-between-micro-nano-objects/68861?camid=4v1a

Low-Power High-Performance Tunnel FET With Analysis for IoT Applications
www.igi-global.com/chapter/low-power-high-performance-tunnel-fet-with-analysis-for-iot-applications/237279?camid=4v1a