

Feature-Based Sentiment Analysis in Online Review with Semi-Supervised Support Vector Machines (S³VMs)

Jessie Setiady, Warih Maharani, Rita Rismala

Department of Informatics Engineering, Faculty of Informatics, Telkom University

Keywords:

Review
Sentiment
Product feature
S³VMs

ABSTRACT

Online reviews provide facility so that internet user can give review about an aspect. Sentiments about a product are useful and have an influence in decision-making by person or organization. As in an opinion, reviewers and provide positive and negative reviews simultaneously. This is due, opinions targets are often not the product as a whole, but rather part of a product called the feature, where there are advantages and disadvantages in the eyes of reviewers.

In this research, sentiment will be identified based on its opinion. Opinion data used in this research is in English, taken from the site www.cnet.com. The product conclusions presented based on product features. Thus, there are two processes undertaken in this research: (1) Extraction of product features in opinion, (2) Sentiment identification for each product feature. Feature extraction is done by searching for phrases that match the relation dependencies template, and then do the filtering feature. In sentiment identification, the positive and negative probability value, and also the target class of the feature opinion, became S³VMs input parameters. In the study by S³VMs, some data are treated as unlabeled data. Results obtained from this study for the evaluation of sentiment identification with F1-Measure at 86% for positive class and 70% for negative class. As for feature identification obtained 82% accuracy. For further development of this research, Improve SVM is suggested to handle the unbalance data problem. Mapping to implicit feature is also advisable to identify more product feature.

*Copyright © 2013 Information Systems International Conference.
All rights reserved.*

Corresponding Author:

Jessie Setiady,
Department of Informatics Engineering, Faculty of Informatics,
Telkom University,
Jalan Telekomunikasi No.1 Ters. Buah Batu, Bandung, Indonesia.
Email: setiady.jessie@live.com

1. INTRODUCTION

Currently the Internet is not only used as a media to access information, but also as a media to share information. Information is categorized into two [5]: knowledge (facts), or opinion. Both types of information can be easily shared by Internet users, or which is known as User Generated Content, through a variety of facilities, such as: blogs, product review sites, social networking, forums, Question and Answer sites, voting sites, etc.

In the survey conducted by comScore (2007), and Horrigan (2008), found that 81% of internet users in the U.S. use the internet to search for products to buy, and more than 30% of internet users provide a review of a product purchased [1]. An online review is one medium that provides facilities so that a reviewer can give reviews or opinion, in the form of thoughts, suggestions or just comments. Reader, with their intuitive abilities, is able to know the sentiments of the reviewer of a topic of discussion, by browse the websites of online reviews available. Overview sentiment on reviewer opinion given can be used as one of the parameters of the analysis, such as the experience of others who have purchased a product determines a person's decision to purchase a particular product. The problem is, so many opinions are available, so that the reader will be overwhelmed if they have to read and analyze one-by-one reviewer's opinion. Another problem is, reviewers often provide an opinion on the positive and negative aspects of the product. Thus, an