**SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining**

Abstract

In this work we describe SENTIWORDNET, a lexical resource in which eachWORDNET synset s is associated to three numerical scores Obj(s), Pos(s) and Neg(s), describing how objective, positive, and negative the terms contained in the synset are. The method used to develop SENTIWORDNET is based on the quantitative analysis of the glosses associated to synsets, and on the use of the resulting vectorial term representations for semi-supervised synset classification. The three scores are derived by combining the results produced by a committee of eight ternary classifiers, all characterized by similar accuracy levels but different classification behaviour. SENTIWORDNET is freely available for research purposes, and is endowed with a Web-based graphical user interface.

1. Introduction

Within OM, several subtasks can be identified, all of them having to do with tagging a given text according to expressed opinion:

1. determining text SO-polarity

2. determining text PN-polarity

3. determining the strength of text PN-polarity

In this paper we describe SENTIWORDNET (version1.0), a lexical resource in which each synset of WORDNET(version 2.0) is associated to three numerical scores Obj(s), Pos(s) and Neg(s), describing how Objective, Positive, and Negative the terms contained in the synset are.

2. Building SENTIWORDNET

2.1. Training a classifier

2.2. Defining the committee of classifiers

2.3. Some statistics

3. Visualizing SENTIWORDNET

4. Evaluating SENTIWORDNET

5. Conclusion and future research

We believe that SentiWordNet can prove a useful tool for opinion mining applications, because of its wide coverage (all WordNet synsets are tagged according to each of the three labels Objective, Positive, Negative) and because of its fine grain, obtained by qualifying the labels by means of numerical scores.