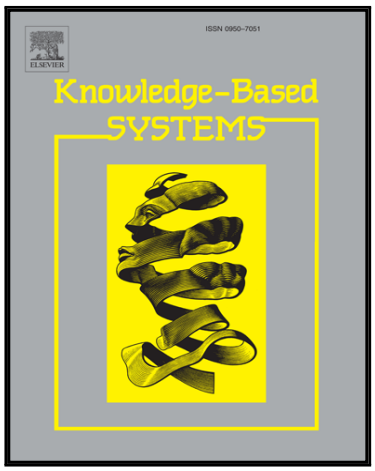

Feature Selection and Ensemble Construction: A Two-step Method for Aspect Based Sentiment Analysis

Md Shad Akhtar, Deepak Gupta, Asif Ekbal, Pushpak Bhattacharyya

PII: S0950-7051(17)30148-X
DOI: 10.1016/j.knosys.2017.03.020
Reference: KNOSYS 3869

To appear in: *Knowledge-Based Systems*

Received date: 21 September 2016
Revised date: 15 February 2017
Accepted date: 25 March 2017

Please cite this article as: Md Shad Akhtar, Deepak Gupta, Asif Ekbal, Pushpak Bhattacharyya, Feature Selection and Ensemble Construction: A Two-step Method for Aspect Based Sentiment Analysis, *Knowledge-Based Systems* (2017), doi: 10.1016/j.knosys.2017.03.020

Feature Selection and Ensemble Construction: A Two-step Method for Aspect Based Sentiment Analysis

Md Shad Akhtar, Deepak Gupta, Asif Ekbal and Pushpak Bhattacharyya

Department of Computer Science and Engineering
Indian Institute of Technology Patna (IIT Patna)
Patna,India

Abstract

In this paper we present a cascaded framework of feature selection and classifier ensemble using particle swarm optimization (PSO) for aspect based sentiment analysis. Aspect based sentiment analysis is performed in two steps, *viz.* aspect term extraction and sentiment classification. The pruned, compact set of features performs better compared to the baseline model that makes use of the complete set of features for aspect term extraction and sentiment classification. We further construct an ensemble based on PSO, and put it in cascade after the feature selection module. We use the features that are identified based on the properties of different classifiers and domains. As base learning algorithms we use three classifiers, namely Maximum Entropy (ME), Conditional Random Field (CRF) and Sipport Vector Machine (SVM). Experiments for aspect term extraction and sentiment analysis on two different kinds of domains show the effectiveness of our proposed approach.

Keywords: Sentiment Analysis, Aspect Term Extraction, Feature Selection, Ensemble, Conditional Random Field, Support Vector Machine, Maximum Entropy, Particle Swarm Optimization.

1. Introduction

Recent past has witnessed a phenomenal growth of internet users globally, and the third world countries like India, China etc. are not the exceptions. Use of social media and messaging applications grew 203 percent year-on-year in 2013, with overall application users rising 115 percent over the same period, as reported by Statista, citing data from Flurry Analytics. This growth means that 1.61 billion people are now active in social media around the world and this is expected to advance to 2 billion users in 2016, led by India. The research shows that consumers are now spending daily approximately 8 hours on digital media including social media and mobile internet usages. This has completely changed the

*Asif Ekbal

Email address: asif.ekbal@gmail.com,asif@iitp.ac.in (Md Shad Akhtar, Deepak Gupta, Asif Ekbal and Pushpak Bhattacharyya)