

Application of Semi-supervised Learning to Evaluative Expression Classification

Yasuhiro Suzuki¹, Hiroya Takamura², and Manabu Okumura²

¹ Interdisciplinary Graduate School of Science and Engineering,
Tokyo Institute of Technology,
4259 Nagatsuta Midori-ku Yokohama, JAPAN, 226-8503
`yasu@lr.pi.titech.ac.jp`

² Precision and Intelligence Laboratory, Tokyo Institute of Technology,
4259 Nagatsuta Midori-ku Yokohama, JAPAN, 226-8503
`{takamura,oku}@pi.titech.ac.jp`

Abstract. We propose to use semi-supervised learning methods to classify evaluative expressions, that is, tuples of subjects, their attributes, and evaluative words, that indicate either favorable or unfavorable opinions towards a specific subject. Due to its characteristics, the semi-supervised method that we use can classify evaluative expressions in a corpus by their polarities. This can be accomplished starting from a very small set of seed training examples and using contextual information in the sentences to which the expressions belong. Our experimental results with actual Weblog data show that this bootstrapping approach can improve the accuracy of methods for classifying favorable and unfavorable opinions.

1 Introduction

An increasing amount of work has been devoted to investigating methods of detecting favorable or unfavorable opinions towards specific subjects (e.g., companies and their products) within online documents such as Weblogs (blogs), messages in a chat room and on bulletin board (BBS) [1, 2, 7, 9, 11, 12, 18]. Areas of application for such an analysis are numerous and varied, ranging from analysis of public opinion, customer feedback, and marketing analysis to detection of unfavorable rumors for risk management. The analyses are potentially useful tools for the commercial activities of both companies and individual consumers who want to know the opinions scattered on the World Wide Web (WWW).

To analyze a huge amount of favorable or unfavorable opinions, we need to automatically detect evaluative expressions in text.

Evaluative expressions are not mere words that indicate unique (favorable or unfavorable) polarity in themselves (such as the adjectives ‘beautiful’ and ‘bad’), but rather they are tuples of the subject to be evaluated, an attribute, and an evaluative word. Tuples are necessary because the evaluative polarity of

[†] Yasuhiro Suzuki currently works at Fujitsu.