Chapter 14
Sentiment Recognition from Bangla Text

K. M. Azharul Hasan
Khulna University of Engineering and Technology (KUET), Bangladesh

Sajidul Islam
Khulna University of Engineering and Technology (KUET), Bangladesh

G. M. Mashrur-E-Elahi
Khulna University of Engineering and Technology (KUET), Bangladesh

Mohammad Navid Izhar
Khulna University of Engineering and Technology (KUET), Bangladesh

ABSTRACT

Sentiment analysis is a very important area of the natural language processing. In general, sentiment classification means the analysis to determine the expression of a speaker whether he or she holds positive or negative opinion to a specific subject. With the rapid growth of e-commerce, sentiment analysis can greatly influence everyone in their real life. For example, product reviews on the Web have become an important source of information for customers' decision making when they want to buy any product. As the reviews are often too many for customers to go through, how to automatically classify and detect the sentiment from them has become an important research problem. In this chapter, the authors present a Sentiment Analyzer that recognizes the Bangla sentiment or opinion about a subject from Bangla text. They construct some phrase patterns and calculate their sentiment orientation. They add tags to words in the Bangla text to construct the phrase pattern for positive and negative sentiment. Then the authors match the phrase pattern in Bangla text with their predefined phrase pattern and cumulate the sentiment orientation of each sentence.

DOI: 10.4018/978-1-4666-3970-6.ch014

INTRODUCTION

Sentiment analysis or opinion mining is a vast area of Natural Language Processing (NLP). Analyzing sentiment of a language is a challenging task in NLP. In general, the motive of sentiment analysis is to determine the expression of a speaker or a writer with respect to some specific topic or story. The expression may be their opinion, feelings, attitude or comments, which indicate the thinking of the writer or the emotional effect the writer wants to have on reader mind (Picard et al., 2011). The fundamental task in sentiment analysis is classifying the sentimental state of a text in a given document whether the expressed opinion in a document is positive, negative, or neutral (Shaikh et al., 2007).

Bangla (or Bengali), one of the more important Indo-Iranian languages, is the sixth-most popular in the world and spoken by a population that now exceeds 250 million. It is the primary language in Bangladesh and second language in India (Das & Bandyopadhyay, 2010). Lots of research on sentiment analysis has been done on different languages such as English (Picard et al., 2011; Sebastiani, 2002; Lu, Liu, & Zhang, 2006), Chinese (Zhang et al. 2011), Urdu (Sayed et al. 2011) etc. But in contrary, sentiment analysis is still an unsolved research problem in Bangla and such kind of research work is very rare due to lake of resource and the complexity of Bangla language.

The Internet has become a rich platform for people to express their opinion, attitude, feeling, and emotion. From this point of view, Web is an important source of product reviews, news reviews, blog reviews, movie review, stock market reviews, travel advice, social issue discussions, consumer complaints, etc. Nowadays, Bangla has been using widely in the Web. Automatic sentiment classification will become very useful in above applications. Sentiment analysis is now a great interest to the social networking media such as Twitter, Facebook, Google+ as well. Using sentiment analysis, they can track their site for some unexpected posts, comments and share. But they require analyzing many languages.

In this chapter, we present a phrase pattern-based (Fei et al. 2004) method in classifying sentiment orientation of Bangla text. That is to analyze whether the text expresses a favorable or unfavorable sentiment for a specific subject. We construct some phrase patterns and calculate their sentiment orientation by unsupervised learning algorithm. When we classify a document, we first add special tags to some words in the text, and then match the tags within a sentence with some phrase patterns to get the sentiment orientation of the sentence. At last, we add up the sentiment orientation of each sentence. We classify the text according to this summation. The research on Bangla sentiment recognition and classification is in very primary stage. Still more research efforts are needed to reach the user satisfaction level and social demand. We present a sentiment analyzer for Bangla language using machine-learning methodology for paragraph level granularity. Along with the proposal, we determine some technical challenges for Bangla language processing. This analyzing technique basically recognizes the Bangla sentiment or opinion about a subject from Bangla text. We implement the corresponding Bangla dictionary in Extensible Markup Language (XML) format using the word as value and its Parts of Speech (POS) as tag name. The XML is used because it is a very efficient technique for data storing and searching. We believe the proposed method can be applied efficiently for detecting sentiment from Bangla text.

TRENDS AND SOLUTIONS OF SENTIMENT RECOGNITION

What is the opinion of other people has always carried a valuable piece of information in any decision-making process which may help to provide better output. There are two popular