



Algorithm for Persian Text Sentiment Analysis in Correspondences on an E- Learning Social Website

Anahid Rais Rohani¹ and A'zam Bastanfard^{1*}

1 Department of Computer, College of Mechatronic, Karaj Branch, Islamic Azad University, Alborz, Iran

Original Article:

Received 22 Aug. 2015 Accepted 26 Jan. 2016 Published 30 Jan. 2016

ABSTRACT

By 2000, sentiment analysis had been only studied based on speech and changes in facial expressions. Since then, studies have been focused on text. Concerning Persian text mining, studies have been conducted on the methods for extracting properties for classification and examination of opinions on social websites with an aim to determine text polarity. The present research was aimed to prepare and implement an algorithm for Persian text sentiment analysis based on the following six basic emotional states: happiness, sadness, fear, anger, surprise, and disgust. In this research, sentiment analysis was carried out using the unsupervised lexical method. Lexicons are divided into four categories, namely the emotional, boosters, negation, and stop lists. The algorithm was written in six different ways using different properties. In the first method, the algorithm was capable of identifying an emotional word in a sentence. The sentiment of the sentence was determined based on the given emotional word. However, it should be noted that the text itself is also important for sentiment analysis because in addition to the emotional words, other factors (such as boosters and negating factors) are also present in the sentence and affect the text sentiment. Hence, the algorithm was enhanced in the subsequent methods to detect the boosters and negating words. Results of running the algorithm using different methods indicated that the algorithm accuracy increased with an increase in the number properties involved. In the sixth method, an algorithm capable of identifying emotional, boosters and negative words was applied to two data samples including sentences written by typical users and sentences written by university students on an electronic learning social website. The accuracy of the algorithm with 100 data samples from typical users and 100 data samples from university students was 80% and 84%, respectively.

Keyword:

- ✓ sentiment analysis
- ✓ opinion mining
- ✓ text mining

* Corresponding author:

*bastanfard03@gmail.com (A'zam Bastanfard),
anahid_r@yahoo.com (Anahid Rais Rohani)*