

Contents lists available at [ScienceDirect](#)

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Sentiment analysis: A review and comparative analysis of web services



Jesus Serrano-Guerrero^{a,*}, Jose A. Olivas^a, Francisco P. Romero^a, Enrique Herrera-Viedma^{b,c}

^a Department of Information Technologies and Systems, University of Castilla-La Mancha, 13071 Ciudad Real, Spain

^b Department of Computer Science and Artificial Intelligence, University of Granada, 18071 Granada, Spain

^c Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Article history:

Received 9 April 2014

Received in revised form 21 February 2015

Accepted 15 March 2015

Available online 21 March 2015

Keywords:

Sentiment analysis

Sentiment classification

Rating prediction

Web services

ABSTRACT

Sentiment Analysis (SA), also called Opinion Mining, is currently one of the most studied research fields. It aims to analyze people's sentiments, opinions, attitudes, emotions, etc., towards elements such as topics, products, individuals, organizations, and services. Different techniques and software tools are being developed to carry out Sentiment Analysis. The goal of this work is to review and compare some free access web services, analyzing their capabilities to classify and score different pieces of text with respect to the sentiments contained therein. For that purpose, three well-known collections have been used to perform several experiments whose results are shown and commented upon, leading to some interesting conclusions about the capabilities of each analyzed tool.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Sentiment Analysis, also called Opinion Mining, is one of the most recent research topics within the field of Information Processing. Textual information retrieval techniques are mainly focused on processing, searching or mining factual information. Facts have an objective component; however, there are other textual elements which express subjective characteristics. These elements are mainly opinions, sentiments, appraisals, attitudes, and emotions, which are the focus of Sentiment Analysis [44]. All of them are closely related, however, they present slight differences. This fact involves the birth of many related tasks in this new research field, such as *opinion mining*, *subjectivity analysis*, *emotion detection* or *opinion spam detection*, among others.

Sentiment Analysis offers many opportunities to develop new applications, especially due to the huge growth of available information in sources such as blogs and social networks. For example, recommendations of items proposed by any *recommender system* can be computed taking into account aspects such as positive or negative opinions about those items. Review- and opinion-aggregation websites could collect information from different sources in order to summary or compose an opinion about a candidate, product, etc., thus replacing systems which require explicitly opinions or summaries. Question answering systems represent another field where opinions play an important role. Detection of opinion-oriented questions and possible answers, and its treatment are essential to compute good answers. Detection of subjective information is really important in fields related to argumentation where objective sentences are usually more valuable. But certainly, one of the most important fields where Sentiment Analysis has a greater impact is in the industrial field. Small and big companies, as

* Corresponding author. Tel.: +34 926 295300x6332; fax: +34 926 295354.

E-mail address: jesus.serrano@uclm.es (J. Serrano-Guerrero).

well as other organizations such as governments, desire to know what people say about their marques, products or members [51,89,15,54,7,28,14,11,59].

Sentiment Analysis is a concept that encompasses many tasks such as extraction of sentiments, sentiment classification, subjectivity classification, opinion summarization or opinion spam detection, among others. To perform any of these activities, Sentiment Analysis has to deal with many challenges. The first one is the definition of the elements involved in this area. Thus, it is necessary to define clearly concepts such as opinion, subjectivity or emotion, however, this task is not really easy. For example, in a simple way a user opinion could be considered as a positive or negative sentiment about an entity or an aspect of that entity. On the other hand, subjectivity does not imply necessarily a sentiment but it allows expressing feelings or beliefs, and specifically, our own feelings or beliefs and our *emotions*.

These definitions have to be represented by mathematical expressions that can be computed and used as inputs for the aforementioned activities. Accordingly, Sentiment Analysis success mainly depends on the ability to extract the necessary features of those definitions from texts to perform those tasks. Thus, Natural Language Processing (NLP) techniques are essential to achieve good results depending on the task that has to be carried out. This is another of the main challenges of this research field, along with all problems related to the adaptation of typical techniques for classifying or summarizing texts in this field, as well as the creation of new techniques and algorithms specialized on opinions.

Despite the complexity and difficulty of this problem, many companies and universities are developing new tools and web services which deal with several of the issues aforementioned. These services could be included, especially for research purposes, into other applications without the need of being expert in Sentiment Analysis, such as other platforms do.

Following this idea and due to the growing number of new services related to Sentiment Analysis, the aim of this work is twofold. On the one hand, to present a detailed description of a set of 15 well-known free access services focused on Sentiment Analysis. These tools might have been developed by private companies or universities, but all of them allow free access to the functionalities that will be analyzed in this work. For that reason, all of them may be especially interesting for research purposes, as it is not necessary to implement services which are already working and are free.

And on the other hand, this work will assess the main functionalities from these 15 services related to Sentiment Analysis and analyze the results obtained. For that purpose, three well-known data collections in the field of Sentiment Analysis will be used. This way, this work will allow the user/researcher to have enough information about the different capabilities provided by each tool, and consequently, the user/researcher can choose the most appropriate one to be included into his own platform.

In summary, this paper presents a comprehensive and in-depth critical assessment of 15 Sentiment Analysis web tools that has never been done before. To properly perform this assessment, a suite of evaluation criteria and well-known data collections from the field of Sentiment Analysis has been selected to allow the reader to look into the pros and cons of the use of these tools regarding aspects such as discovery of sentiments within short and long texts, detection of irony or computation of polarity ratings, among others. Apart from these standard data collections, these tools have also been assessed by emulating a more real scenario, in which the effectiveness for recommending movies from real users' comments has been tested using information collected from the well-known website IMDb.¹

The remainder of the work is organized as follows: Section 2 presents the main concepts related to Sentiment Analysis discussed in several recent works. Section 3 shows the main characteristics of many Web services which allow computing sentiments. Section 4 presents several experiments that have been performed in order to compare the Web services commented on the previous section, as well as the results obtained. Finally, Section 5 points out several conclusions.

2. Background

The concepts Opinion Mining, Sentiment Analysis and Subjectivity Analysis are broadly used as synonyms; however, their origins are not exactly the same and some authors consider that each concept presents different connotations. Pang and Lee [68] present a more detailed review on the origins of these concepts and others closely related, e.g., Affective Analysis, Review Mining or Appraisal Extraction. Therefore, it is necessary to define some concepts to understand the issue dealt with in this work.

2.1. Definition of main concepts

An opinion could be simply defined as a positive or negative sentiment, view, attitude, emotion, or appraisal about an entity (product, person, event, organization or topic) or an aspect of that entity from a user or group of users.

Following that definition, an opinion can be mathematically defined as a 5-tuple $(e_j, a_{jk}, so_{ijkl}, h_i, t_i)$ where e_j represents a target entity and a_{jk} is the k -th aspect/feature of the entity e_j . so_{ijkl} is the sentiment value of the opinion from the opinion holder h_i on aspect a_{jk} of entity e_j at time t_i . That value can be positive, negative, or neutral, or even a more granular rating can be used. h_i is the opinion holder and t_i is the time when the opinion was expressed [44].

Opinions can be classified into different groups, for instance, they could be regular and comparative opinions. Most of opinions are regular, and they can be subdivided into direct or indirect opinions. Direct opinions express an idea on an entity

¹ <http://www.imdb.com/>.

Download English Version:

<https://daneshyari.com/en/article/393247>

Download Persian Version:

<https://daneshyari.com/article/393247>

[Daneshyari.com](https://daneshyari.com)