Chapter 19
Depressive Person Detection using Social Asian Elephants' (SAE) Algorithm over Twitter Posts

Hadj Ahmed Bouarara
https://orcid.org/0000-0002-4973-4385
GeCoDe Laboratory, Saida, Algeria

ABSTRACT

With the advent of the web and the explosion of data sources such as opinion sites, blogs and microblogs appeared the need to analyze millions of posts, tweets or opinions in order to find out what thinks the net surfers. The idea was to produce a new algorithm inspired by the social life of Asian elephants to detect a person in depressive situation through the analysis of twitter social network. The proposal algorithm gives better performance compared to data mining and bioinspired techniques such as naive Bayes, decision tree, heart lungs algorithm, social cockroach's algorithm.

INTRODUCTION AND PROBLEMATIC

Depression is an under-diagnosed disease, too often underestimated and yet widespread. It is as common as other major chronic conditions such as cardiovascular disease or diabetes, as evidenced by the fact that more than 20 million of persons each year live a major depressive situation.

According to the World Health Organization (WHO), approximately 11% of people suffer from major depression in their life. It affects young people aged of 15 to 24 years old. According to the WHO, by 2020, depression will become the second leading cause of disability worldwide, after cardiovascular disease. In very simple terms, it can be said that depression results in some way from an imbalance in brain chemistry. When a person is in depression, his body and mind send him warning signals to tell him that something is wrong. Listening to these signals we can provide a quicker diagnosis and put in place remedial measures (Finegold et al., 2013).

DOI: 10.4018/978-1-6684-6307-9.ch019

For example, being sad after losing a loved one or having a feeling of failure in case of problems at work is normal. But when these states of mind return each day for no particular reason or persist for a long time even with an identifiable cause, it can be a depression. Depression is actually a chronic disease, meeting specific diagnostic criteria. In addition to sadness, the depressed person maintains negative and devaluing thoughts: "I am really bad," "I will never succeed", "I hate who I am." She feels worthless and has trouble projecting herself into the future. She no longer has interest in previously popular activities. It is unclear what is the causes of depression, but it is likely a complex disease involving many factors related to heredity, biology, life events, and environment and habits. of life.

In recent years, we are in a digital world where information is available in large quantities and in various forms. 80% of this mass of information was in textual form. For this reason, we need specific tools to access sentiments and meanings hidden in these data, in order to reduce human intervention.

The sentiment analysis is an automatic process of written and spoken discourses by bringing out the different opinions expressed on a specific subject such as a brand, a news or a product. The importance of this paradigm is presented in several areas, namely policy, marketing, reputation management ...ect. It is a part of a broader area of study called NLP for Natural Language Processing with objective is to render a computer program understanding the codes of human language. The content of this chapter discusses the elaboration of new algorithms by the social life of Asian elephants to the problem of detecting depressed person by decision analysis of tweets.

The general structure of the paper will be as follows: we start with a state of the art for presenting the essential works in this topic, after we go on with a section detailing our approach and proposed components then an experimental and comparative study will be carried out for presenting the best results obtained. Finally, we will finish with a conclusion and describing some lines of thought that remain open and that we want to share them with you.

Stat of the Art

The work of Hatzivassiloglou and McKeown in 1997 consists in using the coordinating conjunctions present between a word already classified and an unclassified word, followed by the contributions of researcher Nasukawa and his team in 2003 who proposed a new method for extracting associated concepts from segments and summing the orientations of the opinion vocabulary present in the same segment (Chauan et al., 2017).

In the same year, researchers Yu and Hatzivassiloglou (2003) used the probability of ranking a word to measure the strength of the orientation of the named entities. In 2006, researchers Kanayama and Nasukawa (2006) as well as Ding and Liu in 2008 proposed, for their part, a learning-based approach that uses the coordination conjunctions present between a word already classified and a word unclassified.

The approaches of Pang et al introduced in 2002, and that of Charton and Acuna-Agost published in 2007 consist of classifying the texts according to a global polarity (positive, negative and neutral). These methods were optimized by Wilson and his research team in 2005. However, the difficulty lies in the constitution of these corpora of learning, which is a manual process to perform for each area studied. Finally, Vernier and his team (2009), have relied on a method of detection and categorization of the evaluations locally expressed in a corpus of multi-domain blogs. The second Dictionary-based Approach has had a lot of work. In 2015, Rosenthal and his team built General Inquiry which contains 3596 words labeled positive or negative. In Nakov and al work published in 2016, they use only adjectives for the detection of opinions. They manually build a list of adjectives they use to predict sentence