Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 83 (2016) 292 – 299

Procedia Computer Science

The 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016)

Using YouTube comments for text-based emotion recognition

Douiji yasmina[a]*, Mousannif Hajar[b], Al Moatassime Hassan[a]

[a]*Faculty of Science and Technology, Abdelkarim Elkhattabi Street, Guéliz, Marrakesh P.C 40549, Morocco*
[b]*Faculty of Semlalia, Prince My Abdellah Street,Marrakesh P.C 42390, Morocco*

Abstract

With the increase of Smartphone use, there is a growing need for advanced features that offer to Smartphone users a smarter interaction. We aim through the presented system to detect users' emotions from their textual exchanges, dealing with the complexity of chat writing style and the evolution of languages. We consider that such a system is a start for interesting applications that exploit users' emotional states. Our system uses an unsupervised machine learning algorithm that performs emotion classification, based on a data corpus built from YouTube comments. The reason behind such a choice is the similarity between YouTube comments and instant messages writing style. To classify a text entry into a particular emotion category, we compute its similarity to each target emotion, using the Pointwise Mutual Information measure. Our method yields a global precision of 92.75%, which reflects the feasibility of our approach.

Peer-review under responsibility of the Conference Program Chairs

Keywords: affective computing; emotion recognition; text; machine learning.

1. Introduction

YouTube comments present a rich resource for publicly available text. They hold different styles of expression, and are in almost all existing languages. They also raise different issues: opinions, stories and Emotions. The empirical research we present in this paper harnesses the huge potential in YouTube comments for text-based emotion detection. Former experiences have shown that it is quite complicated when it comes to emotion extraction

* Corresponding author. Tel.: +212648714788; fax: N/A.
E-mail address: yasmina.douiji@ced.uca.ma

Peer-review under responsibility of the Conference Program Chairs
doi:10.1016/j.procs.2016.04.128

from text as compared to emotion detection using face[1], voice[2] and gestures[3]. Among the difficulties that hinder text-based emotion detection, there are: the complexity of natural language, its continuous evolution (new expressions everyday), and the ambiguous context of the author.

In this paper, we use an unsupervised machine learning algorithm based on the previous work of[4], to which we brought modifications we later discuss in this paper. We classify emotions according to the six basic emotions of Ekman[5]. Each emotion category is represented by a list of expressive words. And to determine the emotion expressed in a piece of text, we first classify its component words. We start by extracting Adjectives, Nouns, Verbs and Adverbs (NAVA words) from text. The other words (pronouns, interjections, prepositions…) are not considered, because they are all the time neutral. We, then, compute the probabilities for each word to belong to each emotion category. The probability of a single word to belong to a particular emotion category is the value of the normalized form of the PMI (Pointwise Mutual Information), between this word and the representative words of the emotion category. In fact, unlike the previous work[4] we prefer to use the normalized form because it gives better insights about the relatedness between two events[6]. When a negation is present, the concerned word is automatically assigned to the category "Neutral" (probability equal to zero). The probabilities of the whole sentence are the average of the obtained probabilities by the number of classified words.

The data corpus that we use to compute the different PMIs is built by importing comments from YouTube using YouTube API version 3. To ensure having enough rich content in the corpus, we browse videos from different YouTube categories (divertissement, Blogs & People …) using keywords relevant to the six emotions of Ekman. Once videos identifiers are retrieved, we import the corresponding comments.

Our system shows satisfying results compared to previous works. First conducted tests give high precision ranging from 91% to 95% for different target emotions. To run tests, we choose two different types of sentences. The first type contains affective words that correspond to each of the target emotions, and the second type does not contain any affective words.

The remainder of this paper is organized as follows: the second section presents previous works in text-based emotion detection, giving their results and accuracy issues. The third section explains, in details, the different steps of our approach, including the process of building the data corpus, the algorithm used and tests organization. Results of our study, as well as, discussions are presented in the fourth section. And finally, we give outlines for possible improvements in the last section.

2. Previous work

Different methods were used to build emotion detection systems from text, which can be grouped into three categories of algorithms[7]: Knowledge-based, machine learning based, and hybrid. In this section, we adopt the same classification scheme, but we present works that are more or less related to the work we develop in this paper.

2.1. Knowledge-based detection

This category of methods consist of using affect lexicons, such as WordNet-affect[8] or General Inquirer[9], and a baseline algorithm in order to check the presence of affective words, present in the lexicon, in the text to classify. The algorithm, then, computes a score that reflects frequency in the text for each detected word. To detect the emotion expressed in a piece of text (e.g. in a sentence), words' scores are aggregated following some linguistic rules: Authors in[10] for example, perform sentiment tagging on news headlines. They consider that the root word, which does not depend on any other word in the headline, has the most important contribution on the global subject of the headline. So they extract the root word using Stanford Dependency Parser[11], and multiply its emotional score by 6. They also search patterns like (noun → subject → verb) and (verb → direct object → noun) in the dependency graph with verbs that increase or decrease a quantity, in order to increase or decrease the score of depending nouns. The presented approach constitutes a viable method for emotion detection, the only concern about such approach is that the rules are not comprehensive and need to be manually defined.