

Sentiment Analysis: Few Approaches

Navreet KaurRana

Abstract

This paper discusses various possible approaches to resolve sentiment analysis. Sentiment, being a human attribute, is complex in nature to deduce, when machine intelligence is considered. Moreover, one approach can not suffice as a solution to all the natural languages at one time. But an attempt is made to provide a solution to some extent.

1. Introduction

Sentiment Analysis is not a new idea. Ever since its inception, it has always been progressing. But given the size of language corpora and the existence of multiple languages, no single exhaustive solution seems to be possible for it.

The fact that sentiment or mood is something abstract in nature, adds to the complexity of the analysis as compared to syntactic and semantic analysis.

Given the nature and the vastness of the languages, providing a single exhaustive algorithm or concept is beyond the scope of this paper. Instead, multiple concepts have been stated to suit to different languages. A concept may work for one or multiple natural languages, depending upon the structure of the languages.

2. Approaches

Based on both human knowledge and computational techniques, following are the few approaches which analyze the sentiments in text both at micro and macro level:

- Creating a set of words, based on their roots, prefixes or suffixes. Certain roots of words have ‘negative only’ or ‘positive only’ meanings. They can definitely be termed as positive or negative in sense. Examples follow:
In English language, the prefix ‘mal’ (French origin), roughly means ‘faulty/unpleasant’. The fact can be established by considering the following examples:
Malafide, malicious, malnourish, malfunction, malpractice, malformed etc.
The same stands true for Spanish. The word ‘mal’ means ‘bad’. Consider the following phrases:

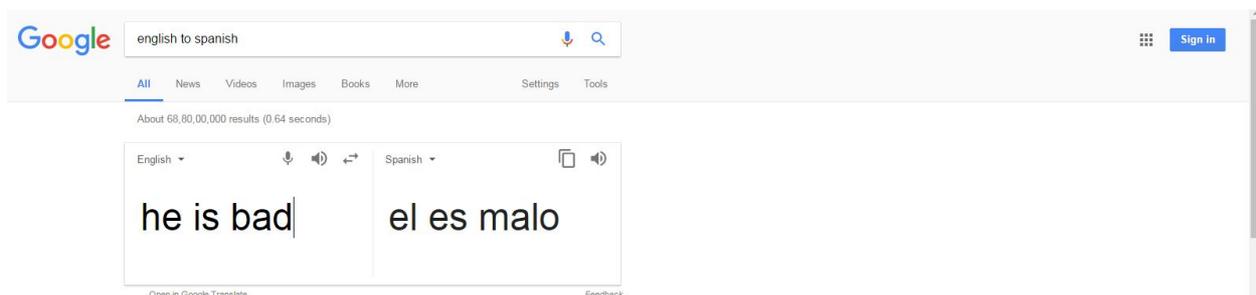


Figure 1(a)



Figure 1(b)

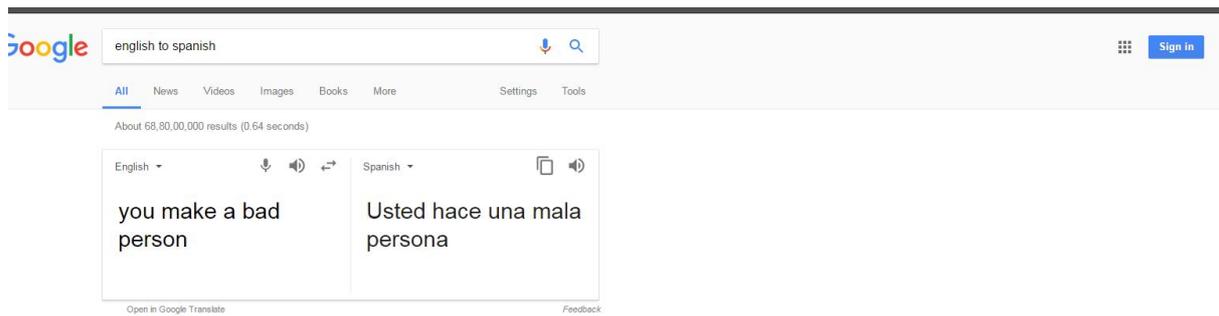


Figure 1(c)

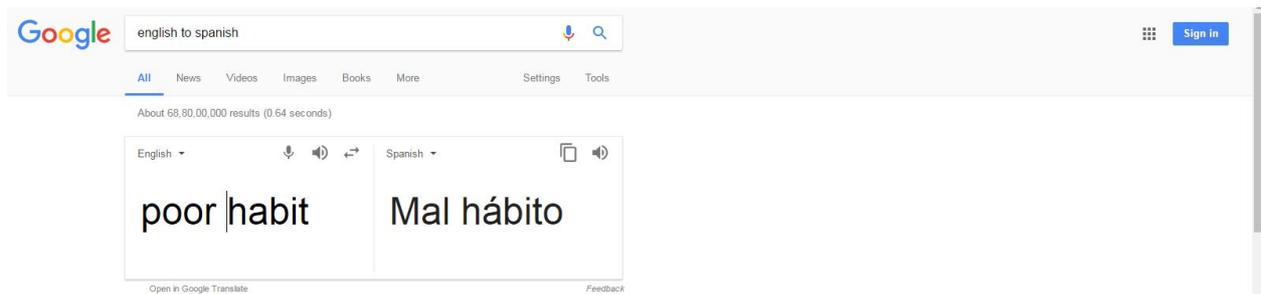


Figure 1(d)

Considering the following examples from Hindi,

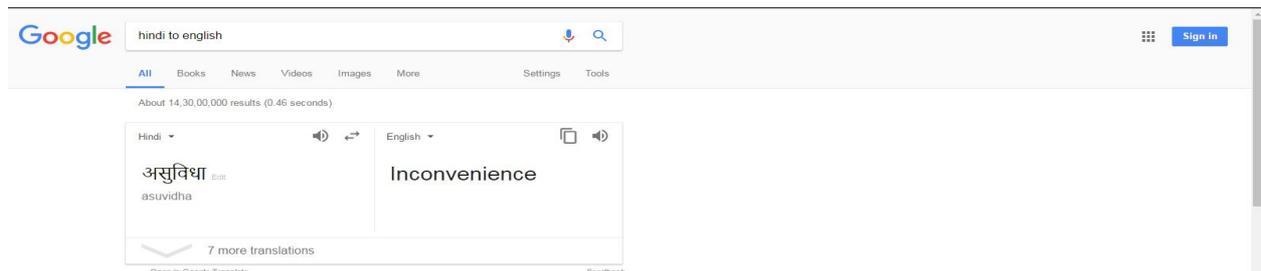


Figure 2(a)



Figure 2(b)



Figure 2(c)

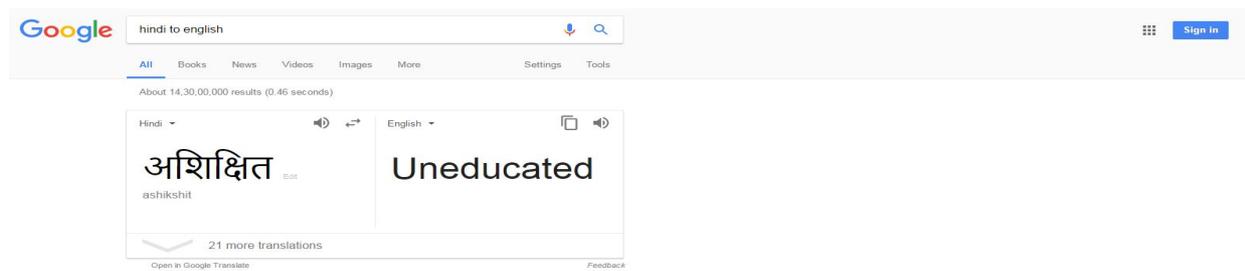


Figure 2(d)

It can be said that 'अ' is used as a prefix in Hindi to negate or in a negative sense.

At a macro level, we can say that the overall opinion about the current text is negative in nature.

The contradictions to this approach definitely exist but as stated at the beginning of the paper, that a solution working for all the corpora of all the languages is not possible, instead a degree of problem resolution can be achieved.

Moreover, this is evident that one solution may work for more than one language.

- The second approach for smaller, informal texts is emoticons. This approach is suitable for texts exchanged among netizens. It is truly a pictorial representation of the facial expression. The expression or the icon absolutely represents the mood or sentiment of the speaker. With emoticons, the need to decode or opine is uncalled for. The icon states the mood it stands for.

A lot has been said about sentiment analysis using emoticons especially when it comes to tweets.

The advantage with emoticons is that, they are independent of language as they are the pictorial representation of the sentiment.

- Thirdly, the use of figures of speech, idioms and adage or proverbs. Proverbs and sayings are nothing but universal truths or advice based on the experience of generations.

Another notable fact is that, proverbs and idioms have always been there. There is no such thing as coinage and new formulation of proverbs. Since they have existed for very long and are not undergoing a continuous process of making, they are finite in number. Taking advantage of the fact that such phrases are countable or finite, it is possible to maintain a set of them.

If a text comes across any of the proverbs stated in the set, it can be termed as neutral.

Another possibility is that if we already tag all the idioms and phrases as positive and negative (since their meaning is already established), the text where they are used can automatically be analyzed for its positive or negative sense.

3. CONCLUSION:

Sentiment is an attribute of living beings who have some thinking mechanism. Computers, on the other hand are machines, which are being made smart by constant human effort. Thus by the law of nature, it is difficult to make machines feel but the attempt is made for machines to deduce the feeling or sentiment of human. None of the approaches completely fit into the solution space but it is an attempt to attain a high degree of problem solution.

References

- [1] Pragya Tripathi, Santosh Kr. Vishwakarma, Ajay Lala, "Sentiment Analysis of English Tweets Using Rapid Miner", <http://ieeexplore.ieee>.
- [2] Haseena Rahmath P, "Opinion Mining and Sentiment Analysis - Challenges and Applications", International Journal of Application or Innovation in Engineering & Management (IJAIEM) Volume 3, Issue 5, May 2014