Gradient Emotional Analysis

Lilia Simeonova
Sofia University “St. Kliment Ohridski”
Lilia.valentinova.simeonova@gmail.com

Abstract

Over the past few years a lot of research has been done on sentiment analysis, however, the emotional analysis, being so subjective, is not a well examined discipline. The main focus of this proposal is to categorize a given sentence in two dimensions - sentiment and arousal. For this purpose two techniques will be combined – Machine Learning approach and Lexicon-based approach. The first dimension will give the sentiment value – positive versus negative. This will be resolved by using Naïve Bayes Classifier. The second and more interesting dimension will determine the level of arousal. This will be achieved by evaluation of given a phrase or sentence based on lexicon with affective ratings for 14 thousand English words.

1 Introduction

A lot of papers related to sentiment analysis show very good results, more than 80% accuracy. This kind of analysis determines whether the sentence is positive or negative (on some occasions neutral as well). Such information can be useful in terms of understanding the intentions of a given user.

The proposed research will try to take this knowledge one level deeper. This means that additional analysis will be made on the level of affection of a given sentence. Not only will this program be able to recognize positive or negative text, but will also determine its arousal level.

This project aims to examine the way people’s reactions are interpreted by machines. Better and more detailed analysis of opinions, reviews, tweets, etc, can be applicable in all kinds of domains, such as commercial, retail, politics and psychological research.

2 Related Work

Hasan, Rundensteiner and Ago (2014) proposed in their paper “EMOTEX: Detecting Emotions in Twitter Messages” that twitter hashtags can be used to automatically extract data suitable for training purposes. They retrieve tweets related to four basic groups (shown below) and apply standard Machine Learning Algorithms on this data.

- Happy-Active
- Happy-Inactive
- Unhappy-Active
- Unhappy-Inactive
They receive good results, however, often people use misleading hashtags in order to express irony. This can lead to wrong assumptions.

3 Gradient Emotional Analysis

3.1 Circumplex Model of Affect

This paper was initially inspired by “The Circumplex model of affect”. James Russell (1980) showed that emotions can be seen in a two-dimensional circular space. This space contains arousal dimension (presented in the vertical axis) and valence dimension (presented in the horizontal axis).

A diagram, showing this representation is provided in Figure 1.

The Circumplex model of affect can be applied in this paper as a starting point for more detailed investigation of emotions.

![Figure 1: Circumplex model of Affect, made by James Russell.](image)

3.2 Emotional Representation

The main idea behind the proposed paper is to analyze the input based on two different dimensions: sentiment and arousal. This is shown in Figure 2 below.

![Figure 2: Emotional Analysis Diagram.](image)

For example the verb “excite” can be categorized as positive with high arousal level, while “calm” goes under classification of positive with low arousal level.

Similar logic puts the verb “frustrate” under negative classification with high arousal, while “bore” should be negative verb with low level of arousal.

Both of the dimensions will be described in this paper.

4 Sentiment Analysis

The first part of the analysis will be achieved with the Naïve Bayes Classifier. It is one of the simplest and most commonly used Machine Learning techniques which is why it will fit perfectly in this solution.

Naïve Bayes is based on the Bayes Theorem showed on Equation 1.

\[
P(A|B) = \frac{P(B|A)P(A)}{P(B)}
\]

Equation 1: Bayes Theorem.

In order to create a classifier which aims to provide a positive or negative value for a given input, it needs to be decided what its features are going to be.
Feature Selection

Feature selection and extraction is a key element in the Naïve Bayes implementation. After several tests, it has been decided that a NLTK implementation of part-of-speech (POS) Stanford tagger, will be used for this experiment.

The steps performed:

- Text Preprocess – tokenization
- POS tag in tuples: (word, POS value)

After the part-of-speech tagging, only verbs, adjectives and adverbs will be extracted as they bring the most value for sentiment analysis.

5 Level of Arousal/Affect

For determining the level of Affect for a given text the Lexicon-based approach will be used.

Almost the same process is followed as the one in the Sentiment Analysis section but with one major difference.

- Text Preprocess – tokenization
- POS tag in tuples: (word, POS value)

**Exclamation mark (!) is not removed from the tokens. It can give a very good understanding of whether the person who wrote the text is feeling an emotion with high arousal.**

- Retrieve the arousal value for every tagged word.
- Apply equation 2.

\[
\text{avg} \left( \sum_{\text{Tagged word}} \right)
\]

Equation 2: Formula to calculate the average arousal level of given text.

6 Experiments and Evaluation

6.1 Data

Three different datasets were used for this experiment.

A dataset with 8 885 annotated tweets has been used to train the Naïve Bayes Classifier –3 593 are labeled as positive and 5 292 are labeled as negative.

For the second part of the research – calculating the level of arousal, a lexicon with affective ratings has been used. It contains 13 915 annotated English words with a number from 1 to 9.

Another dataset has been used to verify the results from arousal level analysis. The number of tweets related with specific emotion are presented in Table 1.

| Emotion                        | Tweets |
|--------------------------------|--------|
| Enthusiasm + Hate              | 700    |
| Relief (Happiness) + Sadness   | 700    |

Table 1: Number of tweets extracted from Dataset.

6.2 Experiment and Results

Sentiment level – Naïve Bayes Classifier

10-fold cross-validation was used to measure the accuracy of the proposed algorithm. The average result from the 10 folds is 72.8%.

Arousal level

In order to verify the accuracy of the proposed arousal evaluation a training data was used with tweets annotated with words which represent active and inactive emotions. The result showed 76% accuracy for active tweets and less than 50% accuracy for inactive ones.

This means that the proposed testing approach is not working.
This failure is because the dataset which was used is automatically generated, based on hashtags. For example the following tweet extracted from the testing set is evaluated as #boredom, because the person used it with his hashtag:

“I'm so tired of being sick ALL the time!!!! #boredom”.

The word “boredom” should be classified as inactive, however the language used in this tweet is showing immense frustration.

Survey

Another way of evaluating the algorithm is to ask people to evaluate manually the level of arousal.

Three people were asked to assign value 0 for inactive tweets and 1 for active.

This survey was based on 100 tweets and the average value of the results is showing 63% accuracy.

The final results are shown in Table 2.

|                  |       |
|------------------|-------|
| Sentiment Analysis | 72.8% |
| Arousal Level Analysis | 63%   |

Table 2: Final Results.

7 Future Work

In order to improve the results in the Arousal level analysis other features need to be taken into consideration. For example, emoticons can be added in the lexicon. Another thing that might improve the results is capturing tweets written with caps lock.

8 Conclusion

This paper explores the possibilities of developing a system, which can have the capability to recognize both sentiment and arousal level of a given text. The research is using the Circumplex Model of Affect (1980) where James Russel states that the emotions can be divided in four main categories – positive-active, positive-inactive, negative-active, negative-inactive.

The implementation is based on the idea to combine two very well-known approaches – Machine Learning Approach and Lexicon Based Approach.

The first approach is to use already labeled dataset. Upon training, it can be used to classify the sentiment of an input. For its realization good results were achieved using Naïve Bayes Classifier. For future work another ML algorithm may be used.

The lexicon-based approach is used for the second part of the project. In this case the work is happening only on sentence level where the text is being tokenized. This is the moment where it is decided whether every token should be included in the analysis. An interesting part of the research was to determine which parts of the speech are best suited to be analyzed for level of affection. The idea to not exclude exclamation marks also increased the correctness of the returned values. After extracting only the needed values the arousal level is retrieved from the lexicon.

9 References

Warriner, A.B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. Behavior Research Methods, 45, 1191-1207

Russell J., (1980). "A Circumplex Model of Affect". Journal of Personality and Social Psychology. 39: 1161–1178

Mika V. Mäntylä Daniel Graziotin, Miikka Kuutila (2015) The Evolution of Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers, arXiv:1612.01556
Twitter Sentiment Analysis Training Data:  
http://thinknook.com/twitter-sentiment-analysis-training-corpus-dataset-2012-09-22/

Hasan M., Rundensteiner E., Agu E.,  
(2014) EMOTEX: Detecting Emotions in Twitter Messages. Academy of Science and Engineering (ASE)

Liu B., Zhang L. (2013), A survey of opinion mining and sentiment analysis, In Mining Text Data (pp. 415-463), Springer US. DOI: 10.1007/978-1-4614-3223-4_13

Lowd D, Domingos P. (2005) - Naive Bayes Models for Probability Estimation, Proceeding ICML '05 Proceedings of the 22nd international conference on Machine learning Pages 529-536

Kanayama, H. and T. Nasukawa (2006) Fully automatic lexicon expansion for domain-oriented sentiment analysis. In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP- 2006)

Mohammad, S., C. Dunne., and B. Dorr. (2009) 
Generating high-coverage semantic orientation lexicons from overly marked words and a Thesaurus. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2009).

Qiu, L., W. Zhang., C. Hu., and K. Zhao. (2009) SELC: A self-supervised model for sentiment classification. In Proceedings of ACM International Conference on Information and knowledge management (CIKM-2009)

Taboada, M., J. Brooke, M. Tofiloski, K. Voll, and M. Stede (2010), Lexicon-based methods for Sentiment analysis. Journal Computational Linguistics archive Volume 37 Issue 2, June 2011 Pages 267-307

Pang B., Lee L., Vaithyanathan S (2002) Thumbs up? Sentiment classification using machine learning techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 79–86, 2002.

Narayanan V., Arora I., Bhatia A. (2013) Fast and Accurate Sentiment Classification Using an Enhanced Naive Bayes Model. In: Yin H. et al. (eds) Intelligent Data Engineering and Automated Learning – IDEAL 2013. IDEAL 2013. Lecture Notes in Computer Science, vol 8206. Springer, Berlin, Heidelberg

Kristina Toutanova and Christopher D. Manning.  
2000. Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger. In proceedings of the Joint SIGDAT Conference on empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC-2000), pp. 63-70

Hernandez-Farias, DI.; Patti, V.; Rosso, P. (2016). Irony Detection in Twitter: The Role of Affective Content. ACM Transactions on Internet Technology. 16(3):19:1-19:24. doi: 10.1145/2930663.