Automatic Emotion Detection from Unstructured/Semi Structured Text

Sajitha N

Assistant Professor, CSE, BNMIT, Bengaluru, Karnataka, India

Abstract- **Automatic emotion detection is a process of detecting emotions from facial expressions, text or audio clips. Collecting and labeling such data is a difficult process and time consuming. This paper proposes detecting emotions from text documents using the Naïve Bayes classifier. Detecting emotions from text can be a used a useful tool in psychological counseling. Different methods are used to detect emotions using signal processing, AI, text processing and Bayesian network. Emotion detection from text is from the written text or speech text.**

Index terms- **Emotion detection, Text detection, Sentiment analysis, Opinion mining, Bayesian classifier, Labelling.**

1. INTRODUCTION

An emotion is caused by a person, consciously or unconsciously, evaluating an event. Psychologists refer to this as appraisal.

Automatic detection and classification of sentiments has several potential areas of application. Marketing research, for instance, can benefit from using such tools for finding out what the general public thinks of their products.

Emotion can be expressed in many ways that can be seen such as facial expression and gestures, speech and by written text. Emotion Detection in text documents is essentially a content based classification problem involving concepts from the domains of opinion mining,

Natural language processing as well as Machine Learning. In this work of emotion recognition based on textual data and the techniques used in emotion detection are discussed.

Emotions can significantly change the message: sometimes it is not what was said that is the most important, but how it was said.

Emotion is expressed as joy, sadness, anger, surprise, hate, fear and so on. Since there is not any standard emotion word hierarchy, focus is on the related research about emotion in cognitive psychology domain.

2. PROPOSED SYSTEM

Emotions are also articulated by written texts. Inspired by works in sentiment analysis, this work explores approaches to automatic detection of emotions in text

The ultimate goal is to combine the annotations with additional affective information collected during experimental learning sessions from different sources such as qualitative, self-reported, physiological, and behavioral information. These data altogether are to train data mining algorithms that serve to automatically identify changes in the learners' affective states when dealing with cognitive tasks which help to provide emotional personalized support.

3. MOTIVATION

Generally individuals and companies are always interested in other's opinion like if someone wants to purchase a new product, then firstly, he/she tries to know the reviews i.e., what other people think about the product and based on those reviews, he/she takes the decision. Similarly, companies also excavate deep for consumer reviews. Digital ecosystem has a plethora for same in the form of blogs, reviews etc.

These solutions include extracting keywords with semantic analysis, and ontology design with emotion theory of appraisal. Furthermore, a case-based reasoning architecture is proposed to combine these solutions.

Web communication can be facilitated through video, voice recordings, images and text. The most prevalent form of communication on the web exists in the form of text which offers a rich platform for expressing emotions. In the absence of face-to-face contact to detect facial expressions and intonations in

voice, the alternative option is to decipher emotions from text in online forums.
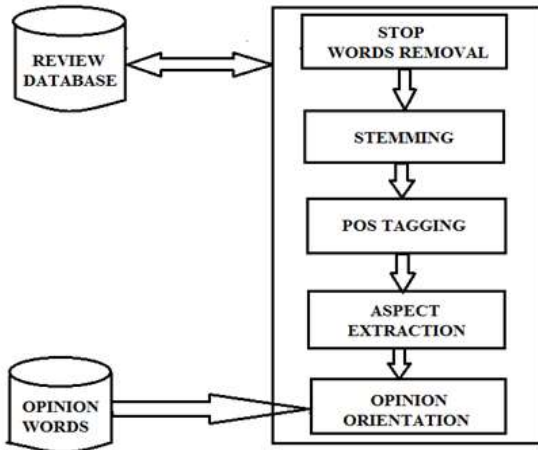
4. SYSTEM ARCHITECTURE

Fig 1: Proposed model

a. Stop word removal

Most frequently used words in a document are called as stop words, which are not useful in text mining.

b. Stemming

Stemming is a process of finding the route word from the actual word, for example longing, longer are the two words then the route word is long.

c. POS tagging

Part of speech (POS) of a sentence describes its linguistic and morphological behaviors. Tagging each word with it POS is important phase in text mining.

d. Aspect extraction

Aspect extraction is used to find the frequent aspect of given sentence. Every noun word is assigned as aspect transaction and finds the minimum count of it.

e. Opinion words

Positive and negative opinion collect from the sentence will be labeled with opinion. For example good, awesome are the positive opinion and bad, ugly are the negative opinion.

f. Opinion orientation

Naïve Bayesian algorithm can be used to classify the positive and negative opinion.

Three approaches currently dominate the emotion detection task; keyword based, learning based and hybrid based approach. These make use of features mainly selected from syntactic and semantic data to detect emotions.

There are three major areas of research under opinion mining.

1   Sentiment classification: This is to label the document in binary form and making an opinion as either an overall positive or negative. Various machine learning algorithms can aid in this segment analysis.

2   Feature based sentiment classification: This enables to perform analysis on the components or attributes of the product for recognizing the sentiments in the document. This requires more refinement than other models.

3   Opinion summarization: It presents a concise view of the large number of opinions of different persons on the social web. It starts from some raw opinionated data up to the generation of human understandable summaries. This includes various methods like NLP, analysis, text mining, sentiment prediction etc.

Opinion mining can be useful in several domains like in search engines, question answering systems, recommendation systems etc.

It also helps in developing better human computer communications. The major application areas are buying or selling products or services, quality control areas, policy or decision making, business intelligence by conducting marketing research and so on.

By applying sentiment analysis on the feedbacks received by people by classifying into positive, negative and neutral groups automatically provides valuable information for further analysis on market reports.

In order to perform large-scale analysis of emotion phenomena and social behaviors on social media, there is a need to first identify the emotions that are expressed in text as the interactions on these platforms are dominantly text-based. With the surging amount of emotional content on social media platforms, it is an impossible task to detect the emotions that are expressed in each message using manual effort.

A broader range of emotions will enable automatic emotion detectors to capture more fine-grained motions that truly reflect actual human emotional experience.

Algorithms to identify sentiment and sentiment strength are needed to help understand the role of emotion in this informal communication and also to identify inappropriate or anomalous affective utterances, potentially associated with threatening behavior to the self or others.

It addresses in two sub-challenges the detection of dimensional act in continuous time and value, and the estimation of self-reported depression.

REFERENCES

[1] E. Cambria, B. Schuller, Y.Q. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," IEEE Intelligent Systems, vol. 28, no. 2, 2013, pp. 15-21

[2] D. Savage , X. Zhang, X. Yu, P. Chou, Q. Wang," Anomaly detection in online social networks",DOI: 10.1016/j.socnet.2014.05.002, Science Direct, 2014.

[3] "Computational Approaches for Emotion Detection in Text" Haji Binali, Chen Wu, Vidyasagar Potdar 4th IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST 2010) © 2010 IEEE.

[4] 4."Data Intensive Review Mining for Sentiment Classification across Heterogeneous Domains" Federica Bisio, Paolo Gastaldo, Chiara Peretti, and Rodolfo Zunino 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.

[5] E. Cambria and A. Hussain, "Sentic Computing: Techniques, Tools, and Applications" Heidelberg: Springer (2012)

[6] D. Bollegala, D. Weir, and J. Carroll, "Cross-Domain Sentiment Classification using a Sentiment Sensitive Thesaurus," IEEE Trans. On Knowledge and Data Engineering, in press, DOI 10.1109/TKDE.2012.103

[7] K. Khan, B. Baharudin, A. Khan, A. Ullah,"Mining opinion components from unstructured reviews: A review", 1319-1578 2014, 2012.

[8] Y.W. Lo, V. Potdar, "A review of opinion mining and sentiment classification framework in social networks" Digital Ecosystems and Technologies, 2009. DEST '09. 3rd IEEE International Conference DOI: 10.1109/DEST. 2009.5276705, Page(s): 396 – 401, IEEE Xplore, 2009.

[9] P. T. Ngoc, M. Yoo," The lexicon-based sentiment analysis for fan page ranking in Facebook", Information Networking (ICOIN), 2014 International Conference, 10-12, 444-448,10.1109/ICOIN.2014. 6799721, IEEE Xplore, Feb 2014.

[10] I.Peñalver-Martinez, F. Garcia-Sanchez, R. Valencia-Garcia, M. Á. Rodríguez-García, V. Moreno, A. Fraga, Jose Luis Sánchez-Cervantes, "Feature-based opinion mining through ontologies", Expert Systems with Applications, Volume 41, Issue 13, 1, Pages 5995-6008, Science Direct, October 2014.