Emotion Analysis On Text

Bhagyashree
Directed by: Prof. U. S. Tiwary

Indian Institute of Information Technology, Allahabad

March 6, 2019

Overview

Emotion
Analysis On
Text

Bhagyashree

Introduction

Problem
statement and
Objective

Methodology
Keyword based
detection
Learning based
detection
Hybrid Detection

Implementation
Removal of Stop
words
Pos tagging

Future Scope
of Project

References

. Introduction
. Problem Definition and Objective
. Methodology
. Implementation
. Future Scope of Project
. References

What are Emotions?

It is difficult to answer the question "How many emotions we have?" But Some of the common emotions we express in our daily lives include Joy, Happiness, Fear, Sadness, Surprise, Anger and so on...which combine to form other emotions.

For example:

Excitement = Joy and Surprise

Disgust = Anger and Sadness

Emotional state of a person can be recognized by:

. Speech

. Facial expressions

. Text

Emotion
Analysis On
Text

Bhagyashree

Introduction

Problem
statement and
Objective

Methodology
Keyword based
detection
Learning based
detection
Hybrid Detection

Implementation
Removal of Stop
words
Pos tagging

Future Scope
of Project

References

Problem statement and Objective

Problem statement

I aim to make machines understand and detect emotions from any textual input.

. Text based emotion recognition system

. To give an overall sentiment of the sentence/paragraph in terms of positivity/negativity and neutrality

. Determine the overall emotion

Objective

Given any textual input such as a word/phrase/sentence the machine should be able to give a percentage of various emotions that constitutes the input.

Methods

There are three basic methods to detect emotions from text:
1. Keyword based detection
2. Learning-based detection
3. Hybrid detection

Tools

For syntactic structure of the word, Continuous Bag of Words(CBOW) and skip-gram architecture is being used. Concept of Word2vec(Developed by Google) is used.

Procedure

1. nltk(natural language toolkit) package is used
2. Assign Emtion labels(positive or negative)
3. Removal of unnecessary words(stop words)
4. Pos tagging is done using nltk.postag(sentence)
5. Stemming and creation of data frame
6. Wordnet is created which find the synonyms and antonyms of the word

Procedure

1. Detect emotions based on a previously trained classifier
2. Word2vec tool is used.
3. To get vectors of words, Continuous Bag of Words and skip-gram architecture is used.
4. Machine learning algorithms such as Naive Bayes classifier, support vector machines and k means clustering is applied

Procedure

1. Combination of 2 methods such as Keyword based and learning based

2. Emotions are detected based on the combination of detected keywords, learned patterns and other required information.

Removal of stop words

1. Stop words are a set of commonly used words in any language.

2. These are few unnecessary words in the sentence, which needs to be removed.

3. Reduces memory overhead

4. Can potentially improve power of prediction

Pos tagging

1. Not all words contribute to the level emotion in a sentence.

2. Certain parts of speech specifically contribute to the emotion level in a sentence.

3. It tokenizes as well as indicate various parts of speech.

In [24]:
```python
#Removing Stop words
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize

s1 = "I am trying to remove stop words from this sentence using stopwo
stop_words = set(stopwords.words('english'))
word_tokens = word_tokenize(s1)

#filtered_sentence = [w for w in word_tokens if not w in stop_words]
filtered_sentence = []

for w in word_tokens:
    if w not in stop_words:
        filtered_sentence.append(w)

print(word_tokens)
print(filtered_sentence)
```

```
['I', 'am', 'trying', 'to', 'remove', 'stop', 'words', 'from', 'this',
f', 'nltk', '.']
['I', 'trying', 'remove', 'stop', 'words', 'sentence', 'using', 'stopw
```

Figure: Removal of stop words from a sentence

In [28]:
```python
# Pos Tagging removing stop words
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize, sent_tokenize

s1 = "I am trying to remove stop words from this sentence using stopwo
stop_words = set(stopwords.words('english'))
word_tokens = sent_tokenize(s1)

#filtered_sentence = [w for w in word_tokens if not w in stop_words]
filtered_sentence = []

for w in word_tokens:
    wordsList = nltk.word_tokenize(w)
    #for i in wordsList:
     # if i not in stop_words:
    wordsList = [w for w in wordsList if not w in stop_words]
    tagged = nltk.pos_tag(wordsList)

print(tagged)
```

```
[('I', 'PRP'), ('trying', 'VBG'), ('remove', 'VB'), ('stop', 'JJ'), (
'VBG'), ('stopword', 'NN'), ('library', 'NN'), ('nltk', 'NN'), ('.', '
```

Figure: Pos tagging

Future work

1. Emotion detection is a growing research field. A lot of work has already been done on sentiment analysis, facial recognition and speech recognition.

2. My project aims to recognize emotions in a word/sentence or paragraph.

3. Determining overall impact of sentence can help in various fields specially in marketing. Consumers state of mind can be read by analysing what they write.

4. More features(emotions) can be added to detect separately like recognizing sarcasm.

📄 Chopade, C. R. (2013).
   Text based emotion recognition: A survey.

📄 Haji Binali, C. w. (2010).
   Computational approaches for emotion detection in text.

📄 V V Ramalingam1, A. P. (2018).
   Emotion detection from text.