

Report  
On  
Emotion Detection in Text



By

Bhagyashree

MIT2018002

Under the supervision of:

Prof. U. S. Tiwari

Professor

IIIT-ALLAHABAD

**INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, ALLAHABAD**

(A UNIVERSITY ESTABLISHED UNDER SEC.3 OF UGC ACT, 1956 VIDE NOTIFICATION NO. F.9-4/99-U.3

DATED 04.08.2000 OF THE GOVT. OF INDIA)

A CENTRE OF EXCELLENCE IN INFORMATION TECHNOLOGY ESTABLISHED BY GOVT. OF INDIA

### **CANDIDATE'S DECLARATION**

I hereby certify that the work which is being presented in the M.Tech. Mini-Project Report entitled “Emotion analysis in text”, being submitted as a part of Mid Semester Project Evaluation to the Department of Information Technology of Indian Institute Of Information Technology, Allahabad, is an authenticated record of my original work under the guidance and supervision of Prof. U.S. Tiwari from February 2019 to June 2019. I have adequately cited and referenced the original sources and have adhered to all principles of academic honesty and integrity.

Date: March 2019

Place: IIIT Allahabad

Signatures:

Bhagyashree

(MIT2018002)

### **CERTIFICATE FROM SUPERVISOR**

This is to certify that the statement made by the candidates is correct to the best of my knowledge and belief. The project entitled “Emotion analysis in text” is a record of candidates’ work carried out by her under my guidance and supervision. I do hereby recommend that it should be accepted in the fulfilment of the requirements of the mini project at IIIT Allahabad.

Prof. U.S. Tiwari, IIIT-Allahabad

March, 2019

## **ACKNOWLEDGEMENT**

I would like to acknowledge and extend our heartfelt gratitude to Prof. U.S. Tiwari, Indian Institute of Information Technology Allahabad, who guided me through this project. His keen vital encouragement, superb guidance, and constant support are the motive force behind this project work. I would like to show my warm thanks to all my friends for their continuous motivation and encouragement during this project. I am very thankful to all the technical and non-technical staffs of the college for their assistance and cooperation.

Bhagyashree

MIT2018002

## Table of Content

---

|   |      |
|---|------|
| 1. Vision & Objective.....                | (6)  |
| 2. Motivation.....                        | (7)  |
| 3. Introduction & Problem definition..... | (8)  |
| 4. Literature Review.....                 | (9)  |
| 5. Methodology.....                       | (10) |
| 6. Implementation.....                    | (12) |
| 7. Future scope of the project.....       | (13) |
| 8. References.....                        | (13) |

## **VISION**

---

In this era of artificial intelligence, intelligent machines are being designed which are expected to perform tasks which can be performed by humans. We are making machines that are intelligent and smart just like human beings are. However, one thing machines can't do is, not able to understand by itself and can't identify various emotions just like humans can do. Despite so much progress in the field of machine learning and natural language processing, machines have still not become capable of detecting emotions clearly.

There are some basic form of expressing the emotions: Through facial expressions, speech or text. A lot work is being done in recognizing emotions through facial expressions and speech, yet we haven't been able to develop machines that can detect emotions in a text successfully.

In the era of internet world, most of the information exists in the form of text. Be it digital encyclopedias, websites, search engines; text constitutes around 70% of the information available to us. Hence, our aim is to make machines detect emotions hidden in every piece of textual information available.

## **OBJECTIVE**

---

Through my project I aim to make machines detect emotions from any form of text, which constitutes about 70% of information available to us. Given any form of text, the machine will identify the specific emotions of happiness, sadness, anger, surprise, fear or love that the text expresses. Understanding emotions exposes us to many possibilities: personalized information generation, like advertisements, search results and so on, development of powerful human computer interaction machines and evolution of more intuitive and emotionally characterized text to speech systems.

## MOTIVATION

---

Emotion Analysis is the broadly studied application of NLP (natural language processing) and Machine learning. The goal of Emotion Analysis is to extract important information regarding public opinion which helps to make the business decisions easily, political campaigns and to increase the product consumption.

A lot of work has been done in the field of sentiment analysis. However, it is a difficult task to study the emotion of a human mind. Thus, I have taken this project to study about the different emotions of a man kind. My goal is to find the emotions of a particular word, and sentence. What the overall impact of sentence is?

I aim to design such a machine which can analyse the emotions in text/sentence or even in paragraph and can tell the overall impact of the paragraph or a chapter of a book. It will help in understanding the human computer interaction.

Also in text classification the purpose of using text classification is to determine the subjective value of a text-document, i.e. how positive or negative is the content of a text document.

## INTRODUCTION

---

Definitions about emotions and its categories have been an important research since long. However, till now no one has been able to answer the question “How many different emotions we have?” because human emotions are complex. Deriving one emotion from other is a quite complex task because it depends on various factors. It is not easy to read one’s state of mind. It can be quite hard to determine that where one emotion starts and where it ends. However, one can have multiple emotion at once, so it becomes even more complex to determine.

According to Paul Ekman model, six basic emotions are concluded which are anger, disgust, fear, happiness, sadness, and surprise. Each emotion has its own neural network in the brain, and corresponding behavioural response.

These basic emotions which compose/combine to form what advanced emotions. For example: the emotion of excitement can be a combination of joy and surprise whereas the emotion of disgust can be a combination of anger and sadness.

There are many challenges while detecting emotion in text. One is context-dependence of emotions within text. A phrase can have element of *anger* without using the word “anger” or any of its synonyms. For example, the phrase “Shut up!”

Another challenge in emotion detection is the lack of a labelled emotion database. Another challenge faced is ambiguity in the text.

## PROBLEM DEFINITION

---

I aim to make machines understand and detect emotions from any textual input. A text based emotion recognition system which give an overall sentiment of the sentence/paragraph in terms of positivity/negativity and neutrality and determine the overall emotion.

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.



## LITERATURE REVIEW

---

The field of emotion recognition has attracted many researchers from computer science, psychology, cognitive sciences, linguistics, and so on. Emotions can be expressed by a person through facial expressions, speech, gestures, and textual conversations. Speech and facial recognition have been the most preferred form for researchers all around the world for detecting emotions as they convey emotion more clearly without any disambiguates. Comparatively, very less work has been done in the field of emotion recognition through text.

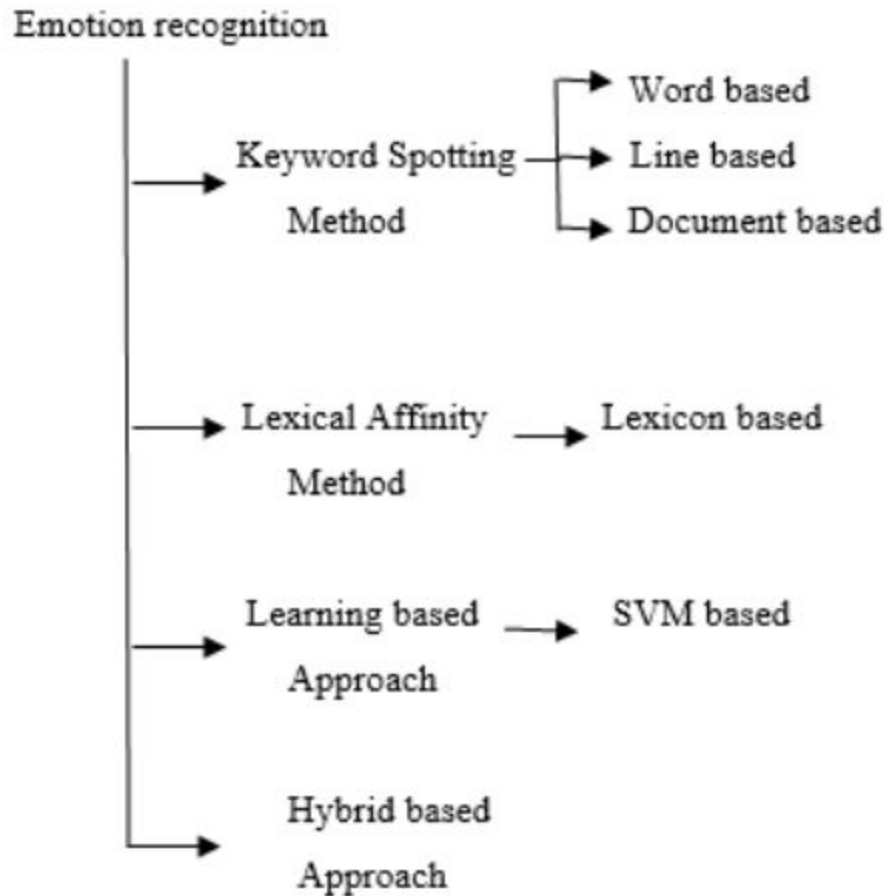
Within text, considerable amount of work has been done in sentiment analysis of movie reviews, twitter feeds and so on. However, nothing significant has been achieved in detecting raw emotions of happiness, sadness, anger, fear, etc in text. The main reason for this is that there are no standard classifications of all human emotions due to complex nature of human minds and any emotion classifications can be seen as ‘labels’ annotated afterwards for different purposes. Since there is not any standard emotion-word hierarchy, focus is on related research about emotion in cognitive psychology domain.

Researchers have used techniques such as keyword spotting, assigning probabilistic affinities to various emotions, use of learning based methods and knowledge based artificial neural networks, rule based system to extract semantics related to specific emotions , detecting emotions based on the cause triggering them and development of situational personalized emotion model.

However, all the present researches do not take into account explicit emotions and analyze only a specific part of the psychological reason that leads to emotion in texts. They have not analyzed lack of keywords, disambiguate, lack of linguistic information and do not take the context into account. The corpus used in the research has been specifically developed such that they suit the main theme of the paper. None of the systems can produce satisfactory results for texts with disambiguates or lack of linguistic information.

## METHODOLOGY

---



There are four different text based emotion recognition techniques: Keyword based detection method, Lexical Affinity Method, learning based method and hybrid methods. These methods are divided into subcategories as shown in above figure.

- 1. Keyword based detection:** This method includes finding occurrences of keywords from a given set as substrings in a given string. It finds the presence of keywords and may involve pre-processing with a parser and emotion dictionary. This can be applied in real time chat systems. It is domain specific, relies on the presence of keywords for accurate results and requires pre-processing for improved accuracy results.

2. **Lexical Affinity Method:** Lexical Affinity method is an extension of keyword spotting method. This assigns a probabilistic affinity for a particular emotion to arbitrary words rather than detecting predefined emotional keywords from text.
3. **Learning based approach:** This approach uses a trained classifier to categorise input text into emotion classes by using keywords as features, which mapped with various machine learning classifiers such as support vector machines, k-means classifier etc.
4. **Hybrid Based Approach:** This approach consists of a combination of the keyword based implementation and learning based implementation. The main advantage of this approach is that it can yield higher accuracy results from training a combination of classifiers and adding knowledge-rich linguistic information from dictionaries. It will balance the high cost involved for information retrieval tasks and minimize difficulties.

## IMPLEMENTATION

---

Given any sentence, the first task is to consider individual words to analyze the emotional content. This is achieved by a process called tokenizing. I consider the tokens in my case to be separated by whitespace characters, such as a space or line break, or by punctuation characters. For the same, I have used nltk(natural language toolkit) package for the removal of stop words and for the pos tagging.

**Removal of Stop words:** Stop words are basically a set of commonly used words in any language. These are few unnecessary words in the sentence, which needs to be removed. Removing the stop words from the sentence reduces memory overhead and also it can potentially improve power of prediction.

**Pos Tagging:** In a sentence, not all words contribute to the level emotion in a sentence. Certain parts of speech specifically contribute to the emotion level in a sentence. It tokenizes as well as indicate various parts of speech.

### Removal of Stop words:

```
In [24]: #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 stopwords library of nltk."
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', 'sentence', 'using', 'stopword', 'library', 'of', 'f', 'nltk', '.']
['I', 'trying', 'remove', 'stop', 'words', 'sentence', 'using', 'stopword', 'library', 'nltk', '.']
```

## Pos Tagging:

```
In [28]: # 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 stopwords library of nltk."
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'), ('am', 'VBG'), ('trying', 'VBG'), ('remove', 'VB'), ('stop', 'JJ'), ('words', 'NNS'), ('sentence', 'NN'), ('using', 'VBG'), ('stopword', 'NN'), ('library', 'NN'), ('nltk', 'NN'), ('.', '.'), ('.', '.')]

```

## **Future scope of the project**

---

Emotion detection is a growing research field. A lot of work has already been done on sentiment analysis, facial recognition and speech recognition. My project aims to recognize emotions in a word/sentence or paragraph. Determining overall impact of sentence can help in various fields especially in marketing. Consumers state of mind can be read by analysing what they write. More features(emotions) can be added to detect separately like recognizing sarcasm.

## **References**

---

1. Emotion detection from text by V V Ramalingam.  
<https://iopscience.iop.org/article/10.1088/1742-6596/1000/1/012027/pdf>
2. Computational Approaches for Emotion Detection in Text by Haji, Chen Wu  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.863.8593&rep=rep1&type=pdf>
3. Text Based Emotion Recognition: A Survey by Chetan R. Chopade  
<https://pdfs.semanticscholar.org/50cc/848978797112a27d124af57991321f545069.pdf>
4. Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text  
<https://ieeexplore.ieee.org/document/6797872>