|  |  |
| --- | --- |
| Name Of The Student | Aman Rai |
| Internship Project Topic | Automate Detection of different emotions from textual comments and feedback |
| Name of the Organization | TCS iON |
| Name of the Industry Mentor | Mr. Debashis Roy |
| Name of the Institute | Institute of Engineering & Management Kolkata |

|  |  |  |
| --- | --- | --- |
| Date | Day # | Hours Spent |
| 5 Aug 2020 | 2 | 4 |
| On Day2 I have done some self-learning and go through some youtube video to know emotion detection in brief, I am going to mention the reference links in the last part of report.  Activities done during the day: The concept of affective computing in 1997 by Since Picard [3] proposed that the role of emotions in human computer interaction. This domain attracted many researchers from computer science, biotechnology, psychology, and cognitive science and so on. Following the trend, the research in the field of emotion detection from textual data emerged to determine human emotions from another point of view. Problem of emotion recognition from text can be formulated as follows: Let E be the set of all emotions, A be the set of all authors, and let T be the set of all possible representations of emotion-expressing texts. Let r be a function to reflect emotion e of author a from text t, i.e., r: A x T ->E, then the function r would be the answer to the problem.  The main problem of emotion recognition systems lies in fact that, although the definitions of E and T may be straightforward, the definitions of individual element, even subsets in both sets of E and T would be rather confusing. On one side, for the set T, new elements may add in as the languages are constantly emerging. Whereas on the other side, currently there are no standard classifications of “all human emotions” due to the complex nature of human minds, and any emotion classifications can only be seen as “labels” annotated afterwards for different purposes.  Methods used for text based emotion recognition system are :   1. Keyword Spotting Technique 2. Lexical Affinity Method 3. Learning-based Methods 4. Hybrid Methods   Keyword Spotting Technique: The keyword pattern matching problem can be described as the problem of finding occurrences of keywords from a given set as substrings in a given string [4]. This problem has been studied in the past and algorithms have been suggested for solving it. In the context of emotion detection this method is based on certain predefined keywords. These words are classified into categories such as disgusted, sad, happy, angry, fearful, surprised etc. Process of Keyword spotting method is shown in the figure    Keyword spotting technique for emotion recognition consists of five steps shown in fig where a text document is taken as input and output is generated as an emotion class. At the very first step text data is converted into tokens, from these tokens emotion words are identified and detected. Initially this technique will take some text as input and in next step we perform tokenization to the input text. Words related to emotions will be identified in the next step afterwards analysis of the intensity of emotion words will be performed. Sentence is checked whether negation is involved in it or not then finally an emotion class will be found as the required output.  Lexical Affinity Method : Detecting emotions based on related keywords is an easy to use and straightforward method. Lexical Affinity approach is an extension of keyword spotting technique; it assigns a probabilistic „affinity‟ for a particular emotion to arbitrary words apart from picking up emotional keywords. These probabilities are often part of linguistic corpora, but have disadvantages; firstly the assigned probabilities are biased toward corpus-specific genre of texts, secondly it misses out emotional content that resides deeper than the word-level on which this technique operates. For example the word „accident‟, having been assigned a high probability of indicating a negative emotion, would not contribute correctly to the emotional assessment of phrases like „I avoided an accident‟ or „I met my girlfriend by accident‟.  Learning-based Methods: Learning-based methods are being used to formulate the problem differently. Originally the problem was to determine emotions from input texts but now the problem is to classify the input texts into different emotions. Unlike keyword-based detection methods, learning-based methods try to detect emotions based on a previously trained classifier, which apply various theories of machine learning such as support vector machines and conditional random fields , to determine which emotion category should the input text belongs.  onditional random fields , to determine which emotion category should the input text belongs.  Hybrid Methods: Since keyword-based methods with thesaurus and naïve learning-based methods could not acquire satisfactory results, some systems use hybrid approach by combining both keyword spotting technique and learning based method, which help to improve accuracy. The most significant hybrid system so far is the work of Wu, Chuang and Lin [11], that utilizes a rulebased approach to extract semantics related to specific emotions and Chinese lexicon ontology to extract attributes. These semantics and attributes are associated with emotions in the form of emotion association rules. As a result, these emotion association rules, replacing original emotion keywords, serve as the training features of their learning module based on separable mixture models. This method outperforms previous approaches, but categories of emotions are still limited.  From above discussion there are few limitations:  Ambiguity in Keyword Definitions: Using emotion keywords is a straightforward way to detect associated emotions, the meanings of keywords could be multiple and vague, as most words could change their meanings according to different usages and contexts. Moreover, even the minimum set of emotion labels (without all their synonyms) could have different emotions in some extreme cases such as ironic or cynical sentences.  Incapability of Recognizing Sentences without Keywords: Keyword-based approach is totally based on the set of emotion keywords. Therefore, sentences without any keyword would imply that they do not contain any emotion at all, which is obviously wrong. For example, “I passed my qualify exam today” and “Hooray! I passed my qualify exam today” should imply the same emotion (joy), but the former without “hooray” could remain undetected if “hooray” is the only keyword to detect this emotion.  Lack of Linguistic Information: Syntax structures and semantics also have influences on expressed emotions. For example, “I laughed at him” and “He laughed at me” would suggest different emotions from the first person‟s perspective. As a result, ignoring linguistic information also poses a problem to keyword-based methods.  Difficulties in Determining Emotion Indicators: Learning-based methods can automatically determine the probabilities between features and emotions but the methods still need keywords, but in the form of features. The most intuitive features may be emoticons which can be seen as author‟s emotion annotations in the texts. The cascading problems would be the same as those in keyword-based methods  Reference Links: <https://www.youtube.com/watch?v=DzAZ2Bq_dgU&t=307s>  <https://towardsdatascience.com/>  <https://ieeexplore.ieee.org/Xplore/home.jsp>  Also read some ieee research papers. | | |