

Text Based Emotion Detection Web Application

P.V. Akash Varun(RA1911003010485) & P. Trishwanth Rohit(RA1911003010482)

Department of Computer Science and Technology, SRM University, Chennai

Abstract:

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 Machine Learning and Natural Language Processing. In this paper text based emotion detection and the techniques used in emotion detection are discussed.

Keywords:

ML,NLP,ED,TF-IDF

Introduction:

Emotion detection (ED) is a branch of sentiment analysis that deals with emotion extraction and analysis. Sentiment Analysis aims to detect positive, neutral, or negative feelings from text, whereas Emotion Analysis aims to detect and recognize types of feelings through the expression of texts, such as anger, neutral, fear, happiness, sadness, and surprise. Emotion Detection in text documents is essentially a content – based classification problem involving concepts from the domains of Natural Language Processing as well as Machine Learning. Emotion Recognition is an important area of work to improve the interaction between human and machine. Complexity of emotion makes the acquisition task more difficult. Recent works with deep learning technique has been performed with different kinds of input of human behavior such as audio-visual inputs, facial expressions, body gestures, EEG signal and related brainwaves. Still many aspects in this area to work on to improve and make a robust system will detect and classify emotions more accurately. In this Project, we tried to explore the relevant significant works, their techniques, and the effectiveness of the methods to detect the emotion in text's and make it into a web app with a scope of the improvement of the result. Emotion detection and recognition from text is a recent research area closely related to sentiment analysis. Sentiment analysis aims to detect positive, neutral, or negative emotions from text.

whereas emotion analysis aims to detect and identify types of emotions through the expression of text, such as purpose, Anger, disgust, fear, happiness, sadness, surprise. Emotion classification attempts to detect the emotional content in the input text and based on different approaches establish what kind of emotional content is present if any. Textual emotion classification relies mainly on linguistic resources, and it introduces many challenges to the assignment of text to emotion represented by a proper model. A crucial part of each emotion detector is the emotion model. Emotion recognition has useful applications. To measure citizen well-being. Different indexes have different definitions. Most develop around economic, environmental, health, and social factors.

Literature Review:

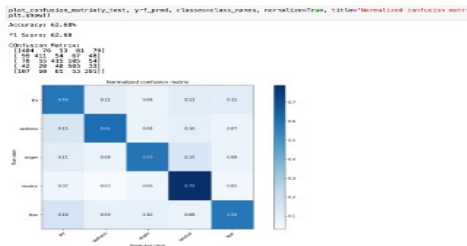
Emotion Recognition is an important area of work to improve the interaction between human and machine. Complexity of emotion makes the acquisition task more difficult. Moreover, deep learning technique with neural network extended the success ratio of machine in respect of emotion recognition. Recent works with deep learning technique has been performed with different kinds of input of human behavior such as audio-

visual inputs, facial expressions, body gestures, EEG signal and related brainwaves. Still many aspects in this area to work on to improve and make a robust system will detect and classify emotions more accurately. In this Project, we tried to explore the relevant significant works, their techniques, and the effectiveness of the methods to detect the emotion in text's and make it into a web app with a scope of the improvement of the result. Emotions are inseparable from human life. These emotions influence human decision-making and help us improve our communication with the world. Since the mid-2000s, governments and organizations around the world have turned their attention to the Happiness Index. Happy Planet Index (HPI) (News, TED Talk). This indicator is defined as an overall index score that ranks countries based on their efficiency and number of happy lives per unit of environmental production. This is unusual as most indices are based on economic indicators. indicators of social welfare. British government measures people's well-being. Other countries and cities such as Seattle, Dubai and South Korea have similar measures. Pervasive computing to better serve individuals. This includes suggesting help when fear is detected in your voice and checking your tone before sending an email. Understand your consumer. Improve customer perception with the ultimate goal of improving brand reputation and sales.

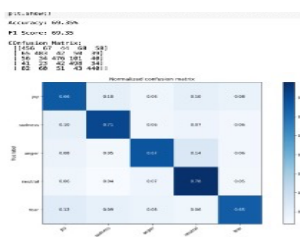
Proposed Methodology :

We've decided to work traditional ML algorithms. After saving our ML model we will then be using Streamlit for connecting with the front-end. As We've decided to go with python as our primarily language. We imported the required Libraries. We have initially done Text Pre-processing such as stopword removal, removing punctuation and removing noise such as html markups, urls. Vectorizing text using Term Frequency technique (Term Frequency(TF) — Inverse Dense Frequency(IDF)), Basically, Text Vectorization is the process of converting text into numerical

representation. We worked on multiple Classifiers: Naive Bayes, Linear Regression, Random Forest, SVM.



We ran our dataset to multiple models. After working on 4 models, We've got to know Logistic Regression model is giving us best accuracy of all. Then We Saved the tf-idf + SVM Model.



Logistic regression is an example of supervised learning. It is used to calculate or predict the probability of a binary (yes/no) event occurring In [linear regression](#), the outcome is continuous and can be any possible value. However in the case of logistic regression, the predicted outcome is discrete and restricted to a limited number of values.

Mathematics behind logistic regression

Probability always ranges between 0 (does not happen) and 1 (happens). Using our Covid-19 example, in the case of binary classification, the probability of testing positive and not testing positive will sum up to 1. We use [logistic function](#) or [sigmoid function](#) to calculate probability in logistic regression. The logistic function is a simple S-shaped curve used to convert data into a value between 0 and 1.

$$h\theta(x) = 1 / 1 + e^{-(\beta_0 + \beta_1 X)}$$

' $h\theta(x)$ ' is output of logistic function , where $0 \leq h\theta(x) \leq 1$

' β_1 ' is the slope

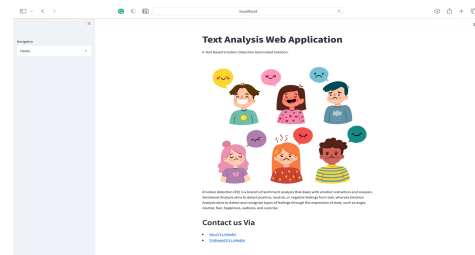
' β_0 ' is the y-intercept

' X ' is the independent variable

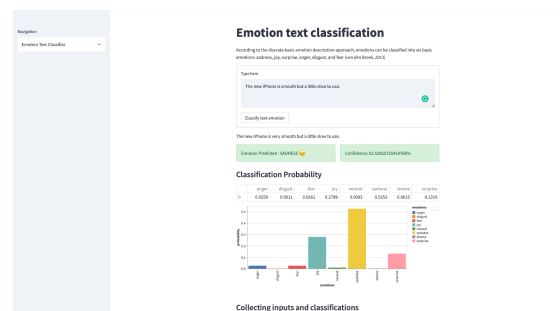
$(\beta_0 + \beta_1 x)$ - derived from equation of a line $Y(\text{predicted}) = (\beta_0 + \beta_1 x) + \text{Error value}$

We are using a fast and lightweight python framework Streamlit for the front-end. We tend to intend to create a simple yet full featured web application. This web application will not only detect or classify the emotions it will collect the inputs and classifications and give an option to download a CSV file. This app also provides Classification probability bar graph. We also have emotion occurrence (The no of times an emotion is classified) and the evolution of the emotion on a day to day basis helpful for exploratory data analysis. To simplify the extra mathematical calculations we provided the percentage of each occurring emotion as pie chart.

Results and Explanation:



This is Home page of our application. It gives a simple overview of what this web application. It describes a simple definition of what text analysis is. This web app is done with the help of python framework called Streamlit.



Once you navigate to Emotion Text Classifier, You will be given a text box to type your given input. Once you have given your input. The code will then take the input and execute the program with the help of saved model and then provide the predicted output. As you can see in the above figure it will provide us the predicted output and confidence of it's prediction along with classification probability with a bar graph.



This web application will not only detect or classify the emotions it will collect the inputs and classifications and give an option to download a CSV file. This app also provides Classification probability bar graph. We also have emotion occurrence (The no of times an emotion is classified) and the evolution of the emotion on a day to day basis helpful for exploratory data analysis. To simplify the extra mathematical calculations we provided the percentage of each occurring emotion as pie chart.

Conclusion:

The Current Paper proposes a text analysis web application using Python, NLP and other frameworks. Previous works provide only the backend of this current end product and most of them detect emotion only via text's and doesn't consider the emoji's on the text. This application provides classify the emotions it will collect the inputs and classifications and give an option to download a CSV file. This app also provides Classification probability bar graph. We also have emotion occurrence (The no of times an emotion is classified) and the evolution of the emotion on a day to day basis helpful for exploratory data analysis. To simplify the extra mathematical calculations we provided the percentage of each occurring emotion as pie chart.

References:

1. Sklearn Documentation: https://scikit-learn.org/stable/getting_started.html
2. Streamlit Documentation: <https://docs.streamlit.io>
3. https://www.researchgate.net/publication/225045375_Emotion_Detection_from_Text
4. Chun-Chieh Liu, Ting-Hao Yang, Chang-Tai Hsieh, Von-Wun Soo, "Towards Text-based Emotion Management and Engineering, 2009.
5. <https://www.wikipedia.org>