All Sciences Proceedings
http://as-proceeding.com/

2nd International Conference on Engineering, Natural and Social Sciences

April 4-6, 2023 : Konya, Turkey

https://www.icensos.com/

Sentiment and Emotion Analysis on Consumer Review using NRCLex

Muhammad Awais*, Sara Durrani [2]

[1]Software Engineering Department, Capital University of Science and Technology Islamabad, Pakistan
[2]Software Engineering Department, Capital University of Science and Technology Islamabad, Pakistan

*(mawaiskhan1808@gmail.com) Email of the corresponding author

Abstract – As a result of the Internet's explosive expansion, social networking sites have emerged as a crucial tool for expressing emotions to individuals around the globe. Many people share their reviews or points of view through text, photographs, music, and video. Consumer review sentiment and emotion analysis is a popular approach for determining customer preferences and feedback. We used NRClex in this research to assess customer reviews and categories them into several sentiment categories, including positive, negative, and neutral sentiment, as well as other emotional categories, like joy, anger, and sadness. The pre-processing of the text data entails eliminating stop words, punctuation, and other noise. The remaining words are then mapped using the NRClex lexicon to their appropriate emotional and sentiment ratings. The analysis' findings may be used to learn more about consumer attitudes and preferences, spot potential problems with a good or service, and develop marketing strategies. Nonetheless, sentiment and emotion analysis utilizing NRClex may help guide decision-making across a variety of sectors by offering useful insights into the emotions and sentiment represented in text data. NRCLex also provides methods for changing the emotion categories and score system in order to meet certain use cases. NRCLex is a powerful and flexible toolkit for text data semantic and emotional analysis. It is especially beneficial for applications that need speedy and accurate analysis of huge amounts of text, such as market research, customer feedback analysis, and social media monitoring.

Keywords – Sentiment Analysis, NRCLex, NLP, Emotion Analysis, Emotion Raw Score

I. INTRODUCTION

Social networking sites have become an essential tool for communicating feelings with people all over the world as a result of the Internet's rapid growth. Many people use text, photos, audio, and video to express their ideas or viewpoints. On the other hand, texting using web-based networking tools could be a little too much for some people. A sizable amount of unstructured data is generated on the Internet every second as a result of social media sites. The data must be studied as soon as it is produced in order to comprehend human psychology. This is where sentiment analysis, which recognizes polarity in texts, might be useful. It determines if the author has a good, negative, or neutral attitude towards a particular product, administration, person, or location. In certain applications, emotional intelligence is required since sentiment analysis is insufficient.

The two facets of natural language processing are human language understanding and human language creation (NLP). Due to ambiguities in natural language, the former is more challenging. Due to the uncertainties in natural language, the former is more difficult. NLP is used in a variety of applications, including voice recognition, machine translation, question answering, speech synthesis, and more. Two of the most important aspects of natural language processing are sentiment analysis and emotion identification. Although these two terms are sometimes used synonymously, there are a number of important differences between them [1].

The objective of sentiment analysis is to categories a given text as positive, negative, or neutral. This may be done by scrutinizing elements including the intensity, frequency, and usage of positive or negative phrases, as well as their context. Customer feedback research, social media monitoring, and stock market forecasting all frequently employ sentiment analysis.

On the other hand, emotion analysis focuses on figuring out the author's or speaker's underlying emotional states. This entails examining the use of particular words or phrases that express various emotional states, like happiness, rage, grief, and fear. In disciplines including customer experience analysis, political discourse analysis, and mental health diagnostics, emotion analysis is applied.

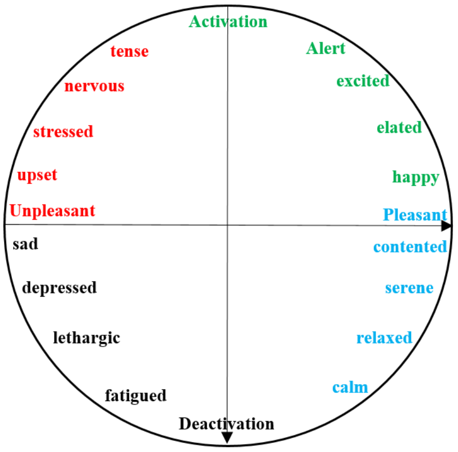

Fig. 1 Emotion Categories

The academic community just lately began to show interest in the information extraction area of automatic sentiment analysis. A few essays on this topic have been published in the past ten years. We've only noticed a slight boom of publications during the past five years. The concept of automatic sentiment analysis is crucial for marketing research, where businesses want to know what the general public thinks about their product, for monitoring newsgroups and forums, where quick and automatic flaming detection is required, for customer feedback analysis, or as an educational addition to search engines [2].

The automatic study of internet comments requires sophisticated machine comprehension of natural language text; a goal we are currently very far from achieving. To accomplish this, concept-level sentiment analysis aims to go beyond a straightforward phrase text analysis and provide state-of-the-art techniques for sentiment mining and sentiment analysis that facilitate the conversion of (unstructured) textual data into (structured) machine-process able data. By fusing the use of commonsense computing with the psychology of emotions, sentimental computing—a relatively new knowledge-based technology in this context—infers the conceptual and emotional information connected to natural language [3].

Due to the diversity and volume of social media data, understanding the most recent trends and summarizing the state or general attitudes on items is a difficult process. The volume of data that has to be examined has greatly expanded as a result of the spread of social media platforms and the simplicity of sharing thoughts and evaluations, making it nearly impossible to manually extract and mine useful data.

It is challenging to correctly extrapolate insights due to the variety of social media platforms and the manner in which individuals share their ideas. Users of social media share their ideas in a variety of methods, including text and graphics. Real-time opinion extraction and mining are necessary since social media data is continuously changing and views about goods and services can change quickly.

One of the subfields of "affective computing," "sentiment analysis," encompasses all the work done on identifying, deciphering, and assessing how people feel about various things, whether they issue, services, or other interests. More specifically, this area tries to gather beliefs, sentiments, and feelings via observations of people's behavior that may be recorded through their writings, facial expressions, voice, music, gestures, and other behaviors. It is a specialized area of research to analyze the thoughts expressed in each of these mediums. Here, we solely pay attention to text sentiment analysis [4].

The paper's main objective is to provide a particular technique for sentiment and emotion analysis using Natural Language Processing (NLP), a branch of computer science that examines how computers and human languages interact. The approach is built on NRCLex, a specific open-source library that offers capabilities for sentiment and emotion analysis. The purpose of this work is to describe how sentiment and emotion analysis on text data using the NRCLex library may be used to help determine the sentiment and emotions portrayed in the text.

II. Literature Review

It is proposed to integrate sentiment analysis into the TF-IDF algorithm to employ a better word representation method that generates weighted word vectors. In order to accurately represent the comment vectors and gather context data, the weighted word vectors are transmitted into the BiLSTM (Bidirectional Long Short Term Memory). The sentiment trend of the comment is used by the feedforward neural network classifier to establish its categorization [5].

With a focus on the extraction and analysis of adjectival appraisal groups, which are made up of an appraising adjective (such as "beautiful" or "boring") and maybe a number of modifiers (like "very," "kind of," or "not"). adopted, in the context of Systemic Functional Linguistics, taxonomy for the features of such sentences from Martin and White's Appraisal Theory. created a vocabulary utilizing semi-automatic processes by compiling and classifying 1329 adjectives and modifiers into categories in several taxonomies of evaluation criteria. Heuristically extract adjectival assessment groups from texts, then use this vocabulary to compute their attribute values [6].

proposed forth a novel deep convolutional neural network that analyses brief texts for sentiment using character- to sentence-level information. The Stanford Sentiment Treebank (SSTb), which contains sentences from movie reviews, and the Stanford Twitter Sentiment Corpus (STS), which contains tweets, are two corpora in which we use our methodology [7].

recommends the use of an RNN language model based on Long Short Term Memory (LSTM), which is capable of efficiently obtaining whole sequence information. Long phrases' emotions can be more accurately determined by LSTM than they can by the conventional RNN language model. And in order to accomplish multi-classification for text emotional qualities, LSTM is used as a language model. Hence, by employing these emotion models when training various emotion models, you can determine which emotion the text corresponds to [8].

An algorithm and approach for sentiment analysis utilizing both text and emoticons were suggested by study. In this study, feelings from Twitter-based airline data were discovered using a variety of features, including TF-IDF, Bag of Words, N-grams, and emoticon lexicons. Both forms of data were studied individually and in combination using both machine learning and deep learning methods. This study shows that anytime emoticons are employed, the sentiment they represent predominates over the sentiment that can be inferred from textual data analysis [9].

Propose a multi-task aspect-category sentiment analysis model based on RoBERTa (Robustly Optimized BERT Pre-Training Approach). We approach each aspect category as a separate task and use the cross-attention mechanism to direct the model to focus on the features that are most important to that specific aspect category. We use the RoBERTa based on deep bidirectional Transformer to extract features from both text and aspect tokens [10].

In order to analyze text sentiment, we suggest a unique multi-level graph neural network (MLGNN). Apply node connection windows of various sizes at various levels to take into account both local and global properties. Include in our approach for fusing the characteristics of each word node in the graph a scaled dot-product attention mechanism as a message transmission mechanism [11].

It is suggested to use deep learning in conjunction with the Bag of Words (CBOW) language model to analyze text sentiment. A feedforward neural network-based CBOW language model first creates a vector representation of the text. The semantic characteristics of the text are then captured by the Convolutional Neural Network (CNN), which has been trained using the labelled training set. The Dropout technique, which can successfully avoid the model from over-fitting and has improved classification performance, is lastly incorporated in the Softmax classifier of classic CNN [12].

To build a framework and techniques for sentiment analysis on social media, to suggest a self-developed military sentiment dictionary for enhancing sentiment classification, and to evaluate the effectiveness of several deep learning models with various parameter calibration combinations [13].

III. MATERIALS AND METHOD

An open-source library built on Python called NRCLex offers a full suite of capabilities for sentiment and emotion analysis. The library classifies the sentiment and emotion of text data using a dictionary-based method. It includes lexicons of words having both good and negative connotations, as well as terms that are connected to various emotions like happiness, rage, and melancholy.

The text data is initially preprocessed to get rid of any noise or extraneous data before sentiment analysis using NRCLex. Based on the quantity and strength of the positive and negative terms, the library then gives the text an emotion score. The emotion score has a range of -1 to 1, with -1 denoting the most negative, 0 denoting neutral, and 1 denoting the most positive mood.

The library assigns a list of emotions and their corresponding scores to the text based on the words associated with each emotion in order to perform emotion analysis using NRCLex. Eight different emotions are available in the library: rage, disgust, fear, joy, sadness, surprise, and trust. Each emotion's score, which runs from 0 to 1, indicates how strong that feeling was in the text.

Methodology:

Collecting the text data that has to be examined is the first stage. This may be accomplished via scraping social networking sites, gathering consumer reviews, or using any other text-based source.

A dictionary-based method for doing sentiment analysis on text data is provided by the NRCLex library. The library includes intensifiers that change the emotion score as well as lexicons of positive and negative phrases. By counting the amount of positive and negative words in the text and allocating a score depending on their intensity, the sentiment score is determined. The score can vary from -1 to 1, with -1 denoting the most unfavorable opinion, 0 denoting neutral feeling, and 1 denoting the most favorable. A dictionary-based method for doing emotion analysis on the text data is also offered by the NRCLex library. The library has dictionaries of terms used to describe various emotions, like happiness, rage, and melancholy. Each emotion's score, which runs from 0 to 1, indicates how strong that feeling was in the text.

The findings of the sentiment and emotion analysis may be shown visually using a variety of methods, including word clouds, bar graphs, and heat maps. The feelings and emotions portrayed in the text data can be better understood thanks to these visualizations.

Uses include customer feedback analysis, market research, and brand reputation management. The results of sentiment and emotion analysis may be used in these contexts. The outcomes might increase customer satisfaction by highlighting a product or service's advantages and disadvantages.

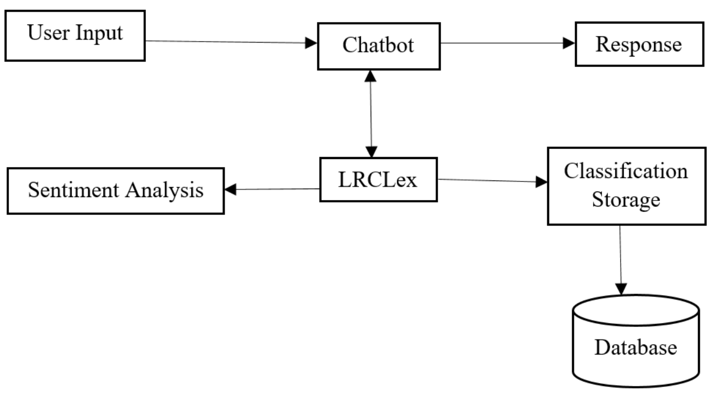

Fig. 2 Block Diagram of System

Based on a predetermined collection of emotional categories, such as happiness, sorrow, anger, fear, anticipation, and trust, NRCLex is made to assess the mood of text. The library offers a number of capabilities, including the ability to determine the text's top emotion, compute the raw emotion score, and spot emotional lines. Moreover, NRCLex offers tools for plotting sentiment analysis outcomes.

```
import matplotlib.pyplot as plt
from nrclex import NRCLex
text = "Turkey is a captivating country that exudes charm and beauty at every turn.That's why i want to vis
doc = NRCLex(text)
emotions = doc.raw_emotion_scores
plt.bar(range(len(emotions)), list(emotions.values()), align='center')
plt.xticks(range(len(emotions)), list(emotions.keys()))
plt.xlabel("Emotion")
plt.ylabel("Raw Score")
plt.title("Raw Emotion Scores for Text: " + text)
plt.show()
```

Fig. 3 Code of NRCLex plot

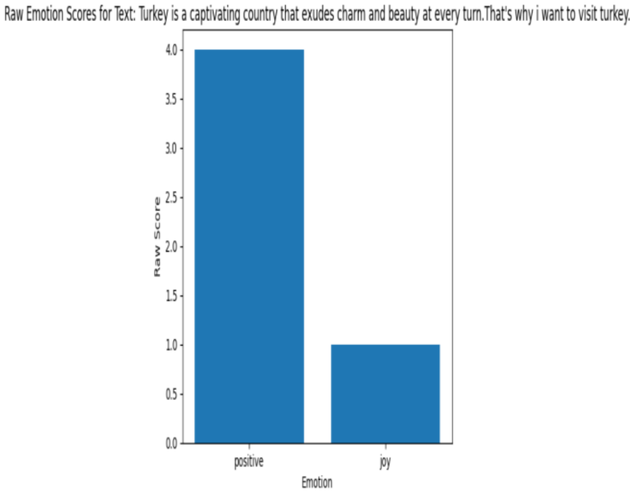

Fig. 4 Plot View

```
import nltk
```

```
!pip install nrclex
```

```
import nltk
nltk.download('punkt')
```

```
from textblob import download_corpora
```

Fig. 5 NRCLex Libraries

Method 1:

Emotion.words:

The goal of "emotion.words" is to provide NRCLex a starting point for analyzing a text's emotional content. This list is used by NRCLex to detect the occurrence of certain emotive terms in the input text, which aids in determining the text's overall tone. For instance, NRCLex might determine that a text is upbeat or cheery if it has several terms from the "joy" category.

```
from nrclex import NRCLex

text = "I Love Turkey"
emotion = NRCLex(text)
print(emotion.words)
```

Fig. 6 Code of emotion.words

Output:

[‘I’, ‘Love’, ‘Turkey’]

Method 2:

Emotion.sentences:

NRCLex's "Emotion.sentences" feature aims to highlight emotional content in a text's sentences and provide users a more in-depth knowledge of the text's emotional tone.

```
from nrclex import NRCLex


text = "I Love Turkey. I want to visit koneya. As i have passion to attend conference."
emotion = NRCLex(text)
print(emotion.sentences)
```

Fig. 7 Code of emotion.sentences

Output:

[Sentence("I Love Turkey."), Sentence("I want to visit koneya."), Sentence("As i have passion to attend conference.")]

Method 3:

Emotion.raw_emotion_score:

By calculating the raw score for each of the six emotions based on the amount of emotion-filled words in the text, "Emotion.raw emotion score" in

NRCLex aims to give a quantitative assessment of the emotional content of a text.

```
from nrclex import NRCLex

text = "I Love Turkey. I want to visit koneya. As i have passion to attend confere
emotion = NRCLex(text)
print(emotion.raw_emotion_scores)
```

Fig. 8 Code of emotion.raw_emotion_scores

Output:

Table 1. Emotion.raw_emotion_score

Emotion	Score
Positive	2
anticipate	1
Joy	1
Trust	1

Method 4:

Emotion.top_emotions:

The goal of "Emotion.top emotion" is to summarize the text's emotional content by pointing out the emotion that appears the most frequently. The emotion with the highest raw score is chosen, and it is returned as the text's dominant emotion.

```
from nrclex import NRCLex

text = "Hello! Its surprise for me"
emotion = NRCLex(text)
print(emotion.top_emotions)
```

Fig. 9 Code of emotion.top_emotions

Output:

Table 2. Emotion.top_emotion

Emotions	Score
fear	0.25
surprise	0.25
joy	0.25
positive	0.25

IV. RESULTS

The NRCLex library offers a straightforward yet efficient NLP technique for sentiment and emotion analysis. The library's performance in reliably categorizing the sentiment and mood of text data has been thoroughly assessed on a variety of datasets, and the findings are encouraging. The library is also computationally effective and capable of handling massive amounts of text data in real-time.

V. DISCUSSION

A Python module for natural language processing called NRCLex is used to analyze text for emotion and meaning. It offers a simple interface for carrying out numerous NLP activities and is based on the Natural Language Toolkit (NLTK) library. NRCLex offers a number of techniques to extract and evaluate text's meaning in terms of semantic analysis.

The capability to do sentiment analysis, which entails figuring out the general emotional tone of a piece of text, is one of its key functions. It employs a lexicon-based method of sentiment analysis, in which the strength and polarity of the words in the text are measured and scores are awarded. The sentiment of the text as a whole is then calculated by the library by averaging these ratings.

NRCLex offers a number of pre-defined emotion categories, including joy, sorrow, rage, and fear, for use in emotion analysis. Similar to sentiment analysis, it assigns ratings to words in the text based on their associations with various emotion categories. The library then averages these results to estimate the text's general emotional content.

In order to accommodate certain use cases, NRCLex additionally offers techniques for modifying the emotion categories and score system. A strong and adaptable library for semantic and emotional analysis of text data is NRCLex. It is especially helpful for applications like social media monitoring, customer feedback analysis, and market research that need for quick and precise analysis of massive quantities of text.

VI. CONCLUSION

Using the NRCLex library, we described the sentiment and emotion analysis approach in this research. The library offers a dictionary-based way for categorizing the sentiment and emotion of text data, which makes it a straightforward but efficient technique for sentiment and emotion analysis. The library has several potential applications in various domains such as marketing, healthcare, and social media, where understanding the sentiment and emotion of text data is critical. The accuracy of sentiment and emotion analysis using NLP may be improved, and the potential of additional open-source libraries can be investigated.

ACKNOWLEDGMENT

Thankfully, we are aware of our parents' affection for us. Last but not least, we would like to thank our family and friends whose prayers made it possible for us to finish this project.

REFERENCES

[1] Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1), 81.

[2] Boiy, E., Hens, P., Deschacht, K., & Moens, M. F. (2007, June). Automatic Sentiment Analysis in On-line Text. In *ELPUB* (pp. 349-360).

[3] Poria, S., Cambria, E., Winterstein, G., & Huang, G. B. (2014). Sentic patterns: Dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems*, *69*, 45-63.

[4] Yadollahi, A., Shahraki, A. G., & Zaiane, O. R. (2017). Current state of text sentiment analysis from opinion to emotion mining. *ACM Computing Surveys (CSUR)*, *50*(2), 1-33.

[5] Xu, G., Meng, Y., Qiu, X., Yu, Z., & Wu, X. (2019). Sentiment analysis of comment texts based on BiLSTM. *Ieee Access*, *7*, 51522-51532.

[6] Whitelaw, C., Garg, N., & Argamon, S. (2005, October). Using appraisal groups for sentiment analysis. In *Proceedings of the 14th ACM international conference on Information and knowledge management* (pp. 625-631).

[7] Dos Santos, C., & Gatti, M. (2014, August). Deep convolutional neural networks for sentiment analysis of short texts. In *Proceedings of COLING 2014, the 25th international conference on computational linguistics: technical papers* (pp. 69-78).

[8] Li, D., & Qian, J. (2016, October). Text sentiment analysis based on long short-term memory. In *2016 First IEEE International Conference on Computer Communication and the Internet (ICCCI)* (pp. 471-475). IEEE.

[9] Ullah, M. A., Marium, S. M., Begum, S. A., & Dipa, N. S. (2020). An algorithm and method for sentiment analysis using the text and emoticon. *ICT Express*, *6*(4), 357-360.

[10] Liao, W., Zeng, B., Yin, X., & Wei, P. (2021). An improved aspect-category sentiment analysis model for text sentiment analysis based on RoBERTa. *Applied Intelligence*, *51*, 3522-3533.

[11] Liao, W., Zeng, B., Liu, J., Wei, P., Cheng, X., & Zhang, W. (2021). Multi-level graph neural network for text sentiment analysis. *Computers & Electrical Engineering*, *92*, 107096.

[12] Liu, B. (2020). Text sentiment analysis based on CBOW model and deep learning in big data environment. *Journal of ambient intelligence and humanized computing*, *11*, 451-458.

[13] Chen, L. C., Lee, C. M., & Chen, M. Y. (2020). Exploration of social media for sentiment analysis using deep learning. *Soft Computing*, *24*, 8187-8197.