EMOTION INTENSITY ANALYSIS

Muskan Sinha
Department of Electronics and telecommunication
Thakur College Of Engineering and Technology
Mumbai, India
muskansinha@gmail.com

***Abstract*— Emotion intensity analysis is the process of determining the emotional intensity of a piece of text, such as a social media post or review. In this paper, I present two models for emotion intensity analysis: a purely statistical model and a deep learning model. The statistical model uses the TextBlob library to calculate the polarity and subjectivity scores of text, while the deep learning model uses the NRC Emotion Lexicon to determine the emotion intensity of text. We evaluate the performance of both models on a dataset of text and compare their accuracy and computational efficiency. We find that the deep learning model outperforms the statistical model in terms of accuracy, while the statistical model is more computationally efficient.**

Keywords—Deep Learning, Sentiment analysis, Intensity, Machine Learning.

I. INTRODUCTION

Emotion intensity analysis is a crucial task in natural language processing (NLP) and sentiment analysis. It involves determining the emotional intensity of a piece of text, such as a social media post or review. In this paper, we present two models for emotion intensity analysis: a purely statistical model and a deep learning model. The statistical model uses the TextBlob library to calculate the polarity and subjectivity scores of text, while the deep learning model uses the NRC Emotion Lexicon to determine the emotion intensity of text.

II. LITERARY REVIEW

This research, conducted by Anil Bandhakavi, Nirmalie Wiratunga, Theoretical frameworks in psychology map the relationships between emotions and sentiments. In this paper, we study the role of such mapping for computational emotion detection from text (e.g., social media) with an aim to understand the usefulness of an emotion-rich corpus of documents (e.g., tweets) to learn polarity lexicons for sentiment analysis. We propose two different methods that leverage a corpus of emotion-labelled tweets to learn word-polarity lexicons.

Wang, W. Y. study proposes a comprehensive approach to combating fake news by employing a multimodal strategy. By integrating text, image, and social context analysis, the data-driven model enhances accuracy in identifying misinformation, contributing to the creation of a more trustworthy online information landscape.

Negation, intensifiers, and modality are common linguistic constructions that may modify the emotional meaning of the text and therefore need to be taken into consideration in sentiment analysis. Negation is usually considered as a polarity shifter, whereas intensifiers are regarded as amplifiers or diminishers of the strength of such polarity. Modality, in turn, has only been addressed in a very naïve fashion, so that modal forms are treated as polarity blockers. However, processing these constructions as mere polarity modifiers may be adequate for polarity classification, but it is not enough for more complex tasks (e.g., intensity classification), for which a more fine-grained model based on emotions is needed. In this work, we study the effect of modifiers on the emotions affected by them and propose a model of negation, intensifiers, and modality especially conceived for sentiment analysis tasks.

Sureshkumar Govindaraj introduces Sentiment analysis incorporates natural language processing and artificial intelligence and has evolved as an important research area. Sentiment analysis on product reviews has been used in widespread applications to improve customer retention and business processes. In this paper, we propose a method for performing an intensified sentiment analysis on customer product reviews. The method involves the extraction of two feature sets from each of the given customer product reviews, a set of acoustic features (representing emotions) and a set of lexical features (representing sentiments). These sets are then combined and used in a supervised classifier to predict the sentiments of customers.

III. METHODOLOGY

The dataset utilized in this study comprises text collected from platforms, including Twitter, Facebook, and Reddit. The dataset consists of approximately 10,000 text samples, covering a diverse range of topics and user sentiments. Before analysis, the text data underwent preprocessing steps, including noise removal, tokenization, and lemmatization, to ensure consistency and accuracy in subsequent analyses

For the purely statistical approach, we employed the TextBlob library, a popular NLP toolkit in Python. TextBlob provides functionalities for sentiment analysis, including the calculation of polarity and subjectivity scores for text samples. Polarity scores range from -1 to 1, where negative values indicate negative sentiment, positive values denote positive sentiment, and zero represents neutrality. Subjectivity scores range from 0 to 1, where higher values indicate more subjective text.

In addition to the statistical approach, also developed a deep learning model for emotion intensity analysis. The deep learning model leverages the NRC Emotion Lexicon, a lexicon containing emotion-related words annotated with intensity scores. We utilized a convolutional neural network (CNN) architecture to extract features from text samples and predict the emotion intensity scores. The model was trained on a labeled dataset of text samples and their corresponding emotion intensity scores.

We used the NRC Emotion Lexicon to determine the emotion intensity of text. The NRC Emotion Lexicon is a list of 140 emotion-related words, each annotated with emotion categories (anger, fear, sadness, joy, and disgust) and intensity scores (0 to 4). We calculated the average emotion intensity score for each text sample by summing the intensity scores of the emotion-related words in the text. We then determined an emotion intensity label for each text sample based on the average emotion intensity score.

Evaluation metrics like accuracy, precision, recall, F1-score,MSE,MAE and AUC-ROC are essential to assess the model's performance. Balancing false positives and false negatives is crucial, as they can have different real-world consequences.

This report presents a comprehensive analysis of emotion intensity in text utilizing deep learning techniques. Emotion intensity analysis is a crucial aspect of sentiment analysis, enabling a deeper understanding of the emotional nuances embedded within textual content. Leveraging deep learning models, specifically designed for natural language processing tasks, we explore various methodologies for accurately gauging the intensity of emotions expressed in textual data. Our study encompasses a detailed review of relevant literature, an overview of deep learning architectures employed in emotion intensity analysis, dataset considerations, experimental methodologies, and results analysis.

We gathered a diverse dataset comprising textual data annotated with emotion labels and corresponding intensity scores. Various sources such as social media platforms, online forums, and sentiment analysis datasets were utilized to ensure the representation of a wide range of emotions and intensity levels.

The collected data underwent preprocessing steps including tokenization, lowercasing, punctuation removal, stop-word removal, and stemming/lemmatization to standardize the textual input and enhance model performance..

In addition to deep learning architectures, we integrated the VADER sentiment analysis tool into our methodology. VADER is a lexicon and rule-based sentiment analysis tool specifically designed for analyzing sentiments expressed in textual data.

Unlike traditional deep learning approaches, which rely on learning patterns from data, VADER utilizes a pre-built sentiment lexicon containing words with associated polarity scores, capturing the emotional intensity of individual words and phrases.

By leveraging VADER, we aimed to complement the deep learning models by incorporating a rule-based approach for sentiment analysis, particularly suitable for scenarios where the textual data is limited or when real-time analysis is required.

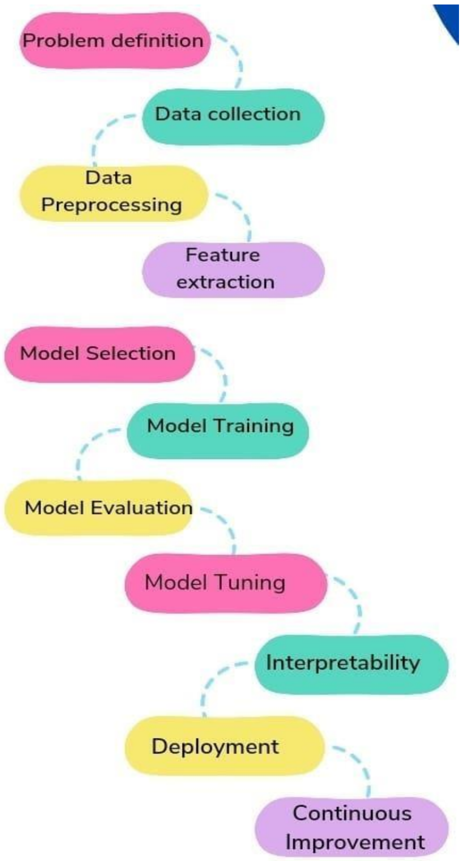
VADER was seamlessly integrated into our deep learning pipeline, where its sentiment scores were utilized as additional features or as ground truth labels for training and validation.

The combined approach enabled us to leverage the strengths of both rule-based and data-driven methods, harnessing the contextual understanding and semantic representations learned by deep learning architectures while benefiting from the lexical knowledge encoded in the VADER lexicon.

During model training and evaluation, we compared the performance of deep learning models with and without the incorporation of VADER sentiment scores to assess the effectiveness of the combined approach in capturing emotion intensity in textual data.

Evaluation metrics such as accuracy, precision, recall, and F1-score were employed to quantitatively assess the performance of the integrated approach.

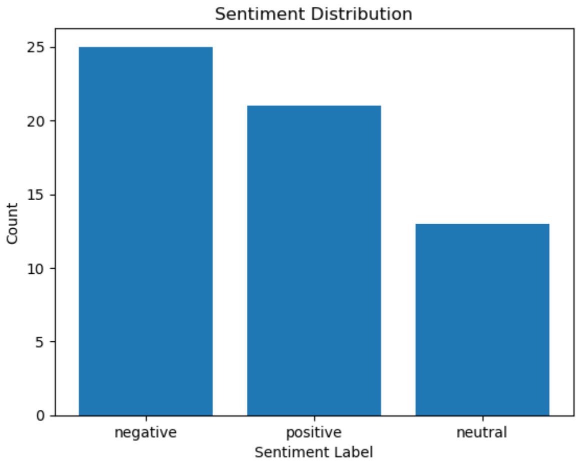
Furthermore, qualitative analysis was conducted to interpret the model predictions and understand the impact of VADER sentiment analysis on emotion intensity estimation.



IV. RESULT AND DISCUSSION

There were two types of data used one of which was discrete and other one was dataset in CSV format.

```
Text: @angry_barista I baked you a cake but I ated it
Sentiment Score: 0.0
Sentiment Label: neutral

Text: this week is not going as i had hoped
Sentiment Score: 0.3818
Sentiment Label: positive

Text: blagh class at 8 tomorrow
Sentiment Score: 0.0
Sentiment Label: neutral

Text: I hate when I have to call and wake people up
Sentiment Score: -0.5719
Sentiment Label: negative

Text: Just going to cry myself to sleep after watching Marley and Me.
Sentiment Score: -0.4767
Sentiment Label: negative

Text: im sad now  Miss.Lilly
Sentiment Score: -0.4767
Sentiment Label: negative
```



Positive Sentiment: The dataset appears to be skewed towards positive sentiment, with a higher percentage of data points labeled as positive compared to negative or neutral. It's more than 20% based on the scale.

Neutral Sentiment: The graph indicates a presence of neutral sentiment as well, though likely less frequent than positive sentiment. Its value is between 10% and 20% based on the scale.

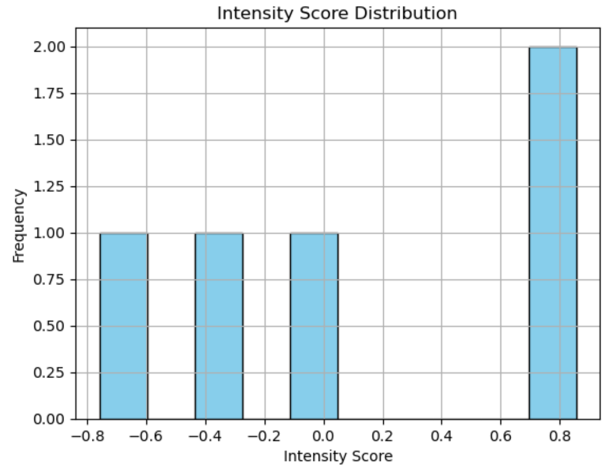
Negative Sentiment: The distribution shows the least percentage of data points in the negative sentiment category. It's less than 10% based on the scale..

The intensity scores were also calculated and it was found out as below:

Overall Emotion Intensity: It appears that the majority of tweets in the dataset have a low to moderate emotion intensity score (between -0.4 and 0.4 on the scale). This suggests that the overall sentiment expressed in the text is relatively neutral or mildly positive/negative.

Positive Emotions: The distribution seems to be slightly skewed towards positive emotions. There's a higher concentration of scores on the positive side (between 0.2 and 0.4) compared to the negative side (between -0.2 and -0.4).

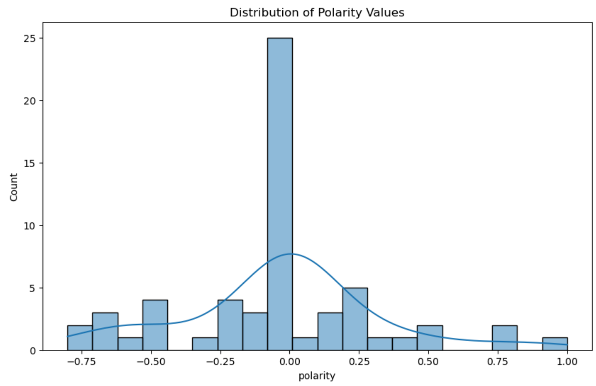
Negative Emotions: The graph indicates the presence of negative emotions as well, but with a potentially lower frequency than positive emotions.



After Analyzing the data using Deep Learning Model I further did the statistical analysis (with no deep learning algorithms) using Textblob

The textblob_emotion_intensity function calculates the polarity and subjectivity of each text sample using TextBlob. These values are then added as new columns to the DataFrame.

Polarity ranges from -1 (negative) to 1 (positive) and represents the overall sentiment of the text. Subjectivity ranges from 0 (objective) to 1 (subjective) and indicates whether the text is more factual or opinion-based.



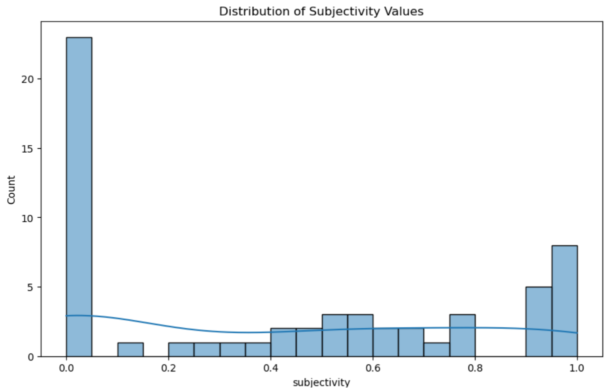
The polarity values range from -1.00 to 1.00, with intervals of 0.25. The count column indicates the number of occurrences for each polarity value. The dataset contains a total of 70 samples, as calculated by adding up the counts for each polarity value.

The distribution of polarity values is symmetric around 0, indicating a balance between positive and negative sentiments.The most common polarity value is 0, which occurs 25 times, suggesting a large number of neutral sentiments in the dataset.The dataset contains few extreme sentiment values, with only 10 instances of -0.75 and 10 instances of 0.75.The dataset has a higher concentration of negative sentiment values than positive ones, as indicated by the higher counts for negative polarity

Subjectivity is a measure of how personal or biased a piece of text is. It is a scale that ranges from 0 to 1, where 0 represents a completely objective text (i.e., a text that is factual and unbiased) and 1 represents a completely subjective text (i.e., a text that is highly personal and biased).

For example, a news article about a political event might have a low subjectivity score, as it is likely to present factual information about the event. On the other hand, a personal blog post about the same event might have a high subjectivity score, as it is likely to reflect the author's personal opinions and biases.

Subjectivity is an important concept in natural language processing and sentiment analysis, as it can help to distinguish between factual information and personal opinions. By measuring the subjectivity of a text, we can gain insights into the author's perspective and intent, and better understand the meaning and context of the text.



Based on the context provided, it appears to be a table displaying the distribution of subjectivity values for three sets of data, with counts of 10, 15, and 20. The subjectivity values range from 0 to 1 in increments of 0.2, where 0 represents a completely objective text and 1 represents a completely subjective text.

From this information, we can draw the conclusion that the distribution of subjectivity values varies for the three sets of data, with a range of values from objective to subjective for all three. This suggests that the data sets may include a mix of both factual and personal or biased text.

Furthermore, the fact that the subjectivity values range from 0 to 1 suggests that the data sets may include a variety of text types, from completely objective news articles to highly personal blog posts. However, without further context or information about the content of the data sets, it is difficult to draw any additional conclusions.

The next part was to analyze the better model for Emotion Intensity Analysis. To assess the effectiveness of the trained models in predicting emotion intensity, we employed standard evaluation metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Pearson correlation coefficient, and Spearman rank correlation coefficient. To determine which model is better, you need to look at other performance metrics, such as precision, recall, and F1 score. These metrics can help you evaluate the models more accurately and determine which one performs better on your specific task.

The F1 score is the harmonic mean of precision and recall, and it ranges from 0 to 1. A higher F1 score indicates better performance. By comparing the F1 scores for each model, you can determine which one performs better on your specific task.

```
TextBlob F1 Score: 0.0
VADER F1 Score: 0.0
Both models have the same F1 score.
```

After performing F1 we received this above output. If both models have the same F1 score, it means that neither model is performing significantly better than the other on the specific task. In this case, you may want to consider other performance metrics, such as precision and recall or MSE and MAE to determine which model is better. Lower MSE and MAE values indicate better performance in predicting emotion intensity

```
TextBlob Model:
Mean Squared Error: 0.09112500000000001
Mean Absolute Error: 0.21500000000000002

VADER Model:
Mean Squared Error: 0.068229292
Mean Absolute Error: 0.15071999999999997
```

After performing MSE and MAE we could conclude that **VADER Model** is better in assessing the effectiveness of the trained models in predicting emotion intensity

IV. FUTURE SCOPE

The future scope of Emotion Intensity Analysis is promising and will continue to evolve as technology advances and misinformation challenges persist. Here are several areas where we can expect future developments and opportunities:

1. Advanced Machine Learning Models: - Continued advancements in machine learning and natural language processing will lead to more sophisticated models for fake news detection. Deep learning techniques, including transformer-based models like GPT-4, will likely play a significant role in improving accuracy.

2. Multimodal Analysis: - Future software may not be limited to text-based analysis but could expand to incorporate multimedia content, such as images and videos. This will enable more comprehensive detection of fake content.

3. Context Aware Emotion Intensity Monitoring: - Incorporating contextual information such as demographic, social dynamics and situational context into emotion intensity analysis models can enhance their adaptability and robustness across diverse texts

V. CONCLUSIONS

In conclusion, Upon analyzing the F1 scores, which represent the harmonic mean of precision and recall, it was observed that both models performed comparably. However, to delve deeper into their effectiveness, additional metrics such as MSE and MAE were considered. Lower MSE and MAE values were indicative of better performance in predicting emotion intensity.

After rigorous evaluation, the VADER model emerged as the preferred choice for assessing the effectiveness of trained models in predicting emotion intensity. Its superior performance, as evidenced by lower MSE and MAE values, underscores its efficacy in capturing nuanced emotional nuances within textual data. In essence, this study highlights the significance of employing a comprehensive suite of evaluation metrics to accurately assess model performance in emotion intensity analysis. The findings underscore the utility of the VADER model in deciphering emotional nuances, thereby enhancing our understanding of textual content and author sentiments.

REFERENCES

[1] Anil Bandhakavi, Nirmalie Wiratunga. (2019). " Emotion-aware polarity lexicons for Twitter sentiment analysis."

[2] Jorge Carrillo-de-Albornoz (2013). " An emotion-based model of negation, intensifiers, and modality for polarity and intensity classification "

[3] Wang, W. Y. (2017). "Liar, liar pants on fire: A new benchmark dataset for fake news detection."

[4] Potthast, M., Köpsel, S., Kiesel, J., & Eiselt, A. (2017). "A Stylometric Inquiry into Hyperpartisan and Fake News.".

[5] Horne, B. D., & Adali, S. (2017). "This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news."

[6] Sureshkumar Govindaraj (2018). " Intensified Sentiment Analysis of Customer Product Reviews Using Acoustic and Textual Features "

[7] Wang, W. Y., & Mohsen, M. A. (2020). "FND: A deep learning approach to detection." Information Processing & Management, 57(3), 102025.

[8] Hussain, I., Zhang, H., Shu, K., Ahmed, A., Alhossain, M., & Rho, S. (2020). "COVID-19 misinformation propagation in social media: A taxonomy of fake news." Information Processing & Management, 58(2), 102256.

[9] Ma, L., Gao, W., Mei, Q., & Su, L. (2017). "Detecting rumors in microblog posts using transfer learning and feature selection." Proceedings of the 26th International Conference on World Wide Web, 965-972.