Using Emotion Word Intensities for Multi-label Emotion Detection of Text

A Special Problem
Presented to the Faculty of
The Institute of Computer Science
University of the Philippines Los Baños

Andrei Francis U. Enrique
May 2024

Using Emotion Word Intensities for Multi-label Emotion Detection of Text

by

Andrei Francis U. Enrique

In partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science

---

Maria Art Antonette D. Clariño
Adviser

Date Signed

Maria Art Antonette D. Clariño
Director
Institute of Computer Science

Date Signed

ABSTRACT

Emotion word intensity values were derived from an emotion intensity lexicon and used along with n-grams as baseline features to create a linear support vector machine classifier using the one-vs-rest strategy for multi-labeled emotion data. The model was trained on a dataset of English subtitles and compared to a similarly configured classifier which only included baseline features. The results showed that the classifier with emotion intensity values performed significantly better than the model without these features, achieving a macro f1 score of 0.4359. Furthermore, emotion intensity values for disgust, fear, joy, and sadness showed very high feature importance for their respective emotion labels. The findings suggest that the use of emotion word intensities can help classification models in identifying the emotions expressed within bodies of text.

ACKNOWLEDGEMENT

I would like to thank my adviser, Maria Art Antonette D. Clariño, as her expertise and guidance was crucial through every part of this study. I would also like to thank all my friends and family for supporting my current and future endeavors.

TABLE OF CONTENTS

I. INTRODUCTION

A. Background of the Study

Emotions play a vital role in human life. They enable us to express our feelings, preferences, and needs to ourselves and others. Emotions can also affect how we assess situations, consider different options, and make decisions (Alzoubi & Aziz, 2021). Our emotions are manifested in various ways, such as our speech, our facial expressions, and the words that we write.

Billions of people from around the world use online social platforms to create and share various types of digital content. These platforms contain a wealth of textual data, including forum discussions, blogs, product reviews, and more (Hartmann et al., 2019). These platforms may also encourage users to express their thoughts and opinions on specific topics, which can allow people to make observations on how sentiments vary across a range of topics. Because of this, the analysis of emotions and sentiments of people from bodies of text can provide valuable insights for fields such as business and psychology. Hence, text-based emotion detection has become an important area of research.

Emotion analysis has been used in many real-life applications. It is able to provide additional granularity to areas that are traditionally served by polarity-based sentiment analysis (Kusal et al., 2022). The emotions derived from a body of text can vary based on the approach that is used. Many existing emotion analysis studies look at the presence of emotion words in bodies of text, and how they relate to other features present in the text. Emotion words are obtained through a lexicon which indicates if a word in the text is associated with one or more emotions. Valuable insight into how emotions influence online behavior and perception has been obtained through this approach, such as in news (Xu & Guo, 2018), e-commerce (Chatterjee, 2020), and social media (Chung & Zeng, 2020).

B. Statement of the Problem

Though lexicon-based approaches are still popular for applications of emotion analysis, they have been eclipsed in performance by more advanced approaches to emotion detection. Multi-labeled datasets are also becoming more widely available, further increasing the need for a more refined approach to the use of emotion lexicons.

More advanced emotion lexicons have been released that may help mitigate particular shortcomings. Among them is the NRC Emotion Intensity Lexicon (Mohammad, 2018), which uses best-worst scaling (BWS) to determine the degree to which a word is associated with a particular emotion. This study uses this lexicon alongside other classification features to develop a method for identifying emotions in text.

C. Significance of the Study

This study was useful in developing a better understanding of how to automatically identify the emotions expressed by a body of text. A better understanding of text-based emotion detection can pave the way for future researchers to develop even more accurate and robust emotion detection classifiers, especially those that are able to identify multiple emotions. This can also help future emotion analysis studies that need to automatically determine the emotions expressed by text by providing a model that is more representative of the true emotions of text, as compared to using an emotion lexicon.

D. Objectives of the Study

The general objective of this study is to utilize emotion word intensities to implement a classification model that can identify the emotions expressed by a body of text. The specific objectives are the following:

- Identify the relationship between the presence of emotion words in a body of text and the emotions expressed within said body of text,
- Determine the importance of emotion word intensities in relation to other features used in classification models to identify the emotions expressed in text,
- Evaluate the performance of the machine learning model in classifying multi-labeled emotion data using metrics such as per-label F1 score and macro F1 score, and test these metrics against a baseline model for statistical significance.

E. Scope and Limitations of the Study

This study is limited only to the development and testing of a classification model that uses values obtained from an emotion word lexicon to identify multi-labeled emotion data. The term ”emotion words” in the context of this study is only used to refer to words present in the utilized lexicon with intensity values greater than zero. The classification model was trained and tested only using text in the English language. The classification performance of the model was evaluated by comparing the emotion labels it identifies with the true emotion labels of the body of text, as identified by the original annotators of the dataset. The emotion labels used in the classification of text was limited only to the list of emotions present in the dataset, excluding the ”neutral” label as this indicates that the dataset does not classify the text under any of the given emotions.

II. REVIEW OF RELATED LITERATURE

A. Sentiment and Emotion Analysis

Sentiment analysis is the process of using natural language processing (NLP) and other computational techniques to identify and extract feelings or opinions expressed by a given source (Wankhade, Rao, & Kulkarni, 2022). Text is by far the most commonly used source, but others include speech, facial expression, body language, and more. Sentiment analysis is generally concerned with identifying the sentiment polarity of a given text, whether it be positive, negative, or neutral. It is a useful technique for applications such as customer feedback, social media analysis, product reviews, and natural language understanding.

Emotion analysis falls under the wider field of sentiment analysis, generally concerned with identifying discrete emotions within a given body of text (Yadollahi, Shahraki, & Zaiane, 2017). This is done with the goal of providing additional granularity to existing applications of sentiment analysis. For example, one may want to differentiate between text that conveys anger to text that conveys sadness, despite both being of negative sentiment. Certain emotion models may also provide labels that do not definitively fall under a certain polarity, such as surprise and anticipation.

Aside from identifying the emotions expressed by a given body of text, one can also look at the intensity in which certain emotions are expressed. Within Plutchik's theory of emotions (Plutchik, 1980) for example, serenity is an emotion that is highly related to joy and ecstasy, but is considered as being more intense than the two. When it comes to bodies of text, most would agree that the word "ecstatic" expresses much more joy than the word "satisfied". By extension, a sentence containing multiple words which express joy can be perceived as being more joyful than a sentence which only contains one or two of these words. Identifying the intensity in which words express emotions provides more nuance to the subject of emotion analysis, which can help provide more accurate information for emotion analysis tasks.

B. Applications of Emotion Analysis

Emotion analysis has seen use in many areas of research over the years. Aslam et al. (2020) used emotion analysis on news headlines during the early months of the COVID-19 pandemic. They found that while the headlines generally had negative sentiment scores, the use of discrete positive and negative emotions varied. Words associated with fear and trust were much more common than those associated with emotions such as disgust or joy. Next, Barnett et al. (2022) looked at the presence of emotions in their analysis of web-based reviews of physicians, divided based on gender and specialization. It was found that for physicians specializing in internal medicine, reviews of male physicians generally had more emotion words than those for female physicians. For physicians specializing in obstetrics and gynecology however, it was found that the reviews of female physicians included more words associated with negative emotions such as disgust and sadness.

The presence of words associated with certain emotions can also be used as features for classification tasks. A study by Chatterjee (2020) found that online reviews that contained words associated with sadness and disgust were more likely to be considered helpful, while reviews with words associated with anger and fear were likely to be considered unhelpful. Palomino et al. (2021) also suggested that the presence of emotions may be used in the identification of toxic content, as they are more likely to contain negative emotions such as disgust and anger.

C. Text-based Emotion Detection

Different approaches are used by studies to identify the emotions present in text. Hakak et al. (2017) divides these approaches into two main categories - lexicon-based and machine learning-based approaches. Others such as Kusal et al. (2022) present additional categories such as rule-based and deep learning-based approaches, though these can also fall under the former two categories.

Lexicon-based approaches, also sometimes referred to as keyword-based approaches, use a lexicon where words are annotated according to their association with one or more emotions. The text is then labeled according to the combined emotion value of each word in the text that is found in the lexicon. Though simple, this approach is far from perfect. Some text can express emotions without containing emotion words, while others can contain emotion words without expressing any emotion at all.

Machine learning-based approaches use machine-learning algorithms to learn from the linguistic features of text. The text usually undergoes various preprocessing steps, such as tokenization, POS tagging, and lemmatization. In supervised learning algorithms, useful features are then extracted from the text and are selected based on their importance. The algorithm is trained using the features and emotion labels and is finally tested according to how well it can predict emotion labels. Machine learning-based approaches are generally more accurate, especially those that utilize pre-trained language models such as BERT (Adoma, Henry, & Chen, 2020).

Hybrid approaches have also been attempted by previous studies that combine the use of a word lexicon and a machine learning model (Murthy & Kumar, 2021). Thanks to the creation of quality multi-labeled datasets such as GoEmotions (Demszky et al., 2020), the future of text-based emotion detection is leading towards the development of models that are able to identify more than one emotion in a body of text. In the original paper presenting the dataset, the researchers used a BERT-based model to achieve an F1 score of 0.46 when attempting to classify across their taxonomy of 27 emotions. When classifying across a condensed list of six emotions, the model was able to achieve a better average F1 score of 0.64.

III. METHODOLOGY

A. Development Tools to be Used

The algorithms that were used to process data and implement models were programmed using the programming language Python. The primary libraries that were used in the study include numpy, pandas, nltk, scikit-learn, and matplotlib. The numpy and pandas libraries were used to ease the handling of the dataset, while nltk was used for preprocessing tasks such as stemming and stop word removal. Scikit-learn (Pedregosa et al., 2011) is a machine learning library that was used in the study for the implementation of classification models. Lastly, matplotlib is a data visualization library which was used to create the figures for this study.

The emotion word lexicon that was used for this study is the NRC Emotion Intensity Lexicon (Mohammad, 2018). The lexicon lists 5,891 English words according to the degree of their association with eight distinct emotions - anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. This lexicon was used to obtain emotion features based on the words present in the lexicon. A summary of the lexicon is provided in table 1:

Table 1: Summary of NRC Emotion Intensity Lexicon

Number of emotions	8
Number of unique words	5891
Number of unique intensity values	9829
Proportion of values per emotion	Anger: 15.97% Anticipation: 8.77% Disgust: 11.11% Fear: 17.94% Joy: 12.86% Sadness: 13.17% Surprise: 6.93% Trust: 15.16%

Table 2 provides a sample of emotion words with their intensity value towards a given emotion. For each word, each emotion is scored from a range of 0 to 1, with values closer to

Table 2: Sample of emotion words with varying intensities

Emotion	Word	Intensity Value
Anger	'outraged'	0.964
	'jerk'	0.621
	'pout'	0.203
Anticipation	'eagerness'	0.812
	'preparation'	0.586
	'latent'	0.273
Disgust	'rancid'	0.859
	'deteriorated'	0.586
	'scoff'	0.375
Fear	'horrifying'	0.906
	'grim'	0.672
	'worrying'	0.484
Joy	'elated'	0.894
	'sweet'	0.625
	'calm'	0.394
Sadness	'misery'	0.938
	'bitterness'	0.578
	'empty'	0.384
Surprise	'thrilling'	0.859
	'electric'	0.578
	'dismissal'	0.312
Trust	'loyalty'	0.805
	'supports'	0.656
	'neutrality'	0.430

1 indicating the word conveys a stronger association with the emotion and values closer to 0 indicating it conveys a weaker association. Any given word may have multiple intensity values for each of the eight emotion labels if it is defined as such by the lexicon.

B. Data Gathering

This study uses the XED dataset (Öhman, Pàmies, Kajava, & Tiedemann, 2020), specifically the English language version of the dataset titled "en-annotated.tsv" within the dataset's repository. As shown in Table 3, this dataset consists of 17,528 movie subtitles annotated according to their association with one or more distinct emotions, totalling 22,422 annotations. This dataset was chosen as it uses the same emotion labels as the NRC Emotion Intensity Lexicon, and because

Table 3: Summary of "en-annotated.tsv" dataset

Number of annotated sentences	17528
Number of annotations	22422
Number of annotations per sentence	1: 60.9% 2: 13.5% 3: 3.2% 4: 0.6%
Proportion of annotations per label	Anger: 17.07% Anticipation: 15.16% Disgust: 10.33% Fear: 10.88% Joy: 12.63% Sadness: 10.99% Surprise: 10.89% Trust: 12.04%

movie subtitles can be used as an analog to real human dialogue. The dataset also separately includes subtitles annotated as "neutral", which were not used in this study. As the dataset has no existing distribution for train and test data, a stratified split of 90:10 for the training and test data was performed.

C. Data Preprocessing

The text that comprises each of the posts underwent various preprocessing steps to clean up the data and to allow for easier feature extraction. The preprocessing steps that were performed are as follows:

- *Removal of special characters*: Special characters as defined by the Python string constant *string.punctuation* were removed from the text. These include symbols such as asterisks, parentheses, punctuation marks, and quotation marks, among others. The newline character will also be removed.
- *Downcasing*: All uppercase letters were converted to lowercase.
- *Tokenization*: Each entry was broken up into a list of "tokens", or words divided by whitespace.

- *Stemming*: Words were truncated to their base form using nltk's SnowballStemmer.

Table 4: Visualization of preprocessing steps

Original	Special Characters Removed and Downcased	Tokenized	Stemmed
We must protect these investments so the area can continue to prosper and grow.	we must protect these investments so the area can continue to prosper and grow	['we', 'must', 'protect', 'these', 'investments', 'so', 'the', 'area', 'can', 'continue', 'to', 'prosper', 'and', 'grow']	['we', 'must', 'protect', 'these', 'invest', 'so', 'the', 'area', 'can', 'continu', 'to', 'prosper', 'and', 'grow']

D. Feature Extraction

Word unigrams, bigrams, and trigrams were used as the baseline features for the classifier. For the emotion word values, each word in the tokenized list before stemming was searched for in the emotion lexicon. The lexicon represents a word's association with eight distinct emotion labels as values from a range of 0 to 1. The emotion intensity values for each word were added up. Table 5 shows the process of obtaining emotion words and the emotion intensity values from a post:

Table 5: Process for obtaining emotion intensity values

Original	Tokenized	Emotion Words	Intensity Values
We must protect these investments so the area can continue to prosper and grow.	['we', 'must', 'protect', 'these', 'investments', 'so', 'the', 'area', 'can', 'continue', 'to', 'prosper', 'and', 'grow']	['continue', 'prosper', 'grow'	{'anger': 0, 'anticipation': 1.547, 'disgust': 0, 'fear': 0 'joy': 0.804, 'sadness': 0, 'surprise': 0, 'trust': 0.914}

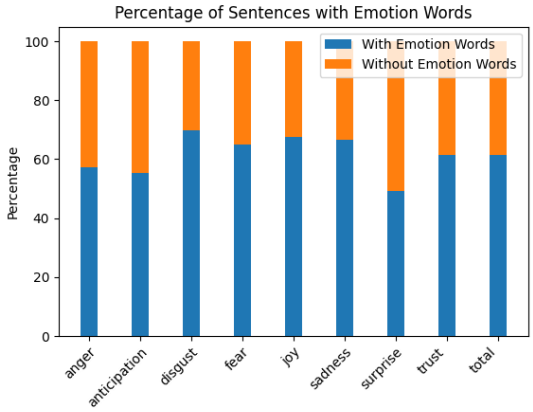

Figure 1. Proportion of sentences with emotion intensity values

Figure 1 shows the percentage of sentences in the dataset contain words with emotion intensity values. About 61.5% of all sentences in the dataset contain emotion words - this indicates that the lexicon used is limited in what words it perceives as being associated with emotions, and that the use of n-grams as baseline features is necessary in order to obtain sufficient feature data for the rest of the sentences.

E. Classification Model

This study utilized a linear Support Vector Machine (SVM) classifier using the one-vs-rest strategy for multi-label classification. Implementation was done using the LinearSVC class from scikit-learn. Linear SVM works by using labeled training data to find an optimal hyperplane that can separate classes in a given space (Shalev-Shwartz & Ben-David, 2014). The one-vs-rest strategy involves splitting a multi-label classification problem into multiple binary classifications (Brownlee, 2021). A linear SVM was trained for each class, which will learn to distinguish its respective class against all other classes. The LinearSVC classifier was initialized with balanced class weights and a regularization parameter of 0.01, as this resulted in the best classification performance.

F. Model Evaluation

Two versions of the classifier were created for this study. One classifier features only the baseline features while the other includes the values derived from emotion word intensities. The accuracy of the models were analyzed based on their F1 score, which is calculated as the harmonic mean of the precision and recall of the model, where precision is the fraction of true positives among all predicted positives for a class, and recall is the fraction of true positives among all actual positives for a class. The F1 scores for each of the emotion labels were obtained and compared, as well as the macro F1 score calculated by getting the average of the F1 scores for each label. The most important n-gram and emotion intensity features were also be analyzed by obtaining the coefficients for each linear SVM. In linear SVMs, a feature having a coefficient value closer to 1.0 indicates a strong positive correlation with the given label, while a value closer to -1.0 indicates a strong negative correlation.

$$\text{Precision}_{class} = \frac{TP_{class}}{TP_{class}+FP_{class}}$$
$$\text{Recall}_{class} = \frac{TP_{class}}{TP_{class}+FN_{class}}$$
$$\text{F1}_{class} = \frac{2*Precision*Recall}{Precision+Recall}$$
$$\text{F1}_{MacroAvg} = \frac{F1_{class1}+F1_{class2}+...+F1_{classX}}{X}$$

Figure 2. Evaluation Metrics

IV. RESULTS AND DISCUSSION

A. Emotion Intensity

Each sentence in the dataset was searched for emotion words and the emotion intensity values obtained were added up. Shown in the figure below are the average emotion intensity values for each emotion label.

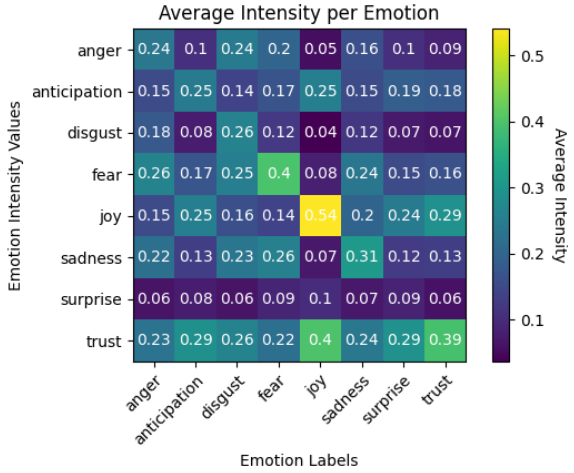

Figure 3. Average emotion intensity values for each emotion label

Based on Figure 3, it can be observed that certain emotion intensity values have a higher average value across all labels compared to others. Words that express trust and joy appear frequently in sentences regardless of label, while words expressing surprise appear very little across all sentences. For most emotion intensities, the label representing a particular emotions contain the highest intensity value for that emotion across all other labels. Sentences labeled with joy in particular have very high intensity values for joy and low values for negative emotions such as anger, disgust, and sadness. The joy label also has highest average intensity values for trust and surprise across all labels.

B. Evaluation of Classification Model

The models were evaluated based on their predicted emotion labels compared to the true labels provided in the dataset. As the dataset is multi-labeled, a given sentence may have multiple true labels. If the model were to predict a label which was not part of the true labels, it would count as a false positive for that particular label. Likewise, if the model failed to identify a label that was part of the true labels, it would be counted as a false negative for that label.

Table 6: F1 scores for each label

Class	F1 score (with emotion intensities)	F1 score (without emotion intensities)	Difference
Anger	0.4840	0.4773	0.0067
Anticipation	0.4873	0.4802	0.0071
Disgust	0.3854	0.3608	0.0246
Fear	0.4091	0.3862	0.0229
Joy	0.5518	0.5224	0.0294
Sadness	0.4338	0.4113	0.0225
Surprise	0.4148	0.3988	0.0160
Trust	0.4170	0.4071	0.0099
Macro Average	0.4479	0.4305	0.0174

Table 6 shows the F1 scores obtained for each class for both classifiers. The model with emotion intensity features has higher F1 scores for all labels compared to the model without these features, showing that the addition of emotion intensity values can improve classification performance. The disgust, fear, joy, and sadness labels show the most improvement in performance, increasing their F1 scores by more than 0.02.

The 'joy' emotion label achieved the highest F1 score at 0.5518, indicating that the model with emotion intensity values performs well at identifying sentences which express joy. All other labels in the model achieve an F1 score lower than 0.5, resulting in a macro average F1 score of 0.4479 across all labels. This indicates that the classifier has much room for improvement, either in the form of additional feature extraction or through the use of a more advanced classification model.

Table 7: Five highest positive coefficient n-grams for each label (model with emotion intensities)

Anger		Anticipation		Disgust		Fear		Joy		Sadness		Surprise		Trust	
n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient
'fuck'	0.60	'ill'	0.59	'stupid'	0.51	'troubl'	0.41	'nice'	0.73	'sorri'	0.68	'thought'	0.43	'worri'	0.46
'shut'	0.37	'readi'	0.38	'drunk'	0.39	'uh'	0.36	'great'	0.59	'poor'	0.58	'dont know'	0.37	'ill'	0.44
'stop'	0.36	'tonight'	0.36	'bullshit'	0.32	'kill'	0.35	'thank'	0.48	'cant'	0.49	'oh'	0.35	'thank'	0.39
'goddamn'	0.35	'let'	0.34	'mother'	0.29	'god'	0.33	'fine'	0.46	'never'	0.46	'whi'	0.35	'dont worri'	0.36
'bullshit'	0.35	'hell'	0.33	'peopl'	0.27	'pleas'	0.33	'love'	0.31	'didnt'	0.34	'surpris'	0.35	'help'	0.33

Table 8: Five highest positive coefficient n-grams for each label (model without emotion intensities)

Anger		Anticipation		Disgust		Fear		Joy		Sadness		Surprise		Trust	
n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient	n-gram	coefficient
'fuck'	0.62	'readi'	0.46	'stupid'	0.51	'kill'	0.55	'love'	0.77	'sorri'	0.66	'thought'	0.43	'thank'	0.40
'hate'	0.39	'tonight'	0.36	'bastard'	0.47	'danger'	0.50	'nice'	0.70	'die'	0.56	'surpris'	0.39	'worri'	0.38
'damn'	0.39	'come'	0.36	'lie'	0.45	'scare'	0.50	'beauti'	0.60	'poor'	0.56	'dont know'	0.38	'promis'	0.35
'shut'	0.38	'let'	0.35	'fool'	0.45	'troubl'	0.41	'wonder'	0.59	'cant'	0.48	'oh'	0.36	'dont worri'	0.32
'stop'	0.38	'wait'	0.35	'shit'	0.42	'afraid'	0.40	'great'	0.59	'cri'	0.44	'what'	0.36	'right'	0.32

C. Feature Importance

The most important features in the model were identified based on their coefficient values in the linear SVMs. A high positive coefficient value indicates a strong positive correlation with the given label, while a high negative coefficient indicates a strong negative correlation with the label. The n-gram and emotion intensity features were then separated to provide a better visualization of the findings.

Table 7 and 8 show the five n-grams with the highest positive coefficient for each label for the two models. It can be observed that certain n-grams with a high coefficient value in the model without emotion intensities do not appear in the table for the model with emotion intensities. These n-grams such as 'scare', 'wonder', and 'die' appear in the emotion lexicon. Words that appear in the emotion lexicon already provide emotion intensity values; as such, their respective unigrams are given a lower coefficient value in the model with emotion intensity values. Table 7 therefore provides insight into words which have a high positive correlation with an emotion label yet have not been included in the used lexicon.

Figure 4 shows the importance of emotion intensity values for each label. Throughout all labels, it can be observed that the the intensity values for the respective emotion label shows the highest importance over the rest of the emotion intensity values. This means that 'anger_int' is the most important for detecting anger, 'joy_int' is the most important for detecting joy, and so on. Of note, the intensity values of disgust, fear, joy, and sadness show very high

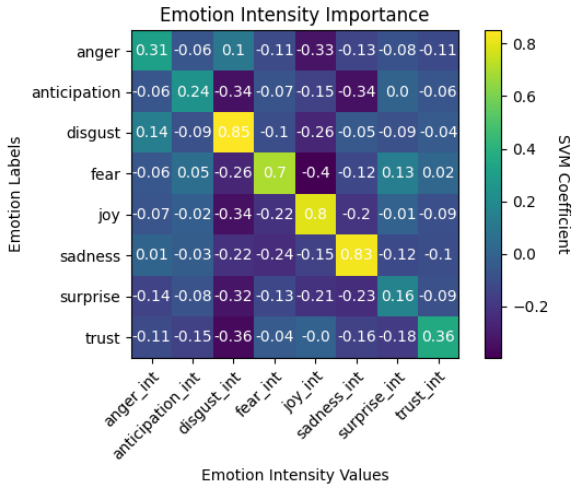

Figure 4. Feature importance for emotion intensity values

importance for their respective emotion labels, while anger, anticipation, surprise, and trust show relatively less importance. This can be interpreted as a limitation of the lexicon used to obtain the intensity values, as words that may be associated with a given emotion were not indicated as such by the lexicon. For example, several unigrams that show high importance with anger (such as profanities) are not included in the lexicon. Negative correlations can also be observed between emotion intensity values and certain labels, such as joy showing a negative correlation with negative emotion labels such as fear, disgust, and sadness, and vice versa.

D. Statistical Significance

In order to determine whether the improvements in F1 score observed between the two models are statistically significant, a 10-fold stratified t-test was performed. This was performed with a t-value of 2.262 and a p-value of 0.05. Shown below are the results of the t-test across the F1 scores of the different classes, with the statistically significant classes highlighted in bold:

The results of the t-test show that every F1 score has a positive t-statistic, indicating that

Table 9: T-statistic and p-value

F1 score	t-statistic	p-value
Macro Average	6.7505	$8.3601e^{-5}$
Anger	1.9331	0.0852
Anticipation	2.6758	0.0254
Disgust	4.1034	0.0027
Fear	3.0154	0.0146
Joy	4.1379	0.0025
Sadness	4.0615	0.0028
Surprise	2.7542	0.0223
Trust	2.2045	0.0549

the F1 score of every class improved when emotion intensity values were added as features for the model. However, the anger and trust classes have t-statistics and p-values that do not meet the defined thresholds, suggesting that the improvement in performance observed in these classes is not statistically significant. The rest of the classes - namely anticipation, disgust, fear, joy, sadness, and surprise, as well as the macro F1 score, are shown to be statistically significant.

V. SUMMARY AND CONCLUSION

The goal of this study was to use emotion word intensities to identify the presence of one or more emotions within a body of text. The results show that the addition of emotion intensities was able to improve the classification performance of the Linear SVM model. In particular, emotion intensity values for disgust, fear, joy, and sadness showed high feature importance for their respective classes, which was reflected in a larger improvement in F1 score. Certain classes also were able to express a negative correlation with several intensity values, such as with the 'joy' label showing a negative correlation with emotion intensity values of fear and disgust.

Based on the classification results of the model, there is still much room for improvement in regards to the identification of multiple emotions in a body of text. A possible area for improvement for future studies would be to expand upon how emotion intensity values are obtained. The use of an emotion intensity lexicon is ultimately limited by what words are included in it. Only 61.5% of the sentences in the dataset included emotion intensity, and so the remaining sentences were only able to be classified based on the presence of certain n-grams. Included in the results is a table of words which show high positive correlations with a given emotion label, but were not included in the utilized lexicon. Future studies may take these results into account when considering a lexicon-based approach. Grammar-based modifiers may also be accounted for in the future, as adjectives and negative words can influence the intensity values derived from emotion words. Lastly, future studies may also explore how emojis, special characters, and other figures of speech can influence the emotions expressed by text.

REFERENCES

ADOMA, A. F., HENRY, N. M., & CHEN, W. (2020). Comparative analyses of bert, roberta, distilbert, and xlnet for text-based emotion recognition. In *2020 17th international computer conference on wavelet active media technology and information processing (iccwamtip)* (p. 117-121).

ALZOUBI, H., & AZIZ, R. (2021). Does emotional intelligence contribute to quality of strategic decisions? the mediating role of open innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(2), 130.

ASLAM, F., AWAN, T. M., SYED, J. H., KASHIF, A., & PARVEEN, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (covid-19) outbreak. *Humanities and Social Sciences Communications*, *7*(1).

BARNETT, J., BJARNADÓTTIR, M., ANDERSON, D., & CHEN, C. (2022). Understanding gender biases and differences in web-based reviews of sanctioned physicians through a machine learning approach: Mixed methods study. *JMIR Form Res*, *6*(9), e34902.

BROWNLEE, J. (2021). *One-vs-rest and one-vs-one for multi-class classification.* Retrieved from `https://machinelearningmastery.com/one-vs-rest-and-one-vs-one-for-multi-class-classification`

CHATTERJEE, S. (2020). Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. *International Journal of Hospitality Management*, *85*, 102356.

CHUNG, W., & ZENG, D. (2020). Dissecting emotion and user influence in social media communities: An interaction modeling approach. *Information Management*, *57*(1), 103108.

DEMSZKY, D., MOVSHOVITZ-ATTIAS, D., KO, J., COWEN, A., NEMADE, G., & RAVI, S. (2020). *Goemotions: A dataset of fine-grained emotions.*

HAKAK, N. M., KIRMANI, M., & MOHD, M. (2017). Emotion analysis: A survey. In *2017 international conference on computer, communications and electronics (comptelix)* (pp. 397–402).

HARTMANN, J., HUPPERTZ, J., SCHAMP, C., & HEITMANN, M. (2019). Comparing automated text classification methods. *International Journal of Research in Marketing*, *36*(1), 20-38.

KUSAL, S., PATIL, S., CHOUDRIE, J., KOTECHA, K., VORA, D., & PAPPAS, I. (2022). *A review on text-based emotion detection – techniques, applications, datasets, and future directions.*

MOHAMMAD, S. (2018). Word affect intensities. In *Proceedings of the 11th edition of the*

language resources and evaluation conference (lrec-2018).

MURTHY, A., & KUMAR, K. A. (2021). A review of different approaches for detecting emotion from text. In *Iop conference series: Materials science and engineering* (Vol. 1110, p. 012009).

ÖHMAN, E., PÀMIES, M., KAJAVA, K., & TIEDEMANN, J. (2020).
In *The 28th international conference on computational linguistics (coling 2020).*

PALOMINO, M., GRAD, D., & BEDWELL, J. (2021). GoldenWind at SemEval-2021 task 5: Orthrus - an ensemble approach to identify toxicity. In *Proceedings of the 15th international workshop on semantic evaluation (semeval-2021)* (pp. 860–864).

PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., ... DUCHESNAY, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

PLUTCHIK, R. (1980). A general psychoevolutionary theory of emotion. In *Theories of emotion* (p. 3-33). Academic Press.

SHALEV-SHWARTZ, S., & BEN-DAVID, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.

WANKHADE, M., RAO, A., & KULKARNI, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780.

XU, Z., & GUO, H. (2018). Using text mining to compare online pro- and anti-vaccine headlines: Word usage, sentiments, and online popularity. *Communication Studies*, *69*(1), 103-122.

YADOLLAHI, A., SHAHRAKI, A. G., & ZAIANE, O. R. (2017). Current state of text sentiment analysis from opinion to emotion mining. *ACM Comput. Surv.*, *50*(2).