How Are You Feeling?
Emotion Recognition in Social Media

Cameron Arnold
Faculty of Science
University of Western Ontario
London, Canada
carnol29@uwo.ca

Ali Al-Tekriti
Faculty of Science
University of Western Ontario
London, Canada
aaltekri@uwo.ca

Lorne Abraham Crewson
Faculty of Science
University of Western Ontario
London, Canada
lcrewson@uwo.ca

Mariah DeMarco
Faculty of Engineering
University of Western Ontario
London, Canada
mdemarc7@uwo.ca

***Abstract*—Social media platforms such as Facebook, Twitter, and Reddit are used every day by millions of users to convey information, facts, and opinions. This project aims to use the GoEmotions dataset made up of human-annotated Reddit comments and apply several different supervised and unsupervised methods to classify the 28 unique emotions of the comments. This project uses Supervised Vector Machines (SVM), K-Means, and Lexicographic Analysis as methods in which to implement for classification. GoEmotions, the dataset used for this project, is preprocessed and cleaned before implementing the methods. The results for each method were similar across the board, as the methods were able to accurately predict and classify emotions in the test set less than half of the time or worse for each method. However, the results did find the reasoning behind the low accuracy to be from similarities between emotion categories and select word categorizations involved with the dataset.**

Keywords—-SVM, K-Means, Lexicographic Analysis, GoEmotions, Reddit, Emotions

I. INTRODUCTION

In the last decade social media has grown immensely in popularity. Along with this popularity came an endless stream of unstructured data representing the opinions, thoughts, and sentiments of the world population [1]. This reality has profound significance as there now exists a view into the general population's consciousness like there never was before. The sentiment analysis of these texts can aid in providing feedback for a specific topic whether it be for political, marketing or any other contextually related purpose.

This project aims to implement an approach to provide this feedback with the use of a dataset containing texts from the social media application, Reddit, labeled with 28 different emotions, and by employing the methods of Supervised Vector Machines (SVM), K-Means, and Lexicographic Analysis.

II. BACKGROUND AND RELATED WORK

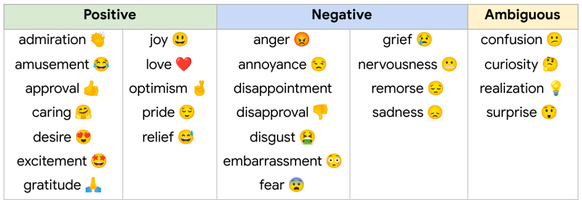
As the usage of both social media and artificial intelligence has continued to increase in recent years, so has the amount of research and applications of machine learning on sentiment analysis in popular applications such as Reddit. Sentiment analysis being the task of detecting, extracting and classifying sentiments in opinionated text documents[2]. There are similar studies that focus on expression analysis of social media posts. Many of these studies divided expressions into categories labeled positive, negative, or neutral, while a lesser number of studies' research aligned closer with what this paper aims to analyze, the emotion analysis of social media text containing more complex emotions such as happy, anger and sadness. However, this paper differs in the way that it aims to categorize a broader range of 28 complex emotions from Reddit comments [3].

Identifying emotions from natural language is difficult, especially when considering the fine-grained approach that requires semantic and syntactic analysis of the sentence[4]. Strapparava et al. did this with an approach that would function based on keyword detection. The implementation included a linguistic resource named WordNet – Affect, a subset of synsets suitable to represent affective concepts correlated with affective words[5]. Emotion classification is then done by mapping emotional keywords that exist in the input sentence to their corresponding WordNet-Affect concepts.

A study published by Rustam et al. had a more narrow focus utilizing sentiment analysis on tweets only regarding COVID-19 for the purpose of handling the current and future pandemics better [6].There have also been many studies that focus on sentiment analysis outside the realm of popular social media such as Pang and Lee's research using Naive Bayes, maximum entropy classification, and support vector machines to classify movie reviews based on negative or positive sentiment [7].

III. DATA

The aim of this project involves identifying the emotion of tweets through fine-grained emotion classification. The dataset used in the project is: ”GoEmotions: A Dataset of Fine-Grained Emotion .” GoEmotions is a human-annotated dataset of 58,000 Reddit comments extracted from popular English-language subreddits and labeled with 27 emotion categories. The 27 emotion categories included in the dataset are admiration, amusement, anger, annoyance, approval, caring, confusion, curiosity, desire, disappointment, disapproval, disgust, embarrassment, excitement, fear, gratitude, grief, joy, love, nervousness, optimism, pride, realization, relief, remorse, sadness, and surprise. Additionally, the category of neutral is included in the dataset. These 27 emotion categories can be further categorized into 12 positive, 11 negative, and 4 ambiguous emotion categories in addition to neutral.



Positive		Negative		Ambiguous
admiration 👏	joy 😃	anger 😡	grief 😢	confusion 😕
amusement 😂	love ❤️	annoyance 😒	nervousness 😬	curiosity 🤔
approval 👍	optimism 🤞	disappointment	remorse 😔	realization 💡
caring 🤗	pride 😌	disapproval 👎	sadness 😞	surprise 😲
desire 😍	relief 😅	disgust 🤮		
excitement 🤩		embarrassment 😳		
gratitude 🙏		fear 😨		

Fig. 1. The 27 classes of emotions.

IV. EXPLORATORY DATA ANALYSIS

Exploratory data analysis helps define which pre-processing steps can be utilized when handling the data, while also revealing significant patterns within the corpus that may be useful when selecting testing models. We will analyze the two primary features within the dataset - ‘target’ and ‘text’ - as follows.

A. Target

The target consists of the 27 emotion categories outlined in Part 3 of this report. By plotting each piece of data based on its emotion category (*see attachment DataEmotionComparison.PNG)*, a significant imbalance between categories is revealed. For instance, the category “neutral” has over 50,000 instances, while the category “embarrassment” has under 3,000. This imbalance may introduce a bias towards “neutral” labels, hinting that an additional pre-processing step to normalize the data may be required.

B. Text

The content of each Reddit post is used to identify its emotional value. Several aspects of the text must be considered. For instance, the sentiment of a post does not depend on its length. A long, 100-word post describing a user’s enjoyable day, and a short post similar to “I am so happy!” will both be labeled under the category “joy”. Furthermore, stopwords within the post also reveal minimal information on the post’s sentiment. This also holds true with the number of user and subreddit references in the post, as users and subreddits may be used in both a positive or negative context.

V. DATA PREPROCESSING

Before the models can be built, text data needs to be preprocessed to remove any ambiguity or inconsistencies present. Some of the text processing techniques include:

A. Conversion to Lowercase

Each uppercase character present in the text was converted to a lower-case format. This allows for words with different representations (such as capitalized and uncapitalised words) to be treated identically, further increasing data consistency.

B. Noise Removal

HTTP hyperlinks, HTML tags, user mentions, and subreddit mentions are removed from the text, as they provide little relevance to the actual sentiment of the user’s post. Furthermore, punctuation was also removed to provide further consistency across data. Punctuation characters also cannot be easily implemented as a vector representation for embedding.

C. Stopwords Removal

A list of common stopwords provided by the *Natural Language Toolkit* were removed from the text. These words (such as “and”, “or”, “the", etc.) tend to have a higher frequency among user posts while providing little sentimental relevance. Removing these stopwords allows for quicker and more effective data processing.

D. Spelling Correction

Misspelled words tend to appear frequently among social media sites such as Reddit, introducing inconsistency across text data. As a solution, a spell-checker was implemented to validate words in the corpus against a dictionary, then allow for an informed correction if needed.

E. Stemming

The Porter Stemmer was used to reduce each word in the corpus to its root form. This further increases consistency across data by allowing for words with different tenses (ex. “loving” and “love”) to be processed identically.

F. Feature Removal

Features included in the original dataset, such as the post identification number, author, and subreddit name, provide no information in terms of the text’s sentiment. These features were removed from the dataset to reflect their negligibility.

G. Row Removal

The feature “Example Very Unclear” represents entries in the dataset where the text’s sentiment could not easily be determined by the rater. Entries with this field marked as true were removed from the dataset to further reduce noise and enable a more confident distinction between different sentiments in text. Furthermore, posts that were classified as “neutral” were also removed from the dataset, as their high prominence and unpredictability significantly reduced accuracy.

VI. RESEARCH OBJECTIVES

Sentiment analysis specifically on social media like Twitter has great significance and boundless applications. The mass amounts of unstructured data generated every second yields a method to collect the feelings and opinions of the general population, something that both governments and corporations strive to procure for the purpose of prevention, manipulation, control, and wealth accumulation. The research in this paper can encourage and better the automation of this process.

A. Riots and Uprisings

Having access to the general public’s social media posts means having a pulse on nations' and communities’ sentiments about current events and feelings toward the government. This can in turn help local or national governments prepare for, control and possibly prevent riots and uprisings. The additional collected variable of location can further increase the achievement of this goal.

B. Political Outcome Prediction/Manipulation

Having the ability to continuously monitor the population’s feelings toward political candidates can help predict who will win the election. This ability can also give feedback to candidates on how they can increase positive public sentiment.

C. Businesses and Marketing Feedback

This research has considerable significance on the market. The availability of accurate and automated sentiment analysis from social media could give businesses the ability to fine tune their marketing strategies to their target markets. Feedback could also help yield new products and services based on general opinion.

VII. RESEARCH METHODOLOGY

The project objective is to train a model using Support Vector Machine (SVM) to categorize and predict which emotion a Reddit comment expresses. TF-IDF technique will be used to represent words as word vectors of numbers. Additionally, K-Means, another unsupervised learning technique will be used to confirm underlying patterns and compare the clusters against the groupings of the labeled dataset.

A. SVM

This supervised method can be used for both classification and regression problems. Support Vector Machines (SVM) aim to find an optimal boundary between possible outputs, which in this case are emotions. The default SVM algorithm does not support multi-class classification, which is required to classify all 28 emotions in the GoEmotions dataset. Two different approaches can allow for multi-class classification using SVM. The first approach is the One-to-One approach which breaks down the multi-class problem into multiple separate binary classification problems. It assigns a binary classifier to each pair of classes. The second approach is the One-to-Rest approach, where the problem is split up into multiple separate binary classification problems. This assigns a binary classifier to each class. For this implementation, we used the LinearSVC model, which implements the One-to-Rest approach for the multi-class strategy. The One-to-Rest approach uses n SVMs where each SVM predicts the membership of one of the n classes. For each SVM, the aim is to find the maximum margin with the closest point for the hyperplane, which divides the set of two classes. Given the formula of the hyperplane:

$$\vec{\omega} \cdot \vec{x} + b = 0$$

If the training data is linearly separable, the hyperplanes of the margin can be described by :

$$\vec{\omega} \cdot \vec{x} + b = 1 \qquad \vec{\omega} \cdot \vec{x} + b = -1$$

These two equations form the boundary for the values of the two separable classes. If the training data is not linearly separable, the SVM must use a soft-margin where the goal is to minimize this equation:

$$\lambda \|\mathbf{w}\|^2 + \left[\frac{1}{n}\sum_{i=1}^{n} \max\left(0, 1 - y_i(\mathbf{w}^T\mathbf{x}_i - b)\right)\right]$$

This supervised method was chosen to train the model using a portion of the dataset and then attempt to predict the classes of the test portion of the dataset.

B. LSTM

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN). The advantage that RNNs give over general neural networks is that RNNs can process sequential information and deal with the sequence change. For instance, the meaning of a given word can differ based on its position in the sentence, and based on the context; to deal with this problem, an RNN is used. However, RNNs are poor when processing long term sequences due to their long-term dependency which can cause the vanishing gradient problem. To solve this problem, LSTM adds three gate units in the hidden layer: forgetting gate, remembering gate, and output gate [citation].

An LSMT network is a
sequence of the modules seen in the figure above. The network works to remember long-term memory and forget trivial information. It does that using the inner gates to manage the transmission state.

In this project, the Reddit comment is embedded using Keras' embedding, then fed as input to the neural network.
The following hyper-parameters were used in the model:

- Each Reddit comment is first converted into a vector of length 130. Then it gets transformed into a 130 x 128 matrix by the embedding layer.
- A spatial dropout layer with a dropout coefficient of 0.2 is added after the embedding layer.
- The output layer has 64 neurons. 20% of the units are dropped for the linear transformation of the inputs and the recurrent states.

- Softmax is used as the activation function for the output layer because the project is about multi-class classification.
- Learning rate?
- The learning rate will start to decrease if the loss on the validation stops improving for 7 rounds.
- If the improvement in loss is smaller than 0.0001, then it is considered negligible, and the training terminates.

The following observations can be made by looking at the figures below:

- Figure 2 shows the loss decreasing on the training and test sets.
- Figure 3 shows the accuracy increasing on the training and test sets.
- In addition, By looking at the figures, we can see that there is no overfitting.

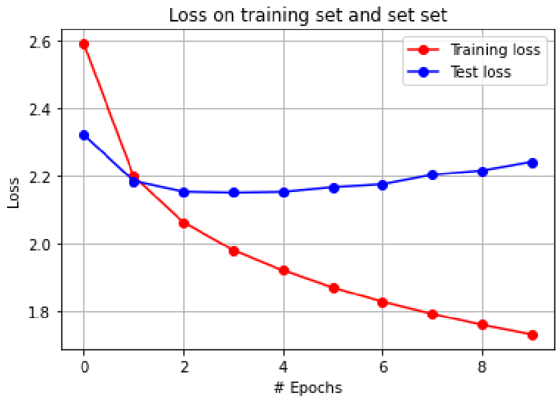


Fig. 2. Loss with Epochs on Training and Test sets.

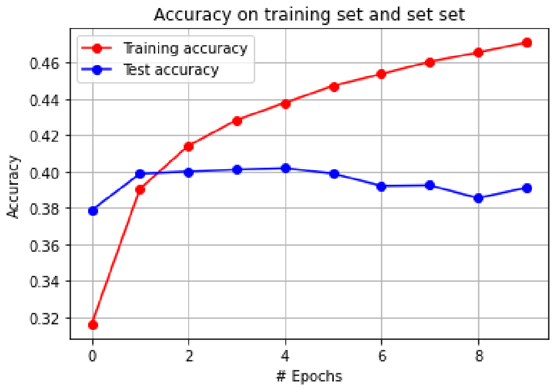


Fig. 3. Accuracy with Epochs on Training and Test sets.

B. Unsupervised K-Means

This unsupervised method makes inferences from the dataset using only the input vectors, being the processed text, without referring to the known labels, in this case the emotions. Instead it will group data points by detecting underlying patterns. This will be done by inputting (k), the number of clusters/centroids needed as a parameter. K is set to 28 since there are 28 known labeled emotions for the dataset. The algorithm will then iteratively assign each data point to one of the clusters based on how close it is to the respective centroid. The K-Means algorithm can be implemented with the following steps:

1. Pick K points as the initial centroids from the dataset.
2. Find the Euclidean distance of each point in the dataset with the identified cluster centroids.

$$d(p,q) \ = \sqrt{(q1-p1)^2+(q2-p2)^2}$$

$$p \ = \ (p1,p2) \ q = \ (q1,q2)$$

3. Assign each data point to the closest centroid using the Euclidean distance.

$$argmin_{c_i \in C}\ dist(c_i x^2)$$

4. Find the new centroid by taking the average of the points in each cluster group.

$$c_i = \frac{1}{|S_i|}\sum_{x_i \in S_i} x_i$$

$$S_{i\ is\ the\ set\ of\ points\ assigned\ to\ the\ ith\ cluster}$$

5. Repeat 2 to 4 for a fixed number of iterations or till the centroids don’t change.

This unsupervised method was chosen to compare the true labeled data against how this method groups the data. Furthermore, it was chosen to visualize the strength of the underlying pattern, and to use it as a factor in justifying the accuracy of the results yielded from alternate supervised methods.

C. TF-IDF and Classifiers

Before classification can begin, the text within the corpus needs to be represented as a word vector of numbers. Traditionally, vectorization is performed using the bag-of-words model, which values each word based on the number of times it appears throughout the text. However, this creates a significant bias towards longer posts. In contrast, the term-frequency inverse-document-frequency (TF-IDF) model allows for document length to be taken into consideration when valuing words. The expression is the product of the following two equations:

$$TF_{t,d} = 1 + log(f_{t,d})\ if\ f_{t,d} > \ 0,\ otherwise\ 0$$
$$IDF_t = log(N/N_t)$$

where $f_{t,d}$ represents the frequency of term t in document d, N represents the number of documents, and N_t represents the number of documents containing term t.

Next, the vector matrix is fed into three different classifiers for comparison. These include the Complement Naive Bayes Classifier, the Bernoulli Naive Bayes Classifier, and the Stochastic Gradient Descent Classifier. The latter minimizes an objective function that has the form of a sum, using the following equation

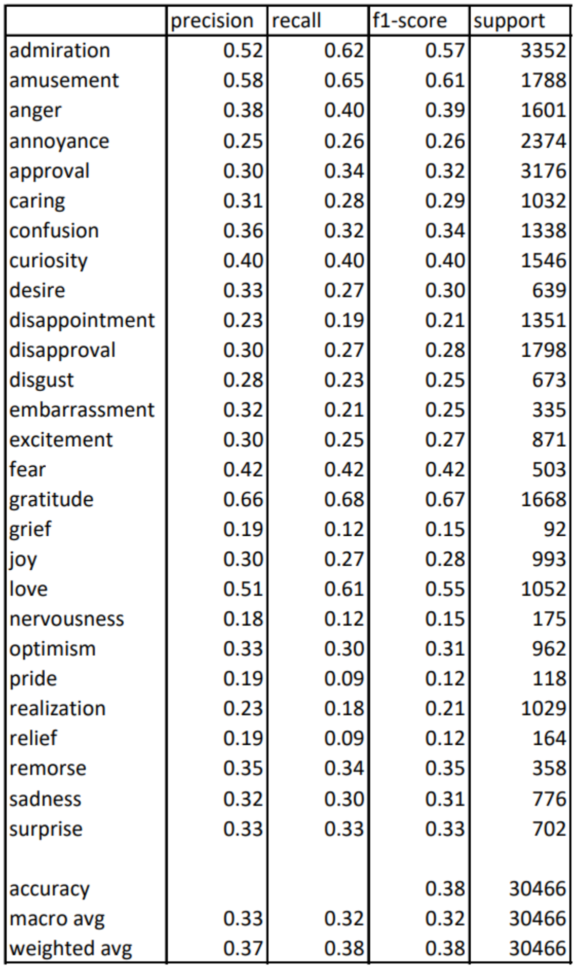
$$Q(w) = \frac{1}{n}\sum_{i=1}^{n} Q_i(w),$$

Where the parameter w that minimizes Q(w) is to be estimated.

VIII. EXPERIMENTAL RESULT

A. SVM

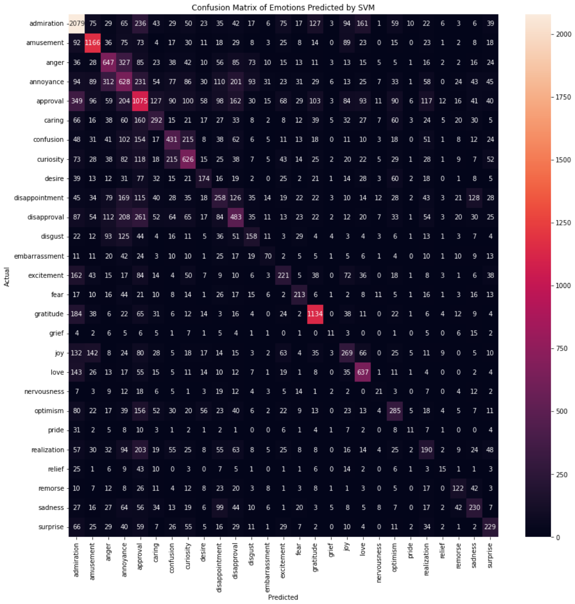
The results of the multiclass SVM on the cleaned and preprocessed data led to poor results in terms of overall accuracy for the model on the dataset. Following the training of the LinearSVC model through a training set of 121861 Reddit comments, the model was evaluated using a test set of 30466 Reddit comments. As shown in Fig. 4, the test's average accuracy came out to approximately 38%. Specific emotion categories fared better than others in terms of precision as larger training sample data was linked to higher precision during this test.



	precision	recall	f1-score	support
admiration	0.52	0.62	0.57	3352
amusement	0.58	0.65	0.61	1788
anger	0.38	0.40	0.39	1601
annoyance	0.25	0.26	0.26	2374
approval	0.30	0.34	0.32	3176
caring	0.31	0.28	0.29	1032
confusion	0.36	0.32	0.34	1338
curiosity	0.40	0.40	0.40	1546
desire	0.33	0.27	0.30	639
disappointment	0.23	0.19	0.21	1351
disapproval	0.30	0.27	0.28	1798
disgust	0.28	0.23	0.25	673
embarrassment	0.32	0.21	0.25	335
excitement	0.30	0.25	0.27	871
fear	0.42	0.42	0.42	503
gratitude	0.66	0.68	0.67	1668
grief	0.19	0.12	0.15	92
joy	0.30	0.27	0.28	993
love	0.51	0.61	0.55	1052
nervousness	0.18	0.12	0.15	175
optimism	0.33	0.30	0.31	962
pride	0.19	0.09	0.12	118
realization	0.23	0.18	0.21	1029
relief	0.19	0.09	0.12	164
remorse	0.35	0.34	0.35	358
sadness	0.32	0.30	0.31	776
surprise	0.33	0.33	0.33	702
accuracy			0.38	30466
macro avg	0.33	0.32	0.32	30466
weighted avg	0.37	0.38	0.38	30466

Fig. 4. Table of SVM testing results on a test set of 30466 Reddit comments by emotion.

As shown in Fig. 5, some emotion categories were more accurately predicted than the rest. These higher precision categories include admiration, amusement, approval, and gratitude. In Fig. 5, there is notable miscategorization between pairs of emotions, including anger and annoyance, confusion and curiosity, disappointment and disapproval, and admiration and approval. Many similarities between emotions caused miscategorization among emotion categories where the pair of emotions both belonged to either positive, negative, or ambiguous subcategories of the dataset. The low overall accuracy of the test was common over all the methods done in this project. The underlying factors causing this poor testing result include the varying amount of data points per emotion, similarities between emotion classifications, and not using a sufficiently complex model to predict the data. For this SVM method, neutral comments were removed prior to training the model as exploratory attempts at using the unaltered data led to higher accuracy and higher miscategorization among all categories except neutral. The miscategorization was due to neutral comments making up the largest sub-portion of the dataset, which caused a large percentage of the test predictions to result in neutral regardless of accuracy.



	admiration	amusement	anger	annoyance	approval	caring	confusion	curiosity	desire	disappointment	disapproval	disgust	embarrassment	excitement	fear	gratitude	grief	joy	love	nervousness	optimism	pride	realization	relief	remorse	sadness	surprise
admiration	2079	75	29	65	236	43	29	50	23	35	42	17	6	75	17	127	3	94	161	1	59	10	22	6	3	6	39
amusement	92	1166	36	75	73	4	17	30	11	18	29	8	3	25	8	14	0	89	23	0	15	0	23	1	2	8	18
anger	36	28	647	327	85	23	38	42	10	56	85	73	10	15	13	11	3	13	15	5	5	1	16	2	2	16	24
annoyance	94	89	312	628	231	54	77	86	30	110	201	93	31	23	31	29	6	13	25	7	33	1	58	0	24	43	45
approval	349	96	59	204	1075	127	90	100	58	98	162	30	15	68	29	103	3	84	93	11	90	6	117	12	16	41	40
caring	66	16	38	60	160	292	15	21	17	27	33	8	2	8	12	39	5	32	27	7	60	3	24	5	20	30	5
confusion	48	31	41	102	154	17	431	215	8	38	62	6	5	11	13	18	0	11	10	3	18	0	51	1	8	12	24
curiosity	73	28	38	82	118	18	215	626	15	25	38	7	5	43	14	25	2	20	22	5	29	1	28	1	9	7	52
desire	39	13	12	31	77	32	15	21	174	16	19	2	0	25	2	21	1	14	28	3	60	2	18	0	1	8	5
disappointment	45	34	79	169	115	40	28	35	18	258	126	35	14	19	22	22	3	10	14	12	28	2	43	3	21	128	28
disapproval	87	54	112	208	261	52	64	65	17	84	483	35	11	13	23	22	2	12	20	7	33	1	54	3	20	30	25
disgust	22	12	93	125	44	4	16	11	5	36	51	158	11	3	29	4	4	3	4	3	6	1	13	1	3	7	4
embarrassment	11	11	20	42	24	3	10	10	1	25	17	19	70	2	5	5	1	5	6	1	4	0	10	1	10	9	13
excitement	162	43	15	17	84	14	4	50	7	9	10	6	3	221	5	38	0	72	36	0	18	1	8	3	1	6	38
fear	17	10	16	44	21	10	8	14	1	26	17	15	6	2	213	6	1	2	8	11	5	1	16	1	3	16	13
gratitude	184	38	6	22	65	31	6	12	14	3	16	4	0	24	2	1134	0	38	11	0	22	1	6	4	12	9	4
grief	4	2	6	5	6	5	1	7	1	5	4	1	1	0	1	0	11	3	0	0	1	0	5	0	6	15	2
joy	132	142	8	24	80	28	5	18	17	14	15	3	2	63	4	35	3	269	66	0	25	5	11	9	0	5	10
love	143	26	13	17	55	15	5	11	14	10	12	7	1	19	1	8	0	35	637	1	11	1	4	0	0	2	4
nervousness	7	3	9	12	18	6	5	1	3	19	12	4	3	5	14	1	2	2	0	21	3	0	7	0	4	12	2
optimism	80	22	17	39	156	52	30	20	56	23	40	6	2	22	9	13	0	23	13	4	285	5	18	4	5	7	11
pride	31	2	5	8	10	3	1	2	1	2	1	0	0	6	1	4	1	7	2	0	8	11	7	1	0	0	4
realization	57	30	32	94	203	19	55	25	8	55	63	8	5	25	8	8	0	16	14	4	25	2	190	2	9	24	48
relief	25	1	6	9	43	10	0	3	0	7	5	1	0	1	1	6	0	14	2	0	6	1	3	15	1	1	3
remorse	10	7	12	8	26	11	4	12	8	23	20	3	8	1	3	8	1	1	3	0	5	0	17	0	122	42	3
sadness	27	16	27	64	56	34	13	19	6	99	44	10	6	1	20	3	5	8	5	8	7	0	17	2	42	230	7
surprise	66	25	29	40	59	7	26	55	5	16	29	11	1	29	7	2	0	10	4	0	11	2	34	2	1	2	229

Fig. 5. Confusion matrix of SVM testing results on a test set of 30466 Reddit comments by emotion.

B. LSTM

C. Unsupervised K-Means

This means of unsupervised learning did not yield good nor helpful results. Partitioning in K-Means is better suited for when there are spherical clusters. The graph demonstrates that many data points do not produce clear clusters and overlap with many data points from other clusters. These results also precede other techniques and portend the accuracy of further classification methods.

Clustering of data with K=28

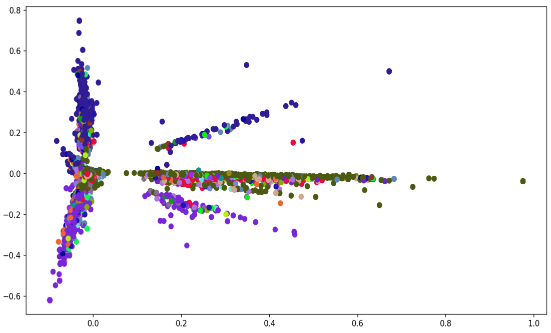
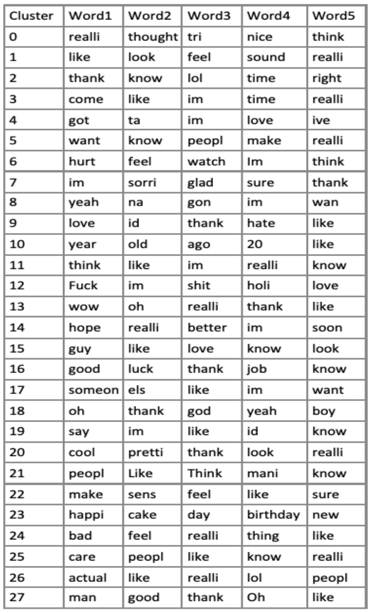


Fig 6. Clustering of emotion with 28 centroid.

It can also be noted that in the top 5 words produced for each cluster that some are contradictory to each other in the context of being associated with their labeled emotion. Clearer partitions could be expected with fewer clusters. Many words are able to be attributed to different emotions depending on the context. K-means could have yielded more helpful results despite this, if the complexity and scope of emotions were reduced.



Cluster	Word1	Word2	Word3	Word4	Word5
0	realli	thought	tri	nice	think
1	like	look	feel	sound	realli
2	thank	know	lol	time	right
3	come	like	im	time	realli
4	got	ta	im	love	ive
5	want	know	peopl	make	realli
6	hurt	feel	watch	Im	think
7	im	sorri	glad	sure	thank
8	yeah	na	gon	im	wan
9	love	id	thank	hate	like
10	year	old	ago	20	like
11	think	like	im	realli	know
12	Fuck	im	shit	holi	love
13	wow	oh	realli	thank	like
14	hope	realli	better	im	soon
15	guy	like	love	know	look
16	good	luck	thank	job	know
17	someon	els	like	im	want
18	oh	thank	god	yeah	boy
19	say	im	like	id	know
20	cool	pretti	thank	look	realli
21	peopl	Like	Think	mani	know
22	make	sens	feel	like	sure
23	happi	cake	day	birthday	new
24	bad	feel	realli	thing	like
25	care	peopl	like	know	realli
26	actual	like	realli	lol	peopl
27	man	good	thank	Oh	like

Fig 7. Top 5 words for each cluster.

Looking at the results, it should also be noted that another reason for the undefined and unexpected clustering is due to the fact that without labels the method can classify the text in any way. There are far more ways to categorize text than just by emotions.

D. TF-IDF and Classifiers

After experimenting with the parameters test size and random state it was determined that the values of 0.5 and 3 resulted in the highest scores respectively. In this instance, TF-IDF applied with the Complement Naive Bayes Classifier resulted in an accuracy score of 35.03%, and applied with the Bernouilli Naive Bayes Classifier resulted in an accuracy score of 33.94%. In contrast, the non-Bayesian classifier - the Stochastic Gradient Descent Classifier - scored an impressive 37.02% (*see attached file SGDCFinalConfusionMatrix.JPG for reference).* Furthermore, in all applied classifiers, emotion categories with similar semantic meaning tended to produce the most inconsistencies. For instance, the implemented models commonly confused “admiration” with “gratitude”, “love”, and “approval”, and “annoyance” with “disapproval”, “disgust”, and “disappointment”.

The success of the Stochastic Gradient Descent Classifier may be a result of the gradient descent optimization, which calculates the gradient as opposed to a general probability as seen with the Bayesian classifiers. In posts with semantic context, gradients in text data with reveal more significant results as opposed to generalized probabilities.

IX. CONCLUSIONS AND FUTURE WORKS

As people continue to use the internet in their everyday lives and the use of social media platforms such as Reddit increases, the need for evaluation of online comments is becoming more critical. The project has explored three different classification methods on the GoEmotions dataset. The results of the method testing were lackluster, but the lessons learned from the attempt to classify a dataset with a large number of classes can be used moving forward in other research. The difficulties of classifying a multiclass dataset of 28 classes were shown clearly in this project, as the methods used were not accurate in classifying the Reddit comments from the dataset. For future research studies on this topic, alterations such as the number of classifications, the distribution of data among the classes, and the similarity between classes can be made to increase the effectiveness of future results using these three methods. Understanding and categorizing comments can allow a deeper understanding of how people use social media and what they choose to write on them.

REFERENCES

[1] - Özkent Y (2022) Social media usage to share information in communication journals: An analysis of social media activity and article citations. PLoS ONE 17(2): e0263725. https://doi.org/10.1371/journal.pone.0263725

[2] - Montoyo, A., MartíNez-Barco, P., & Balahur, A. (2012). Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments. Decision Support Systems, 53(4), 675–679. doi:10.1016/j.dss.2012.05.022.

[3] - Baali, M., Ghneim, N. Emotion analysis of Arabic tweets using deep learning approach. J Big Data 6, 89 (2019). https://doi.org/10.1186/s40537-019-0252-x

[4] - El-Hajj, Wassim. "Emotion Recognition from Text Based on Automatically Generated Rules" 2014.12.14.

[5] - Carlo Strapparava, and Alessandro Valitutti. "WordNet Affect: an Affective Extension of WordNet." LREC. Vol. 4. 2004.

[6] - Rustam F, Khalid M, Aslam W, Rupapara V, Mehmood A, Choi GS (2021) A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis. PLoS ONE 16(2): e0245909. https://doi.org/10.1371/journal.pone.0245909

[7] - Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. In Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002), pages 79–86. Association for Computational Linguistics.