ENHANCING EMOTION CLASSIFICATION

Embeddings Evaluation for Emotion Classification in Social Media Text

"The ability to understand and interpret emotions is the essence of emotional intelligence, and is fundamental to personal and professional success." - Daniel Goleman

"Understanding emotions is not just about self-awareness; it's also about empathy, compassion, and connection with others." - Brené Brown

"Emotions are the compass of life, and understanding them leads to a deeper understanding of ourselves and others." - Tara Brach

AGENDA

- ❖ Introduction
- ❖ Background
- ❖ Key Contributions
- ❖ Contribution 1: Data Collection to Expanding Emotional Dataset
  - ❖ Methods, Results and Analysis
- ❖ Contribution 2: Evaluation of Vectorization Techniques
  - ❖ Methods, Results and Analysis
- ❖ Evaluation
- ❖ Conclusion

INTRODUCTION

- Emotions are central to communication, and understanding them online provides insights into user sentiments, behaviors, and societal trends.

- **Approach:** Using Machine Learning, our system processes text data from platforms like Twitter and Reddit, aiming to classify underlying emotions conveyed in the text.

BACKGROUND

- **Prior Work:**
  - **Ji Ho Park's** paper introduced emotional word vectors (EVEC) learned from a convolutional neural network model with emotion-labeled Twitter data.
  - Emotion **labels were annotated using hashtags**, demonstrating the effectiveness of distant supervision.
- **Dataset Exploration:**
  - Utilized the **Kaggle Emotions dataset** containing English Tweets annotated with six fundamental emotions.
  - Dataset aligned with Park's methodology, **validating distant supervision** for emotion annotation.

- **Expanding Emotion Labels:**
  - Recognized the need to **account for additional emotions** prevalent in social media data.
  - Expanded dataset to include emotions like loneliness, jealousy, and awkwardness.
- **Vectorization Techniques:**
  - Experimented with **various vectorization** techniques to enhance model performance.
  - **Evaluated effectiveness** in capturing nuanced emotional expressions in text.

KEY STEPS:

1.**<u>Data Collection</u>:** Initially, we utilized the Kaggle dataset, Emotions [1], comprising English Tweets demonstrating six emotions. We expanded this dataset to include feelings of loneliness, jealousy, and awkwardness by scraping text posts from relevant subreddits.

2.**<u>Data Preprocessing:</u>** We **ensured compatibility** between the two datasets through preprocessing techniques.

3.**<u>Model Training</u>:** Classification **models were trained on the labeled dataset** to predict associated emotions or feelings based on text features.

4.**<u>Evaluation</u>:** Performance **evaluation of the system** was conducted using metrics such as accuracy and F1 score to assess classification effectiveness.

KEY CONTRIBUTIONS:

1. **Expanding Emotional Spectrum:**

   - *Recognizing the* ***limitations of conventional emotion labels****, we expanded the emotion spectrum to include additional feelings prevalent in social media discourse.*
   - *By* ***incorporating emotions*** *like loneliness, jealousy, and awkwardness, we aimed to provide a more nuanced understanding of human emotions expressed online.*

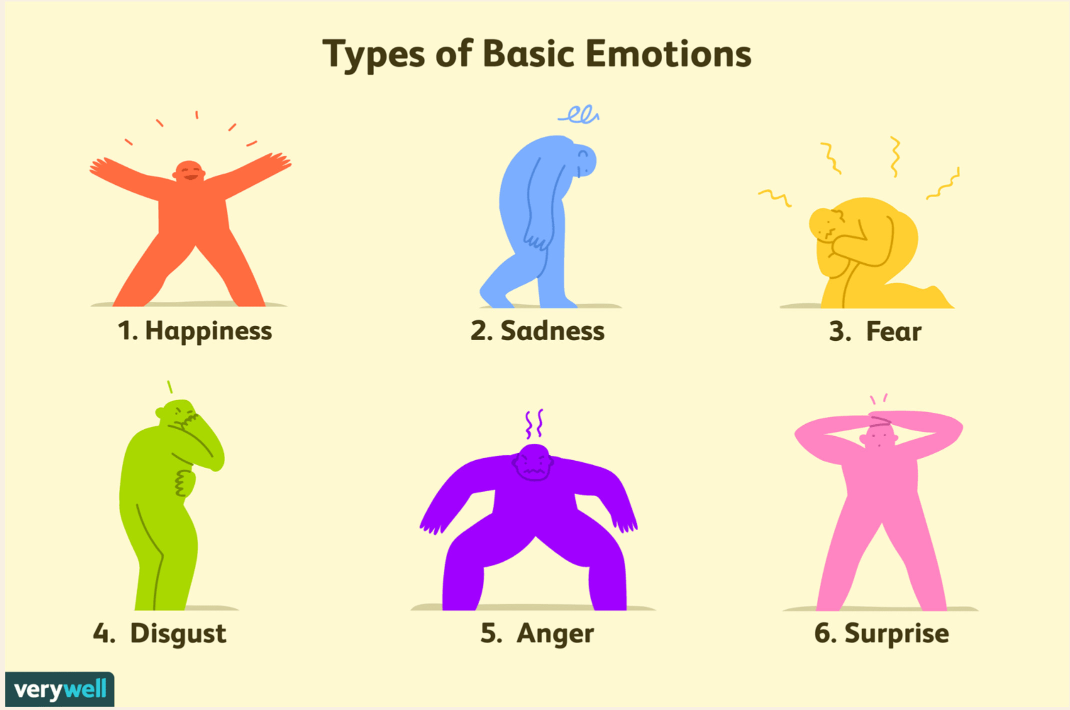
2. **Evaluation of Vectorization Techniques:**

   - *Conducted a comprehensive evaluation of various* ***vectorization techniques*** *to assess their effectiveness in enhancing emotion classification models.*
   - ***Identified factors contributing to the performance of these techniques*** *and aimed to improve the accuracy and robustness of emotion classification in real-world scenarios.*

CONTRIBUTION 1

Expanding Emotional Spectrum

ORIGINAL DATASET



Emotions is a dataset on Kaggle that provides us with a collection of English Tweets annotated with six fundamental emotions: anger, fear, joy, love, sadness, and surprise.

Data set insufficient to account for other more nuanced emotions.

Expanded labels by adding more feelings to classify like loneliness, jealousy, and awkwardness.

DATA COLLECTION

- **Utilized the Reddit API to scrape text posts** from relevant subreddits ("r/loneliness", "r/retroactivejealousy", "r/socialskills") to capture expressions of the identified feelings.
- Due to the challenge of accessing Twitter data due to changes in the Twitter API, we naturally gravitated towards Reddit, since **many users also express their emotions through posts** and the platform allows for a **larger character limit** so people are more likely to express their thoughts more.
- Similarly to how **distant supervision** was used in the Twitter data where the proxy for knowing whether a tweet talked about a particular emotion was the use of a hashtag, our proxy for the feeling was the subreddit name.

COMPATIBILITY VERIFICATION

- The twitter data we sampled for our models has as the maximum length of a **tweet 328 characters**.

- In order to make both the process of tokenizing and embedding text more computationally feasible and the results more comparable, we decided to only scrape comments for posts rather than the entire post body written by the original poster. This made our data collection process more efficient given that **scraping comments was faster** and there are more comments than posts.

- Additionally, we filtered comments so that we only **include comments that have 280 characters or less**. We ended up with a dataset containing **9000 values, 1000 texts per each emotion.**

	text	label
0	im sick with allergies and feeling horrible	0
1	i feel the music hit me in a vain attempt to k...	0
2	i feel terribly helpless and thus i am putting...	0
3	im feeling like ive missed you all this time s...	0
4	im finding it harder and harder every day to c...	0
...	...	...
8995	I thought I was the only one! I'm currently go...	8
8996	Sometimes it's a survival skill, health wise I...	8
8997	Quite dudes i hang shit on..usually it be shut...	8
8998	I could say I've been in similar situations wh...	8
8999	I'm 80 next month. All I ask of anyone I inter...	8

9000 rows × 2 columns

CONTRIBUTION 2

Evaluation of Vectorization Techniques

AIM TO ADDRESS KEY QUESTIONS:

1.How do different vectorization techniques—namely **TF-IDF, Word2Vec, BERT, DistilBERT, ELECTRA**—perform when integrated with traditional classifiers such as Support Vector Machines (SVM) and Logistic Regression?

2.Upon determining the most effective architecture, what **specific components or features of this architecture contribute to its superior performance** over other vectorization and classification combinations?

METHODS

The first analysis will be conducted on vectorization methods and traditional classifiers.
In this analysis, six feature extractor methods will be utilized:

Feature Extractor	Purpose
TF-IDF Vectorization	Measures word importance based on frequency in a document and rarity across all documents. Able to highlight distinctive words that may indicate specific emotions.
Word2Vec Vectorization	Represents words as vectors based on contextual similarity in a text corpus. Able to capture semantic relationships and potentially improve the detection of emotional subtext.

METHODS

Feature Extractor	Purpose
BERT Embeddings	BERT is a pretrained model that uses two objectives: masked language modeling (MLM) and Next Sentence Prediction (NSP). It creates deep contextual representations by considering both the left and right context of each word in a sentence.
DistilBERT Embeddings	DistilBERT was introduced as a "smaller, faster, cheaper, lighter" and distilled version of BERT. It is a Transformer model trained by distilling BERT base, which has 40% less parameters than BERT, runs 60% faster while preserving over 95% of BERT's performances.
ELECTRA Embeddings	Unlike BERT, ELECTRA uses a different pre-training objective called replaced token detection. This involves replacing a small fraction of input tokens with potential alternatives and training the model to distinguish the replaced tokens from the original ones. Thus, ELECTRA embeddings also capture contextual information but with a more efficient training process.

METHOD

- For classification tasks with Transformers models, such as text classification or sentiment analysis, it's **common to use the pooled representation for making predictions**. This is because the pooled representation is derived from the output of the last layer of the models.

- By using it, **we effectively summarize the entire input sequence in a single vector**, which can then be fed into a classifier as input to make predictions.

- The vectors produced by each of these extractors **will then be inputted into each of the following two classification models**, in order to create a combination between each feature extractor and each model.

Classifier	Purpose
Logistic Regression	An efficient baseline for binary and multiclass classification.
SVM	Able to handle high-dimensional data from text, making it robust in distinguishing between emotional categories, even when data is not linearly separable.

REUSLTS

❖ Each model was trained on the training set, comprising of **70% of the data**.

❖ **Hyperparameter tuning**, such as the number of maximum iterations for logistic regression, was conducted as needed.

❖ **Accuracy and F-1 score** will be used to evaluate the models.

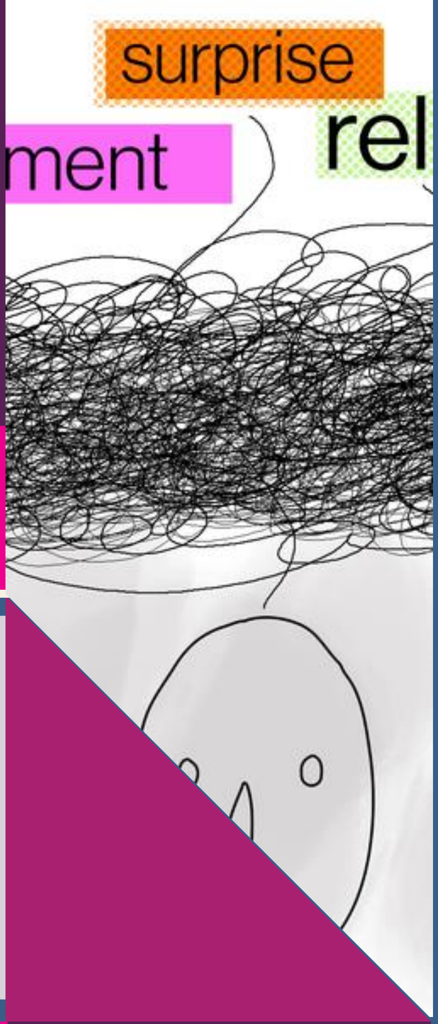
Our results are summarized next.

RESULTS

Feature Extractor	Classification Model	Accuracy	F1 Score
TF-IDF Vectorization	Logistic Regression	80.286	80.271
	SVM	79.762	79.650
Word2Vec Vectorization	Logistic Regression	67.571	67.602
	SVM	70.333	70.373
BERT Embeddings	Logistic Regression	85.333	85.585
	SVM	78.619	79.637
DistilBERT Embeddings	Logistic Regression	**86.333**	**86.370**
	SVM	**82.857**	**83.210**
ELECTRA Embeddings	Logistic Regression	85.095	85.209
	SVM	79.952	80.434

EVALUATION

- **DistilBERT + Logistic Regression outperforms all other models**:
  - Accuracy: 86.33%, F1 score: 86.37%.
  - Attributes success to DistilBERT's architecture.
- **Transformer models excel due to contextual understanding**:
  - Capture semantic nuances better than traditional methods.
  - Process entire sequences for contextual meaning.
- **Attention mechanism enhances emotion detection**:
  - Focuses on contextually significant words.
  - Helps differentiate similar yet distinct emotions.
- **BERT underperformed due to complexity and limited data**:
  - Hypothesized limited data affected BERT's generalization.
  - Overfitting evidenced by high training accuracy (82.63%).
- **DistilBERT balances performance and efficiency**:
  - Utilizes knowledge distillation from BERT.
  - Generalizes better with limited data, avoids overfitting.
- **ELECTRA's unique training objective boosts performance**:
  - Trained as a discriminator with a replaced token detection task.
  - Leverages more input, leading to efficient learning representations.
- **TF-IDF serves as a competitive baseline**:
  - Robust in capturing term frequency and document specificity.
  - Sparse and interpretable vectors reduce overfitting risks.



CONCLUSION

- **Computational Challenges**: Training Transformer models requires substantial computational resources, highlighting the importance of leveraging pre-trained models.
- **Ethical Considerations**: While social media data offers valuable insights, ethical concerns regarding privacy and data sensitivity must be addressed when training AI models for emotion detection.
- **Practical Applications**: Emotion detection models can **benefit various domains**, including customer feedback analysis, content creation, and virtual assistant interactions.
- **Future Directions**: Exploration **of ensemble methods, increased data sampling, and incorporation of more emotions** for robust analysis are potential avenues for improvement.

THANK YOU

Andrés Mondragón

- *amt00@seas.upenn.edu*

Ana Bernal

- *anbernal@seas.upenn.edu*

Leah Tesfa

- *ltesfa@seas.upenn.edu*