



# GoEmotion

Team 8: Salsabil Arabi, Ana Klabjan and Yu Zhang



Computer Sciences

SCHOOL OF COMPUTER, DATA & INFORMATION SCIENCES

UNIVERSITY OF WISCONSIN-MADISON

# Research Problem

- Classifying emotions in text
- GoEmotion: dataset of ~58,000 Reddit comments
- Labeled with 27 emotion categories
- Original Google research paper created a fine-tuned pre-trained BERT-base model that labels emotions to text
- Aim to address the limitations of the existing baseline model
- Aim to explore novel approaches to improve results

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  
neutral



# Significant To The Community

- Enhancing Mental Health Support
- Empathy in Artificial Agents
- Content Moderation and Harmful Behavior Detection
- Practical Applications in Sentiment Analysis
- Advancing the State of NLP



# Technical Challenges

- Ambiguity in Emotion Expression
- Cultural and Contextual Sensitivity
- Imbalanced Datasets
- Fine-Grained Emotion Recognition
- Temporal Dynamics
- User-specific Variability



# Recent Approaches And Their Limitations

- GoEmotions Baseline Model
  - Approach: BERT, dense layer, sigmoid
  - Limitations: baseline F1 score improvement needed
- BERT-Based Emotion Detection
  - Approach: Utilized BERT for ISEAR emotions.
  - Limitations: Complexity in BERT fine-tuning
- Neural Network Architectures
  - Approach: Used CNNs, LSTMs for dialogue
  - Limitations: Data Imbalance in neural network
- Dimensional Emotion Detection
  - Approach: Model predicted nuanced emotions.
  - Limitations: challenges in dimensional model mapping





# Our Experimental Methods

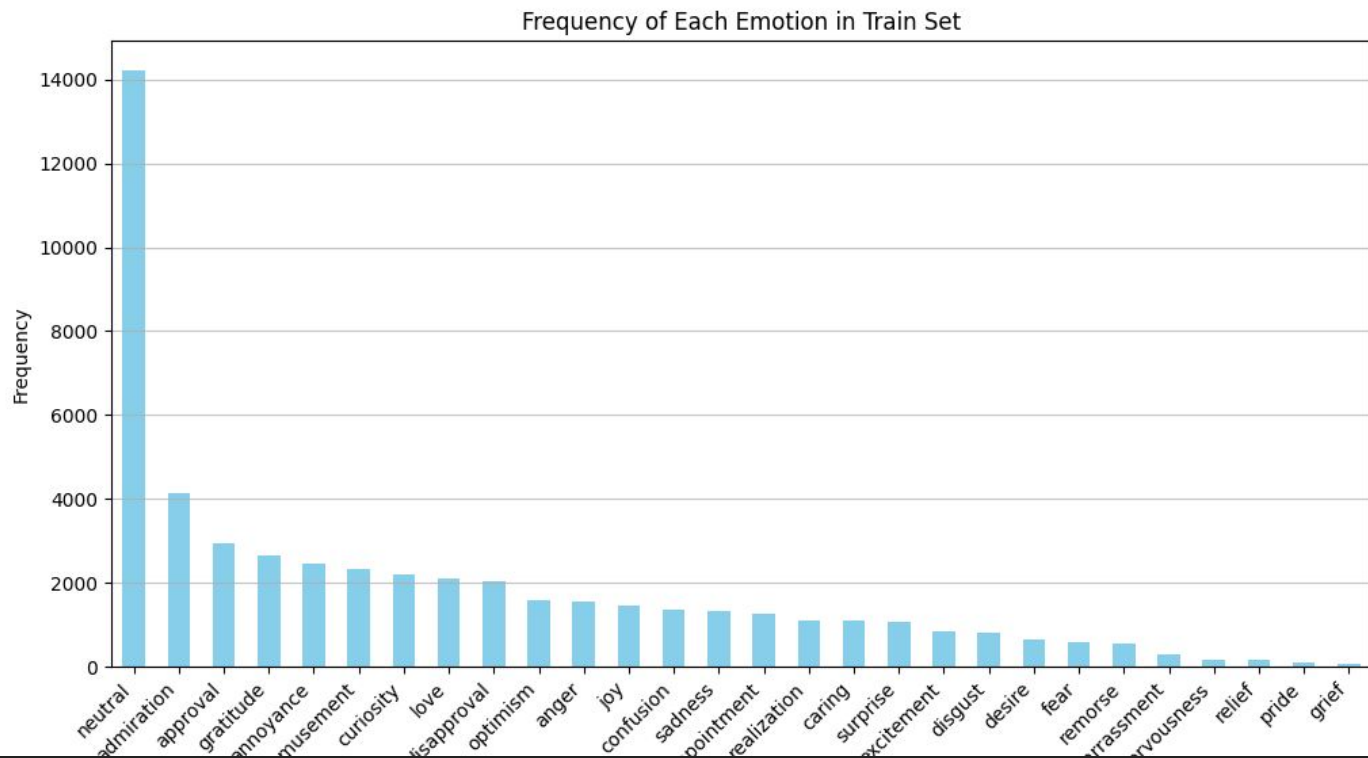
# Data Augmentation: Back Translation

- 43410 train data entries -> 78245 train data entries
- Labels stayed the same

Original	Back Translated
To make her feel threatened	So that she feels a threat
You are going to do the dishes now	You are going to make dishes now
Happy to be able to help.	I am glad that I could help.
It might be linked to the trust factor of your friend.	This may be due to the factor of your friend's trust.

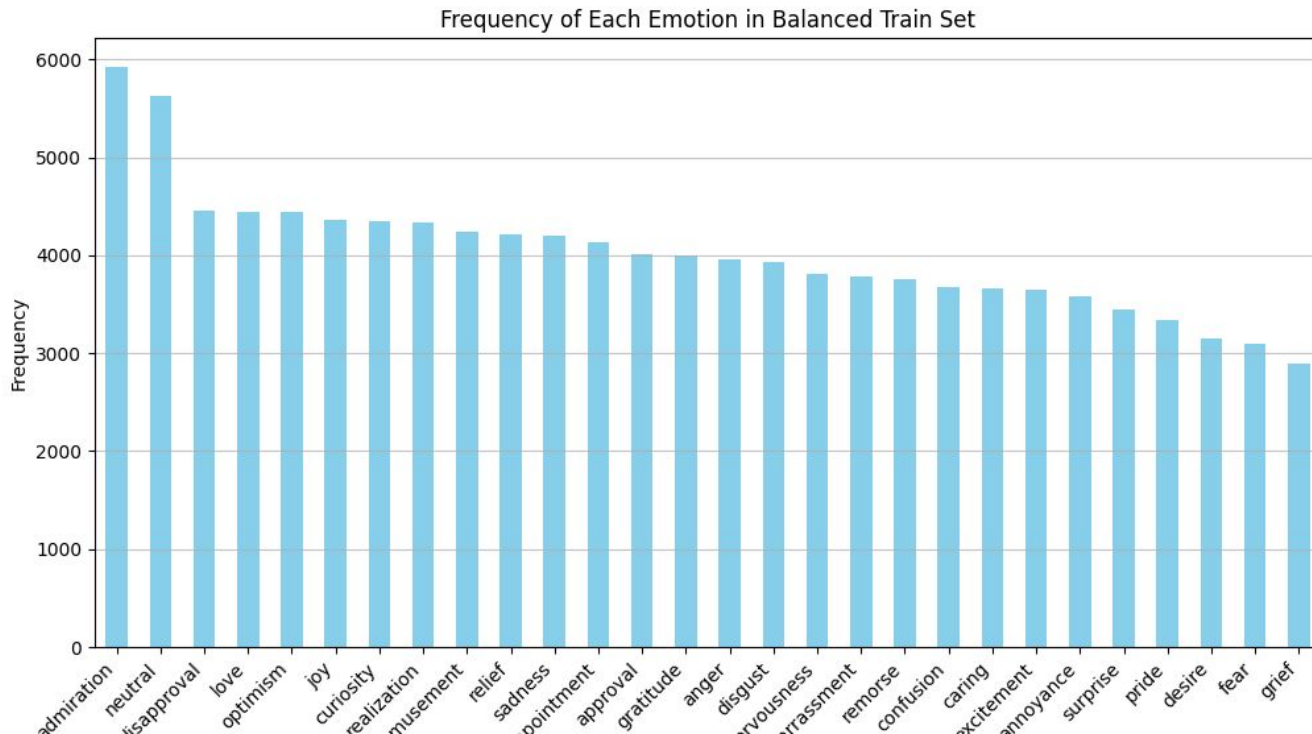
Table 1: The original text entries in GoEmotion and the corresponding new entries for train after being processed with back-translation using Russian.

# Data Augmentation: Resampling

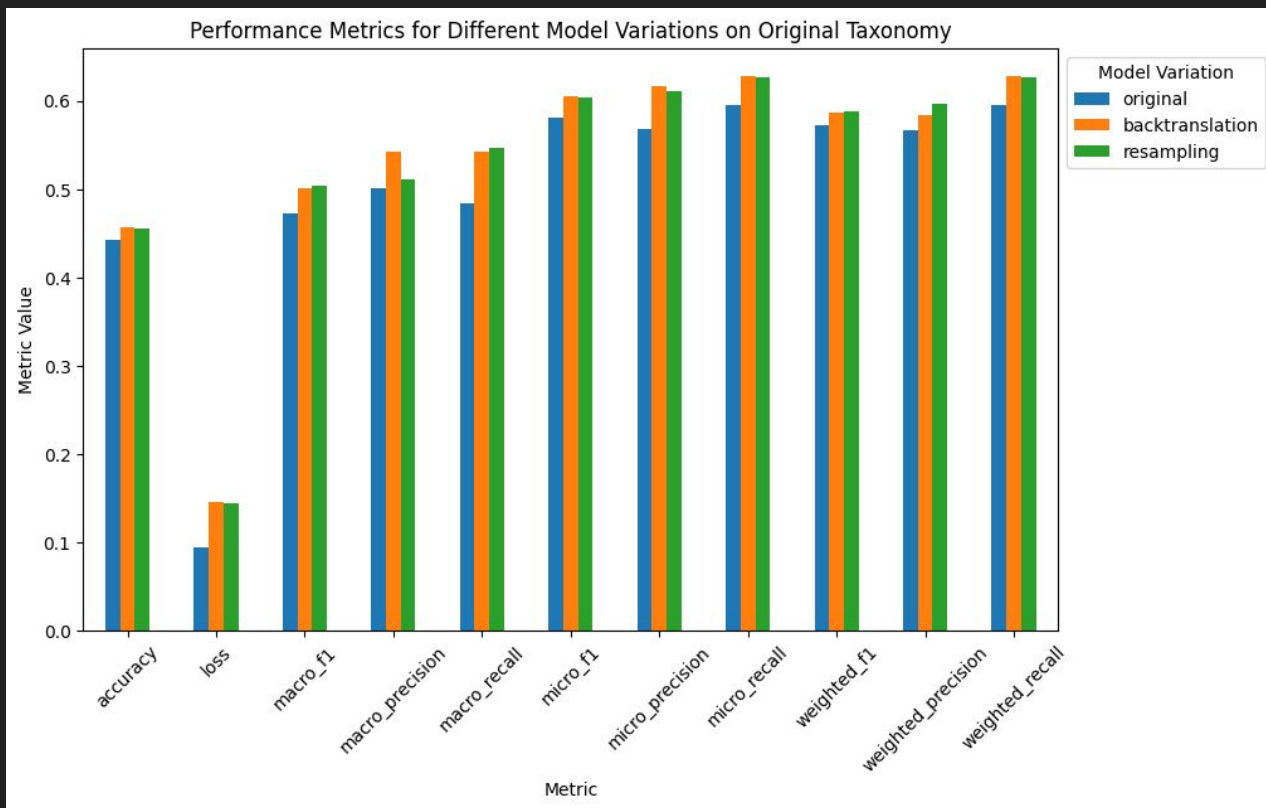




# Data Augmentation: Resampling



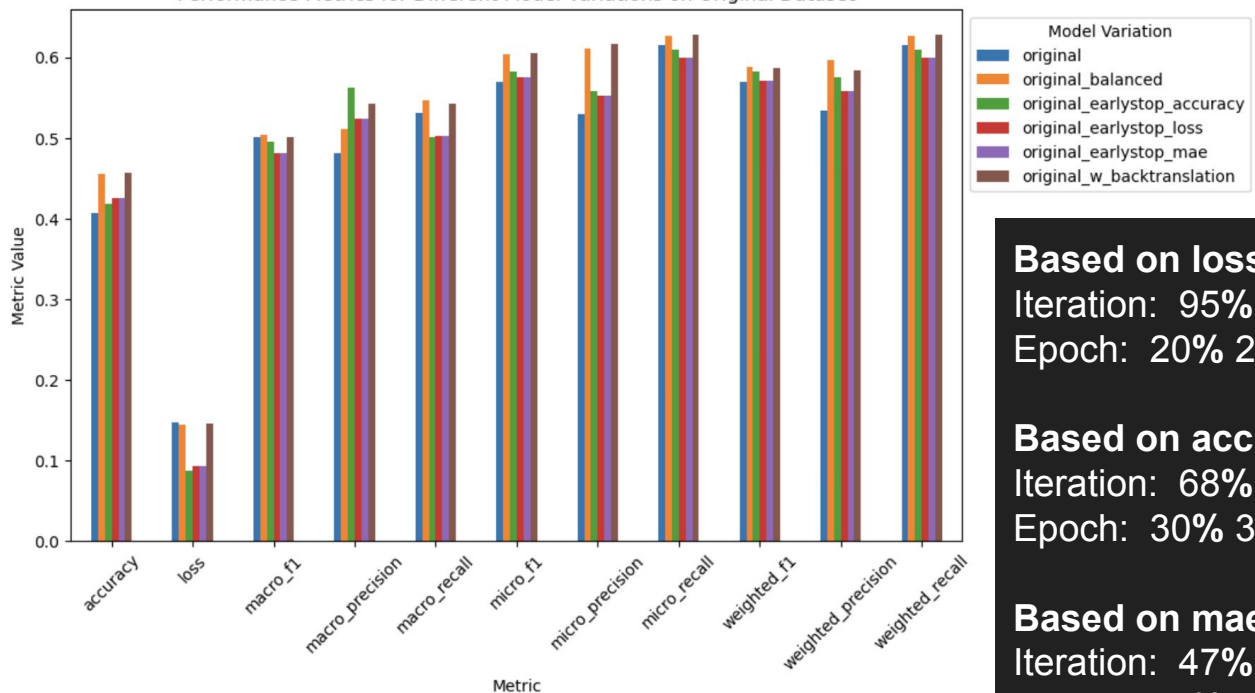
# Results: Data Augmentation





# Results: Early Stop

Performance Metrics for Different Model Variations on Original Dataset



## Based on loss:

Iteration: 95% 2571/2714 [02:56<00:09, 14.58it/s]

Epoch: 20% 2/10 [08:57<35:51, 268.98s/it]

## Based on accuracy:

Iteration: 68% 1857/2714 [02:02<00:56, 15.21it/s]

Epoch: 30% 3/10 [10:40<24:55, 213.63s/it]

## Based on mae:

Iteration: 47% 1285/2714 [01:30<01:41, 14.13it/s]

Epoch: 10% 1/10 [04:27<40:11, 267.92s/it]

# Weighted Loss/Density Loss/Multi-label Soft Margin Class

- Cost sensitive Learning
  - Inverse class frequency
- Multilabel Soft Margin Loss
- Dense + Dropout
- LSTM + Dropout

