CS 626 - Speech, Natural Language Processing, and the Web

Final Project

Emotion Classification And Sentiment Analysis Using Retrofitted Supervised Contrastive Learning

Group ID - 16

Team Members :

Leader : Aryan Khilwani (24M0781)
MTech-CSE 1st Yr

Member 2 : Rajshekhar K (24M0793)
MTech-CSE 1st Yr

Member 3 : Aditya Hol (24M0778)
MTech-CSE 1st Yr

Member 4 : Sejal Kadamdhad (24M0806)
MTech-CSE 1st Yr

Date - 29/11/2024

Problem Statement

- **Objective:** Given a sentence, predict the conveyed emotion and sentiment, along with the corresponding Valence, Arousal, and Dominance (VAD) values.
- **Input:** A text sentence that expresses a sentiment that is either positive or negative.
- **Output:**
    - Emotion: A label indicating the primary emotion expressed (e.g., happiness, sadness, anger, fear, surprise, disgust).
    - Sentiment: A label indicating the overall sentiment of the sentence (e.g., positive or negative).
    - VAD Values: Numerical values for valence, arousal, and dominance, typically ranging from 0 to 1.

- **Example:**
  - Input: “I am very happy.”
  - Output:
    - Emotion: Joy
    - Sentiment: Positive
    - Sentiment Probability: 0.99
    - Valence: 0.99
    - Arousal: 0.85
    - Dominance: 0.81

Motivation

Motivation Behind Our Choice:

- **Relevance of Emotion Recognition:**
  - Understanding emotions in text is essential for applications in customer service, mental health, and social media.
- **Challenges with Current Models:**
  - While models like BERT and RoBERTa are efficient, they often fail to capture the nuanced emotional content in language, leading to inadequate emotion representation.

- **Need for Improvement:**
  - Observations show that embeddings for similar emotions do not cluster as expected, indicating a significant gap in emotion detection.
- **Innovative Approach:**
  - Our team aimed to enhance these models through Retrofitted Supervised Contrastive Learning, bridging the gap between general language understanding and emotional context.

We are motivated by the need for better emotion classification in NLP, aiming to contribute innovative solutions to this critical area.

Literature Review

Emotion Detection and Sentiment Analysis in NLP

- **Introduction:** Emotion detection and sentiment analysis play vital roles in understanding human communication, especially in applications like social media monitoring and customer feedback analysis.
- **Machine Learning Advancements:** Techniques such as Support Vector Machines (SVMs) and deep learning models (RNNs, CNNs) improved classification performance, yet often required extensive feature engineering (Manning et al., 2008; Zhang et al., 2018).
- **Limitations of Pre-trained Models:** While transformer models like BERT achieved state-of-the-art results (Devlin et al., 2019), they often fail to represent emotional content accurately, as their embeddings do not consistently reflect semantic similarities among texts with the same emotional context (Khan et al., 2021).

Retrofitted Supervised Contrastive Learning

- **Need for Enhanced Models:** There is a growing recognition that PLMs need augmentation to better capture affective aspects of language (Kumar et al., 2020).
- **Retrofitting Techniques:** Recent studies suggest combining retrofitting techniques with supervised contrastive learning to refine embeddings, making them more sensitive to emotional nuances (Khosla et al., 2020).
- **Benefits:** This approach aims to improve the performance of emotion classification and sentiment analysis, enabling more effective applications in areas such as empathetic AI and sarcasm detection.
- **Conclusion:** Enhancing PLMs through retrofitting can significantly advance the field of affective NLP, making models more adept at understanding and responding to human emotions.

Datasets

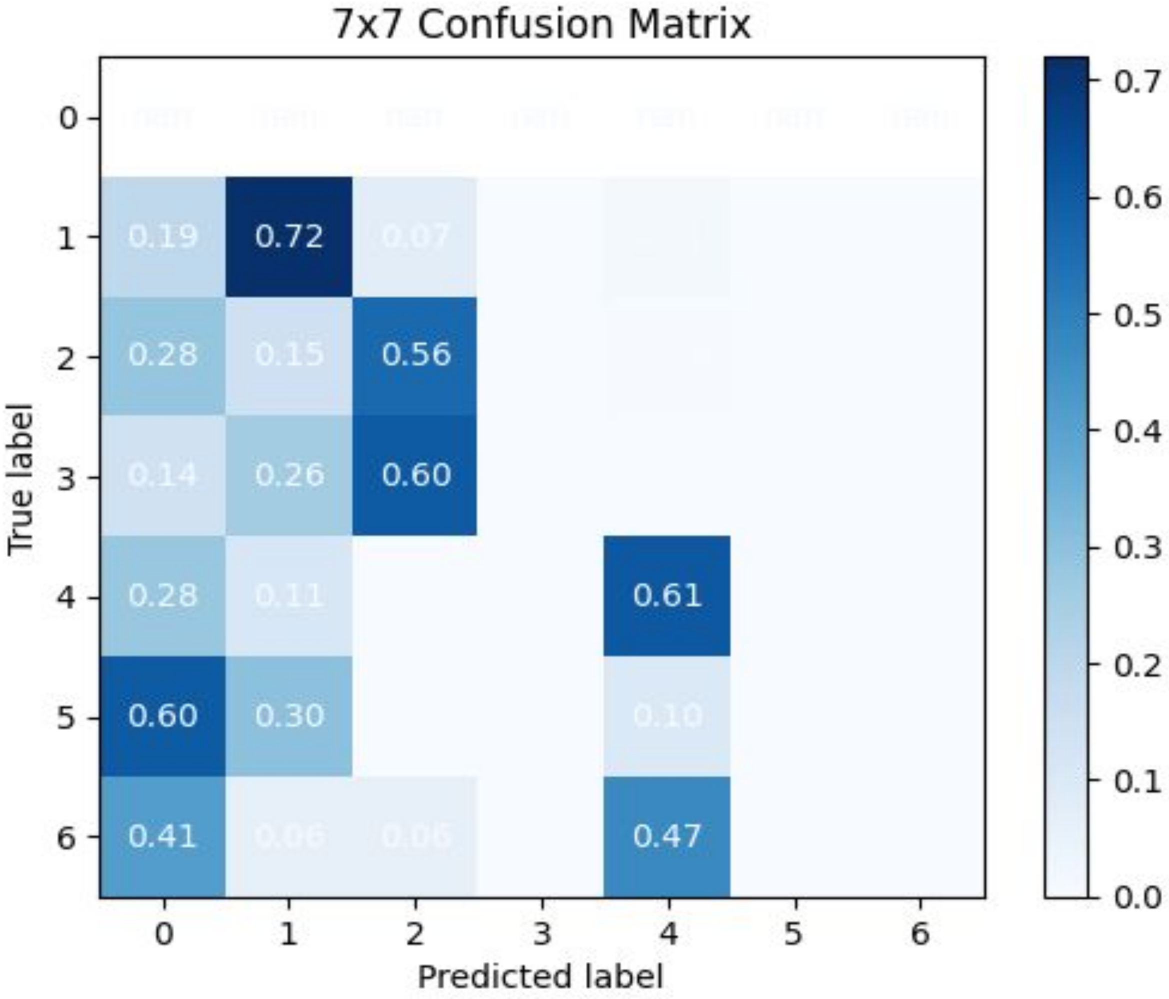
1. **Go_Emotions Dataset:**
    a. Description: A large-scale dataset designed for emotion recognition, containing over 58,000 annotated examples sourced from online discussions.
    b. Labels: Includes 28 emotion labels, reflecting a wide range of human emotions and sentiments.
    c. Custom Adaptation: We further refined this dataset to focus on 7 key emotions based on Ekman's emotion model: Joy, Anger, Sadness, Surprise, Fear, Disgust, Neutral.
    d. Purpose: This custom set allows for focused emotion classification relevant to our research objectives.
2. **Stanford Sentiment Treebank (SST-2):** A dataset specifically designed for sentiment analysis, featuring fine-grained sentiment labels (positive and negative) derived from movie reviews.
3. **SumEval Dataset:** A sentiment analysis dataset that includes various summaries of text documents, annotated for sentiment.

Method/Technique

1. **Retrofitted Supervised Contrastive Learning:**
    a. We employed a retrofitting technique that enhances pre-trained language models (PLMs) with a supervised contrastive learning approach to better capture emotions and sentiments in text.
2. **Model Fine-Tuning:**
    a. Pre-training: Utilize existing PLMs trained on large datasets to leverage their general language understanding capabilities.
    b. Fine-tuning: Adapt these models on the custom Go_Emotions dataset (7 emotion labels) and sentiment datasets (SST-2, SumEval) using a supervised learning framework.
3. **Contrastive Learning Technique:**
    a. Mechanism: Pairing similar emotion-label text examples to create a contrastive loss function that minimizes the distance between embeddings of the same class while maximizing the distance between different classes.

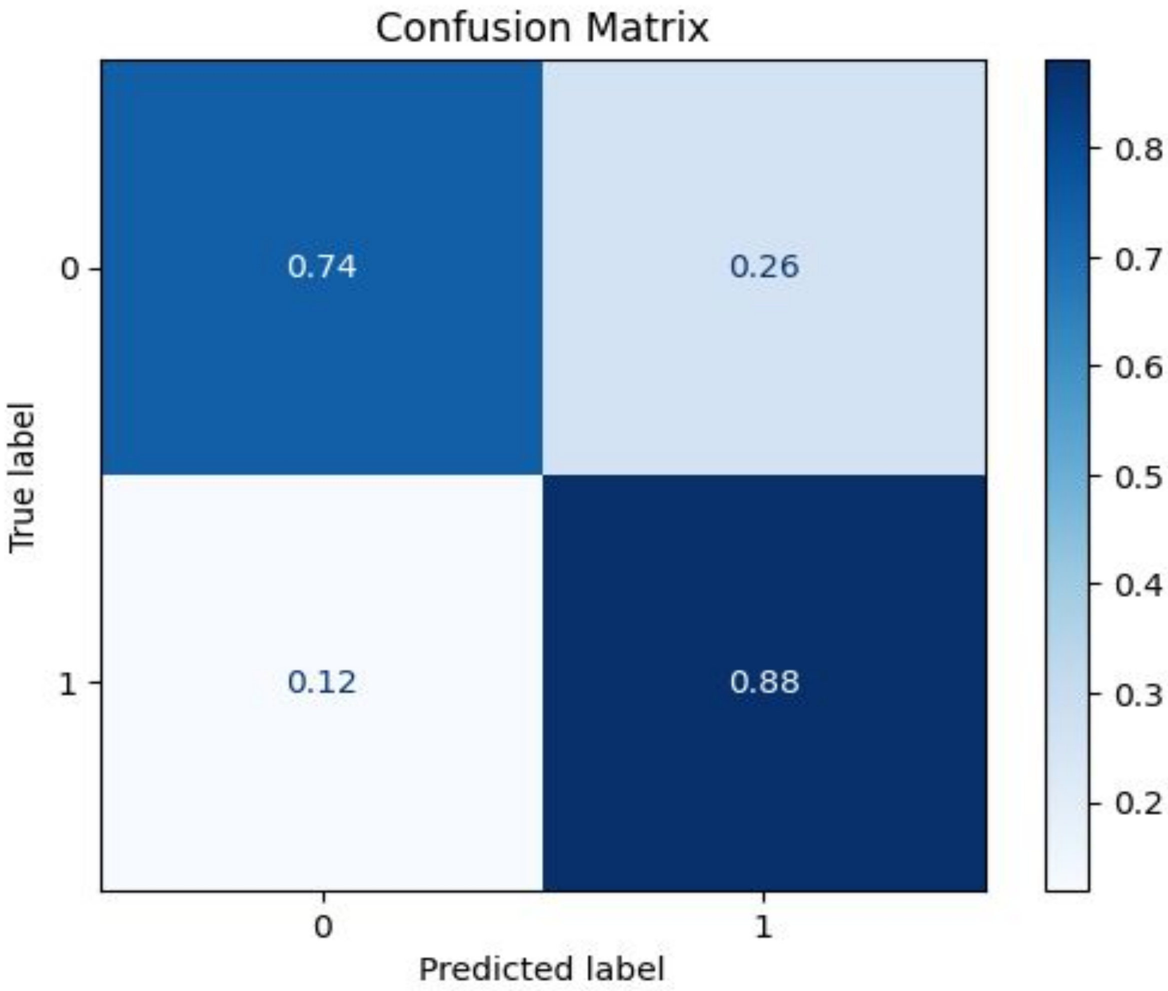
Results

F1 - score : 46 %
Accuracy - 57 %



Emotion Classification(7 labels)
Go_emotion dataset

F1 - score - 82 %
Accuracy - 81.3%



Sentiment Analysis(2 labels)
SST-2 dataset

Analysis

We evaluated our model's sentiment analysis performance by comparing it with ChatGPT on different types of sentences. Below are the key findings:

- **Performance on Clear Sentiments:**
  - For sentences with explicit and unambiguous sentiment, such as "I am very happy" or "I am very sad", our model performed exceptionally well.
  - It assigned probabilities close to 1 for positive and 0 for negative sentiments, demonstrating high confidence and accuracy for such cases.
- **Challenges with Ambiguity:**
  - The model struggled to correctly classify sentences with ambiguous sentiment. For example, sentences with mixed emotions or unclear tonal indicators often led to less accurate predictions.

- **Difficulty with Sarcasm:**
    - Sarcastic sentences, especially those where the conveyed sentiment is the opposite of what appears at face value (e.g., "Oh, great, another traffic jam!"), posed a significant challenge.
    - The model often failed to identify the sarcasm and instead predicted sentiment based on the literal words, resulting in incorrect classifications.
- **Comparison with ChatGPT:**
    - ChatGPT, with its advanced contextual understanding, performed better in handling ambiguous and sarcastic sentences, showcasing its ability to infer deeper meanings from context.
    - This gap highlights the limitations of our model's current architecture in capturing complex linguistic phenomena like sarcasm.

Error Analysis

Sentences	Predicted	Actual
The food tasted great, but the portions were too small.	Positive	Negative
I’m happy about the promotion, but sad to leave my old team.	Positive	Negative
Oh no, another five-star review for your book. How terrible!	Negative	Positive
This movie was an absolute masterpiece—if you enjoy being bored.	Positive	Negative
Wow, I just love being stuck in traffic for hours.	Positive	Negative
Must be awful to have everyone love your cooking so much!	Negative	Positive
Oh great, another software update to ruin my day.	Positive	Negative

Improvements over the paper

Overview of Our Work

- **Adopted Methodology:**
  - Implemented the constraint learning techniques proposed in the paper to fine-tune an existing BERT model.
  - Used the GoEmotions dataset for training the retrofitted BERT for emotion classification.
- **Extended Functionality:**
  - Added two neural networks on top of the fine-tuned BERT model:
    - One for sentiment analysis (positive/negative).
    - Another to compute Valence-Arousal-Dominance (VAD) values for given sentences.
    - Incorporated the suggestions from the "Future Work" section of the paper.

Key Insights and Challenges:

- **Paper's Concern:**
  - The paper was skeptical about using only the Go_Emotions dataset due to potential performance limitations outside the dataset.
- **Our Findings:**
  - The model performed well within the GoEmotions dataset but struggled with:
  - Complex sentences (e.g., sarcasm, ambiguous tone).
  - Out-of-distribution data (unseen phrases or contexts).
- **Observation:**
  - Sole reliance on Go_Emotions for fine-tuning limits generalizability and robustness.

Learning

- **Understanding Retrofitting with Constraint Learning**
    - Gained in-depth knowledge about retrofitting pre-trained language models (PLMs) like BERT using constraint learning techniques for emotion classification.
- **Challenges of Dataset Dependency**
    - Recognized the limitations of relying solely on the Go_Emotions dataset, as it struggled with:
    - Complex sentence structures (e.g., sarcasm, mixed emotions).
    - Generalization to unseen or domain-specific texts.
- **Integration of Multi-Task Learning**
    - Successfully extended the BERT model with two neural networks:
    - One for sentiment analysis (positive/neutral/negative).
    - Another for predicting Valence-Arousal-Dominance (VAD) values.