SentiMind: A Emotion Classifier

Zixi Qu, Jinyan Yi, Kangzhi Gao

UNIVERSITY OF TORONTO

About

- 28 emotion classes. Can detect multiple emotions in one sentence
- Implemented using both CNN approach and RNN (LSTM) approach, to compare which is more suitable for this task
- Input is preprocessed with a pre-trained encoder transformer model.

Dataset

- Go_emotions from Google. Available on Hugging Face
- Include 58k data-points. Split into 60:20:20
- Provides following fields:
  1. Text: Input to our system, a sentence to be detect emotions.
  2. labels: a list of one or more emotions in the sentence.
- Produced by 3 English-speaking crowdworkers

Sample Examples From the Dataset

1. Text: `Three words, no subtlety. "Dude. Stop. Seriously."`
   Labels: `[3]` (Annoyance)
2. Text: `This report is going to be bullet proof`
   Labels: `[20, 27]` (Optimism, Neutral)
3. Text: `Lmaooo I'm sooo gonna use "As asap as possible."`
   Labels: `[1]` (Amusement)
4. Text: `Thank you so much! I love Germany! I was in Berlin 2 years ago for research work.`
   Labels: `[15, 18]` (Gratitude, Love)

Pre-trained Encoder Transformer

A number of model tested,

Model	Top Accuracy	Training Speed
roberta-base	38%	8it/s
albert-base-v2	40%	4it/s
distilroberta-base	24%	10.75it/s
huawei-noah/TinyBERT	38%	41.80it/s

Table 1. Encoder Model Performance

Architecture

The process of our project is as following:

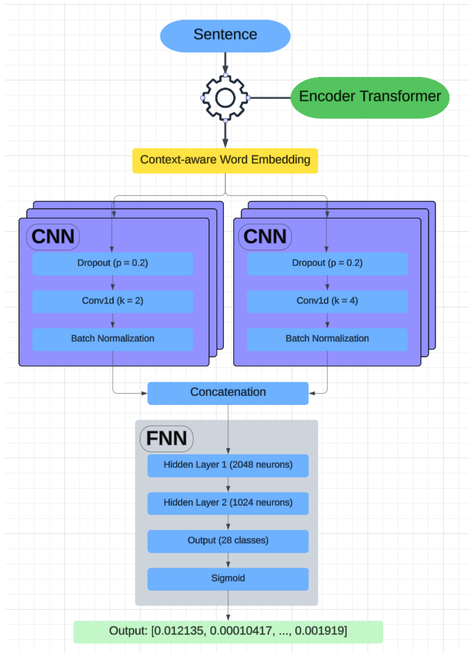

Figure 1. CNN Approach

Alternative Approach: Replace CNN with LSTM

Replace the CNN layer with LSTM model

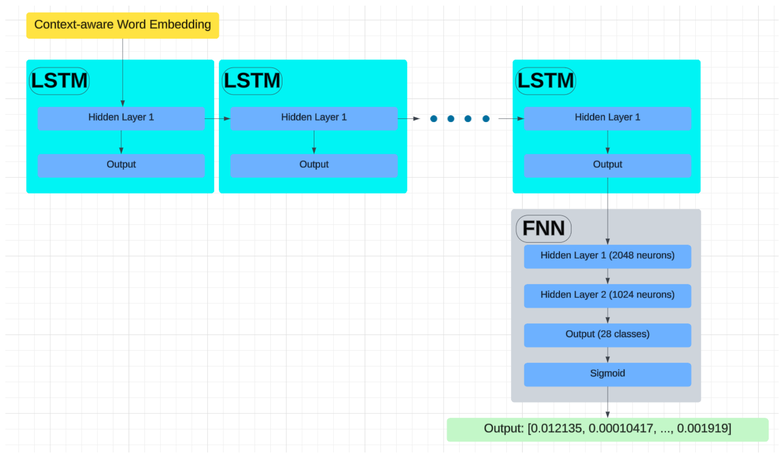

Figure 2. LSTM Approach

Accuracy Metric: Jaccard Index

The **Jaccard Index** is defined as:

$$J(Label, Prediction) = \frac{|Label \cap Prediction|}{|Label \cup Prediction|}$$

Where:

- $Label$ and $Prediction$ are two sets of emotion labels.
- $|Label \cap Prediction|$ is the size of the intersection of $Label$ and $Prediction$.
- $|Label \cup Prediction|$ is the size of the union of $Label$ and $Prediction$.

Architecture Performance

Performance for various tested architecture

Model	Train Accuracy	Test Accuracy
1-layer CNN	87.72%	40.38%
2-layer CNN	44.71%	37.78%
1-Layer LSTM	54.02%	38.46%
3-Layer LSTM	54.85%	37.09%

Table 2. A table caption.

Discussion

- CNN-based Architecture is more suited for this task
- Increase the depth of both models does not improve the performance for this task.
- Among the decoder models, albert-base-v2 has a slight better performance.

References

[1] Dana Alon and Jeongwoo Ko.
Goemotions: A dataset for fine-grained emotion classification, October 28 2021.
Google Research Blog.

[2] Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi.
GoEmotions: A Dataset of Fine-Grained Emotions.