

NLP Project Report

Introduction: Motivation

Understanding the underlying causes of emotions expressed in conversations is crucial for various applications, from improving dialogue systems to enhancing mental health support. Emotion-cause pair extraction serves as a pivotal step in this endeavor. By identifying the precise textual spans that trigger specific emotions, we can gain deeper insights into the dynamics of human interaction. This project aims to contribute to this field by developing a robust system capable of automatically extracting emotion-cause pairs from conversational data. By doing so, we not only facilitate the analysis of emotional expressions but also pave the way for more nuanced dialogue understanding and emotional intelligence in artificial systems. This task holds promise for advancing both research and practical applications in fields such as natural language processing, affective computing, and human-computer interaction.

Related Work

Literature Review of Emotion-Cause Pair Extraction in Conversations: A Two-Step Multi-Task Approach

The paper delves into the methodology of identifying emotions and their corresponding causes in textual conversations. This approach typically involves a dual-stage process where emotions and causes are first identified independently and then paired together. Existing research in this area often employs advanced techniques such as attention networks, dual-questioning mechanisms, and context awareness to enhance the accuracy and efficiency of emotion-causing pair extraction. By breaking down the extraction process into distinct stages, this multi-task approach aims to improve the precision of identifying emotional cues and their underlying triggers within conversational contexts, thereby contributing to a more nuanced understanding of emotional dynamics in text.

Literature Review of ECPEC: Emotion-Cause Pair Extraction in Conversation

The literature review on Emotion-Cause Pair Extraction in Conversations (ECPEC) focuses on the extraction of emotion-cause pairs within conversational text. This area of research has gained significant attention due to its relevance in understanding emotional dynamics in various domains. Scholars have proposed innovative approaches, such as neural networks, emotion-aware word embeddings, and Bi-LSTM layers, to enhance the accuracy of identifying emotional cues and their underlying causes. By combining previous research with novel methodologies, the ECPEC literature review aims to advance the field of Natural Language Processing by providing insights into the nuanced relationship between emotions and their triggers in textual conversations.

Methodology

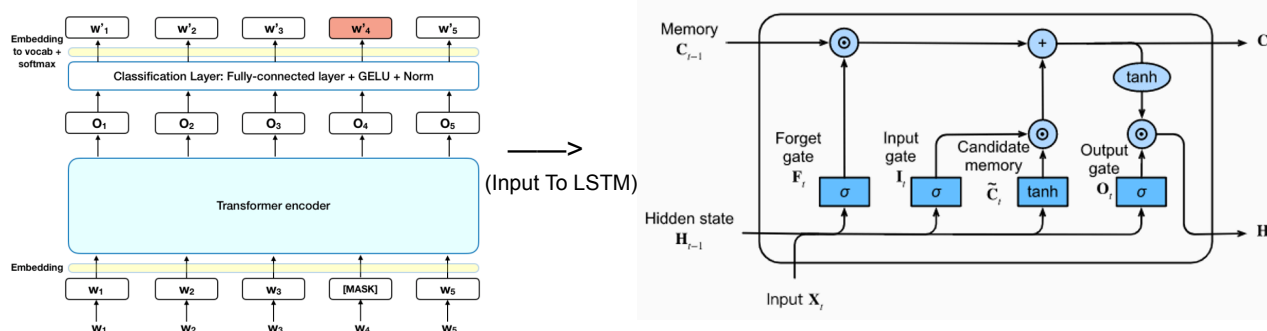
We have also utilized a two layer phase for our emotion-cause pair extraction task. We have devised a model which based on contextual learning from the next utterance predicts the emotion associated with the sentence. The dataset being derived from the TV show “Friends” has a lot of ups and downs in the emotional context of the sentence. These changes in the emotions are further studied by a distinct model based on its specialization and identifies parts of utterances which act as trigger for the emotion cause or the emotion flip.

- First Phase

The first phase of the task is used to generate emotion for each of the utterances in the entire conversation. This is done by first generating a sentence embedding using the BERT Model. These embeddings contain all the information that was present in the sentence using a 768 dimensional vector. Now each utterance is dependent on the previous sentence spoken and therefore there is a requirement of contextual learning and memory to appropriately judge the emotion of the sentence. For this task we further process the word embeddings by feeding them into a LSTM Model. This LSTM model identifies the dependencies between the sentences and finally outputs the emotion for each of the utterances.

- Second Phase

The second phase begins with the emotions generated by the first phase. This task involves first identifying the change in the emotion from neutral to some other emotion depicting utterance of a sentence which has caused this change of emotion. We identify the responsible part of the sentence by fine tuning the BertForTokenClassification approach. Basically we create 6 different models, one for each emotion, to identify the triggers that have caused the respective emotion. We employ the token classifier as we view the span prediction task under the lens of modifier NER, where the 0 label marks as non-trigger and 1 label as the trigger for that certain emotion.



Dataset

The dataset is based on TAFFC 2022 paper: Multimodal Emotion-Cause Pair Extraction in Conversations. It contains conversation from the popular TV show "Friends".

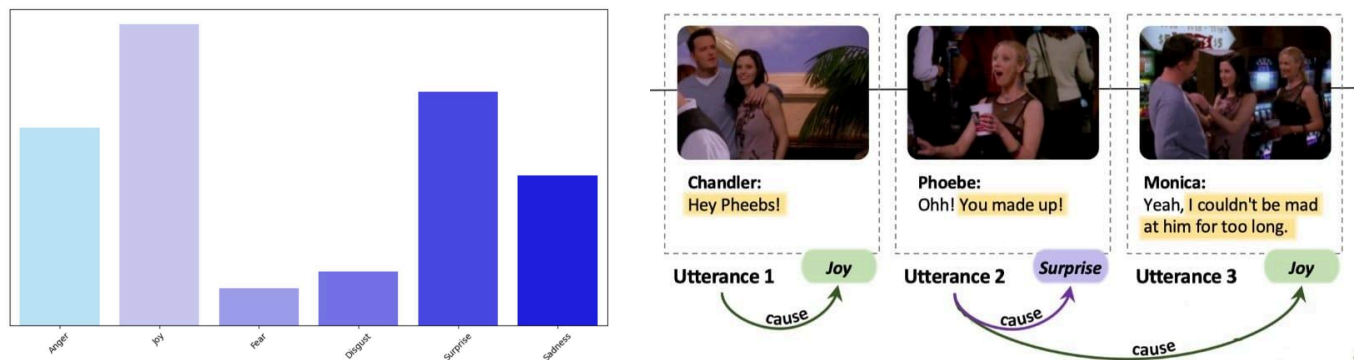


Figure: Emotion classes distribution per samples and the emotion cause extraction task.

Observations

- One of the critical aspects confirming the validity of existing theories is the influence of bias within the training data. This is evident through the varying F1 scores obtained for different emotions, wherein a higher volume of training data correlates with better F1 scores in general. Consequently, the observed lower F1 score for the "fear" emotion underscores the impact of this bias.
- Furthermore, an intriguing revelation post phase 2 training is the inadequacy in describing emotions such as "sadness" and "anger" within the dataset. Despite having a greater number of data points compared to "disgust," the model struggled to grasp these emotions effectively. There are two plausible scenarios to consider in explaining this phenomenon:
 - 1) Firstly, the dataset from the "Friends" TV show exhibits a notable bias towards emotions like "joy" and "surprise," a characteristic evident upon viewing the show. Consequently, this bias presents a challenge for generic BERT models, resulting in diminished performance for both "sadness" and "anger."
 - 2) Alternatively, it's plausible that the textual expression of emotions such as "sadness" and "anger" lacks the nuanced portrayal facilitated by multimodal data. Unlike "joy" and "surprise," which can be conveyed through punctuation marks like exclamation points in conjunction with a diverse lexicon, the textual medium may inherently limit the effective communication of "sadness" and "anger."

Experimental Setup

Phase	Model	Learning Rate	Batch Size	No. Of Epochs	Activation Function	Optimizer
1	LSTM	0.001	35	10	Softmax	Adam
1	BERT Model	2e-5 to 5e-5	35	10	Softmax	Adam
2	BERT for Token Classification	2e-5	16	10	argmax	AdamW

Result & Findings

We have proposed a novel method for emotion cause extraction in conversation using an approach similar to NER according to our knowledge. We have modified the training data to form binary encoded labels with 0 for a trigger and 1 for non trigger which is being fed into a BertForTokenClassification model, where we are without freezing it, fine-tuning it.

Emotions	Score
Overall	0.42
Anger	0.33
Fear	0.12
Disgust	0.28
Sadness	0.39
Surprise	0.40
Joy	0.45
Neutral	0.97

Model	Score
Overall	0.65
Fear	0.51
Sadness	0.62
Joy	0.58
Anger	0.60
Disgust	0.52
Surprise	0.59

Figure: Phase 1 and Phase 2 F1 Scores

Conclusion

The project aimed to extract emotion-cause pairs from conversations for diverse applications. We have utilized a two layer phase for our emotion-cause pair extraction task. We have devised a model which based on contextual learning from the next utterance predicts the emotion associated with the sentence. The dataset being derived from the TV show “Friends” has a lot of ups and downs in the emotional context of the sentence. Our two-layered approach addressed context and specialized models. Despite challenges like biased data and expressing certain emotions through text, our results underscored the importance of data distribution and innovation in advancing emotional intelligence and NLP.

Future Scope

The low F1 scores for the entire task gives a clear cut indication to improve the existing models and architecture. Utilizing multi-modal inputs and more efficient inference time are vital for the real time applications of these models in enhancing human computer interaction and mental health support. The following two tasks are worthy for further study.

- How to effectively model the speaker's relevance for both emotion recognition and cause extraction in conversations?
- How to utilize the external commonsense knowledge to bridge the gap between emotion and cause that are not explicitly reflected in the conversation?

References

Literature Review of Emotion-Cause Pair Extraction in Conversations: A Two-Step Multi-Task Approach : Lee, Jaehyeok & Jeong, DongJin & Bak, JinYeong. (2023).

Literature Review of ECPEC: Emotion-Cause Pair Extraction in Conversation: Wei, Li & Li, Yang & Pandelea, Vlad & Ge, Mengshi & Zhu, Luyao & Cambria, Erik. (2022).