

# Multi-Modal Emotion Cause Detection in Classroom

Understanding Feelings Through Artificial Intelligence

Team EMOED

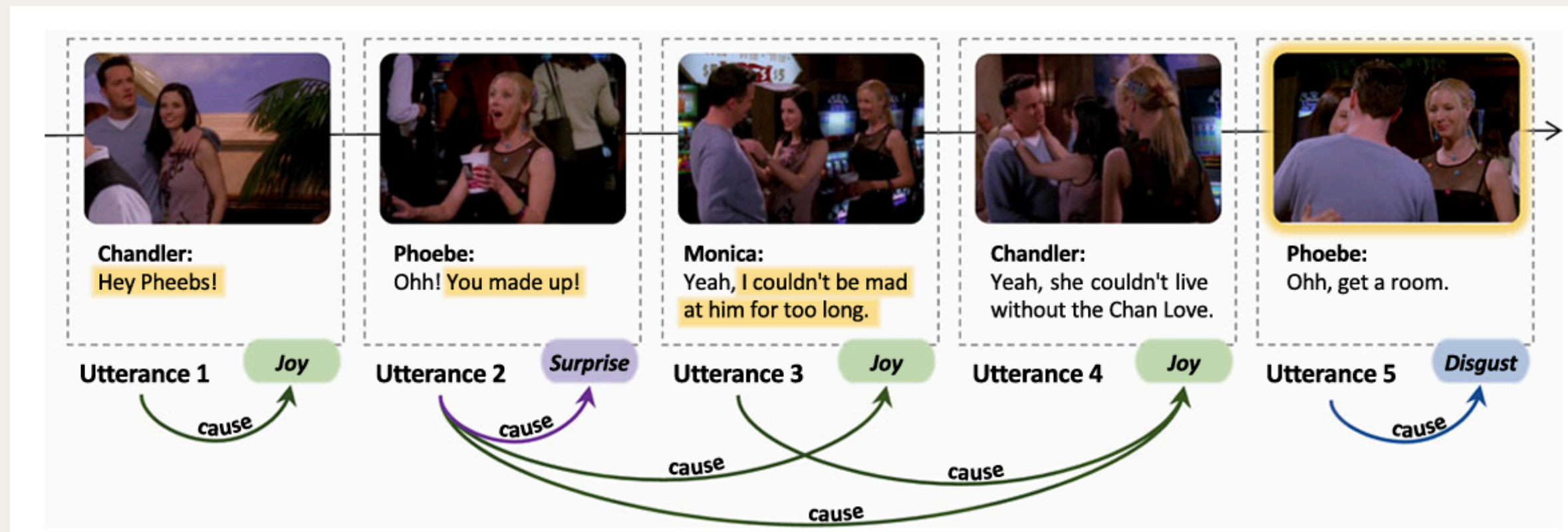
Ganta Naga Jaswanth  
Chelluboina Siri

DS510:AI ML LAB



# Introduction and Task Definition

- Emotion-Cause Pair Extraction (ECPE): The goal is to identify pairs of utterances in a conversation where one utterance expresses an emotion and another provides its cause. This involves analyzing data across three modalities: text, audio, and video.
- Example:  
In a conversation, if "U3" expresses joy and "U2" explains the reason behind it, the identified pair would be `(U3, U2)`



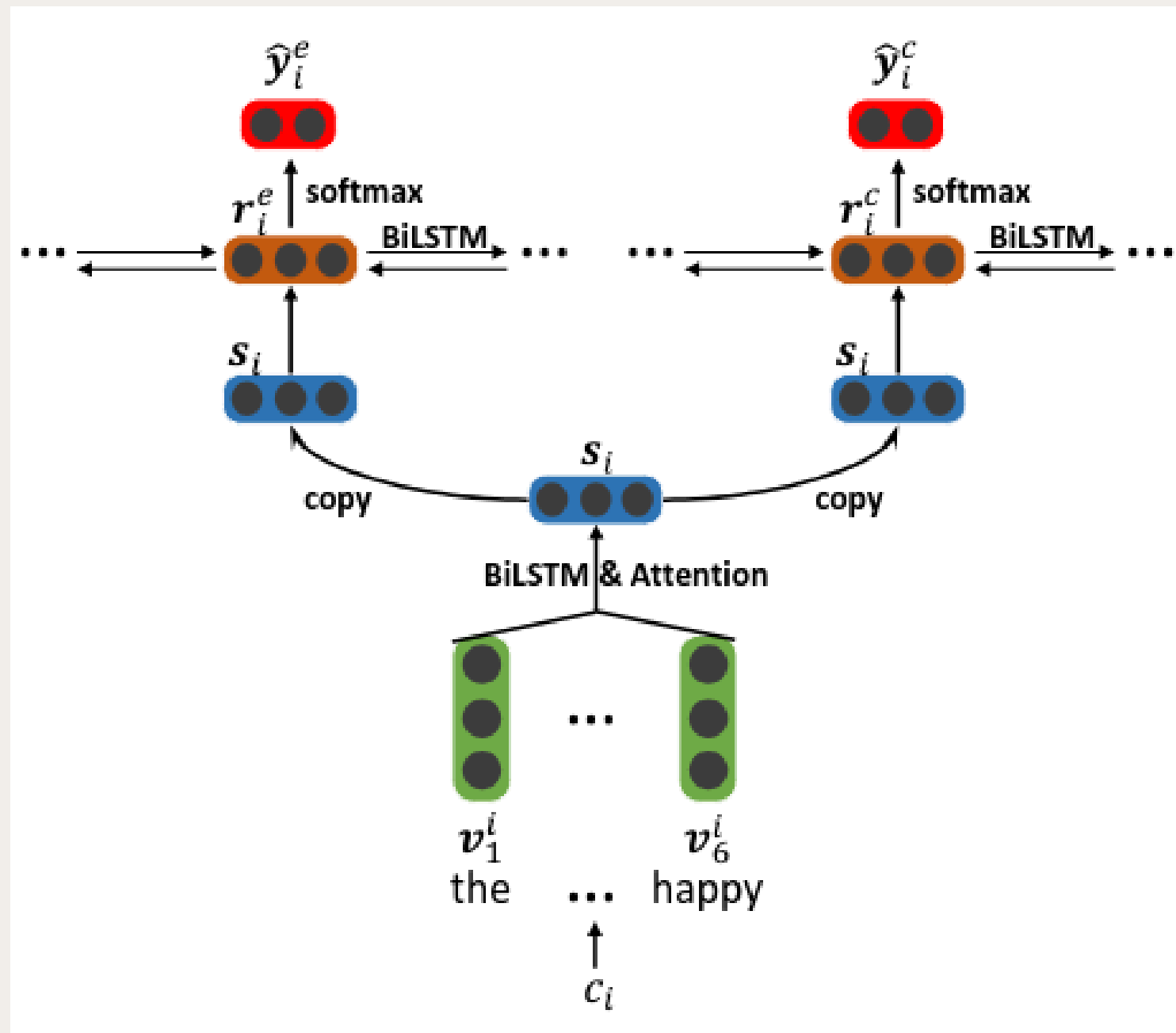


# MOTIVATION

- In classrooms, identifying emotions and their causes can:
  - Help teachers adapt to student needs.
  - Highlight engagement levels and confusion points.
- Existing systems lack this multimodal emotion analysis. Need for multimodal analysis to better capture emotions and their causes.



# Previous Work and Novelty

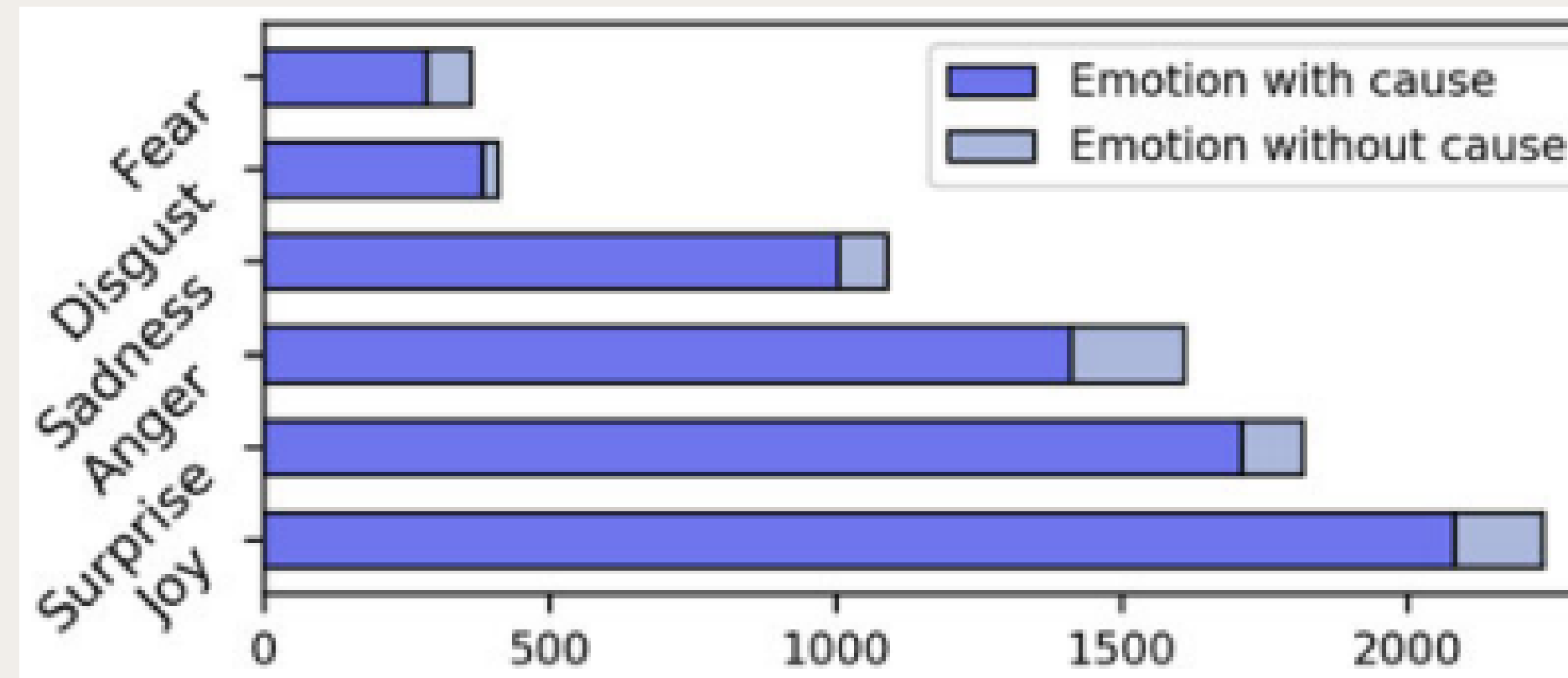


- Many studies have been undertaken on multimodal emotion recognition in talks; however, there is a substantial paucity of research on multimodal emotion cause analysis, which is critical for understanding the underlying reasons for emotions in discussions.
- This paper introduces a **new task**, Multimodal Emotion-Cause Pair Extraction in Conversations, along with a novel dataset, **Emotion-Cause-in-Friends**, and establishes benchmark systems, including deep learning approaches, making it the **first to investigate** joint emotion-cause extraction across **text, audio, and video modalities**.

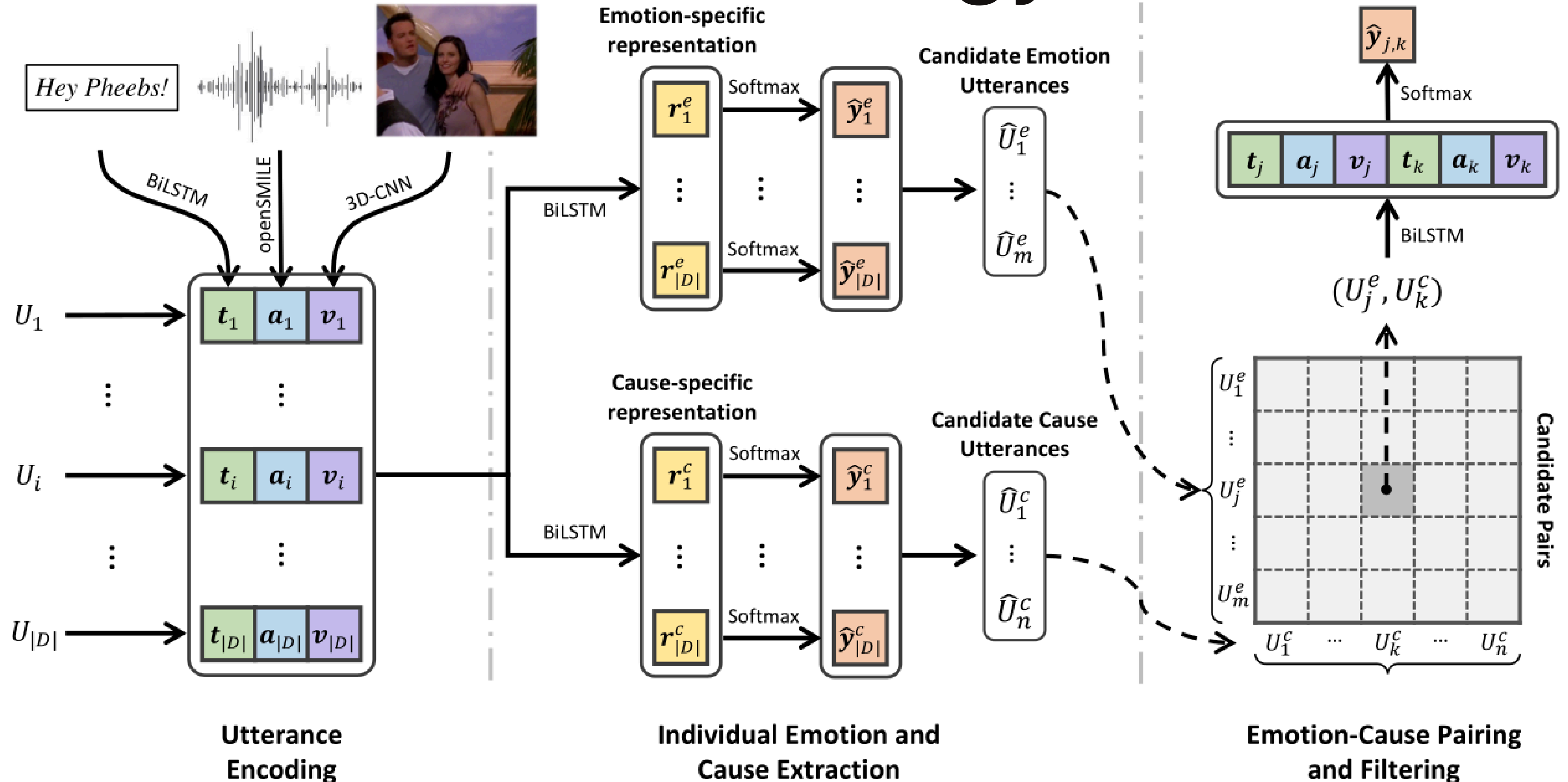
Rui Xia and Zixiang Ding. 2019. Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 1003–1012, Florence, Italy. Association for Computational Linguistics.

# Proposed Dataset

- Dataset: Emotion-Cause-in-Friends (ECF).
- Statistics:
  - 1,344 conversations from Friends.
  - 13,509 utterances with 9,272 annotated emotion-cause pairs.
- Modalities:
  - Text: Dialogues.
  - Audio: Intonation and acoustic cues.
  - Video: Facial expressions and visual context.



# Methodology



# Frame Work

## 1) Multimodal Feature Extraction:

- Text: GloVe embeddings + BiLSTM.
- Audio: Acoustic features via openSMILE.
- Video: Scene features using 3D-CNN.

## 2) Concatenation of Modalities

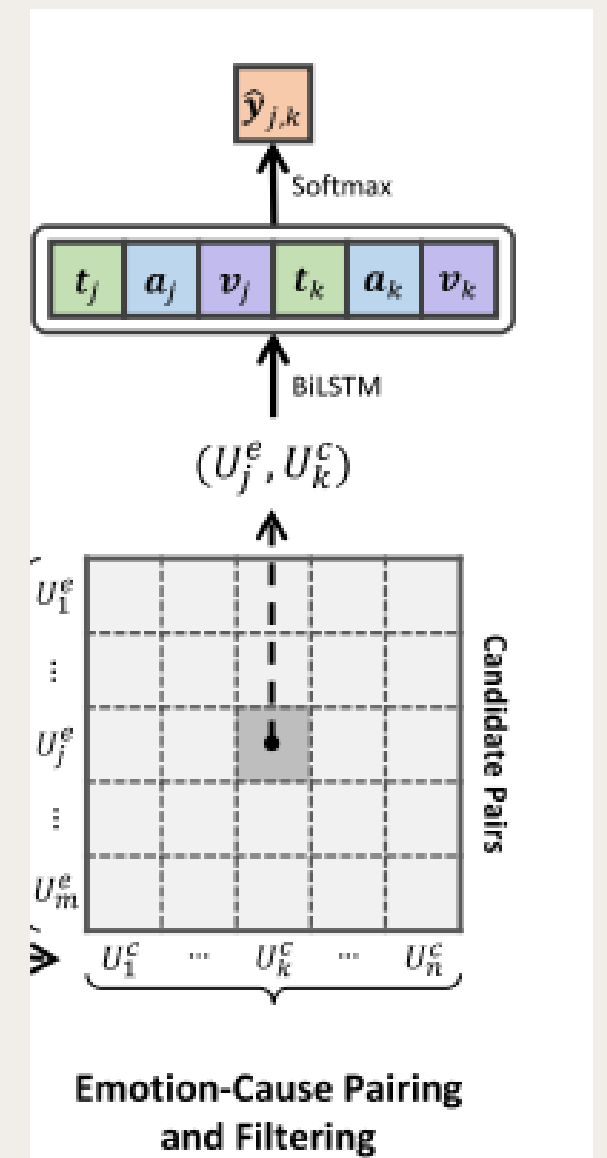
## 3) Emotion and Cause Extraction:

- Uses BiLSTM for emotion/cause utterance representation.

## 4) Pairing and Filtering:

- Combines pairs using attention mechanisms.
- Filters non-causal pairs using a feed-forward network.

$$\hat{\mathbf{y}}_{j,k} = \text{softmax} \left( \mathbf{W} \mathbf{x} \left( U_j^e, U_k^c \right) + \mathbf{b} \right).$$



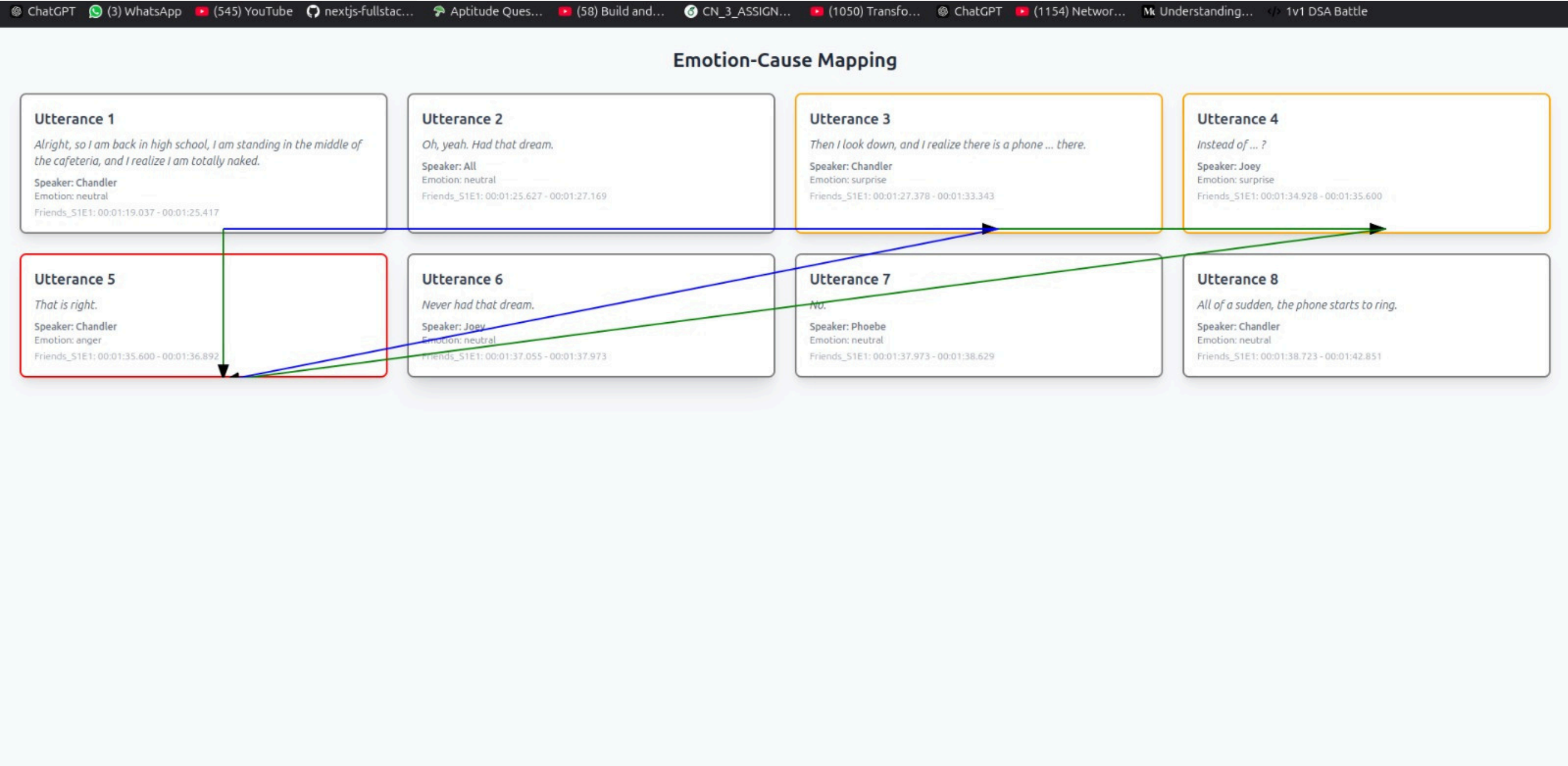
# Experimental Setting and Results

Results for the Replication of MECPE (Text\_LSTM + Audio + Video)

	<b>Precision</b>	<b>Recall</b>	<b>F1 Score</b>
MECPE	0.5538	0.4702	<b>0.5070</b>
Ours	0.53	0.44	0.475



# DEPLOYMENT



# Application in Education

- Adapting MECPE:
  - Fine-tune for classroom conversations.
  - Focus on emotions like confusion, excitement, and frustration.
- Potential Uses:
  - Identifying moments of student engagement or confusion.
  - Providing actionable insights for teachers.
- Steps for Deployment:
  - a. Annotate classroom data for fine-tuning.
  - b. Preprocess text, audio, and video inputs.
  - c. Validate model on real classroom scenarios.

# FUTURE WORK

- Explore advanced fusion techniques like early fusion, late fusion, or joint fusion to better integrate the different modalities (Text, Audio, Video).
- Investigate attention mechanisms that dynamically weigh the importance of each modality for more accurate emotion-cause pair extraction.
- Address emotion category imbalance.
- Develop real-time MECPE applications for intelligent systems.

# Conclusion

## Summary:

- MECPE introduces a novel multimodal approach for emotion-cause extraction.
- Demonstrates the importance of combining text, audio, and video data.

## Vision:

- Extend MECPE to real-world applications like classrooms, customer support, and mental health monitoring.

**Thank You**