

# Emotion Discovery and Emotion Flip Detection in Conversation

Dhriti Singh; 40278439; dhriti.singh@mail.concordia.ca  
and  
Karthik Dammu; 40275326; K\_dammu@live.concordia.ca

November 4, 2024

## Contents

<b>1</b>	<b>Team Info</b>	<b>1</b>
<b>2</b>	<b>A description of the goal of the project</b>	<b>1</b>
<b>3</b>	<b>A proposed methodology to implement the project</b>	<b>1</b>
3.1	Data Collection and Preprocessing: . . . . .	1
3.2	Model Development: . . . . .	2
<b>4</b>	<b>The evaluation methodology</b>	<b>2</b>
<b>5</b>	<b>A timeline and role of each team member</b>	<b>3</b>

## 1 Team Info

Your cool team name and the names & ID of the students in your team.

- Team Name: Binary Squad
- Team Details:
  - Dhriti Singh - 40278439
  - Karthik Dammu - 40275326

## 2 A description of the goal of the project

The aim of this project is to develop system capable of understanding emotions conveyed during conversations and detecting significant emotional shifts or "flips". Specifically, this project will focus on:

- Emotion Discovery: Classifying each utterance within a dialogue into emotion categories joy, anger, sadness, surprise, neutral, disgust and fear.
- Emotion Flip Detection: Identifying moments when the emotional state changes significantly from one utterance to another.

For this project we are using the English version of the MELD dataset. It is a collection of multi-party conversations that provides valuable context for emotion recognition.[2]

Dataset	Anger	Joy	Sadness	Surprise	Neutral	Disgust	Fear
Train	1120	1756	683	1220	4710	275	272
Test	347	409	208	284	1256	68	54

There are three Tasks in the paper "SemEval 2024 – Task 10: Emotion Discovery and Reasoning its Flip in Conversation (EDiReF)", out of which this project will be focusing and working on two tasks, Task A (Emotion Discovery) and Task B (Emotion Flip Detection)[1].

## 3 A proposed methodology to implement the project

### 3.1 Data Collection and Preprocessing:

- Dataset: Utilize the English version of the MELD dataset, which consists of over 1,400 dialogues with approximately 13,000 utterances, each annotated with emotion labels.

- **Preprocessing:** Clean data to handle inconsistencies, null values, and duplicate entries. Tokenize text using a transformer-compatible tokenizer (e.g., BERT tokenizer). Normalize text to ensure consistency across the dataset. Augment data by adding contextual features, such as sentiment scores, to aid in detecting emotion shifts.

### 3.2 Model Development:

- **Task A: Emotion Discovery** Fine-tune a pre-trained transformer model, BERT, to classify emotions in each utterance.
- **Task B: Flip Detection**  
Use the continuous output of Task A to identify transitions between emotional states.
- **Training and Optimization:** Train models on preprocessed data, implementing techniques like early stopping to prevent overfitting if needed. Experiment with model hyper-parameters and fine-tune for optimal performance.

## 4 The evaluation methodology

Following will be done as a part of evaluation:

1. **Task A (Emotion Discovery):** Evaluate model performance using metrics such as accuracy, precision, recall, and F1-score. Compare predictions with the labeled data to assess accuracy.
2. **Task B (Emotion Flip Detection):** Measure precision and recall to understand the model's ability to accurately detect true emotional shifts while minimizing false positives and negatives.
3. **Cross-Validation:** Implement K-Fold Cross-Validation to ensure reliable evaluation and avoid overfitting. This method involves splitting the dataset into K subsets, training on K-1 subsets, and validating on the remaining subset, repeated K times.
4. **Visualization Tools:** Utilize confusion matrices, ROC curves, and precision-recall curves to visualize model performance and facilitate comparison across different model versions.

## 5 A timeline and role of each team member

Phase	Tasks
<b>Phase 1</b>	Project Setup and Data Preprocessing <ul style="list-style-type: none"> <li>- Set up the development environment.</li> <li>- Load and clean the MELD dataset.</li> <li>- Preprocess data and split into training, validation, and test sets.</li> <li>- Study emotion detection models.</li> </ul>
<b>Phase 2</b>	Model Selection and Training <ul style="list-style-type: none"> <li>- Develop and train models (e.g., BERT, RoBERTa).</li> <li>- Fine-tune transformer models for emotion classification.</li> <li>- Implement models for emotion flip detection.</li> </ul>
<b>Phase 3</b>	Model Evaluation and Refinement <ul style="list-style-type: none"> <li>- Hyperparameter tuning and optimization.</li> <li>- Evaluate models with K-Fold Cross-Validation.</li> <li>- Analyze and refine based on results.</li> </ul>
<b>Phase 4</b>	Documentation and Presentation Preparation <ul style="list-style-type: none"> <li>- Prepare project report and poster.</li> <li>- Final demo preparation.</li> </ul>
<b>Throughout the Project</b>	<ul style="list-style-type: none"> <li>- Review and adjust the timeline.</li> <li>- Stay updated on research.</li> <li>- Document progress and maintain a GitHub repository.</li> </ul>

## References

- [1] KUMAR, S., AKHTAR, M. S., AND T., E. Semeval 2024 – task 10: Emotion discovery and reasoning its flip in conversation. *arXiv preprint arXiv:2402.18944* (2024).
- [2] PORIA, S., HAZARIKA, D., MAJUMDER, N., AND MIHALCEA, R. Meld: A multimodal multi-party dataset for emotion recognition in conversations. *arXiv preprint arXiv:1810.02508* (2018).