SentEmoContext: Context-Aware Metric Learning for Efficient Emotion Recognition in Conversation

ACL 2024 Bangkok, Thailand
WASSA

Barbara Gendron[1,2] and Gaël Guibon[1]
[1] LORIA, Université de Lorraine, CNRS, [2] Université du Luxembourg

Motivations

- Information from the conversational context is essential to ensure an accurate understanding of emotions in a dialog between humans.
- There already exist models that perform Emotion Recognition in Conversation (ERC), but they often consist in intricate architectures that require high computing power.

We propose a lightweight yet efficient model based on contrastive learning to perform ERC

Research Questions

- **RQ1:** How to use the information from the conversational context to guide ERC?
- **RQ2:** Does taking into account the conversational context help to improve ERC in dyadic dialogues?

Approach - the Training Procedure

(1) builds contextual utterance representations by deploying attention, (2) specifies contextual information for emotion prediction and (3) preforms prediction on utterance triplets $(\boldsymbol{A}, \boldsymbol{P}, \boldsymbol{N})$ where $\boldsymbol{A}$ and $\boldsymbol{P}$ are from the same class and $\boldsymbol{N}$ is from another one.

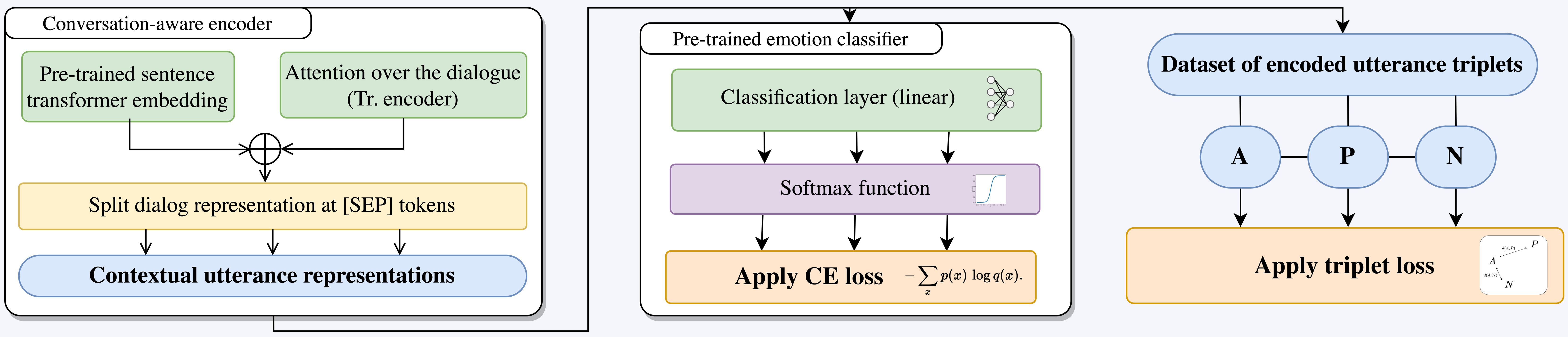

Figure 1: Training procedure for conversation-aware emotion predictions. Both losses (CE and triplet) backpropagate to train the encoder.

Experimental Setup

The DailyDialog dataset (Li et al., 2017)

- 13,118 dyadic dialogues in English about daily life concerns
- Utterance-wise emotion annotation (`no emotion`, `happiness`, `anger`, `disgust`, `fear`, `surprise`, `sadness`)

Metric learning with Siamese Networks (Koch et al., 2015)

- Flexible classification framework: learns relations between labels, adapts to unseen classes
- Intrinsic adaptation to few-shot learning to learn little-represented emotions
- Weight updates using triplet loss help mitigate class imbalance issue

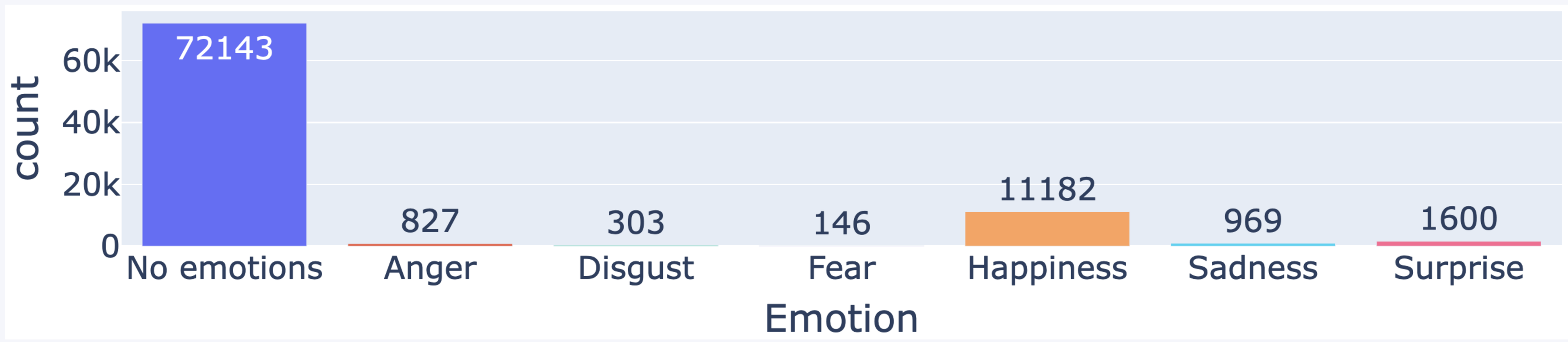

Figure 2: Label distribution in DailyDialog training set.

Evaluation metrics in addition to microF1

- Macro F1: more demanding metric that favors versatile emotion recognition
- MCC (*Matthews Correlation Coefficient*): correlation coefficient between actual and predicted samples for a given class

Results

Model name	macro F1*	micro F1*	MCC
State-of-the-art models on ERC			
CNN+cLSTM (Poria et al., 2017)	–	50.24	–
KET (Zhong et al., 2019)	–	53.37	–
COSMIC (Ghosal et al., 2020)	51.05	58.48	–
RoBERTa (Ghosal et al., 2020)	48.20	55.16	–
Rpe-RGAT (Ishiwatari et al., 2020)	–	54.31	–
Glove-DRNN (Ghosal et al., 2021)	41.8	55.95	–
roBERTa-DRNN (Ghosal et al., 2021)	49.65	57.32	–
CNN (Ghosal et al., 2021)	36.87	50.32	–
DAG-ERC (Shen et al., 2021)	–	59.33	–
TODKAT (Zhu et al., 2021)	<u>52.56</u>	58.47	–
SKAIG (Li et al., 2021)	51.95	59.75	–
Sentic GAT (Tu et al., 2022)	–	54.45	–
CauAIN (Zhao et al., 2022)	–	58.21	–
DialogueRole (Ong et al., 2022)	–	60.95	–
S+PAGE (Liang et al., 2022)	–	**64.07**	–
DualGAT (Zhang et al., 2023)	–	<u>61.84</u>	–
CD-ERC (Pereira et al., 2023)	51.23	–	–
MCM-CSD (Xu and Yang, 2024)	–	60.70	–
Ours			
`SentEmoContext`	**57.71**	57.75	**0.49**

Table 1: All results for ERC on DailyDialog. * indicates metrics do not include neutral label.

Comparison with Large Language Models (LLMs)

Model name	macro F1*	micro F1*	MCC
`llama2-7b-chat-hf`	9.70	24.92	0.08
`llama2-13b-chat-hf`	22.26	43.37	0.15
`falcon-7b-instruct`	07.54	42.75	0.01
Ours			
`SentEmoContext`	**57.71**	57.75	**0.49**

- The tested LLMs performed poorly in zero-shot emotion prediction, while being far bigger that `SentEmoContext`
- Finding the right prompting strategy is challenging
- The MCC values suggest that the emotion predictions are close to random in that case: the task is not well understood in such setting

Take Home Messages

- We propose a novel metric-based approach to perform emotion recognition in conversation
- `SentEmoContext` is state-of-the-art in Macro F1, and consists in a versatile and efficient emotion prediction
- Incorporating context information leads to instability issues because of noise propagation

Check the code!
https://github.com/B-Gendron/sentEmoContext