EDITOR: Erik Cambria, cambria@ntu.edu.sg

DEPARTMENT: AFFECTIVE COMPUTING AND SENTIMENT ANALYSIS

AdaCLF: An Adaptive Curriculum Learning Framework for Emotional Support Conversation

Geng Tu, Taiyu Niu, Ruifeng Xu, *Harbin Institute of Technology, Shenzhen, China*

Bin Liang, *The Chinese University of Hong Kong, Hong Kong, China*

Erik Cambria, *Nanyang Technological University, Singapore*

Abstract—Emotional Support Conversation (ESC) aims to alleviate emotional distress using data-driven approaches trained on human-generated responses. However, the subjective and open-ended nature of human conversations presents challenges in training ESC models due to uneven complexities in query-response pairs. This uneven complexity impedes the efficiency and effectiveness of learning in ESC models. Based on this, we propose an Adaptive Curriculum Learning Framework (AdaCLF) to dynamically choose courses of varying complexity according to the learning status of the ESC model. AdaCLF consists of two main components: the student model (referred to as the ESC model) and the teacher model (responsible for selecting appropriate data to enhance the student model's training). The framework operates within the reinforcement learning paradigm, where the teacher model utilizes feedback from the student model to optimize its teaching strategy, fostering collaborative evolution. Both automatic and human evaluations on benchmark datasets demonstrate that our framework significantly improves existing ESC methods, generating more effective supportive responses.

INTRODUCTION

Emotional Support Conversation (ESC) aims to alleviate emotional distress and foster positive psychological changes [1]. This capability holds significant importance for interactive chatbots and shows potential in diverse applications such as mental health support [2], [3], customer service platforms [4], and aspect-based sentiment analysis [5], [6].

Existing efforts have attempted to endow the model with empathy from the perspective of fine-grained user state modeling [7] and hierarchical psychological relation modeling [8]. Some researchers have also explored manual annotation and data augmentation techniques via policy annotations [4].

1541-1672 © 2024 IEEE
Digital Object Identifier 10.1109/MIS.2024.3411369

Despite these efforts, they remain rooted in the data-driven paradigm, wherein the model is trained on numerous query-response pairs, striving to emulate human conversations. Being a data-driven nature, the caliber of responses generated in ESC is greatly contingent on the quality of the training data [9]. Given the subjective and open-ended nature inherent in human conversations, the complexity of training dialogues varies greatly. In Fig. 1. the response of the third sample appears peculiar given the query provided, whereas the first sample is evidently more straightforward to comprehend. The disparate complexity levels of query-response pairs hinder the learning efficiency and efficacy of ESC models. Inspired by human learning behaviors, we introduce Curriculum Learning (CL) [10] into the ESC model to facilitate gradual learning from simple to complex dialogue scenarios.

 # 1
Seeker: I am sooo pissed at my ex.
Supporter: [Affirmation and Reassurance] why? Any problem? (Anger)

 # 2
Seeker: Hi
Supporter: [Affirmation and Reassurance] Hi! How are you feeling today? (Anxiety)

 # 3
Seeker: It sounds like you are encouraging me to have hope against all odds?
Supporter: [Providing Suggestions] You are a warrior and you need to fight off the worrisome and motivate yourself to help you kiddo. (Sadness)

FIGURE 1. An example illustrates the uneven complexity of query-response pairs in ESC. The supporter's emotions and response strategies are in red and blue font, respectively.

However, designing a curriculum with escalating difficulty encounters significant challenges. Firstly, **there is no standard method for automatically evaluating dialogue complexity**. Previous studies have attempted to address this issue by analyzing variables like sentence length, word rarity, or objective function value, yet a unified approach remains elusive [9].

Additionally, **different models may vary in their alignment with such complexity measures**. For instance, while some models excel in handling long sentences but struggle with sparse word responses, others exhibit the opposite pattern. This disparity complicates the selection of an appropriate curriculum for different ESC models. Secondly, unlike tasks with singular complexity metrics, **dialogue complexity encompasses various attributes** such as response specificity, repetitiveness, and relevance to the query [11].

Based on the aforementioned, we present an Adaptive Curriculum Learning Framework (AdaCLF) to dynamically choose select of varying complexity based on the learning status of ESC models. AdaCLF comprises two primary components: the student model (referred to as the ESC model) and the teacher model (which selects suitable data to enrich the student model's training). Operating within the reinforcement learning (RL) paradigm, AdaCLF leverages feedback from the student model to optimize its teaching strategy, thereby fostering collaborative evolution.

- The first exploration of CL in ESC.
- Introducing the teacher-student structure within the reinforcement learning paradigm, dynamically selecting appropriate data to enhance the ESC model's training.
- Experimental results on the benchmark dataset show that models using AdaCLF are effective in selecting the right strategy and generating better supportive responses.

RELATED WORK

Emotional Support Conversation

Since the proposal of the ESC task and the release of the ESConv dataset, this field has been widely studied. Tu et al. [7] integrate fine-grained emotional status and mixed strategies into emotional support conversation. Zhao et al. [4] use turn-level state Transitions of ESC to make the conversation smoother and more natural. Furthermore, to better extract dialogue features and improve model performance, some works use graph networks to capture the psychological knowledge [8].

Curriculum Learning in Dialogue Generation

CL is one of the machine learning strategies that train the model from simple samples to difficult samples [10]. CL has been widely applied in dialogue generation tasks [12]. These works usually calculate data complexity and design a hierarchical curriculum for student models to enhance their performance [9], [11]. However, previous CL approaches have often overlooked the interaction between teacher and student models. Some recent works have integrated reinforcement learning into CL to improve student model feedback to the teacher. Cai et al. [9] pioneered the application of an RL framework to dialogue generation. Currently, CL's application in the ESC task remains unexplored.

METHODOLOGY

In this section, we provide a comprehensive introduction to the proposed AdaCLF, as depicted in Fig. 2.

Task Definition

The goal of ESC is to produce the subsequent response Y_t using the dialogue history $D = [X_i]^{t-1}$ between the seeker and the supporter, to progressively alleviate the seeker's distress. Each $X_i = [w_i^k]^m$ denotes the i-th utterance, comprising a sequence of m words. Alongside, supplementary details such as the supporter's employed support strategy S_i and the seeker's emotional state label E_i are provided.

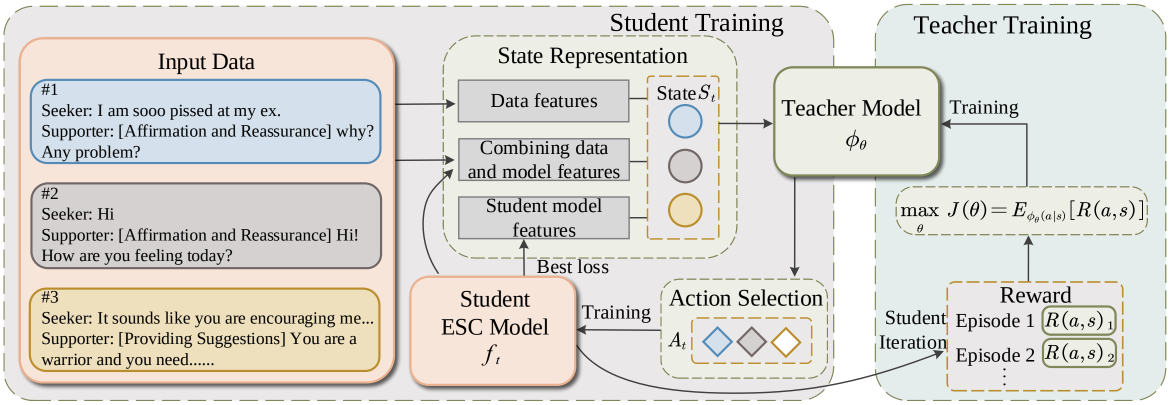

FIGURE 2. Illustration of the AdaCLF.

Overview

The general structure of the proposed AdaCLF framework is depicted in Figure 2. The teacher model ϕ_θ interacts with the environment (composed of training data and student ESC model f_t) to acquire the environmental state s_t at time t. Then the teacher model selects the action a_t, which is a set of mini-batch data. The student ESC model iterates based on a_t, updating its state to s_{t+1}. It then furnishes the teacher model a reward r_t. The teacher model updates itself based on maximizing the expected rewards r_T.

Adaptive Curriculum Learning

General ESC models learn based on the dialogue sample provided by the data loader, but it is important to select data of different difficulty levels during training. For example, it is very hard to learn when providing hard samples in the early stages of training, also there is very little improvement When the ESC model performs well but the data loader still provides simple samples. The selection of data for training is especially crucial when the training set just has a small amount of data. Our AdaCLF trains a teacher model that can adaptively select appropriate data provided for ESC model training at different learning stages based on the current state of the model.

State Representation State s_t refers to the training process information from the beginning of training to the current time step t, for example, the performance of the current ESC model. The teacher model ϕ_θ, a multilayer perceptron, selects the action a_t adapted to the current environmental state s_t to more effectively improve the performance of the ESC model. So an appropriate state representation is crucial for the teacher model to grasp the current learning status. We use three strategies to represent state features:

Data features, representing the feature details of the current training data. Short and broad responses in unspecific conversations are easy to learn, but long responses and responses with many rare words in specific conversations are very difficult to learn. Based on this, we employed two metrics to delineate the characteristics of data complexity: dialogue length and specificity. The specificity is calculated as the mean Normalized Inverse Document Frequency (NIDF) of the words in the response:

$$NIDF(w) = \frac{log(\frac{N_r}{N_w}) - idf_{min}}{idf_{max} - idf_{min}} \tag{1}$$

where N_r denotes the total number of responses in the training dataset and N_w represents the count of responses containing the term w. idf_{max} and idf_{min} signify the maximum and minimum IDF values, respectively.

Student model features represent the performance of the student ESC model. We collect several features to represent the learning status: (1) The current iteration steps; (2) The average training loss across previous iterations; and (3) The best validation loss observed in past iterations.

Combining data and model features, representing how well the model performs with the given input data. We use three types of losses in the student model (like TransESC [4]) to represent the data and model combined features: response prediction loss, strategy prediction loss, and emotion prediction loss.

Action Selection Teacher model $\phi_\theta(a_t|s_t)$ selects the action $a_t = \{a_t^m\}_{m=1}^M \in \{0,1\}^M$ based on the state s_t, where $a_t^m \in \{0,1\}$ indicates whether to retain the m-th data or not. The student ESC model updates itself using the data selected by the teacher and upgrades the state to s_{t+1}. If the accuracy of the student model reaches a pre-set threshold τ, the interaction between the student and the teacher model stops and the teacher model upgrades itself according to the reward $r_T = -log(i_\tau / T^{'})$ given by the student model, where $T^{'}$ is a predefined maximum iteration count.

Teacher Training The objective of the training teacher model is to maximize the expected reward:

$$\max_{\theta} J(\theta) = E_{\phi_\theta(a|s)}[R(a, s)] \tag{2}$$

where the reward function $R(a, s)$ for state s and action a is not differentiable with respect to θ. Therefore, we utilize the policy gradient algorithm based on the likelihood ratio, to optimize $J(\theta)$.

$$\nabla_\theta J(\theta) \approx \sum_{t=1}^{n} \nabla_\theta log\phi_\theta(a_t|s_t)v_T \tag{3}$$

where v_t denotes the sampled estimation of $R(a, s)$. Specifically, v_t is calculated as the reward r_T from one episode. Once the teacher model has been adequately trained, it is utilized to train a new student ESC model by providing selected training data.

EXPERIMENTS

ESConv Dataset

Our research focuses on the ESC dataset, specifically ESConv [1], where conversations involve a seeker in distress seeking help and a supporter aiming to identify, console, and offer suggestions to overcome the seeker's problems. The dataset annotates eight support strategies employed by the supporter (e.g., questioning, reflecting feelings, and providing suggestions). Nevertheless, the ESConv dataset does not possess emotion labels for the seeker's turns and keyword sets for individual utterances. To address this gap, we utilize external tools for automatic annotation, following the methodology outlined in [4].

Experiment Settings

We use three ESC models for students: BlenderBot-Joint [1], MISC [7], and TransESC [4] with their released source codes and identical settings as their original papers. The teacher model comprises three layers, with dimensions 6×4×2, using *tanh* activation for the middle layer and batch normalization to address gradient vanishing. We initialize weights uniformly (-0.01, 0.01) and biases to 0, except the last layer's bias set to 2. Training is conducted using a mini-batch size of 16, leveraging an NVIDIA Tesla A100 GPU. The Adam optimizer is utilized with a learning rate set at 0.001. Training stops when perplexity on the validation set is less than 17.0 or the strategy predicts accuracy higher than 0.30.

Comparison Methods

Our method is compared with several competitive baselines, including two empathetic response generators: Multi-Task Transformer (Multi-TRS) [13] and MIME [14]; four state-of-the-art models: BlenderBotJoint [1], MISC [7], TransESC [4]; and ChatGPT[1] using the same prompt template as in prior work [15].

Evaluation Metrics

Automatic Evaluation: Drawing from [4], our assessment uses various automated metrics to evaluate response quality, diversity, and strategy selection accuracy. Perplexity (PPL) offers an overall measure of response quality, while BLEU-2 (B-2), BLEU-4 (B-4), ROUGE-L (R-L), and Distinct-n (Dist-n) delve into lexical, semantic fidelity, and diversity aspects. Additionally, the accuracy of strategy prediction evaluates the model's capability in selecting appropriate supportive strategies, reflecting its understanding of context and intended response strategies.

Human Evaluation: Three professional annotators evaluated TransESC + AdaCLF against another model using 100 random dialogues from the ESConv dataset's test set. They acted as seekers, comparing the models on five criteria: **Empathy** - understanding of seeker's feelings; **Fluency** - coherence and smoothness of responses; **Suggestion** - helpfulness of suggestions; **Identification** - exploration of seeker's problems; **Overall** - effectiveness of emotional support.

Results and Analysis

Automatic Evaluation: As depicted in Table 1, the superior performance of ESC models over empathetic response generators is evident, emphasizing the importance of addressing the seeker's problems and providing helpful suggestions. TransESC excels due to its explicit turn-level strategy transition modeling, effectively capturing dependencies among different strategies used in each supporter's turn. This leads to superior performance in strategy selection compared to BlenderBotJoint and MISC. Notably, TransESC + AdaCLF demonstrates cutting-edge performance in

[1] https://chat.openai.com/

TABLE 1. Comparison results on different methods in automated evaluation metrics.

Model	Acc	PPL	D-1	D-2	B-1	B-2	B-3	B-4	R-L
Multi-TRS	-	89.52	1.28	7.12	-	6.58	-	1.47	14.75
MIME	-	47.51	2.11	10.94	-	5.23	-	1.17	14.74
BlenderBot-Joint	17.69	17.39	2.96	17.87	18.78	7.02	3.20	1.63	14.92
MISC	31.67	16.27	4.62	20.17	16.31	6.57	3.26	1.83	17.24
TransESC	34.71	15.85	**4.73**	20.48	17.92	7.64	4.01	2.43	17.51
ChatGPT 2-shot	-	-	4.17	**23.23**	12.99	4.07	1.90	1.05	12.70
TransESC + AdaCLF (Ours)	**35.05**	**15.62**	4.45	22.82	**20.33**	**7.99**	**4.08**	**2.48**	**17.59**

TABLE 2. Human evaluation results (%).

Ours vs.	TransESC			ChatGPT 2-shot		
	Win	Loss	Tie	Win	Loss	Tie
Fluency	43.9	15.5	40.6	10.2	6.2	83.7
Identification	39.0	12.2	48.8	17.7	14.1	68.2
Empathy	46.6	15.2	38.3	43.4	21.1	35.5
Suggestion	42.7	13.1	44.1	10.7	50.6	38.8
Overall	47.2	15.9	37.0	27.8	20.4	51.8

TABLE 3. Results of ablation study.

Model	D-1	B-2	B-4	R-L
Ours	4.73	7.99	2.48	17.59
w/o Data features	4.67	6.64	1.90	17.40
w/o Student model features	4.21	6.99	2.01	16.91
w/o Combining data and model features	4.43	6.91	2.14	17.06

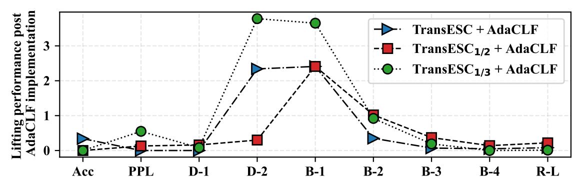

FIGURE 3. Lifting performance post-AdaCLF implementation.

automated evaluation, eliciting more effective emotional responses across all metrics when compared to other ESC models.

Human Evaluation: Table 2 illustrates that TransESC surpasses ChatGPT in empathetic capabilities because ChatGPT tends to rush into providing advice and solutions once it identifies the user's problem, neglecting to offer emotional support and comfort [15]. While ChatGPT may offer a broader range of more effective suggestions, its eagerness to provide solutions overshadows the need for empathetic responses. In contrast, TransESC, especially when enhanced with AdaCLF, outperforms the original model across all evaluation metrics. It generates smoother and more fluent responses, demonstrating the benefits of dynamically selecting appropriate data to enhance the ESC model's training. Moreover, while all three models can identify the user's problems comparably, TransESC + AdaCLF excels in eliciting empathetic responses to comfort the user, thus providing more beneficial suggestions.

Ablation Study

Removal of any type of transition information in AdaCLF results in a noticeable drop in automatic evaluation scores, underscoring the effectiveness of each. Specifically, Table 3 illustrates that removing Data features yields the most significant performance decline in B-2 and B-3, highlighting the pivotal role of selecting appropriate data features based on the data's characteristics in model learning. Removing Student model features leads to the most substantial performance drop in R-L and D-n, likely due to the model's enhanced understanding and representation learning from input data during training, including better comprehension of text semantics and structure, and deeper language model learning. The impact of Combining data and model features is relatively minor, possibly influenced by noise in annotated data. Poor performance in a current batch suggests a need for further training, but if the batch's poor performance is due to noise, it could mislead model learning.

Performance Analysis across Data Sizes

To verify the effectiveness of our AdaCLF with limited data, we investigate the model performance using less data. Table 4 reports notable improvements at 1/2 and 1/3 data proportions compared to the full training set. This can be attributed to the heightened dependence on data when training with fewer samples. Therefore, the importance of selecting suitable data at various training stages becomes more apparent. Notably, diversity and improvements in BLEU scores are most pronounced, as supported by Fig. 3.

TABLE 4. Reimplemented experimental results of generalizability analysis. 1/2 and 1/3 denote training data proportions.

Model	Acc	PPL	D-1	D-2	B-1	B-2	B-3	B-4	R-L
TransESC	34.71	15.85	4.73	20.48	17.92	7.64	4.01	2.43	17.51
TransESC + AdaCLF	35.05	15.62	4.45	22.82	20.33	7.99	4.08	2.48	17.59
TransESC$_{1/2}$	32.90	16.84	3.78	16.95	14.91	5.90	3.02	1.80	16.69
TransESC$_{1/2}$ + AdaCLF	31.37	16.97	3.94	17.25	17.32	6.92	3.39	1.94	16.91
TransESC$_{1/3}$	31.93	17.63	3.42	14.18	14.68	5.98	3.09	1.83	16.68
TransESC$_{1/3}$ + AdaCLF	31.59	18.18	3.50	17.96	18.33	6.90	3.28	1.82	16.69
MISC	31.67	16.27	4.62	20.17	16.31	6.57	3.26	1.83	17.24
MISC + AdaCLF	32.09	17.02	4.57	20.16	17.78	7.23	3.64	2.15	17.57
MISC$_{1/2}$	29.82	17.76	4.06	17.99	16.53	6.65	3.37	1.97	17.26
MISC$_{1/2}$ + AdaCLF	32.26	18.38	4.13	18.29	17.30	6.82	3.26	1.81	17.44
MISC$_{1/3}$	27.49	18.79	4.10	18.94	17.04	6.61	3.15	1.73	16.72
MISC$_{1/3}$ + AdaCLF	29.42	20.16	4.17	18.43	18.03	7.16	3.57	2.04	16.75
BlenderBot-Joint	17.69	17.39	2.96	17.87	18.78	7.02	3.20	1.63	14.92
BlenderBot-Joint + AdaCLF	30.27	16.34	3.78	16.91	18.32	7.61	3.90	2.28	17.65
BlenderBot-Joint$_{1/2}$	26.42	17.25	3.67	16.87	18.71	7.20	3.38	1.82	16.45
BlenderBot-Joint$_{1/2}$ + AdaCLF	28.74	17.57	3.83	17.97	19.53	7.72	3.72	2.04	16.91
BlenderBot-Joint$_{1/3}$	28.85	17.31	3.56	16.27	18.70	7.58	3.72	2.09	17.17
BlenderBot-Joint$_{1/3}$ + AdaCLF	28.40	18.30	3.92	18.24	19.56	7.76	3.76	2.13	17.06

Generalizability Analysis

To validate AdaCLF's generalizability, we experimented with various ESC models, as depicted in Table 4. Significant improvements were observed across almost all evaluation metrics compared to other methods. This underscores the robust generalization capability of AdaCLF. It is noteworthy that there are occasional dips in the D-n metric, indicating a decrease in diversity. This decline may be attributed to the ESC model's selective training, which utilizes up to 85% of the available data rather than the entire dataset.

CONCLUSION

In this paper, we present AdaCLF, a novel solution tailored to tackle the inherent complexities present in ESC training data, which may hinder model learning. Through a dynamic teacher-student framework within a reinforcement learning framework, AdaCLF effectively identifies and utilizes training data to enhance the learning process of ESC models. Experimental findings, supported by both automated metrics and human evaluations, showcase the significant performance improvements achieved by AdaCLF across various ESC models, surpassing existing state-of-the-art methods.

ACKNOWLEDGMENTS

We thank all anonymous reviewers for their helpful comments. This work was supported in part by the National Natural Science Foundation of China under Grants 62176076, the Natural Science Foundation of Guangdong under Grant 2023A1515012922, Shenzhen Foundational Research Funding under Grant JCYJ20220818102415032, The Major Key Project of PCL under Project PCL2023A09, and Guangdong Provincial Key Laboratory of Novel Security Intelligence Technologies under Grant 2022B1212010005k.

REFERENCES

1. S. Liu et al., "Towards Emotional Support Dialog Systems," *Proceedings of ACL-IJCNLP*, 2021, pp. 3469–3483.
2. S. Ji et al., "Domain-specific continued pretraining of language models for capturing long context in mental health," *arXiv preprint arXiv:2304.10447*, 2023.
3. S. Ji et al., "Rethinking large language models in mental health applications," *arXiv preprint arXiv:2311.11267*, 2023.
4. W. Zhao et al., "TransESC: Smoothing Emotional Support Conversation via Turn-Level State Transition," *Findings of ACL*, 2023, pp. 6725–6739.
5. Q. Wang et al., "Image-to-Text Conversion and Aspect-Oriented Filtration for Multimodal Aspect-Based Sentiment Analysis," *IEEE Transactions on Affective Computing*, , no. 01, 2023, pp. 1–15.
6. E. Cambria et al., "SenticNet 8: Fusing Emotion AI and Commonsense AI for Interpretable, Trustworthy, and Explainable Affective Computing," *International Conference on Human-Computer Interaction (HCII)*, 2024.
7. Q. Tu et al., "MISC: A Mixed Strategy-Aware Model integrating COMET for Emotional Support Conversation," *Proceedings of ACL*, 2022, pp. 308–319.
8. W. Peng et al., "Control Globally, Understand Locally: A Global-to-Local Hierarchical Graph Network

for Emotional Support Conversation," *Proceedings of IJCAI*, 2022, pp. 4324–4330.
9. H. Cai et al., "Learning from easy to complex: Adaptive multi-curricula learning for neural dialogue generation," *Proceedings of AAAI, vol. 34*, 2020, pp. 7472–7479.
10. Y. Bengio et al., "Curriculum learning," *Proceedings of ICML*, 2009, pp. 41–48.
11. M. Sachan and E. Xing, "Easy questions first? a case study on curriculum learning for question answering," *Proceedings of ACL*, 2016, pp. 453–463.
12. X. Zhang, R. Mao, and E. Cambria, "Granular Syntax Processing with Multi-task and Curriculum Learning," 2023.
13. H. Rashkin et al., "Towards Empathetic Open-domain Conversation Models: A New Benchmark and Dataset," *Proceedings of ACL*, 2019, pp. 5370–5381.
14. N. Majumder et al., "MIME: MIMicking Emotions for Empathetic Response Generation," *Proceedings of EMNLP*, 2020, pp. 8968–8979.
15. W. Zhao et al., "Is chatgpt equipped with emotional dialogue capabilities?" *arXiv preprint arXiv:2304.09582*, 2023.

Geng Tu is a Ph.D. Student at the School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, China. His current research interests include affective computing and deep learning. Contact him at 22b951011@stu.hit.edu.cn.

Taiyu Niu received the B.S. degree from Xi'an Jiaotong University, Xi'an, China, in 2019. His research interests include natural language processing and multimodal sentiment analysis. Taiyu Niu and the first author contributed equally to this work. Contact him at 23s051050@stu.hit.edu.cn.

Ruifeng Xu received the Ph.D. degree in computer science from The Hong Kong Polytechnic University. He is currently a Professor at the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, China. He has published more than 100 papers in natural language processing, sentiment analysis, and social media analysis. Ruifeng Xu is the corresponding author of this work. Contact him at xuruifeng@hit.edu.cn

Bin Liang is a postdoctoral fellow at the Department of Systems Engineering and Engineering Management in the Chinese University of Hong Kong. He received Ph.D. degree in Computer Science from Harbin Institute of Technology, China. His current research interests include natural language processing, sentiment analysis, deep learning, and machine learning. Contact him at bin.liang@cuhk.edu.hk.

Erik Cambria is currently a Professor at Nanyang Technological University, where he also holds the appointment of Provost Chair in Computer Science and Engineering. His research focuses on neurosymbolic AI for interpretable, trustworthy, and explainable affective computing in domains like social media monitoring, financial forecasting, and AI for social good. Contact him at cambria@ntu.edu.sg