

NLU Project Proposal - Emotional Chatting Machine

Team Member: Bofei Zhang(bz1030), Shaoling Chen(sc6995), Yu Cao(yc3651), Yuxuan He(yh2857)

Motivation: On the current stage, machine learning (focus purely on the context of the words) for dialogue generation has been widely studied. Indeed, existing studies of Prendinger et al. show that addressing emotion in dialogue system can enhance user satisfaction, contribute to a more positive perception of the interaction, and lead to fewer breakdowns in dialogues. Therefore, to improve conversational agent performance, our goal is to utilize and improve Emotional Chatting Machine(ECM) that can generate appropriate responses not only in content but also in emotion.

Work Plan:

1. **Data preparation:** Train sentiment classifiers on an annotated dataset and choose the best classifier for automatic annotation of our dialogue dataset
2. **ECM implementation:** train emotional chatting machine(ECM) using the approach described in Zhou et al.
 - a. **Task:** Given a post $X = (x_1, x_2, \dots, x_n)$ and an emotion category e of the response to be generated, the goal is to generate a response $Y = (y_1, y_2, \dots, y_m)$ that is coherent with the emotion category e
 - b. **Approach:** Use a variant of seq2seq model, ECM, that incorporates Emotion Embedding, Internal Memory, and External Memory. (Framework shown below)

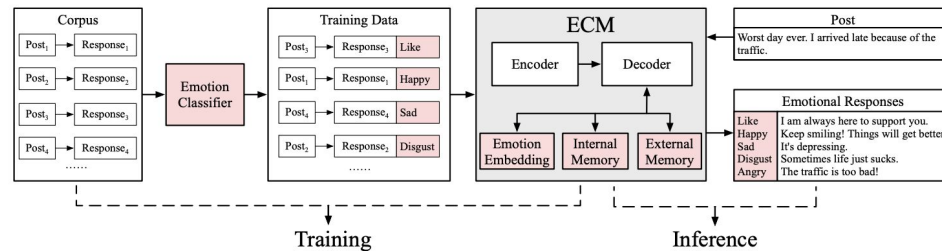


Figure 1: Overview of ECM (the grey unit). The pink units are used to model emotion factors in the framework.

3. **ECM improvement:** experiment the following approaches to improve the model performance
 - a. Incorporate the emotion of posts into the encoder in the training phase to allow empathic responses in the inference phase (Rashkin et al.)
 - b. Use global attention mechanism (Luong et al.)
 - c. Utilize Transformer-based model to replace seq2seq (Devlin et al)

Data and Tools

- [Dialogue Datasets](#): BNC Corpus, Twitter Dialogue dataset, Cornell Movie Corpus, Daily Dialog
- Training data for sentiment classifier: LiveJournal tweets (large text corpus with emotional tag)
- Sentiment Classifiers: There are multiple machine learning models introduced by Bostan et al.

Work Distribution | Collaboration

- Research individually on finding good sentimental classifier to annotate dialogue data
- Implement the ECM according to the paper approach together
- Research individually on ways to improve our ECM model
- Distribute the work to write the paper together

All members agree with this work distribution. All tasks will be distributed evenly.

References

- [1]
L. A. M. Bostan and R. Klinger, “An Analysis of Annotated Corpora for Emotion Classification in Text,” in *Proceedings of the 27th International Conference on Computational Linguistics*, Santa Fe, New Mexico, USA, 2018, pp. 2104–2119.
- [2]
H. Zhou, M. Huang, T. Zhang, X. Zhu, and B. Liu, “Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory,” *arXiv:1704.01074 [cs]*, Apr. 2017.
- [3]
H. Wei, Y. Zhao, and J. Ke, “Building Chatbot with Emotions,” p. 8.
- [4]
I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to Sequence Learning with Neural Networks,” *arXiv:1409.3215 [cs]*, Sep. 2014.
- [5]
H. Rashkin, E. M. Smith, M. Li, and Y.-L. Boureau, “I Know the Feeling: Learning to Converse with Empathy,” *arXiv:1811.00207 [cs]*, Oct. 2018.
- [6]
H. Prendinger and M. Ishizuka, “The Empathic Companion: A Character-Based Interface That Addresses Users’ Affective States,” *Applied Artificial Intelligence*, vol. 19, no. 3–4, pp. 267–285, Mar. 2005.
- [7]
M.-T. Luong, H. Pham, and C. D. Manning, “Effective Approaches to Attention-based Neural Machine Translation,” *arXiv:1508.04025 [cs]*, Aug. 2015.
- [8]
J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” *arXiv:1810.04805 [cs]*, Oct. 2018.