Emotion Recognition in Conversations

Junyin Chen Hanshu Ding

Zoe Fang Yifan Jiang

Task description

- Primary task
  - Detect Emotion based on English dialogue inputs
- Dataset: MELD (Multimodal EmotionLines Dataset)
  - 13000 utterances from 1433 dialogues from TV-series Friends
  - Multiple speakers
    - 6 main characters and others
  - 7 emotions
    - Joy, surprise, sadness, angry, disgust, fear, neutral

neutral	0.471519
joy	0.174492
surprise	0.120633
anger	0.111022
sadness	0.068375
disgust	0.027130
fear	0.026830

No.,Utterance,Speaker,Emotion,Sentiment,Dialogue_ID,Utterance_ID
1,"Oh my God, he’s lost it. He’s totally lost it.",Phoebe,sadness,negative,0,0
2,What?,Monica,surprise,negative,0,1

Task description

1. Data Augmentation: EmoryNLP dataset
   - Emotion Detection on TV Show Transcripts with Sequence-based Convolutional Neural Networks
   - 12606 utterances from 897 dialogues from TV-series Friends
   - Multiple speakers
     - 6 main characters and others
   - 7 emotions (Willcox (1982)'s feeling wheels)
     - neutral, joyful, peaceful, sad, mad, scared, powerful
     - (neutral, joy, surprise, sadness, anger, fear, disgust)

Neutral	0.305416
Joyful	0.219851
Scared	0.129354
Mad	0.108315
Peaceful	0.090598
Powerful	0.078921
Sad	0.067546

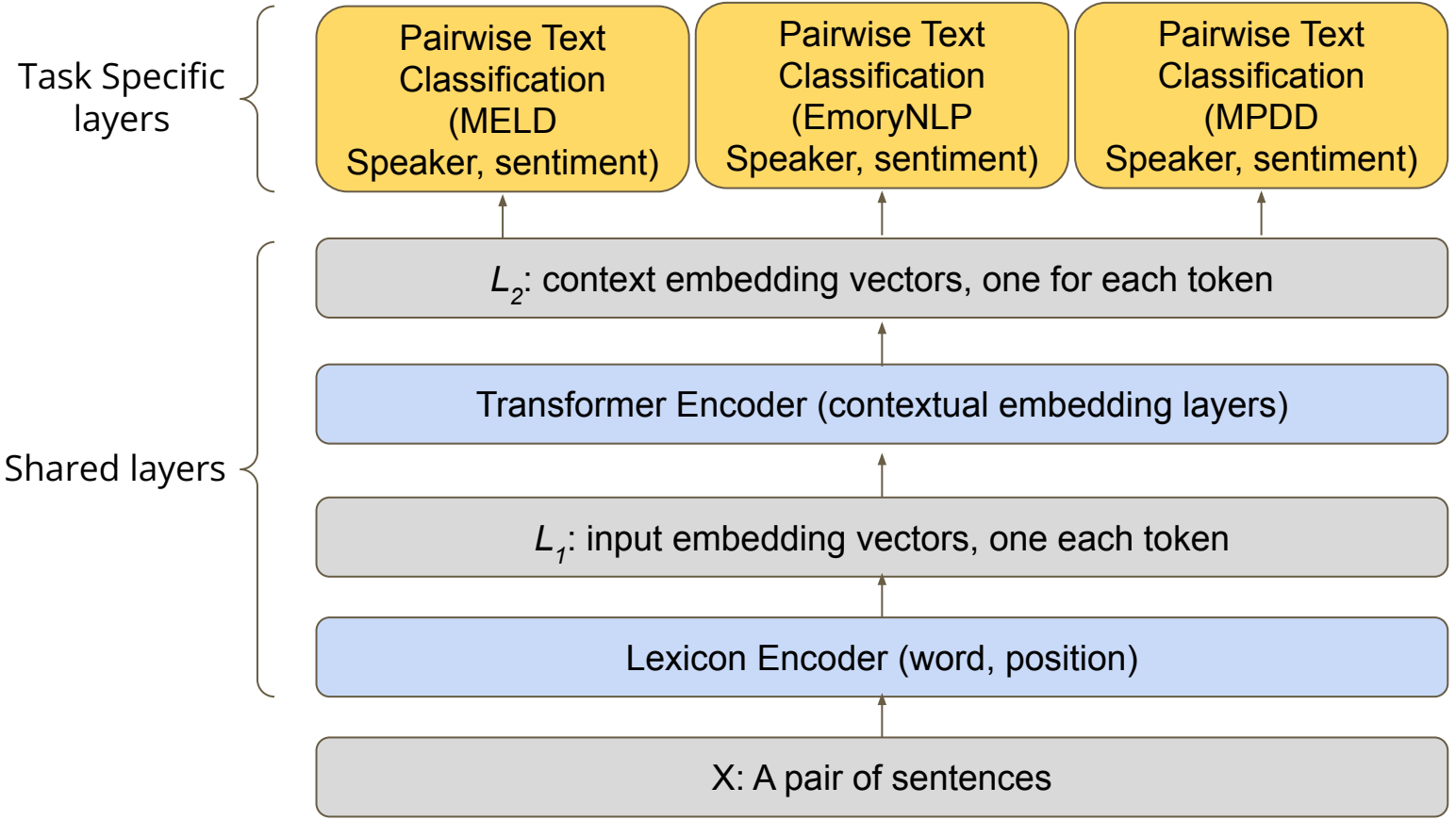
Utterance,Speaker,Emotion,Dialogue_ID,Utterance_ID
Coffee.,Rachel,Neutral,1,0
Thank you.,Joey,Neutral,1,1
Cappuccino.,Rachel,Neutral,1,2

Task description

- Adaptation task
  - Detect Emotion based on Chinese dialogue inputs
- Dataset: MPDD (Multi-party Dialogue Dataset)
  - 25548 dialogues from 4142 dialogues from TV-series
    - Manual random split
  - Multiple speakers
  - 7 emotions
    - Joy, surprise, sadness, angry, disgust, fear, neutral

No.,Utterance,Speaker,Emotion,Dialogue_ID,Utterance_ID
22377,這個，放在這存一下，我們上那邊買衣服，很快回來，行嗎？ ,亞琳,joy,3313,0
22378,謝謝老闆！,亞琳,joy,3313,1

System Architecture



Core Approach

- Multi-Task Deep Neural Network
  - Shared encoder: **RoBERTa (D3)** → **xlm-roberta-longformer-base**
  - Task Heads:
    - Main Task: ERC on MELD -> MPDD
    - Auxiliary Task
      - Speaker Classification (D3) -> not helping
      - Data Augmentation -> MELD, EmoryNLP

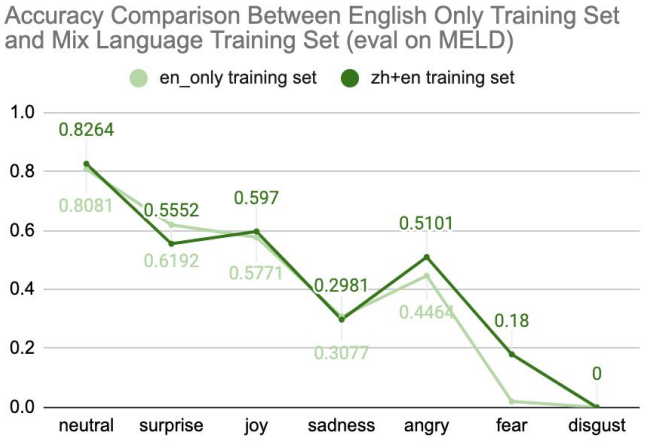
- Input Preprocessing
  - Speaker information
    - “<speaker name>: “ + utterance (D3) -> not helping
  - Utterance Context
    - past/future utterances

D4 Result

Setting	Weighted F1
Training: MELD, EmoryNLP; Test: MELD	0.6118
Training: MELD, EmoryNLP, MPDD; Test: MELD	0.6288
Training: MPDD Test: MPDD	0.5793
Training: MELD, EmoryNLP, MPDD; Test: MPDD	0.5973

D4 – Accuracy Comparison

- The one with mix language training sets present better than the one with mono-language training set.



Issue and Success – Primary Task

<s> Look what I got! Look what I got! Look what I got! Can you believe they make these for little people? Little village people. Okay, look at this one. This is my favourite. </s></s> Oh, that is so sweet!</s>

Predict: joy

Truth: joy

<s> And you're upset because you didn't make your best friend cry? I mean, all I'm asking for is just a little emotion! Is that too much to ask after six years?! </s></s> I mean what?</s>

Predict: surprise

Truth: anger

Issue and Success – Adaptation Task

<s> “美玉,你看咱們是不是該回去了? 是該回去了,學校快開學了,我也有點想家了。</s></s> 是,你姐說她也想咱們了。我們出來了十來天,和你叔叔他們也見了面,他們都挺好,咱們也放心了。</s>

Predict: joy

Truth: joy

<s> 八萬是多少? 八萬就是八萬。八萬就八萬吧。咱們就去取吧,在哪兒取? 在鎮政府。那好吧,咱們這就去。要帶上您的身份證和戶口本。</s></s> 行。</s>

Predict: joy

Truth: neutral

Related readings

- Yi-Ting Chen, Hen-Hsen Huang, and Hsin-Hsi Chen. 2020. MPDD: A multi-party dialogue dataset for analysis of emotions and interpersonal relationships. n *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 610–614, Marseille, France. European Language Resources Association.

- Chao-Chun Hsu, Sheng-Yeh Chen, Chuan-Chun Kuo, Ting-Hao Huang, and Lun-Wei Ku. 2018. Emotion-Lines: An emotion corpus of multi-party conversa-ions. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

- Taewoon Kim and Piek Vossen. 2021. EmoBERTa: Speaker-Aware Emotion Recognition in Conversation with RoBERTa. *arXiv e-prints, page arXiv:2108.12009.*

- Xiaodong Liu, Pengcheng He, Weizhu Chen, and Jianfeng Gao. 2019. Multi-task deep neural networks for natural language understanding. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 4487–4496, Florence, Italy. Association for Computational Linguistics.*

- Soujanya Poria, Devamanyu Hazarika, Navonil Majumder, Gautam Naik, Erik Cambria, and Rada Mihalcea. 2019. MELD: A multimodal multi-party dataset for emotion recognition in conversations. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 527–536, Florence, Italy. Association for Computational Linguistics.*
- Sayyed M. Zahiri and Jinho D. Choi. 2017. Emotion detection on TV show transcripts with sequence-based convolutional neural networks. *arXiv e-prints, page arXiv:1708.04299*

Thank you!