# 自然语言处理技术报告

# Empathetic Dialogue Generation via Sensitive Emotion Recognition and Sensible Knowledge Selection

# 通过敏感情感识别和合理知识选择的同理心对话生成

## 一、项目内容

Empathetic Dialogue Generation via Sensitive Emotion Recognition and Sensible Knowledge Selection

通过敏感情感识别和合理知识选择的同理心对话生成

同理心在心理咨询中被广泛使用，是日常人际对话中的一个关键特质。目前的同理心回应生成方法依赖于常识知识，并着重捕捉对话语境中的隐含情绪，将情绪视为对话过程中的静态变量。然而，情绪在话语之间是动态变化的，这使得之前的方法难以感知情绪的流动并预测目标回应的正确情绪，导致回应不当。此外，简单地导入常识知识而没有进行协调可能会引发知识与情绪之间的冲突，使模型在生成过程中选择不正确的信息。

为了解决上述问题，我们提出了一种序列编码和情感知识交互（SEEK）方法来生成同理心对话。我们采用了细粒度编码策略，对对话中的情绪动态（情绪流）更为敏感，以预测回应的情绪意图特征。此外，我们设计了一个新的框架来模拟知识和情绪之间的交互，生成更加合理的回应。在EMPATHETICDIALOGUES上进行了大量实验，结果显示，SEEK在自动评估和手动评估中均优于强基线模型。

## 二、国内外相关工作

MIME: Majumder等人(2020)提出了一个基于transformer的模型，采用模仿策略根据检测到的用户情绪对目标响应的情绪进行采样。情绪分为两类(积极的和消极的)。该模型利用虚拟声发射来获得模仿和非模仿情绪的表示。

EmpDG (Li et al ., 2019):对抗训练框架由共情生成器和语义情感鉴别器组成。鉴别器确保生成器生成的响应与上下文相关，并且具有同理心。在对抗性框架上训练的收敛生成器可以产生具有高度多样性的共情反应。

KEMP: Li等人(2020)采用图形编码器提取上下文和概念信息基于外部知识构建的上下文图。知识丰富的语境图包含情感依赖，有助于理解会话的情感特征。

CEM: Sabour等人(2021)使用COMET根据说话人在对话中的最后一句话生成常识性知识。作者使用五个特定的前缀(xIntent, xEffect, xWant, xNeed, xReact)来获得与最后一句话相对应的五种知识类型。该模型可以产生更多的信息共情反应。

# 三、方法描述与已有工作的区别

• 据我们所知，我们的工作是第一个在同理心对话生成任务中模拟情感流动，包括情感动态过程。除了对话级别的粗略情感外，我们还引入了话语级别的细粒度情感。

• 通过模拟常识知识和情感之间的双向交互选择过程，我们不仅提高了识别上下文情感的能力，还提高了过滤不合理外部知识的能力，使模型能够生成更加合理的同理心回应。

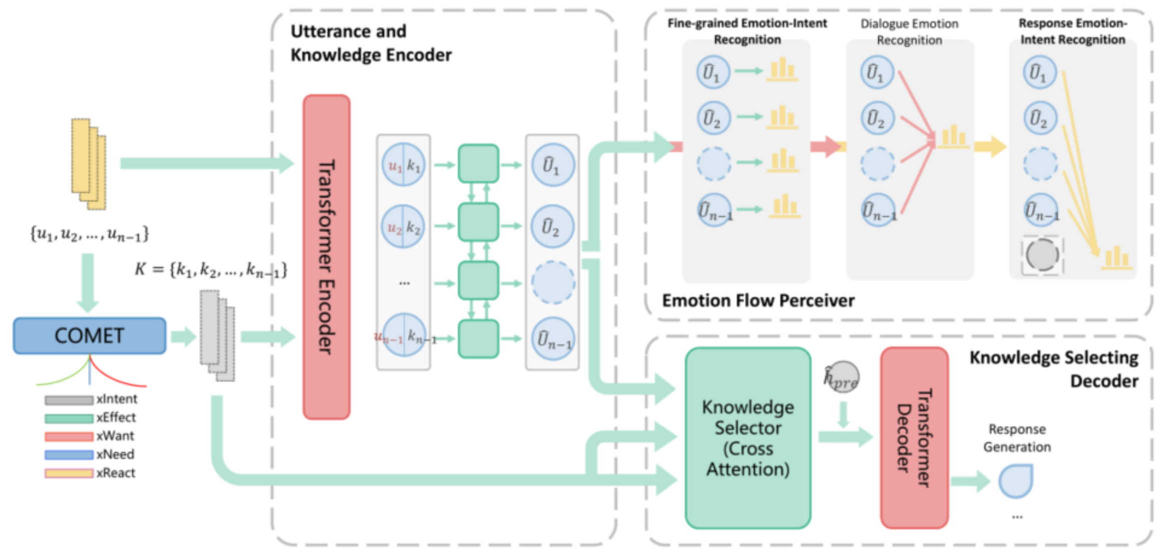
• 在注释的ED数据集上进行的自动评估和手动评估显示，我们提出的模型优于强基线模型，能够生成更多样化和合理的同理心回应。

# 四、算法描述

## 1 Task Formulation 任务制定

共情对话生成的任务是基于历史语境产生共情反应。给定对话D，其中上下文和目标响应表示为C = [C1，…]， CN−1]和Y，带有整个语境的情感标签等。此外，给定的情感意图标签序列EI = [ei1，…]，eiN−1,eiY]，其中包含32个情感类别和9个常见意图类。我们的目标是生成下一个话语Y，它与上下文连贯流畅，并对说话者的处境和感受表达同情。

模型总体架构如下:




## 2 utterance and Knowledge Encoder 话语和知识编码器

### Utterance Encoding 话语编码

为了得到每个话语的精确表示，我们首先在话语层面对上下文进行编码，提取上下文信息。我们使用Transformer 对话语进行编码。输入的嵌入是单词嵌入、位置嵌入和对话状态嵌入的总和。

根据前面的工作，我们在话语ui前加上[CLS]令牌，得到话语输入Ci = [wCLS, w1, w2，…], " wLi]。然后将嵌入输入到Transformer中，从中获得表示。

### Knowledge Encoder 知识编码

为了为相应的上下文生成高质量的常识推理，我们使用COMET (Bosselut等人，2019)，这是一种预训练的GPT (Radford等人，2018)语言模型，并在TOMIC (Sap等人，2019)上进行了微调，以生成五种类型的常识知识:人的效果(xEffect)、说相应句子的人的反应(xReact)、说话前的意图(xIntent)、人的需求(xNeed)以及说完句子后的需求(xWant)。将这5个特殊关系符号附加在话语之后，并将它们输入COMET，我们得到每个输入话语关系的5个常识性推理文本，然后将它们连接到Ki。同样，我们使用相同的Transformer Encoder对知识文本进行编码，并通过平均池化对编码的隐藏状态进行平均。

## 3 Emotion Flow Perceiver 情感流动感知器

将每个话语的情感理解任务作为标记任务，我们使用Bi-LSTM来模拟上下文理解过程中不同话语之间的情感动态和相互作用。

Bi-LSTM的输入是经过编码的话语和知识的拼接:

$$\boldsymbol{a}_i = [\boldsymbol{U}_i; \boldsymbol{K}_i],$$
$$\hat{\boldsymbol{U}}_i = \textbf{BiLSTM}(\boldsymbol{W}_a \boldsymbol{a}_i),$$

其中，$\boldsymbol{W}_a \in \mathbb{R}^{2d \times d}$ 为可训练权值，$\hat{\boldsymbol{U}}_i \in \mathbb{R}^{2d}$ 为处理后的话语表示。

### 3.1 Fine-grained Emotion Recognition 细粒度情感识别

为了更好地理解对话，我们将 $\hat{\boldsymbol{U}}_i$ 传递给标记分类器，生成一个细粒度的情感意图标记分布 $P_{tag} \in \mathbb{R}^t$:

$$P_{tag}(\boldsymbol{ei}_i) = \text{Softmax}(\boldsymbol{W_e} \hat{\boldsymbol{U}}_i)$$

其中t是情绪意图类别的数量。

我们使用会话上下文的预测分布和真实标签之间的交叉熵损失来训练标记模块:

$$\mathcal{L}_{emo} = -\sum_{i=1}^{N-1} log(P_{tag}(\boldsymbol{ei}_i)).$$

### 3.2 Response Emotion-Intent Prediction 回应的情感意图预测

移情对话中情感和意图的转变符合直觉模式。我们使用注意机制来学习话语之间情绪和意图的转换模式。

$$\hat{\mathbf{h}}_{pre} = \text{attention}([\hat{\boldsymbol{U}}_1, \hat{\boldsymbol{U}}_2, ..., \hat{\boldsymbol{U}}_{N-1}]),$$
$$P_{pre} = \text{Softmax}(\boldsymbol{W_p} \hat{\mathbf{h}}_{pre}),$$

在训练过程中，我们将预测反应Ppre的情绪意图分布与目标反应的真实标签eiN之间的交叉熵损失最小化:

$$\mathcal{L}_{pre} = -log(P_{pre}(\boldsymbol{ei}_N)).$$

### 3.3 Dialogue Emotion Recognition 对话情感识别

话语序列的表示不仅包含了话语本身的上下文信息，还表明了整个对话的情感特征。同样地，我们采用注意力机制来总结全局情感标签，基于序列[ˆU1,ˆU2,...,ˆUN−1]。

$$\hat{\mathbf{h}}_{dia} = \text{attention}([\hat{\boldsymbol{U}}_1, \hat{\boldsymbol{U}}_2, ..., \hat{\boldsymbol{U}}_{N-1}]),$$
$$P_{dia} = \text{Softmax}(\boldsymbol{W_d}\hat{\mathbf{h}}_{dia}),$$

### 4 Knowledge Selecting Decoder 知识选择解码器

仅仅将常识性知识引入移情模型而不进行情感上的逻辑选择是不理想的。Sabour等人(2021)通过隐式程序选择常识性推论。相反，我们的方法模拟了对话中相应话语的情感和知识之间的双向互动过程。

我们采用了5层的交叉注意转换器来实现情感和知识的和谐。由于话语表示序列[uu1, uu2，...][U N−1]通过了情感的三个任务，它包含了相应话语的情感特征。交叉注意知识选择器的输入由作为查询向量的话语表示序列组成，键向量和值向量都是由COMET模型K = [K1，...KN−1]生成的知识文本。所选知识的隐式表示如下:

$$\boldsymbol{\mathcal{S}} = \text{Cross-Attention}(\hat{\boldsymbol{U}}, \boldsymbol{\mathcal{K}}, \boldsymbol{\mathcal{K}}),$$

我们以变压器解码器作为解码器的骨干。我们在平均协调知识S和响应表示的预测之间执行串联操作，以获得这两种类型的信息的混合物来表示[SOS]令牌:

$$[\text{SOS}] = \boldsymbol{W}_k([\boldsymbol{S}; \hat{\mathbf{h}}_{pre}])$$

### 5 Training Objectives 训练目标

在训练过程中，我们需要最小化三个分类损失和一个响应生成损失。分类损失的权重相等:

$$\mathcal{L}_{cls} = \mathcal{L}_{tag} + \mathcal{L}_{pre} + \mathcal{L}_{dia}.$$

为了提高生成响应的多样性，我们采用频率感知交叉熵(FACE) 作为惩罚高频令牌的额外损失：

$$\mathcal{L}_{div} = -\sum_{t=1}^{T}\sum_{i=1}^{V} w_i\delta_t(c_i)log(P(y_t|C, y_{<t}),$$

其中wi是词汇表V中第i个标记的频率权重值，ci表示词汇表中的候选标记，δt(ci)是表示ci是否等于基本真实标记yt的函数。

最后，通过最小化上述三种损失的加权和，对我们提出的模型的所有参数进行联合训练和优化：

$$\mathcal{L} = \alpha\mathcal{L}_{nll} + \beta\mathcal{L}_{cls} + \gamma\mathcal{L}_{div},$$

其中α， β和γ是用于平衡三个损失的超参数。在我们的实验中，我们设置α= 1， β= 1， γ= 1.5。

## 五、实验结果及分析

## 1 Dataset数据集

我们的实验是在话语水平注释的empathticdialogues (ED)上进行的(Rashkin et al .， 2019;Welivita and Pu, 2020)。ED是一个大规模的多回合对话数据集，包含说话者和听者之间的25k次移情对话。ED提供了日常聊天中常见的32个均匀分布的情绪标签。然而，ED数据集的情感标签是在语境层面上的，对于话语层面的情感没有明确的信号。Welivita和Pu(2020)用41个新的话语级情感和意向标签类别对ED数据集进行了注释，这些标签提供了关于ED数据集中共情对话的细粒度信息。

## 2 Baselines基本

我们选择了几个强大的基线模型进行比较，包括:

| Models | PPL | Dist-1 | Dist-2 | DE Acc. | UEI Acc. | REI Acc. |
|---|---|---|---|---|---|---|
| MIME | 37.08 | 0.31 | 1.03 | 29.38 | - | - |
| EmpDG | 37.77 | 0.59 | 2.48 | 30.03 | - | - |
| KEMP | **36.89** | 0.61 | 2.65 | 37.58 | - | - |
| CEM | 37.03 | 0.66 | 2.99 | 36.44 | - | - |
| SEEK | 37.09 | **0.73** | **3.23** | **41.85** | **34.08** | **25.67** |

表1：基线和我们模型的自动评估结果。SEEK对四条强基线的改善具有统计学意义(配对t检验，p值< 0.05)。

| Models | PPL | Dist-1 | Dist-2 | DE Acc. | UEI Acc. | REI Acc. |
|---|---|---|---|---|---|---|
| **SEEK** | **37.09** | **0.73** | **3.23** | **41.85** | 34.08 | 25.67 |
| **w/o Utter** | 37.37 | 0.70 | 3.13 | 38.9 | - | **30.41** |
| **w/o Res** | 37.97 | 0.63 | 2.74 | 40.82 | **50.48** | - |
| **w/o Utter & Res** | 38.48 | 0.60 | 2.70 | 39.7 | - | - |
| **w/o Emo** | 37.67 | 0.61 | 2.66 | 41.27 | 35.88 | 23.37 |
| **w/o Know** | 37.35 | 0.31 | 1.19 | 41.07 | 33.53 | 25.58 |
| **+ Others know** | 37.50 | 6.90 | 2.88 | 38.25 | 34.43 | 24.32 |
| **+ Context Enc** | 38.68 | 0.67 | 2.60 | 41.81 | 32.86 | 24.45 |

表2：我们提出的模型SEEK的消融研究。最好的结果用粗体标出。

## 3 Implementation Details实现细节

我们使用Pytorch (Paszke等人，2019)实现我们的模型，并使用Adam (Kingma和Ba, 2015)优化器来优化模型。我们使用300维预训练的GloVE向量(Pennington et al .， 2014)初始化词嵌入，这些词嵌入在编码器和解码器之间共享。在训练阶段，学习率初始值为0.0001，我们根据V aswani et al(2017)改变学习率。我们的模型在一个NVIDIA Geforce RTX 3090 GPU上使用批处理大小为32和早期停止策略进行训练。对于其他设置，例如辍学率，最大解码步骤等，我们与Sabour等人(2021)保持相同。SEEK的训练时间约为3小时，迭代次数约为27000次。

## 4 Automatic Evaluation自动评估

由于Liu et al(2016)已经证明了一些基于词重叠的自动度量可能不适合评估对话系统，例如BLEU (Papineni et al .， 2002)和ROUGE (Lin, 2004)，因此我们采用Perplexity (PPL)和Distinct-n (distn) (Li et al .， 2016)作为生成质量的主要自动度量。对于会话情绪识别和我们新引入的细粒度情绪意图标记和反应情绪意图预测两个任务，我们采用了对话情绪准确性(DE Acc)、话语情绪意图准确性(UEI Acc)和反应情绪意图准确性(REI Acc)。

为了检验SEEK是否可以通过细粒度情感识别产生更明智的响应，我们将模型的性能与强基线进行了比较。如表1所示，SEEK的多样性得分(dist1和dist2)优于所有基线，这表明我们的模型可以基于外部知识产生更多的信息响应。我们将这种改进归因于知识选择器和目标反应的预测情绪，其中交叉注意机制有助于根据话语的上下文信息选择相关知识，并提供预测向量生成过程的附加信息。

为了证明SEEK是否对对话情感有更好的理解，我们列出了基线和我们提出的模型的准确性。值得注意的是，SEEK大大超过了所有的基线，我们将性能的提高归功于我们引入的两个细粒度任务。对对话中的话语理解得越好，对话的准确性就越高。在UEI准确率和REI准确率这两个新的准确率分数上，SEEK达到了令人满意的表现，因为这两个任务的类别数量为41个。

| Models | Coh. | Emp. | Flu. |
|---|---|---|---|
| MIME | 2.84 | 2.97 | 2.87 |
| EmpDG | 2.85 | 2.78 | 2.76 |
| KEMP | 2.73 | 2.80 | 2.80 |
| CEM | 2.82 | 2.99 | 2.75 |
| **SEEK** | **2.91** | **3.02** | **3.07** |

表3：评价结果。我们使用Fleiss's Kappa(表示为κ)来衡量注释者间的一致性，其中$0.4 < \kappa < 0.6$表示中度一致性。

## 5 Human Evaluation人工评估

根据之前的工作，我们基于三个方面进行了人类评估:连贯性(Coh.):回应与上下文的相关性有多大?共情(Emp.):模型对说话人的处境和情绪特征了解多少?模型是否有足够的共鸣或给出建议?流畅性(流感):生成的回答在多大程度上符合语法?我们随机选择100个对话，并将模型生成的响应分配给三位众包工作者进行评估。

每个方面都在1到5的范围内。此外，考虑到不同个体之间的差异，我们进行了另一次人类A/B测试，直接将我们的方法与其他基线进行比较。三名专业注释员对回答对的问卷进行评分，随机选择一个回答，或在提供的句子质量难以区分时选择“Tie”。人力评级和A/B测试的结果如表3和表4所示，SEEK在所有这三个方面都优于基线。

| Comparisons | Aspects | Win | Lose | Tie |
|---|---|---|---|---|
| SEEK vs. MIME | Coh. | **24.3** | 17.1 | 58.6 |
| | Emp. | **31.4** | 22.2 | 46.4 |
| | Flu. | **28.6** | 25.9 | 45.5 |
| SEEK vs. EmpDG | Coh. | **32.1** | 26.3 | 41.6 |
| | Emp. | **35.5** | 27.4 | 37.1 |
| | Flu. | **26.9** | 22.3 | 50.8 |
| SEEK vs. KEMP | Coh. | **29.2** | 25.2 | 45.6 |
| | Emp. | **28.8** | 19.9 | 51.3 |
| | Flu. | **38.7** | 15.6 | 45.7 |
| SEEK vs. CEM | Coh. | **27.3** | 24.8 | 47.9 |
| | Emp. | **33.4** | 27.5 | 39.1 |
| | Flu. | **35.7** | 21.6 | 42.7 |

表4：人类A/B测试(%)在三个方面:连贯性、同理心和流畅性。对比结果直接表明SEEK模型优于四种基线模型。

## 6 Ablation Studies消融实验

为了研究模型中任务和模块的影响，我们删除了新引入的任务和情感与知识之间的交互过程。此外，我们还分别替换了知识类型和编码策略。

结果如表2所示。

去除细粒度的话语情感意图标注和响应情感意图预测任务(w/o Utter、w/o Res和w/o Utter & Res)会导致对话情感识别的准确性和生成质量下降，因为这些变体失去了对对话的细粒度理解和预测目标响应的情感意图特征的能力。

没有情感协调的知识变体(w/o Emo)与SEEK之间的余量从模型的知识选择模块证明了知识与情感意图交互的重要性。无知识的变体(w/o Know)表明外部知识对模型生成的响应多样性的重要性。

此外，替换知识类型+ other Know和编码策略+ Context Enc的性能下降表明了我们的方法的优越性。在我们的模型中使用Others类型的知识而不是PersonX导致所有性能的显著下降，这表明PersonX类型的常识有助于模型更有效地理解话语。采用的编码策略by baselines(如变体+ Context Enc所使用的)强调对整个对话的整体理解，而忽略了对话语的准确把握，这导致了表现的下降。

值得注意的是，w/o Utter的ui精度和w/o Res的REI精度均高于SEEK。

这可能是由于带注释的ED数据集的话语标签的噪声和意图类别之间的细微差异(如同意和承认，咨询和质疑)，这意味着话语或响应的分类监督信号会使注意模块的输入向量变得更加困难，并丢失一些其他类别的信息。关于隐藏状态的信息的丢失可能会混淆其他分类器并导致准确性的降低。在任何情况下，虽然这两个任务之间存在权衡，但它们可以同时提高模型通过建模情绪流来产生更合理的共情反应的能力。

## 7 Case Study案例分析

图1中的第一个例子说明了在多回合对话中情绪是如何变化的。为了更好地比较我们的模型和基线生成的响应，我们在表5中显示了我们的模型和基线生成的两个结果。在第一种情况下，基线没有给出带有怀旧色彩的反应，类似于图1所示的常识知识，CEM选择了错误的知识来产生带有快乐情绪和娱乐意图的反应。相反，SEEK成功地给出了一个更敏感和准确的情绪感知的回应。同样，在第二种情况下，所有的基

线都是基于明确的内疚情绪产生的反应，没有细致的理解哪个更准确。与基线不同，SEEK以同情的意图敏感地回应。

| | |
|---|---|
| **Context** | *Speaker:* I love YouTube. I've been listening to all my **classic** tracks. Tupac forever. (**Nostalgic**) |
| | *Listener:* I love me some Tupac. Real talk. (**Acknowledging**) |
| | *Speaker:* I started out with *One Hit Wonders* but ended up at Pac. **I miss my youth lol**. (**Nostalgic**) |
| **Speaker Emotion** | Nostalgic |
| **MIME** | I am sure you will be able to get a new one. |
| **EmpDG** | I am sure you will be fine. |
| **KEMP** | I am sure you will be fine. |
| **CEM** | That sounds like fun! |
| **SEEK** | I do not like that, but i think they will be able to get the old times . (**Nostalgic**) |
| **Golden** | Me too. I always end up on the 90s hip hop station on Pandora. (**Nostalgic**) |

| | |
|---|---|
| **Context** | *Speaker:* Yeah about 10 years ago I had a horrifying experience. It was 100% their fault, but they hit the water barrels and survived. They had no injuries, but they almost ran me off the road. (**Guilty**) |
| | *Listener:* Did you suffer any injuries? (**Questioning**) |
| | *Speaker:* No, I was not hit. It turned out they were drunk. I felt guilty, but realized it was his fault. (**Guilty**) |
| **Speaker Emotion** | Guilty |
| **MIME** | I would be too. I would be so scared. |
| **EmpDG** | I am sorry to hear that. |
| **KEMP** | I am sorry to hear that. |
| **CEM** | That is so sad. |
| **SEEK** | That is pretty scary! I am glad that you were able to get in out! (**Sympathizing**) |
| **Golden** | Why did you feel guilty? People really shouldn't drive drunk. (**Questioning**) |

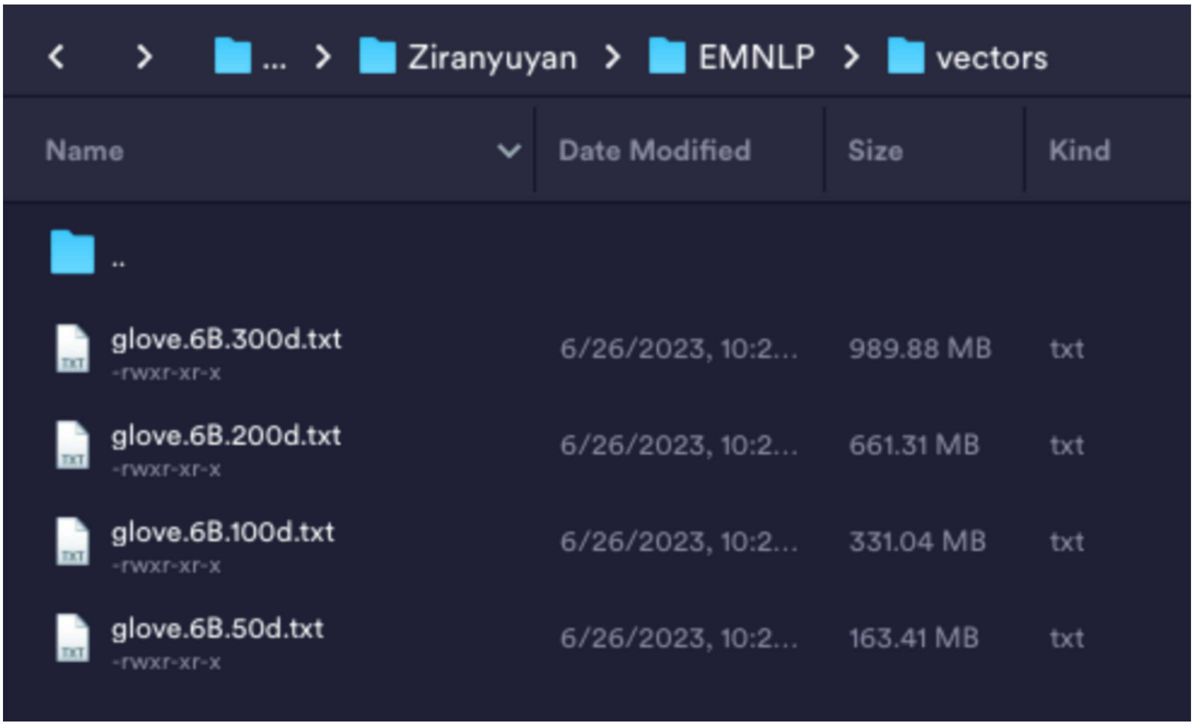
表5：SEEK和基线生成的响应的两个案例。我们在每句话的末尾用情感或意图标签标注每一个回合。SEEK响应中与预测标签相关的单词以红色突出显示。

## 8 代码实现过程

### 1、安装所需的库

```
transformers==4.22.1
tqdm==4.64.0
nltk==3.7
numpy==1.19.2
torch==1.11.0
tensorboardX==2.2
pip install -r requirements.txt
```

### 2、下载预训练模型 Pretrained GloVe Embeddings ， 嵌入并保存在 /vectors 中。

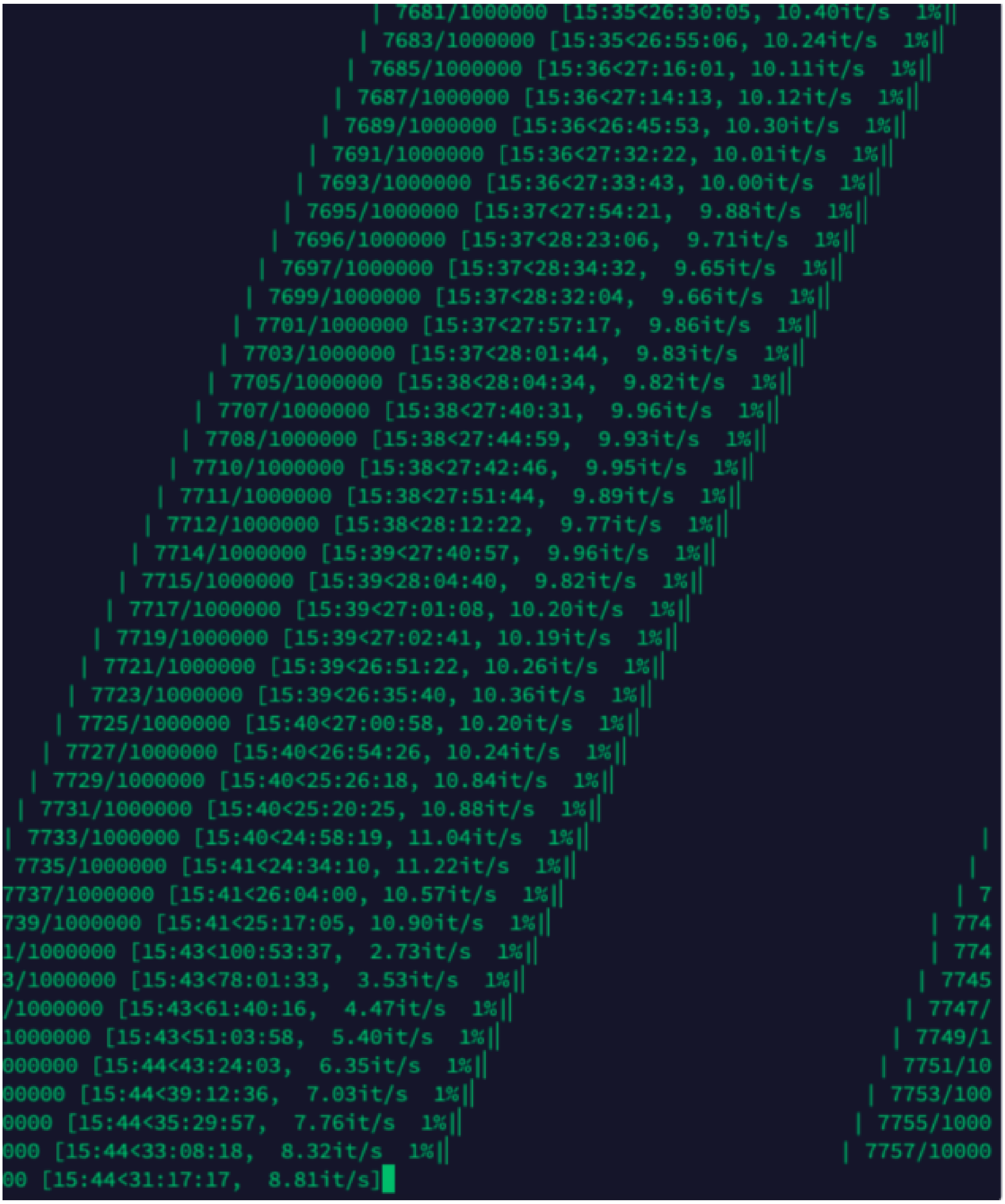


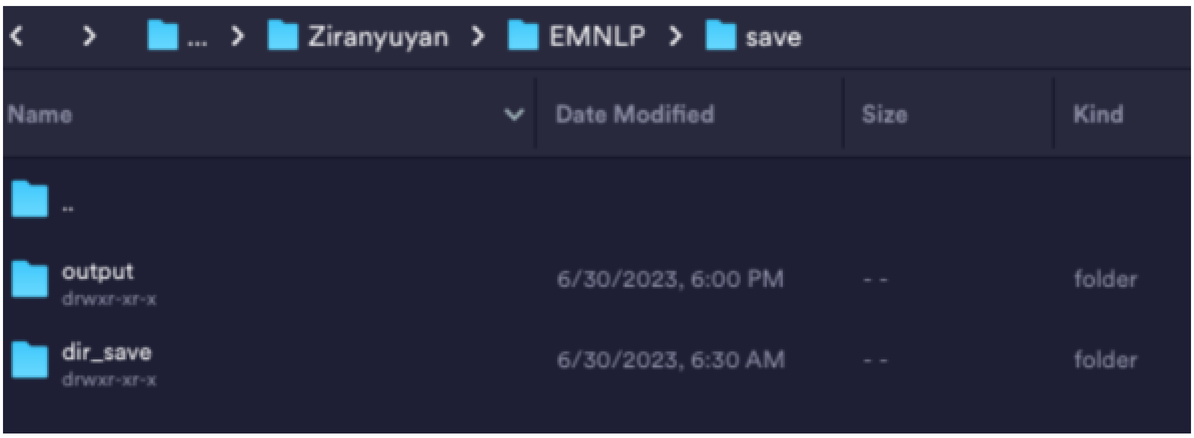

**3、预处理后的数据保存为“/data/ED/dataset_preproc.p”。预处理后的数据集将在训练脚本之后**生**成。**

**4、训练**

```
python main.py --cuda --save_path save/output
```

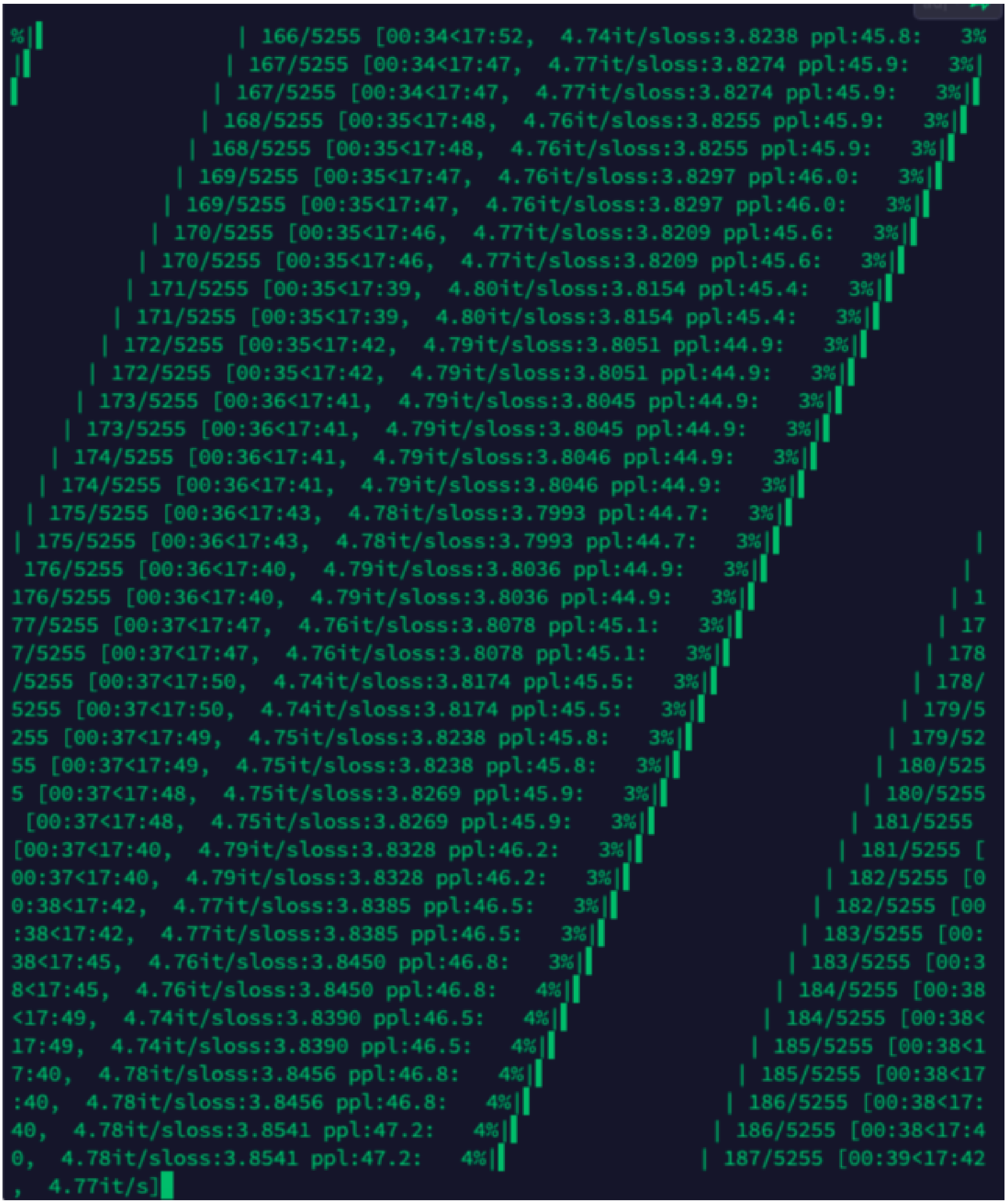
```
[dia_emotion]: furious
[emotion]: ['furious', 'questioning', 'prepared', 'neutral']
[context]: ['i spent hours shopping and getting everything i needed . i was fi
nally done and went to the car and dropped my cookies all over the floor . i w
as so mad at myself !', 'that sounds frustrating . did you drop anything else
?', 'thankfully no . and next time i went , i bought more cookies !']
[target]: well , that is good that it was just the cookies ? it could have bee
n much worse .

[dia_emotion]: nostalgic
[emotion]: ['nostalgic', 'questioning', 'nostalgic', 'nostalgic']
[context]: ['my best friend from high school and i pulled out our old year boo
ks the other day .', 'thats awesome ! i was actually looking for mine the othe
r day . were you looking up anyone in particular or just for fun ?', 'just for
 fun . we had a blast telling stories about our high school days .']
[target]: i still see my high school buddies on a somewhat regular basis but i
t is a great time telling those stories over and over .

[dia_emotion]: devastated
[emotion]: ['devastated', 'sympathizing']
[context]: ['i had yet another death in my family . this year has been the wor
st .']
[target]: im sorry , theres nothing worse than that . i hope you are doing oka
y

[dia_emotion]: embarrassed
[emotion]: ['embarrassed', 'acknowledging']
[context]: ['i was sad when i tripped in public . it was not a good look']
[target]: must 've been embarassing .

[dia_emotion]: surprised
[emotion]: ['surprised', 'questioning']
[context]: ['i found $ 20 on the side walk today']
[target]: did you try to find who it belonged to ?

06-30 17:59 Vocab  22322
Embeddings: 22322 x 300
Loading embedding file: vectors/glove.6B.300d.txt
Pre-trained: 18717 (83.85%)
# PARAMETERS 17370171
  0%|                                     | 256/1000000 [00:26<29:39:06,  9.37it/s]
```


| 7681/1000000 [15:35<26:30:05, 10.40it/s 1%||
| 7683/1000000 [15:35<26:55:06, 10.24it/s 1%||
| 7685/1000000 [15:36<27:16:01, 10.11it/s 1%||
| 7687/1000000 [15:36<27:14:13, 10.12it/s 1%||
| 7689/1000000 [15:36<26:45:53, 10.30it/s 1%||
| 7691/1000000 [15:36<27:32:22, 10.01it/s 1%||
| 7693/1000000 [15:36<27:33:43, 10.00it/s 1%||
| 7695/1000000 [15:37<27:54:21, 9.88it/s 1%||
| 7696/1000000 [15:37<28:23:06, 9.71it/s 1%||
| 7697/1000000 [15:37<28:34:32, 9.65it/s 1%||
| 7699/1000000 [15:37<28:32:04, 9.66it/s 1%||
| 7701/1000000 [15:37<27:57:17, 9.86it/s 1%||
| 7703/1000000 [15:37<28:01:44, 9.83it/s 1%||
| 7705/1000000 [15:38<28:04:34, 9.82it/s 1%||
| 7707/1000000 [15:38<27:40:31, 9.96it/s 1%||
| 7708/1000000 [15:38<27:44:59, 9.93it/s 1%||
| 7710/1000000 [15:38<27:42:46, 9.95it/s 1%||
| 7711/1000000 [15:38<27:51:44, 9.89it/s 1%||
| 7712/1000000 [15:38<28:12:22, 9.77it/s 1%||
| 7714/1000000 [15:39<27:40:57, 9.96it/s 1%||
| 7715/1000000 [15:39<28:04:40, 9.82it/s 1%||
| 7717/1000000 [15:39<27:01:08, 10.20it/s 1%||
| 7719/1000000 [15:39<27:02:41, 10.19it/s 1%||
| 7721/1000000 [15:39<26:51:22, 10.26it/s 1%||
| 7723/1000000 [15:39<26:35:40, 10.36it/s 1%||
| 7725/1000000 [15:40<27:00:58, 10.20it/s 1%||
| 7727/1000000 [15:40<26:54:26, 10.24it/s 1%||
| 7729/1000000 [15:40<25:26:18, 10.84it/s 1%||
| 7731/1000000 [15:40<25:20:25, 10.88it/s 1%||
| 7733/1000000 [15:40<24:58:19, 11.04it/s 1%|| |
7735/1000000 [15:41<24:34:10, 11.22it/s 1%|| |
7737/1000000 [15:41<26:04:00, 10.57it/s 1%|| | 7
739/1000000 [15:41<25:17:05, 10.90it/s 1%|| | 774
1/1000000 [15:43<100:53:37, 2.73it/s 1%|| | 774
3/1000000 [15:43<78:01:33, 3.53it/s 1%|| | 7745
/1000000 [15:43<61:40:16, 4.47it/s 1%|| | 7747/
1000000 [15:43<51:03:58, 5.40it/s 1%|| | 7749/1
000000 [15:44<43:24:03, 6.35it/s 1%|| | 7751/10
00000 [15:44<39:12:36, 7.03it/s 1%|| | 7753/100
0000 [15:44<35:29:57, 7.76it/s 1%|| | 7755/1000
000 [15:44<33:08:18, 8.32it/s 1%|| | 7757/10000
00 [15:44<31:17:17, 8.81it/s]


... Ziranyuyan EMNLP save
Name
Date Modified
Size
Kind
..
output
drwxr-xr-x
6/30/2023, 6:00 PM
- -
folder
dir_save
drwxr-xr-x
6/30/2023, 6:30 AM
- -
folder

**5、测试**

```
python main.py --cuda --test --save_path save/dir_save --model_path save/output/SEEK_19999_1000.0000
```



```
ppl:43.9:  18%|█          | 926/5255 [03:13<15:01,  4.80itloss:3.7825 ppl
:43.9:  18%|█          | 927/5255 [03:14<14:59,  4.81itloss:3.7854 ppl:44
.1:  18%|█          | 927/5255 [03:14<14:59,  4.81itloss:3.7854 ppl:44.1:
  18%|█          | 928/5255 [03:14<14:59,  4.81itloss:3.7832 ppl:44.0:  1
8%|█          | 928/5255 [03:14<14:59,  4.81itloss:3.7832 ppl:44.0:  18%|
█          | 929/5255 [03:14<14:59,  4.81itloss:3.7808 ppl:43.9:  18%|█|
 929/5255 [03:14<14:59,  4.81itloss:3.7808 ppl:43.9:  18%|█          | 93
0/5255 [03:14<15:00,  4.80itloss:3.7849 ppl:44.0:  18%|█          | 930/5
255 [03:14<15:00,  4.80itloss:3.7849 ppl:44.0:  18%|█          | 931/5255
 [03:14<14:59,  4.81itloss:3.7880 ppl:44.2:  18%|█          | 931/5255 [0
3:14<14:59,  4.81itloss:3.7880 ppl:44.2:  18%|█          | 932/5255 [03:1
5<15:02,  4.79itloss:3.7875 ppl:44.1:  18%|█          | 932/5255 [03:15<1
5:02,  4.79itloss:3.7875 ppl:44.1:  18%|█          | 933/5255 [03:15<14:5
6,  4.82itloss:3.7877 ppl:44.2:  18%|█          | 933/5255 [03:15<14:56,
 4.82itloss:3.7877 ppl:44.2:  18%|█          | 934/5255 [03:15<14:58,  4.
81itloss:3.7879 ppl:44.2:  18%|█          | 934/5255 [03:15<14:58,  4.81i
tloss:3.7879 ppl:44.2:  18%|█          | 935/5255 [03:15<14:58,  4.81itlo
ss:3.7879 ppl:44.2:  18%|█          | 935/5255 [03:15<14:58,  4.81itloss:
3.7879 ppl:44.2:  18%|█          | 936/5255 [03:15<15:00,  4.80itloss:3.7
871 ppl:44.1:  18%|█          | 936/5255 [03:15<15:00,  4.80itloss:3.7871
 ppl:44.1:  18%|█          | 937/5255 [03:16<14:59,  4.80itloss:3.7861 pp
l:44.1:  18%|█          | 937/5255 [03:16<14:59,  4.80itloss:3.7861 ppl:4
4.1:  18%|█          | 938/5255 [03:16<14:58,  4.80itloss:3.7851 ppl:44.0
:  18%|█          | 938/5255 [03:16<14:58,  4.80itloss:3.7851 ppl:44.0:
18%|█          | 939/5255 [03:16<14:56,  4.81itloss:3.7853 ppl:44.0:  18%
|█          | 939/5255 [03:16<14:56,  4.81itloss:3.7853 ppl:44.0:  18%|█
▏          | 940/5255 [03:16<14:57,  4.81itloss:3.7841 ppl:44.0:  18%|█ |
 940/5255 [03:16<14:57,  4.81itloss:3.7841 ppl:44.0:  18%|█          | 94
1/5255 [03:16<14:58,  4.80itloss:3.7833 ppl:44.0:  18%|█          | 941/5
255 [03:16<14:58,  4.80itloss:3.7833 ppl:44.0:  18%|█          | 942/5255
 [03:17<14:59,  4.79itloss:3.7808 ppl:43.9:  18%|█          | 942/5255 [0
3:17<14:59,  4.79itloss:3.7808 ppl:43.9:  18%|█          | 943/5255 [03:1
7<14:55,  4.81itloss:3.7827 ppl:43.9:  18%|█          | 943/5255 [03:17<1
4:55,  4.81itloss:3.7827 ppl:43.9:  18%|█          | 944/5255 [03:17<14:5
5,  4.81itloss:3.7836 ppl:44.0:  18%|█          | 945/5255 [03:17<14:52,
 4.83itloss:3.7878 ppl:44.2:  18%|█          | 945/5255 [03:17<14:52,  4.
loss:3.6062 ppl:36.8: 100%|██████████| 5255/5255 [18:20<00:00,  4.77it/s]
```

测试结果指标评价：

| Models | PPL | DE Acc. | UEI Acc. | REI Acc. | Dist-1 | Dist-2 |
|---|---|---|---|---|---|---|
| **论文** | 37.09 | 41.85 | 34.08 | 25.67 | 0.73 | 3.23 |
| **我们复现** | 36.83 | 41.16 | 37.23 | 23.39 | 0.80 | 3.61 |

6、代码复现结果及分析

```
Loss    Accuracy
3.6062  36.8254 0.4116  0.3723  0.2339  0.0060  0.0233  7.06
Emotion: ['terrified', 'questioning']
Pred Emotions: agreeing, acknowledging, neutral
x_intent:['none', 'to escape', 'scared', 'sad', 'hurt']
x_need:['make sure they are ok', 'make sure they are safe', 'make sure they are okay', 'to be in a car', 'to have a
car']
x_want:['to call the police', 'to apologize to them', 'to get revenge', 'to apologize', 'none']
x_effect:['they get a ticket', 'none', 'they are ok', 'they got hurt', 'they are safe']
x_react:['scared', 'sad', 'nervous', 'hurt', 'angry']
Context:['yeah about 10 years ago i had a horrifying experience . it was 100 % their fault but they hit the water
barrels and survived . they had no injuries but they almost ran me off the road .']
Greedy:i was thinking of my husband and i was in the same thing .
Ref:did you suffer any injuries ?
---------------------------------------------------------------
Emotion: ['terrified', 'questioning', 'guilty', 'questioning']
Pred Emotions: neutral, questioning, acknowledging
x_intent:['none', 'to escape', 'scared', 'sad', 'hurt']
x_need:['make sure they are ok', 'make sure they are safe', 'make sure they are okay', 'to be in a car', 'to have a
car']
x_want:['to call the police', 'to apologize to them', 'to get revenge', 'to apologize', 'none']
x_effect:['they get a ticket', 'none', 'they are ok', 'they got hurt', 'they are safe']
x_react:['scared', 'sad', 'nervous', 'hurt', 'angry']
Context:['yeah about 10 years ago i had a horrifying experience . it was 100 % their fault but they hit the water
barrels and survived . they had no injuries but they almost ran me off the road .', 'did you suffer any injuries ?',
'no i was not hit . it turned out they were drunk . i felt guilty but realized it was his fault .']
Greedy:well , i am glad they were able to get a while on the next time .
Ref:why did you feel guilty ? people really should n't drive drunk .
---------------------------------------------------------------
Emotion: ['questioning', 'grateful']
Pred Emotions: questioning, trusting, acknowledging
x_intent:['to know what happened', 'to share their experience', 'to know about something', 'to be understood', 'to
share their experiences']
x_need:['to be in the same place', 'to be in the same situation', 'to be in a different place', 'to think about what
happened', 'to think about it']
x_want:['to talk about the experience', 'to talk to someone else', 'to tell me about it', 'to tell the truth', 'to
talk about it']
x_effect:['get a new job', 'get a new experience', 'gets a new job', 'gets a new experience', 'is asked to leave']
x_react:['happy', 'satisfied', 'relieved', 'informed', 'good']
Context:['well , can you tell me about your experience ? i think we swapped places']
Greedy:that is so annoying !
Ref:yeah i wanted to tell you about the time i was hit by a drunk driver im so happy to still be alive after that
experience
---------------------------------------------------------------
```

results.txt

```
-----------------------------------------------------------------------
Emotion: ['impressed', 'questioning', 'acknowledging', 'acknowledging']
Pred Emotions: questioning, acknowledging, impressed
x_intent:['none', 'to be entertained', 'to show off', 'have fun', 'to learn']
x_need:['to know about minecraft', 'to learn about minecraft', 'to know about it', 'to learn about it', 'none']
x_want:['to show it to others', 'to take pictures of it', 'to learn more about it', 'to show it off', 'to have fun']
x_effect:['gets tired', 'none', 'is tired', 'gets dirty', 'gets excited']
x_react:['happy', 'proud', 'amazed', 'excited', 'satisfied']
Context:['my friend showed me his minecraft world and it was amazing ! looked like he spent days on it', 'what had he
made in the world ?', 'he had created one giant hole and filled it up with water , super cool']
Greedy:that is so annoying !
Ref:that is super cool , i played a few years back but eventually quit .
-----------------------------------------------------------------------
Emotion: ['disappointed', 'sympathizing']
Pred Emotions: questioning, sympathizing, disappointed
x_intent:['to be with family', 'to go on vacation', 'to save money', 'none', 'to have a vacation']
x_need:['to have a family vacation', 'to find out the reason', 'to find out what happened', 'to make a plan', 'to
make a decision']
x_want:['to find a new vacation', 'to go to the airport', 'to go to the beach', 'to go on vacation', 'to find another
vacation']
x_effect:['has to pay for it', 'has to find another vacation', 'has less money', 'is sad', 'none']
x_react:['sad', 'regretful', 'sad .', 'upset', 'disappointed']
Context:['i had to cancel our family vacation coming up next month .']
Greedy:oh wow , that is so exciting !
Ref:i am really sorry to hear that . i hope everything is alright .
-----------------------------------------------------------------------
Emotion: ['disappointed', 'sympathizing', 'disappointed', 'annoyed']
Pred Emotions: sympathizing, questioning, consoling
x_intent:['to be with family', 'to go on vacation', 'to save money', 'none', 'to have a vacation']
x_need:['to have a family vacation', 'to find out the reason', 'to find out what happened', 'to make a plan', 'to
make a decision']
x_want:['to find a new vacation', 'to go to the airport', 'to go to the beach', 'to go on vacation', 'to find another
vacation']
x_effect:['has to pay for it', 'has to find another vacation', 'has less money', 'is sad', 'none']
x_react:['sad', 'regretful', 'sad .', 'upset', 'disappointed']
Context:['i had to cancel our family vacation coming up next month .', 'i am really sorry to hear that . i hope
everything is alright .', 'yes , his work told him he could not go after they already approved the time off . guess
next month is going to be busy at his work : (']
Greedy:that is so annoying .
Ref:well , that is just terrible . i hate it when companies jerk you around like that . i hope you will be able to
reschedule it , although i know that would n't make up for the frustration .
-----------------------------------------------------------------------
Emotion: ['content', 'agreeing']
Pred Emotions: content, acknowledging, questioning
x_intent:['none', 'to be happy', 'life is good', 'to be loved', 'happiness']
x_need:['have a good life', 'have a good time', 'to be happy', 'to enjoy life', 'none']
x_want:['to be happy', 'to enjoy the moment', 'to have a party', 'to enjoy the gift', 'to enjoy life']
x_effect:['none', 'gets thanked', 'is loved', 'is happy', 'smiles']
x_react:['happy', 'good', 'satisfied', 'joyful', 'content']
Context:['life is good , it is the gift that keeps on giving']
Greedy:i know what you mean , but i am sure you will be able to get them
Ref:that is a great way to look at it
-----------------------------------------------------------------------
Emotion: ['content', 'agreeing', 'acknowledging', 'jealous']
Pred Emotions: content, acknowledging, grateful
x_intent:['none', 'to be happy', 'life is good', 'to be loved', 'happiness']
x_need:['have a good life', 'have a good time', 'to be happy', 'to enjoy life', 'none']
x_want:['to be happy', 'to enjoy the moment', 'to have a party', 'to enjoy the gift', 'to enjoy life']
x_effect:['none', 'gets thanked', 'is loved', 'is happy', 'smiles']
x_react:['happy', 'good', 'satisfied', 'joyful', 'content']
```

上图节选自result.txt，以一段生成的对话为例：

提供对话内容 `['my friend showed me his minecraft world and it was amazing ! looked like he spent days on it', 'what had he made in the world ?', 'he had created one giant hole and filled it up with water , super cool']` `（“我的朋友向我展示了他的minecraft世界，真是太棒了！”看起来他在上面花了好几天’，‘他在这个世界上创造了什么？’，‘他创造了一个巨大的洞，并把它填满了水，超级酷’]）`，

**模型成功生成具有合理同理心的对话回复：** **`that is super cool , i played a few years back but eventually quit`** . `（这太酷了，我几年前玩过，但最终放弃了。）`

以及不具备同理心的回复：`that is so annoying ! 太烦了!`

我们训练出的模型实现了成功生成有**同理心的对话内容。**

7、算法架构改进

**改进思路**

1. 引入更精细的情感表示：当前的方法通常使用离散的情感标签来表示对话情感，可能无法捕捉到情感的连续性和细微差别。可以探索使用连续的情感表示方法，尝试基于向量空间的情感表示或情感维度的建模，以更准确地捕捉和表达情感。
2. 改进知识选择机制：当前的方法通常使用基于注意力机制的知识选择模块，但这可能存在选择错误或不准确的问题。我们探索更高级的知识选择机制，使用基于图模型的知识选择方法（图注意力网络Graph Attention Network，GAT），以提高对相关知识的选择准确性。
3. 融合上下文的建模：当前的方法通常将对话历史视为一个序列，忽略了上下文之间的复杂关系。可以探索更强大的上下文建模方法，我们引入图神经网络（Graph Convolutional Networks，GCN），以更好地捕捉上下文之间的依赖关系和长期依赖。
4. 结合情感生成和知识生成：当前的方法通常将情感生成和知识生成作为两个独立的模块处理。可以探索将情感生成和知识生成相互融合的方法，我们在生成过程中引入情感感知的知识选择机制。
    a. 结合注意力机制来加强情感感知的知识选择。使用自注意力机制（Self-Attention），来计算知识库中每个文本与查询词之间的相关性权重，然后将这些权重应用于知识选择过程。
    b. 我们还使用词向量模型（如Word2Vec、GloVe）来计算词汇的语义相似度。通过计算查询词与知识文本之间的相似度，可以选择最相关的知识作为系统的输出。

这些优化改进旨在提高方法的表达能力、选择准确性和上下文建模能力，以实现更准确、更连贯和更具情感共鸣的共情对话生成。具体的改进方法需要根据具体问题和数据集的特点进行细致的研究和实验。

考虑完成时间有限，我们仅成功实现4中的改进思路，将结合情感生成和知识生成结合起来，并在本文的原有评价指标下进行定量实验，实验结果优于未改进结果，实验结果如下：

原文采用**困惑度（Perplexity，PPL）和不同性（Distinct-n，Dist-n）**（Li等，2016）作为主要的自动评估指标来评估生成质量。对于对话情感识别和我们新引入的两个任务，包括细粒度情感-意图标记和回应情感-意图预测，我们采用对**话情感准确度（DE Acc.）、话语情感-意图准确度（UEI Acc.）和回应情感-意图准确度（REI Acc.）**作为评估指标。

| Models | PPL | Dist-1 | Dist-2 | DE Acc. | UEI Acc. | REI Acc. |
|---|---|---|---|---|---|---|
| **original model** | 37.09 | 0.73 | 3.23 | 41.85 | 34.08 | 25.67 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **revised model** | 36.49 | 0.8 | 3.77 | 43.38 | 34.71 | 26.22 |

p.s.上述指标除PPL外，结果越大表明生成的具同理心内容越理想。

**实验证明了我们提出的修改方案的可行性。模型结合情感生成和知识生成后，生成具同理心的对话内容更理想。**

# 六、思考及本门课程的建议和改进

## 1、总结：

在本文中，我们研究共情对话生成的任务。强基线忽略了对话的情感流。因此，我们提出了一种用于共情对话生成的序列编码和情绪-知识交互(SEEK)方法，通过感知情境的情绪流动，协调常识性知识和细粒度情绪以避免冲突，从而预测目标反应的正确情绪。

在话语级注释的EMPATHETICDIALOGUES上的实验表明，我们的模型优于基线，并且烧烧研究表明，我们的模型的所有组成部分，编码策略和常识知识都有效。

在未来，我们将专注于进一步的用法，例如:为共情系统提供在线情感援助，并尝试提高我们的模型在其他数据集上的规范化能力。

## 2、限制

我们工作的局限性主要来自于移情对话生成任务中数据集的缺乏。虽然有几个新发布的大规模数据集(Liu et al .， 2021;Welivita et al, 2021)，大多数研究只能在英语语料库empathticdialogues上进行。另一个限制是评估指标的问题。如Liu等人所述

(2016)，标准自动评价指标得分与人工评价结果不一致。缺乏针对特定任务的自动度量使得评估移情对话生成变得困难。

## 3、本门课程的建议和改进

课程十分完满，希望以后的课程可以多些实验内容。