Affective Neural Response Generation

Nabiha Asghar, Pascal Poupart, Jesse Hoey, Xin Jiang, Lili Mou
Cheriton School of Computer Science, University of Waterloo, Canada Noah's Ark Lab, Huawei Technologies, Hong Kong

Motivation

Emotion information is very important for generating reasonable dialogs.
eg.

- I hage you! hehe (no)
- I hate you! I still love you (good)

Approaches

- Affective Word Embeddings
- Affective training objectives
- Affectively diverse beam search

Affective Word Embeddings

VAD dictionary(3D): Valence (V, the pleasantness of a stimulus), Arousal (A, the intensity of emotion produced), Dominance(D, the degree of power/control exerted by a stimulus)

eg.

ecstatic [6.45,6.95,5.63]

bored [2.95,3.65,4.96]

$$\texttt{W2AV}(w) = \begin{cases} \text{VAD}(l(w)), & \text{if } l(w) \in dict \\ \vec{\eta} = [5, 1, 5], & \text{otherwise} \end{cases}$$

$\hat{\eta}$ means OOV (neutral word)

Affective Loss Functions

1.Minimzing Affective Dissonance

$$L^i_{\texttt{DMIN}}(\theta) = -(1-\lambda)\log p(y_i|y_1,\cdots,y_{i-1},X) + \lambda\,\hat{p}(y_i)\left\|\sum_{j=1}^{|X|}\frac{\texttt{W2AV}(x_j)}{|X|} - \sum_{k=1}^{i}\frac{\texttt{W2AV}(y_k)}{i}\right\|_2$$

Minimzing the dissonance between the prompts and the responses.

2.Maximizing Affective Dissonance

Maximizing the dissonance between the prompts and the responses.

3.Maximizing Affective Content

$$L^i_{\texttt{AC}}(\theta) = -(1-\lambda)\log p(y_i|y_1,\cdots,y_{i-1},X) - \lambda\,\hat{p}(y_i)\left\|\texttt{W2AV}(y_i) - \vec{\eta}\right\|_2$$

Maximizing the dissonance between the responses and the neutral sentences, namely making the response more emotional.

Affectively Diverse Beam Search

这个解码是基于 diverse beam search 的，diverse beam search 是在 beam search 的基础上，把 top B 序列分成 G 个 group，然后加一个额外的惩罚项，这个惩罚项的目的在于让 group A 中下一时刻的 search 出来的东西尽可能和其他 group 下一时刻 search 出来的东西不一样，这就达到了 diverse 的目的。这里也从两个 level 来衡量，一个是单个词的情绪，一个是整个句的情绪，标准就用向量的 cosine 函数来衡量相似度。

Experiments

First phase, using the general cross-entropy loss for 40 epochs. In the second phase, each model was fine-tuned using the affective loss functions for 10 epochs.