

面向一致性对话生成的对抗匹配网络与目标侧注意力机制研究^①蔡恒毅^{②*} 王成瑞^{**} 宋永浩^{③*} 袁旭^{**} 张程^{*} 赵晓芳^{*}

(* 中国科学院计算技术研究所 北京 100190)

(** 中国科学院大学 北京 100049)

摘 要 序列到序列(seq2seq)方法在开放域对话生成领域中备受研究学者的关注。然而,标准的序列到序列模型容易产生语义冲突和不连贯的对话回复,这种不一致性是现有系统生成的回复显著有别于人类真实对话的重要原因之一。对话生成中的一致性既包括回复内部的语义一致性,也包括上文与其回复之间的外部关联性。本文提出了一个新的对话生成框架,称为基于张量匹配的生成式对抗网络(MatchGAN),以提高对话回复与其上文之间的外部关联性。与传统的基于最大似然估计的方法不同,该框架通过基于序列到序列模型的生成器和基于张量匹配网络的判别器之间的对抗学习来生成与上文相关的回复。通过使用匹配网络对上文与回复之间的多维关系进行建模,该模型所产生的回复更加符合人类对话的特点。此外,本研究进一步引入了目标侧注意力机制来增强所产生回复的内部语义一致性。实验结果表明,本文提出的框架能够产生高质量的对话回复,在量化指标评价和人工评测方面均优于其他基线方法。

关键词 生成式对话模型;神经张量网络;对抗学习;目标侧注意力机制

0 引 言

近年来,作为人工智能领域中的一重要课题,如何构建一个与人类进行自然沟通的开放域对话系统受到了众多研究者的关注。通常此类开放域对话系统可以采用基于序列到序列(sequence-to-sequence, seq2seq)方法的模型来进行构建^[1-8]。然而,此类模型的效果仍远未达到预期。具体而言,基于序列到序列方法的对话模型倾向于生成一些较短且无实际意义的回复,如“I don’t know”,或者一些较长但内部语义矛盾的回复,如“I love to cook, I’m a vegan so I don’t like cooking”^[9-10]。由于这些回复往往存在上文相关性较差且句内语义不一致的问题,尽管模型给出的回复看起来流畅且语法正确,但

是基于序列到序列模型的人机对话系统仍缺乏实用性。

导致上述问题的部分原因在于人类对话上文与其回复之间普遍存在的一对多映射关系^[11]。在对话任务中,即使对于相同的上文,也可能存在多个具有不同语义和表达方式的回复与之对应。尽管对话任务普遍采用翻译任务中的编码—解码器模型,但对话任务中上文 X 和回复 Y 之间的松耦合关系使得其显著区别于翻译任务。在翻译任务中,源语言和目标语言在语义上是一致且对齐的,而对话任务中上文与回复之间则无此对齐关系,即目标分布 $P(Y|X)$ 具有更高的条件熵,这也意味着对话任务中,仅靠编码—解码器模型无法充分刻画上文与回复之间的关系。另外,序列到序列模型很容易退化为缺乏对上文关注的语言模型^[12](在其他任务中也

① 国家自然科学基金(U1836111, U1736106)和国家重点研发计划(2018YFB0904503)资助项目。

② 男,1995年生,博士生;研究方向:自然语言生成,对话系统;E-mail: caihengyi@ict.ac.cn。

③ 通信作者, E-mail: songyonghao@ict.ac.cn。

(收稿日期:2020-12-15)

有类似发现,例如故事生成^[13])。这些现象表明,以最大似然估计作为其训练准则的序列到序列模型,不足以充分捕获对话上文和回复之间的多维语义关系,而这种上下文之间的复杂依赖关系对于建模人人对话至关重要。

文献[14]尝试通过提升上文和回复之间的语义相似度来增强它们之间的关联性。然而,对话上文与其相应的回复并不一定要在语义上相同。用语义相似度来刻画对话上文与其回复之间复杂的、多维度的关系是不够的。在现有的对话生成研究中,仍然没有一种有效的方法来刻画这种隐式关系。此外,以往的模型很少考虑生成回复的内部语义一致性,而这种句内语义一致性的有效建模有助于避免模型产生语义自相矛盾的回复。

本文面向对话生成任务中的一致性问题,旨在提高对话生成中回复的上文相关性和内部语义一致性。具体而言,对于上文相关性(外部一致性),本研究提出了一种基于多维度关系匹配的生成式对抗网络(matching-based generative adversarial network, MatchGAN)。不同于其他基于生成式对抗网络(generative adversarial networks, GANs)^[15]的回复生成模型^[16-17]中使用二分类器来作为判别器的做法,本文引入了一种用于多维度关系匹配的张量网络来学习对话上文和其回复之间关系的分布式表示,并以此来判别给定的回复是否与上文匹配。这里假设来自训练集中人类对话的上文与回复之间存在匹配关系,而产生自机器的回复则与其上文构成相应的负样本对。给定一条上文及模型生成的回复,本文中的判别器可以输出它们之间的匹配度分数,然后利用该匹配度来指导生成器生成与上文更为相关的回复。此外,本文中的判别器也能够显式地利用一对多的语料来更好地建模对话中的多映射关系。对于内部语义一致性(内部一致性),本研究将目标侧注意力机制引入到对话生成器中,使得模型在回复生成的过程中能够更为细粒度地关注到已经生成的前文回复,从而更好地保持句内语义的一致性。简而言之,本文中的张量匹配网络用于鼓励生成器产生与上文更为相关的回复(外部一致性),而目标侧注意力机制则使得模型能够产生语义更为一致的回

复(内部一致性)。

本研究在 3 个大规模的对话数据集上评估了本文的方法,包括 Baidu-Tieba、StackOverflow 和 PERSONA-CHAT。实验结果表明,本文方法在量化指标评测和人工评测下均优于多个基线模型。本文的贡献可以总结为以下 3 点:(1)本文提出了一种新的对话生成框架 MatchGAN,该框架能够显式地建模对话上文和其相应回复之间的多维语义关系,并基于此生成与上文更为相关的对话回复;(2)本研究在对话解码器中引入了目标侧注意力机制,以更好地建模所生成回复的内部语义一致性;(3)本研究采用了 3 组具有不同特征的对话数据集进行了详尽的实证研究。实验结果证实了本文方法的有效性。

本文的整体组织如下,第 1 节是相关研究工作,第 2 节描述本文方法的整体设计,第 3 节给出实验结果及分析,最后在第 4 节进行总结。

1 相关工作

开放域下的对话建模可以看作是一个序列到序列的生成问题。文献[18]将对话生成建模为统计机器翻译的问题,并采用数据驱动的方式进行模型训练。近年来,深度学习在机器翻译和语言模型中的成功应用进一步启发了众多学者对于神经对话生成的研究^[1,7,19-20]。

尽管起源自机器翻译任务的序列到序列模型在对话建模中有着不错的表现,然而,文献[11]指出,对话生成任务中的回复往往具有较高的多样性,这一特点使得其区别于上下文具有严格语义对齐关系的机器翻译任务。这也一定程度上解释了为何将序列到序列模型应用于对话生成任务后,模型总是倾向于给出通用甚至无意义的回复。为了解决这一问题,文献[10]使用互信息最大化这一准则来训练对话模型,旨在减少通用回复的生成。文献[21]则显式地利用对话上文来进行实体预测,并通过使用一个外部的知识库来挖掘相关实体以提供给模型更多的信息用作对话回复生成。文献[22]则进一步提出在解码过程中使用基于门控的线索词融合方法来提高生成回复的信息量。条件变分自编码器^[23]近年

来也被广泛用于对话生成模型的训练并表现出不错的效果。与这些工作相比,本文工作旨在更加充分地挖掘对话上文与回复之间的复杂语义关系并用以指导对话模型的学习。

除了提高对话回复信息量的研究工作之外,对话生成中的一致性问题的研究也受到了关注。文献[24]使用基于注意力机制的语言模型来生成更为连贯的对话。与之不同,本文工作使用一个独立的匹配模型用以学习上文与回复之间的连贯性。文献[14]使用强化学习的训练方式以使得模型生成的回复与相应的上文在语义上要尽可能地相似,然而,上下文的关系往往涉及多个维度的因素,如话题转移、情感交互、下文引导等,语义相似度并不能充分地建模对话中复杂的上下文关系。本文引入的张量匹配网络旨在从多个维度更为充分地建模上下文之间的关系。

本文也与对话建模中的生成式对抗网络^[16-17]等工作相关。在这些基于 GAN 的回复生成模型中,其判别器以生成的回复作为输入来完成二分类过程,忽略了对话上文与回复之间的匹配关系。本文工作则设计了一种匹配模型来显式地建模这种匹配关系。对于目标侧注意力机制的相关工作,文献[25]将其应用于机器翻译任务中,然而,其在对话建模中的作用尚未被充分研究。文献[26]将部分回复语句与对话上文进行合并以使得模型能够在回复生成的过程中考虑前文信息,相比于这种数据预处理的方式,本文以一种更为自然的方式设计了回复生成

中的目标侧注意力机制。

2 本文方法

2.1 模型概览

MatchGAN 遵循生成式对抗训练的学习范式来显式地建模所生成的回复和其上文之间的匹配模式。该框架的整体概览见图 1,其中包含一个基于序列到序列模型的回复生成器和一个基于神经张量网络的匹配模型来刻画上下文之间细粒度的语义关系。对于给定的上文 $X = x_1, x_2, \dots, x_{T_X}$, 生成器网络 G 首先联合使用源端注意力机制和目标端注意力机制采样出回复 $\hat{Y}$ 。匹配器网络 M 在训练的过程中,将上文与生成的回复作为其负样本,而上文与真实回复则作为其正样本。匹配器网络 M 所给出的匹配度分值则被用来指导生成器 G 的学习。整个框架使用对抗学习的训练范式来进行优化,并使用策略梯度方法来解决因回复采样导致的不可微问题。需要说明的是,这里的生成器虽然以基于循环神经网络(recurrent neural network, RNN)单元^[27]的编码—解码器作为示例,其他的结构如 Transformer^[28]也是适用的。本文首先介绍用以保持回复语义内部一致性的目标侧注意力机制,然后介绍基于神经张量网络的匹配式判别器,该判别器通过区分上文回复对是人人对话还是人机对话来显式地建模生成回复的外部相关性。

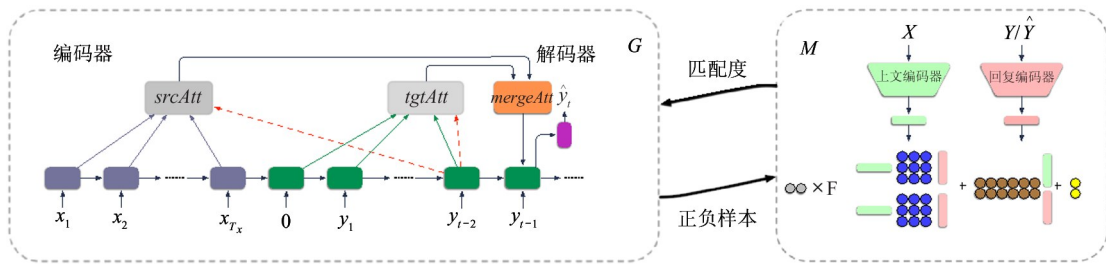


图1 本文框架 MatchGAN 的模型概览图

2.2 回复生成中的目标侧注意力机制

在本文所提出的生成式对抗学习框架中,以序列到序列模型作为其中的生成器 G 。对于该生成器,它的输入为长度为 T_X 的对话上文 $X = x_1, x_2,$

$\dots, x_{T_X}$, 其中单词 x_i 均来自于词表 Γ 。生成器 G 的输出为长度为 T_Y 的对话回复 $Y = y_1, y_2, \dots, y_{T_Y}$, 其中单词 y_i 来自于词表 Λ 。具体地,给定上文输入 X , 采用 RNN 单元实现的序列到序列模型按照编码—

解码器的建模方式来生成其对应的回复 Y 。首先,编码器循环地读入 X 中的单词,并将其编码为一个上文向量 $\mathbf{c}$,然后,解码器网络根据该上文向量 $\mathbf{c}$ 再逐词逐句地生成回复 Y 。典型地,使用长短期记忆(long short-term memory, LSTM)单元^[27]构建的序列到序列模型按照式(1)来建模条件概率 $P(Y|X)$:

$$\begin{aligned} P(Y|X) &= \prod_{t=1}^{T_Y} P(y_t | y_{[0:t-1]}; X) \\ &= \prod_{t=1}^{T_Y} P(y_t | y_{t-1}, h'_{t-1}, \mathbf{c}) \end{aligned} \quad (1)$$

其中, h'_{t-1} 是对应于 y_{t-1} 的隐状态。

为了更好地建模该条件概率 $P(Y|X)$, 文献[29]提出了在解码过程中使用注意力机制的方法来使得解码器在生成每个单词的时候能够动态地关注到上文序列的不同部分。具体而言,使用注意力机制的序列到序列模型按照式(2)来定义 y_t 的条件生成概率:

$$\begin{aligned} P(y_t | y_{[0:t-1]}; X) &= \\ &LSTM(y_{t-1}, h'_{t-1}, \text{Attention}(h'_{t-1}, \mathbf{x})) \end{aligned} \quad (2)$$

其中, $\text{Attention}(h'_{t-1}, \mathbf{x}) = \sum_{i=1}^{T_X} a_{t,i} h_i$, $a_{t,i}$ 是时间步 t 时刻对应于 h_i 的注意力权重。 $a_{t,i}$ 可以按照式(3)进行计算:

$$e_{t,i} = g(h'_{t-1}, h_i); a_t = \text{softmax}(e_t) \quad (3)$$

其中, h_i 是对应于 x_i 的隐藏层状态。 $g(\cdot)$ 是用于计算 h_i 和 h'_{t-1} 之间对齐关系的相似度函数。本文使用双线性函数来实现函数 $g(\cdot)$, 即: $g(h'_{t-1}, h_i) = v^T \tanh(W_h h_i + W_{h'} h'_{t-1})$, 其中, v , W_h 和 $W_{h'}$ 是可学习的网络参数。

上述的这种源端注意力机制在翻译任务中显著提高了模型的表达能力,然而,对话生成任务中的上下文之间并不像翻译任务中源语言和目标语言之间那样有严格的语义对齐关系。在对话生成过程中,为了保持所生成回复的内部语义一致性和可读性,模型不仅应该关注于给定的对话上文,更应该关注于在解码过程中已经生成的前缀回复部分。因此,本文提出目标侧注意力机制来增强现有的源端注意力模型,使得模型在回复生成过程中解码每一个回复词的时候也对所生成的前文语句执行注意力操

作。如图1中的左半部分所示,生成器 G 中的解码器在时间步 $t-1$ 时刻接收到的输入有:上一步预测词 y_{t-1} , 上一步隐藏层状态 h'_{t-1} , 源端的注意力表示 $\text{srcAtt} = \text{Attention}(h'_{t-1}, \mathbf{x})$ 和目标侧注意力表示 $\text{tgtAtt} = \text{Attention}(h'_{t-1}, \mathbf{y}_{<t-1})$, 其中, $\mathbf{y}_{<t-1}$ 代表时间步 $t-1$ 之前的解码器输入。此时,回复词 y_t 的生成概率可以建模为

$$\begin{aligned} P(y_t | y_{[0:t-1]}; X) &= \\ &LSTM(y_{t-1}, h'_{t-1}, \text{mergeAtt}(\text{srcAtt}, \text{tgtAtt})) \end{aligned} \quad (4)$$

其中,函数 $\text{mergeAtt}(\text{srcAtt}, \text{tgtAtt})$ 用于合并源端的注意力表示和目标侧的注意力表示。本文将其实为一个多层感知器模型。通过引入目标侧注意力机制,生成器 G 在解码回复词的过程中能够显式地考虑到已经生成的前部分语句片段,因此能够缓解以往回复生成中语句内部语义冲突的问题。

2.3 基于张量网络的多维度上下文关系建模机制

以往工作中通常使用语义相似度来表征对话上文与回复之间的关系,这样做的问题在于,对话上文与其回复之间的关系并不具备严格的语义对齐等特点,即对于某条对话上文,合适的回复并不一定要与其语义相同,此外,这种上下文的关系往往涉及多个维度的因素,如话题转移、情感交互、下文引导等方面。由于这些原因,基于句子嵌入的语义相似度方法并不能充分地建模对话中复杂的上下文关系。

为了有效建模对话上文与其回复之间复杂的语义关系,本文设计了一种基于张量网络的匹配模型来对给定的上文—回复对之间的关系进行判别。生成器 G 进而使用该判别信号来进行训练。如图1中的右半部分所示,判别器采用类似于孪生网络的结构进行实现,其中包括2个对称的RNN用于对上文和回复进行编码以及一个神经张量网络用于学习对话上文和回复之间的匹配关系。具体而言,给定上文 X 和回复 Y , 本文使用各自对应的RNN网络来对其进行编码,编码得到的结果为上文的向量表示 $\mathbf{x} \in R^{n_x}$ 和回复的向量表示 $\mathbf{y} \in R^{n_y}$ 。受文献[30]启发,本文使用一个由双线性函数实现的神经张量网络从多个维度建模 $\mathbf{x}$ 和 $\mathbf{y}$ 之间的关系,其中,上文 $\mathbf{x}$ 和回复 $\mathbf{y}$ 之间的多维交互关系按式(5)进行建模:

$$m(\mathbf{x}, \mathbf{y}) = uf(\mathbf{x}^T \mathbf{W}^{[1:d]} \mathbf{y} + \mathbf{V} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} + \mathbf{b}) \quad (5)$$

其中, f 是一个对输入进行逐元素非线性变换的函数。 $\mathbf{W}^{[1:d]} \in R^{n_x \times n_y \times d}$ 是一个张量参数, 双线性张量乘积 $\mathbf{x}^T \mathbf{W}^{[1:d]} \mathbf{y}$ 的结果是一个 d 维的向量 $\mathbf{h} \in R^d$, 其中 $\mathbf{h}$ 的每一个元素 h_i 可由张量 $\mathbf{W}^{[1:d]}$ 的一个分片计算求得: $h_i = \mathbf{x}^T \mathbf{W}^i \mathbf{y}$ 。其他的参数为标准的神经网络参数, 包括 $\mathbf{V} \in R^{d \times (n_x + n_y)}$ 、 $\mathbf{b} \in R^d$ 和 $\mathbf{u} \in R^d$ 。

利用上述的匹配模型, 网络可以对上文和回复之间的交互关系和各自的表示进行联合学习: 上文和回复的向量表示由各自的 RNN 网络学得, 而它们之间的交互关系则由神经张量网络进行建模。该匹配模型 M 的输出则为上文与回复之间匹配关系的量化得分, 记为 $M(\{X, Y\})$ 。

为了使得生成器 G 产生与上文 X 更为契合的回复 Y , 在对抗训练的过程中, 对于采样于训练集合的 $\langle$ 上文, 回复 $\rangle$ 对, 系统要求匹配器 M 将其作为正样本并给出较高的匹配度分数, 对于来自模型输出的 $\langle$ 上文, 回复 $\rangle$ 对, 则要求匹配器 M 将其作为负样本并给出较低的匹配度分数。对于匹配器的学习, 本文采用下面的损失函数:

$$L = -\mathbb{E}_{X \sim P_h, Y \sim P_h} [M_\phi(\{X, Y\})] + \mathbb{E}_{X \sim P_h, Y \sim G_\theta} [M_\phi(\{X, Y\})] \quad (6)$$

其中, 生成器 G 的参数为 θ , 匹配器 M 的参数为 ϕ 。 P_h 代表给定训练集的数据分布。 $Y \sim P_h$ 表示回复 Y 从训练集中采样, 而 $Y \sim G_\theta$ 则表示回复 Y 从生成器所定义的分中采样。

2.4 基于策略梯度方法的联合训练

由于自然语言的离散性特点, 使用判别器模型得到的学习信号来指导生成器训练的方法在应用于对话建模的时候存在梯度无法回传的问题。因此, 本文使用策略梯度^[31]的方法来解决这一问题。给定上文输入 X , 本文模型通过从生成器 G 中进行采样从而得到回复 Y 。这里将生成器 G_θ 视作一个策略网络。匹配模型 M_ϕ 以给定的上文 X 和生成的回复 Y 作为输入, 并输出其对应的匹配度分数 $M_\phi(\{X, Y\})$ 。该分数即作为从策略网络 G_θ 中进行采样所对应的奖励值。本文框架的训练目标则是训练该策略网络 G_θ 以最大化所生成回复的期望奖励值, 如

式(7)所示:

$$J(\theta) = \mathbb{E}_{Y \sim G_\theta} (M_\phi(\{X, Y\}) \mid \theta) \quad (7)$$

由于 Y 是从 G_θ 中采样而来, 使用该目标函数无法直接对参数 θ 求导, 因此, 本文使用式(8)来对参数 θ 的梯度进行估计:

$$\begin{aligned} \nabla_\theta J(\theta) &\approx [M_\phi(\{X, Y\}) - b(\{X, Y\})] \nabla \log G_\theta(Y \mid X) \\ &= [M_\phi(\{X, Y\}) - b(\{X, Y\})] \end{aligned}$$

$$\nabla \sum_{t=1}^{T_Y} \log P(y_t \mid y_{[0, t-1]}; X) \quad (8)$$

其中, $G_\theta(Y \mid X)$ 为给定 X 时回复 Y 的条件概率。 $b(\{X, Y\})$ 是为了降低训练方差而引入的基线值函数。给定状态值和其对应的奖励值, 该基线值函数可以使用 L_1 损失函数来训练从而预测基线奖励值。对于匹配器 M_ϕ , 它以训练集合中的 $\langle$ 上文, 回复 $\rangle$ 对作为正样本, 以模型输出的 $\langle$ 上文, 回复 $\rangle$ 对作为负样本, 通过最小化式(6)中的损失函数来进行训练。生成器 G_θ 的训练和匹配器 M_ϕ 的训练交替进行, 即训练其中一项的时候固定另外一项的参数, 以保证模型学习的稳定性。

采用上述的策略梯度方法进行训练时, 由于匹配器的输入是对话上文和整句回复, 因此生成器 G_θ 只能在整个回复都解码完毕之后才能接收到匹配器求得的奖励信号, 这导致用于训练生成器 G_θ 的奖励信号非常稀疏, 模型的学习因此也不够有效。

本文采用蒙特卡洛采样^[32]的方法来缓解这一问题, 使得生成器 G_θ 对于解码过程中的每一步都能接收到相应的奖励信号。具体而言, 对于一个长度为 T 的序列, 在解码到第 t 个时间步时, 蒙特卡洛采样使用一个临时的策略 G'_θ 来解码出剩余的 $T-t$ 个单词, 该采样过程重复 N 次。如此操作之后, 对于相同的上文 X , 便得到了 N 条完整且不同的回复 $\{Y^1, Y^2, \dots, Y^N\}$ 。然后将这 N 对 $\langle$ 上文, 回复 $\rangle$ 样本送入匹配器 M 便可求得 N 个匹配度分数。本文使用这些匹配度分数的均值来近似估计第 t 个时间步时解码单词的奖励值。用作采样的临时策略网络 G'_θ 可以使用和生成器 G_θ 相同的参数, 并随着 G_θ 的训练也相应地更新其网络参数。

上述的蒙特卡洛采样估计奖励值的方法存在一个显著问题, 即每一步解码均需执行 N 次采样, 计算开销是原方法的 NT 倍。本文中, 令 N 值随时间

步 t 逐级减小从而缓解计算开销过大的问题,即 $N = \left\lfloor N_0 \times (1 - \frac{t-1}{T}) \right\rfloor$ 。本文中, N_0 设置为 10。

使用对抗学习的方法对本文框架 MatchGAN 进行训练时,由于匹配器 M_ϕ 相比于生成器 G_θ 更容易学习,因此训练后期 G_θ 的学习往往存在梯度消失的问题。换句话说,由于匹配器 M_ϕ 很快收敛,因此对于生成器 G_θ 产生的回复,匹配器 M_ϕ 给出的奖励值往往过小,使得生成器 G_θ 无法继续训练。针对这一问题,本文借鉴文献[33,34]中的思想,在将匹配器给出的奖励值送入生成器之前,对其进行缩放操作。具体地,对于生成器给出的 B 条长度为 T 的回复,匹配器相应地给出了 $B \times T$ 个奖励值,记为 $\mathbf{Z}_{B \times T}$ 。对于奖励值矩阵中的第 t 列 $\mathbf{Z}'$ 中的每一个元素对其做如下的缩放操作:

$$\mathbf{Z}'_i = \sigma(\delta \cdot (0.5 - \frac{\text{rank}(i)}{B})) \quad (9)$$

其中, $\text{rank}(i)$ 是第 i 个元素在该列中按照由高到低顺序的排名。 δ 是一个用以控制缩放强度的超参数,本文中设置为 12。 $\sigma(\cdot)$ 是 sigmoid 激活函数。

3 实验与分析

本节描述本文框架的实验仿真,包括数据集设置、对比模型与实验参数设置、实验结果和实验分析 4 部分内容。

3.1 数据集设置

实验在 3 类人人对话数据集上进行,其中包括中文多轮对话数据集 Baidu-Tieba、英文一对多单轮对话 StackOverflow 和英文多轮对话数据集 PERSONA-CHAT。

Baidu-Tieba 数据集 该数据集包含了从百度贴吧抓取的百万条多轮对话,其中,100 万条语料用于训练,5000 条语料用于验证,10 000 条语料用于测试。该数据集中句子的长度被控制在 6 ~ 80 中文字符之间。经过词化器处理后,词表大小为 5564。

StackOverflow 数据集 该数据集是一个一对多的对话数据集,即一条对话上文有多条回复与之对应。使用该数据集来验证本文模型能否有效利用人人对话中的多映射关系。本文从 StackOverflow

论坛上抓取了 50 多万条对话语料,其中,50 万条语料用作训练集,3000 条语料用作验证集,5000 条语料用作测试集。语料中的句子长度被控制在 10 ~ 100 中文字符之间,词表大小为 24 012。每条对话上文平均对应于 2.2 条回复。

PERSONA-CHAT 数据集 该数据集共包含 139 239 条使用众包平台 Amazon mturk 收集而来的对话语料^[35],其中,131 438 条数据被用作训练集,7801 条被用作验证集和测试集,词表大小为 9293。

3.2 实验设置

基线对比方法。将本文所提出的方法与下列模型进行了对比,包括:(1) Seq2Seq-attn:一种基于源端注意力机制的序列到序列模型^[29],该模型使用极大似然估计方法进行训练。(2) Seq2Seq-tgt-attn:一种基于源端注意力机制和目标侧注意力机制的序列到序列模型,该模型用来验证目标侧注意力机制。(3) Adver-REGS:一种基于对抗训练的对话生成模型^[16],其中生成器为一个序列到序列模型,判别器是一个二分类器。(4) PG-BLEU^[36]:该模型使用策略梯度方法直接优化特定任务评价指标,如 BLEU^[37]。(5) MMI:一种基于最大化对话上文和回复之间互信息的对话生成模型^[10]。(6) RL-Cos:使用强化学习的训练方式以使得模型生成的回复与相应的上文在语义上要尽可能地相似^[14]。

为了更好地评估本文的方法,本文设计了模型的变体 MatchGAN w/o Tgt-attn,其中,MatchGAN w/o Tgt-attn 模型的生成器没有使用目标侧注意力机制。

实验参数设置。本研究使用 ParlAI^[38]对本文所提出的对话学习框架和相关基线模型进行了实现。所有模型的训练均使用英伟达 P40 型显卡。模型的具体实现细节如下:对于数据集 Baidu-Tieba,由于其词语分布长尾效应和词切分错误等问题^[39],本文使用字级别的向量嵌入来对其进行表征学习。对于数据集 StackOverflow 和 PERSONA-CHAT,本文使用 GloVe 预训练词向量^[40]来初始化模型的词嵌入层。对于训练过程,本文使用了双层的 LSTM 网络来构建本文框架中的生成器和判别器中的编码器。由于策略梯度方法的不稳定性,首先使用极大

似然估计方法来预训练本文框架中的生成器,而后
再进一步训练整个模型。为了避免过拟合,本文采
用早期停止 (early-stopping) 策略来选择模型终止
点。对于模型参数设置,LSTM 网络的隐藏层单元
大小为 256,词向量的大小为 300。本文使用 Adam
优化器^[41]来优化模型参数,初始学习率为 0.001。
Dropout 的丢弃率为 0.2。

3.3 实验结果

3.3.1 量化评估结果

使用量化指标来评估开放域对话系统的性能一
直以来都是一个备受关注的问题。为了衡量回复的
信息丰富度、语言流畅性及其与参考回复差异性,本
文采用了多样性指标 $\text{dist}^{[10]}$ 、困惑度指标 $\text{PPL}^{[42]}$ 和

BLEU 指标^[37]。除此之外,本文也采用了对话回复
评测中常用的召回率指标 $R_c@k^{[35]}$ 。该指标的计算
方式为:给定对话上文,使用待评测模型对 C 条候
选回复进行重排序,如果正确的回复被排到了前 k
名,则输出评测结果为正确。 $R_c@k$ 计算了测试集中
评测结果为正确的回复所占的比例。使用该指标对
模型进行评测的潜在假设是:如果对话模型在生成
任务上能够有好的表现,则它在排序任务中也应该
表现良好。本文实验中使用了 $R_{20}@1$ 和 $R_{20}@5$ 这
2 个指标。

首先考察目标侧注意力机制的作用。如表 1 所
示,本文发现,相比于模型 Seq2Seq-attn,增加了目标
侧注意力机制的 Seq2Seq-attn-tgt 模型在 3 类数据集

表 1 在 3 个数据集上的量化评估结果

Baidu-Tieba	PPL	$R_{20}@1$	$R_{20}@5$	BLEU	dist-1
Seq2Seq-attn	44.26	7.63	33.4	11.03	0.3633
Seq2Seq-tgt-attn	43.77	7.67	33.7	11.62	0.3679
Adver-REGS	43.65	7.87	34.6	11.55	0.2343
PG-BLEU	44.11	7.82	34.3	11.52	0.2575
MMI	44.23	7.53	33.8	12.21	0.4561
RL-Cos	44.84	7.46	34.0	12.19	0.4127
MatchGAN w/o Tgt-attn	43.62	7.74	33.9	13.19	0.7766
MatchGAN	42.64	8.33	35.7	13.77	0.8157
StackOverflow	PPL	$R_{20}@1$	$R_{20}@5$	BLEU	dist-1
Seq2Seq-attn	52.54	8.48	37.3	11.21	0.4319
Seq2Seq-tgt-attn	50.97	9.10	39.9	12.57	1.1760
Adver-REGS	49.10	9.04	42.4	11.07	0.6269
PG-BLEU	49.75	9.29	40.8	11.49	0.4392
MMI	53.35	7.73	35.7	11.68	1.5350
RL-Cos	43.62	7.74	33.9	13.19	0.7766
MatchGAN w/o Tgt-attn	48.3	9.91	42.1	13.09	1.6780
MatchGAN	48.03	9.41	42.7	13.42	1.3070
PERSONA-CHAT	PPL	$R_{20}@1$	$R_{20}@5$	BLEU	dist-1
Seq2Seq-attn	37.28	9.27	33.9	16.41	0.4943
Seq2Seq-tgt-attn	32.55	9.60	34.2	16.99	0.7179
Adver-REGS	33.79	9.51	34.4	16.94	0.5999
PG-BLEU	33.94	9.11	34.4	17.36	0.5561
MMI	35.13	8.42	32.1	17.44	0.6231
RL-Cos	33.40	9.55	34.8	16.16	0.8835
MatchGAN w/o Tgt-attn	32.74	9.96	34.9	17.53	1.0280
MatchGAN	32.58	10.10	35.2	17.62	1.2450

的 5 个指标下都有效果上的提升。这说明目标侧注意力机制的确能够提升对话生成模型的效果。相似的提升也可以在表 1 所展示的 MatchGAN 和 MatchGAN w/o Tgt-attn 模型实验结果对比中看出。

对于本文所提出的对抗匹配网络,从表 1 所示的实验结果可以观察到,移除了该对抗匹配网络的 Seq2Seq-tgt-attn 模型相较于 MatchGAN 模型在 3 类数据集上的多项指标下均有明显的性能下降。移除了对抗匹配网络的 Seq2Seq-attn 模型相较于 MatchGAN w/o Tgt-attn 在各评测数据集上的量化指标结果也呈现出类似的趋势。这些实验结果说明,本文所提出的基于张量匹配的对抗学习方法确实有助于提高对话生成回复的质量。

对于基线模型来说,由于 MMI 直接以提升多样性作为其训练目标,其在 dist-1 指标上的得分较高。然而,MMI 模型在 $R_{20}@1$ 和 $R_{20}@5$ 指标上的表现一般,这意味着该模型不能很好地从候选集中识别出目标回复。本文的模型在 dist-1 指标上有着较高的得分,这说明多维度地建模并利用对话上文与回复之间的关系来增强模型的上下文感知能力有助于提高生成回复的多样性及信息丰富度,这与本文的假设一致。

本文所提出的 MatchGAN 框架在 3 类数据集上的多个指标下几乎都带来了明显的效果提升,尤其是在数据集 StackOverflow 上。相比于 Seq2Seq-attn 模型,MatchGAN 在数据集 StackOverflow 上对于 $R_{20}@1$ 指标有 12.87% 的相对提高,对于 dist-1 指标则带来了超过 200% 的相对提高。这些提升的潜在原因在于,StackOverflow 数据集中的一对多关系使

得本文框架中的匹配网络能够更为显式地学习对话中的这种多映射关系,从而使得模型能够生成更为连贯的回复。相比于模型 RL-Cos,由于本文方法对上下文之间关系的建模没有局限于语义相似这一维度,而是试图挖掘多个方面的隐式关系,因此,本文方法在效果上也优于 RL-Cos。对于 $R_{20}@1$ 、 $R_{20}@5$ 、BLEU、dist-1 和 PPL 这些指标,本文框架相比于基线模型均有明显的效果提升,这进一步说明了本文方法中所设计的用以学习上下文多维度关系匹配的张量网络和目标侧注意力机制能够提升生成对话的质量。

3.3.2 人工评测

已有的研究工作^[43]指出,仅依靠自动指标来评价开放域对话系统是不充分的。因此,本文进一步使用了人工评测的方法来评估所提出方法的表现。具体而言,本研究从以下 3 个方面对回复的质量进行人工评测:上文连贯性(连贯性)、内部语义一致性(一致性)以及该回复是否能让与试人员更容易做出进一步应答(启发性)。这 3 个方面的量化评测采用四分制,由低到高分别为 0~3 分。本文随机抽取了 200 条对话样例,并邀请了 3 位本专业的研究人员来对各个系统所产生的回复作出评估。在评测过程中,某条回复所来自的系统对于标注人员是不可见的。给定某条对话上文,各系统分别生成对应的回复,然后标注人员对这些回复进行独立评分。

表 2 展示了人工评测的实验结果。可以看到,本文方法在上文连贯性上的得分要明显高于其他基线模型,这说明,本文提出的多维度上下文关系建模机制能够有效学习并利用对话上文与其回复之间的

表 2 实验数据集上的人工评测结果

模型	Baidu-Tieba			StackOverflow			PERSONA-CHAT		
	连贯性	一致性	启发性	连贯性	一致性	启发性	连贯性	一致性	启发性
Seq2Seq-attn	1.11	1.18	1.11	1.14	1.04	1.10	1.85	1.77	1.81
Seq2Seq-tgt-attn	1.42	1.56	1.30	1.42	1.50	1.20	2.09	2.13	1.97
Adver-REGS	1.26	1.37	1.22	1.32	1.14	1.26	2.01	1.86	1.87
PG-BLEU	1.39	1.33	1.33	1.38	1.24	1.30	1.93	1.77	1.87
MMI	1.22	0.60	0.88	0.88	0.32	0.66	1.20	1.10	1.13
RL-Cos	1.85	1.99	1.84	1.44	1.44	1.38	2.14	2.22	2.15
MatchGAN	2.03	2.07	1.87	1.52	1.90	1.54	2.32	2.47	2.38

匹配关系,使得生成的回复与给定的上文更为契合。另外值得注意的是,MatchGAN 在内部语义一致性指标上也有着较高的得分。相比于 Seq2Seq-attn, Seq2Seq-tgt-attn 模型在内部语义一致性指标上也有明显的提升。这些提升得益于目标侧注意力机制的引入,回复句子在解码时可以关注到序列内部的信息,因而缓解了回复句子内部诸如片段重复与逻辑冲突这些问题。评测结果中的 Cohen's Kappa 分数^[44]表明了标注人员评分的相对一致性。

3.4 实验分析

3.4.1 上文轮数的影响

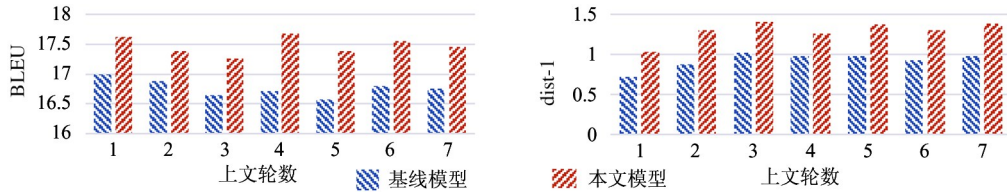


图2 在数据集 PERSONA-CHAT 中上文轮数对模型性能的影响分析

3.4.2 案例分析

本节给出实际的对话样例来展示模型效果。表3提供了若干条测试样例与对应的模型输出。可以看到:(1) Seq2Seq-tgt-attn 模型相比于 Seq2Seq-attn 生成了信息更为丰富且语义一致的回复,所生成的回复中连接词“然而”“但是”的使用更为合适,子句之间也没有逻辑冲突。这说明目标侧注意力机制能够更为细粒度地捕捉回复侧的语义信息并提高句子内部的一致性。(2)相比于基线模型,MatchGAN 模型生成回复的上文相关性更强,信息量也更为丰富。值得一提的是,本文模型所生成的回复往往包含与对应上文内容密切相关的线索词,例如,Baidu-Tieba 测试样例中,球星“林佳德”对应于上文中的球星“德佩”;PERSONA-CHAT 测试样例中,歌星“the Beatles”对应于歌星“Hank Williams”,这反映了有效建模对话上下文关系能够增强谈论话题的可持续性,而这种话题的延展能力对于构建一个用户乐于与之交互的对话模型来说非常重要。这一现象也符合本文的预期,即张量匹配网络使得对话上下文之间的关系建模不仅局限于语义相似,而是能够扩展到更为隐式的上下文关联关系。(3)本文模型所产

在本文所提出的方法中,模型通过显式地建模并利用上文与回复之间多维度的语义关系来增强对话生成模型的上文感知能力。本节探讨对话上文轮数这一因素对本文方法的影响。如图2所示,相比于基线模型,本文模型在 BLEU 和 dist-1 上的表现随着上文轮数的增加而进一步提升,而且这种提升在上文轮数超过3时表现得更为明显。这种现象的部分原因在于,更长的上文使得匹配网络能够更为充分地建模其与回复之间的关系,进而为对话生成模型提供更为准确的反馈信息并提高其回复生成的能力。

生的回复在内部语义一致性和外部上文相关性上均有明显提升。模型 Adver-REGS 和 MMI 相比 Seq2Seq-attn 能够生成更长的回复,然而,这些回复要么与上文的相关性较弱,要么其内部子句存在语义不一致的现象。RL-Cos 模型往往给出与上文语义相近的回复,无法进一步拓展话题。本文模型得益于目标侧注意力机制和对抗式匹配模型的引入,所生成的对话回复中子句之间语义冲突的问题及上文相关性较差的问题得到了缓解。样例展示的回复质量的提升并非个例,本研究在其他测试样例上也观察到了类似的质量提升。综合前文中量化指标评测与案例分析结果,本文发现,融合了目标侧注意力机制和张量匹配网络的对抗式对话生成模型确实能够有效提高对话生成的内部语义一致性和外部上文相关性。

4 结论

本文介绍了面向一致性对话生成的对抗性张量匹配网络与目标侧注意力机制。为了增强对话回复的内部语义一致性,本文引入了目标侧注意力机制,

表 3 在数据集 Baidu-Tieba 和 PERSONA-CHAT 上各模型的生成样例展示

对话上文	个人觉得,上半场德佩是继叔叔之后第二大进攻威胁。
参考回复	我也觉得孟菲斯除了那两脚毒了点,其他的还是可圈可点的。
Seq2Seq-attn	德佩的传球是什么?
Seq2Seq-tgt-attn	这个赛季的表现不错,而且他的位置感不错。
Adver-REGS	我觉得他是个好球员,但是他不是一个人,而是他的位置。
PG-BLEU	德佩的表现不错。
MMI	不好好看球啊。我看球没用啊。我说过他们不用他踢中卫。
RL-Cos	我也觉得德佩不错。
MatchGAN	我觉得林佳德踢得挺好,毕竟上赛季踢得不错,毕竟他是铁腰,虽然很容易受伤。
对话上文	Hank Williams... I grew up on country music.
参考回复	I like country too! I am kinda tired tonight from a long day at the shipyard.
Seq2Seq-attn	I love it, I love it.
Seq2Seq-tgt-attn	I love the country, but I'm not a fan of them.
Adver-REGS	I've never heard of that. I've never heard of that.
PG-BLEU	I'm a UNK. I'm going to see the sunset.
MMI	Well what about that band. . lol ? ? !
RL-Cos	I'm a country boy!
MatchGAN	I love to sing. I love the Beatles.

使得模型在自回归解码的过程中能够关注到已经生成的前文回复。为了提高对话回复的外部上文相关性,本文基于张量网络设计了一种对抗式的匹配模型来显式地建模上下文之间的交互关系并增强回复生成模型的上文感知能力。在 3 个大规模真实对话数据集上的实验表明,相较于其他模型,本文方法有效地提升了对话模型所生成回复的内部语义一致性和外部上文相关性。本文中的对抗式匹配模型并不局限于特定的生成器结构,未来工作中,本研究计划进一步探讨其在其他生成器结构如 Transformer^[28]中的作用。

参考文献

[1] SERBAN I V, SORDONI A, BENGIO Y, et al. Building end-to-end dialogue systems using generative hierarchical neural network models[C] // Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, USA, 2016: 3776-3783

[2] ROLLER S, DINAN E, GOYAL N, et al. Recipes for building an open-domain chatbot[EB/OL]. <https://arxiv.org/pdf/2004.13637v2.pdf>; arXiv, (2020-04-30), [2020-12-15]

[3] CAI H, CHEN H, SONG Y, et al. Exemplar guided neural dialogue generation[C] // Proceedings of the International Joint Conference on Artificial Intelligence, Yokohama, Japan, 2020: 3601-3607

[4] HUANG M, ZHU X, GAO J. Challenges in building intelligent open-domain dialog systems[J]. *ACM Transactions on Information Systems*, 2020, 38(3): 1-32

[5] 陈晨,朱晴晴,严睿,等. 基于深度学习的开放领域对话系统研究综述[J]. *计算机学报*, 2019, 42(7): 1439-1466

[6] 王孟宇,俞鼎耀,严睿,等. 基于 HRED 模型的中文多轮对话任务方法研究[J]. *中文信息学报*, 2020, 34(8): 78-85

[7] SHANG L, LU Z, LI H. Neural responding machine for short-text conversation[C] // Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, 2015: 1577-1586

[8] LI J, MONROE W, RITTER A, et al. Deep reinforcement learning for dialogue generation[C] // Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, USA, 2016: 1192-1202

[9] SORDONI A, GALLEY M, AULI M, et al. A neural network approach to context-sensitive generation of conversational responses[C] // Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Denver, USA, 2015: 196-205

[10] LI J, GALLEY M, BROCKETT C, et al. A diversity-promoting objective function for neural conversation models[C] // Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, USA, 2016: 110-119

[11] ZHOU G, LUO P, CAO R, et al. Mechanism-aware neural machine for dialogue response generation[C] // Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, USA, 2017: 3400-3407

- [12] WANG D, JOJIC N, BROCKETT C, et al. Steering output style and topic in neural response generation[C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 2017: 2140-2150
- [13] FAN A, LEWIS M, DAUPHIN Y. Hierarchical neural story generation[C] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics Melbourne, Australia, 2018: 889-898
- [14] ZHANG H, LAN Y, GUO J, et al. Reinforcing coherence for sequence to sequence model in dialogue generation[C] // Proceedings of the International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 2018: 4567-4573
- [15] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C] // Advances in neural information processing systems, Montreal, Canada, 2014: 2672-2680
- [16] LI J, MONROE W, SHI T, et al. Adversarial learning for neural dialogue generation[C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 2017: 2157-2169
- [17] XU Z, LIU B, WANG B, et al. Neural response generation via GAN with an approximate embedding layer[C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 2017: 617-626
- [18] RITTER A, CHERRY C, DOLAN W B. Data-driven response generation in social media[C] // Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, Edinburgh, UK, 2011: 583-593
- [19] VINYALS O, LE Q. A neural conversational model[EB/OL]. <https://arxiv.org/pdf/1506.05869v3.pdf>; arXiv, (2015-07-22), [2020-12-15]
- [20] CHEN H, REN Z, TANG J, et al. Hierarchical variational memory network for dialogue generation[C] // Proceedings of the 2018 World Wide Web Conference, Lyon, France, 2018: 1653-1662
- [21] LI X, MOU L, YAN R, et al. StalemateBreaker: a proactive content-introducing approach to automatic human-computer conversation[C] // Proceedings of the 25th International Joint Conference on Artificial Intelligence, New York, USA, 2016: 2845-2851
- [22] YAO L, ZHANG Y, FENG Y, et al. Towards implicit content-introducing for generative short-text conversation systems[C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 2017: 2190-2199
- [23] ZHAO T, ZHAO R, ESKENAZI M. Learning discourse-level diversity for neural dialog models using conditional variational autoencoders[C] // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, Canada, 2017: 654-664
- [24] MEI H, BANSAL M, WALTER M R. Coherent dialogue with attention-based language models[C] // Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, USA, 2017: 3252-3258
- [25] XIA Y, TIAN F, QIN T, et al. Sequence generation with target attention[C] // Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Skopje, Macedonia, 2017: 816-831
- [26] SHAO Y, GOUWS S, BRITZ D, et al. Generating high-quality and informative conversation responses with sequence-to-sequence models[C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 2017: 2210-2219
- [27] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780
- [28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 5998-6008
- [29] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[C] // Proceedings of the 3rd International Conference on Learning Representations, San Diego, USA, 2015: 1-15
- [30] SOCHER R, CHEN D, MANNING C D, et al. Reasoning with neural tensor networks for knowledge base completion[C] // Advances in Neural Information Processing Systems, Lake Tahoe, USA, 2013: 926-934
- [31] WILLIAMS R J. Simple statistical gradient-following algorithms for connectionist reinforcement learning[J]. *Machine Learning*, 1992, 8(3-4): 229-256
- [32] RUBINSTEIN R Y, KROESE D P. Simulation and the Monte Carlo Method[M]. Hoboken: John Wiley & Sons, 2016
- [33] LIN K, LI D, HE X, et al. Adversarial ranking for language generation[C] // Advances in Neural Information Processing Systems, Long Beach, USA, 2017: 3155-3165
- [34] GUO J, LU S, CAI H, et al. Long text generation via adversarial training with leaked information[C] // Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, USA, 2018: 5141-5148
- [35] ZHANG S, DINAN E, URBANEK J, et al. Personalizing

- Dialogue Agents: I have a dog, do you have pets too? [C] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, 2018: 2204-2213
- [36] BAHDANAU D, BRAKEL P, XU K, et al. An actor-critic algorithm for sequence prediction [EB/OL]. <https://arXiv.org/pdf/1607.07086.pdf>; arXiv, (2017-03-03), [2020-12-15]
- [37] PAPINENI K, ROUKOS S, WARD T, et al. BLEU: a method for automatic evaluation of machine translation [C] // Proceedings of the 40th annual meeting of the Association for Computational Linguistics, Philadelphia, USA, 2002: 311-318
- [38] MILLER A, FENG W, BATRA D, et al. ParlAI: a dialog research software platform [C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Copenhagen, Denmark, 2017: 79-84
- [39] HU B, CHEN Q, ZHU F. LCSTS: a large scale Chinese short text summarization dataset [C] // Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 2015: 1967-1972
- [40] PENNINGTON J, SOCHER R, MANNING C D. Glove: global vectors for word representation [C] // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 2014: 1532-1543
- [41] KINGMA D P, BA J. Adam: a method for stochastic optimization [EB/OL]. <https://arXiv.org/pdf/1412.6980.pdf>; arXiv, (2017-01-30), [2020-12-15]
- [42] BROWN P F, DELLA PIETRA S A, DELLA PIETRA V J, et al. An estimate of an upper bound for the entropy of English [J]. *Computational Linguistics*, 1992, 18(1): 31-40
- [43] LIU C W, LOWE R, SERBAN I V, et al. How NOT to evaluate your dialogue system: an empirical study of unsupervised evaluation metrics for dialogue response generation [C] // Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, USA, 2016: 2122-2132
- [44] MCHUGH M L. Interrater reliability: the kappa statistic [J]. *Biochemia Medica*, 2012, 22(3): 276-282

Research on towards coherent dialogue generation via adversarial matching network and target-side attention

CAI Hengyi^{***}, WANG Chengrui^{***}, SONG Yonghao^{*}, YUAN Xu^{***}, ZHANG Cheng^{*}, ZHAO Xiaofang^{*}

(^{*} Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

(^{**} University of Chinese Academy of Sciences, Beijing 100049)

Abstract

The sequence-to-sequence (seq2seq) approach has received great attention in the field of open-domain dialogue generation. However, the standard seq2seq model is prone to generate meaningless and incoherent responses, making it distinguish clearly from the human-human conversations. This coherence includes both the internal consistency of the long response and the external relevance between the post and its response. This work proposes a novel dialogue generation framework called matching-based generative adversarial network (MatchGAN) to improve external relevance. Instead of imitating the ground truth with supervised learning, this model can generate post-relevant responses through the generative-adversarial learning with a seq2seq-based generator and a matching-based discriminator, allowing the generated human-machine conversation more like a human-human conversation by discriminating whether a response is matching with the post. Furthermore the target-side attention mechanism is introduced to maintain the internal consistency of the generated responses. This new framework is able to generate coherent responses with high quality. Experimental results show that the proposed model can achieve substantial improvements in both metric-based and human evaluations among various baselines.

Key words: neural dialogue generation, neural tensor layer, adversarial learning, target-side attention mechanism