

基于多奖励强化学习的半监督文本风格迁移方法

李静文, 叶琪, 阮彤, 林宇翩, 薛万东

引用本文

李静文, 叶琪, 阮彤, 林宇翩, 薛万东. 基于多奖励强化学习的半监督文本风格迁移方法[J]. 计算机科学, 2024, 51(8): 263-271.

LI Jingwen, YE Qi, RUAN Tong, LIN Yupian, XUE Wandong. [Semi-supervised Text Style Transfer Method Based on Multi-reward Reinforcement Learning](#) [J]. Computer Science, 2024, 51(8): 263-271.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多样化标签矩阵的医学影像报告生成](#)

Diversified Label Matrix Based Medical Image Report Generation

计算机科学, 2024, 51(8): 200-208. <https://doi.org/10.11896/jsjcx.230600018>

[基于深度确定性策略梯度与注意力Critic的多智能体协同清障算法](#)

Multi-agent Cooperative Algorithm for Obstacle Clearance Based on Deep Deterministic Policy Gradient and Attention Critic

计算机科学, 2024, 51(7): 319-326. <https://doi.org/10.11896/jsjcx.230600129>

[计及风电的发电商报价多智能体模型](#)

Multi-agent Based Bidding Strategy Model Considering Wind Power

计算机科学, 2024, 51(6A): 230600179-8. <https://doi.org/10.11896/jsjcx.230600179>

[基于深度强化学习的二进制代码模糊测试方法](#)

Fuzz Testing Method of Binary Code Based on Deep Reinforcement Learning

计算机科学, 2024, 51(6A): 230800078-7. <https://doi.org/10.11896/jsjcx.230800078>

[基于深度强化学习的数据中心热感知能耗优化方法](#)

Deep Reinforcement Learning Based Thermal Awareness Energy Consumption Optimization Method for Data Centers

计算机科学, 2024, 51(6A): 230500109-8. <https://doi.org/10.11896/jsjcx.230500109>

基于多奖励强化学习的半监督文本风格迁移方法

李静文 叶琪 阮彤 林宇翱 薛万东

华东理工大学计算机科学与工程学院 上海 200237

(y30211022@mail.ecust.edu.cn)

摘要 文本风格迁移是自然语言处理中的一项重要任务,其主要目的在于改变文本的风格属性,同时保留必要的语义信息。然而,在许多任务缺乏大规模平行语料库的情况下,现有的无监督方法存在文本多样性不足和语义一致性较差的问题。针对这些问题,文中提出了一种半监督的多阶段训练框架。该框架首先利用风格标注模型和掩码语言模型构造伪平行语料库,以有监督的方式引导模型学习多样性的迁移方式。其次,设计了对抗性相似奖励、Mis奖励和风格奖励,从未标记的数据中进行强化学习以增强模型的语义一致性、逻辑一致性和风格准确性。在基于 YELP 数据集的情感极性转换任务中,该方法的 BLEURT 分数提升了 3.1%,Mis 分数提升了 2.5%,BLEU 分数提升了 9.5%;在基于 GYAFC 数据集的正式文体转换实验中,该方法的 BLEURT 分数提高了 6.2%,BLEU 分数提高了 3%。

关键词: 文本生成;文本风格迁移;多阶段训练;风格标注模型;强化学习

中图分类号 TP391

Semi-supervised Text Style Transfer Method Based on Multi-reward Reinforcement Learning

LI Jingwen, YE Qi, RUAN Tong, LIN Yupian and XUE Wandong

School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China

Abstract Text style transfer is an important task in natural language processing that aims to change the stylistic attributes of text while preserving necessary semantic information. However, in many tasks where large-scale parallel corpora are lacking, existing unsupervised methods suffer from issues such as insufficient text diversity and poor semantic consistency. To address these problems, this paper proposes a semi-supervised multi-stage training framework. It first constructs a pseudo-parallel corpus using a style labeling model and a masked language model to guide the model to learn diverse transfer styles in a supervised manner. Then, adversarial similarity reward, Mis reward, and style reward are designed to conduct reinforcement learning on unlabeled data to enhance the model's semantic consistency, logical consistency, and accuracy of style transfer. In the sentiment polarity conversion task based on the YELP dataset, the proposed method's BLEURT score increases by 3.1%, the Mis score increases by 2.5%, and the BLEU score increases by 9.5%. In the formal style conversion experiment based on the GYAFC dataset, its BLEURT score increases by 6.2%, and the BLEU score increases by 3%.

Keywords Text generation, Text style transfer, Multi-stage training, Style labeling model, Reinforcement learning

1 引言

文本风格迁移在自然语言生成中具有重要作用,可用于转换口头情绪、调整文本的正式程度以及消除文本中的偏见^[1]。以情感极性转换任务为例,该技术可应用于情感调节训练或品牌形象创作。执行此类任务需要将消极句子转化为积极句子,同时确保必要的独立于风格的语义信息不会丢失。因此,一个成功的风格迁移模型必须同时满足以下 3 个条件^[2]:1)生成的文本属于目标风格;2)转换前后的文本具有较高的语义保留度;3)生成的文本流畅自然且具有多样性。

由于许多风格迁移任务缺乏平行数据,目前的研究通常采用无监督训练方式^[3]。其中一种主流方法是基于风格词替换的显示解耦,需要识别原始句子中的风格词并将其替换为与目标风格相关的内容^[4]。该方法非常直观且可解释性较强,但不适用于某些风格隐藏在句法结构中的任务。第二种方法则隐式进行文本和风格的解耦,模型需要学习独立于风格的文本表示以及独立于内容的风格表示^[5-6]。该方法显著提高了生成文本的质量,但在解耦过程中可能会丢失一部分语义信息,从而降低输出文本内容的保真度。第三种方法不进行文本和风格的解耦,直接学习原始风格和目标风格之间

到稿日期:2023-06-22 返修日期:2023-11-20

基金项目:上海市促进产业高质量发展专项资金(2021-GZL-RGZN-01018);国家重点研发计划(2021YFC2701800,2021YFC2701801)

This work was supported by the Shanghai Municipal Special Fund for Promoting High-quality Development of Industries(2021-GZL-RGZN-01018) and National Key Research and Development Program of China(2021YFC2701800,2021YFC2701801).

通信作者:阮彤(ruantong@ecust.edu.cn)

的映射函数。这种方法可以更好地保持原始输入文本内容的信息,同时还能将目标风格应用到生成的输出文本中。因此,生成的结果更加自然、准确^[7-9]。使用强化学习不进行文本和风格解耦已被证明是一种有效的方法^[10-11],可通过设定文本奖励和风格奖励来优化句子生成过程。

在使用强化学习实现文本风格迁移时,通常需要先构建伪平行语料库进行预热。现有模型通常采用基于频率比或注意力分数的方法来筛选风格词以构建伪平行语料,但这些方法筛选出的风格词准确性较差。基于频率比的方法仅根据统计信息筛选高频风格词,无法考虑实际上下文语境,从而存在遗漏单词的问题^[12]。而对于基于注意力分数的方法,在文本中单词的注意力分布变化不显著时,其阈值难以设定,可能会导致误删^[13]。此外,获取显示去除风格词的句子后,还需要进一步生成目标风格句子。现有的工作一般采用基于检索或生成的方式进行风格词替换,但这种方法无法保证生成文本的多样性和语义一致性,需要额外花费时间和成本进行人工调整^[14]。

强化学习的核心是奖励设置。在设置奖励时,需要考虑多种类型的奖励,例如风格奖励和文本奖励,并对每种奖励类型进行明确定义。然而,在缺乏平行数据的情况下,由于风格迁移前后风格词的差异,模型无法直接计算迁移前后文本的相似度。尤其在情感极性转换任务中,这种差异使得计算文本奖励变得更加困难和具有挑战性。现有的模型通过循环重构的方式计算文本奖励^[10],但这会极大地增加模型优化过程的计算量,导致模型收敛速度变慢,影响模型的稳定性和泛化能力。此外,现有用于计算文本奖励的指标(如 BLUE^[15] 和 ROUGE^[16]等),都是基于词重叠或者词向量的角度计算两个文本之间的语义相似度。然而,这些指标并不能准确地捕捉生成文本与原文本之间的逻辑一致性和语义关系。

为了解决上述问题,本文提出了基于多奖励强化学习的半监督文本风格迁移方法(Semi-Supervised Text Style Transfer Method based on Multi-Reward Reinforcement Learning, MRRL),该方法构造了高质量的伪平行数据集,并优化了无监督训练过程中奖励函数的设置。为了提高风格词识别的准确率,MRRL方法首先通过风格标注模型以及风格熵的协同配合,识别出文本中的情感词并进行掩码处理。其次,使用基于 BART^[17]的风格掩码语言模型预测替代的风格词,构造出富有多样性的文本。最后,利用风格分类器和 BLEURT^[18]文本指标共同过滤出高质量的平行数据。为了优化强化学习模型在无监督的情况下计算文本奖励的过程,本文提出了对抗相似度奖励(Adversarial Similarity Reward, ASR),能够对两种不同风格的句子直接进行文本比较。此外,本文还提出使用 MIS^[19]文本指标作为新的奖励函数,可以更加准确地捕捉生成文本和原文本之间的双向相似性和语义关联。

本文使用了 MRRL 方法,在两个流行任务上进行了实验,分别是情感极性转换任务和正式文体转换任务。本文的主要贡献包括以下 3 点:

1)充分利用有监督和无监督学习范式,提出了一种半监督的多阶段训练框架;

2)将风格标注模型和风格掩码语言模型相结合,生成高质量的伪平行数据集;

3)设计了对抗性相似奖励和 MIS 文本奖励,提高了风格迁移模型在语义保留程度和逻辑一致性方面的能力。

2 相关工作

文本风格迁移的目的是转变文本的风格属性,并且需要保留独立于风格的必要语义信息。针对这个目标,研究者提出了多种方法和思路。总体而言,可以将主流的研究思路归纳为以下 3 类:

1)基于风格词替换的显示解耦,需要在原语句中识别风格词并将其替换成内容相关的目标风格词语。Tokpo 等^[20]通过风格分类器为每个单词分配权重,设定阈值识别风格词,并将其与潜在内容表示相结合来生成目标文本。Reid 等^[21]将注意力分数高于平均值的短语作为情感词并将其删除,使用掩码语言模型预测目标风格词。Wang 等^[14]通过掩码模型和风格分类器的对抗显示地去除风格词,再利用 BERT 模型来预测风格词。这种方法非常直观且可解释性强,但是不适用于风格隐藏在句法结构中的任务,例如正式文体转换任务。

2)隐式进行文本跟风格的解耦,模型需要学习给定序列独立于风格的文本表示以及独立于内容的风格表示。例如, Kashyap 等^[6]使用风格分类器和编码器之间的对抗学习,隐式去除文本中的风格,并使用特定风格的解码器生成相对应风格的句子。另外, Lee 等^[13]使用反向注意力机制隐式去除风格,并且利用条件层规范化来构建依赖于内容的风格表示,随后输入解码器生成目标文本。Li 等^[22]将条件先验引入潜在的内容空间和风格空间中进行文本跟风格的解耦,再将独立于风格的文本表示与目标风格表示相结合,指导解码器生成具有目标风格的句子。该方法显著提高了生成文本的质量,但是解耦过程中可能会损失一部分的语义信息,这可能会导致在保持内容和风格均衡的效果时存在一定难度。

3)不进行文本跟风格的解耦,直接学习从原风格到目标风格的映射函数。例如, Yi 等^[23]通过流模型在风格空间中构造了复杂的风格分布,并将其与原文本的潜在表示一起输入解码器。Nouri^[24]使用了风格损失和循环重构损失,同时引入了最优传输距离,以保证文本之间语法和语义信息的一致性。Deng 等^[25]采用了文本奖励和风格奖励来指导模型进行文本风格迁移,并且使用策略梯度算法进行优化。Liu 等^[26]使用半监督学习的方法,在强化学习训练之前,利用自动构造的伪平行数据对模型进行预热,以提高模型训练的稳定性。这类方法可以更好地保留原始输入文本的内容信息,并将风格应用到生成的输出文本中,从而获得更加准确和自然的结果。

在这项工作中,本文采用不进行文本和风格解耦的方式,并基于强化学习模型设计了一种半监督的多阶段训练框架。为了提高模型的收敛速度,选用基于风格词替换的显示解耦方法构造伪平行语料库,对模型进行预热。接着,使用多个奖励项从各个方面指导模型进行文本风格迁移。

3 MRRL 方法

假设 S_x 代表原风格标签, S_y 代表目标风格标签。例如,

在情感极性转换任务中, $S_X = negative$, $S_Y = positive$ 。给定长度为 n 的 S_X 风格数据集 $D_X = \{x_1, \dots, x_n\}$, 文本风格迁移的目的就是生成风格为 S_Y 的对应数据 $D_Y = \{y_1, \dots, y_n\}$ 。其中, x_i 代表一系列 S_X 风格的句子, y_i 则代表与 x_i 文本信息相同的一系列 S_Y 风格的句子。本文一共定义了两个 BART 模型^[17], 分别用于实现两个方向的迁移任务。Model A 的目标是实现风格 S_X 到风格 S_Y 之间的迁移任务, Model B 的目标是实现风格 S_Y 到风格 S_X 之间的迁移任务。形式上, 该任务的目标是最小化以下负对数似然函数:

$$L(\phi) = -\sum_i \log(p(y_i | y_{1:i-1}, x; \phi)) \quad (1)$$

其中, ϕ 代表风格迁移模型的参数。

3.1 框架概述

本文提出的 MRRL 方法采用了基于强化学习框架的多阶段训练模式, 如图 1 所示。首先, 使用风格标注模型和掩码语言模型自动构建伪平行语料库, 并对预训练语言模型进行预热。其次, 通过迭代反向翻译的方法进行无监督文本风格迁移训练, 包括使用风格分类器、对抗编码器以及循环重构模式, 分别计算生成文本的风格奖励、对抗相似度奖励以及 MIS 文本奖励, 用于指导模型进行强化学习训练。最后, 使用第二阶段模型生成的一系列高质量数据, 在新的预训练语言模型上进行最后的有监督训练。

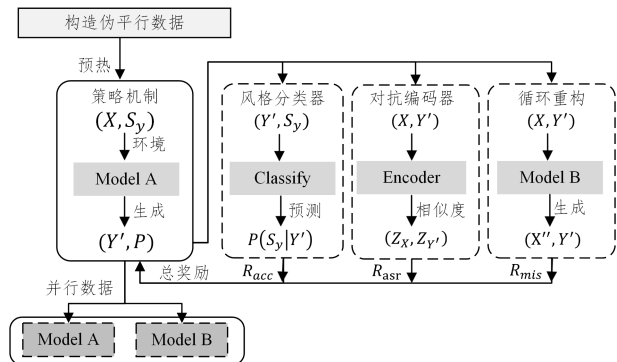


图 1 基于强化学习框架的多阶段训练框架图

Fig. 1 Multi-stage training framework based on reinforcement learning

3.2 伪平行数据集的构造

为了提高强化学习训练过程的学习速度和稳定性, MRRL 采用显示风格词替换的方法构造伪平行语料库。具体的构造过程包括 3 个步骤, 如图 2 所示。第一步, 使用风格标注模型来预测掩码向量, 并且通过计算风格熵进一步筛选出风格词; 第二步, 在特定风格语料库中微调掩码语言模型, 并用其预测第一步中掩码处的短语, 从而构造伪平行语料库; 第三步, 利用分类器以及 BLEURT 指标的协同过滤筛选出高质量的平行数据, 用于后续的强化学习。

在第一个步骤中, 风格标注模型的目的是预测输入文本中具有风格特征的标注词位置。给定一个 S_X 风格的原文本 X , 风格标注模型会生成一个掩码向量 $v = [v_1, v_2, \dots, v_n]$, 其中 $v_i \in \{0, 1\}$ 代表 X 中第 i 个文本是否是风格标注词。 $v_i = 0$ 代表第 i 个单词与风格属性的相关性较大, 在进行风格转换过程中需要被掩码。 $v_i = 1$ 代表第 i 个单词跟文本主题的相关性较大, 因此不是风格词, 应该被保留。随后再使用特殊标记 [mask] 代替风格标注词, 并计算掩码序列 \tilde{X} 的风格熵。

如果风格熵超过阈值, 则说明已成功去除风格词; 否则, 依次掩码文本中的剩余单词, 直到风格熵超过阈值, 并更新掩码向量 v 。为了确保文本的内容质量, 本文设定了文本中单词的最大掩码率 α 。例如, 如果文本 X 的内容是: “The food was sweet and tasty”, 风格标注模型以及风格熵所确定的掩码向量 $v = [1, 1, 1, 0, 1, 0]$, 则代表单词 “sweet” 和 “tasty” 是风格词, 使用 [mask] 替换, 最终掩码序列 \tilde{X} 可以表示为 “The food was [mask] and [mask]”。

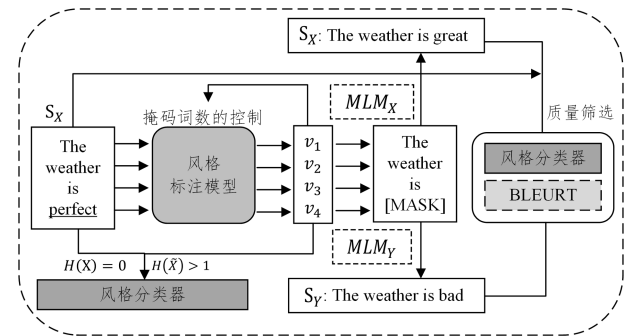


图 2 伪平行数据构造过程图

Fig. 2 Diagram of pseudo-parallel data construction process

在训练风格标注模型之前, MRRL 首先利用外部统计信息预先构建了一个风格词典^[27], 该库包含了所有的风格属性词。随后, 将原始文本 X 转换为 token 序列, 并将其映射为数字表示。在此过程中, 需要查询风格词典以确定每个 token 是否属于某种特定的风格属性。若是, 则标注为 “0”, 否则标注为 “1”。接着, 使用预训练语言模型学习输入句子的上下文信息, 并提取每个 token 的特征向量。再将这些特征向量与风格标注序列输入条件随机场 (CRF)^[28] 中, 使其根据已知的标注序列, 分析不同文本序列中风格标记的统计规律, 从而预测出未知文本的风格标记序列。

为了进一步保证风格词识别的准确率, 本文引入了风格熵的概念, 代表一段文本中含有的风格属性信息量。在计算风格熵时, 通过分类器预测的结果和置信度得到各个风格类别的概率分布, 并对各个类别的熵值进行加权平均, 从而得到文本的总体风格熵。风格熵的计算式如下:

$$H(x) = -\sum_{i=0}^k p_i \log_2 p_i \quad (2)$$

其中, k 代表风格标签的个数, p_i 代表当前文本被预测为风格 i 的概率值。具体来说, 假设有两个不同的风格 S_X 和 S_Y , 当文本中无明显风格词, 即被预测为 S_X 风格或者 S_Y 风格的概率都很小时, 此时风格熵取得最大值。

对于第二个步骤, 使用风格掩码语言模型进行掩码词的预测, 首先需要在特定属性的数据集上微调预训练语言模型。以 MLM_X 为例, 需要随机丢弃 S_X 风格语料库中的短语, 要求模型 MLM_X 预测最有可能出现的短语来填充原句, 并且最小化预测结果和真实结果之间的误差。通过这种训练方式, 对于每一个掩码句 \tilde{X} , 分别输入两个不同的风格掩码语言模型, 即可生成两个新的分别带有特定属性 S_X 和 S_Y 的文本 \hat{X} 和 \hat{Y} 。

$$\hat{X} = MLM_X(\tilde{X}; \theta_X) \quad (3)$$

$$\hat{Y} = MLM_Y(\tilde{X}; \theta_Y) \quad (4)$$

其中, θ_X 和 θ_Y 分别代表两种风格掩码语言模型的参数。

根据上述伪平行数据集的构造方法,对于给定非平行的语料库 X 和 Y ,通过风格标注模型识别风格词并进行掩码,再将其放入 MLM_X 和 MLM_Y 中进行预测,可以生成两种风格的预测句,如表 1 所列。

表 1 伪平行数据对构造实例

消极原句	absolutely the worst care in all my experience with vets
掩码句	absolutely the <mask> in all my experience with vets
积极预测句	absolutely the best care i have ever had in all my experience with vets
消极预测句	absolutely the worst customer service i have had in all my experience with vets
积极原句	not to mention they area mazing at what they do
掩码句	not to mention they are <mask> at what they do
积极预测句	not to mention they are very good at what they do
消极预测句	not to mention they are very unprofessional at what they do

在第三个步骤中,需要对构造的伪平行数据集做风格质量和文本质量的筛选。筛选计算式如下:

$$D_n = \{ (x, y) \mid BLEURT(x, y) > \sigma_c \text{ and } \frac{p(s_x|x) + p(s_y|y)}{2} > \sigma_s \} \quad (5)$$

其中, σ_c 代表伪平行数据对之间 BLEURT 分数的筛选阈值, $p(s_i | *)$ 代表文本被预测为风格 s_i 的概率值, σ_s 代表伪平行数据对中 x 和 y 分别被预测为 s_x 风格和 s_y 风格的概率平均数的阈值。

按照以上方法, MRRL 方法构造出两种伪平行语料库。第一种伪平行语料库是由原始风格数据以及生成的目标风格预测数据组成。第二种伪平行语料库充分利用了原风格预测数据,增加了由原风格预测数据以及目标风格预测数据组成的平行数据对,有 3/4 的数据是由模型生成的。

3.3 基于强化学习的无监督训练

在使用伪平行数据集预热之后, MRRL 通过强化学习无监督地迭代训练两个迁移方向的语言模型,如图 3 所示。在此过程中, MRRL 使用对抗性相似奖励 (ASR), 确保一对文本中关于风格无关的内容信息之间的一致性。同时, 还使用 MIS 文本奖励来保证迁移前后文本之间的内容信息蕴含关系。此外, 还引入了风格奖励 (ACC) 来促使模型提高风格转化率。

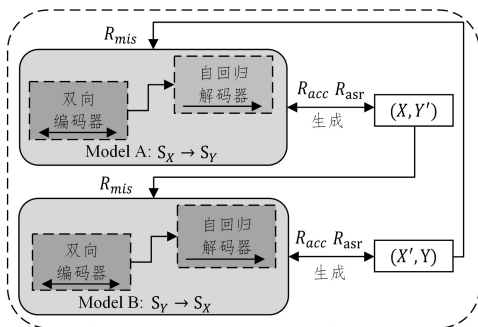


图 3 基于多奖励机制的迭代反向翻译训练框架

Fig. 3 Multi-reward mechanism-based iterative back translation training framework

1) 对抗性相似奖励

文本和风格往往都是相互纠缠、密不可分,风格属性改变的同时也会引起文本内容属性的变化^[29]。而在风格迁移领域平行数据十分稀缺,现有方法都很难衡量原风格文本和生成的目标风格文本之间的语义一致性。因此,本文提出了对抗性相似奖励,可以直接计算两种不同风格文本之间的语义相关度。

使用对抗性相似奖励需要优先训练一个对抗编码器,该编码器可以将句子编码到潜在空间中,从而隐式进行文本和风格的解耦,在潜在空间中去除风格信息。Zhao 等^[30]提出的 ARAE 框架是生成式对抗网络 (GAN)^[31] 自编码器的一种变体。该框架主要通过施加隐含的先验信息来约束潜在空间以获得平滑的数据分布,并简化从中采样的过程。本文基于此模型提出了对抗性相似奖励。首先使用编码器和风格鉴别器的对抗,训练编码器模型的参数,使得它编码之后的潜在表示风格熵最大,即无法被鉴别器预测出句子的风格。与此同时,通过训练一个生成器模型来生成虚假的潜在表示,以欺骗文本鉴别器并使其无法准确区分潜在表示的来源。随着生成器和文本鉴别器的不断对抗学习,文本鉴别器可以逐渐提高辨别真假数据的能力,从而确保编码器在隐式去除文本风格的同时不引入噪声数据。

$$L_{adv1}(\theta_c) = -\sum_{i=0}^k H(p(i | Encoder(x; \theta_c); \theta_c)) \quad (6)$$

$$L_{adv2}(\psi) = \mathbb{E}_{z \sim P_z} [crc_\xi(z)] - \mathbb{E}_{z \sim P_z} [crc_\xi(\bar{z})] \quad (7)$$

$$L_{crc}(\xi) = -\mathbb{E}_{z \sim P_z} [crc_\xi(z)] + \mathbb{E}_{z \sim P_z} [crc_\xi(\bar{z})] \quad (8)$$

其中, θ_c 代表对抗编码器的模型参数, θ_c 代表风格鉴别器的模型参数, crc_ξ 代表鉴别器模型, z 和 \bar{z} 分别代表真实的潜在表示和虚假的潜在表示。

在进行文本风格迁移的过程中, MRRL 将原数据 X 和生成的目标风格数据 Y' 同时输入训练好的对抗编码器中,得到两个隐式去除风格之后的潜在表示 Z_X 和 $Z_{Y'}$,接着计算这两个向量之间的余弦相似度,即为对抗性相似奖励,计算式如下:

$$R_{csr} = \lambda_{csr} [SIM(Z_X, Z_{Y'})] \quad (9)$$

2) MIS 文本奖励

MIS (Mutual Implication Score)^[19] 是一种基于语言学上的逻辑和语义关系来判断相似度的文本指标,它通过两个方向上的 NLI 模型计算不同文本之间的互蕴含分数。对于具有相同含义的两个句子,它们在推理性质上应该是等价的,无论输入模型的顺序如何,都会返回相似的蕴含分数。相比其他基于词重叠或者基于词向量的文本质量评估方法, MIS 指标更加关注两段文本中词汇或者语义之间的关联程度,在处理复杂语义场景下的文本度量效率更高。而在文本风格迁移任务中,数据集中大多包含否定句或者条件语句这种复杂的语义结构,在迁移过程中语句中的词汇和句法结构通常会发生变化。因此,使用 MIS 指标可以更准确地捕捉句子之间的逻辑和语义关系,更高效地判断转换前后句子语义之间的关联程度。

如图 3 所示,以风格 s_x 到风格 s_y 的迁移过程为例,将模型 A 生成的句子 Y' 重新输入模型 B 中,即可得到循环重构文本

X'' 。此时的 X'' 与原文本 X 同属于 s_x 风格,可以直接进行文本信息的对比,得到一个循环重构文本奖励^[32]。MIS指标中的正向隐含概率将用来衡量 X'' 中的所有信息都保存在 X 中的概率,反向隐含概率用来衡量 X 中的所有信息都保存在 X'' 中的概率。当两个方向的隐含概率相似且均较高时,即代表这一对文本之间的内容一致性很高。该奖励机制有助于确保模型生成的文本准确表达原始文本的语义,成功完成风格转换任务。MIS文本奖励的计算式如下:

$$R_{\text{mis}} = \lambda_{\text{mis}} [\text{MIS}(X'', X)] \quad (10)$$

3) 风格奖励

为了给模型提供一个明确的风格信号,MRRL方法中使用一个二进制风格分类器来计算风格奖励函数。风格分类器是用带有风格标签的语料库进行训练,可以评估生成的文本 y' 符合目标风格的概率有多大。每个迁移方向的风格标签预测概率的计算式如下:

$$p(s_i | y') = \text{softmax}(\text{TextCNN}(y', \theta)) \quad (11)$$

其中, s_i 代表可能的风格标签, θ 代表风格分类器的参数,在进行风格迁移的过程中,参数固定不变。风格奖励的定义如下:

$$R_{\text{acc}} = \lambda_{\text{acc}} [p(s_y | y') - p(s_x | y')] \quad (12)$$

其中, s_x 和 s_y 分别代表原语句风格标签和目标风格标签, y' 代表从模型输出分布中采样出来的目标语句。MRRL鼓励模型生成的目标句被预测为 s_y 风格的概率变大。

4) 梯度和目标

MRRL方法使用策略梯度算法^[33]最大化生成句子的期望奖励,风格迁移模型参数 ϕ 的梯度计算抽样方式估计如下:

$$\nabla_{\phi} J(\phi) = E[R \cdot \nabla_{\phi} \log(P(y' | x; \phi))] \quad (13)$$

其中, $\nabla_{\phi} J(\phi)$ 代表目标函数 $J(\phi)$ 关于模型参数 ϕ 的梯度, $E(\ast)$ 代表期望, R 代表针对每一个解码时间步从模型的输出分布中抽样出的句子 y' 计算出的奖励值。整个无监督训练过程是通过结合基础模型损失和策略梯度奖励来实现的。

3.4 并行数据生成与模型微调

在模型训练的最后一个阶段,MRRL方法使用第二阶段中表现最优秀的模型生成数据对,以构造静态平行数据。具体来说,将原先训练集中的非平行数据作为输入,要求模型生成相应的目标风格文本,并组成数据对。然后,使用式(5)介绍的条件筛选的方法,利用BLEURT文本分数和风格分类器输出的风格分数对数据对进行过滤,筛选得到高质量的数据对。

最终,使用这些高质量的平行数据,采用3.3节中所有的奖励策略微调训练语言模型,以有监督的方式进行文本风格迁移训练。通过最后一个阶段的训练,可以充分利用有监督训练的优势,提升模型的稳定性和准确率。

4 实验

4.1 实验数据集

MRRL方法在Yelp数据集^[4]上完成了情感极性转换任务,在GYAFC数据集^[34]上完成了正式文体转换任务。Yelp数据集中包括餐馆和家庭服务在内的企业的评论数据集,由积极和消极两种情绪的语句组成。Yelp数据集中无平行语料,在测试时使用人工编写的参考语句进行比较。GYAFC

[F&R]数据集包括家庭与关系领域的的数据,由对齐的正式文体和非正式文体句子组成。两个任务的训练集、验证集和测试集遵从文献[10,24]中的划分方式,任务类型和统计数据如表中的2-5行所列。由于情感极性转换任务缺少平行数据集,利用3.1节中的方法构造出两种伪平行语料库,统计信息如表2中的第6-9行所列。

表2 数据集统计

Table 2 Dataset statistics

数据集	属性	训练集	验证集	测试集
YELP (unpair)	消极	177 218	2 000	500
	积极	266 041	2 000	500
GYAFC (pair)	非正式	51 967	2 788	1 332
	正式	51 967	2 247	1 019
Pseudo YELP (5 万 Pairs)	消极	50 000	1 691	500
	积极	50 000	1 691	500
Pseudo YELP (10 万 Pairs)	消极	105 028	2 000	500
	积极	105 028	2 000	500

4.2 实验参数

MRRL方法在PyTorch框架下使用Huggingface Transformers^[35]中的BART base模型进行了风格掩码语言模型和风格迁移生成器的实验,并使用TextCNN^[36]模型训练风格分类器。在训练过程中,使用Adam优化器^[37],将初始化学学习率设定为 10^{-5} 、批次大小设定为64进行训练。构造伪平行数据集时,风格熵的阈值为1,文本中单词的最大掩码率 $\alpha = 0.5$ 。对抗性相似奖励、MIS文本奖励、风格准确性奖励的权重都设定为1。

在筛选平行数据时,风格阈值和文本阈值具有至关重要的作用。风格分数阈值 σ_c 均设定为0.95,而文本阈值的最佳设定取决于具体的任务类型。因此,在相同的模型参数下,本文选择了多个不同大小的文本阈值进行测试,以研究其在不同任务中的影响差异,并确定最佳的文本阈值。如图4所示,以针对情感极性转换任务的YELP数据集为例,当 σ_c 设为0时,模型达到最大的HM值;而针对正式文体转换任务的GYAFC数据集,当 σ_c 取0.15时,模型才能够实现最佳的HM值。基于以上结果,本文将情感极性转换任务中的文本阈值 σ_c 设定为0,而将正式文体转换任务中的文本阈值 σ_c 设定为0.15。

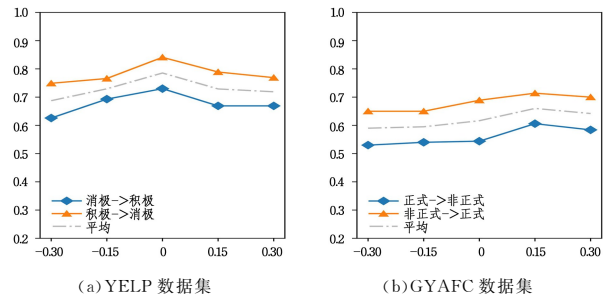


图4 文本阈值 σ_c 对模型性能的影响

Fig. 4 Influence of text threshold σ_c on model performance

4.3 评估指标

为了评估生成文本的效果,本文考虑从文本相似度和风格准确性两个角度出发,采用通用度量标准进行综合评估。

对于文本相似度,本文使用了基于词向量的可学习指标 BLEURT、基于文本蕴含的指标 MIS 和基于词重叠的指标 BLEU,来分别衡量生成文本与人工注释文本之间的文本相似性。

对于风格准确性,本文使用训练过的风格分类器来自动评估生成的迁移文本关于目标风格的准确性。风格分类器在 YELP 和 GYAFC 测试集上的准确率分别为 98% 和 96%。为了与前人的工作进行比较,本文又计算了风格准确率和 BLEU 分数之间的调和平均值 HM,用于综合评价生成文本的质量。

4.4 结果分析

4.4.1 对比实验

在情感极性转换任务中,本文选择具有最好综合效果的

两个模型 M3.1 和 M3.3,与模型 DP-Latent^[38],CP-VAE^[39],NAST^[40],GenRes^[11] 以及 RL-prompt^[25] 进行对比,如表 3 (左)所列。模型 M3.1 在文本保留方面达到了最优,BLEURT 分数、Mis 分数以及 BLEU 分数分别领先于最优模型 3.1%,2.5%,9.5%;模型 M3.3 在风格方面领先于最优模型 1.7%。在正式文体转换任务中,本文选择文本效果最好的模型 M3.2 和风格效果最好的模型 M3.4,与模型 DP-Latent,StyIns^[23],NAST,GenRes 以及 OT4^[24] 进行对比,结果如表 3 (右)所列。模型 M3.2 的 BLEURT 分数以及 Mis 分数分别领先于最优模型 6.2%,3%;模型 M3.4 在风格方面领先于最优模型 1%。

结果表明,MRRL 在两种最流行的任务中不仅可以提高文本风格的准确率,同时也能够保证文本内容的完整性。

表 3 与其他模型效果的对比

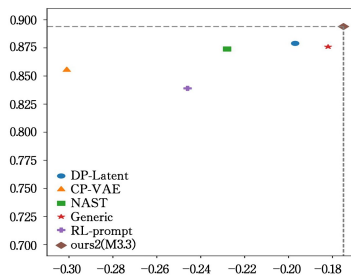
Table 3 Comparison of results with other models

YELP(情感极性转换)						GYAFC(正式文体转换)					
模型	BLEURT	Mis	BLEU	ACC	HM	模型	BLEURT	Mis	BLEU	ACC	HM
输入	-0.383	0.225	0.312	0.019	0.036	输入	-0.114	0.842	0.474	0.120	0.192
DP-Latent	-0.197	0.248	0.491	0.879	0.63	DP-Latent	-0.324	0.265	0.264	0.804	0.397
CP-VAE	-0.301	0.238	0.452	0.856	0.592	StyIns	-0.098	0.65	0.478	0.706	0.57
NAST	-0.228	0.411	0.546	0.874	0.672	NAST	-0.234	0.574	0.428	0.729	0.539
GenRes	-0.182	0.52	0.479	0.876	0.619	GenRes	0.09	0.811	0.507	0.908	0.651
RL-prompt	-0.246	0.485	0.421	0.839	0.561	OT4	0.112	-	0.618	0.942	0.746
Ours(M3.1)	-0.138	0.533	0.598	0.850	0.702	Ours(M3.2)	0.119	0.836	0.515	0.898	0.655
Ours(M3.3)	-0.175	0.517	0.578	0.894	0.702	Ours(M3.4)	0.11	0.828	0.505	0.952	0.66

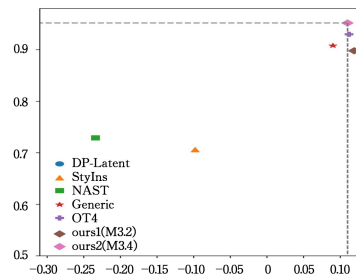
本文进一步探索和分析了各个模型在文本指标和风格指标之间的平衡效果,如图 5 所示,横坐标表示模型的文本保留 BLEURT 分数,纵坐标则表示模型的风格准确率 ACC 分数。

具体来说,在情感极性转换任务中,各个模型在风格准确率上的表现相当,但是文本分数相差很大。很多模型为了达到最佳的风格准确率而选择牺牲文本内容度量的效果,如 CP-VAE。模型 M3.3 不仅在风格分数方面领先于最优模型 1.7%,还在文本分数方面领先于最优模型 0.7%,达到了

最佳的均衡效果。在正式文体转换任务上,各个模型的效果参差不齐,有些模型则为了保证文本内容最佳而牺牲了风格准确率,如 StyIns 模型。虽然 GenRes 模型和 OT4 模型与 MRRL 模型在文本内容保存分数上接近,但 MRRL 模型在风格准确率方面的表现更加优秀,其中 M3.4 领先于 OT4 模型 1%。综上所述,本文中的 M3.3 模型在情感极性转换任务中表现出最佳的均衡效果,而 M3.4 模型在正式文体转换任务中也达到了最佳的均衡效果。



(a) YELP 数据集上的实现结果



(b) GYAFC 数据集上的实现结果

图 5 不同模型风格和内容权衡效果对比图

Fig. 5 Comparison of style and content trade-offs in different models

4.4.2 消融实验

表 4 详细列出了风格迁移模型在每个步骤中的效果。步骤 1 展示了使用平行数据进一步预热 BART 模型的过程。在情感极性转换任务中,由于缺乏平行数据,本文使用 3.1 节中提出的方法构造出两种不同的伪平行数据集,分别包含了约 5 万和 10 万的数据量用于预训练。在正式文体转换任务中,本文使用原始数据中提供的高质量平行数据集,分别使用

3 万和 5 万的平行数据进行预训练。由 M1.1 和 M1.2 的对比可以得出,使用更多的平行数据进行预训练,将会得到更好的预热效果。

在步骤 2 中,第一部分(M2.1—M2.2)本文使用了原始的 BART 模型直接进行文本风格迁移训练,这一步可以充分证明进一步预热对于提升模型效果的重要性。在第二部分(M2.3—M2.4)和第三部分(M2.5—M2.6),本文分别采用

M1.1 和 M1.2 中预训练过的模型进行迭代反向翻译训练。由第二部分和第三部分的对比可以观察到,正式文体转换任务中使用 M1.2 预热的模型效果更好,但是情感极性转换任务中使用 M1.1 预热的模型效果更好,即使用更多的自然存在、人工

编写的文本效果更佳。表 4 中, R_1 代表由风格奖励和对抗性相似奖励组成的奖励函数组合, R_2 是在 R_1 的基础上增加了 MIS 文本奖励。实验结果表明,奖励组合 R_1 更加适用于情感极性转换任务,奖励组合 R_2 则更加适用于正式文体转换任务。

表 4 模型实现过程中不同步骤的结果

Table 4 Results of different steps in model implementation

数据集(任务)	YELP(情感极性转换)					GYAFC(正式文体转换)					
	模型	BLEURT	Mis	BLEU	ACC	HM	BLEURT	Mis	BLEU	ACC	HM
M0:原始 BART 模型		-0.388	0.226	0.309	0.022	0.041	-0.116	0.840	0.414	0.333	0.369
Step 1:进一步地进行预训练											
M1.1:预训练(5万/3万)		-0.186	0.469	0.476	0.689	0.563	0.106	0.823	0.548	0.884	0.676
M1.2:预训练(10万/5万)		-0.176	0.471	0.552	0.749	0.636	0.114	0.825	0.553	0.889	0.682
Step 2:使用基于强化学习的无监督迭代反向翻译训练											
M2.1:M0+IBT+ R_1		-0.275	0.36	0.332	0.682	0.446	0.075	0.818	0.467	0.814	0.594
M2.2:M0+IBT+ R_2		-0.298	0.37	0.335	0.749	0.463	0.082	0.823	0.457	0.778	0.576
M2.3:M1.1+IBT+ R_1		-0.245	0.489	0.523	0.885	0.657	0.116	0.808	0.519	0.911	0.661
M2.4:M1.1+IBT+ R_2		-0.224	0.502	0.531	0.876	0.661	0.116	0.818	0.524	0.921	0.668
M2.5:M1.2+IBT+ R_1		-0.247	0.469	0.525	0.893	0.664	0.113	0.82	0.524	0.935	0.672
M2.6:M1.2+IBT+ R_2		-0.264	0.488	0.491	0.896	0.637	0.118	0.822	0.53	0.934	0.676
Step 3:并行数据集生成与模型微调											
M3.1:M2.3数据集+ R_1		-0.138	0.533	0.598	0.850	0.702	0.121	0.835	0.511	0.896	0.651
M3.2:M2.4数据集+ R_2		-0.146	0.531	0.586	0.851	0.694	0.119	0.836	0.515	0.898	0.655
M3.3:M2.5数据集+ R_1		-0.175	0.517	0.578	0.894	0.702	0.114	0.828	0.485	0.942	0.64
M3.4:M2.6数据集+ R_2		-0.212	0.517	0.553	0.896	0.685	0.11	0.828	0.505	0.952	0.66

在步骤 3 中,本文采用了步骤 2 中训练得到的模型,生成高质量的平行数据,以进行静态有监督训练。总体而言,各方面的效果都获得了显著提升。两个任务均在第一部分(M3.1—M3.2)中获得了不错的文本效果;在第二部分(M3.3—M3.4)中获得了不错的风格效果。

4.4.3 实验结果反思

根据上述实验结果,在文本风格迁移过程中采用 MIS 指标显著提高了生成效果,尤其在正式文体转换任务中表现突出。表 5 列出了额外的论证,展示了在 N 个不同系统上, BLEURT, MIS 和 BLEU 这 3 个文本指标在两个文本风格迁移任务中的 Pearson 相关系数。Pearson 相关系数可用于衡量两个评价指标之间线性相关程度的强度和方向。如表 5 所列,评价指标之间都是正相关,并且相关性非常显著。其中,使用 MIS 指标作为奖励函数时,与其他两个指标的相关性都很高,特别是在正式文体转换任务上。这充分证明了在正式文体转换任务中使用 MIS 指标作为奖励函数可以更好地衡量生成结果。同时也表明 MIS 指标在文本风格迁移应用中具有较高的可靠性和有效性,可以作为一种重要的评估指标来衡量生成模型的性能。

表 5 文本指标之间的 Pearson 相关系数

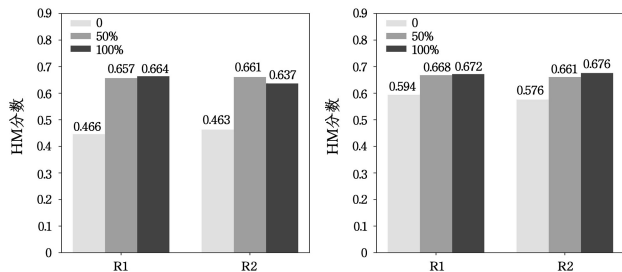
Table 5 Pearson correlation coefficients between text metrics

任务	N	BLEURT	MIS	BLEU
		MIS	BLEU	BLEURT
情感极性转换任务	13	0.88	0.927	0.862
		($P<0.01$)	($P<0.01$)	($P=0.01$)
正式文体转换任务	16	0.959	0.913	0.864
		($P<0.01$)	($P<0.01$)	($P<0.01$)

此外,本文使用了不同比例的平行数据集进行模型测试,以考察预热数据对模型最终结果的影响。如图 6 所示,在两个文本风格迁移任务中,未提前使用预热数据的模型表现

最差。在正式文体转换任务中,随着平行数据集比例的增大,两种奖励方法对应的 HM 分数都有所提高。在情感转换任务中,使用伪平行数据集进行预热能够显著提升模型性能,这证明在 MRRL 方法中采用的伪平行数据集构造方法非常有效。然而,当伪平行数据集的比例达到 100% 时,使用奖励组合 R_1 的模型性能仅得到微弱提升,而使用奖励组合 R_2 的模型性能则出现轻微下降。由此可见,使用伪平行数据集进行预热的方法在情感极性转换和正式文体转换两个任务上的表现不一致。

在情感极性转换任务中,本文使用了风格标注模型和掩码语言模型来自动构造伪平行语料库。虽然构造的平行数据在风格指标和文本指标上表现出色,但它们都属于模型生成的数据,与人类的相关性水平不够理想。而在正式文体转换任务中,伪平行语料库的构造过程参考了一部分人工编写的平行数据进行微调,这导致生成的数据效果有了显著的提升。因此,模型自动生成的伪平行数据仍然会受到限制,距离人工编写的平行语料库还有一定的差距。



(a) 情感极性转换任务

(b) 正式文体转换任务

图 6 平行数据集规模对模型测试结果的影响

Fig. 6 Influence of parallel dataset size on model performance

结束语 本文提出了基于多奖励强化学习的半监督文本风格迁移方法。训练过程中使用自动构造的伪平行语料库

引导语言模型进行进一步的预热,实验表明该构造方法是有效的。此外,利用多奖励函数联合训练语言模型,显著地增强了模型生成文本的语义保留度和逻辑一致性。最终再使用最优模型生成的高质量数据进行静态有监督训练,提高模型的稳定性。实验结果表明,该方法在两个文本风格迁移任务中获得了很好的性能,情感极性转化任务使用对抗性相似奖励和风格奖励效果更佳,正式文体转换任务中搭配使用 MIS 文本奖励的效果更好。

本文提出了构造伪平行数据的方法,虽然该方法带来了显著的性能提升,但与人工编写的文本仍存在一定差距。未来有望能进一步提高生成文本的人类相关性水平。此外,我们也考虑把目前的工作与大语言模型相结合,继续增强文本风格迁移领域的模型性能。

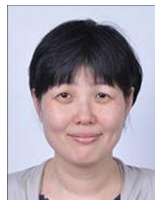
参考文献

- [1] HU Z, LEE R K W, AGGARWAL C C, et al. Text style transfer: A review and experimental evaluation[J]. ACM SIGKDD Explorations Newsletter, 2022, 24(1): 14-45.
- [2] TOSHEVSKA M, GIEVSKA S. A review of text style transfer using deep learning[J]. IEEE Transactions on Artificial Intelligence, 2021, 3(5): 669-684.
- [3] JIN D, JIN Z, HU Z, et al. Deep learning for text style transfer: A survey[J]. Computational Linguistics, 2022, 48(1): 155-205.
- [4] LI J, JIA R, HE H, et al. Delete, retrieve, generate: a simple approach to sentiment and style transfer[J]. arXiv: 1804. 06437, 2018.
- [5] LYU Y, LIANG P P, PHAM H, et al. StylePTB: A Compositional Benchmark for Fine-grained Controllable Text Style Transfer[C]//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2021: 2116-2138.
- [6] KASHYAP A R, HAZARIKA D, KAN M Y, et al. So Different Yet So Alike! Constrained Unsupervised Text Style Transfer [C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2022: 416-431.
- [7] LIU D, FU J, ZHANG Y, et al. Revision in continuous space: Unsupervised text style transfer without adversarial learning [C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020: 8376-8383.
- [8] RILEY P, CONSTANT N, GUO M, et al. TextSETTR: Few-Shot Text Style Extraction and Tunable Targeted Restyling [C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2021: 3786-3800.
- [9] NARASIMHAN S, DEY S, DESARKAR M. Towards Robust and Semantically Organised Latent Representations for Unsupervised Text Style Transfer[C]//Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2022: 456-474.
- [10] LUO F, LI P, ZHOU J, et al. A Dual Reinforcement Learning Framework for Unsupervised Text Style Transfer[C]// Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, 2019.
- [11] LAI H, TORAL A, NISSIM M. Generic resources are what you need: Style transfer tasks without task-specific parallel training data[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. 2021: 4241-4254.
- [12] LEE J. Stable Style Transformer: Delete and Generate Approach with Encoder-Decoder for Text Style Transfer[C]//Proceedings of the 13th International Conference on Natural Language Generation. 2020: 195-204.
- [13] LEE D, TIAN Z, XUE L, et al. Enhancing Content Preservation in Text Style Transfer Using Reverse Attention and Conditional Layer Normalization [C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2021: 93-102.
- [14] WANG J, ZHANG R, CHEN J, et al. Text Style Transferring via Adversarial Masking and Styled Filling[C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. 2022: 7654-7663.
- [15] 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. 2002: 311-318.
- [16] LIN C Y. ROUGE: A Package for Automatic Evaluation of summaries[C]//Proceedings of the Workshop on Text Summarization Branches Out (WAS 2004). 2004.
- [17] LEWIS M, LIU Y, GOYAL N, et al. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 7871-7880.
- [18] SELLAM T, DAS D, PARIKH A. BLEURT: Learning Robust Metrics for Text Generation[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 7881-7892.
- [19] BABAKOV N, DALE D, LOGACHEVA V, et al. A large-scale computational study of content preservation measures for text style transfer and paraphrase generation[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop. 2022: 300-321.
- [20] TOKPO E K, CALDERS T. Text Style Transfer for Bias Mitigation using Masked Language Modeling[C]//Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop. 2022: 163-171.
- [21] REID M, ZHONG V, LEWIS: Levenshtein Editing for Unsupervised Text Style Transfer[C]//Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. 2021: 3932-3944.
- [22] LI Z, QU L, XU Q, et al. Variational autoencoder with disentanglement priors for low-resource task-specific natural language

- generation[C]//2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022). Association for Computational Linguistics,2022;10335-10356.
- [23] YI X,LIU Z,LI W, et al. Text style transfer via learning style instance supported latent space[C]//Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence. 2021;3801-3807.
- [24] NOURIN. Text Style Transfer via Optimal Transport. [C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Abu Dhabi, United Arab Emirates; Association for Computational Linguistics, 2022; 2532-2541.
- [25] DENG M, WANG J, HSIEH C P, et al. RLPrompt: Optimizing Discrete Text Prompts with Reinforcement Learning[C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. 2022;3369-3391.
- [26] LIU Z,CHEN N. Learning from Bootstrapping and Stepwise Reinforcement Reward; A Semi-Supervised Framework for Text Style Transfer[C]//Findings of the Association for Computational Linguistics; NAACL 2022. 2022;2633-2648.
- [27] KRISHNA K, WIETING J, IYYER M. Reformulating Unsupervised Style Transfer as Paraphrase Generation[C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing(EMNLP). 2020;737-762.
- [28] LAFFERTY J, MCCALLUM A, PEREIRA F. Conditional random fields; Probabilistic models for segmenting and labeling sequence data[C]//ICML. 2001.
- [29] CHEN K J, FEI Z Y, CHEN J Q, et al. A survey on text style transfer[J]. Journal of Software,2022,33(12):20.
- [30] ZHAO J, KIM Y, ZHANG K, et al. Adversarially regularized autoencoders[C]//International Conference on Machine Learning. PMLR,2018;5902-5911.
- [31] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems-Volume 2. 2014;2672-2680.
- [32] HUANG Y, ZHU W, XIONG D, et al. Cycle-Consistent Adversarial Autoencoders for Unsupervised Text Style Transfer[C]//Proceedings of the 28th International Conference on Computational Linguistics. 2020;2213-2223.
- [33] WILLIAMS R J. Simple statistical gradient-following algorithms for connectionist reinforcement learning[J]. Machine Learning, 1992,8;229-256.
- [34] RAO S, TETREAULT J. Dear sir or madam, may I introduce the GYAF dataset; corpus, benchmarks and metrics for formality style transfer[C]//Proceedings of the ACL. 2018;129-140.
- [35] WOLF T, DEBUT L, SANH V, et al. Transformers; State-of-the-art natural language processing [C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing; System Demonstrations. 2020;38-45.
- [36] KIM Y. Convolutional Neural Networks for Sentence Classification[J]. arXiv:1408.5882,2014.
- [37] KINGMA D P, BA J. Adam; A method for stochastic optimization[J]. arXiv:1412.6980,2014.
- [38] HE J, WANG X, NEUBIG G, et al. A Probabilistic Formulation of Unsupervised Text Style Transfer[C]//International Conference on Learning Representations. 2019.
- [39] XU P, CHEUNG J C K, CAO Y. On variational learning of controllable representations for text without supervision[C]//International Conference on Machine Learning. PMLR, 2020; 10534-10543.
- [40] HUANG F, CHEN Z, WU C H, et al. NAST; A Non-Autoregressive Generator with Word Alignment for Unsupervised Text Style Transfer[C]//Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. 2021;1577-1590.



LI Jingwen, born in 1998, postgraduate. Her main research interests include text generation and text style transfer.



RUAN Tong, born in 1973, professor, Ph.D supervisor. Her main research interests include medical big data, natural language processing and text generation.

(责任编辑:喻藜)