计算机研究与发展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202220713
60(12): 2844−2863, 2023

基于中间域语义传导的跨领域文本生成方法

马廷淮[1] 于 信[1] 荣 欢[2]
[1](南京信息工程大学软件学院 南京 210044)
[2](南京信息工程大学人工智能学院(未来技术学院) 南京 210044)
(thma@nuist.edu.cn)

Cross-Domain Text Generation Method Based on Semantic Conduction of Intermediate Domains

Ma Tinghuai[1], Yu Xin[1], and Rong Huan[2]
[1](*School of Software*, *Nanjing University of Information Science & Technology*, *Nanjing* 210044)
[2](*School of Artificial Intelligence* (*School of Future Technology*), *Nanjing University of Information Science & Technology*, *Nanjing* 210044)

Abstract The deep neural network has been widely used in natural language processing. In text generation tasks with multi-domain data, there is often a discrepancy of data in different domains. And the introduction of new domains can simultaneously bring about the problem of data deficiency. The supervised methods require a large amount of data containing ground-truth in the domain of the task to train a deep neural network text generation model, and the trained model cannot achieve good generalization in a new domain. To address the problems of data distribution differences and data deficiency in multi-domain tasks, a comprehensive transfer text generation method inspired by transfer learning methods is designed to reduce the data distribution differences in text data between different domains while leveraging the semantic correlation on text data between source domain and target domain to help deep neural network text generation models generalize over new domains. The effectiveness of the proposed method for domain transfer is verified through experiments on a publicly available dataset, and the transfer deep neural network text generation model has a better performance in text generation on new domains. Also, the proposed method improves in all text generation evaluation metrics compared with other existing transfer text generation methods.

Key words deep neural network; text generation model; data distribution alignment; maximum mean discrepancy; zero-shot learning; semantic conduction

摘 要 在多领域数据的文本生成场景中,不同领域中的数据通常存在差异性,而新领域的引入会同时带来数据缺失的问题.传统的有监督方法,需要目标领域中大量包含标记的数据来训练深度神经网络文本生成模型,而且训练好的模型无法在新领域中取得良好的泛化效果.针对多领域场景中数据差异和数据缺失的问题,受到迁移学习方法的启发,设计了一种综合性的迁移式文本生成方法,减少了不同领域之间文本数据的差异性,同时借助已有领域和新领域之间文本数据上的语义关联性,帮助深度神经网络文本生成模型在新领域上进行泛化.通过在公开数据集上的实验,验证了所提方法在多领域场景下领域迁移的有效性,模型在新领域上进行文本生成时具有较好的表现,对比现有的其他迁移式文本生成方法,在各

收稿日期: 2022-08-16; **修回日期**: 2023-01-11
基金项目: 国家自然科学基金项目 (62102187, 62372243); 江苏省自然科学基金 (基础研究计划) 项目 (BK20210639); 国家重点研发计划项目 (2021YFE0104400)
This work was supported by the National Natural Science Foundation of China (62102187, 62372243), the Natural Science Foundation of Jiangsu Province (Basic Research Program) (BK20210639), and the National Key Research and Development Program (2021YFE0104400).
通信作者: 荣欢(ronghuan@nuist.edu.cn)

项文本生成评价指标上均有提升.

关键词 深度神经网络；文本生成模型；数据分布对齐；最大均值差异；零次学习；语义要素传导

中图法分类号 TP183

21 世纪以来，随着互联网的快速发展，出现了大批的互联网媒体平台，例如新闻传媒机构、网络购物网站、社交网络平台等，这些平台的出现使得互联网中的数据呈指数级增长. 在这其中，文本数据由于其编写容易、传播方便的特性成为了这些平台中数据的主要组成. 大量文本数据的涌现，导致平台中的用户很难在短时间内获取到自己想要的信息，这既不利于互联网平台的发展同时又降低了用户的浏览体验，为此需要快速有效的方法从海量文本中提炼出关键的信息. 文本生成方法作为自然语言处理领域的重要研究内容之一，利用深度神经网络模型可以实现自动化的文本摘要（automatic text summarization）生成，例如给长文章生成相应的摘要内容，或者给新闻生成对应的标题等. 通过自动文摘技术可以从海量文本数据中生成能准确反映原文中心内容的简短文本，这既帮助用户快速筛选出了有价值的文本信息内容，又降低了各个平台的人工编辑成本，提升了内容的传播速率，因此具有重要的现实意义[1].

然而，传统的基于深度神经网络的自动文摘生成模型依赖于大量的含有标注的数据进行模型的训练[2]，且训练出来的模型只适用于单一的任务领域，无法在其他领域中有效地泛化. 但在实际的应用场景中，文本数据往往存在多主题、多领域的特点[1]，且一个新领域出现时，很难在短时间内获得该领域中大量含有标注的数据对文本生成模型进行传统有监督地训练. 因此，在目标领域参考真值标注数据缺失的情况下，如何有效训练深度神经网络文本生成模型，以达到较好的领域泛化效果值得进一步研究.

为了解决上述的问题，现有工作多采用迁移学习中的“预训练-微调”（pre-train & fine-tune）方法，来缓解目标任务领域中已标注真值数据缺失的限制[2]，即针对给定的深度神经网络文本生成模型，由相关源域中大量已标注的文本数据对生成模型进行预训练；在此基础上，基于从源域学习到的模型参数，通过目标域中少量已标注的文本数据对模型进行微调[2]，以使生成模型由源域有效迁移至目标域，从而达到领域适应的目的. 由此，通过引用相关源域的先验知识，辅助标注数据量较少的目标域完成摘要文本的生成.

然而，“预训练-微调”的迁移学习范式仍存在不足. 首先，源域和目标领域之间存在较明显的数据差异，除通过微调手段外，仍需进一步从数据分布的角度消除数据差异对领域迁移效果的负面影响. 其次，当目标域中缺少足够或不存在任何可用于微调的标注数据时，所给定深度文本生成模型无法通过微调有效适应至目标领域，进而导致迁移式文本生成性能欠佳，直接削弱了文本生成模型在目标领域上的适应性.

对此，零次学习（zero-shot learning）提供了较好的思路启发[3]，通过特征属性为各领域构建“领域要素”（domain prototype）以描述该领域下的数据语义，通过不同领域要素之间的语义关联性，由最相关源域的“已标注样本”辅助处理目标域“未标注样本”（即语义要素传导），进而针对自动文本摘要生成任务. 即便没有给定任何目标域人工标注数据，仍可借助深度文本生成模型，根据零次学习语义要素传导原理，为目标域中大量未标注原始文本产生领域适应性较好的目标领域摘要文本[4].

综上所述，本文提出了一种基于中间域语义传导的跨领域文本生成方法，旨在通过源域和目标域数据之间的语义关联，由最为相关的源域已标注样本指导目标域文本生成，从而克服新领域标注样本稀缺的限制，提升深度文本生成模型在真实场景中的可用性. 本文的主要贡献有 5 点：

1）为源域数据和目标域数据构建文本数据语义要素；

2）改进深度神经网络文本生成模型内部结构，强化模型编码和解码过程，使模型可以接收文本语义要素的各个要素，从模型结构上提升领域间的可迁移性；

3）在核空间中，对源域数据和目标域数据进行数据表示分布对齐，缓解不同领域间数据表示的分布差异对领域间迁移所带来的负面影响，在数据表示层面增强了领域间的可迁移性；

4）将源域数据和目标域数据按照文本相似性综合指标划分至 K 个中间过渡域中，由此目标域数据可以通过更为恰当的领域数据选择，在生成过程中参考更具有语义相似性的源域数据；

5)基于改进后的文本生成模型,为文本语义要素中的不同要素构建相应的文本生成损失函数,以此引导模型捕捉跨领域数据在语义要素上的近似参考关系,进而学习到跨领域数据间的语义关联,从而在中间域内将相关新源域已标注文本作为目标域无标注原始文本的可参考真值.

1 相关工作

自动文本摘要生成技术属于自然语言处理领域中文本生成任务的一个分支[1]. 当前主流的自动文本摘要生成模型主要依赖于大量已标注真值摘要样本对生成模型进行有监督训练,从而得到具有较好生成性能的模型. 但在实际应用场景中常出现真值文本缺失的问题,由此引入了迁移学习相关方法用于解决此问题. 现对自动文本摘要生成方法、文本生成任务中传统的迁移学习方法以及零次学习方法相关工作进行归纳总结.

1.1 自动文本摘要生成方法

自动文本摘要生成是指利用计算机通过算法自动地将文本或文本集合转换成简短摘要,帮助用户通过摘要全面准确了解原始文献的中心内容[1],此类自动文本摘要生成任务的变体包括论文生成摘要、新闻生成标题[5]、海量社交媒体文本生成的关键内容.

当前主流的自动文本摘要生成方法可分为抽取式(extractive)和生成式(abstractive). 抽取式方法是从原始文章中提取突出的句子或短语[1];而生成式方法则产生新的词语或短语,这些词语可能会改写或使用原始文章中没有的词语[6]. 在本文中,主要研究生成式文本摘要生成模型,具体是根据给定原始文本产生相应的标题.

近年来,许多研究者采用序列到序列(sequence to sequence)的模型结构建立生成式文本摘要生成模型. Rush 等人[7]在“编码器-解码器”的形式中,将包含注意力(attention)机制的循环神经网络(recurrent neural network, RNN)应用于生成式摘要任务,与传统的方法相比,该方法的性能得到了有效的提升;吴仁守等人[8]同样基于“编码器-解码器”的形式,但在编码器端引入全局自匹配机制,根据文本中每个单词的语义和文本整体语义的匹配程度,寻找出文本的核心内容为给定文本生成核心摘要内容;Narayan 等人[9]使用指针生成器网络[10]在输入文档中识别突出的句子和关键词,将句子和关键词结合以形成最终的摘要. 此外,文本摘要生成模型也可以通过基于自注意力(self-attention)机制的神经网络组件进行构建,如 Transformer[11]. 基于 Transformer 的文本生成模型同样以“编码器-解码器”的形式进行构建,解决了传统 RNN 架构不能并行计算的问题,提高了文本生成的效率. 劳南新等人[12]将改进的预训练语言模型作为编码器,用于提取词级粒度的信息特征,同时采用多层 Transformer 作为解码器,以字为粒度生成混合字词特征的中文文本摘要.

由此可见,目前主流的文本生成模型结构仍为“编码器-解码器”的形式. 目前采用 RNN 或 Transformer 对其构建,结构为“编码器-解码器”的生成式文本摘要生成模型通常采用传统有监督方式进行训练[13],并不适用于目标域已标注真值样本缺失的应用场景[14],这意味着需要研究针对此类场景下的迁移式文本生成方法,以克服目标域已标注真值数据稀缺的限制.

1.2 文本生成中的迁移学习方法

对于迁移学习方法在文本生成任务中的应用,已有研究工作表明,使用特定语料数据训练的模型不能跨领域通用[15]. 目前,传统迁移学习方法侧重于通过某种迁移策略,由源域数据辅助目标域完成特定任务[13]. 典型的迁移策略包括 3 个方面:

1)基于参数的迁移策略. 先从源域数据中学习模型参数;再基于全部或部分已学习到的模型参数,在目标域数据上进行微调;最后使用微调后的模型完成目标任务. 这也是目前最常见的迁移学习策略.

2)基于特征的迁移策略. 侧重于寻找“好的”特征表示,以减少源域和目标域之间数据的表示差异.

3)基于关系的迁移策略. 根据领域语义关联在源域和目标域之间建立映射.

在基于参数的迁移策略研究方面,随着深度学习的不断发展,预训练模型被引入到自然语言生成任务中并获得了广泛的应用. 通过使用大规模语料库获得预训练模型,并使用目标域中相对少量的训练数据对预训练模型进行微调,实现从源域到目标域的迁移[16]. 按照“预训练-微调”模式,多种预训练语言模型被提出. 具体地:Raffel 等人[17]提出了预训练文本生成模型 T5,通过使用包含多个领域数据的大规模 common crawl 数据库来进行不同跨度掩码填充任务的预训练;Lewis 等人[18]使用去噪自动编码器预训练了序列到序列的模型 BART,在预训练过程中采用噪声函数来掩码随机跨度的文本,引导模型学习如何重建原始文本;Zhang 等人[19]提出的预训练文本生成模型 PEGASUS 在语料库中学习如何重新填充

多个被掩码的句子以进行预训练.

在基于特征的迁移策略研究方面，有研究者提出了用多种方法来获得文本或特征上的可迁移表示，从而在不同特征空间的领域之间转移知识. 由于不同特征空间之间通常没有对应关系，因此需要额外的信息来连接各个领域[20]. 通过将不同领域之间的数据联系起来，在尽可能保留数据原始特征信息的同时，减少源域和目标域之间的数据特征差异，从而达到领域适应目的. 具体地，Chen 等人[21] 设计了一种广义协变量迁移假设方法对无监督领域适应问题进行建模，通过在子空间中应用分布适应函数并使用凸优化损失函数，使源域数据分布适应于目标域数据分布，从而解决当领域差异较大时，传统特征转换方法不能使转换后的源域分布和目标域分布近似的问题；Li 等人[22] 提出一种基于矩阵分解的半监督异构域适应方法，在再生希尔伯特核空间（reproducing kernel Hilbert space, RKHS）内进行矩阵分解，利用特征和数据实例之间的非线性关系学习源域和目标域的异质特征，以弥补核空间中源域和目标域之间的特征差异；Zellinger 等人[23] 提出了基于度量的正则化方法，该方法通过最大化不同领域中特定激活分布之间的相似性，来表示不同领域中相似的潜在特征，以实现无监督的领域自适应；王文琦等人[24] 和 Deng 等人[25] 没有直接将不同领域的数据表示进行对齐，而是利用生成对抗网络，将源域和目标域中的原始文档输入到生成器中生成新的文本，使判别器无法区分生成文本所属领域，从而获得不同领域数据潜在的迁移式文本表示.

现有的研究表明[16]，一方面，通过少量目标域数据微调预训练语言模型，可以有效地进行语言模型的领域适应. 但另一方面，将预训练语言模型应用到目标领域时，仍需通过一定量的数据对模型进行微调才能达到较好的领域适应效果[26]. 若目标域缺乏已标注真值数据，会直接影响模型在目标域中的泛化效果，新领域标注数据缺失的限制仍然存在. 因此越来越多的研究者开始关注在目标域缺乏已标注数据的情况下，研究更有效的方法将文本生成模型从源域向目标域迁移，从而在目标域中达到较好的文本生成效果.

1.3 文本生成中的零次学习方法

在基于关系的迁移策略方面，近年来，许多研究者将零次学习[27] 相关方法应用于迁移式文本生成任务中. 零次学习方法相比于传统的迁移学习方法，更加针对于解决目标域已标注样本缺失的问题. 在目标域可参考真值数据缺失的条件下，零次学习方法通常会给每个领域构建相应的“要素描述”. 由此，即使输入数据是未标注的，但若输入数据的一组属性“接近”某个领域的“要素描述”，就可以推断出给定输入数据的类别标签[4]. 由此，目标域中缺乏可参考真值数据的问题就可以通过领域要素传导的方式解决. 具体地：Zhao 等人[28] 通过从各领域数据选择若干具有代表性的对话文本，将相应的真值文本作为种子，以及将代表性对话文本中的关键实体词作为注释，使用跨域编码器对源域和目标域之间共享的领域要素进行编码，再通过解码器生成对话文本，由此根据不同领域间领域要素的相似性实现了从源域到目标域的迁移；Liu 等人[29] 在多语言场景下的源语言和目标语言中收集语义相似的术语（包括从目标语言真值文本中所收集的词汇）作为领域语义要素，并在此基础上，使用隐变量模型处理不同语言间相似句子的领域分布差异；Ayana 等人[30] 和 Duan 等人[31] 提出的迁移式文本生成模型将源域的原始文档作为输入，直接为目标域生成文本，并采用目标域真值文本训练生成模型，并通过建立结构相同的精简文本生成模型，模仿“输入→输出”过程，建立从源域到目标域的语义要素映射，最终将目标域的原始文档作为输入，以产生目标域对应的文本生成结果. 由此可见，目前已有大量的零次学习方法用于解决跨域的文本生成任务，但目前应用在跨域文本生成任务中的零次学习方法通常会使用目标域真值数据参与领域语义要素构建. 但是当目标域真值数据缺失时，相关工作仍存在限制.

综上所述，通过对现有迁移式文本生成方法的归纳总结，发现仍有 3 个方面需进一步研究：首先，通过大规模语料库预训练的语言模型应用到目标域上时，仍然需要目标域中一定量的已标注数据进行微调，从而使模型适应到目标域，这意味着目标域中可参考真值数据缺失的限制依然存在；其次，不同领域间数据在数据表示分布上的差异性会对模型产生跨域的负面影响[15]，这意味需要通过有效的方法减少不同领域数据表示之间的差异性；最后，在进行跨域的模型生成过程中，目标域数据要尽可能地借助源域数据进行辅助，以提升文本生成效果，这意味需要从已有源域数据中挖掘出对目标域数据有帮助的信息，通过获取数据间信息的关联性改进模型获取关联信息的能力，针对目标域数据找出最有帮助的源域数据，从而辅助目标域数据生成.

2 方法设计与实现

采用基于零次学习方法进行迁移式文本生成的任务，主要的挑战是如何充分借助源域中已有的标注数据，帮助无参考真值的目标域数据进行文本生成.

本文要解决的问题可以定义为：给定源域的原始正文 X_{source}、源域真值文本 Y_{source} 和目标域的原始文本 X_{target}. 在目标域没有可参考真值文本 Y_{target} 的情况下，通过提出的基于零次学习语义要素传导的文本生成方法，生成出目标域的相应摘要文本 Y_{target}.

本节将分别从文本语义原型构建、迁移式文本生成模型构建、领域数据分布对齐、中间域重划分和零次学习语义要素传导这 5 个方面阐述所提出的迁移式文本生成方法.

1)在各个中间域中，为不同领域形如(新闻 x, 标题 y)的数据构建“语义要素”.

2)针对跨域迁移式的文本生成场景，改进“编码器-解码器”结构的文本生成模型，以适用于零次学习中的语义要素传导方法，实现从源域到目标域的迁移.

3)将源域和目标域数据的文本表示投射到再生希尔伯特核空间中，将源域的数据分布与目标域的数据分布对齐，从而减少不同领域之间数据分布差异所带来的负面影响，从数据表示层面提升领域间的可迁移性.

4)建立中间域，将源域和目标域中的数据根据文本相似性的综合指标重新划分至若干中间域中，使得在中间域内进行更为恰当的领域数据选择，为目标域数据分配了更具有语义相似性的源域数据.

5)通过零次学习语义要素传导，将中间域中的目标域无标注原始文本与新源领域中最相关的标题进行语义关联，根据语义要素上的相似或接近，为目标域原始文本迁移式生成摘要文本.

最终，在迁移式文本生成过程中，相关源域中的真值文本将充当目标域文本生成的参考真值，从而不再依赖于对目标域数据进行人工标注.

2.1 文本语义要素构建

首先，利用原始文本 x、相应的真值文本 y 和基于原始文本 x 得到的语义注释 a 这 3 个要素，为源域和目标域中各个数据(原始文本 x, 摘要文本 y)构建一个语义要素，记为 $z=(x^d, y^d, a^d)$, 其中，d 表示领域(domains)，$d\in\{\text{src, tar}\}$. 表示数据来自源域(source domains, src)或目标域(target domain, tar). 语义要素 z 中源域和目标域的原始文本表示为 x^{src} 和 x^{tar}; 源域的摘要文本表示为 y^{src}. 在涉及到目标域的摘要文本数据 y^{tar} 时，将根据相应的原始文本 x^{tar} 中每个子句与整个原始文本 x^{tar} 之间的 ROUGE-L指标得分，从原始文本 x^{tar} 中抽取得分最高的前 n 个子句作为当前目标域原始文本的“伪真值”y^{tar}(即目标域伪摘要文本). 此处，抽取的子句数量 n 由当前目标域原始文本 x^{tar} 所属中间域内源域(原始文本 x, 摘要文本 y)数据的平均长度压缩率决定; 源域和目标域的语义注释 a^{src} 和 a^{tar} 是将源域和目标域的原始文本 x^{src} 和 x^{tar} 分词转换为关键词序列得到的，该关键词序列中各词汇词性属于名词、动词、形容词或副词中的一种，并且各词汇均被赋予相应的情感极性值(即在 [−1,1] 之间). 由此，通过上述过程为源域和目标域中各“原始文本 x-(伪)摘要文本 y”对构建了数据级语义要素，记为 $z=(x^d, y^d, a^d)$, $d\in\{\text{src, tar}\}$.

2.2 迁移式文本生成模型构建

迁移式文本生成模型可以有效应对生成过程中目标域缺少参考真值的问题，本文设计了基于中间域的零次学习语义要素传导迁移式文本生成模型. 通过语义要素传导策略，迁移式文本生成模型可以学习到不同领域之间的文本语义关联，这样的语义关联可以被认为是所涉及领域的先验知识. 当为目标领域生成文本时，若无可供参考的真值数据，可将领域先验知识作为参考.

本文提出的迁移式文本生成模型基于“编码器-解码器”的形式进行构建，如图 1 所示.

图 1 中，编码器端由 2 个结构相同的编码器模块 E_1 和 E_2 组成. E_1 和 E_2 以及解码器端的解码器模块 D 是将 Transformer 模型[11] 与双向长短期记忆网络

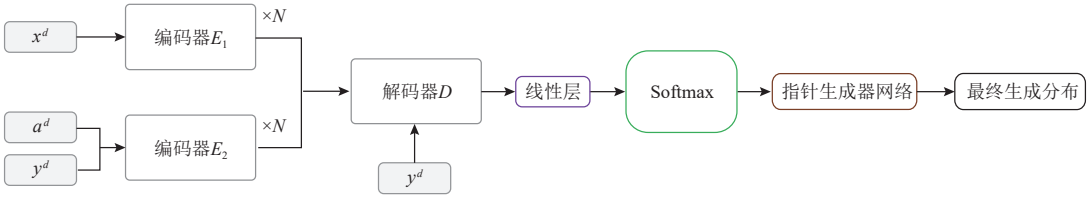
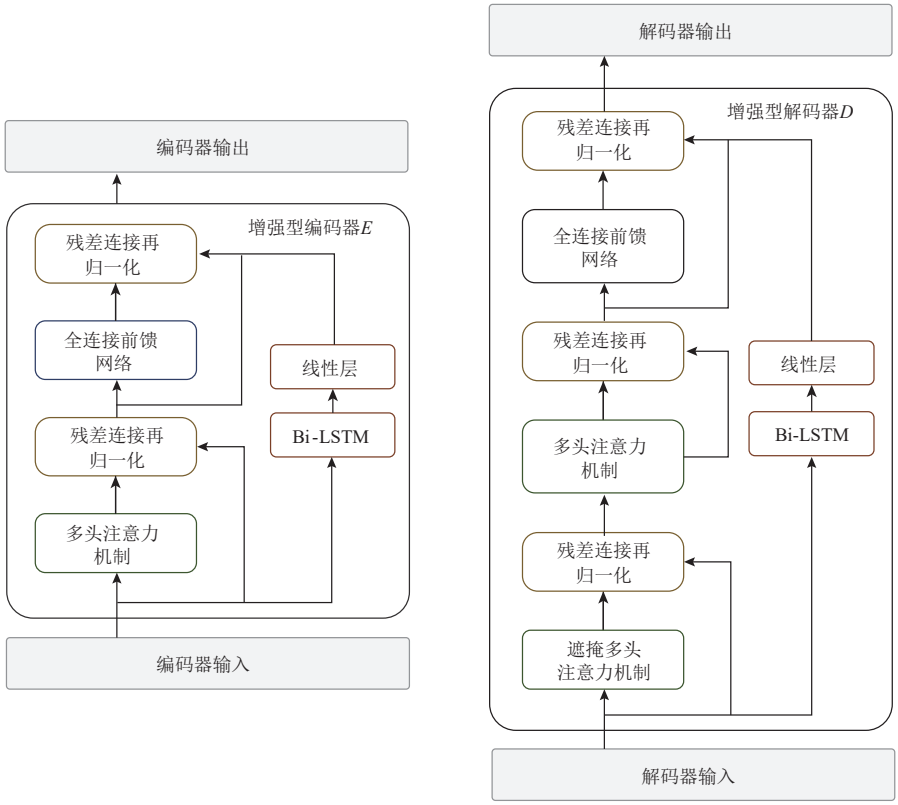


Fig. 1 Structure of the transferable text generation model

图 1 迁移式文本生成模型结构

(bidirectional long-short term memory, Bi-LSTM)相结合构建的，这样的设计使得迁移式文本生成模型可以整合自注意力机制与循环神经网络. 此外在模型解码端添加了指针生成器网络[10]，以解决文本生成任务中的未登录词(out-of-vocabulary, OOV)问题.

图 2 中迁移式文本生成模型的编码器模块 E 以及解码器模块 D 参考原始的 Transformer 模型[11]设计，每个模块中都包括了 N 个堆叠的子层，每一个子层中由多头注意力机制(multi-head attention)与全连接前馈(feed forward)网络组成，同时都采用了残差连接再归一化的处理. 将 Bi-LSTM 层添加到 E 和 D 的每个子层中，构建增强型的编码器与解码器. 在这样设计的每个子层中，Bi-LSTM 层的输入与子层的原输入相同，而输出在子层最后的归一化之前，与子层的原输出相加. 此外，如果 Bi-LSTM 使用与 Transformer 模型相同数量的隐藏单元数 h，就会得到维度为 $2h$ 的 Bi-LSTM 输出，因此设计添加一个线性层(linear layer)，将 Bi-LSTM 的输出维度 $2h$ 投射到维度 h，以便与 Transformer 的输出维度相匹配.



(a) 编码器 (b) 解码器

Fig. 2 Internal structure of encoder E and decoder D

图 2 编码器 E 和解码器 D 内部结构

由此，输入数据中的语义关联性(由 Transformer 中的自注意力机制提供)和时序依赖性(由 Bi-LSTM 提供)可以同时得到保留. 在模型训练过程中编码器端的编码器模块 E_1 用于接收原始文本 x^d 作为输入，另一个编码编码器模块 E_2 用于接收摘要文本 y^d 或语义注释 a^d 作为输入，而解码器端模块 D 会接收摘要文本 y^d 参与模型训练. 当摘要文本 y^d 是来自源域时，使用源域的真值摘要文本 y^{src}；当摘要文本 y^d 来自目标域时，则使用目标域的伪摘要文本 y^{tar}.

通过上述方式，将源域和目标域的原始文本 x^d 和摘要文本 y^d 同时反馈给编码器和解码器，从而在零次学习语义要素传导阶段建立源域和目标域数据之间的语义关联. 由此，在迁移式文本生成模型的训练过程中，解码器模块会分别和 2 个编码器模块的输出进行多头注意力计算[11]，在编码器端和解码器端捕捉原始文本 x^d、语义注释 a^d 和摘要文本 y^d 之间的全局依赖性. 此外，由于指针生成器网络的加入，解码器在生成文本的过程中，会使用指针生成器网络提供的“复制机制”[10]，在生成摘要文本的每个时间步上决定是从编码器端的输入文本中复制词汇或是从词表中生成词汇，从而完成最终的摘要文本生成.

本文构建的适用于语义要素传导的文本生成模型, 接收语义要素 $z=(x^d, y^d, a^d)$, $d\in\{\text{src}, \text{tar}\}$作为输入, 输出生成的摘要文本 $y^{d'}$. 具体地, 模型编码器接收语义要素 $z=(x^d, y^d, a^d)$, $d\in\{\text{src}, \text{tar}\}$作为输入, 在编码阶段, 编码器接收输入 $\boldsymbol{v}=(\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n)$得到编码器隐藏状态 $\boldsymbol{h}=(\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_n)$. 在解码阶段, 给定输入 $\boldsymbol{x}_t$ 后, 可以得出时间步骤 t 的解码隐藏状态 $\boldsymbol{s}_t$, 并计算出编码器隐藏状态 $\boldsymbol{h}$ 的注意力分布 $\boldsymbol{a}_t$, 以结合编码器隐藏状态 $\boldsymbol{h}$ 和解码器状态 $\boldsymbol{s}_t$ 的线性转换. 接下来, 在时间步骤 t, 由编码器隐藏状态对注意力分布的加权和计算得出上下文向量表示 $\boldsymbol{c}_t$. 于是可以得到词汇分布 $P_{\text{vocab}}(\boldsymbol{w}_t)$, 而 $P_{\text{vocab}}(\boldsymbol{w}_t)$表示在时间步骤 t 预测单词时词表中所有单词的概率分布.

此外, 使用指针生成器网络在解码的时间步骤 t 采用指针 p^t_{gen} 作为软开关, 以选择是按概率 $P_{\text{vocab}}(\boldsymbol{w}_t)$ 从词汇表中选择生成一个词汇, 或根据注意力权重 $\boldsymbol{a}_t$ 从输入的文本中复制一个词汇. 因此, 得到最终扩展词表的概率分布 $P(\boldsymbol{w}_t)$. 其中, p^t_{gen} 是根据上下文向量 $\boldsymbol{c}_t$、解码器状态 $\boldsymbol{s}_t$ 和解码器输入 $\boldsymbol{x}_t$ 计算得到的. 图 1 所示模型生成摘要文本 y^d_{gen} 的具体过程如式(1)所示:

$$\begin{cases} \boldsymbol{a}_t = softmax\left(\boldsymbol{v}^{\mathrm{T}}\tanh\left(\boldsymbol{W}_h\boldsymbol{h}_i+\boldsymbol{W}_s\boldsymbol{s}_t+\boldsymbol{b}_{\text{att}}\right)\right), \\ P_{\text{vocab}}(\boldsymbol{w}_t)=\tanh\left(\boldsymbol{V}_{\text{p}}[\boldsymbol{s}_t, \boldsymbol{c}_t]+\boldsymbol{b}_{\text{vocab}}\right), \\ P^t_{\text{gen}}=softmax\left(\tanh\left(\boldsymbol{W}_c\boldsymbol{c}_t+\boldsymbol{W}_s\boldsymbol{s}_t+\boldsymbol{W}_x\boldsymbol{x}_t+\boldsymbol{b}_{\text{gen}}\right)\right), \\ P(\boldsymbol{w}_t)=p^t_{\text{gen}}P_{\text{vocab}}(\boldsymbol{w}_t)+\left(1-p^t_{\text{gen}}\right)\sum\limits_{\boldsymbol{w}_t}\boldsymbol{a}_{t,j}, \end{cases} \tag{1}$$

其中 $\boldsymbol{v}$, $\boldsymbol{W}_h$, $\boldsymbol{W}_s$, $\boldsymbol{b}_{\text{att}}$, $\boldsymbol{V}_{\text{p}}$, $\boldsymbol{b}_{\text{v}}$, $\boldsymbol{W}_c$, $\boldsymbol{W}_x$, $\boldsymbol{b}_{\text{gen}}$ 都是可学习的参数.

由此, 在图 1 所示模型的训练过程中, 模型接收输入 x^d, y^d, a^d, 并按式(1)将词汇生成概率分布 P_{vocab} 和注意力概率分布 $\boldsymbol{a}_t$ 与指针开关 p^t_{gen} 加权求和获得最终的词序分布概率 $P(\boldsymbol{w}_t)$, 以生成相应的摘要文本 y^d_{gen}.

2.3 领域数据分布对齐

一般而言, 2 个领域的特征空间存在相似性与差异性[3]. 具体地, 不同的领域间有一些共同的特征, 但每个领域也有自己域的特有特征. 在领域适应的过程中, 利用不同领域的共同特征将不同的领域联系起来, 可以有效减少不同领域数据分布之间的差异性. 如图 3 所示, 2 个领域间会存在一些共同特征 $\boldsymbol{S}_c$ 和 $\boldsymbol{T}_c$, 其中 $\boldsymbol{S}_c$ 表示源域内部所包含的源域和目标域的共同特征, $\boldsymbol{T}_c$ 表示目标域内部所包含的源域和目标域的共同特征. 同时每个领域中也存在各自特有的领域特征 $\boldsymbol{S}_s$ 和 $\boldsymbol{T}_t$, 其中 $\boldsymbol{S}_s$ 表示源域特有特征, $\boldsymbol{T}_t$ 表示目标域特有特征. 因此, 为了在迁移式文本生成上取得更好的性能指标, 首先要对齐源域和目标域之间的数据分布表示, 以减小不同领域间数据表示的分布差异对迁移式文本生成造成的影响.

具体地, 通过预训练语言模型 BERT[32] 分别输出源域和目标域的文本词嵌入(word embedding)表示. 将源域原始文本表示为 $\boldsymbol{X}_{\text{src}}$, 输入特征的词嵌入表示为 $\boldsymbol{X}_{\text{src}}=[\boldsymbol{S}_c; \boldsymbol{S}_s]$, 其中 $\boldsymbol{S}_c$ 表示 $\boldsymbol{X}_{\text{src}}$ 中包含 c 个共同特征的特征矩阵, $\boldsymbol{S}_s$ 表示 $\boldsymbol{X}_{\text{src}}$ 中包含 s 个源域特有特征的特征矩阵; 目标域原始文本数据表示为 $\boldsymbol{X}_{\text{tar}}$, 输入特征的词嵌入表示为 $\boldsymbol{X}_{\text{tar}}=[\boldsymbol{T}_c; \boldsymbol{T}_t]$, 其中 $\boldsymbol{T}_c$ 表示 $\boldsymbol{X}_{\text{tar}}$ 中包含c个共同特征的特征矩阵, $\boldsymbol{T}_t$ 表示 $\boldsymbol{X}_{\text{tar}}$ 中包含 t 个目标域特有特征的特征矩阵, 如图 3 所示.

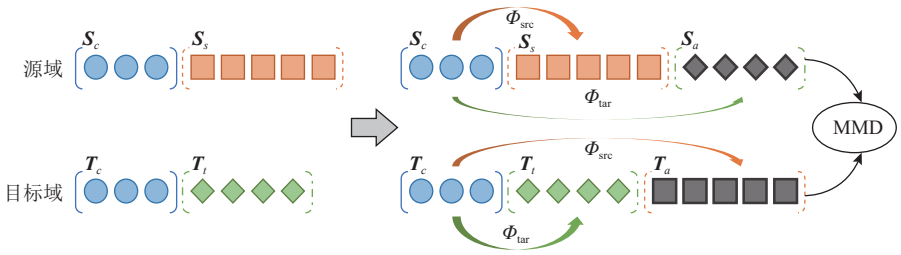


Fig. 3 Feature fill alignment

图 3 特征填充对齐

图 3 中, $\boldsymbol{X}_{\text{src}}$ 和 $\boldsymbol{X}_{\text{tar}}$ 之间的数据分布首先通过类交叉填充的方式实现特征填充对齐, 减小领域特有特征影响; 在此基础上, 使用最大均值差异(maximum mean discrepancy, MMD)在再生希尔伯特核空间内通过最小化最大均值差异以减小填充后的领域数据分布差异, 从数据分布层面对齐填充后的源域和目标域数据.

具体地:

1)特征映射函数 Φ_{src} 和 Φ_{tar} 将源域和目标域中的共同特征与各自领域中的特有特征进行映射联系, 如式(2)所示:

$$\begin{cases} \min_{\Phi_{\text{src}}}\|\Phi_{\text{src}}(\boldsymbol{S}_c)-\boldsymbol{S}_s\|^2, \\ \min_{\Phi_{\text{tar}}}\|\Phi_{\text{tar}}(\boldsymbol{T}_c)-\boldsymbol{T}_t\|^2. \end{cases} \tag{2}$$

2)将所得特征映射 Φ_{src} 和 Φ_{tar} 交叉作用于 $\boldsymbol{T}_c$ 和 $\boldsymbol{S}_c$ 上以进行特征填充, 如图 3 所示, 将从目标域得到

的特征映射 Φ_{tar} 应用到源域的共同特征 $\boldsymbol{S}_c$ 上，得到领域适应化特征矩阵 $\boldsymbol{S}_a$. 为目标域做相同的交叉操作，得到领域适应化特有特征矩阵 $\boldsymbol{T}_a$：

$$\begin{cases}\boldsymbol{S}_a=\Phi_{\text{tar}}(\boldsymbol{S}_c),\\ \boldsymbol{T}_a=\Phi_{\text{src}}(\boldsymbol{T}_c).\end{cases}\tag{3}$$

3)将源域和目标域的原始特征矩阵 $\boldsymbol{S}_c$、特有特征矩阵 $\boldsymbol{S}_s$ 和适应化特征矩阵 $\boldsymbol{S}_a$ 进行填充，分别得到填充后的特征矩阵$\boldsymbol{X}_{\text{s}_\text{f}}$和$\boldsymbol{X}_{\text{t}_\text{f}}$，如式(4)所示：

$$\begin{cases}\boldsymbol{X}_{\text{s}_\text{f}}=(\boldsymbol{S}_c;\boldsymbol{S}_s;\boldsymbol{S}_a),\\ \boldsymbol{X}_{\text{t}_\text{f}}=(\boldsymbol{T}_c;\boldsymbol{T}_t;\boldsymbol{T}_a).\end{cases}\tag{4}$$

特别地，式(3)中的 2 个特征映射 Φ_{src} 和 Φ_{tar} 可以分别表示为 $\Phi_{\text{src}}(\boldsymbol{S}_c)=\boldsymbol{W}_{\boldsymbol{S}}^{\text{T}}\boldsymbol{S}_c$ 和 $\Phi_{\text{tar}}(\boldsymbol{T}_c)=\boldsymbol{W}_{\boldsymbol{T}}^{\text{T}}\boldsymbol{T}_c$，则 $\boldsymbol{S}_a=\boldsymbol{W}_{\boldsymbol{T}}^{\text{T}}\boldsymbol{T}_c$，$\boldsymbol{T}_a=\boldsymbol{W}_{\boldsymbol{S}}^{\text{T}}\boldsymbol{S}_c$. 于是式(2)可以进一步推导为式(5)：

$$\begin{aligned}&\min_{\Phi_{\text{src}},\Phi_{\text{tar}}}\|\Phi_{\text{src}}(\boldsymbol{S}_c)-\boldsymbol{S}_s\|^2+\|\Phi_{\text{tar}}(\boldsymbol{T}_c)-\boldsymbol{T}_t\|^2=\\ &\min_{\Phi_{\text{src}},\Phi_{\text{tar}}}\left\|\boldsymbol{W}_{\boldsymbol{S}}^{\text{T}}\boldsymbol{S}_c-\boldsymbol{S}_s\right\|^2+\left\|\boldsymbol{W}_{\boldsymbol{T}}^{\text{T}}\boldsymbol{T}_c-\boldsymbol{T}_t\right\|^2=\\ &\min_{\Phi_{\text{src}},\Phi_{\text{tar}}}\text{tr}\left(\boldsymbol{S}_c^{\text{T}}\boldsymbol{W}_{\boldsymbol{S}}\boldsymbol{W}_{\boldsymbol{S}}^{\text{T}}\boldsymbol{S}_c-2\boldsymbol{S}_c^{\text{T}}\boldsymbol{W}_{\boldsymbol{S}}\boldsymbol{S}_s+\boldsymbol{S}_s^{\text{T}}\boldsymbol{S}_s\right)+\\ &\text{tr}\left(\boldsymbol{T}_c^{\text{T}}\boldsymbol{W}_{\boldsymbol{T}}\boldsymbol{W}_{\boldsymbol{T}}^{\text{T}}\boldsymbol{T}_c-2\boldsymbol{T}_c^{\text{T}}\boldsymbol{W}_{\boldsymbol{T}}\boldsymbol{T}_t+\boldsymbol{T}_t^{\text{T}}\boldsymbol{T}_t\right).\end{aligned}\tag{5}$$

4)为了使源域更好地适应于目标域，还需要确保式(4)所输出源域和目标域的特征矩阵$\boldsymbol{X}_{\text{s}_\text{f}}$和$\boldsymbol{X}_{\text{t}_\text{f}}$在分布上尽可能接近. 将填充对齐后的表示映射到再生希尔伯特核空间中；在此核空间中，通过最大均值差异来度量不同领域数据映射到核空间后的分布距离 *Dist*. 通过缩小$\boldsymbol{X}_{\text{s}_\text{f}}$和$\boldsymbol{X}_{\text{t}_\text{f}}$映射结果之间的分布距离 *Dist* 从而减小源域和目标域数据的分布差异，如式(6)所示：

$$\min Dist\left(\boldsymbol{X}_{\text{s}_\text{f}},\boldsymbol{X}_{\text{t}_\text{f}}\right)=\left\|\frac{1}{n_1}\sum_{i=1}^{n_1}\boldsymbol{X}_{\text{s}_\text{f}}-\frac{1}{n_2}\sum_{i=1}^{n_2}\boldsymbol{X}_{\text{t}_\text{f}}\right\|^2.\tag{6}$$

最后，源域文本词嵌入表示通过全连接层与激活函数 sigmoid 进行特征变换，再将其结果投射到核空间中，而目标域的文本词嵌入表示则直接投射到核空间中，如图 4 所示.

通过最小化式(6)中的目标函数 $Dist(\boldsymbol{X}_{\text{s}_\text{f}},\boldsymbol{X}_{\text{t}_\text{f}})$使源域与目标域的数据分布接近. 由此，图 4 中全连接层的参数将在式(6)目标函数最小化的过程中被更新.

按式(6)训练后，将源域全连接层映射 FC_Φ 输出的源域文本表示 $\boldsymbol{X}'_{\text{src}}$ 作为与目标域分布对齐的表示结果. 而目标域自身的文本表示 $\boldsymbol{X}'_{\text{tar}}$ 则是通过将目标域的原始词嵌入表示输入至源域映射 FC_Φ 中计算所得，如式(7)所示：

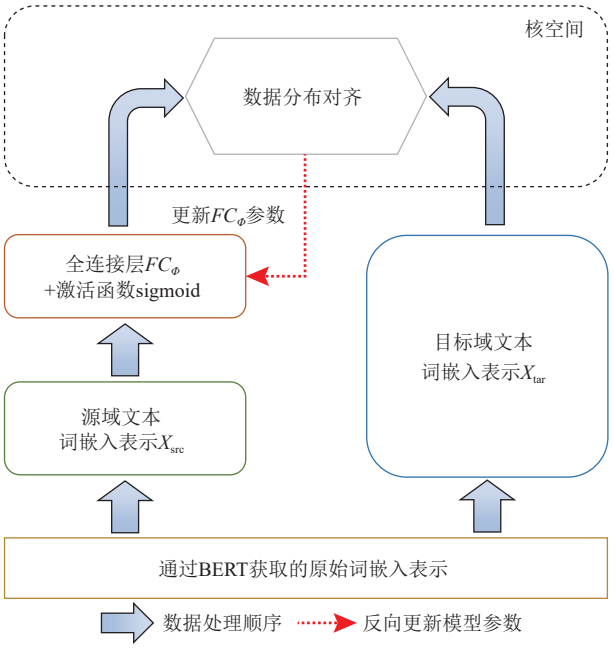


Fig. 4 Data distribution alignment schematic diagram

图 4 数据分布对齐示意图

$$\begin{cases}\boldsymbol{X}'_{\text{src}}=FC_\Phi\left(\boldsymbol{X}_{\text{s}_\text{f}}\right),\\ \boldsymbol{X}'_{\text{tar}}=\dfrac{1}{N-1}\sum_{N-1}FC_\Phi\left(\boldsymbol{X}_{\text{t}_\text{f}}\right).\end{cases}\tag{7}$$

当有多个源域时，如式(7)所示，则目标域的文本表示将为多个源域上的平均表示. 此处，式(7)中 N 表示所有领域的总数量. 综上，针对源域原始文本 $\boldsymbol{X}_{\text{src}}$ 和目标域原始文本 $\boldsymbol{X}_{\text{tar}}$ 的领域数据分布对齐总体过程如算法 1 所示.

算法 1. 领域数据分布对齐过程.

输入：源域原始文本 $\boldsymbol{X}_{\text{src}}$，目标域原始文本 $\boldsymbol{X}_{\text{tar}}$；源域特征表示 $\boldsymbol{X}_{\text{src}}=[\boldsymbol{S}_c;\boldsymbol{S}_s]$，目标域特征表示 $\boldsymbol{X}_{\text{tar}}=[\boldsymbol{T}_c;\boldsymbol{T}_t]$；

输出：源域分布对齐表示 $\boldsymbol{X}'_{\text{src}}$，目标域分布对齐表示 $\boldsymbol{X}'_{\text{tar}}$.

① 通过最小化式(2)的目标函数，获取特征映射函数 Φ_{src} 和 Φ_{tar}；

② 将特征映射 Φ_{src} 和 Φ_{tar} 交叉作用于 $\boldsymbol{T}_c$ 和 $\boldsymbol{S}_c$ 上获取式(3)中的领域适应化特征矩阵和 $\boldsymbol{T}_a$；

③ 进行式(4)中的特征填充操作，获取源域和目标域填充对齐后的特征矩阵$\boldsymbol{X}_{\text{s}_\text{f}}$和$\boldsymbol{X}_{\text{t}_\text{f}}$；

④ 通过最小化式(6)中的最大均值差异 *Dist* 来减小分布差异，获取源域全连接层映射 FC_Φ；

⑤ 将③中得到的$\boldsymbol{X}_{\text{s}_\text{f}}$输入式(7)中源域全连接层映射 FC_Φ，获取对齐后的源域分布对齐表示 $\boldsymbol{X}'_{\text{src}}$；

⑥ 将③中得到的$\boldsymbol{X}_{\text{t}_\text{f}}$输入式(7)中源域全连接层映射 FC_Φ，获取对齐后的目标域分布对齐表示 $\boldsymbol{X}'_{\text{tar}}$. 如果有多个源域则取平均表示.

2.4 中间过渡域重划分

为加强源域和目标域之间的可迁移性，提高迁移过程中领域数据的相关性，从而为目标域原始文本寻找更为适配的源域摘要文本作为生成参考，本文进一步将源域和目标域中所有数据根据文本相似性综合指标归纳成簇，重新划分至 K 个中间过渡域中，从而在中间域中，为目标域数据分配更为合适的源域数据，即更为恰当的领域数据选择，如图 5 所示.

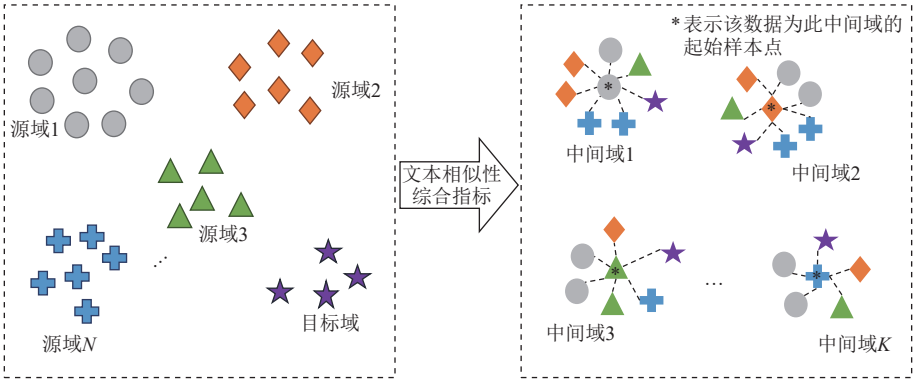


Fig. 5 Intermediate domain redistribution schematic diagram

图 5 中间域重划分示意图

具体地，每个重划分的中间域内包含了最具有相似性的源域和目标域数据. 由于不同领域数据之间具有语义差异，不恰当的中间域划分会导致其所包含的源域和目标域数据之间产生负迁移问题[3]. 因此，各中间域内的数据应拥有尽可能多的相似特征.

首先，由式(7)得到各源域和目标域的分布对齐表示 $\boldsymbol{X}'_{\text{src}}$ 和 $\boldsymbol{X}'_{\text{tar}}$ 之后，对每个源域中所有数据的分布对齐表示取平均，得到各源域内的平均分布对齐表示向量. 接着，将各源域内与平均分布对齐表示向量距离最相近的数据点作为各中间域的起始点，由此得到源域个数 N−1 个中间域起始点. 最后，本文研究并选择了 4 个相似性计算指标，从文本内容相似性角度进行中间域重划分：

1)特定词重合度 S_{overlap}. 计算给定文本对的相似度，即文本中特定用词的重合度越高，表示文本传达的主要信息越相似. 使用余弦相似度来量化这一指标，如式(8)所示：

$$S_{\text{overlap}}=\frac{\sum_{i=1}^{n}(\boldsymbol{x}_i\times\boldsymbol{y}_i)}{\sqrt{\sum_{i=1}^{n}(\boldsymbol{x}_i)^2}\times\sqrt{\sum_{i=1}^{n}(\boldsymbol{y}_i)^2}},\tag{8}$$

其中 $\boldsymbol{x}_i$ 和 $\boldsymbol{y}_i$ 表示源域文本和目标域文本经过 OneHot 编码后，词频向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 在同位 i 上的值，即每个分词出现的次数.

2)用词覆盖率 S_{coverage}. 将给定文本对中重合词的数量除以目标域文本中的词数量，即文本中相同用词越多表明源域文本与目标域文本越相似. 根据召回率(recall)来衡量源域文本和目标域文本在单个词语上的共现性，如式(9)所示：

$$S_{\text{coverage}}=\frac{\sum_{S\in\{reference\}}\sum_{gram_1\in S}Count_{\text{match}}(gram_1)}{\sum_{S\in\{reference\}}\sum_{gram_1\in S}Count(gram_1)},\tag{9}$$

其中 $gram_1$ 表示共现词的词粒度为 1，式(9)中分子部分表示源域文本与目标域文本中同时出现 $gram_1$ 的个数，式(9)中分母部分表示目标域文本中出现的 $gram_1$ 个数.

3)信息密度 S_{density}. 将给定文本对中的重合词数量除以源域文本中的词数量，即高信息密度表明源域文本中有大量可迁移至目标域的信息. 根据信息密度(density)来衡量源域文本和目标域文本在词语上的重复度，如式(10)所示：

$$S_{\text{density}}=\frac{\sum_{gram_1\in C}Count_{\text{clip}}(gram_1)}{\sum_{gram'_1\in C'}Count(gram'_1)},\tag{10}$$

其中 $gram_1$ 表示共现词的词粒度为 1，式(10)分子部分表示源域文本与目标域文本中同时出现的 $gram_1$ 个数，式(10)分母部分表示源域文本中出现的 $gram_1$ 个数.

4)文本长度 S_{length}. 文本长度可以反映出所包含信息量的多少，即拥有相似长度的文本对所包含的信息量大致相同. 使用源域文本和目标域文本标记长度绝对差值与文本标记长度和比值的负值来量化这一指标，如式(11)所示：

$$S_{\text{length}} = -\frac{\left|S_{\text{tar\_len}} - S_{\text{src\_len}}\right|}{S_{\text{tar\_len}} + S_{\text{src\_len}}}, \tag{11}$$

其中 $S_{\text{tar\_len}}$ 表示目标域文本经过分词后得到的词序列中的词数量，$S_{\text{src\_len}}$ 表示源域文本经过分词后得到的词序列中的词数量.

最终如式(12)所示，将特定词重合度 S_{overlap}、用词覆盖率 S_{coverage}、信息密度 S_{density} 和文本长度 S_{length} 相加，得到用于计算源域文本和目标域文本内容相似性的综合指标 S:

$$S = S_{\text{overlap}} + S_{\text{coverage}} + S_{\text{density}} + S_{\text{length}}. \tag{12}$$

然后，在得到源域个数 N−1 个中间域起始点后，使用聚类方法中常用的轮廓系数(silhouette coefficient)[33]对起始点个数进行评价，从而从 N−1 个中间域起始点中确定最佳的 K 个中间域起始点. 假设已经将源域和目标域数据按照文本内容相似性的综合指标 S 划分为源域数量个中间域，对于每个中间域中的每个样本点 i，分别计算其轮廓系数. 具体地，需要对每个样本点 i 计算 2 个指标：$a(i)$表示样本点 i 到同一中间域中其他样本点距离的平均值；$b(i)$表示样本点 i 到其他中间域 C_j 中所有样本的距离的平均值 $b_{i,j}$，其中 $b(i)=\min\{b_{i1}, b_{i2}, \cdots, b_{ik}\}$. 则样本点 i 的轮廓系数如式(13)所示：

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}. \tag{13}$$

中间域中所有样本点 i 的轮廓系数的平均值，即为该中间域总的轮廓系数 $S\in[-1, 1]$，S 越接近于 1，说明中间域划分效果越好. 接着将每个中间域的轮廓系数进行相加排名，获得轮廓系数总和得分最高的中间域组合，此时组合的中间域个数即为中间域划分最优 K 取值. 最后，将源域和目标域剩余的原始文本分别与 K 个中间域起始点所对应的原始文本，通过式(12)进行内容相似性指标计算，按所得综合相似性指标评分排序，逐个将源域和目标域剩余的原始文本划分到得分排名第 1 的中间域中，由此将所有领域文本划分到各自最相似的中间域中，如图 5 所示，形成 K 个中间域D_i'. 每个中间域均同时包含了最相似的源域和目标域数据，由此在后续利用语义要素传导策略进行迁移时，中间域内的目标域原始文本可按照语义要素的相似性将最为相关的源域摘要文本作为模型训练参考真值. 图 5 基于文本相似性指标的领域文本中间域重划分总体过程如算法 2 所示.

算法 2. 中间域重划分过程.

输入：源域原始文本，源域数量为 N−1，目标域原始文本，目标域数量为 1；

输出：重新划分为 K 个(不超过 N−1 个)中间域 D_i'的新源域原始文本和目标域原始文本.

① 对式(7)获取的源域分布对齐词嵌入表示取平均，获取源域中的平均分布对齐表示；

② 获取源域中与平均分布对齐表示最相近的原始文本作为起始文本，获取 N−1 个中间域起始点新闻文本数据；

③ 根据式(13)的轮廓系数，获得每个起始点为中心的新中间域轮廓系数 s；

④ 根据 N−1 个轮廓系数，得出排名最高的中间域廓系数 s 的得分组合，此时的中间域个数即为最佳 K 取值；

⑤ 将剩余的源域和目标域中的数据分别与 K 个中间域起始新闻文本通过式(12)计算文本相似性综合指标 S，并根据得分进行排序，根据指标得分，将文本划分到得分最高的中间域中；

⑥ 对源域和目标域剩余的原始文本重复⑤操作，直到所有数据被划分到新的 K 个中间域D_i'中.

2.5 基于中间域的语义要素传导

基于图 1 中构建的迁移式文本生成模型、分布对齐后的源域数据表示 $\boldsymbol{X}'_{\text{src}}$ 和目标域数据表示 $\boldsymbol{X}'_{\text{tar}}$，以及图 5 中重新划分的 K 个中间域D_i'中的数据，本文设计了一种基于中间域的语义要素传导方法，训练迁移式的文本生成模型，从而有效解决新领域存在的数据缺失问题.

值得注意的是：1)原始文本 x^d、摘要文本 y^d 和语义注释 a^d(包含关键词序列及关键词情感极性值)均通过 BERT 模型获取其词嵌入表示；2)在构建语义要素 $z=(x^d, y^d, a^d)$，$d\in\{\text{src, tar}\}$时，所有领域数据均已遵循图 5 所示的领域重划分原则被划分至 K 个中间域中，并且原始文本表示 x^d 已按式(7)进行了领域数据分布对齐；3)所构建语义要素 $z=(x^d, y^d, a^d)$，$d\in\{\text{src, tar}\}$将会输入至如图 1 所示的适用于语义要素传导的迁移式文本生成模型中.

具体地，基于式(1)所示的生成过程，针对零次学习语义要素传导，按式(14)为语义要素 z 中的(x^d,y^d)设计损失函数 $Loss_1$，从而使所输入原始文本 x^d 生成的摘要文本$\hat{y}^d$“接近于”x^d 对应的参考摘要文本 y^d，以此推导出原始文本 x^d、真值摘要文本 y^d 和所生成摘要文本$\hat{y}^d$ 三者间的语义转导关系.

$$\begin{cases} Loss_1\left(x^d, y^d\right) = -\log P\left(\hat{y}^d \mid E_1\left(x^d\right)\right) + \\ \qquad MSE\left[E_1\left(x^d\right) \| E_2\left(y^d\right)\right], \\ d\in\{\text{src, tar}\}, \end{cases} \tag{14}$$

具体地，如式(14)所示，$E_1(x^d)$表示将原始文本 x^d 输

入到编码器端的编码器模块 E_1 中; $E_2(y^d)$ 表示将摘要文本 y^d 输入到编码器端的另一个编码器模块 E_2 中. 当最小化式(14)中所定义的损失函数 $Loss_1$ 时, $D[E_1(x^d)\|E_2(y^d)]$ 表示 $E_1(x^d)$ 输出的隐藏状态应该“接近于” $E_2(y^d)$ 输出的隐藏状态, 此处均方误差(mean square error, *MSE*)作为距离度量. 由此, 对于中间域 D_i' 所包含的领域数据而言, 给定语义要素 $z=(x^d, y^d, a^d)$, $d\in\{\text{src, tar}\}$, 通过最小化损失函数 $Loss_1$, 可以在中间域 D_i' 内建立隐式的语义转导关系 $y^d\approx x^d\rightarrow\hat{y}^d\approx y^d$.

类似地, 基于式(1)所示的生成过程, 针对零次学习语义要素传导, 按式(15)为语义要素 z 中的 (a^d, y^d) 设计损失函数 $Loss_2$, 从而使所输入语义注释 a^d 生成的标题 $\hat{y}^d$ “接近于” a^d 对应的真值摘要文本 y^d, 以此推导出语义注释 a^d、摘要文本 y^d 和所生成摘要文本 $\hat{y}^d$ 三者间的语义转导关系.

$$\begin{cases} Loss_2\left(a^d, y^d\right) = -\log P\left(\hat{y}^d \mid E_2\left(a^d\right)\right) + \\ \qquad MSE\left[E_2\left(a^d\right)\|E_2\left(y^d\right)\right], \\ d\in\{\text{src, tar}\}, \end{cases} \tag{15}$$

具体地, 如式(15)所示, 将原始文本 x^d 对应的语义注释 a^d 输入到编码器模块 E_2 后, 仍然令模型生成摘要文本 $\hat{y}^d$. 与此同时, 通过最小化 $MSE[E_2(a^d)\|E_2(y^d)]$, 引导编码器模块 E_2 输出的隐藏状态 $E_2(a^d)$ “接近于” $E_2(y^d)$ 输出的隐藏状态. 最终, 对于中间域 D_i' 所包含的领域数据而言, 给定数据语义要素 $z=(x^d, y^d, a^d)$, $d\in\{\text{src, tar}\}$, 通过最小化损失函数 $Loss_2$, 可以在中间域 D_i' 内建立隐式的语义转导关系 $y^d\approx a^d\rightarrow\hat{y}^d\approx y^d$.

最后, 如式(16)所示, 通过将损失函数 $Loss_1$ 和 $Loss_2$ 相结合, 构建了复合生成损失函数 $Loss_{co}$, 从而间接反映了基于语义要素传导的迁移式文本生成原理, 即当输入语义要素 $z=(x^d, y^d, a^d)$, $d\in\{\text{src, tar}\}$ 时, 图 1 中迁移式文本生成模型的参数将通过式(16)中的复合损失函数 $Loss_{co}$ 进行训练, 从而如图 6 所示, 在中间域 D_i' 内建立语义转导关系 $x^d\approx y^d\approx a^d\rightarrow\hat{y}^d\approx y^d$.

$$Loss_{co} = Loss_1 + Loss_2. \tag{16}$$

因此, 在每个中间域 D_i' 中, 当给定来自新源域的语义要素 $z^{src}=(x^{src}, y^{src}, a^{src})$ 时, 新源域内可建立语义关联 $x^{src}\approx y^{src}\approx a^{src}\rightarrow y^{src}$. 接着, 当给定来自目标域的语义要素 $z^{tar}=(x^{tar}, y^{tar}, a^{tar})$ 时, 目标域内可建立语义关联 $x^{tar}\approx y^{tar}\approx a^{tar}$. 当涉及新源域和目标域之间的语义要素传导时, 如图 6 所示, 如果在一个中间域 D_i' 中, 存在任何一对(原始文本 x, 摘要文本 y)的语义要素 $z^{tar}=(x^{tar}, y^{tar}, a^{tar})$ 与 $z^{src}=(x^{src}, y^{src}, a^{src})$ 接近或相似, 则会产生一个跨域的语义关联 $x^{tar}\approx y^{src}\approx a^{src}\rightarrow y^{src}$, 即为 $x^{tar}\rightarrow y^{src}$, 如图 6 所示.

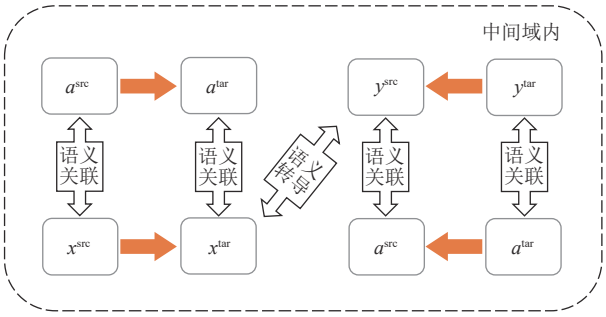


Fig. 6 Text semantic prototype conduction schematic diagram

图 6 文本语义要素传导示意图

因此, 当给定目标域原始文本 x^{tar} 时, 可以参考新源域中相关的真值文本 y^{src} 来辅助生成目标域中的摘要文本 y^{tar}. 由此, 即使目标域中没有真值文本数据, 也可以通过零次学习语义要素传导的方式借助新源域数据帮助目标域中的原始文本生成摘要文本, 整体过程如算法 3 所示.

算法 3. 基于零次学习语义要素传导的迁移式文本生成过程.

输入: 源域语义要素 $z^{src}=(x^{src}, y^{src}, a^{src})$, 目标域语义要素 $z^{tar}=(x^{tar}, y^{tar}, a^{tar})$;

输出: 生成摘要文本 $\hat{y}^d$, $d\in\{\text{src, tar}\}$.

① 在中间域 D_i' 内, 通过式(14)中 $Loss_1$ 训练迁移式文本生成模型, 构建源域内语义关联:

$x^{src}\approx y^{src}\approx a^{src}\rightarrow y^{src}$;

② 在中间域 D_i' 内, 通过式(15)中 $Loss_2$ 训练迁移式文本生成模型, 构建目标域内语义关联:

$x^{tar}\approx y^{src}\approx a^{src}\rightarrow y^{src}$;

③ 在中间域 D_i' 内, 通过式(16)中 $Loss_{co}$ 训练迁移式文本生成模型, 构建跨域语义关联:

$x^{tar}\approx y^{src}\approx a^{src}\rightarrow y^{src}$, 即为 $x^{tar}\rightarrow y^{src}$;

④ 模型通过式(1)生成摘要文本 $\hat{y}^d$, $d\in\{\text{src, tar}\}$. 生成过程中更新迁移式文本生成模型参数.

3 实验及分析

3.1 实验数据与实验设置

在实验中, 针对本文设计的多领域场景下的迁移式文本生成任务, 因为新闻天然地具有多领域、多主题的特点, 所以选择了新闻标题生成任务进行实验. 本文选取了公开数据集 PENS(personalized news headlines)[5] 个性化新闻标题生成数据集. PENS 中包含 113 762 篇新闻, 分为 15 个主题, 每篇新闻包含标

题和正文. 本文从 PENS 数据集中随机选择 8 个新闻主题作为不同领域，包括体育(sports)、金融(finance)、音乐(music)、天气(weather)、汽车(auto)、电影(movie)、健康(health)和儿童(kid). 在每一个领域中，随机选择 8 000 条新闻数据作为训练数据集.

表 1 中描述了实验所使用数据集的相关信息. 其中，“平均长度”和“最大长度”表示每个领域中，所有新闻正文和新闻标题通过预训练 BERT 模型进行分词后，所得词序列的最大长度与平均长度. “压缩率”表示一个领域中新闻标题的文本平均长度与新闻正文文本平均长度的比率.

Table 1 Statistical Information on the News Data Extracted from PENS Dataset

表 1 PENS 数据集中提取的新闻数据的统计信息

序号	主题	新闻正文		新闻标题		压缩率/%
		平均长度数目	最大长度数目	平均长度数目	最大长度数目	
1	体育	480.5	537	12.9	17	2.22
2	金融	482.7	588	8.9	19	1.84
3	音乐	528.0	557	10.5	18	1.99
4	天气	484.9	566	12.3	19	2.53
5	汽车	511.7	580	9.1	15	1.77
6	电影	483.1	636	10.1	19	2.09
7	健康	483.8	544	9.0	15	1.53
8	儿童	509.4	560	13.3	16	2.61

在实验中，图 6 中迁移式文本生成模型编码器模块和解码器模块的子层数量均为 4，子层的输入输出维度为 512，多头注意力的注意力头数量为 8；用于获取词嵌入表示的预训练 BERT 模型采用维度大小为 512 的 BERT-Medium；Bi-LSTM 的隐藏单元数量为 512；模型训练采用带有自定义学习率的 Adam 优化器[11]；在每个领域上训练的迭代次数(epochs)为 1 000；本文所有实验均采用 Python 3.8 和 tensorflow-gpu 2.5.0 实现，实验平台配置为 Windows 10 操作系统，GPU 为 NVIDIA 2080Ti 显卡，内存为 32GB RAM，CPU 为 Intel Core i7-11700K 处理器.

3.2 评价指标及基准模型

为了评估本文提出的迁移式文本生成模型应用到新闻标题生成任务时的有效性，将本文提出的迁移式文本生成模型与现有性能表现出众的预训练语言模型和零样本数据或小样本数据学习相关的文本生成模型进行比较.

本实验选择 T5[17]，BART[18]，PEGASUS[19]，BertSum[34] 预训练语言模型. 这 4 个预训练语言模型均使用预训练参数作为模型的初始参数，在不改变其他超参数情况下，使用表 1 中的数据对这 4 个模型在预训练初始参数的基础上继续进行训练.

对于零样本数据或小样本数据文本生成模型，选择 ZSDG[28]，TransferRL[35]，DAML[36]，MTL-ABS[37]. 其中，ZSDG 通过将“种子级别”的数据描述投射到一个子空间中，再在领域层面上进行语义描述迁移，从而使用零次学习方法通过领域描述进行目标域零数据的迁移式文本生成. TransferRL 包含一个在不同领域之间共享的解码器，并通过强化学习自我批评(self-critic)策略最大化解码器泛化至不同领域的“奖励”，提升模型的领域适应性，从而只需要在小批量数据上进行微调便可快速适应至目标领域. DAML 和 MTL-ABS 均根据元学习(meta-learning)原理，使用序列到序列的形式构建生成模型，但 DAML 使用门控循环神经网络作为编码器和解码器，而 MTL-ABS 以 Transformer 作为编码器和解码器. DAML 和 MTL-ABS 通过元学习方式从梯度优化层面，为模型搜索最具潜力的参数取值，使模型对目标域少样本数据反应更加灵敏，提升模型的领域泛化性. 与预训练语言模型相比，零样本数据或少样本数据学习模型都直接使用表 1 中的数据，并根据各自的迁移策略对模型进行训练.

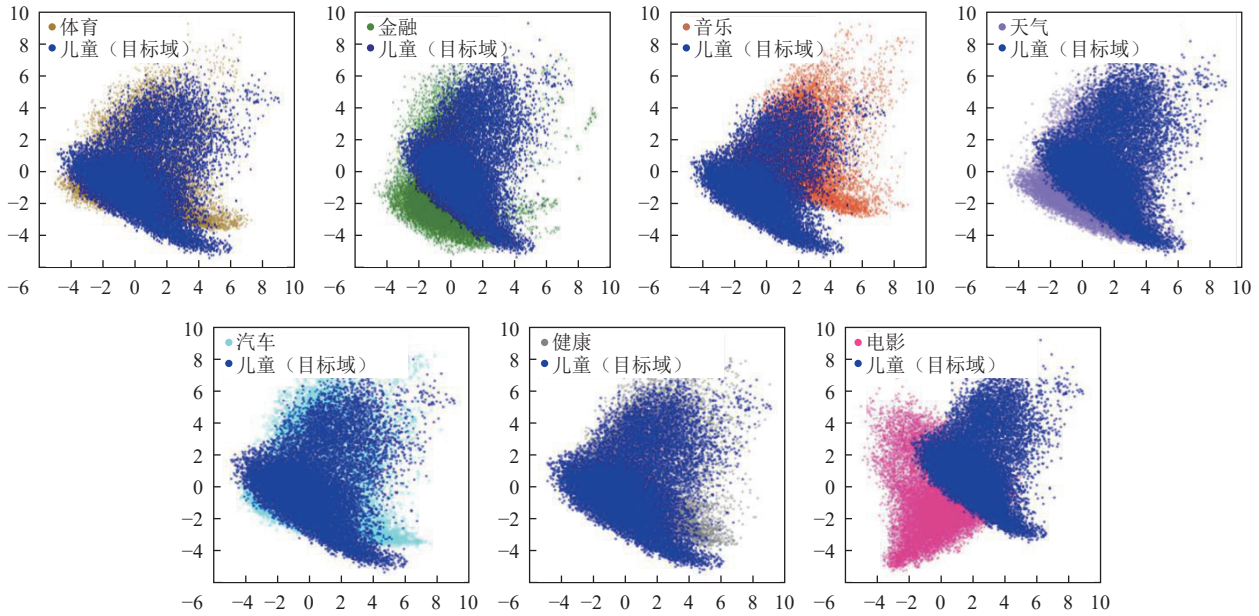
本文对比模型的生成效果采用文本生成任务中常用的评价指标 ROUGE-1/2/L[38]，BLEU[38]，METEOR[38] 来评估. 将目标域中的新闻正文输入至训练后的模型中，计算模型生成的新闻标题与相应的真值新闻标题之间的评价指标得分. 其中，目标域中的真值新闻标题仅用于评估而不参与模型训练过程. 基于上述指标得分，考察本文提出的迁移式文本生成模型能否有效地从源域数据中获取相关的可借鉴知识，从而在不给定目标域文本参考真值的前提下，有效辅助目标域完成文本生成任务.

3.3 实验结果与分析

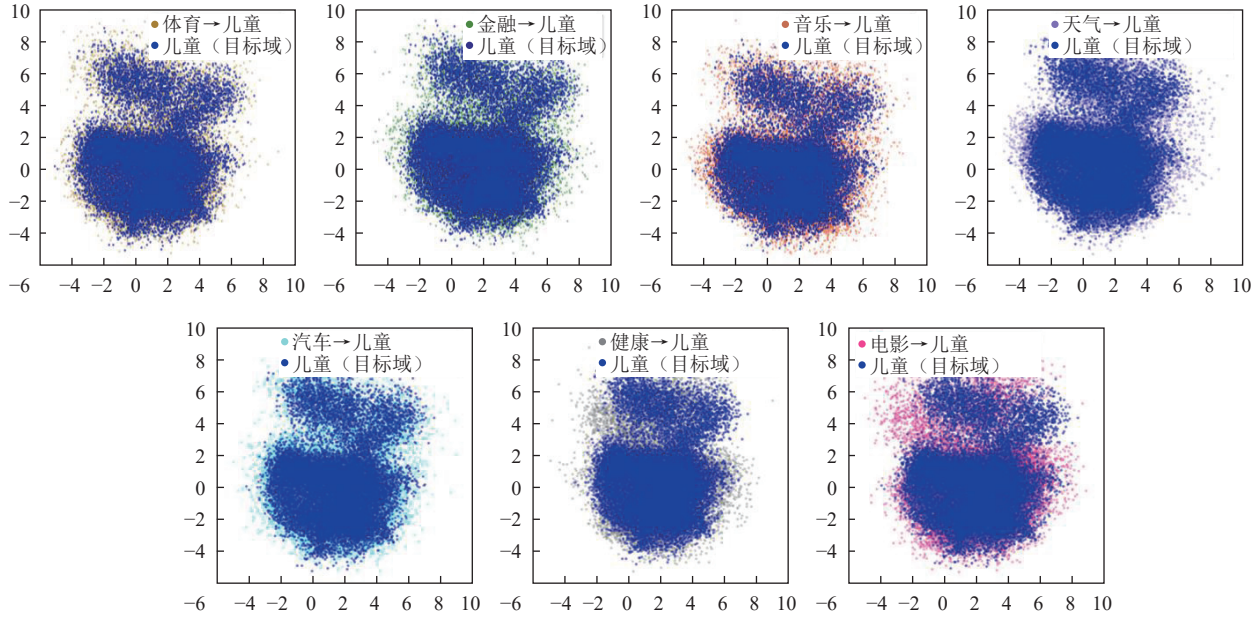
3.3.1 数据分布对齐效果

为了更直接展示本文所提出迁移式文本生成模型各阶段内部机制实际效果，如图 7 所示，以“儿童”新闻主题作为目标域，进一步展示领域数据分布对齐效果. 其中源域与目标域数据按式(7)进行映射训练. 图 7(a)中源域和目标域的原始词嵌入表示 $\boldsymbol{X}_{\text{src}}$ 与 $\boldsymbol{X}_{\text{tar}}$，以及图 7(b)中通过式(7)获得的对齐后表示 $\boldsymbol{X}'_{\text{src}}$ 和 $\boldsymbol{X}'_{\text{tar}}$ 均采用主成分分析(principal component analysis, PCA)方法进行降维表示.

具体地，在图 7 中，不同领域的数据表示采用不



(a) 数据分布对齐前的词嵌入表示分布图



(b) 数据分布对齐后的词嵌入表示分布图

Fig. 7 Visualization of the alignment effect of the data distribution after dimensionality reduction

图 7 降维后的数据分布对齐效果可视化

同颜色进行显示，位于上层的深蓝色区域表示“儿童”新闻主题作为目标域时，领域中数据的词嵌入分布表示.

图 7(a)中展示了 8 个领域的文本数据通过预训练 BERT 模型输出的原始表示分布，此时的原始表示分布没有经过任何交叉特征填充和数据分布对齐处理. 可以发现，所给定的 8 个领域的原始表示分布存在明显差异. 其次，如图 7(b)所示，将除了“儿童”以外的其他 7 个领域作为源域. 源域中的数据与目标域“儿童”领域新闻数据首先按式(2)~(6)进行源域和目标域之间的交叉特征填充；在此基础上，按图 4 所示过程由式(7)做领域数据分布对齐处理，最终结果如图 7(b)所示. 可以发现，经领域数据分布对齐后，源域和目标域数据之间虽然仍有轻微差异，但不同领域间数据的分布差异已明显缩小. 将对齐前的图 7(a)和对齐后的图 7(b)进行对比可以发现，本文所提出模型涉及的领域数据分布对齐在不同领域间先采用交叉填充为源域和目标域数据填充特征，再用最小化源域与目标域间的最大均值差异距离度量，有效降低了源域和目标域之间的数据分布差异.

3.3.2 目标域轮循实验

针对零次学习语义要素传导，依次将表 1 列出的 8 个域中的 1 个域选作目标域，其余的 7 个域作为源域. 根据中间域重划分方法将 7 个源域和 1 个目标域组成如图 5 所示的 K 个中间域进行实验. 在目标域轮循过程中，通过式(13)，即 K-聚类(K-means)方法中常用的轮廓系数(silhouette coefficient)[33]来评价不同 K 取值下的中间域划分效果，从而确定 K 的取值，此时 K 的取值不超过源域数量 7. 轮廓系数的取值范围为 [−1,1]，若轮廓系数的值越趋近于 1，代表内聚度和分离度相对较优，聚类效果较好，由此确定中间域个数 K.

图 8 表示通过算法 2 确定在每个领域作为目标域时，不同的 K 值取值下轮廓系数的大小. 取轮廓系数最大的 K 值点作为该领域下的中间域最佳个数 K. 在得到每个领域作为目标域时的最佳中间域个数 K 的取值后，表 2 中 ROUGE-1/2/L, BLEU, METEOR 指标得分是轮循实验中每次确定目标域后，在相应的中间域划分方案下，由模型生成的新闻标题和相应的标题参考真值计算得出的. 具体地，首先评估每个目标域中的文本生成效果. 在这种情况下，只有源域的真值新闻标题文本数据参与了模型训练，目标域中没有标题真值数据参与，目标域仅使用从新闻正文抽取的伪新闻标题文本. 由此，基于式(7)获得的领域数据分布对齐表示和按式(14)(15)进行的零次学习语义要素传导，每个目标域中的新闻正文可以不依赖于任何人工标注的参考真值，直接生成新闻标题.

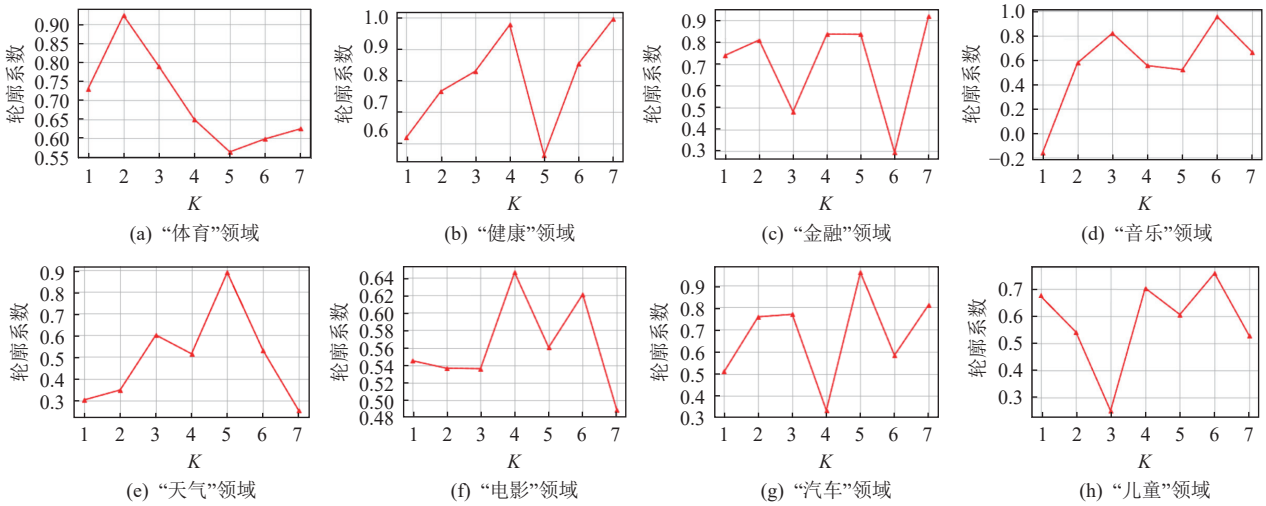


Fig. 8 The silhouette coefficients corresponding to different K values in different fields

图 8 不同领域中不同 K 值对应的轮廓系数

Table 2 Different Evaluating Indicator Scores in Different Target Domains

表 2 不同目标域中各项评价指标的得分

编号	目标域	最优中间域数 K	评价指标				
			ROUGE-1/%	ROUGE-2/%	ROUGE-L/%	BLEU/%	METEOR/%
1	体育	6	74.52	59.33	73.83	43.51	68.84
2	健康	7	79.03	64.38	78.25	49.95	73.34
3	金融	7	76.36	60.99	75.82	48.36	70.02
4	音乐	6	72.54	61.53	72.24	47.02	67.76
5	天气	5	77.60	62.00	76.92	45.45	70.80
6	电影	2	63.65	41.70	62.19	24.51	55.04
7	汽车	5	78.20	64.05	77.68	49.59	73.04
8	儿童	4	75.15	59.44	74.28	45.28	69.48

表 2 列出了本文提出的适用于语义要素传导的迁移式文本生成模型在不同目标域中的新闻标题生成性能. 可以看出，除了“电影”领域外，其余各领域的指标表现相对稳定；“健康”“汽车”“天气”领域的

指标表现综合来看排在前 3 位. 由此, 虽然模型在生成训练过程中没有参考目标域中的标题真值数据, 但通过图 4 中根据式(7)所采用的领域数据分布对齐和图 6 中基于(新闻 x, 标题 y)进行的语义要素传导迁移, 获取到不同领域之间的数据语义关联性, 从而在不同目标域轮循的过程中和各评价指标上都能获得较好的得分. 该现象可以归因于: 首先基于图 4 在领域数据分布对齐后, 数据在不同领域间的分布差异被缩小, 因此可以在模型从源域迁移至目标域的过程中, 减少不同领域数据分布差异所带来的负面影响; 接着通过零次学习语义要素传导, 本文提出的迁移式文本生成模型通过图 2 中增强型编码器与解码器中的注意力机制与时序依赖性来同时获取不同领域数据之间的语义关联性, 从而调整模型参数以提高模型领域迁移效果.

更进一步, 图 9 展示了全部领域作为目标域时在零次学习语义要素传导阶段, 文本生成模型的训练表现. 在该阶段中, 模型通过式(16)定义的损失函数 $Loss_{co}$ 经过 1 000 次迭代进行训练. 词汇准确率是计算生成文本在每个时间步上生成的文本与参考真值文本之间相同词汇的比率. 从图 9 可以看出, 即使是文本生成评价指标最低的 3 个领域, 训练中的损失函数 $Loss_{co}$ 也在逐渐减小, 证明了模型在目标域无参考真值情况下, 能够通过为语义要素 z 中 (x^d, y^d) 设计的损失函数 $Loss_1$ 和 (a^d, y^d) 设计的损失函数 $Loss_2$, 使得编码器和解码器按零次学习语义要素传导方法充分解析各领域数据的语义要素, 使模型在生成过程中捕捉到不同领域数据语义要素间的关联性, 从而进行从源域至目标域的有效迁移; 而词汇准确率的平稳上升, 证明了本文提出的迁移式文本生成模型在从源域迁移至目标域后所生成文本的准确性, 其中指针生成器网络负责处理未登录词问题, 进一步提升了文本质量.

3.3.3 消融性实验

从表 2 可以看出, 当“健康”“汽车”“天气”这 3 个领域作为目标域时, 迁移式文本生成性能最佳. 因

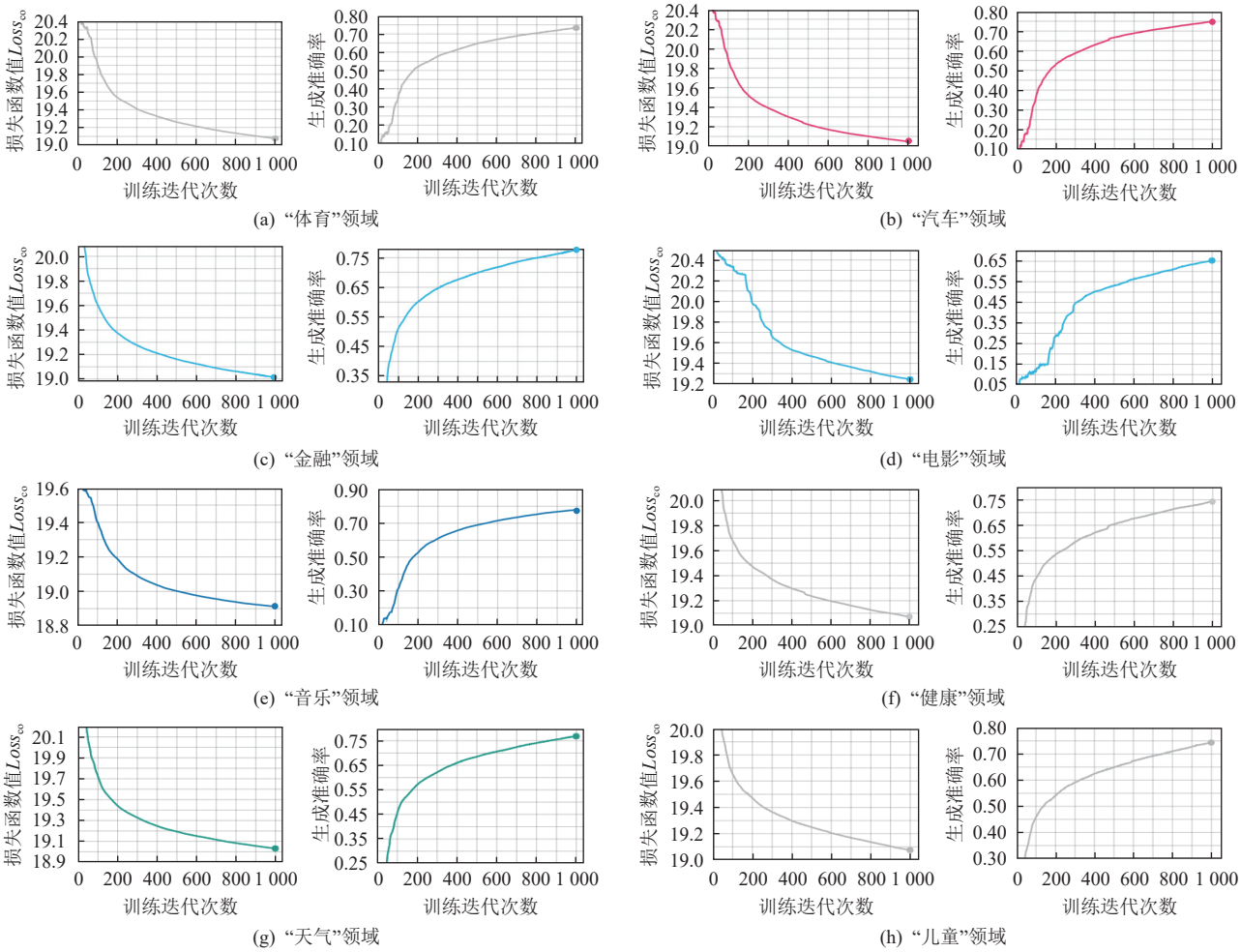


Fig. 9 Loss function curves and word accuracy curves in different target domains

图 9 不同目标域中的损失函数曲线与词汇准确率曲线

此，使用这 3 个域进一步对本文提出的迁移式文本生成方法进行消融实验，结果如表 3 所示.

表 3 中，“语义转导”表示直接采用预训练 BERT 模型输出的原始词嵌入表示，不进行中间域划分，直接使用图 6 中基于式(14)~(16) 的语义要素传导进行模型训练；“中间域划分+语义传导”表示直接采用预训练 BERT 模型输出的原始表示，按最佳中间域个数 K 取值进行中间域划分后，再使用图 6 中基于式(14)~(16) 的语义要素传导进行模型训练；“分布对齐+中间域划分+语义转导”表示基于图 4 中按式(7)采用分布对齐后的数据表示，按最佳中间域个数 K 取值进行中间域划分后，再进行图 6 中基于式(14)~(16) 的语义要素传导训练.

从表 3 可以看出，在每个目标域中采用了分布表示对齐方法后，其文本生成效果要优于直接使用原始表示的方法，这意味着通过领域数据分布对齐可以有效消除领域间的数据分布差异，提升从源域向目标域的可迁移性. 此外，将表 3 与表 4 对比可以看出，本文提出的模型仅使用语义要素传导方法进行训练，与多数其他的迁移式文本生成模型相比，也可以获得更高的评价指标得分. 该现象表明了在本文提出的迁移方案中，零次学习语义要素传导在不同领域间探索数据语义关联性，通过“编码器-解码器”结构中增强型编码器与解码器使目标领域中的无标注新闻正文与源领域中最相关的新闻标题进行关联，根据注意力机制与时序依赖性获得语义要素上的相似性或接近性，得出目标域在文本生成时对源域数据的参考，从而提升了迁移的文本生成效果.

Table 3 Results of Ablation Experiments

表 3 消融性实验结果 %

目标域	消融组合	最优中间域数 K	评价指标				
			ROUGE-1/%	ROUGE-2/%	ROUGE-L/%	BLEU/%	METEOR/%
健康	语义转导	0	72.36	59.85	71.34	42.85	61.47
	中间域划分+语义转导	7	77.38	63.12	76.43	46.62	68.67
	分布对齐+中间域划分+语义转导	7	79.03	64.38	78.25	49.95	73.34
汽车	语义转导	0	57.19	55.72	56.53	36.90	46.45
	中间域划分+语义转导	5	68.98	60.78	68.36	43.31	60.63
	分布对齐+中间域划分+语义转导	5	78.20	64.05	77.68	49.59	73.04
天气	语义转导	0	57.04	48.43	55.71	39.72	39.52
	中间域划分+语义转导	5	68.58	55.87	67.53	42.62	55.82
	分布对齐+中间域划分+语义转导	5	77.60	62.00	76.92	45.45	70.80

Table 4 Comparison of Experimental Results

表 4 实验结果对比

模型分类	模型/方法	评价指标				
		ROUGE-1/%	ROUGE-2/%	ROUGE-L/%	BLEU/%	METEOR/%
预训练语言模型	T5[17]	45.34	30.25	43.27	23.63	41.06
	BART[18]	35.73	20.06	33.78	18.81	33.91
	PEGASUS[19]	35.23	29.62	33.29	18.48	33.49
	BertSum[34]	39.57	20.70	27.29	14.51	38.83
零样本数据/小样本数据学习模型	MTL-ABS[37]	39.56	34.74	38.68	19.16	32.55
	TransferRL[35]	32.79	31.32	28.10	14.39	26.40
	ZSDG[28]	30.89	33.03	23.91	11.67	30.98
	DAML[36]	34.90	32.09	22.14	13.86	27.33
迁移式文本生成模型（本文模型）	语义转导	56.98	37.17	55.28	17.41	43.17
	中间域划分+语义转导	62.00	40.44	60.37	11.18	50.37
	分布对齐+中间域划分+语义转导	63.65	41.70	62.19	24.51	55.04
性能提升率		18.31	11.45	18.92	0.88	13.99

另外, 从图 10 可以看出, 采用了“中间域划分+语义转导”组合的方法相比仅采用“语义转导”的方法获得了更高的评价指标得分, 说明了在通过内容相似性综合指标划分的中间域中, 目标域文本在生成过程中根据更具有语义相似性的相关源域数据, 实现了更好的迁移式文本生成性能. 同时, 完整采用表 3 中的“分布对齐+中间域划分+语义转导”的方法能够取得模型最优的文本生成效果, 意味着模型在获得式(7)的领域数据分布对齐表示和通过式(16)进行零次学习语义要素传导的复合迁移策略时, 能在目标域没有参考真值数据的情况下, 在中间域中从相关源域中获取有帮助的信息, 从而在目标域上带来最优的迁移式文本生成性能, 同时指针生成器网络也会提升生成文本的准确性.

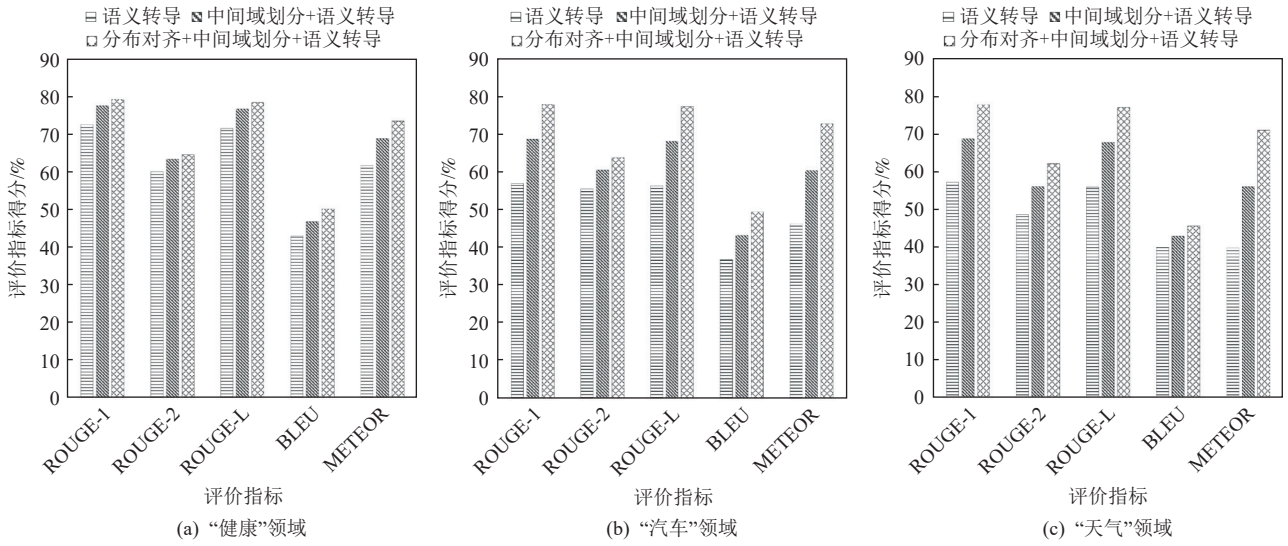
3.3.4 对比实验

如表 2 所示, “电影”域作为目标域时模型的文本生成性能最差, 因此针对“电影”领域, 从预训练语言模型(即 T5, BART, PEGASUS, BertSum)和“零数据/小数据学习模型”(即 TransferRL, ZSDG, DAML, MTL-ABS)2 方面, 进一步比较本文提出的适用于零次学习语义要素传导的文本生成模型方法与其他迁移式文本生成模型方法之间的性能, 结果如表 4 所示.

在经过领域数据分布对齐后, 表 4 中所有模型均采用图 5 所示的中间域数据进行训练, 且所有模型在训练过程中都未使用目标域中的真值数据. 其中, 性能提升率是指本文提出的“分布对齐+中间域划分+语义转导”方法在各项性能评价指标得分上相较于对比模型中最高得分的提升差值.

具体地, 如图 11 所示, 在本文方法效果最差的“电影”领域作为目标域的情况下, 首先, 根据各项评价指标得分, 本文提出的迁移式文本生成模型在对比中取得了最佳性能表现, 其次是预训练语言模型的方法, 最后是零样本数据/小样本数据学习模型的方法. 该现象可归因于本文提出的迁移式方案首先基于图 4 按式(7)在文本表示层面通过领域数据分布对齐, 缓解了领域间的数据分布差异, 然后基于图 1 通过改进文本生成模型结构, 使其更加适用于式(16)进行的零次学习语义要素传导, 从而模型可以更为有效地从相关源域中获取有助于迁移的先验知识, 提高模型在目标域中的文本生成性能.

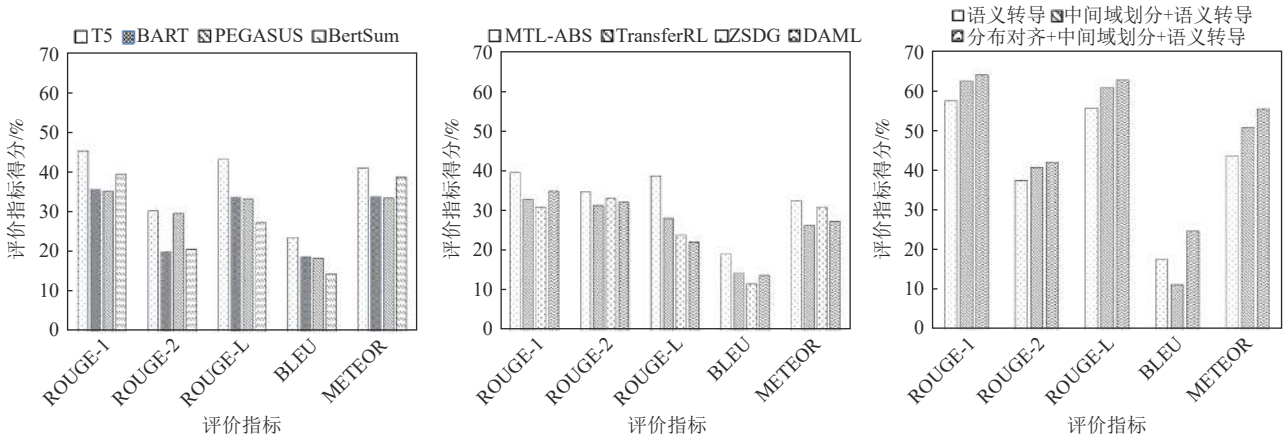
表 4 中的预训练语言模型 T5, BART, PEGASUS, BertSum 已经在大规模语料库中进行了预训练, 因此更多的先验知识已经提前被纳入此类预训练语言模型的参数中. 但是通过表 4 可以看出, T5, BART, PEGASUS, BertSum 的各项评价指标得分均低于迁移式方法. 由此可以发现, 迁移式文本生成模型在领域可迁移性方面优于通过大规模语料训练的预训练语言模型, 此现象可归因为虽然预训练语言模型通过大规模语料库预训练已经获得了大量的领域先验知识, 但这些知识并不针对特定的目标领域及其任务. 相比之下, 迁移式文本生成模型首先通过领域数据分布对齐, 从目标域角度降低了与其他相关源域数据在数据表示上的分布差异, 并通过零次学习语义要素传导, 根据语义要素 $z^{\mathrm{src}}=(x^{\mathrm{src}}, y^{\mathrm{src}}, a^{\mathrm{src}})$ 与 $z^{\mathrm{tar}}=(x^{\mathrm{tar}}, y^{\mathrm{tar}}, a^{\mathrm{tar}})$, 建立跨域语义关联 $x^{\mathrm{tar}}\rightarrow y^{\mathrm{src}}$, 最大程度挖掘了不同领域数据间的语义相关性, 确保目标域即



(a) “健康”领域 (b) “汽车”领域 (c) “天气”领域

Fig. 10 Comparison results of the ablation experiments

图 10 消融性实验对比结果



(a) 预训练语言模型 (b) 零样本数据/小样本数据学习模型 (c) 迁移式文本生成模型

Fig. 11 Results of comparative experiments

图 11 对比实验结果

使没有参考真值数据，也可以通过语义要素传导的方式，借助源域数据帮助目标域生成文本，从而针对特定的目标领域及其下任务有更好的领域迁移适应性.

最后，对于表 4 中的零样本数据/小样本数据学习模型 TransferRL, ZSDG, DAML/MTL-ABS 而言，这些模型分别采用了强化学习、零次学习或元学习方法进行迁移. 但从图 11 可以看到，这些方法的各项评价指标得分均低于迁移式文本生成模型. 该现象可归因于本文在图 1 中对迁移式文本生成模型所采取的结构改进. 具体地，如图 2 所示，改进后的文本生成模型通过加入 Bi-LSTM 层解析文本序列化依赖关系，同时由 Transformer 多头注意力机加大对文本内部上下文观察，借助指针生成器网络处理未登录词汇，故模型可更大程度挖掘文本蕴含的语义；在此基础上，通过构建数据级语义要素，将目标域中无标注新闻正文与源域中最相关的新闻标题进行关联，并根据语义要素上的近似捕捉跨域文本的语义关联性；由此，当给定目标域新闻正文 x^{tar} 时，将参考源域中最为相关的真值新闻标题 y^{src} 以辅助生成目标域中的新闻标题 y^{tar}，因而在 ROUGE-1/2/L, BLEU, METEOR 这些评价指标上也就有了更高的得分表现.

4 总结与展望

本文针对多领域的文本生成任务，提出了基于领域数据分布对齐和零次学习语义要素传导的跨域迁移式文本生成模型，其主要原理是借助相关源域的已标注数据辅助目标域进行文本生成，以克服目标域中参考真值数据缺失的问题. 本文提出的方法在传统文本生成模型的基础上主要改进了 5 个方面：

1)从原始文本、摘要文本和正文语义注释 3 个方面，构建数据级语义要素；

2)在适用于语义要素传导的生成模型结构上，构建增强型“编码器-解码器”，通过为不同语义要素构建的损失函数，从而使模型在生成过程中捕捉不同领域数据语义要素间的关联性，同时在文本生成过程中通过指针生成器网络提高生成文本的准确度；

3)在文本数据表示上，通过特征填充与分布对齐使数据在表示层面减少分布差异性；

4)通过文本相似性综合指标将源域和目标域数据划分为中间域，从而为目标域数据进行更为合适的源域数据选择；

5)在基于语义要素的语义转导方法上，由语义要素之间的相似性使目标域数据在文本生成过程中参考最具关联性的源域已标注数据，由此不依赖目标域自身的已标注真值.

实验结果表明，本文提出的迁移式方法可以有效地应用于实际的新闻标题生成场景中，通过领域数据迁移解决目标域真值数据缺失问题.

未来工作有 2 个方面值得进一步探讨：1)当给定一个目标域时，相关源域的选择对最终迁移式生成性能来说非常关键. 因此，需要进一步研究更具有关联性的领域数据选择方法. 2)源域数据在迁移过程中往往也会提供与目标域不相关的噪声信息，从而影响迁移效果导致“负迁移”. 因此如何避免“负迁移”问题，也是值得进一步研究的方向.

作者贡献声明：马廷淮提出指导意见并修改论文；于信负责完成实验，并撰写、修改论文；荣欢提出实验方案设计和写作思路.

参 考 文 献

[1] Li Jinpeng, Zhang Chuang, Chen Xiaojun, et al. Survey on automatic text summarization[J]. Journal of Computer Research and Development, 2021, 58(1): 1-21 (in Chinese)
(李金鹏, 张闯, 陈小军, 等. 自动文本摘要研究综述[J]. 计算机研究与发展, 2021, 58(1): 1-21)

[2] Zhuang Fuzhen, Luo Ping, He Qing, et al. Survey on transfer learning[J]. Journal of Software, 2015, 26(1): 26-39 (in Chinese)
(庄福振, 罗平, 何清, 等. 迁移学习研究进展[J]. 软件学报, 2015, 26(1): 26-39)

[3] Choi H, Kim J, Joe S, et al. Analyzing zero-shot cross-lingual transfer in supervised NLP tasks [C] //Proc of the 25th Int Conf on Pattern Recognition. Piscataway , NJ : IEEE, 2021: 9608-9613

[4] Wang Wei, Zheng Wenchen, Yu Han, et al. A survey of zero-shot learning: Settings, methods, and applications[J]. ACM Transactions on Intelligent Systems and Technology, 2019, 10(2): 1-37

[5] Ao Xiang, Wang Xiting, Luo Ling, et al. PENS: A dataset and generic framework for personalized news headline generation [C] //Proc of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int Joint Conf on Natural Language Processing (Volume 1: Long Papers). Stroudsburg, PA: ACL, 2021: 82-92

[6] Bae S, Kim T, Kim J, et al. Summary level training of sentence rewriting for abstractive summarization [J]. arXiv preprint, arXiv: 1909.08752, 2019

[7] Rush A M, Chopra S, Weston J. A neural attention model for abstractive sentence summarization [C] //Proc of the 20th Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2015: 379-389

[8] Wu Renshou, Wang Hongling, Wang Zhongqing, et al. Short text summary generation with global self-matching mechanism[J]. Journal of Software, 2019, 30(9): 2705-2717 (in Chinese)
(吴仁守, 王红玲, 王中卿, 等. 全局自匹配机制的短文本摘要生成方法[J]. 软件学报, 2019, 30(9): 2705-2717)

[9] Narayan S, Maynez J, Adamek J, et al. Stepwise extractive summarization and planning with structured transformers[C] //Proc of the 25th Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2020: 4143-4159

[10] See A, Liu P J, Manning C D. Get to the point: Summarization with pointer-generator networks [C] //Proc of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg, PA: ACL, 2017: 1073-1083

[11] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C] // Proc of the 31st Conf on Neural Information Processing Systems. Cambridge, MA: MIT, 2017: 5998-6008

[12] Lao Nanxin, Wang Banghai. Hybrid word character model for Chinese summarization based on BERT[J]. Computer Applications and Software, 2022, 39(6): 258-269 (in Chinese)
(劳南新, 王帮海. 基于BERT的混合字词特征中文文本摘要模型[J]. 计算机应用与软件, 2022, 39(6): 258-269)

[13] Golovanov S, Kurbanov R, Nikolenko S, et al. Large-scale transfer learning for natural language generation[C] //Proc of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2019: 6053–6058

[14] Huang Jiajia, Li Pengwei, Peng Min, et al. Research on deep learning-based topic models[J]. Chinese Journal of Computers, 2020, 43(5): 827-855 (in Chinese)
(黄佳佳, 李鹏伟, 彭敏, 等. 基于深度学习的主题模型研究[J]. 计算机学报, 2020, 43(5): 827-855)

[15] Dethlefs N. Domain transfer for deep natural language generation from abstract meaning representations[J]. IEEE Computational Intelligence Magazine, 2017, 12(3): 18-28

[16] Qiu Xipeng, Sun Tianxiang, Xu Yige, et al. Pre-trained models for natural language processing: A survey[J]. Science China Technological Sciences, 2020, 63(10): 1872-1897

[17] Raffel C, Shazeer N, Roberts A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer[J]. Journal of Machine Learning Research, 2020, 21(140): 1-67

[18] Lewis M, Liu Yinhan, Goyal N, et al. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension[C] //Proc of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2020: 7871-7880

[19] Zhang Jingqiang, Zhao Yao, Saleh M, et al. PEGASUS: Pre-training with extracted gap-sentences for abstractive summarization[C] //Proc of the 37th Int Conf on Machine Learning. Cambridge, MA: PMLR, 2020: 11328-11339

[20] Wang Chang, Mahadevan S. Heterogeneous domain adaptation using manifold alignment[C] //Proc of the 22nd Int Joint Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2011: 1440–1451

[21] Chen Sentao, Han Le, Liu Xiaolan, et al. Subspace distribution adaptation frameworks for domain adaptation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(12): 5204-5218

[22] Li Haoliang, Pan Jialin, Wang Shiqi, et al. Heterogeneous domain adaptation via nonlinear matrix factorization[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(3): 984-996

[23] Zellinger W, Moser B A, Grubinger T, et al. Robust unsupervised domain adaptation for neural networks via moment alignment[J]. Information Sciences, 2019, 483(1): 174-191

[24] Wang Wenqi, Wang Run, Wang Lina, et al. An adversarial sample generation method for Chinese text propensity classification[J]. Journal of Software, 2019, 30(8): 2415-2427 (in Chinese)
(王文琦, 汪润, 王丽娜, 等. 面向中文文本倾向性分类的对抗数据生成方法[J]. 软件学报, 2019, 30(8): 2415-2427)

[25] Deng Zhennong, Ma Fuxin, Lan Rushi, et al. A two-stage chinese text summarization algorithm using keyword information and adversarial learning[J]. Neurocomputing, 2021, 425(1): 117-126

[26] Houlsby N, Giurgiu A, Jastrzebski S, et al. Parameter-efficient transfer learning for NLP[C] //Proc of the 36th Int Conf on Machine Learning. Cambridge, MA: PMLR, 2019: 2790-2799

[27] Zhang Haofeng, Liu Li, Long Yang, et al. Deep transductive network for generalized zero shot learning[J]. Pattern Recognition, 2020, 105(1): 107370-107402

[28] Zhao Tiancheng, Eskenazi M. Zero-shot dialog generation with cross-domain latent actions [C] //Proc of the 19th Annual SIGdial Meeting on Discourse and Dialogue. Stroudsburg, PA: ACL, 2018: 1−10

[29] Liu Zihan, Shin J, Xu Yan, et al. Zero-shot cross-lingual dialogue systems with transferable latent variables [C] //Proc of the 24th Conf on Empirical Methods in Natural Language Processing and the 9th Int Joint Conf on Natural Language Processing. Stroudsburg, PA: ACL, 2019: 1297−1303

[30] Ayana, Shen Shiqi, Chen Yun, et al. Zero-shot cross-lingual neural headline generation[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2018, 26(12): 2319−2327

[31] Duan Xiangyu, Yin Mingming, Zhang Min, et al. Zero-shot cross-lingual abstractive sentence summarization through teaching generation and attention[C] //Proc of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2019: 3162−3172

[32] Devlin J, Chang Mingwei, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding [J]. arXiv preprint, arXiv: 1810.04805, 2018

[33] Rousseeuw P J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis[J]. Journal of Computational and Applied Mathematics, 1987, 20(1): 53−65

[34] Liu Yang, Lapata M. Text summarization with pretrained encoders [C] //Proc of the 2019 Conf on Empirical Methods in Natural Language Processing and the 9th Int Joint Conf on Natural Language Processing. Stroudsburg, PA: ACL, 2019: 3730−3740

[35] Keneshloo Y, Ramakrishnan N, Reddy C K. Deep transfer reinforcement learning for text summarization [J]. arXiv preprint, arXiv: 1810.06667, 2018

[36] Qian Kun, Yu Zhou. Domain adaptive dialog generation via meta learning[C] //Proc of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2019: 2639−2649

[37] YiSyuan C, Shuai Honghan. Meta-transfer learning for low-resource abstractive summarization [C] //Proc of the 36th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2021: 12692−12700

[38] Song Xiaotao, Sun Hailong. A review of neural network-based automatic source code abstraction techniques[J]. Journal of Software, 2022, 33(1): 55−77 (in Chinese)
(宋晓涛, 孙海龙. 基于神经网络的自动源代码摘要技术综述[J]. 软件学报, 2022, 33(1): 55−77)

Ma Tinghuai, born in 1974. PhD, professor, PhD supervisor. Senior member of CCF. His main research interests include social network privacy protection, big data mining, and text emotion computing.

马廷淮, 1974 年生. 博士, 教授, 博士生导师. CCF 高级会员.主要研究方向为社交网络隐私保护、大数据挖掘、文本情感计算.

Yu Xin, born in 1994. PhD candidate. Student member of CCF. His main research interests include nature language processing and transfer learning.

于　信, 1994 年生. 博士研究生. CCF 学生会员. 主要研究方向为自然语言处理、迁移学习.

Rong Huan, born in 1990. PhD, lecturer. His main research interests include social media mining, content security on social network, and knowledge engineering.

荣　欢, 1990 年生. 博士, 讲师. 主要研究方向为社交媒体挖掘、社交网络内容安全、知识工程.