

面向汉语作为第二语言学习的个性化语法纠错

张生盛^{1,3}, 庞桂娜^{2,3}, 杨麟儿^{2,3}, 王辰成^{4,3}, 杜永萍⁴, 杨尔弘³, 黄雅平¹

¹北京交通大学, 计算机与信息技术学院

²北京语言大学, 信息科学学院

³北京语言大学, 语言资源高精尖创新中心

⁴北京工业大学, 信息学部

摘要

语法纠错任务旨在通过自然语言处理技术自动检测并纠正文本中的语序、拼写等语法错误。当前许多针对汉语的语法纠错方法已取得较好的效果, 但往往忽略了学习者的个性化特征, 如二语等级、母语背景等。因此, 本文面向汉语作为第二语言的学习者, 提出个性化语法纠错, 对不同特征的学习者所犯的错误分别进行纠正, 并构建了不同领域汉语学习者的数据集进行实验。实验结果表明, 将语法纠错模型适应到学习者的各个领域后, 性能得到明显提升。

关键词: 语法纠错; 个性化; 汉语学习者; 领域适应

Personalizing Grammatical Error Correction for Chinese as a Second Language

Shengsheng Zhang^{1,3}, Guina Pang^{2,3}, Liner Yang^{2,3},
Chencheng Wang^{4,3}, Yongping Du⁴, Erhong Yang³, Yaping Huang¹

¹Beijing Jiaotong University, School of Computer and Information Technology

²Beijing Language and Culture University, School of Information Science

³Beijing Language and Culture University,

Beijing Advanced Innovation Center for Language Resources

⁴Beijing University of Technology, Faculty of Information Technology

Abstract

The Grammatical Error Correction (GEC) task is to realize automatic error detection and correction of text through natural language processing technology, such as word order, spelling and other grammatical errors. Many existing Chinese GEC methods have achieved good results, but these methods have not taken into account the characteristics of learners, such as level, native language and so on. Therefore, this paper proposes to personalize the GEC model to the characteristics of Chinese as a Second Language (CSL) learners and correct the mistakes made by CSL learners with different characteristics. To verify our method, we construct domain adaptation datasets. Experiment results on the domain adaptation datasets demonstrate that the performance of the GEC model is greatly improved after adapting to various domains of CSL learners.

Keywords: Grammatical Error Correction, Personalizing, Chinese as a Second Language Learners, Domain Adaptation

1 引言

语法纠错(Grammatical Error Correction, GEC)任务旨在自动检测文本中存在的标点、语序等语法错误, 识别错误的类型并对错误进行自动改正。语法纠错系统的输入是一个可能有语法错误的句子, 输出是其相应的修改后的句子。如图1所示, 第一行表示系统的输入, 第二行表示系统的输出, 其中加粗部分表示修改的内容。随着人工智能越来越深入地影响人们的日常生活, 而自然语言处理作为语言学和计算机学科完美融合的一个学科, 在人工智能领域扮演着重要的角色。语法纠错任务是自然语言处理领域的一个重要分支, 不论是在实际生活还是科研领域都有着举足轻重的地位, 吸引了大量的研究者。

输入: 虽然不能完整的解决代沟的问题, 但能减少代沟之宽度。

输出: 虽然不能**彻底地**解决代沟的问题, 但能**缩短**代沟之**距离**。

图 1. 语法纠错系统的输入输出示例

Lee and Seneff (2008)发现, 二语学习者犯的语法错误经常受他们的母语因素的影响。例如, 母语是日语且把英语作为第二语言的学习者常滥用冠词和介词。对母语是日语的二语学习者常犯的这类错误建模, 可以有效地提高语言学习系统的性能。许多基于深度学习的方法 (Chollampatt and Ng, 2018; Junczys-Dowmunt et al., 2018)将语法纠错视为序列到序列(seq2seq)的任务, 因此神经机器翻译(Neural Machine Translation, NMT)的方法被成功地运用到语法纠错中, 通过将一个错误的句子翻译为正确的句子来实现纠错, 并且在一般的领域获得了很好的性能。但是这些基于seq2seq的方法在特定的领域并不能获得鲁棒的性能, 其主要原因是并未对特定的领域建模从而出现了领域漂移的现象。例如, Nadejde and Tetreault (2019)使用一般领域训练的模型在特定领域做测试, 发现性能明显下降。进而他们将语法纠错模型适应到英语作为第二语言的学习者的母语和等级上, 发现语法纠错系统的性能得到了明显的提升。但是, 当前针对汉语的语法纠错方法 (Ren et al., 2018; Fu et al., 2018; Zhou et al., 2018; Zhao and Wang, 2020)都集中在一般的领域, 并未对特定的领域建模。由于汉语语法与英语语法不同, 汉语水平考试的等级与英语水平考试的等级评判标准不一致, 二者之间没有直接的联系, 故本文面向汉语作为第二语言的学习者, 提出个性化语法纠错, 通过迁移学习方法将一般的语法纠错系统适应到汉语学习者不同的领域, 如汉语学习者的等级、母语等, 并对不同等级、不同母语的汉语学习者犯的错误分别进行纠正。

为验证提出的方法是否合理, 首先, 本文选择汉语学习者的等级和母语作为领域适应的设置, 并构建了不同领域汉语学习者的数据集。然后, 本文将语法纠错任务视为翻译任务, 通过将错误的句子翻译为正确的句子实现纠错, 并选择基于Transformer增强架构的中文语法纠错模型 (王辰成 et al., 2019)作为实验模型。最后, 在不同领域的数据集上展开了实验, 将一般的语法纠错系统适应到相应的领域, 并对不同领域的汉语学习者所犯的错误分别进行纠错。实验结果表明, 语法纠错模型适应到学习者的各个领域后, 纠错性能得到显著提升。其中, 学习者等级领域适应模型、母语领域适应模型以及母语-等级领域适应模型分别比基线模型高出1.92、1.73、1.76个百分点。

本文的主要贡献如下:

- 1) 首次提出面向汉语作为第二语言学习的个性化语法纠错, 对不同等级、不同母语的汉语学习者分别进行纠错;
- 2) 构建了不同领域汉语学习者的数据集, 用来训练和测试语法纠错模型适应到汉语学习者不同领域后的性能;
- 3) 将语法纠错模型适应到汉语学习者不同的领域后, 纠错性能得到显著提升, 整体实验结果均超越基线模型, 验证了提出的方法的合理性。

论文的整体结构: 第一节是引言; 第二节介绍了使用的语法纠错模型和领域适应的方法; 第三节是详细的实验设置; 第四节给出了实验细节和各个模型对样例的纠错结果; 第五节是相关工作, 主要介绍了与本文有关的工作; 最后是论文的总结部分。

2 方法

在给定一个长度为 M 的错误句子 $X = \{x_1, \dots, x_M\}$ 和一个学习者的领域 d 后, 基于神经机器翻译的语法纠错模型使用神经网络对输出句子 $Y = \{y_1, \dots, y_N\}$ 的条件概率建模, 如公式1:

$$p(Y|X, d; \theta) = \prod_{t=1}^N p(y_t|y_{1:t-1}, x_{1:M}, d; \theta), \quad (1)$$

其中 θ 是模型的参数。根据Madotto et al. (2019)的工作, 首先将模型的参数 θ 适应到学习者的领域 d , 然后使用错误的句子对输出句子的条件概率建模, 如公式2:

$$p(Y|X; \theta_d) = \prod_{t=1}^N p(y_t|y_{1:t-1}, x_{1:M}; \theta_d), \quad (2)$$

其中 θ_d 是对学习者的领域 d 建模以后的参数。学习者的领域有多种定义的方式, 如等级、母语、母语和等级的组合等。

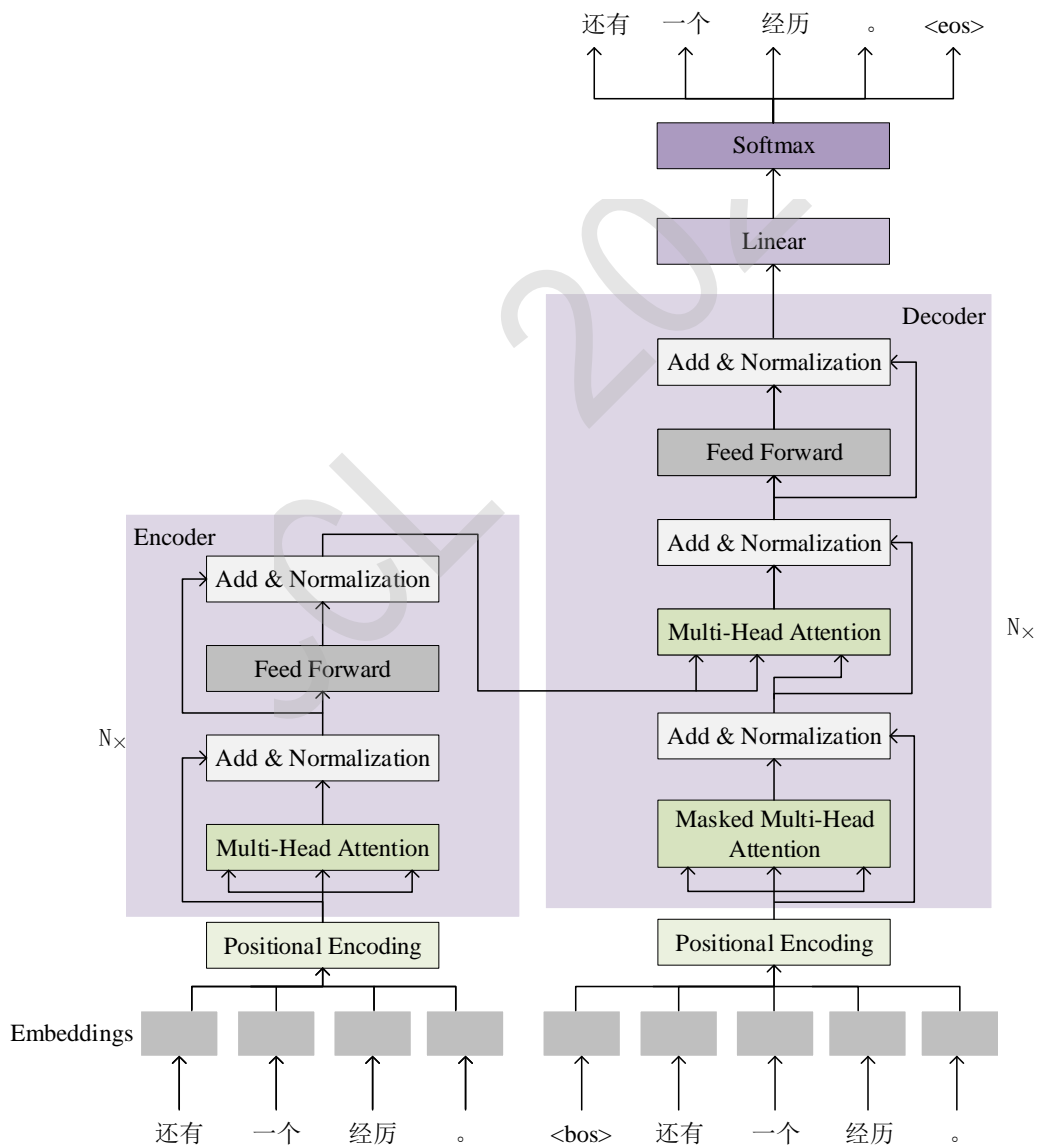


图 2. Transformer语法纠错模型

本文采用王辰成et al. (2019)实现的基于Transformer增强架构的中文语法纠错模型, 该模型不仅可以捕获丰富的语义信息, 还可以减少因为网络过深出现的梯度消失问题。

Transformer (Vaswani et al., 2017)模型是基于多头注意力机制的seq2seq生成模型，如图2所示，它由编码器(Encoder)和解码器(Decoder)组成。其中编码器由 N 个相同的模块构成，即图2左侧部分，每个模块由两个网络层构成，分别是多头自注意力机制和全连接的前馈网络，两者之间都使用了归一化和残差连接。解码器同样是由 N 个相同的模块构成，即图2右侧部分，与编码器不同的是解码器中包含使用编码器输出进行运算的多头注意力层。编码器的作用是将输入序列编码为高维隐含语义向量，解码器根据上一时间步的输出，解码隐含语义向量作为当前时间步的输出，每个时间步的输出对应序列中的一个元素，所有时间步的输出拼接在一起作为最终的输出序列。王辰成et al. (2019)将各个模块的输出动态地结合到一起，以此来增强模型对语义信息的表达能力。

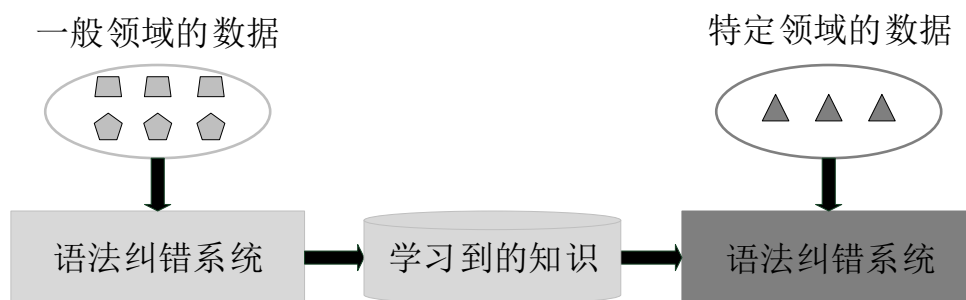


图 3. 语法纠错系统领域适应框架

为实现个性化语法纠错，本文采用迁移学习方法，将一般领域的语法纠错模型适应到学习者的特征领域。采用迁移学习方法的主要原因是一般领域的数据较多，学习者特征领域的数据较少，一般领域的数据与学习者特征领域的数据可以共享模型的参数，更好地帮助模型适应到学习者的特征领域。具体做法是：首先使用一般领域的的数据对语法纠错模型做预训练，再利用学习者特征领域的的数据对模型进行微调，使其适应到相应的领域，整体的框架如图3所示。

3 实验设置

3.1 数据集

本文中首先在Lang-8⁰数据集上对模型进行预训练，然后使用HSK¹作为领域适应的数据集。两个数据集均由汉语作为第二语言的学习者书写，并由母语是汉语的人进行了纠错。我们从两个数据集中抽取平行句对，去掉未修改的句对。使用jieba²分词工具对所有句子进行分词，并且运用字节对编码算法(Byte Pair Encoding, BPE)³ (Sennrich et al., 2016)进一步限制词表大小，以缓解罕见词和未登录词(Out of Vocabulary, OOV)的问题。

原始句子	好像我的疲劳感也飞过去了。
分词后的句子	好像 我 的 疲 劳 感 也 飞 过 去 了 。
BPE拆分后的句子	好像 我 的 疲@@ 劳@@ 感 也 飞 过 去 了 。

表 1. 一个句子经过jieba分词和BPE的示例

数据集	句子数目	原始句子的词语	修改后句子的词语
Lang-8	1,095,985	14,352,734	15,090,960
HSK	88,509	1,783,487	1,768,823

表 2. 数据集详情

⁰<https://lang-8.com>

¹<http://hsk.blcu.edu.cn/>

²<https://github.com/fxsjy/jieba>

³<https://github.com/rsennrich/subword-nmt>

如表1所示，给出了一个句子经过分词和BPE以后的结果，其中‘@@’符号表示当前单元与下一个单元同属一个词语。经过处理后两个数据集的详细情况如表2所示。

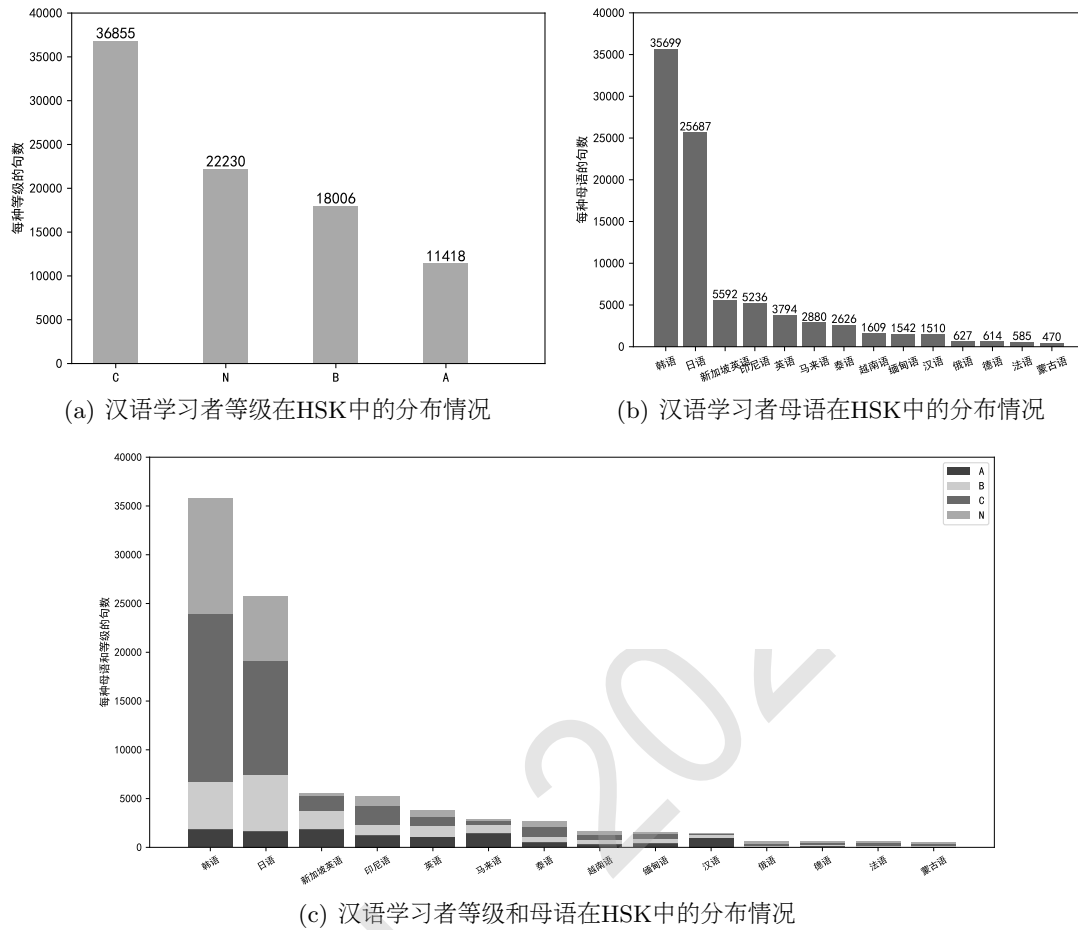


图 4. HSK中各个等级、母语以及等级和母语的的句子数分布

我们在HSK数据集上研究个性化语法纠错，该数据集由14种不同母语和4种不同等级的汉语学习者所写的考试作文组成。我们根据学习者的母语和等级信息抽取了平行句对，其中各个等级所包含的句子数如图4(a)所示，各个母语所包含的句子数如图4(b)所示，各个等级和母语的的句子数如图4(c)所示。

3.2 超参数

在实验中，编码器词嵌入矩阵和解码器词嵌入矩阵维度为512，解码器的输入和输出词嵌入矩阵共享权重。编码器和解码器各包含6个模块，每个模块的多头注意力层有8个注意力头，前馈层的维度大小为2048。优化器使用Adam，动量设置为(0.9, 0.98)，warm-up为4000，学习率的更新策略为初始值是 1×10^{-7} ，在前4000步的训练中，线性增长到 5×10^{-4} ，之后逐步指数下降到训练结束，Dropout为0.3，柱搜索的大小为12，最大的token数为4000。在预训练阶段，我们根据Ren et al. (2018)的方法，随机从Lang-8中抽取5000个句对作为验证集，并选取验证集上最优的模型作为最终预训练的模型。在微调阶段，选取验证集上最优的前5个模型的权重并计算权重的平均值作为领域适应模型的权重。

3.3 基线模型

本文的基线模型有两个，分别为：

- 1) 无微调 (Nadejde and Tetreault, 2019): 该方法直接使用预训练的模型对各个领域测试集中的句子进行纠错;

- 2) 随机 (Nadejde and Tetreault, 2019): 该方法随机地从HSK数据集中抽取与各个领域相同数量的训练集和开发集, 并使用这些数据对预训练的模型进行微调, 然后对各个领域的测试集中的句子进行纠错。领域适应的各个模型超越这个基线模型可以帮助我们验证语法纠错模型性能的提升不仅是因为模型适应到HSK数据集上, 还因为对各个领域成功地建模。

3.4 评价指标

正如王辰成 et al. (2019)提到的应该重点关注模型对错误编辑的准确性而非编辑的数量, 所以为了评价语法纠错模型的性能, 我们采用MaxMatch(M^2)工具包计算 $F_{0.5}$ 分数, 根据 $F_{0.5}$ 的大小判断模型的性能。计算 $F_{0.5}$ 需要有 $m2$ 格式的文件, 它是根据原句和修改句来生成最佳的黄金编辑集合。因此, 我们使用ERRANT⁴工具包制作了各个测试集 $m2$ 格式的文件。

4 实验结果

为了验证提出的方法是否合理, 我们构建了各个领域适应的数据集, 根据Nadejde and Tetreault (2019)的工作中的实验设置, 我们将各个领域适应的数据集按照8:1:2的比例划分为训练集、开发集和测试集。

4.1 针对汉语学习者等级的语法纠错模型

HSK数据集将汉语学习者的水平分为A、B、C和N(N表示无)四个等级, 我们为每个等级随机抽取11,000个平行句对; 为随机基线模型在不考虑等级的条件下随机抽取8,000个平行句对作为训练集, 1,000个平行句对作为开发集。

等级	无微调	随机	适应到等级
A	18.53	38.45	42.04
B	22.44	41.02	43.19
C	24.56	44.76	44.91
N	23.62	45.84	47.63
平均	22.29	42.52	44.44

表 3. 不同等级上的结果

实验结果如表3所示: 我们的结果在等级A上超出随机基线模型3.59个百分点, 从图4(a)可以看出, 等级A的句子是HSK中数量最少的, 却是提升最高的, 这表明了个性化语法纠错是合理且有效的。在等级C上, 相比随机基线模型我们的方法提升了0.15个百分点, 从图4(a)可以看出, 等级C的句子是HSK中数量最多的, 所以在不考虑等级的情况下随机抽到的句对中等级C的数量也是最多的, 这很好地解释了在等级C上的提升是最小的。从整体结果来看, 我们的方法比随机基线模型高出1.92个百分点。

4.2 针对汉语学习者母语的语法纠错模型

母语	无微调	随机	适应到母语
韩语	23.84	44.25	45.58
日语	28.31	46.19	46.47
新加坡英语	14.74	30.25	35.23
印尼语	18.61	37.93	38.28
平均	21.38	39.66	41.39

表 4. 不同母语上的结果

HSK数据集由14种母语的汉语学习者书写, 本实验选用其中数据最多的4种母语: 韩语、日语、新加坡英语、印尼语。我们为这4种不同的母语分别随机抽取4,950个平行句对; 为随机

⁴<https://github.com/chrisjbryant/errant>

基线模型在不考虑母语的条件下随机抽取3,600个平行句对作为训练集, 450个平行句对作为开发集。

实验结果如表4所示: 我们的方法在新加坡英语上比随机基线模型高出4.98个百分点, 超越了其他母语。此外, 在新加坡英语上, 从无微调基线模型到随机基线模型的提升为15.51个百分点, 低于其他母语。以上结果说明领域适应模型性能的提升不仅是因为将模型适应到HSK数据集上, 还因为对学习者的领域进行了建模。从整体结果来看, 我们的方法比随机基线模型超出1.73个百分点。

4.3 针对汉语学习者母语和等级的语法纠错模型

HSK数据集中包含4种等级和14种母语的组, 共有56种母语-等级组合的情况, 本实验选用其中数据最多的5种母语-等级组合: 韩语-C、韩语-N、日语-B、日语-C、日语-N。我们为这5种不同的母语-等级组合分别随机抽取4,950个平行句对; 为韩语和日语每种语言随机抽取3,600个平行句对作为训练集, 450个平行句对作为开发集; 为B、C、N每种等级随机抽取3,600个平行句对作为训练集, 450个平行句对作为开发集; 为随机基线模型在不考虑等级、母语的条件下随机抽3,600个平行句对作为训练集, 450个平行句对作为开发集。

	韩语-C	韩语-N	日语-B	日语-C	日语-N	平均
无微调	23.43	25.00	24.01	23.78	25.11	24.27
随机	44.67	44.64	43.13	44.71	46.37	44.70
等级	44.50	47.80	42.33	45.72	44.66	45.00
母语	44.92	46.69	43.96	45.27	46.83	45.53
母语-等级	45.83	48.26	43.34	46.62	48.26	46.46

表 5. 不同母语-等级上的结果

实验结果如表5所示: 我们的方法在韩语-N上比随机基线模型提升了3.62个百分点, 比韩语提升了1.57个百分点, 超越了其他母语-等级, 是母语-等级与随机、母语相比提升效果最明显的组合; 在日语-N上比等级N提升了3.60个百分点, 是母语-等级与等级相比提升效果最明显的组合。从整体结果来看, 适应到母语-等级的模型比随机基线模型高出1.76个百分点, 比适应到等级的模型高出1.46个百分点, 比适应到母语的模型高出0.93个百分点。

4.4 实验结果分析

原始句子	第一当一个人在公共场所内抽烟当他抽完烟, 如果他找不到丢烟蒂的地方, 他可能会随手丢掉,
标准答案	第一, 一个人在公共场所内抽烟, 当他抽完烟后, 如果他找不到丢烟蒂的地方, 他可能会随手丢掉。
无微调	第一, 当一个人在公共场所内抽烟时, 他抽完烟, 如果他找不到丢心烦的地方, 他可能会随手丢掉,
随机	第一, 当一个人在公共场所内抽烟时, 他抽完烟, 如果他找不到丢拾水的地方, 他可能会随手丢掉。
适应到等级A	第一, 当一个人在公共场所内抽烟时, 他抽完烟, 如果他找不到丢的地方, 他可能会随手丢掉。

表 6. 不同模型的纠错结果

从表3可知, 相比随机基线模型, 领域适应的模型性能在等级A上的提升最明显。因此本文针对等级A的纠错样例进行分析, 从而直观地反映将一般的语法纠错模型适应到相应领域后纠错能力的表现。如表6所示, 给出了各个模型对一个有语法错误句子的修改结果, 其中加粗部分表示修改内容。从表6观察发现相比随机基线模型, 领域适应的模型修改更符合汉语用语习惯。

因此, 将语法纠错模型适应到汉语学习者的特征领域以后, 模型纠错的结果更接近汉语的习惯表达。

5 相关工作

传统的语法纠错方法可以分为两类：1) 基于规则的 (Bustamante and León, 1996)，这些方法只关注文本中的几种错误类型；2) 基于统计机器翻译的 (Brockett et al., 2006)，这些方法将语法纠错任务视为翻译任务，并使用统计器翻译方法进行纠错，极大地提升了语法纠错系统的性能。

随着深度学习的发展，许多序列到序列的方法成功地应用到语法纠错中，这些方法将语法纠错视为一般的seq2seq任务，即系统的输入是一个序列，输出也是一个序列。Yuan and Briscoe (2016)第一次将神经机器翻译模型应用到语法纠错任务中，他们使用双向的递归神经网络编码器和一个基于注意力的解码器对错误进行纠错，性能超越了基于统计机器翻译的纠错模型。Ji et al. (2017)使用嵌套注意力神经混合模型纠错，该模型通过合并单词和字符级别的信息来纠正两种类型的错误。Chollampatt and Ng (2018)使用一个多层卷积编码-解码器神经网络模型进行纠错，并结合语言模型对纠错的结果进行重排序。Junczys-Dowmunt et al. (2018)将语法纠错视作低资源的机器翻译任务，并使用Transformer作为纠错模型。

针对英语的语法纠错方法层出不穷，并且取得了很好的效果。但汉语语法纠错方兴未艾，自中国计算机协会举办的国际自然语言处理与中文计算会议(NLPCC)在2018年加入了汉语语法纠错评测任务后，出现了许多汉语语法纠错方法。Fu et al. (2018)采用简单到复杂的分阶段纠错方法，使用语言模型纠正简单的错误，字、词级的Transformer模型纠正复杂的错误。Zhou et al. (2018)使用多个模型纠错，分别是基于规则、统计和神经网络，通过模型组合的方式得到最终的纠错结果。Ren et al. (2018)使用基于卷积神经网络的seq2seq模型纠错，还采用了subword算法 (Sennrich et al., 2016)来缩小词表和缓解未登录词的问题。王辰成 et al. (2019)提出了基于Transformer增强架构的中文语法纠错模型，该模型使用动态残差结构结合不同神经模块的输出来增强模型捕获语义信息的能力。Zhao and Wang (2020)采用动态掩码的方式提高模型的纠错性能，在训练步骤中动态地向原始的句子加入掩码来增加更多的平行句对。但以上这些汉语语法纠错方法均没有对汉语学习者的个性化特征进行建模。

周小兵 et al. (2007)在对汉语作为第二语言的学习者的教学研究中发现，母语迁移是造成二语学习者语法偏误的一项主要原因。如有的汉语学习者会写“*我见面我的老师。”这样的错句，因为法语、韩语、日语、越南语等语言中“见面”可以带宾语，但是汉语中“见面”后面是不可以带宾语的。此外，Swan and Smith (2001)在对二语学习者的教学研究中发现不同母语写作者会犯不同类型的错误，他们将其中的某些错误归因于语言之间的“转移”或“干扰”，即母语的“负迁移”。因此，已有许多针对二语学习者的研究，如Rozovskaya and Roth (2011)使用朴素贝叶斯对不同母语的英语学习者所犯的介词错误进行纠错；Mizumoto et al. (2011)发现当训练数据和测试数据用相同的母语时，语法纠错系统的表现会更好；Nadejde and Tetreault (2019)针对英语作为第二语言的学习者提出了个性化语法纠错，发现将语法纠错模型适应到学习者的不同特征时表现会更好。

6 总结

本文针对汉语作为第二语言的学习者，首次提出了个性化语法纠错。并将语法纠错任务视为翻译任务，使用基于Transformer增强架构的中文语法纠错模型对错误进行纠正。为了验证提出的方法的合理性，构建了不同领域的数据集，并使用迁移学习方法将语法纠错模型适应到学习者不同的领域，实现个性化语法纠错。各个领域测试集上的平均结果都超越了未做领域适应的基线模型，表明语法纠错系统在对学习者的特征建模以后可以有效地改进纠错的效果。

致谢

感谢北京语言大学语言资源高精尖创新中心项目(TYZ19005)和国家语委信息化项目(ZDI135-105)对本研究的支持。

参考文献

Chris Brockett, William B. Dolan, and Michael Gamon. 2006. Correcting ESL errors using phrasal SMT techniques. In *Proceedings of the 21st International Conference on Computational Linguistics*

- and *44th Annual Meeting of the Association for Computational Linguistics*, pages 249–256, Sydney, Australia, July. Association for Computational Linguistics.
- Flora Ramírez Bustamante and Fernando Sánchez León. 1996. GramCheck: A grammar and style checker. In *COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics*.
- Shamil Chollampatt and Hwee Tou Ng. 2018. A multilayer convolutional encoder-decoder neural network for grammatical error correction. *CoRR*, abs/1801.08831.
- Kai Fu, Jin Huang, and Yitao Duan. 2018. Youdao’s winning solution to the NLPCC-2018 task 2 challenge: A neural machine translation approach to chinese grammatical error correction. In Min Zhang, Vincent Ng, Dongyan Zhao, Sujian Li, and Hongying Zan, editors, *Natural Language Processing and Chinese Computing - 7th CCF International Conference, NLPCC 2018, Hohhot, China, August 26-30, 2018, Proceedings, Part I*, volume 11108 of *Lecture Notes in Computer Science*, pages 341–350. Springer.
- Jianshu Ji, Qinlong Wang, Kristina Toutanova, Yongen Gong, Steven Truong, and Jianfeng Gao. 2017. A nested attention neural hybrid model for grammatical error correction. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 753–762, Vancouver, Canada, July. Association for Computational Linguistics.
- Marcin Junczys-Dowmunt, Roman Grundkiewicz, Shubha Guha, and Kenneth Heafield. 2018. Approaching neural grammatical error correction as a low-resource machine translation task. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 595–606, New Orleans, Louisiana, June. Association for Computational Linguistics.
- J. Lee and S. Seneff. 2008. An analysis of grammatical errors in non-native speech in english. In *2008 IEEE Spoken Language Technology Workshop*, pages 89–92.
- Andrea Madotto, Zhaojiang Lin, Chien-Sheng Wu, and Pascale Fung. 2019. Personalizing dialogue agents via meta-learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5459, Florence, Italy, July. Association for Computational Linguistics.
- Tomoya Mizumoto, Mamoru Komachi, Masaaki Nagata, and Yuji Matsumoto. 2011. Mining revision log of language learning SNS for automated Japanese error correction of second language learners. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 147–155, Chiang Mai, Thailand, November. Asian Federation of Natural Language Processing.
- Maria Nadejde and Joel Tetreault. 2019. Personalizing grammatical error correction: Adaptation to proficiency level and l1. In *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, pages 27–33.
- Hongkai Ren, Liner Yang, and Endong Xun. 2018. A sequence to sequence learning for chinese grammatical error correction. In Min Zhang, Vincent Ng, Dongyan Zhao, Sujian Li, and Hongying Zan, editors, *Natural Language Processing and Chinese Computing - 7th CCF International Conference, NLPCC 2018, Hohhot, China, August 26-30, 2018, Proceedings, Part II*, volume 11109 of *Lecture Notes in Computer Science*, pages 401–410. Springer.
- Alla Rozovskaya and Dan Roth. 2011. Algorithm selection and model adaptation for ESL correction tasks. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 924–933, Portland, Oregon, USA, June. Association for Computational Linguistics.
- Rico Sennrich, Barry Haddow, and Alexandra Birch. 2016. Neural machine translation of rare words with subword units. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1715–1725, Berlin, Germany, August. Association for Computational Linguistics.
- M. Swan and B. Smith. 2001. *Learner English: A Teacher’s Guide to Interference and Other Problems*. Cambridge handbooks for language teachers. Cambridge University Press.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *CoRR*, abs/1706.03762.

- Zheng Yuan and Ted Briscoe. 2016. Grammatical error correction using neural machine translation. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 380–386, San Diego, California, June. Association for Computational Linguistics.
- Zewei Zhao and Houfeng Wang. 2020. Maskgec: Improving neural grammatical error correction via dynamic masking. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 1226–1233. AAAI Press.
- Junpei Zhou, Chen Li, Hengyou Liu, Zuyi Bao, Guangwei Xu, and Linlin Li. 2018. Chinese grammatical error correction using statistical and neural models. In Min Zhang, Vincent Ng, Dongyan Zhao, Sujian Li, and Hongying Zan, editors, *Natural Language Processing and Chinese Computing - 7th CCF International Conference, NLPCC 2018, Hohhot, China, August 26-30, 2018, Proceedings, Part II*, volume 11109 of *Lecture Notes in Computer Science*, pages 117–128. Springer.
- 周小兵, 朱其智, and 邓小宁. 2007. 外国人学汉语语法偏误研究. 北京语言文化大学出版社.
- 王辰成, 杨麟儿, 王莹莹, 杜永萍, and 杨尔弘. 2019. 基于Transformer增强架构的中文语法纠错方法. 中国计算语言学大会.