

## ◆ 算法研究与应用 ◆

结合多注意力和条件变分自编码器的宋词生成模型<sup>\*</sup>梁 骁<sup>1</sup>, 黄文明<sup>1,2</sup>, 姚 俊<sup>3</sup>, 温雅媛<sup>4</sup>, 邓珍荣<sup>1,2\*\*</sup>

(1. 桂林电子科技大学计算机与信息安全学院, 广西桂林 541004; 2. 广西图形图像与智能处理重点实验室, 广西桂林 541004; 3. 广西壮族自治区高级人民法院, 广西南宁 530000; 4. 广西师范大学电子工程学院, 广西桂林 541004)

**摘要:** 现有的诗词生成方法忽略了风格的重要性。另外, 由于宋词大部分词牌词句较多, 逐句生成宋词的过程中容易产生上下文缺乏连贯性的现象, 在上下文连贯性方面仍存在提升空间。针对这两个问题, 在编码解码的文本生成框架基础上, 引入自注意力机制的句子表示算法计算多注意力权重矩阵, 用于提取词句的多种重要语义特征, 让模型更多地关注上文的显著信息来提高上下文连贯性。引入条件变分自编码器(CVAE)将每条宋词数据转化为隐空间中不同风格特征的高维高斯分布, 从各自的分布中采样隐变量来控制宋词的風格。由于自构建的宋词语料库缺少完整风格分类标签, 使用具有风格标签的宋词微调 BERT 模型, 将其作为风格分类器标注全部的宋词数据。在上述关键技术的基础上实现了宋词生成模型, 生成上下文连贯的婉约词以及豪放词。通过与其他 4 种基准方法进行对比实验, 结果表明引入自注意力机制的句子表示算法和条件变分自编码器, 在上下文连贯性和风格控制方面有一定的提升。

**关键词:** 条件变分自编码器 宋词风格 宋词生成 Bi-GRU 自注意力机制

中图分类号: TP391.1; I207.22 文献标识码: A 文章编号: 1005-9164(2022)02-0308-08

DOI: 10.13656/j.cnki.gxkx.20220526.011

词源于民间, 兴于五代, 盛于两宋。宋词是人们对文化生活强烈追求的产物, 也是中国传统文化最重要的文学形式之一。宋词自动生成研究能帮助普罗大众学习和自动生成宋词, 对进一步弘扬中华民族传统文化具有重要意义。近年来, 神经生成模型在自然语言处理上得到广泛应用, 其中的古诗词自动生成研

究受到众多学者的青睐, 引起了研究者们浓厚的兴趣。

将现有的古诗词自动生成研究分为传统方法和基于深度学习的方法。传统方法通过诗词领域的专家人为设计规则和约束来生成古诗词。其中基于模板的方法类似于完形填空, 从诗歌中去除一些单词,

收稿日期: 2021-04-16

<sup>\*</sup> 广西科技计划项目(桂科 AB20238013), 广西图像图形与智能处理重点实验室培育基地(桂林电子科技大学)开放基金项目(GIIP2011)和广西高校中青年教师科研基础能力提升项目(2019KY0238)资助。

## 【作者简介】

梁 骁(1995-), 男, 硕士研究生, 主要从事自然语言处理研究, E-mail: 2024605849@qq.com。

## 【\*\*通信作者】

邓珍荣(1978-), 女, 研究员, 主要从事自然语言处理研究, E-mail: 799349175@qq.com。

## 【引用本文】

梁骁, 黄文明, 姚俊, 等. 结合多注意力和条件变分自编码器的宋词生成模型[J]. 广西科学, 2022, 29(2): 308-315.

LIANG X, HUANG W M, YAO J, et al. Song Ci Generation Model Based on Multi-attention and Conditional Variational Auto-encoder [J]. Guangxi Sciences, 2022, 29(2): 308-315.

再从单词库中选择合适的单词放入空白处形成新的诗歌。周昌乐等<sup>[1]</sup>提出基于遗传算法的宋词生成方法,从语料库中随机选择一句词句,用设计好的评估函数评价词句,并不断迭代直到生成整首宋词,但该方法缺乏语义连贯性。Yan等<sup>[2]</sup>将古诗生成看成给定写作意图的摘要生成问题,同时加入一些优化约束。He等<sup>[3]</sup>将诗歌生成看成统计机器翻译问题,人为加上平仄押韵等约束。Tosa等<sup>[4]</sup>和Wu等<sup>[5]</sup>采用短语搜索的方法生成日本诗。Netzer等<sup>[6]</sup>采用基于词联想的方法生成俳句。Oliveira等<sup>[7,8]</sup>提出了一种基于语义和语法模板的西班牙诗生成方法。

随着深度学习技术的发展,国内外的古诗词生成研究进入新的阶段。国外基于深度学习的机器翻译和诗歌生成方法如下:Cho等<sup>[9]</sup>提出由两个递归神经网络RNN组成的编码解码器来实现机器翻译。Gulcehre等<sup>[10]</sup>采用内存寻址方案实现动态神经机器翻译。Hopkins等<sup>[11]</sup>集成了一个循环神经网络(RNN)和一个有限状态接受者(FSA)。它生成给定任意主题的英文4行诗。Ghazvininejad等<sup>[12]</sup>使用一个经过语音编码训练的神经语言模型来学习英语诗歌的形式和隐含表达的内容,生成英文韵律诗。这种模式可以有效地学习常见的诗歌音律,如押韵、节奏等。但这些方法使用单向量不能保存上下文的全部信息,并且随着输入序列长度的增加会导致循环神经网络的性能下降。国内基于深度学习的方法如下:首先,基于循环神经网络<sup>[13]</sup>的古诗生成方法将每首古诗的诗句通过拼接的方式形成语料库,用于训练循环神经网络,最终得到一个输出单词字典大小的概率分布模型;然后,给定初始内容,不断从概率分布模型中采样出诗句中的词语,直到生成完整的古诗。基于循环神经网络模型<sup>[14]</sup>给定关键词逐行生成古诗,它由3个部分组成:用于获取句子的局部表示的卷积神经网络;获取已生成诗句的上下文向量的句子级别的循环神经网络;根据历史诗句上下文和该句已经生成的字符,输出下一个字符概率向量的字符级别循环神经网络。该模型捕获诗句的全局信息和局部信息,提高模型的特征提取能力。Wang等<sup>[15]</sup>基于注意力的编码解码框架宋词生成模型(ICG),将已生成诗句拼接起来作为输入传递给编码器,由解码器输出下一个词句。该方法基于注意力机制配合长短期记忆神经网络,可以学习更长的诗歌,在一定程度上提高前后语义的连贯性。Wang等<sup>[16]</sup>先给每一句诗歌规划关键词大纲,让生成的诗句与关键词相关,提高了古诗的主题一致

性。Yan<sup>[17]</sup>基于编码解码框架,其中编码器是卷积神经网络,编码用户输入的意图关键词,解码器由句子级别和词级别两个循环神经网络组成。Ghazvininejad等<sup>[18]</sup>的主题诗生成模型(TPG)与规划模型相似,不同的是它先生成4个押韵的关键词,将其作为诗句的结尾,保证古诗的押韵。基于序列到序列的古诗生成模型(SPG)<sup>[19]</sup>在规划模型基础上增加了基于注意力机制的关键词扩展模型,用于匹配古诗生成模型,提高了古诗的主题一致性。Yi等<sup>[20]</sup>基于显著信息的古诗生成模型(SCG),通过注意力权重提取已生成诗句的显著语义特征,用于指导解码器生成上下文连贯的诗句。Liu等<sup>[21]</sup>通过引入条件变分自编码器(CVAE)生成具有比喻或拟人修辞手法的现代诗,首次在生成的诗歌中加入修辞元素,让诗歌变得更加生动形象。

风格作为宋词的重要元素,对提高意境起着重要作用。宋词的风格主要有两类:豪放与婉约。豪放词直抒胸臆,描述了广阔视野,恢宏的气象;婉约词词调蕴藉,婉约含蓄,侧重细节描写。现有的研究主要关注诗词的通顺性和主题一致性方面,对于宋词风格方面的研究较少,在很大程度上忽视了风格的重要性。另外,宋词属于长文本的体裁,词句较多,现有的生成方法在上下文连贯性上仍存在提升空间。

为此,本文在编码解码的文本生成框架基础上,引入多注意力组件和条件变分自编码器,提出结合多注意力和条件变分自编码器的风格宋词生成方法。在自制的语料库上进行实验,探讨本文方法对于提高上下文连贯性以及实现风格控制的效果。

## 1 概述

### 1.1 设计思想

在采用门控循环单元(GRU)的编码解码文本生成框架基础上,加入基于自注意力机制的句子表示算法的多注意力组件,用于提取词句中的显著信息,提高上下文连贯性。另外,引入条件变分自编码器,将每条宋词数据转化为隐空间中不同风格特征的高维高斯分布,从各自的分布中采样隐变量来控制宋词的风格。本文将从带有注意力机制的基础框架、多注意力组件和条件变分自编码器等方面介绍方法的具体实现。

### 1.2 宋词生成过程

借鉴规划模型的生成过程,用户输入 $n$ 个关键词 $(k_1, k_2, k_3, k_4, \dots, k_n)$ 作为宋词的大纲,其中 $k_n$ 是

第  $n$  行词句的关键词。输入关键词  $k_n$  和第  $n-1$  行的词句, 输出第  $n$  行的词句。

### 1.3 关键词提取

TextRank 算法<sup>[22]</sup>是评估句子或段落中词语重要性的方法。可使用该算法提取出词句中最重要词语作为关键词。首先将宋词数据集中所有词句都拼接起来, 对每个词句分词。根据分词后的词句构造图, 即句子中的每个词语与它自身左右一定范围内的词语连接构成一个无向图。将图中边的权重初始化为 1, 通过公式(1)进行迭代计算得分, 直到收敛为止。最终得到所有关键词的重要性得分表。在训练阶段, 根据关键词重要性得分表从现有的宋词数据集中构建出关键词数据集, 即从每个词句中提取出最重要的关键词。在预测阶段, 对用户输入文本进行分词, 根据关键词得分表提取出关键词。

$$S(V_i) = (1-d) + d \sum_{V_j \in E(V_i)} \frac{\omega_{ji}}{\sum_{V_k \in E(V_j)} \omega_{jk}} S(V_j), \quad (1)$$

式中,  $\omega_{ji}$  为词语  $V_j$  连接边的权值;  $E(V_i)$  为连接到词语  $V_j$  的边;  $d$  为阻尼因子, 通常设置为 0.8;  $S(V_i)$  为词语  $V_i$  的分数, 初始分数设为 1.0。

## 2 模型整体结构

模型的整体结构如图 1 所示, 其中橙色部分包括分别处理关键词、源序列的两个编码器和先验网络 (Prior network); 绿色部分是处理目标序列的编码器和后验网络 (Recognition network), 它们共同组成条件变分自编码器; 黄色部分是多注意力组件; 粉色部分是注意力机制; 紫色部分是解码器。

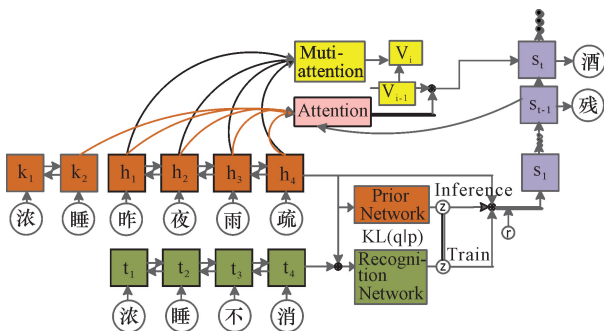


图 1 模型整体结构

Fig. 1 Overall structure of the model

### 2.1 基础框架

本文将带有注意力机制的编码解码框架<sup>[14]</sup>作为基础。编码部分包括 3 个编码器, 其中一个编码器将

关键词转化成隐藏状态  $K = \{k_1, k_2, \dots, k_n\}$ 。另一个编码器将目标序列文本  $Y = \{y_1, y_2, y_3, \dots, y_n\}$  转化成隐藏状态序列  $T = \{t_1, t_2, t_3, \dots, t_n\}$ , 计算公式如下:

$$t_j = GRU(e(y_j), t_{j-1}), \quad (2)$$

式中,  $e(y_j)$  表示目标序列文本  $Y$  中第  $j$  个词的词嵌入。

编码器将源序列文本  $X = \{x_1, x_2, x_3, \dots, x_n\}$  转换成隐藏状态序列  $H = \{h_1, h_2, h_3, \dots, h_n\}$ , 计算公式如下:

$$h_j = GRU(e(x_j), h_{j-1}), \quad (3)$$

式中,  $e(x_j)$  表示源序列文本  $X$  中第  $j$  个词语的词嵌入。

解码器将目标序列文本  $Y_{in} = \{y_s, y_1, y_2, y_3, \dots, y_n\}$  转化为隐藏状态序列  $S = \{s_1, s_2, s_3, \dots, s_n\}$ , 其中  $y_s$  表示词牌指示向量<sup>[15]</sup>。计算公式如下:

$$s_t = GRU(e(y_{t-1}), v_{t-1}, c_t, s_{t-1}), \quad (4)$$

式中, 将隐藏状态  $h_n$ 、条件变分自编码器(2.3 小节)提取到的隐变量  $z$  和风格标签  $r$  的拼接作为解码器的初始隐藏状态。  $v_{t-1}$  是多注意力组件(2.2 小节)提取的语义向量;  $c_t$  是通过注意力机制计算的上下文向量, 用于提供源序列和关键词的全局信息, 计算公式如下:

$$c_t = \sum_{j=0}^{T_n} a_{tj} h_j, \quad (5)$$

式中, 将关键词编码器的最后一个隐藏状态  $k_n$  作为式中  $h_0$ ;  $h_j$  为  $H$  第  $j$  个隐藏状态;  $a_{tj}$  为该隐藏状态的注意力权重, 计算公式如下:

$$a_{tj} = \frac{\exp(e_{tj})}{\sum_{k=1}^{T_n} \exp(e_{tk})}, \quad (6)$$

式中,  $e_{tj}$  的计算公式如下:

$$e_{tj} = V_a^T \tanh(W_a s_{t-1} + U_a h_j), \quad (7)$$

式中,  $V_a, W_a, U_a$  都是需要训练的参数。

### 2.2 多注意力组件

常见的上下文处理方法是将词句通过编码器转换成单个向量, 并通过后续生成的词句来更新该向量。单个向量不能存储宋词长文本的全部特征。无意义的词也会混入该向量中, 如虚词、停用词等。为此, 引入多注意力组件提取词句的多种语义特征, 过滤非显著特征, 从而提高上下文的连贯性。

多注意力组件基于自注意力机制的句子表示算法<sup>[23,24]</sup>, 计算词句的多注意力权重矩阵, 提取词句的 3 种重要语义特征并保存在指定的语义向量  $v$  中。

多注意力权重矩阵的计算公式如下:

$$A = W_{s_2} \tanh(W_{s_1} H^T), \quad (8)$$

式中,  $W_{s_1}$  是参数矩阵, 维度为  $[d_a, 2u]$ ;  $d_a$  是超参数,  $2u$  表示隐藏状态的维度;  $W_{s_2}$  是参数矩阵, 维度为  $[r, d_a]$ ;  $r$  是超参数, 设置为 3;  $A = [a_1, a_2, a_3]$ , 是多注意力权重矩阵, 每行表示词句某种语义的权重向量。为避免  $A$  的每行权重向量相似度高, 导致每个注意力权重向量提取同一语义特征的问题, 在损失函数中引入一个惩罚项<sup>[20]</sup>, 迫使每个注意力权重向量都具有明显的差异。惩罚项公式如下:

$$P = \|AA^T - I\|_F^2, \quad (9)$$

将多注意力权重矩阵  $A$  按行加和并归一化得到向量  $a$ :

$$a = \text{softmax}(\sum_{k=1}^r a_k), \quad (10)$$

在得到  $a$  后, 通过式(11)计算存储多种语义的句子表示:

$$M = aH, \quad (11)$$

$M$  是包含多种语义的句子表示, 根据新生成的  $M$  更新语义向量  $v_i$ :

$$v_i = \partial(v_{i-1}, M), v_0 = 0. \quad (12)$$

多注意力权重矩阵  $A$  在词句“花色烘晴草色浮”上的表现如图 2 所示。

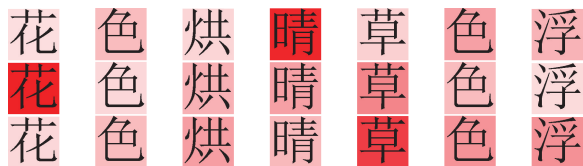


图 2 多注意力权重

Fig. 2 Multi-attention weight

图 2 的 3 个注意力权重主要用于提取“晴草花”的语义特征。 $a$  在词句上的表现如图 3 所示。



图 3 加和并归一化的注意力权重

Fig. 3 Attention weight added and normalized

## 2.3 条件变分自编码器

受 Zhao 等<sup>[25]</sup>和 Sohn 等<sup>[26]</sup>的多样性文本生成模型的启发, 为了控制生成的文本风格, 在基础框架上引入条件变分自编码器, 将每条宋词数据转化成隐空间的风格特征分布。在训练阶段, 输入源序列、关键词、目标序列和风格标签。通过误差反向传播, 降低 KL 散度(KL divergence)损失, 让先验分布接近后验分布, 从后验分布中采样隐变量  $z$ 。在预测阶段, 输入源序列、关键词和风格标签, 从先验分布中采

样隐变量  $z$ 。

### 2.3.1 先验网络和后验网络

构建先验网络和后验网络, 分别计算先验分布和后验分布的均值和方差, 如式(13)和(14)所示:

$$[\mu, \sigma^2] = MLP(t_n, h_n), \quad (13)$$

$$[\mu', \sigma'^2] = MLP(h_n), \quad (14)$$

式中,  $\mu, \sigma^2$  分别是后验分布的均值和方差;  $\mu', \sigma'^2$  分别是先验分布的均值和方差。最后通过 KL 散度估计两个高维高斯分布的信息损失。

为解决采样过程不可导的问题, 采用了重参数技巧<sup>[27]</sup>, 从均值  $\mu$  和偏差  $\sigma^2$  的高维高斯分布中采样隐变量  $z$  等价于从标准正态分布中采样一个  $\beta$ , 再对其进行线性变换, 如式(15)所示:

$$z = \mu + \beta * \sigma, \quad (15)$$

式中, 从  $\beta$  到  $z$  只涉及线性变换,  $\beta$  是固定的。

### 2.3.2 变分下界

通过 KL 散度公式衡量先验分布和后验分布的相似度来判断生成质量。但这两个分布的 KL 散度的计算涉及数据的真实分布, 难以直接计算。为此, 通过最大化变分下界(ELBO), 替代直接计算两个分布的 KL 散度。变分下界如下式所示:

$$\text{ELBO} = E_{q_\varphi(z|y, h_n)}[\log p_\theta(y|z, h_n, r)] - \text{KL}(q_\varphi(z|y, h_n) \| p_\theta(z|h_n)), \quad (16)$$

式中,  $\varphi$  是后验网络的参数,  $\theta$  是先验网络的参数。 $z$  是隐变量,  $y$  是目标序列,  $r$  是风格标签的词嵌入。 $q_\varphi(z|y, h_n)$  是后验分布,  $p_\theta(z|h_n)$  是先验分布。公式右边的第一项是重构的似然估计, 目的是让生成的词句和真实的词句尽可能地接近。第二项是后验分布和先验分布的 KL 散度, 用于衡量后验分布相比先验分布的信息损失。

### 2.3.3 弱化解码器

GRU 解码器因解码能力强容易忽略隐变量  $z$ , 无法控制文本风格。为此, 在训练阶段, 用未知词语(UNK)替换解码器的部分输入单词, 迫使解码器无法仅通过先前生成的单词来预测下一个单词, 需要更多地依赖隐变量  $z$ 。

## 2.4 目标函数和损失函数

### 2.4.1 目标函数

解码器生成目标序列  $Y_{out} = \{y_1, y_2, y_3, \dots, y_n, y_e\}$ , 其中  $y_e$  表示结束字符。目标函数如下式所示:

$$y_t = \text{argmax}_y P(y|s_t), \quad (17)$$

式中,  $s_t$  为在  $t$  时刻的隐藏状态,  $y_{t-1}$  为上一时刻的输出, 最终生成  $t$  时刻最可能的字符  $y_t$ 。

### 2.4.2 损失函数

损失函数在重构损失的基础上,加入多注意力的惩罚项和KL散度损失。损失函数如下式所示:

$$loss = l_{decoder} - l_{kl} + l_p, \quad (18)$$

式中,  $l_{decoder}$  是重构损失,采用交叉熵损失函数,保证了数据压损的质量;  $l_{kl}$  是KL散度损失正则项;  $l_p$  是多注意力的惩罚项。

## 3 验证实验

### 3.1 数据处理

采用 Request 库和 BeautifulSoup 库,从古诗文网和 Github 上公开的中文诗歌库中爬取 20 000 首宋词,对其进行数据清洗,剔除存在乱码的词句,最终形成包含 15 200 首宋词的语料库。将其中的 1 000 首作为验证集,1 000 首作为测试集,剩余的作为训练集。由于语料库的宋词数量较少,首先使用 60 000 首古诗作为预训练模型。

语料库仅存在少部分具有风格标签的宋词,其中婉约词 452 首,豪放词 105 首。从中随机选择 50 首婉约诗词和 50 首豪放诗词作为测试集,剩余的宋词作为训练集微调中文 BERT 模型,在测试集上的准确率为 0.81。通过微调后的 BERT 模型对所有宋词数据做风格标注。

### 3.2 实验设计

本实验的目的是验证本文方法能否生成婉约或豪放的宋词,以及是否在上下文连贯性上有所提高。实验环境如下:硬件环境采用 Intel® Xeon® Bronze 3104 CPU 1.7 GHz 处理器;内存 32 GB; Nvidia-aQuadro P2000 显卡,显存 8 GB。软件环境主要采用 python, cuda, tensorflow 深度学习框架和 jieba 分词。

鉴于宋词数据较小,模型参数设置过大会导致过拟合,所以模型设置的词嵌入、隐藏状态、隐变量、语义向量、上下文向量的维度分别为 300, 256, 256, 128, 128。优化器采用小批量(64)随机梯度下降算法,并采用 Adadelta 算法<sup>[28]</sup>调整学习速率。预测阶段采用集束搜索算法,集束宽度为 10。

宋词包含多种格式的词牌。为此,需要定义指示矩阵,矩阵的每一行向量分别表示不同的词牌指示向量,作为解码器的初始输入,告知解码器正在生成何种词牌的宋词。

### 3.3 实验结果评估和对比

由于诗歌生成任务与机器翻译任务相似,因此使

用 BLEU 算法<sup>[29]</sup>自动评估生成的宋词。计算基准方法和本文方法的 BLEU-1, BLEU-2, BLEU-3, BLEU-4 的分数,然后取平均值作为最终的 BLEU 分数。

借鉴文献[1-3]的人工评估方法,从意境、语句流畅性、上下文连贯性和意义等方面判断古诗的质量。每方面满分 5 分,分数越高表示宋词在该评估方面的质量越高。让系统生成 50 首宋词,邀请 10 位硕士及以上的学者为生成的宋词打分,最后取平均值作为最终得分。

本文方法(MACVAE)与 4 个基准方法 ICG<sup>[15]</sup>, TPG<sup>[18]</sup>, SPG<sup>[19]</sup>, SCG<sup>[20]</sup>的自动评估和人工评估结果见表 1。3 个基准方法 TPG, SPG, SCG 都在规划模型基础上改进。首先规划在语义上相互关联的关键词作为主题大纲,依据大纲生成宋词 ICG 未加入主题大纲,通过一个词句生成整首宋词。从表 1 可观察到,这 3 个方法比 ICG 方法的上下文连贯性高,表明加入关键词主题大纲的生成方法对一致性具有重要影响。

表 1 BLEU 评估和人工评估结果

Table 1 BLEU evaluation and artificial evaluation results

方法 Approach	BLEU	人工评估 Artificial evaluation			
		意境 Artistic	流畅性 Fluency	连贯性 Coherence	意义 Meaning
ICG	0.105	2.78	2.65	2.81	2.74
TPG	0.323	3.26	3.05	3.19	3.20
SPG	0.328	3.18	3.16	3.21	3.34
SCG	0.351	3.56	3.23	3.45	3.42
MACVAE	0.363	3.62	3.40	3.66	3.52

TPG 先根据用户输入的关键词得到每句话的最后一个词语,这些词语都押韵且与用户输入相关,旨在提高押韵性。SPG 在规划模型的基础上,另外训练了基于注意力机制的关键词扩展模型来匹配生成模型。SPG 在上下文连贯性上较 TPG 有了进一步提高,表明增加的关键词扩展模型对提高上下文连贯性有一定帮助。

SCG 通过注意力机制的权重提取输入序列的显著信息,用于指导生成的下文与上文连贯。SCG 相比于 SPG 在连贯性上的评分有所提升,表明提取上文词句中的重要词语用于指导下文的生成,对提高上下文连贯性具有重要帮助。

相较于基准方法,本文方法在序列到序列模型的

基础上,引入多注意力组件提取词句中的重要语义,在上下文连贯性上有一定的提升;引入 CVAE 控制文本风格,基本实现了婉约类和豪放类宋词的生成,提高了宋词的意境。

### 3.4 生成样例

生成的婉约词示例如图 4 所示。整首宋词侧重于细节描写且上下文连贯。但第 4 句“留我青山作暮眠”和最后一句“海岸江海与苏烟”描绘广阔的场景,属于豪放词句,原因是风格标注的不完全准确性。

生成的豪放词示例如图 5 所示。整首宋词描绘了广阔的场景,整体上下文连贯,符合豪放风格的宋词。

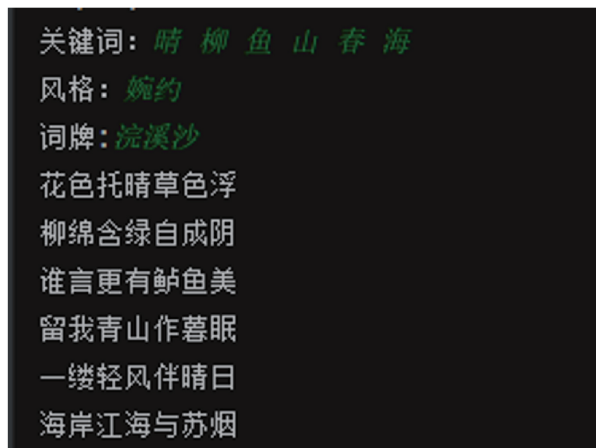


图 4 生成示例 1

Fig. 4 Generated example 1

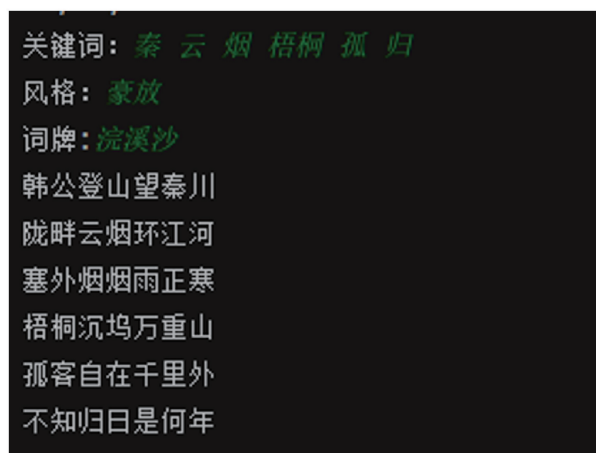


图 5 生成示例 2

Fig. 5 Generated example 2

## 4 结束语

为提高宋词的上下文连贯性及实现风格控制,引入多注意力组件提取词句的多种重要语义,突出词句中的重要词语。另外,引入 CVAE 控制生成宋词风

格。在自制的语料库上进行基准方法和本文方法的对比实验,本文的方法在上下文连贯性上有一定提升。实现了宋词风格的控制,在意境上有所提升。通过分类器标注的风格分类标签并非完全准确,后续期望通过大量的语料库和人工标注方式,进一步提高风格宋词的生成质量。

### 参考文献

- [1] 周昌乐,游维,丁晓君.一种宋词自动生成的遗传算法及其机器实现[J].软件学报,2010,21(3):427-437.
- [2] YAN R,JIANG H,LAPATA M,et al. I,poet:Automatic Chinese poetry composition through a generative summarization framework under constrained optimization [C]//ROSSI F. Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, Beijing: AAAI Press,2013:2197-2203.
- [3] HE J,ZHOU M,JIANG L. Generating Chinese classical poems with statistical machine translation models [C]//HOFFMANN J,SELMAN B. Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, Toronto: AAAI Press,2012:1650-1656.
- [4] TOSA N,OBARA H,MINOH M. Hitch haiku: An interactive supporting system for composing haiku poem [C]//STEVENS S M,SALDAMARCO S J. Entertainment Computing - ICEC 2008,7th International Conference,Pittsburgh:ICEC,2008:209-216.
- [5] WU X F,TOSA N,NAKATSU R. New Hitch Haiku: An interactive Renku poem composition supporting tool applied for sightseeing navigation system [C]//NATKIN S,DUPIRE J. Entertainment Computing - ICEC 2009,8th International Conference, Paris: ICEC,2009:191-196.
- [6] NETZER Y,GABAY D,GOLDBERG Y,et al. Gaiku: Generating Haiku with word associations norms [C]//FELDMAN A,LÖNNEKER-RODMAN B. Proceedings of the Workshop on Computational Approaches to Linguistic Creativity,Stroudsburg,CALC,2009:32-39.
- [7] OLIVEIRA H G. PoeTryMe:A versatile platform for poetry generation [C]//ECAI 2012 Workshop on Computational Creativity,Concept Invention,and General Intelligence (C3GD),Montpellier,France;2012,1:21.
- [8] OLIVEIRA H G,CARDOSO A. Poetry generation with PoeTryMe [M]. Paris:Atlantis Press,2015.
- [9] CHO K,MERRIËNBOER B,GULCEHRE C,et al. Learning phrase representations using RNN encoder - decoder for statistical machine translation [C]//Proceedings of the 2014 Conference on Empirical Methods

- in Natural Language Processing (EMNLP). Doha, Qatar: Association for Computational Linguistics, 2014: 1724-1734.
- [10] GULCEHRE C, CHANDAR S, CHO K, et al. Dynamic neural turing machine with continuous and discrete addressing schemes [J]. *Neural Comput*, 2018, 30(4): 857-884.
- [11] HOPKINS J, KIELA D. Automatically generating rhythmic verse with neural networks [C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vancouver, Canada: Association for Computational Linguistics, 2017: 168-178.
- [12] GHAZVININEJAD M, SHI X, PRIYADARSHI J, et al. Hafez: An interactive poetry generation system [C]//Proceedings of ACL 2017, System Demonstrations ACL. Vancouver, Canada: Association for Computational Linguistics, 2017: 43-48.
- [13] MIKOLOV T, KARAFIÁT M, BURGET L, et al. Recurrent neural network based language model [C]//INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, 2010: 26-30.
- [14] ZHANG X X, LAPATA M. Chinese poetry generation with recurrent neural networks [C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar: Association for Computational Linguistics, 2014: 670-680.
- [15] WANG Q X, LUO T Y, WANG D, et al. Chinese song iambs generation with neural attention-based model [C]//BREWKA G. Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16). New York, NY, United States: AAAI Press, 2016: 2943-2949.
- [16] WANG Z, HE W, WU H, et al. Chinese poetry generation with planning based neural network [C]//Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka, Japan: The COLING 2016 Organizing Committee, 2016: 1051-1060.
- [17] YAN R. I, poet: Automatic poetry composition through recurrent neural networks with iterative polishing schema [C]//BREWKA G. Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. New York, NY, United States: AAAI Press, 2016: 2238-2244.
- [18] GHAZVININEJAD M, SHI X, CHOI Y, et al. Generating topical poetry [C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, TX, United States: Association for Computational Linguistics, 2016: 1183-1191.
- [19] 黄文明, 卫万成, 邓珍荣. 基于序列到序列神经网络模型的古诗自动生成方法[J]. *计算机应用研究*, 2019, 36(12): 3539-3543.
- [20] YI X Y, LI R Y, SUN M S. Chinese poetry generation with a salient-clue mechanism [C]//Proceedings of the 22nd Conference on Computational Natural Language Learning. Brussels, Belgium: Association for Computational Linguistics, 2018: 241-250.
- [21] LIU Z Q, FU Z H, CAO J, et al. Rhetorically controlled encoder-decoder for modern Chinese poetry generation [C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: Association for Computational Linguistics, 2019: 1992-2001.
- [22] MIHALCEA R, TARAU P. TextRank: Bringing order into text [J]//Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. Barcelona, Spain: Association for Computational Linguistics, 2013: 404-411.
- [23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, United States: Curran Associates Inc, 2017: 6000-6010.
- [24] LIN Z H, FENG M W, SANTOS C N, et al. A structured self-attentive sentence embedding [C]//5th International Conference on Learning Representations (ICLR). Toulon, France: 2017: 24-26.
- [25] ZHAO T C, ZHAO R, ESKENAZI M. Learning discourse-level diversity for neural dialog models using conditional variational autoencoders [C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vancouver, BC, Canada: Association for Computational Linguistics, 2017: 654-664.
- [26] SOHN K, LEE H, YAN X C. Learning structured output representation using deep conditional generative models [C]//CORTES C, LEE D D, SUGIYAMA M. Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2. Cambridge, MA, United States: MIT Press, 2015: 3483-3491.
- [27] KINGMA D P, WELLING M. Auto-encoding varia-

- tional bayes [C]//International Conference on Learning Representations, Banff, AB, Canada, 2014; 14-16.
- [28] ZEILER M D. ADADELTA: An adaptive learning rate method [J]. Neural Computing and Applications, 2019, 31(10): 6685-6698.
- [29] PAPANENI K, ROUKOS S, WARD T, et al. BLEU: A method for automatic evaluation of machine translation [C]//Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, Philadelphia, PA, United States; Association for Computational Linguistics, 2002; 311-318.

## Song Ci Generation Model Based on Multi-attention and Conditional Variational Auto-encoder

LIANG Xiao<sup>1</sup>, HUANG Wenming<sup>1,2</sup>, YAO Jun<sup>3</sup>, WEN Yayuan<sup>4</sup>, DENG Zhenrong<sup>1,2</sup>

(1. College of Computer Science & Information Security, Guilin University of Electronic Technology, Guilin, Guangxi, 541004, China; 2. Guangxi Key Laboratory of Image and Graphic, Intelligent Processing, Guilin, Guangxi, 541004, China; 3. Guangxi Zhuang Autonomous Region Higher People Court, Nanning, Guangxi, 530000, China; 4. College of Electronic Engineering, Guangxi Normal University, Guilin, Guangxi, 541004, China)

**Abstract:** The existing poetry generation methods ignore the importance of style. In addition, since there are the large number of Cipai and phrases in most of Song Ci, it is easy to produce the phenomenon that the context lacks coherence in the process of generating Song Ci sentence by sentence, and there is still space for improvement in context coherence. To solve these two problems, based on the encoding and decoding text generation framework, a sentence representation algorithm based on self-attention mechanism is introduced to calculate multi-attention weight matrix, which is used to extract a variety of important semantic features of words and sentences, so that the model can pay more attention to the significant information above to improve context coherence. The conditional variational self-encoder is introduced to transform each Song Ci data into a high-dimensional Gaussian distribution with different style features in the hidden space, and the hidden variables are sampled from the respective distributions to control the style of Song Ci. Since the self-constructed corpus of Song Ci lacks a complete style classification label, a fine-tuning BERT model of Song Ci with style label is used as a style classifier to label all the data of Song Ci. Based on the above key technologies, the generation model of Song Ci is realized, which generates euphemistic poetry and bold poetry with coherent context. Compared with the other four benchmark methods, the results show that the sentence representation algorithm with self-attention mechanism and the conditional variation self-encoder have certain improvements in context coherence and style control.

**Key words:** conditional variational auto-encoder; Song Ci style; generation of Song Ci; Bi-GRU; self-attention mechanism

责任编辑:陆媛峰



微信公众号投稿更便捷

联系电话:0771-2503923

邮箱:gxxk@gxas.cn

投稿系统网址:http://gxxk.ijournal.cn/gxxk/ch