

## 基于 BERT 提示的矿产资源管理规则检测方法研究<sup>①</sup>

胡容波<sup>②</sup>\* \*\* \*\*\* 张广发\* \*\* \*\* 王雅雯\* \*\* \*\* 方金云<sup>③</sup>\*

(\* 中国科学院计算技术研究所 北京 100190)

(\*\* 自然资源部信息中心 北京 100036)

(\*\*\* 中国科学院大学 北京 100190)

**摘要** 政策文本中管理规则检测是一个新兴的自然语言处理任务,在政策冲突检测、政策智能检索、事项合规性检查以及政务系统需求工程等方面具有重要应用价值。本文以矿产资源管理规则检测为研究目标,提出基于转换器的双向编码表征(BERT)提示的政策文本管理规则检测方法。该方法通过构建融入管理规则信息、带有[MASK]标记的提示模板,可以充分发挥掩码语言模型的自编码优势,有效激发 BERT 模型提取与管理规则相关的文本特征,增加模型稳定性;提出基于 BERT 模型进行管理规则检测的新应用模式,放弃使用[CLS]隐向量而采用[MASK]隐向量进行分类预测;在矿产资源管理规则数据集上的实验结果表明,该方法的准确率、宏平均  $F_1$  值、加权平均  $F_1$  值均优于基线方法,在公开数据集上的实验结果也表明了该方法的有效性。

**关键词** 矿产资源;管理规则;文本分类;基于转换器的双向编码表征(BERT);提示学习

### 0 引言

管理规则是法律法规、规章规定等政策文本的基本要素之一,是对各种权利、义务以及相关后果的具体规定。从政策文本中提取管理规则对政策冲突检测<sup>[1]</sup>、政策智能检索<sup>[2]</sup>、事项合规性检查<sup>[3]</sup>、政务系统需求工程<sup>[4]</sup>等均具有重要意义。然而,识别、提取和形式化管理规则是一项知识密集型和劳动密集型的任务<sup>[5]</sup>。为此,可以采用多阶段的管理规则自动抽取方法:(1)对政策文本自动分类,识别目标类别句子;(2)检测目标句子中是否存在管理规则以及判定规则类别;(3)对存在管理规则的政策文本句子根据规则类别抽取相应的规则元素,最终形成规则库。

此前的工作<sup>[6]</sup>探讨了第 1 阶段任务,可实现自然资源政策文本句子自动分类(土地管理、矿产资

源管理、海洋管理等)。本文聚焦于第 2 阶段任务,选取矿产资源管理政策文本句子,对其中的命令类、禁止类、允许类、处罚类等管理规则进行自动检测。

近年来,法律智能<sup>[7]</sup>已成为研究热点,但主要是对法律文书、案件描述等进行解析处理,以辅助开展罪名预测、法条推荐、刑期预测等。目前直接针对政策文本进行管理规则检测的工作还不太多。自动检测政策文本中的管理规则是一个新兴的自然语言处理(natural language processing, NLP)任务,主要挑战在于管理规则大多包含义务、许可、禁止等道义模态(deontic modality)<sup>[8]</sup>,但以自然语言表达的道义模态常常存在模糊和歧义<sup>[9]</sup>。有些有明确的道义词,有些有多个道义词,有些则未出现道义词;有些虽然出现了道义词,但不一定具有道义指示含义;由于自然语言的复杂性,有些道义词形式多变。如“对可以由本地方开发的矿产资源,优先合理开发

① 北京科技攻关项目(A201908230146)和河北省重点研发计划(20310106D)资助项目。

② 男,1979 年生,博士生;研究方向:智能搜索技术;E-mail: hurongbo@sina.com。

③ 通信作者,E-mail: fangjy@ict.ac.cn。

(收稿日期:2022-04-18)

利用”中虽然有“可以”道义词,但该管理规则在语义上并非允许类,而是命令类,但义务道义词“应该”并未出现。另外,根据矿产资源管理需求,本文将处罚类规则单独提出,并将文本处理范围从语言较为严谨的法律扩大到普通政策文件,进一步增加了任务的挑战性。

研究者已提出基于模式匹配<sup>[10]</sup>、基于传统机器学习<sup>[11-12]</sup>以及基于深度学习<sup>[9,13]</sup>的政策文本规则检测方法。近年来,以基于转换器的双向编码表征(bidirectional encoder representation from transformers, BERT)<sup>[14]</sup>为代表的预训练语言模型(pre-trained language model, PLM)在广泛的下游任务中显示出强大性能,已成为 NLP 领域的主流模型之一。在政策文本规则检测任务中,研究者也提出了基于 BERT 模型的应用<sup>[15]</sup>,结果优于其他已有模型。

BERT 模型的优异性能既来源于带有自注意力机制的多层双向 Transformer 编码器,也来源于 2 个新颖的在大规模语料上应用的无监督预训练任务:掩码语言模型(masked language model, MLM)和下一句预测(next sentence prediction, NSP)。其中,MLM 类似完形填空,可实现双向语言建模,模型学习到的掩码(mask)词表示(representation)可以较为精准地融合上下文语义信息。NSP 则使 BERT 具备建模两段文本间关系的能力。

然而,BERT 模型在预训练阶段和下游任务微调阶段的不一致性会影响 BERT 性能的发挥:一方面,预训练阶段引入的[MASK]标记在下游任务中并不会出现,BERT 特意设计了 3 种掩码方式以降低这种不一致带来的影响;另一方面,预训练阶段 BERT 模型输出的[CLS]隐向量主要用于建模文本对之间的关系,而在下游其他句子级任务中进行句子表示时通常也使用[CLS]隐向量。已有研究<sup>[16]</sup>表明,词语频率等带来的词嵌入偏差和对 BERT 网络层的低效利用导致[CLS]隐向量并不能很好地表示句子语义,但基于提示(prompt)的句子表示方法可以缓解这种嵌入偏差。

为此,本文提出基于 BERT 提示的矿产资源管理规则检测方法。该方法在不引入其他机制的情况下,通过改变模型输入层和输出分类层实现对

BERT 模型特点的充分利用。在模型输入层,构建带有[MASK]标记的提示模板,如“应该不得可以惩罚[MASK]:”,再将该提示模板与原始政策文本句子拼接后输入 BERT。提示模板虽然不是流畅的自然语言,但融入了管理规则道义模态知识,可借助 BERT 自身机制探测政策文本句子中与道义模态相关的信息。在分类输出层,使用经过多层 Transformer 编码、蕴含丰富上下文语义信息的[MASK]隐向量进行分类预测。在矿产资源管理规则数据集上的实验结果表明,该方法在模型准确率、宏平均  $F_1$  值、加权平均  $F_1$  值上均优于其他基线方法。在公开数据集上的实验结果也显示了该方法的有效性。

本文主要贡献如下。

(1)提出构建带有[MASK]标记和管理规则信息的提示模板,可以充分发挥掩码语言模型的自编码优势,激发 BERT 模型更好地提取与管理规则相关的文本特征。

(2)提出基于 BERT 模型进行管理规则检测的新应用模式,放弃使用[CLS]隐向量而采用[MASK]隐向量进行分类预测。由于未改变 BERT 的原始网络结构,所提方法几乎不增加计算量。

(3)在矿产资源管理规则数据集上的大量实验结果显示,所提方法优于其他基线方法,在模型稳定性上也更具优势。在公开数据集上的实验结果也显示该方法可以进一步提升 BERT 模型在相关下游任务中的性能。

## 1 相关工作

本节详细阐述与本文相关的研究工作,包括政策文本中管理规则检测方法、BERT 模型应用模式以及 BERT 模型在分类任务中的标签嵌入方法。

### 1.1 政策文本中管理规则检测方法

已有的政策文本中管理规则检测方法包括模式匹配、传统机器学习以及深度学习方法。文献[9]比较了传统机器学习方法与深度学习方法在法律规则(义务、禁止和许可)检测上的性能,其采用的传统机器学习方法包括逻辑回归(logistic regression, LR)、支持向量机(support vector machines, SVM)、决

策树(decision trees, DT)、随机森林(random forest, RF);深度学习方法包括卷积神经网络(convolutional neural networks, CNN)、长短时记忆网络(long short-term memory, LSTM)、双向长短时记忆网络(bi-directional long short-term memory, BiLSTM)等,结果显示深度学习模型性能更优,BiLSTM 取得了最高性能。文献[15]比较了基于模式匹配、基于 BiLSTM 以及基于 BERT 的方法在法规规则(义务、禁止、许可)检测上的性能。结果显示,BERT 优于基于模式匹配(需要人工构建匹配模式,模式过少或过多都会影响性能,且通用性较差)和基于 BiLSTM 的方法,是目前在管理规则检测任务上的最先进方法。

本文在上述工作的基础上,进一步提出基于 BERT 模型进行管理规则检测的新方法。另外,本文针对中文政策领域进行研究,情形更为复杂,这方面的工作目前还较为欠缺。

### 1.2 BERT 模型应用模式

以文本分类任务为例,目前对 BERT 模型有 2 种典型的应用模式。一种是基于微调(fine-tuning)<sup>[14]</sup>的方法,即在 BERT 模型上增加额外的分类器,结合下游具体任务进行微调,这种模式在监督任务上可取得优异性能;另一种是基于提示学习(prompt-based learning)<sup>[17]</sup>的方法,即在输入文本中插入带有[MASK]标记的提示模板,将分类问题转化为 MLM 建模问题。这种应用模式在少样本和零样本场景中取得了更好的性能。然而,基于提示学习的方法对预训练模型的知识要求更高,同时需要进行复杂的模板工程(prompt engineering)和答案工程(answer engineering),模型的性能受这些因素影响较大。

本文提出的方法是一种将 2 种应用模式进行部分结合的方法,既可以充分利用 MLM 建模的优势,又避免了复杂的答案工程。

### 1.3 BERT 模型在分类任务中的标签嵌入方法

对标签信息的利用吸引了不少研究者的关注。文献[18]通过在训练过程中计算实例和标签之间的相似性捕获标签之间的语义重叠,生成更好的标签分布以取代原来的独热标签向量,最终提高分类

性能。但该方法需要增加额外的标签编码器、标签分布计算层等组件。

文献[19]提出了一种更为简洁的标签嵌入方法,通过将标签文本与输入文本拼接后一起送入 BERT 模型进行编码,在不改变原始编码器结构和增加其他机制的情况下,直接利用 BERT 固有的自注意力机制实现标签信息与输入文本的交互,增强文本的上下文表示,提高分类性能。然而,这是一种对标签信息的“硬嵌入”方法,分类性能受标签文本的影响较大。如果标签文本对不同类别数据的区分度不大,则增加标签信息后不仅不能增强文本表示,反而会破坏文本表示,从而导致精度降低。

本文提出的是一种对标签信息“软嵌入”的方法,通过将不同管理规则有代表性的道义词引入提示模板,借助 BERT 模型的 MLM 建模机制,可以更好地利用标签信息。

## 2 模型与方法

基于 BERT 提示的矿产资源管理规则检测方法并不改变 BERT 模型的基本网络结构,即保持 BERT 编码层不变,主要改变是调整输入层和分类输出层。接下来将详细介绍本文提出的模型与方法。

### 2.1 任务定义

本研究任务属于分类任务,可形式化定义为:对于输入的矿产资源政策文本句子  $x = (x_1, x_2, \dots, x_n)$ , 预测其管理规则类别  $y \in \mathbf{y}$ 。其中,  $n$  为政策文本句子长度,  $\mathbf{y}$  为管理规则类别标签集合。

### 2.2 模型架构

图1为基于BERT提示的矿产资源管理规则检

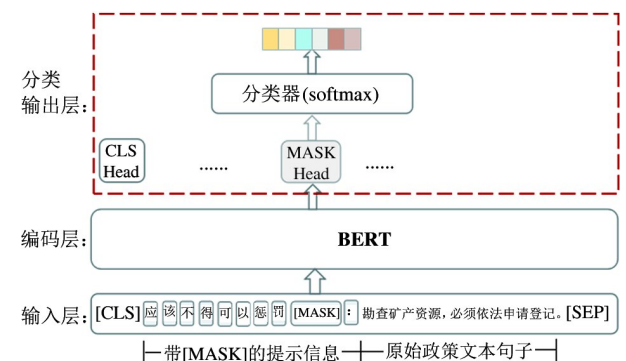


图1 基于 BERT 提示的矿产资源管理规则检测方法整体框架

测方法的整体框架,由输入层、编码层、分类输出层构成。其中,输入层将政策文本句子使用提示模板包装后输入模型。编码层为 BERT 基本网络结构,是由多个双向 Transformer 堆叠而成的深层神经网络,可将输入文本编码为深层上下文语义表示。分类输出层由 Softmax 分类器构成,用于输出模型对政策文本句子中管理规则检测的结果。

### 2.3 建模方法

(1)输入层。对于给定的矿产资源政策文本句子  $x$ ,使用带有[MASK]标记,且融合了管理规则信息(代表性道义词)的提示模板进行包装。如政策文本句子  $x$  为“勘查矿产资源,必须依法申请登记。”,提示模板为“应该不得可以惩罚[MASK]:”,则输入为

$x_p = [\text{CLS}]$  应该不得可以惩罚[MASK];勘查矿产资源,必须依法申请登记。[SEP]

其中,[CLS]是 BERT 模型专为分类任务设定的特殊标记,[SEP]是文本序列之间的分隔标记,[MASK]为提示模板中人为设计的掩码标记。

文本序列  $x_p$  经过分词器分词及映射,获得每个标记(分词单元)的词向量(token embedding)、块向量(segment embedding)以及位置向量(position embedding),3个向量相加得到每个标记的输入向量,拼接后得到整个文本序列  $x_p$  的输入向量  $\mathbf{v}$ 。

(2)编码层。BERT 模型中的多层 Transformer 对输入向量  $\mathbf{v}$  进行逐层编码,通过多头自注意力机制充分学习句子中每个词之间的语义关联<sup>[20]</sup>,最终获得政策文本句子的上下文语义表示  $\mathbf{h}$ :

$$\mathbf{h} = \text{BERT}(\mathbf{v}) \quad (1)$$

式中, $\mathbf{h} \in \mathbb{R}^{N \times d}$ ,为 BERT 模型中最后一层 Transformer 的输出;其中, $N$ 为 BERT 输入文本的最大长度, $d$ 为 BERT 隐藏层维度。

如果仅输出最后一层编码结果,BERT 模型可分别输出[CLS]标记对应的隐向量  $\mathbf{h}_{[\text{CLS}]}$  以及整个文本序列所有标记对应的隐向量  $\mathbf{h}$ 。其中, $\mathbf{h}_{[\text{CLS}]}$  为  $\mathbf{h}$  的首个分量经过全连接层并使用 tanh 函数激活后的结果,通常作为句子表示用于下游分类任务。

(3)分类输出层。本文提出的方法放弃使用  $\mathbf{h}_{[\text{CLS}]}$  作为句子表示,而代之以[MASK]标记对应的

隐向量  $\mathbf{h}_{[\text{MASK}]}$ 。将  $\mathbf{h}_{[\text{MASK}]}$  作为分类输出层的输入,经过 Softmax 分类层,最后输出政策文本句子中管理规则类别的概率分布  $\mathbf{p}$ :

$$\mathbf{p} = \text{Softmax}(\mathbf{W}\mathbf{h}_{[\text{MASK}]} + \mathbf{b}) \quad (2)$$

其中, $\mathbf{p} \in \mathbb{R}^K$ , $\mathbf{W} \in \mathbb{R}^{d \times K}$  为全连接层的权重矩阵, $\mathbf{b} \in \mathbb{R}^K$  为全连接层的偏置, $K$  表示管理规则类别个数。

对于第  $i$  个样本,取概率最大值所对应的管理规则类别作为模型预测类别:

$$\tilde{y}_i = \arg \max_y (\{p_i, \forall y \in y\}) \quad (3)$$

其中, $p_i$  为模型对第  $i$  个样本的预测概率。

模型训练时,以交叉熵损失作为模型优化的目标函数:

$$CE = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^K y_j^{(i)} \log p_j^{(i)} \quad (4)$$

其中, $m$  为样本个数, $y_j^{(i)}$  表示第  $i$  个样本在第  $j$  类上的真实结果,属于该类为 1,否则为 0; $p_j^{(i)}$  表示模型对第  $i$  个样本属于第  $j$  类的预测概率。

## 3 实验与结果分析

本节详细介绍对基于 BERT 提示的矿产资源管理规则检测方法的评估实验,并给出相应分析。

### 3.1 实验数据集

本文从此前工作<sup>[6]</sup>所构建的自然资源政策文本分类数据集中选取业务类别为矿产资源管理的一部分政策文本句子,由领域专家进行人工标注,形成矿产资源管理规则数据集。其中,管理规则类别为:命令类规则、禁止类规则、允许类规则、处罚类规则、其他类规则。表 1 为数据集统计信息。

表 1 矿产资源管理规则数据集统计

管理规则类别	政策文本句子个数
命令类	912
禁止类	168
允许类	198
处罚类	169
其他类	352

本文对数据集中的每一类别样本按 6:2:2 的比例进行划分,组合成训练集、验证集、测试集。随

机划分 10 次,形成 10 组数据集。对所有模型,在这 10 组数据集上进行 10 次评估。

### 3.2 对比模型和方法

本文选择以下用于管理规则检测任务的典型方法及最新 BERT 标签嵌入方法等进行对比分析。

(1) SVM<sup>[21]</sup>: SVM 是传统机器学习的代表性技术之一。文献[9]在英文金融法规道义模态检测上使用了该方法,在所比较的传统机器学习方法中, SVM 性能较高。

(2) 基于卷积神经网络的文本分类学习 (convolutional neural networks for text classification, TextCNN)<sup>[22]</sup>: TextCNN 的基本结构由输入层、卷积层 (convolution layer)、池化层 (pooling layer)、全连接层 (fully connected layer) 和输出层构成。文献[9]在英文金融法规道义模态检测上使用该方法取得了超越传统机器学习方法的效果。

(3) BiLSTM<sup>[23]</sup>: 在应用于文本分类任务时, BiLSTM 可以从 2 个方向 (从前往后、从后往前) 对文本序列进行编码表示。文献[9]在英文金融法规道义模态检测上使用了该方法,在神经网络模型中取得了较高性能。

(4) BERT<sup>[14]</sup>: 基于 BERT 模型的常规微调方法,即输入层的输入为原始政策文本句子,分类输出层将 [CLS] 隐向量作为分类器的输入向量。文献[15]在英文合同条款义务类、许可类、禁止类句子检测任务上应用了该方法,结果显示 BERT 性能明显优于 BiLSTM。该方法也是此前管理规则检测任务的最优方法。

(5) BERT-label<sup>[19]</sup>: 将标签信息融入 BERT 的方法,即输入层的输入为管理规则各类别标签词拼接原始政策文本句子组成的新文本序列,分类输出层将 [CLS] 隐向量作为分类器的输入向量。该方法可以利用标签嵌入提高 BERT 在文本分类中性能。

(6) BERT-tfidf<sup>[19]</sup>: 在 BERT-label 的基础上,计算各类别语料分词后的词频-逆文档频率 (tf-idf) 值,每个类别各取 5 个 tf-idf 值最高的分词对标签词进行扩展,输入层拼接方法及分类输出层与 BERT-label 相同。这是文献[19]提出的进一步优化方法。

(7) BERT-提示 [CLS]: 输入层的输入为提示模板去

掉 “[MASK]:” 后与原始政策文本句子拼接组成的新文本序列,分类输出层将 [CLS] 隐向量作为分类器输入向量。本方法专为检验使用 [CLS] 隐向量和 [MASK] 隐向量进行管理规则检测的性能差异设计。

以上基于 BERT 模型的方法分类输出层均采用 Softmax 分类器。

### 3.3 评价指标

对于单个管理规则类别的检测性能,采用  $F_1$  值作为评价指标。

$$R_i = \frac{TP_i}{TP_i + FN_i} \quad (5)$$

$$P_i = \frac{TP_i}{TP_i + FP_i} \quad (6)$$

$$F_1 = \frac{2P_i \times R_i}{P_i + R_i} \quad (7)$$

其中,  $R_i, P_i, F_{1i}$  分别表示第  $i$  类的召回率、精确率和  $F_1$  值,  $TP_i, FP_i, FN_i$  分别表示模型预测的第  $i$  类真正例、假正例、假负例个数。

对于模型整体性能,采用准确率 (accuracy)、宏平均  $F_1$  值和加权平均  $F_1$  值进行评价。

$$Accuracy = \frac{\sum_{i=1}^N (TP_i + TN_i)}{\sum_{i=1}^C (TP_i + FP_i + TN_i + FN_i)} \quad (8)$$

$$Macro\_F_1 = \frac{2P_{macro} \times R_{macro}}{P_{macro} + R_{macro}} \quad (9)$$

其中,  $Accuracy, Macro\_F_1$  分别表示模型的准确率和宏平均  $F_1$  值。  $P_{macro} = \frac{\sum_{i=1}^N P_i}{N}, R_{macro} = \frac{\sum_{i=1}^N R_i}{N}$ ,  $TN_i$  表示模型预测的第  $i$  类真负例个数,  $N$  为类别个数。

$$Weighted\_F_1 = \frac{2P_{weighted} \times R_{weighted}}{P_{weighted} + R_{weighted}} \quad (10)$$

其中,  $Weighted\_F_1$  表示模型的加权平均  $F_1$  值。

$$P_{weighted} = \frac{\sum_{i=1}^N (P_i \times w_i)}{N}, R_{weighted} = \frac{\sum_{i=1}^N (R_i \times w_i)}{N},$$

$w_i = \frac{Number_i}{Number}$ ,  $Number_i$  为第  $i$  类样本个数,  $Number$  为样本总数,  $N$  为类别个数。

### 3.4 实验设置

SVM 模型使用 word2vec 生成文本向量表示,核函数使用 RBF,  $C = 10$ ,  $\gamma$  取默认值。

TextCNN 及 BiLSTM 模型使用文献[24]开源的中文词向量(人民日报 Word + Character + Ngram 300d)进行初始化。学习率(learning\_rate)为 0.001,文本固定长度(pad\_size)为 128,批大小(batch\_size)为 64。TextCNN 的卷积核大小设置为 2、3、4,每个尺寸的卷积核数量为 256,迭代次数(epoch)为 20。BiLSTM 的隐藏层大小为 384,epoch 为 60。

BERT 模型使用 BERT-Base-Chinese 预训练模型,隐藏层大小为 768,丢弃率(dropout)为 0.1, batch-

size 大小为 32, pad\_size 为 128,选择 AdamW<sup>[25]</sup> 作为优化器, learning\_rate 为 0.00005, epoch 为 10。主实验提示模板为“应该不得可以惩罚[MASK]”。

实验环境:操作系统为 Linux, CPU 为 12 核 Intel(R) Xeon(R) Gold 5320 CPU@2.20 GHz, 内存为 32 GB, GPU 为 1 块 RTX A4000, 显存为 16 GB。

### 3.5 实验结果及与基线模型/方法对比

本文报告了基于 BERT 提示的矿产资源管理规则检测方法和其他基线方法在 10 组随机划分的矿产资源管理规则数据集上的详细测试性能(表 2)。表中数值为各方法 10 次评估的模型准确率、宏平均  $F_1$  值、加权平均  $F_1$  值的平均值  $\pm$  标准差,粗体字表示较好结果。可以得出如下结论。

表 2 BERT 提示方法与基线方法的对比实验结果

方法	准确率	宏平均 $F_1$ 值	加权平均 $F_1$ 值
SVM	0.714 53 $\pm$ 0.000 05	0.555 79 $\pm$ 0.001 24	0.673 13 $\pm$ 0.000 27
TextCNN	0.954 37 $\pm$ 0.000 33	0.944 70 $\pm$ 0.000 43	0.954 11 $\pm$ 0.000 32
BiLSTM	0.890 15 $\pm$ 0.001 12	0.859 83 $\pm$ 0.001 81	0.890 21 $\pm$ 0.001 10
BERT	0.986 56 $\pm$ 0.000 06	0.982 91 $\pm$ 0.000 08	0.986 54 $\pm$ 0.000 06
BERT-label	0.986 56 $\pm$ 0.000 02	0.982 25 $\pm$ 0.000 06	0.986 54 $\pm$ 0.000 02
BERT-tfidf	0.982 92 $\pm$ 0.000 10	0.977 81 $\pm$ 0.000 18	0.982 90 $\pm$ 0.000 10
BERT-提示[CLS]	0.986 00 $\pm$ 0.000 03	0.980 71 $\pm$ 0.000 11	0.985 95 $\pm$ 0.000 04
BERT 提示	<b>0.991 60 <math>\pm</math> 0.000 05</b>	<b>0.989 75 <math>\pm</math> 0.000 07</b>	<b>0.991 60 <math>\pm</math> 0.000 05</b>

(1)总体上,基于 CNN、循环神经网络(recurrent neural network, RNN)的深度学习模型在管理规则检测任务上的性能显著优于传统机器学习方法,这主要是因为传统机器学习方法捕获政策文本语义的能力有限。基于 BERT 的方法又明显优于基于 CNN、RNN 的方法,这主要得益于 BERT 模型有更强大的双向语言表征能力,可以有效捕获政策文本句子中的道义模态特征。

(2)本文提出的 BERT 提示方法在准确率、宏平均  $F_1$  值、加权平均  $F_1$  值上均优于已有方法。其中, BERT-提示[CLS] 方法的性能不仅低于 BERT 提示方法,还低于 BERT 方法。这说明即使是相同的提示信息,不同的利用方式甚至会对模型性能带来完全不同的影响。同时说明 BERT 提示方法的有效性不仅来自于提示信息,也来自于对[MASK]隐向量

的利用。

另外,所有方法的宏平均  $F_1$  值均低于准确率,说明各方法对小样本类别(禁止类、允许类、处罚类)的预测准确性都低于相对多样本类别(命令类、其他类)。但 BERT 提示方法在宏平均  $F_1$  值和准确率上的差值最小,显示 BERT 提示方法处理样本不平衡问题的能力更强。

(3)将标签词融入 BERT 的方法(BERT-label)相对于常规 BERT 方法几乎没有提升,仅在准确率、加权平均  $F_1$  值的方差上比 BERT 方法略小,而在宏平均  $F_1$  值上反而低于 BERT 方法。这与标签词(“命令类规则、禁止类规则、允许类规则、处罚类规则、其他类规则”)对管理规则的区分度不大有关。文献[19]指出,如果标签词对类别的区分度不足,则该方法可能会破坏文本表示而不是增强文本表示,

从而导致分类精度降低。

(4)使用 tf-idf 值较高的词扩展标签词的方法 (BERT<sub>-idf</sub>) 也没有获得性能提升,反而在准确率、加权平均  $F_1$  值、宏平均  $F_1$  值上均有不同程度下降。这些词的加入显然进一步破坏了政策文本表示。该方法在应用于情感极性分类任务时也有类似性能下降的情况<sup>[19]</sup>。

以上对比实验中,BERT 提示方法与其他基于 BERT 的基线方法最大的区别在于使用不同的隐向量进行分类预测。BERT 模型的自注意力机制以及预训练任务让 [CLS] 隐向量擅长捕捉句子级上下文信息,[MASK] 隐向量擅长捕捉标记(词)级上下文信息。文献[16]指出,通过使用提示模板的方法来获取 BERT 中的句子表示可以避免嵌入偏差,并且能够更好地利用 BERT 中的网络结构,从而可以在相关下游任务中获得更好的性能。文献[16]显示 BERT 模型的 [CLS] 隐向量与基于提示的 [MASK] 隐向量在语义相似度任务中性能不同,本文实验则显示两者在应用于下游分类任务时也存在差异。合理构建提示模板有助于 [MASK] 隐向量捕获更多与具体任务相关的特征,因而有望进一步提升 BERT 模型的性能。

### 3.6 不同提示模板的影响

为了探索不同提示模板对模型性能的影响,本文设计了多组硬提示 (hard prompt)、软提示 (soft prompt) 模板进行实验。表 3 列出其中有代表性的几组实验结果。其中“[unused1]”、“[unused2]”、“[unused3]”为 BERT-base 预训练模型词汇表中预留的特殊标记,准确率为 10 次评估结果的平均值,粗体字表示较好结果。

表 3 使用不同提示模板实验结果

模板编号	提示模板	准确率
1	“这种行为是[MASK]的:”	0.989 08
2	“应该不得可以惩罚[MASK]:”	0.991 60
3	“命令禁止允许处罚其他[MASK]:”	0.987 68
4	“[unused1][MASK]:”	0.989 36
5	“[unused1][unused2][MASK]:”	0.986 56
6	“[unused1][unused2][unused3][MASK]:”	0.990 48

模板 1 更符合自然语言的流畅性,但并未取得最高性能。模板 2 只是将有代表性的道义词进行拼接,反而获得较好结果。这些道义词借助 MLM 机制激发 BERT 模型通过 [MASK] 隐向量更好地捕获上下文中与规则类别相关的信息。模板 3 的文字部分为标签词,性能不及模板 1,再次显示标签词对管理规则的区分度有限。模板 4~6 为所谓软提示模板,实验结果总体性能尚可,但并未取得最高性能。

这几组提示模板的准确率均高于常规 BERT 方法,但如果提示模板构建不当也会导致准确率不及 BERT 方法。

### 3.7 不同优化器的影响

本文实验数据集的样本量较少。文献[26]指出,BERT 模型中使用的优化器 ( BertAdam) 未进行梯度偏差校正,导致 BERT 模型在小数据集上的训练效率较低,模型不稳定。AdamW<sup>[25]</sup> 优化器则可以对梯度偏差进行校正,使模型训练时能更快收敛,性能更加稳定。本文在矿产资源规则数据集上对 BERT 方法和基于 BERT 提示的方法分别使用 2 种优化器进行实验对比。表 4 为 10 次评估结果的准确率平均值。

表 4 使用不同优化器实验结果

方法	BertAdam 优化器	AdamW 优化器	提升
BERT	0.972 00	0.986 56	1.46%
BERT 提示	0.986 84	0.991 60	0.48%

可以看出,AdamW 优化器为 2 种方法都带来了性能提升。其中,对 BERT 方法的提升更为明显。但是,即使采用 BertAdam 优化器,BERT 提示方法的准确率也高于 BERT 方法采用 AdamW 优化器的准确率。这显示 BERT 提示方法在模型稳定性上更具优势。

### 3.8 案例分析

表 5 显示了 BERT 方法与 BERT 提示方法对矿产资源管理规则各类别的具体检测性能,各指标均为 10 次评估结果的平均值。在矿产资源管理规则检测上,无论是 BERT 方法还是 BERT 提示方法都具有较高性能。在使用 BERT 提示方法后,所有类

别的  $F_1$  值均有提升,其中处罚类规则和允许类规则的  $F_1$  值提升相对较大,分别提升 1.03% 和 0.95%,命令类规则和其他类规则的  $F_1$  值提升相对较小,分别提升 0.33% 和 0.35%。

表 5 2 种方法对不同类别管理规则实验结果对比 ( $F_1$  值)

	BERT	BERT 提示	提升
命令类规则	0.991 22	0.994 50	0.33%
禁止类规则	0.980 32	0.987 92	0.76%
允许类规则	0.971 44	0.980 97	0.95%
处罚类规则	0.985 16	0.995 43	1.03%
其他类规则	0.986 39	0.989 90	0.35%

案例分析发现,BERT 提示方法的有效性在语义特征较强、样本数较少的类别上更为明显。如“整合后形成的矿井只能有一套生产系统,选用先进开采技术和先进装备,杜绝一矿多井或一矿多坑。”,由于“杜绝”一词在样例中不多,BERT 方法难以提取到相关特征,在多次评估中有时将其预测为允许类规则,有时将其预测为命令类规则,而 BERT 提示方法则能将其正确预测为禁止类规则。再如“油气探矿权人发现可供开采的油气资源的,在报告有登记权限的自然资源主管部门后即可进行开采。”,句子中的 2 个“可”对管理规则类别的意义不同,BERT 方法将其错误预测为命令类规则,BERT 提示方法正确预测为允许类规则。

### 3.9 在公开数据集上的实验

为进一步验证 BERT 提示方法的有效性,本文在公开数据集上进行了实验。管理规则检测任务场景具有 2 个主要特点:(1)数据量不够充足;(2)管理规则具有道义模态,但道义词存在缺失、模糊和歧义情形,需要深层语义理解。由于缺乏管理规则检测公开数据集,本文选择与管理规则检测场景相对接近的 ChnSentiCorp 数据集<sup>[27]</sup>和垃圾短信检测数据集<sup>[28]</sup>进行实验。

本文重点对比了 BERT 提示方法和常规 BERT 方法在分类任务上的性能,其中 BERT 提示方法分别构建硬提示、软提示 2 类提示模板。采用准确率作为性能评估指标,准确率为采用 10 个随机种子进行 10 次评估的平均值。

(1) ChnSentiCorp 数据集上的实验。该数据集为经典的句子级情感分类数据集,训练集、开发集、测试集大小分别为 9600、1200、1200,包含积极、消极 2 个类别。文献[19]实验显示,在此类情感极性分类任务上将标签信息融入 BERT 的方法难以发挥作用。

表 6 为在 ChnSentiCorp 数据集上的评估结果,BERT 提示方法的准确率均高于 BERT 方法,其中软提示模板带来的提升更大。

表 6 在 ChnSentiCorp 数据集上的实验结果

方法	提示模板	准确率
BERT		0.931 58
BERT 提示-1	“对此感觉[MASK]:”	0.937 75
BERT 提示-2	“[unused1][unused2][MASK]:”	0.939 83

(2) 垃圾短信检测数据集上的实验。本文从“带标签短信”中随机采样 5000 个正常短信和 5000 个垃圾短信,按 8:1:1 的比例划分为训练集、开发集、测试集。

表 7 为在垃圾短信检测数据集上的评估结果,BERT 提示方法的准确率同样高于 BERT 方法,其中硬提示模板带来的提升更大。但如果将提示模板改为“垃圾信息[MASK]:”,则准确率只有 0.988 40,反而略低于 BERT 方法。这显示需要合理构建提示模板才能充分发挥 BERT 提示方法的作用。

表 7 在垃圾短信检测数据集上的实验结果

方法	提示模板	准确率
BERT		0.988 80
BERT 提示-1	“公司优惠欢迎联系[MASK]:”	0.991 60
BERT 提示-2	“[unused1][unused2][MASK]:”	0.991 40

在 2 个公开数据集上的实验结果表明,本文提出的 BERT 提示方法依然有效。

## 4 结论

本文提出了基于 BERT 提示的矿产资源管理规



则检测新方法。该方法未改变 BERT 的原始网络结构,仅对模型输入层和输出分类层进行改造,几乎不增加计算量。通过引入带有 [MASK] 标记和管理规则信息的提示模板,一定程度上弥合了 BERT 预训练和下游任务之间的不一致性,可以充分发挥掩码语言模型的自编码优势,更好地利用标签信息。在分类输出层,放弃使用 [CLS] 隐向量而采用 [MASK] 隐向量进行分类预测,可以激发 BERT 模型更好地提取与管理规则相关的文本特征,从而进一步提高检测精度。在矿产资源管理规则数据集上的广泛实验结果表明,本文所提方法优于已有方法,而且在处理样本不均衡问题以及模型稳定性上更具优势。在公开数据集上的进一步实验结果也显示了本文方法的有效性,可以为相关工作提供参考和借鉴。

#### 参考文献

- [ 1 ] LI T, BALKE T, DE VOS M, et al. Legal conflict detection in interacting legal systems [ M ]. Amsterdam: IOS Press, 2013:107-116.
- [ 2 ] BOELLA G, DI CARO L, HUMPHREYS L, et al. NLP challenges for eunomos, a tool to build and manage legal knowledge [ C ] // The 8th International Conference on Language Resources and Evaluation. Istanbul: LREC, 2012:3672-3678.
- [ 3 ] SONG J, LEE J K, CHOI J, et al. Deep learning-based extraction of predicate-argument structure ( PAS ) in building design rule sentences [ J ]. Journal of Computational Design and Engineering, 2020,7(5):563-576.
- [ 4 ] SLEIMI A, SANNIER N, SABETZADEH M, et al. An automated framework for the extraction of semantic legal metadata from legal texts [ J ]. Empirical Software Engineering, 2021,26(3):1-50.
- [ 5 ] WYNER A, PETERS W. On rule extraction from regulations [ M ]. Amsterdam: IOS Press, 2011:113-122.
- [ 6 ] 胡容波,郭诚,王锦浩,等. 基于深度学习的自然资源政策文本分类研究 [ J ]. 高技术通讯, 2023,33(7):692-703.
- [ 7 ] ZHONG H, XIAO C, TU C, et al. How does NLP benefit legal system: a summary of legal artificial intelligence [ C ] // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Seattle: Association for Computational Linguistics, 2020:5218-5230.
- [ 8 ] 王涌. 道义逻辑、人工智能与法律——霍菲尔德法律  
关系形式理论的应用 [ J ]. 经贸法律评论, 2020(2):73-83.
- [ 9 ] NEILL J O, BUTTELAAR P, ROBIN C, et al. Classifying sentential modality in legal language: a use case in financial regulations, acts and directives [ C ] // Proceedings of the 16th Edition of the International Conference on Artificial Intelligence and Law. London: Association for Computing Machinery, 2017:159-168.
- [ 10 ] BREAUX T D, ANTON A I, VAIL M W. Towards regulatory compliance: extracting rights and obligations to align requirements with regulations [ C ] // Proceedings of the IEEE International Conference on Requirements Engineering. Minneapolis/St. Paul: IEEE, 2006:49-58.
- [ 11 ] FRANCESCONI E, PASSERINI A. Automatic classification of provisions in legislative texts [ J ]. Artificial Intelligence and Law, 2007,15(1):1-17.
- [ 12 ] ESPEJO-GARCIA B, MARTINEZ-GUANTER J, PÉREZ-RUIZ M, et al. Machine learning for automatic rule classification of agricultural regulations: a case study in Spain [ J ]. Computers and Electronics in Agriculture, 2018,150:343-352.
- [ 13 ] CHALKIDIS I, ANDROUTSOPOULOS I, MICHOS A. Obligation and prohibition extraction using hierarchical RNNs [ C ] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne: Association for Computational Linguistics, 2018:254-259.
- [ 14 ] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [ C ] // 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: Association for Computational Linguistics, 2019:4171-4186.
- [ 15 ] JOSHI V, ANISH P R, GHASIAS S. Domain adaptation for an automated classification of deontic modalities in software engineering contracts [ C ] // Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. Athens: Association for Computing Machinery, 2021:1275-1280.
- [ 16 ] JIANG T, HUANG S, ZHANG Z, et al. PromptBERT: improving BERT sentence embeddings with prompts [ EB/OL ]. ( 2022-01-12 ) [ 2022-02-27 ]. <http://arxiv.org/pdf/2201.04337.pdf>.
- [ 17 ] DING N, CHEN Y, HAN X, et al. Prompt-learning for fine-grained entity typing [ EB/OL ]. ( 2021-08-24 ) [ 2022-02-27 ]. <http://arxiv.org/pdf/2108.10604.pdf>.

- [18] GUO B, HAN S, HAN X, et al. Label confusion learning to enhance text classification models [ C ] // The 35th AAAI Conference on Artificial Intelligence. Virtual: AAAI Press, 2021 :12929-12936.
- [19] XIONG Y, FENG Y, WU H, et al. Fusing label embedding into BERT: an efficient improvement for text classification [ C ] // Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. Virtual: Association for Computational Linguistics, 2021 :1743-1750.
- [20] 车万翔, 郭江, 崔一鸣. 自然语言处理: 基于预训练模型的方法 [ M ]. 北京: 电子工业出版社, 2021 :2-3.
- [21] CORTES C, VAPNIK V. Support-vector networks [ J ]. Machine Learning, 1995, 20(3) :273-297.
- [22] KIM Y. Convolutional neural networks for sentence classification [ C ] // 2014 Conference on Empirical Methods in Natural Language Processing. Doha: Association for Computational Linguistics, 2014 :1746-1751.
- [23] GRAVES A, JAITLY N, MOHAMED A. Hybrid speech recognition with deep bidirectional LSTM [ C ] // 2013 IEEE workshop on Automatic Speech Recognition and Understanding. Olomouc: IEEE, 2013 :273-278.
- [24] LI S, ZHAO Z, HU R, et al. Analogical reasoning on Chinese morphological and semantic relations [ C ] // The 56th Annual Meeting of the Association for Computational Linguistics. Melbourne: Association for Computational Linguistics, 2018 :138-143.
- [25] LOSHCHILOV I, HUTTER F. Fixing weight decay regularization in Adam [ EB/OL ]. ( 2017-11-14 ) [ 2022-02-27 ]. <http://arxiv.org/pdf/1711.05101.pdf>.
- [26] ZHANG T, WU F, KATIYAR A, et al. Revisiting few-sample BERT fine-tuning [ EB/OL ]. ( 2017-03-11 ) [ 2022-02-27 ]. <http://arxiv.org/pdf/2006.05987.pdf>.
- [27] TAN S, ZHANG J. An empirical study of sentiment analysis for chinesedocuments [ J ]. Expert Systems with applications, 2008, 34(4) :2622-2629.
- [28] Hrwhisper. 中文垃圾短信识别 ( 手写分类器 ) [ EB/OL ]. ( 2018-12-08 ) [ 2022-03-25 ]. <http://github.com/hrwhisper/SpamMessage>.

## Research on detection method of mineral resources regulatory rules based on BERT with prompts

HU Rongbo<sup>\* \*\* \*\*\*</sup>, ZHANG Guangfa<sup>\* \*\* \*\*</sup>, WANG Yawen<sup>\* \*\* \*\*</sup>, FANG Jinyun<sup>\*</sup>

(<sup>\*</sup> Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

(<sup>\*\*</sup> Information Center of Ministry of Natural Resources, Beijing 100036)

(<sup>\*\*\*</sup> University of Chinese Academy of Sciences, Beijing 100190)

### Abstract

Regulatory rules detection in policy text is an emerging natural language processing task, which has important application value for policy conflict detection, policy intelligent retrieval, regulatory compliance inspection, and e-government system requirements engineering. This paper takes the detection of mineral resources regulatory rules as the research goal, and proposes a detection method based on bidirectional encoder representation from transformers (BERT) with prompts. By constructing a prompt template with [ MASK ], which incorporates regulatory rules information, the proposed method can give full play to the auto-encoding advantages of the mask language model, effectively stimulate BERT model to extract text features related to regulatory rules and increase the stability of the model. A new application mode of regulatory rules detection based on BERT model is proposed, which uses the [ MASK ] hidden vector instead of the [ CLS ] hidden vector for classification and prediction. The experimental results on the dataset of mineral resources regulatory rules show that the accuracy, macro-average  $F_1$  score and weighted-average  $F_1$  score of this method are better than the baseline methods. The experimental results on the public dataset also show the effectiveness of the proposed method.

**Key words:** mineral resources, regulatory rule, text classification, bidirectional encoder representation from transformers (BERT), prompt-based learning