

面向法律文书的量刑预测方法研究

谭红叶 张博文 张虎 李茹

(山西大学 计算机与信息技术学院, 山西省 太原市 030006)

摘要: 大规模法律文书数据为智能司法审判研究提供了重要的数据基础。量刑预测是智能司法审判中的一个关键环节,对维护司法审判的公平与公正具有重要意义。该文首先基于区间划分和多模型投票方法进行了量刑预测初探,发现区间划分策略可以有效缓解刑期类别众多和数据不平衡问题;在此基础上,该文又采用基于量刑属性的预测方法来充分理解量刑情节。在 CAIL2018 评测数据上的实验表明:该文所提两种方法性能明显超过其他基线系统。

关键词: 量刑预测; 区间划分; 多模型投票; 量刑属性

Automatic Sentencing Prediction for Legal Texts

Abstract: Large-scale legal documents provide the data base for Intelligent Judicial Adjudication research. Sentencing prediction is a key task in intelligent judicial adjudication, which is significant for maintaining legal justice. Firstly, we probes into sentencing prediction based on interval partition and multi-models voting method. Then we found the strategy of interval division can effectively alleviate the problems of excessive types of sentence and data imbalance. Further, we also use the prediction method based on sentencing attributes to fully understand the circumstances of sentencing on this basis. Experiments on the dataset, provided by 2018 Competition of AI and Law in China (CAIL2018), show that our models obviously outperform all the other baselines.

Key words: Sentencing Prediction; Interval Partition; Multi-models Voting; Penalty Attributes

0 引言

法律人工智能的研究实践主要包括三方面:

(1) 利用人工智能技术完成耗时耗力的简单重复性工作,提高司法人员的工作效率。如:庭审笔录的自动生成、证据的检索与展示、电子卷宗的随案生成、审结报告的自动回填。(2) 利用人工智能技术辅助进行司法审判,如:类案推送、量刑建议、证据审查等,实现最高人民法院提倡的同案同判,提升司法审判的公平与公正。(3) 利用人工智能技术开展司法相关便民服务。如:概念化指引服务、诉讼风险评估等,改善公共法律服务水平。在这些研究中,最受关注的是智能司法审判。2018 年由中国最高人民法院信息部和清华大学共同举办的“2018 中国法研杯法律智能挑战赛”(CAIL2018),吸引了学术界和工业界 300

余个团队参加。该评测包括司法审判中的 3 个核心子任务:适用法条预测、罪名预测与刑罚预测,旨在促进刑事案件审判的自动化与智能化。

在有关政策与相关评测的推动下,人工智能技术在司法审判方面已取得一些进展,但司法审判是一个复杂的问题,不仅因为语言本身存在模糊性、法律规范存在灵活解释等原因,而且大量法律问题不存在标准答案,必须结合法律规范的解释、案件的具体情况以及当事人的特殊性等方面作出分析与权衡,才能做出合理公正的审判。

本文以自动量刑(刑罚预测)为研究对象,按照法官量刑的逻辑,基于裁判文书大数据,给定案情描述及罪名,利用深度学习模型预测量刑建议,探索深度学习在智慧司法审判中的应用。本文方法主要思想与特点如下:(1) 采用区间划分和多模型投票模型方法对量刑预测进行初步探索,具体来说将区间划分是利用真实数据的刑

收稿日期:; 定稿日期:

基金项目:国家社会科学基金项目(18BYY074)。

期分布和刑期预测评价指标,将刑期区间划分为若干区间并取区间中位数作为类别,多模型投票结合了不同模型的特点。(2)为了使得量刑预测模型更加合理,本文还提出基于刑期属性预测模型,具体来说选取了34个罪名并对这些罪名进行刑期属性标注,然后将这三个判案依据结合起来进行模型训练。模型基于CAIL2018中公布的裁判文书大数据训练得到。相关实验表明:本文所提方法预测效果超过所有基线系统。

1 相关研究概述

研究者关于人工智能与法的研究已经开展了几十年。从上世纪五、六十年代开始,研究者探索用数学方法对司法行为进行定量分析及预测的可能性。Kort^[1]分析了案件中影响决策的事实要素,并利用数学公式对这些要素打分。Ulmer^[2]从司法决策与情境因素之间的关系,合议庭决策的有效性、正确性与决策成员数目之间的关系出发,探讨了使用定量方法与概率理论进行司法分析的可能性。Keown^[3]总结了司法判决预测中的数学模型,重点讨论了前人基于线性模型方法与最近邻方法在小规模特定案件集上的研究进展。Segal^[4]针对上世纪1962年至1981年间的搜查与扣押案件,提出一个带有变量的法律模型对法院做出搜查或扣押决定的合理性进行解释。

在上世纪八、九十年代,研究者实现了一些基于规则的专家系统来辅助进行司法行为预测。Shapira^[5]为以色列的缓刑监督官员实现了一个青少年缓刑决策专家系统。Hasset^[6]针对英国司法中嫌疑人保释决定缺乏统一性的局限,用规则描述了法官进行保释决定的方式,实现了一个基于规则的保释决策系统。总体来说,基于规则的方法主要局限为:系统使用的规则只反映了构建规则的人对法律的理解,与其他人的法律解释可能不一致;抽取的规则存在语义模糊现象,随着规则数目增加,规则冲突严重。

随着机器学习技术的发展,越来越多的学者通过提取一些特征采用文本分类的思想进行司法预测。Lin等^[7]针对抢劫罪和恐吓罪定义了21个法律要素来刻画犯罪构成要件(犯罪客体、客观方面、主体、主观方面),利用逻辑模型树与SVM模型进行罪名预测。Aletras等^[8]基于

n-gram与主题特征,利用SVM二元分类模型对案件是否违反人权公约条款进行预测。Sulea等^[9]基于大规模案件语料库,采用多个SVM的集成模型对法国最高法院案件的裁判结果、案件类别、裁判日期进行预测。Katz等^[10]从案件要素(法院ID、裁判周期、辩论月、申诉人、答辩人、法院管辖方式等)中提取潜在特征信息,构建了一个随时间演化的随机森林分类器。

近年来随着大数据的拥有与计算能力的提高,深度学习在图像识别及自然语言处理领域得到成功应用,因此也被研究者用来解决司法审判问题,尤其是定罪与量刑这些司法审判的关键环节问题。Luo等人^[11]提出了一种基于注意力机制的神经网络方法,在统一框架内对罪名预测与法律条款的提取联合进行建模,对不同表达风格的案件的罪名进行有效预测。Hu等人^[12]定义了10个具有判别性的法律属性为易混淆罪名提供有效信息,在此基础上提出一个属性-注意力罪名预测模型,显著提升低频罪名与易混淆罪名的预测精度。Ye等人^[13]提出一个带有注意力机制的序列到序列(seq2seq)模型,基于刑事案件事实描述与罪名标签编码,生成可区分罪名与可解释的法院观点。Zhong^[14]等人提出拓扑多任务学习框架,将法条、罪名与刑期之间的依赖关系融合到统一的司法判决框架中,在所有预测子任务上取得了一致和显著的提升。

可以看出,已有的智能司法审判研究大多关注罪名预测,仅有少数学者做了自动量刑研究。如:Zhong^[14]同时解决法条、罪名与刑罚预测三个任务,但刑期预测的准确率最好只有38.3%。

2 方法

本文将自动量刑任务形式化描述为一个分类任务:给定一个案情描述序列 $X = \{x_1, x_2, \dots, x_m\}$,预测与序列 X 相对应的刑罚类别 $y \in Y$ 。其中 m 是序列 X 的长度, x_i 是序列中第 i 个词, $Y = \{y_1, y_2, \dots, y_l\}$ 为区间划分后的刑罚标签集合, l 为刑罚类别总数。任务定义中,考虑了案件的罪名,因为根据上述的量刑定义可知,定罪是量刑的前提,只有定罪准确,量刑才能有正确的可能。本文尝试了两种量刑预测方法,一种是基于区间划分和多模型投票方法(Interval Partition and

Multi-models Voting, IP&MV), 通过实验发现区间划分可以缓解刑期类别过多、数据分布不平衡问题;针对量刑中未能考虑罪名和相关法条信息又提出基于刑期属性的量刑预测方法 (Sentencing based on Penalty Attributes, SPA), 使得量刑模型更符合法官判案逻辑。

2.1 基于区间划分和多模型投票模型

2.1.1 偏差区间划分

通过分析数据, 我们发现: 在 CAIL 2018 数据中刑期是以月份为单位, 因此有期徒刑 0-25 年会有将近 300 个类别。由于刑期是具有顺序和大小关系的一种数据, 因此对于刑期的整体区间可以结合真实数据分布和评价指标进行划分。

数据集中刑期分布如图 1 所示。其中, 横坐标为刑期, 纵坐标为该刑期的案例个数。由图 1 可以看出, 刑期分布属于典型的“2-8”分布, 即大约 80% 的案例都集中在 0-18 个月这些类别上。

本文通过区间划分来减少类别和平衡数据分布。常见的区间划分方法有等间距法、极大熵法与等信息量法。其中, 等间距法是将整体区间划分成若干个间隔相等的区间; 极大熵法是使用了一个使熵最大化、信息损失最小化的标准来划分区间; 等信息量法是由用户提出适当的区间个数, 然后将获得的所有信息均匀地分布在每个区间中, 从而确定最佳的区间边界。

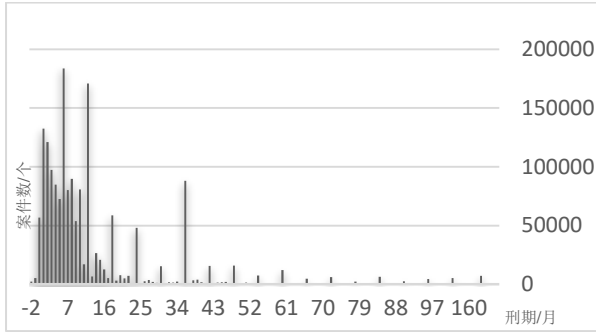


图 1 数据集中的刑期分布

本文提出一种偏差区间划分方法, 结合刑期分布和评价指标对刑期进行划分, 不仅在一定程度上平衡了数据分布, 而且可以将预测结果和真实结果的偏差降低到最小。

刑期预测的评价指标偏差 v 计算公式如下:

$$v$$

其中, l_p 为预测的刑期, l_a 为真实刑期。具体计算得分时采用如下公式:

$$score_i$$

其中, i 为案件标号。可以看出, v 越小, 量刑预测效果越好。当系统预测结果与真实刑期之间偏差在 0.2 以内, 预测结果为满分。因此, 在划分刑期区间 $[l_m, l_n]$ 过程中, 中位数 l_p 需要满足 $|\log(l_p + 1) - \log(l_m + 1)| \leq 0.2$ 且 $|\log(l_p + 1) - \log(l_n + 1)| \leq 0.2$, 因此要保证 $\log(l_n + 1) - \log(l_m + 1) \leq 0.4$ 。其中 l_m 和 l_n 分别为划分区间的左边界和右边界, 真实刑期 $l_a \in [l_m, l_n]$ 。在数据集中, 最小刑期是 0 个月, 最大刑期是 300 个月 ($\log 300 \approx 8$)。从 0 月开始, 根据 $\log(l_n + 1) - \log(l_m + 1) = 0.4$ 可以将刑期划分为 20 个区间 ($8/0.4 = 20$)。具体划分时, 由于刑期大于 180 个月的案件数几乎为 0, 因此将 180-300 个月划分为一个区间。

区间划分后的数据分布如图 2 所示。划分完可以看出, 刑期类别由 236 类变为 20 类, 虽然案件大部分还是集中在 0-18 个月的类别上, 但是可以看到相比区间划分前数据分布较为平衡。

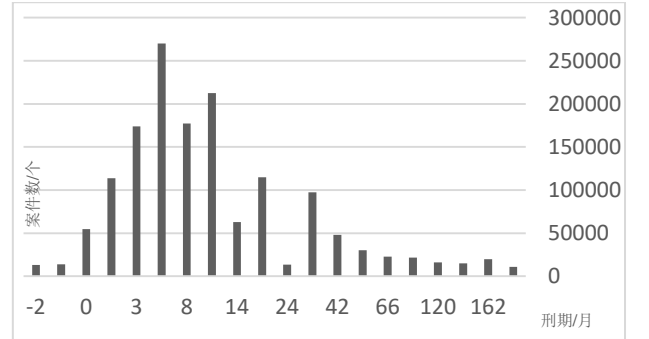


图 2 区间划分刑期分布

2.1.2 多模型投票

多模型投票框架具体采用如图 3 所示的基于深度学习的多模型投票框架, 框架中采用不同参数下的两种模型: 卷积神经网络 (CNN) 和长短时记忆模型 (LSTM); CNN 通过卷积和池化操作可以提取关键特征, LSTM 能够解决长距离依赖问题, 通过对两种模型的输出概率进行累加并取最大值得对应类别作为输出。本框架由三个模块组成, 包括文本表示模块、多模型投票和输出模块。

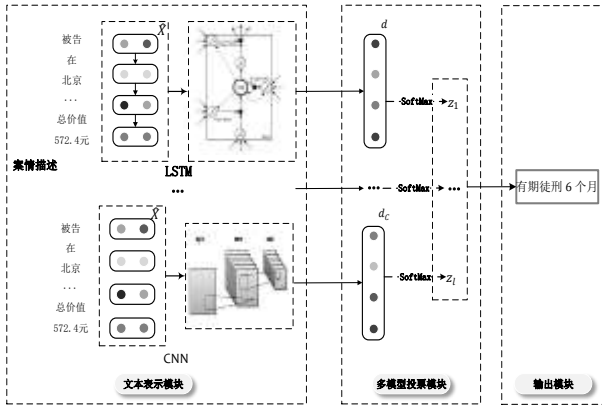


图3 IP&MV 模型框架

(1) 文本表示模块。其输入是案情描述 $X = \{x_1, x_2, \dots, x_m\}$ 对应的词向量序列 $\hat{X} = \{X_1, X_2, \dots, X_m\}$ ，输出是词向量序列 \hat{X} 通过 CNN 与 LSTM 模型得到的向量表示。

利用 CNN 的卷积、池化和全连接操作得到序列 \hat{X} 的最终表示 $d_c = \{d_{c1}, \dots, d_{cm}\}$ 。

采用 LSTM 的输入门、遗忘门和输出门得到序列 \hat{X} 的最终表示 $d = \{d_1, \dots, d_m\}$ 。

(2) 多模型投票及输出模块。多模型投票模块是将所有深度学习模型输出的类别概率累加，然后取最大概率的类别作为输出。具体采用多个不同参数下的卷积神经网络模型和长短时记忆模型，这样可以捕捉到更多文本信息，提高模型的预测效果和稳定性。

由文本表示模块可以得到案情描述隐层表示，然后经过 softmax 函数，得到案情描述在类别集 Y 上的概率分布 Z ；最后取概率累加最大值的类别 y 。计算公式如下：

$$Z$$

$$y$$

其中， $Z = (z_1, z_2, \dots, z_l)$ ， l 为类别个数， W_n^Y 、 b_n^Y 与 d_n 分别是第 n 个模型在输出层的权重矩阵和偏置矩阵与第 n 个模型的隐层输出。

2.2 基于刑期属性的量刑预测模型

2.2.1 刑期属性定义及标注

受到了 Zhong^[14] 基于罪名属性进行罪名预测工作的启发，我们在量刑预测中引入刑期属性，实现对犯罪事实与量刑情节进一步理解，为量刑提供更多有效信息。本文所提方法与该文献的主

要不同在于：(1) 研究目标不同，其主要解决低频罪名与易混淆罪名的预测，而本文主要针对自动量刑进行研究，重点解决量刑情节理解问题。

(2) 属性复杂程度不同。量刑问题中除了考虑犯罪事实，还要考虑犯罪事实之外的情节，如：社会危害程度、被告人所犯罪行的轻重等对刑罚轻重有影响的属性。因此，确定量刑属性时考虑的范围更宽泛，属性标注也更为复杂。

本文根据《刑法》与《量刑指导意见》，针对常见的 31 个罪名总结归纳出 29 个属性，如：盗窃罪有是否入户、是否多次、是否携带凶器和涉案金额 4 个属性。有些罪名会涉及到多个属性，同时属性有的是二值属性，取值为 0、1，分别表示对应属性不发生或发生；有的是多值属性，如：涉案金额、涉案毒品量、损失严重程度等属性，取值一般是 0、1、2、3，表示该属性的不同严重程度。如：对“涉案金额”属性来说，分别对应数额不大、数额较大、数额巨大、数额特别巨大等情况。

具体采用两种策略标注：(1) 通过分析大量文书内容，总结规则进行属性标注；(2) 根据案件刑期对属性进行反推标注。我们随机抽取 100 个标注后的案例进行人工统计分析，属性标注正确率为 92%，覆盖率约为 95%。

2.2.2 基于量刑属性的刑期预测框架

具体采用如图 4 所示的刑期预测框架，框架中采用长短时记忆模型 (LSTM) 对事实描述编码，在得到未带刑期属性和带有刑期属性的表示后将两部分表示拼接起来，最终得到预测类别。本框架由三个模块组成，包括文本表示模块、拼接模块和输出模块，文本表示模块与 3.1.2 节中 LSTM 表示部分相同。

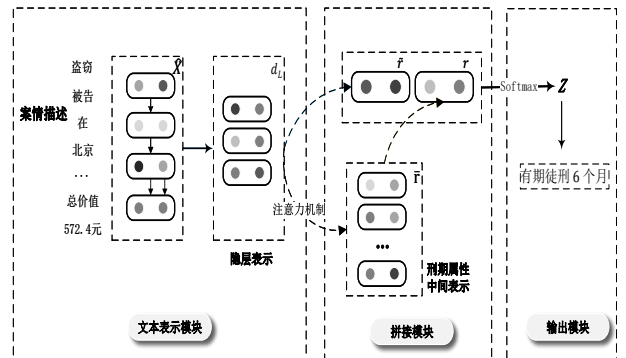


图4 SPA 模型框架

(1) 属性表示拼接模块。拼接模块主要是将

文本表示模块得到的隐层表示分别通过最大和平均池化操作，得到未带和带有刑期属性的表示，然后将两部分表示线性拼接得到最终表示。

由文本表示模块得到 LSTM 编码后的隐层表示 $d = \{d_1, \dots, d_m\}$ ，未带刑期属性的表示 $\tilde{r} = [\tilde{r}_1, \dots, \tilde{r}_t]$ 通过最大池化得到，具体计算如下：

$$\tilde{r}_i = \max(d_{1,i}, \dots, d_{m,i}), \forall i$$

其中， t 为隐层的维度。

带有刑期属性的表示 r 需要通过注意力权重 $a = \{a_1, \dots, a_k\}$ 得到中间表示 $\tilde{r} = [\tilde{r}_1, \dots, \tilde{r}_k]$ ，然后通过平均池化得到，其中 k 为刑期属性个数，具体计算如下：

$$a_{i,j}$$

$$\tilde{r}_i$$

$$r$$

其中， $i \in [1, k], j \in [1, m], a_i = [a_{i,1}, \dots, a_{i,m}]$, W^a 为所有刑期属性共享权重矩阵。

在得到表示 \tilde{r} 和 r 后，将两部分拼接得到最终表示 R ，计算公式如下：

$$R$$

其中， \oplus 表示线性连接。

(2) 量刑预测模块。输出模块是将拼接模块得到的最终表示映射为刑期类别概率分布 p ，最后输出预测类别 y ，具体如下：

$$p$$

$$y$$

其中， W^p 和 b^p 是输出层权重矩阵和偏置向量。

3 实验设置

(1) IP&MV 模型实验设置。对于所有的基线模型，将隐层维度设置为 256 维，最大句子长度设置为 300 词。训练过程中，学习率设置为 0.001，dropout 值为 0.9。在 batchsize 的选取上，综合考虑训练速度和训练结果设置为 500，epoch 设置为 30。CNN 模型和 LSTM 模型均采用交叉熵作为损失函数。

评价指标。CAIL2018 采用的是基于偏差的评分，具体计算如公式 (1) 所示。若案件刑期的标准答案为死刑则 $l_p = 2$ ，无期则 $l_p = -1$ ，才计

分。最后，将所有测试点的分数相加并除以测试点总数乘以 100 作为刑期预测的评价得分如公式 (12) 所示，其中的 $score_i$ 按照公式 (2) 计算。

$$Score$$

实验中分别使用区间划分前的数据和区间划分后的数据作为输入，用以证实区间划分的优势。

(2) SPA 模型实验设置。对于所有的模型，隐层维度设置为 100 维，最大句子长度设置为 500 词。训练中，学习率设置为 10^{-3} ，dropout 值为 1.0，batchsize 设置为 512，epoch 设置为 70，模型损失函数为交叉熵损失函数。采用的评价指标与 IP&MV 模型实验采用相同评价指标。

4 实验结果及分析

为了评价提出的刑期预测模型，本文在两个不同规模的刑事案件数据集上，对基于区间划分和多模型投票模型进行了一系列的实验，并选取一些主流模型进行对比实验。同时在涉及 34 个常见罪名的刑事案件数据集上，对结合案情描述和刑期属性的模型进行单独实验。

4.1 相关数据集

实验数据为“2018 中国‘法研杯’法律智能挑战赛”评测数据，这些数据均来自“中国裁判文书网”公开的刑事法律文书，其中每份文书包括：案情描述、所涉及的法条、被告人被判的罪名和刑期长短等要素。

表 1 相关数据集规模

数据集	CAIL-S	CAIL-L	CAIL-A
案件数	196,000	1,700,000	1,229,225
刑期类	214	236	221

表 2 数据集划分

数据集	CAIL-S	CAIL-L	CAIL-A
训练集	150,000	1,350,000	990,881
验证集	16,000	150,000	127,942
测试集	30,000	200,000	110,402

数据集分为三个子集：CAIL-S、CAIL-L 和 CAIL-A，分别包含 19,6000、170,0000 和 55471

份文书，共涉及 202 条法条，3 类刑期：无期、死刑、有期（时间长短在 0-300 月之间）。其中 CAIL-S 和 CAIL-L 为 IP&MV 模型数据集，均涉及 183 个罪名。CAIL-A 为 SPA 模型数据集，涉及 34 个罪名。数据集规模和划分情况如表 1、表 2 所示。

为了更好的进行对比，本文还采用等距区间划分将刑期划分为 18 个区间，取区间中位数作为区间类别，加上死刑和无期两个特殊类别得到 20 类。等距区间划分和偏差区间划分后数据集分别用 CAIL-S_{等距}、CAIL-L_{等距}和 CAIL-S_{偏差}、CAIL-L_{偏差}表示。

4.2 基线模型

本文采用以下几个模型作为基线模型：

TFIDF+MEM^[15]:该模型主要思想是将 tf-idf 值作为词项特征添加到最大熵模型（MEM）中，是基于传统机器学习文本分类的一种代表方法。

TFIDF+SVM^[16]:该模型主要思想是将 tf-idf 值作为词项特征添加到支持向量机模型（SVM）中，实现文本分类。

CNN^[17]:该模型使用卷积神经网络（CNN）进行事实编码和文本分类。

Multi-Filter CNN^[18]:该模型使用不同大小窗口的滤波器，使用多窗口卷积神经网络（Multi-Filter CNN）进行事实编码和文本分类。

LSTM^[19]:该模型利用一个多层的长短时记忆模型（LSTM）来实现事实编码和文本分类。

4.3 实验结果及分析

表 3 IP&MV 及基线模型实验结果

方法	CAIL-S	CAIL-L	CAIL-S _{等距}	CAIL-L _{等距}	CAIL-S _{偏差}	CAIL-L _{偏差}
TFIDF+MEM	46.35	48.50	47.00	44.28	52.82	53.02
TFIDF+SVM	52.15	54.33	52.26	47.54	65.45	67.23
CNN	56.78	59.01	56.40	50.89	71.00	72.40
Multi-Filter CNN	57.22	60.15	57.31	54.01	72.87	74.54
LSTM	57.41	59.23	56.59	55.53	71.54	73.01
IP&MV	58.25	61.35	57.63	60.46	74.85	78.38

实验结果如表 3、表 4 和表 5 所示，其中表 3、表 4 中的指标为基于偏差的评分。从表中可以看出：

表 4 SPA 模型实验结果

方法	CAIL-A	CAIL-A _{等距}	CAIL-A _{偏差}
TFIDF-SVM	54.33	56.10	60.51
CNN	59.01	64.20	72.54
Multi-Filter CNN	60.15	65.97	73.68
LSTM	59.23	67.99	74.58
SPA	63.00	68.72	76.80

表 5 SPA 模型在数据集 CAIL-A 消融实验结果

模型	Acc.	MP	MR	macro-F1
DL&ATT	46.2	39.3	31.1	32.1
-罪名	44.8	45.3	28.3	29.0
-刑期属性	45.7	42.2	27.7	28.7
-罪名-刑期属性	44.7	41.0	27.0	28.2

（1）无论采用哪种区间划分方法，所有模型在区间划分后的数据集上表现都有提高，充分表明：采用区间划分策略可以缓解数据不平衡问题，减少类别数，使系统性能有效提升。同时也可以看出，在基于评价指标的区间划分策略下，每个模型性能获得了更大的提升，说明了该划分策略更加有效。

（2）整体来看，IP&MV 模型和 SPA 模型得分均高于所有的基线模型。这些结果表明，多模型投票方法有效提升模型预测效果；加入罪名信息和量刑属性后，模型可以捕捉到更多量刑的关键信息，使得模型的预测效果得到提升。同时基线模型中的 Multi-Filter CNN 模型表现优于普通 CNN 模型，主要原因在于拥有多个不同尺寸的滤波窗口，能够获取词语、短语、短句等不同粒度的更丰富的信息。LSTM 模型比 CNN 和 SVM 模型效果好，主要由于其可以解决序列的长距离依赖问题，而案情描述又是相对比较长的文本，因此 LSTM 的表现较好些。

(3) 表 5 第一行是 SPA 模型在 CAIL-A 数据集上表现。Zhong^[4] 提出的拓扑学习量刑预测模型在 CAIL 数据集上获得的性能为 Acc. (38.3), MP (36.1), MR (33.1), F1 (32.1)。尽管我们无法获得该研究采用的训练集与测试集*, 但实验中采用的数据集与本文数据集都来自 CAIL2018 评测公开数据集, 特点与风格一致, 因此实验结果具有一定的参考价值。可以看出, SPA 模型在 Acc. 与 MP 指标上比该模型要好, F1 指标上的结果相当, 但 MR 指标上略低一些。原因是: 模型在个别类别上召回率较低, 特别是刑期为 1 个月的召回率仅为 1%, 使得 MR 值大幅降低。通过数据分析, 发现错误集中在故意伤害罪和盗窃罪上。其中, 故意伤害罪在刑期属性定义时未考虑到当事人双方谅解或者刑事和解的因素; 盗窃罪未考虑到退还赃款和双方谅解的情况, 同时盗窃金额对刑期的影响受到时间、地域、和特殊情况的影响。

(4) SPA 模型的消融实验结果如表 5 所示, 可以看出, 模型分别去掉罪名、刑期属性后, 预测结果大部分降低, 但刑期属性对模型性能影响更大, 表明: 刑期属性能够实现对案情描述的一定理解, 并为量刑提供一定的信息。但是获得的性能提升没有预期的高, 可能的原因是: 罪名与属性信息的利用方式比较单一, 使得这两部分信息没有真正融合到深度学习模型中。与我们预期不符的是: 加入罪名后 MP 一定幅度降低。分析数据后, 主要原因为加入罪名后, 死刑类别识别的准确率较低, 导致 MP 值较大程度变化。通过进一步分析《刑法》, 发现对判处死刑的描述为“死刑只适用于罪行极其严重的犯罪分子”, 非常笼统, 可计算性差。

5 结论

在本文中, 为了解决刑期预测面临的几个挑战, 提出了区间划分和多模型投票的方法来解决。在进行大量对比实验后, 实验结果表明本文提出的偏差区间划分方法一定程度上缓解了刑期类别过多和分布不平衡问题, 同时多模型投票策略在提高模型的预测能力和稳定性。本文主要贡献如下:

(1) 在类别定义中, 引入了区间划分的思想, 综合利用刑期的真实分布和刑期预测评价指标, 将刑期区间切分为 18 个区间并取区间中位数作为该区间类别标签, 有效地克服了类别过多和数据分布不平衡的问题。采用多模型投票策略可以简单有效的结合了不同模型的优点, 提升模型的预测效果, 同时对于模型的稳定性也有了一定的提升。

(2) 在量刑预测中利用罪名信息并从法条信息中总结出量刑属性, 提升模型对于量刑细节的捕获能力, 同时也更加符合法官判案逻辑。

(3) 采用数据集来源于中国裁判文书网数据, 是真实世界的案件信息, 在这些数据上进行研究对当前的司法智能化研究具有一定的借鉴意义。

未来我们将探索如何定义更加合理的刑期属性和对刑期属性进行更为有效的编码和拼接, 进一步提升量刑准确率, 并着手研究量刑预测的可解释性问题。

参考文献

- [1] Fred Kort. 1957. Predicting supreme court decisions mathematically: A quantitative analysis of the "right to counsel" cases [J]. American Political Science Review, 51(1):1-12.
- [2] S Sidney Ulmer. 1963. Quantitative analysis of judicial processes: Some practical and theoretical applications [J]. Law and Contemporary Problems, 28:164.
- [3] R Keown. 1980. Mathematical models for legal prediction [J]. Computer/LJ, 2:829.
- [4] Jeffrey A Segal. 1984. Predicting supreme court cases probabilistically: The search and seizure cases, 1962-1981. American Political Science Review, 78(4):891-900.
- [5] Shapira, M. 1990. Computerized decision technology in social service, International Journal of Sociology and Social Policy 10, 138-164.
- [6] Hassett, P. 1993. Can expert system technology contribute to improved bail conditions, Int. J. of Law and Information Technology 1, 144-188.
- [7] Wanchen Lin, Tsung Ting Kuo, and Tung Jia Chang. 2012. Exploiting machine learning models for chinese legal documents labeling, case classification, and sentencing prediction [C]. In Processdings of ROCLING, page 140.
- [8] Nikolaos Aletras, Dimitrios Tsarapatsanis, Daniel Preotiuc-Pietro, and Vasileios Lampsos. 2016. Predicting judicial decisions of the European court of human rights: A natural language processing perspective [J]. PeerJ Computer Science, 2.
- [9] Octavia Maria Sulea, Marcos Zampieri, Mihaela

* 文献[4]在 github 上公布的系统未能正常运行。

Vela, and Josef Van Genabith. 2017. Exploring the use of text classification in the legal domain[C]. In Proceedings of ASAIL workshop.

- [10] Katz DM, Bommarito Ii MJ, Blackman J. Predicting the Behavior of the Supreme Court of the United States: A General Approach [J]. Plos One, 2017, 12(4).
- [11] Bingfeng Luo, Yansong Feng, Jianbo Xu, Xiang Zhang, and Dongyan Zhao. 2017. Learning to predict charges for criminal cases with legal basis [C]. In Proceedings of EMNLP.
- [12] Zikun Hu, Xiang Li, Cunchao Tu, Zhiyuan Liu, and Maosong Sun. 2018. Few-shot charge prediction with discriminative legal attributes [C]. In Proceedings of COLING.
- [13] Hai Ye, Xin Jiang, Zhunchen Luo, and Wenhan Chao. 2018. Interpretable charge predictions for criminal cases: Learning to generate court views from fact descriptions [C]. In Proceedings of NAACL.
- [14] Haoxi Zhong, Guo Zhipeng, Cunchao Tu, Chaojun Xiao, Zhiyuan Liu, Maosong Sun. Legal Judgment Prediction via Topological Learning [C]. The Conference on Empirical Methods in Natural Language Processing (EMNLP 2018).
- [15] N. Kamal, L. John, M. Andrew. 1999. Using maximum entropy for text classification [C]. Proceedings of the International Joint Conference on Artificial Intelligence Workshop on Machine Learning for Information Filtering.
- [16] Johan AK Suykens and Joos Vandewalle. 1999. Least squares support vector machine classifiers [J]. Neural processing letters, 9(3):293-300.
- [17] Johnson R, Zhang T. Effective Use of Word Order for Text Categorization with Convolutional Neural Networks [J]. Eprint Arxiv, 2014.
- [18] Yoon Kim. 2014. Convolutional neural networks for sentence classification [C]. In Proceedings of EMNLP.
- [19] Ang D, Qin B, Feng X, et al. Effective LSTMs for Target-Dependent Sentiment Classification [J]. Computer Science, 2015.



张博文 (1996-) 硕士研究生, 主要研究方向: 数据挖掘与人工智能。
Email: 853681976@qq.com

张虎 (1979-), 副教授, 硕士生导师, 主要研究方向: 人工智能与自然语言处理。

Email: zhanghu@sxu.edu.cn



李茹 (1963-), 教授, 博士生导师, 主要研究方向: 中文信息处理、数据智能预测。

Email: liru@sxu.edu.cn



谭红叶 (1971-), 副教授, 硕士生导师, 主要研究方向: 人工智能与自然语言处理。

Email: hytan_2006@126.com

