

人工智能在司法量刑中的应用沿革与技术演进

于淼 陆娇 管政翔 陈柳娅

摘要:近年来,人工智能在智慧司法领域的应用方兴未艾。智慧量刑是人工智能为智慧司法赋能的重要环节。国内外实践案例从应用场景、算法设计上有所突破,但仍存在重点难攻克、算法不明确、效果不明晰的困境。本文聚焦智慧量刑问题,深入挖掘以随机森林算法、IP&MV算法、BERT算法为核心的解决方案,打通算法体系与法律体系的屏障。未来,智慧量刑系统将坚持司法审判的辅助性角色,实现人工智能在司法实践的全流程渗透,维护司法审判公平正义。

关键词:智慧量刑;随机森林算法;IP&MV算法;BERT算法

中图分类号: D926

文献标识码: A

文章编号: 1009-2447(2021)03-0088-08

智慧量刑是智慧司法的核心环节,指利用人工智能算法学习司法数据知识,提出精准化量刑建议的过程。自《国家信息化发展战略纲要》和《“十三五”国家信息化规划》发布以来,司法领域从国家战略层面上开始关注人工智能。最高人民法院办公厅发布《关于做好2019年智慧法院建设工作的通知》,将智慧司法从高层的战略部署落地为具体的实践指南。近年来,多地也都对智慧司法的实践问题进行了尝试与深入思考。智慧量刑属于智慧司法的一部分,无论是从量刑科学化、精准化和个性化角度上,都能够展现独特的优势。

智慧量刑的科学化优势,在于妥善处理法律规定与自由裁量的关系。从法律规范上看,《人民法院量刑指导意见(试行)》《关于规范量刑程序若干问题的意见》对量刑步骤和程序进行说明,体现了量刑规范本身“可计算”的特点。从司法实践来看,量刑过程呈现出程序性、重复性的特点,这与智慧量刑的算法本质不谋而合。人工智能通过对法律规则和大量司法案例的学习,可以迅速提取案件

要素并得出量刑建议,将机械的重复劳动简单化,优化司法业务流程,减少司法工作人员的诉累负担,大大提高司法效率。

智慧量刑的精确化优势,体现在尊重同案同判的司法精神。法律面前人人平等是现代司法的重要原则之一,同等案情同等情节要求获得同等对待。传统的司法量刑中,案件的审理除了需要依据成文的法律规范,还受到形势政策、价值判断、法官的认知背景等多种因素的影响。智慧量刑系统通过算法对法律法规进行学习,对不同的量刑因素进行考量,能够有效减少同案异判的司法不公现象,客观上使量刑结果更加公正、均衡。

智慧量刑的个性化优势体现在对案件具体情节的全面考量。随着大数据、云计算、人工智能等技术的发展,我国司法系统实现了法条检索、类案参考、文书生成等功能,开始向人工智能算法判断案情、法律咨询等功能进化。^①传统司法实践中,过于规范化的量刑程序和实体规定可能会忽略刑罚的个别性,而人工智能赋能下的量刑系统通过全面考

基金项目: 国家科技部技术预测专项“技术预测方法与政策实践国际比较研究”(1061-23318064)

作者简介: 于淼,男,北京人,中国政法大学商学院教授、博士生导师,研究方向为大数据与人工智能、物流与供应链管理、智慧量刑、电子商务;陆娇,女,江苏常州人,中国政法大学光明新闻传播学院硕士生,研究方向为传播法。

察案件,避免了遗漏特殊情节。

本文立足于未来人工智能可能为智慧量刑提供的积极效用,扎根法律领域的自然语言处理与量刑预测问题,总结国内外的实践经验,探讨人工智能解决智慧量刑的多种可能方案,为日后智慧量刑的实现提供路径参考。

一、国内外智慧量刑实践经验

近年来,国内在智慧司法、智慧量刑领域产生了很多实践案例。2017年,海南省高级人民法院的“量刑规范化智能辅助系统”实现了推送关联法条、建立智能审判数据库等功能。2018年,上海市虹口区检察院建立“量刑过程可视化系统”,勾选案件在各个量刑因素上的表现即可生成量刑结果。2019年,河北省保定市检察院开发“智能比对系统”,通过智能抓取并对比文书中的关键词,呈现既判类案中刑期、缓刑适用比例、罚金数额的最多判决结果,供法官参考。2020年,安徽省怀宁县检察院建立“小包公智能量刑辅助系统”,通过选择区域和罪名、敲定量刑情节,即可生成量刑预测报告。

国外对于智慧司法与智慧量刑的实践研究先于我国。2014年,澳大利亚“SIS量刑系统”能够提供法律数据查询等服务。2016年,人工智能律师“ROSS”出现,能高效地进行案例检索、回答法律问题等。2017年“E-DISCOVERY系统”出现,它能提供服务更快、价格更低廉的法律分析服务,主要用于卷宗分析及电子取证。2018年出现的“COMPAS信息系统”通过罪犯实施犯罪行为的画像,提取诱发犯罪行为因素并评估罪犯再犯的可能性。

人工智能赋能智慧司法已经在多个环节起到有效的推动作用,从以上示例可以看出,国外对于智慧量刑的尝试主要聚焦于“法条检索”“类案推荐”“文书自动生成”“电子证据与电子取证”等;国内除以上提及的应用外,另有“量刑计算器”等产品的开发。不难发现,智慧量刑中鲜少涉及有关量刑结果预测的内容。原因主要有二,一是诸多算法的选择不明确、效果不明晰。人工智能发

展至今,能够处理智慧量刑问题的算法有很多,但机制不同、逻辑不同,效果自然也不同。二是量刑预测的实现难度较大,这里的难度主要包含两大方面,不仅是算法实现的难度,还涉及背后的法律问题。以下,作者将提出三种采用人工智能算法解决量刑结果预测问题的方案。

二、随机森林算法解决智慧量刑问题

随机森林(Random Forest)算法由Leo Breiman在2001年提出,算法主要基于决策树算法进行改进。决策树,顾名思义,是利用树形分类进行决策。决策树根据决策因素或特征进行分叉,每一次分叉的叶子标识为决策结果。随机森林即建立数百棵决策树组成的“森林”,每一棵决策树都会对同一问题进行根节点的选择并进行结节分裂,生成决策结果。最后统计每棵决策树的决策结果,投票决定该问题的最终决策结果。随机森林算法的应用较为广泛,能够解决多分类问题、预测问题等,在金融领域、市场营销领域、视觉跟踪领域等均有应用。以下将介绍随机森林原理中主要涉及的算法核心:

一是根节点的选择。决策树算法中由于通常仅有一棵树,每次分支都会选择最优的特征作为根节点。而随机森林在根节点的选择上,往往随机选择数据中的样本特征。该特征节点往往能够产生较大的信息增益,其判断方式为信息增益^②或信息增益比^③,涉及公式如下所示:

$$H(D) = -\sum_{k=1}^K \frac{|C_k|}{|D|} \log_2 \frac{|C_k|}{|D|} \quad (1)$$

$$H(D|A) = -\sum_{i=1}^n \frac{|D_i|}{|D|} H(D_i) = -\sum_{i=1}^n \frac{|D_i|}{|D|} \sum_{k=1}^K \frac{|C_k|}{|D_i|} \log_2 \frac{|C_k|}{|D_i|} \quad (2)$$

其中D为训练数据集,|D|表示样本个数,|C_k|为属于类别C_k的样本个数;设特征A有n个不同的取值a₁、a₂、a₃……a_n,根据特征A的取值将D分为n个子集D₁、D₂、D₃……D_n,其中|D_i|为对应样本个数,|D_{ik}|为D_i中属于C_k类的样本集合个数。

二是投票机制的融入。随机森林是由上百棵决策树组成,需要对决策树的结果进行统计处理得到最终的算法结果。投票机制主要分为三种,一是绝

对多数投票^④；二是相对多数投票^⑤；三是加权投票法。^⑥

目前，随机森林算法组合应用于智慧量刑系统构建的技术已经较为成熟。得益于随机森林的算法特征，系统可以高效处理大样本数据，拥有较为明显的速度优势。采用随机森林算法构建智慧量刑系统可以分为三大步，即获取裁判文书、挖掘量刑因素与裁判结果预测。

（一）获取裁判文书

案件的审理过程涉及海量文件与数据，此处的数据源将决定未来算法学习的效果，通常选择裁判文书作为系统的训练文本。其原因有二：一是裁判文书的格式较为标准；二是裁判文书书记录案件基本情况、具有法律效力的证据链等全面的案件信息，有利于提取完整的案件信息从而提高系统准确率。裁判文书的获取渠道主要为中国裁判文书网^⑦，网站上具有上亿篇文本可供下载。下载后的文本储存到数据库中，便于采用文本挖掘算法进行预处理。

（二）挖掘量刑因素

以上提取的文本为非结构化数据，算法无法进行学习。因此，需要建立词库利用文本挖掘对其进行分词、去停用词、去冗余化等处理，构建词向量，而后挖掘主要的量刑因素。构建案件量刑词库指寻找在法律案件中经常出现的作案时间、作案空间、相关人、相关行为以及相关物等词汇。以抢劫类刑事案件为例，其所构建的词库中应当包含犯罪人信息与抢劫案件特定量刑因素两大部分。犯罪人基本信息中包括年龄、受教育程度等，而特定量刑因素中则包含抢劫次数、抢劫方式、抢劫金额等差异化的量刑因素。

分词、去停用词、去冗余化、标注词性等方式都是对非结构化文本的预处理手段，经由以上处理的非结构化文本将转化为词汇或短语。分词过程经常采用Jieba分词，标注词性经常采用POS-Tagging算法。通常抽取10%作为训练集，训练适用于不同案件类型的POS算法进行词库的读取，对训练集的文本进行标记。

至此，长篇文本已经转化为一个个具有词性的词汇或短语。为能够利用随机森林进行预测学习，

需要将非结构化文本转变为结构化数据。这里通常用到的方式为Word2vec词向量生成算法。^⑧Word2vec将生成的词汇或短语转化为计算机理解的稠密向量，词向量可以包含更丰富的信息并且体现词之间的相似关系。

（三）裁判结果预测

量刑结果预测阶段主要利用以上生成的词向量并结合随机森林算法实现。该部分经历两大步骤：

一是训练随机森林生长。历史的裁判文书生成的词向量组成样本集，分为训练集与测试集。训练集用于随机森林算法的训练，测试集用于准确率的测试。由于每棵决策树的生长都需要训练集，因此进行随机有放回的抽样工作，生成若干训练集分别进行决策树的训练。决策树读取历史裁判文书词向量进行生长，生长过程遵循信息增益原理随机生成根节点，逐渐提高裁判能力。决策树长成后形成完整的随机森林，森林中每棵决策树都相当于一位法官。

二是测试随机森林的预测准确率。该部分通过输入测试集的案例信息部分词向量至训练好的随机森林算法中，各位“法官”采用投票机制，票数最高者为最终的预测结果。最终该方案的准确度由测试集生成准确结果的比率决定。

通过以上处理步骤，成熟的随机森林算法能够被利用至司法实践中。法官将案件信息输入至算法系统中，系统将给予在该案件背景下可能的量刑结果。

三、IP&MV算法解决智慧量刑问题

基于区间划分和多模型投票方法（Interval Partition and Multi-Models Voting, IP&MV）是一种结合卷积神经网络（Convolutional Neural Network, CNN）与长短时记忆模型（Long Short-Term Memory, LSTM）两种模型优点的深度学习模型。IP&MV模型应用于辅助司法量刑时，其本质是构建一种映射，使文书信息和量刑结果之间建立一一对应的关系。IP模型主要针对映射对象的处理，通过划分区间解决映射结果的问题，建立包含各种数据可能性的集合。MV模型主要针对原始文本信息处理的问题，通过学习大量训练集数据，将多个模型的学习预测结果的概率进行累加，以概率形式输出预测结

果。MV模型主要分为三大模块:文本表示模块、多模型投票模块、输出模块。文本表示模块的主要任务是将文本转化为可读取接收的输入。多模型投票模块由LSTM和CNN对文本特征进行提取和预测。输出模块中,算法将综合预测出的不同结果进行投票决策。以下将具体叙述该算法应用至智慧量刑问题时的推进步骤。

(一) IP模型平衡化刑期分布

该算法的数据源一般为裁判文书。在非“全样本”的中国裁判文书的大数据基础上,刑期、刑罚类型等数据不平衡,存在“长尾现象”。个别刑罚类型的数据样本可能极少,构建映射时难以保证样本选择的典型性与全面性,带来“选择性偏误”。IP模型能够平衡数据样本分布,解决“选择性偏误”问题。

(二) MV模型预测刑期

MV模型深度学习大量裁判文书中蕴含的文本信息,将多种模型的学习预测结果的概率进行累加,以概率形式输出量刑预测结果。文本表示模块将原始的法律文书转化为结构化数据。该过程首先需要建立相应的司法词典并进行分词,建立文书中的字词与字符串之间的联系,将原始的法律自然语言转化成词向量。词向量分别平行进入多模型投票模块中的LSTM和CNN进行特征的提取和预测。在LSTM模型中,裁判文书形成的词向量需要先后通过遗忘门、输入门和输出门,过滤冗余信息,将输入的重要信息与隐藏状态的信息有选择性地保留在“记忆细胞”中。之后,LSTM模型会根据保留的“记忆”,通过不同的参数共享提取特征,沿着长链的序列结构进行预测。

与LSTM模型不同,CNN模型不注重词序排列,而更擅长抽取局部特征。在CNN模型中,词向量将经过卷积、池化和全连接等模型的一系列处理,提取出文本的特征信息。进入模型后,CNN会进一步“凝视”,对向量卷积,把滤波器与文本转化来的n维向量的各个部分逐一累乘,以提取出文本特征。但是,直接由卷积计算出的特征往往有较高的维数,需要通过池化,减小表示输入微小变化的空间维数,避免算力浪费。全连接层将在卷积层中创建的“局部的”特征组合到一个分类器中,得

到整个法律文本的特征矩阵,能够解决可能产生的“死节点”问题,提高全局搜索的能力。特征矩阵反映了法律文书总体内容。将特征矩阵导入输出层后进行预测,经激活函数处理,得到不同量刑结果的概率分布。

最后,在多模型投票模块的末端,将CNN和LSTM预测出的不同量刑结果出现的概率量刑结果依照相同的罪名和刑期进行累加,并做归一化以得到最终的预测。最高概率值的结果可被最终输出。

(三) 微调MV模型

通过对法律规定和相关指导性文件的解读,总结规律以制定属性。MV模型能够根据评估属性对文书信息的覆盖度以验证算法的效度。将评估属性进行编码后进行排列形成向量,将此向量与原始输入转化而来的词向量进行拼接,拼接后形成的新词向量将带有属性供模型识别,成为最终被模型调用的预测序列。这也为自由裁量的空间匹配了可操作性。

四、BERT算法解决智慧量刑问题

BERT,全称Pre-training of Deep Bidirectional Transformers for Language Understanding(基于深度双向转换器的自然语言理解预训练模型),是由Google人工智能团队于2018年提出的自然语言处理预训练模型,在多项自然语言处理任务上取得突破性成绩。BERT由Transformer的编译器结构堆叠而成,内部可以被分为输入层、位置嵌入层、多头注意力层、前馈层、残差连接与归一化层、输出层六层结构。通过这种多层结构,算法得以提取出文书中文本的特征,将其数理化表达为隐藏层的数组。

在输入层,BERT算法将输入的自然语言文本转化为结构化数据。这种转化是以单字为单位进行的,每一个单字都对应着一个行向量,行向量的每一列都由字符串构成,代表单字的一层含义。通过这种转化,文本的每一句话都可以被表达为一个二维矩阵,自然语言文本整体则可以被表达为一个三维数组。特殊的输入层使得BERT算法能够摆脱传统自然语言模型机械死板的分词操作,这是BERT其他结构成立的基础。

在位置嵌入层,BERT算法赋予自然文本位置信

息。由于BERT算法的输入没有迭代操作,所有单字同时输入而不分先后,因此需要这一结构赋予位置信息。这些位置信息一般用三角函数计算确定,并以数组的形式组织起来,与原数组相叠加以实现位置信息的添加。BERT算法对位置信息的处理方法有别于传统模型中以排序表达位置信息的方法,使得每一个单字在全文中的相对位置都能被准确表达。具体的计算公式如下所示,其中,POS指向句中字的位置, i 指所求值的列坐标,列坐标奇偶性不同的值适用不同的公式。

$$PE_{(POS,2i)} = \sin(POS/10000^{2i/d_{model}}) \quad (3)$$

$$PE_{(POS,2i+1)} = \cos(POS/10000^{2i/d_{model}}) \quad (4)$$

在多头注意力层, BERT算法可以模仿人类的注意力机制挖掘自然文本的深层含义。人类可以通过在不同事物间分配注意力以挖掘事物间联系,注意力分配的方式不同,人类对事物联系的认知也不同,基于多角度的综合的认知往往相对接近客观事实。在这一层模型将对表达自然语言文本的数组进行切割,并在不同切片的基础上建立数理化单字间的联系,从而模仿出人类的注意力机制,从多个角度实现对自然文本深层语义的提取。

在残差连接与归一化层, BERT算法对层间的输入和输出进行规范和调整。残差连接与归一化层不独立存在,而是附属于多头注意力层和前馈层之后,起防止梯度消失和爆炸的作用。残差连接是指将上一层的输入和输出进行拼接,从而将梯度传递下去,方便训练反转,防止原始信息在传递过程中丢失;归一化则是将残差连接后的矩阵转变为标准正态分布,以加快模型训练和收敛的速度,提高模型运算效率。

在前馈层, BERT算法对文本特征信息做进一步的加工。前馈层一般由线性映射和激活函数嵌套构成,能够对多头注意力层输出的形式进行调整。前馈层及类似的结构在众多神经网络模型的搭建中都有运用,为常规结构,不做展开说明。

一般的BERT预训练预测机制主要包括两种:掩盖预测和句间预测。掩盖预测指掩盖掉输入的自然文本中的一部分,训练模型预测出被掩盖的部分。句间预测指以两句为一组构建若干组句对,然后训练模型判断句对是否满足上下文关系。

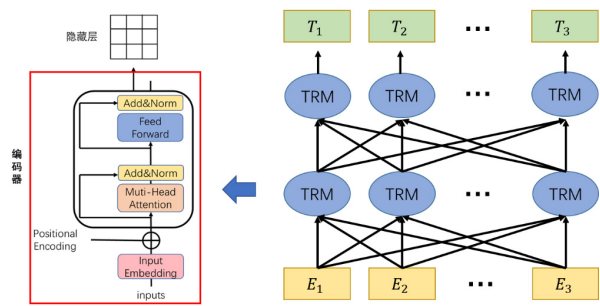


图1 BERT算法流程

通过以上对BERT算法原理的剖析,可以发现, BERT算法能够摆脱词典对一般自然语言模型的限制,从字词句关系中学习文本的真正含义,这一优势若用于智慧量刑中效果将十分显著。

(一) 训练集文本输入

在BERT算法应用于智慧量刑的训练阶段,数据源为带有案件信息的法律文书,一般为裁判文书。每份裁判文书会在模型的输入层以单字为单位进行数理化,并在位置嵌入层获得位置信息,最终以三维数组的形式送入模型的多头注意力层。

(二) 数组转换,识别语义

在多头注意力层,数组会分别与三个权重矩阵 W_Q 、 W_K 、 W_V 相乘,得到三个新的数组Q、K、V,这个运算相当于对原数组做了三次不同的线性变换,运算所使用的权重矩阵由模型预训练确定。得到新数组后,模型将对三个新数组做切片操作。在对输入层的阐述中,原数组中每个单字都由行向量表达,行向量的每一位都是能独立表达单字某一含义的字符串。因此,表达自然文本的数组和其线性变换做切片后仍然可以表达完整的文本含义,每一个切片都可以从一个角度完整地表达法律文书的含义。将单个切片称之为头,切片数也即头数。

做切片后,模型分别取Q的一头和K的一头,对Q的一头和K的一头的转置做点积运算,根据向量点积的定义,点积值越大,参与点积的两个向量的相似程度越大。例如“故”和“意”这两个字在法律文书中相邻时常组合为词语“故意”。因此在训练完成的模型的运算下,表示这两个单字的行向量的点积值也较高,可以被模型识别为一个词。由此模型得到两个单字之间关系的数理表达。

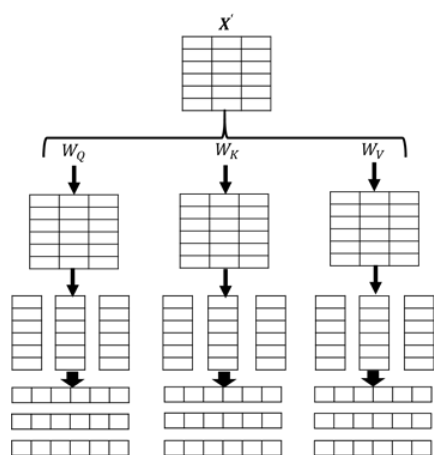


图2 多头注意力机制

将该运算重复进行并将结果整理，模型就可以得到自然文本中任意两个单字间的关系，并以数组 QK^T 表达。将这一数组与未被调用数组 V 相组合，可以得到一个新的数组，这个数组一般被称为注意力矩阵，既能够反映单字间的联系，又保留了原本矩阵行向量与单字之间的关系。对于量刑预测，这个矩阵包含了罪名的成立要件等关键信息。注意力矩阵形成后会被送入残差连接与归一化层、前馈层。

（三）掩盖机制辅助预测与准确度测试

通过残差连接与归一化层、前馈层，模型最终在输出层得到能够完整表达文本特征信息的“隐藏层”数组。得到隐藏层后，只需使用训练集重复掩盖预测的步骤，即可使模型依据隐藏层中包含的法律文书特征信息进行量刑预测，输出罪名和刑期等量刑信息。

具体地，对于训练集，算法使用掩盖预测的机制，将有关判决结果的文本做掩盖处理，训练模型根据法律文书的未掩盖部分预测量刑结果，通过比对预测出的量刑结果和真实的量刑结果调整模型，直到模型达到预设的准确度，并通过测试集的有效性测试。在训练完成后，模型便可以准确地从文书中提取案件信息，并基于此实现对量刑结果的预测。

五、总结

（一）三大解决方案的优劣势

1. 随机森林算法解决智慧量刑问题的优劣势
随机森林算法处理智慧量刑问题的优势在于，

一是原理较为简单，处理量刑问题的逻辑也较为清晰；二是该算法在训练速度上有较大的优势，能够高效率处理大样本；三是算法的泛化能力强，这得益于根节点的选择采用随机机制。

但从司法量刑预测的实践上来说具有一些局限性，一是随机森林算法一般需要结合文本挖掘使用，需要建立词库，而词库的建立过程将耗费大量的人力、物力、财力；二是从准确率方面来说，随着问题的复杂度增强，算法的准确率会下降。量刑预测是复杂度较高的司法实践问题，最终精度不会太高；三是从算法运行逻辑上说，为简单的特征选择结合统计分析，无法真正学习法律文本中的判案逻辑。

2. IP&MV模型解决智慧量刑问题的优劣势

IP&MV模型一是通过IP模型平衡了法律文书的数据分布，减小“选择性偏误”带来的预测偏差；二是通过MV模型融LSTM与CNN优点于一身，模型吸纳了LSTM的整体性机制，有利于从数理逻辑入手提取作为大前提的构成要件要素事实模型与作为小前提的相关案件事实要素。在CNN的基础上，模型从局部特征方面进行突破，把握文本概念的语义边缘，更准确地提取法律文书中的案件“事实场景”。

但IP&MV模型也具有一些局限性，一是依赖于较为成熟的词库，而法律语言系统庞杂且处于变化之中，体量也较大，因此构建词库难度较大，需要不断更新校准；二是较难从法律的角度对LSTM模型的遗忘机制进行点对点的逻辑解释，算法的审判依据较难剥离；三是CNN模型的局部特征提取时，可能会因为池化忽略重要的审判信息。

3. BERT算法解决智慧量刑问题的优劣势

凭借其特有的结构设置，BERT算法能够以单字为基本单位进行自然语言的输入和处理，摆脱词典对一般自然语言模型的限制。通过注意力机制学习单字之间的联系，从而深入理解文本的深层含义，以更加连续、更加灵活的方式自主对文本进行分词，克服了文书用词的不规范等因素造成的困难，省去了词典搭建成本的同时提高自然语言处理的准确率。美中不足的是，BERT算法的搭建难度较高。

（二）人工智能智慧量刑风险问题

人工智能处理智慧量刑过程中，除算法的设计难度外，在实践中也存在一些法律风险，这些风险

一直被学界探讨。一是算法黑箱问题,算法应用于司法实践主要依靠两大条件,分别是数据源和复杂的程序。数据源的质量影响着最终的实践效果,实现最终目的的程序设计方式也受程序设计者风格、实践目的的影响。因此,算法黑箱问题需要从法律法规的层面对两大方面进行规制,对数据源的适用性以及算法设计给予法律明文限制,以保证智慧量刑的公正性,从而促进智慧量刑系统的健康发展;二是更新机制问题,随着法律的更新和案件的积累,算法的数据源需要设置一定的更新机制,从而保证量刑过程对法律法规的高度匹配,实现量刑结果的准确性与规范性。

(三) 应用展望

法治问题本质上是人的问题,运用智慧量刑系统是为司法提供更多的参考材料及参考价值,应当明确智慧量刑在司法中仍然处于辅助性地位,而非主体性地位。未来智慧量刑应用于司法审判的流程是,法官接收案件后,对其案件信息进行初步梳理,并将整理出的文本输入至智慧量刑系统中;而后系统根据案件信息中包含的量刑因素输出量刑预测结果及相应的建议书(主刑建议书、缓刑建议书、罚金刑建议书等)。未来智慧量刑系统还可以将功能延伸,对接智慧司法的前端(电子取证、电子证据存储、智慧法庭等)和后端(司法审判案件分析、风险评估、业绩评估等),人工智能和大数据将为司法实践全流程赋能。

注释

- ① 钱大军:《司法人工智能的中国进程:功能替代与结构强化》,法学评论,2018年第5期,第143页。
- ② 信息增益指 $H(D)-H(D|A)$ 。
- ③ 信息增益比指 $\frac{H(D)-H(D|A)}{H(A)}$ 。
- ④ 即统计每类结果的分类票数,超过半数则将该预测结果列为标记结果。
- ⑤ 即首先选择分类结果较多的类别,并随机选择一类作为最终结果。
- ⑥ 即根据每类结果的票数和所占比例计算综合结果,最高值则为最终的投票结果。
- ⑦ 中国裁判文书网网址: <https://wenshu.court.gov.cn/>。

⑧ 这是Google提出基于上下文语境来获取词向量的算法。

参考文献

- [1] Guo Z, Yu B, Hao M, et al. A Novel Hybrid Method for Flight Departure Delay Prediction using Random Forest Regression and Maximal Information Coefficient[J]. Aerospace Science and Technology, 2021(334):106822.
- [2] Choi J, Lee S, Jamal T. Smart Korea: Governance for Smart Justice during a Global Pandemic[J]. Journal of Sustainable Tourism, 2021(2-3):540-549.
- [3] Suykens J, Vandewalle J. Least Squares Support Vector Machine Classifiers[J]. Neural Processing Letters, 1999(3):293-300.
- [4] 李春楠, 王雷, 孙媛媛, 等. 基于BERT的盗窃罪法律文书命名实体识别方法[J]. 中文信息学报, 2021(8):73-81.
- [5] 刘雁鹏. 智慧司法中的忧虑: 想象, 剖析与展望[J]. 理论与改革, 2020(3):168-181.
- [6] 钱大军. 司法人工智能的中国进程: 功能替代与结构强化[J]. 法学评论, 2018(5):138-152.
- [7] 张玉洁. 智能量刑算法的司法适用: 逻辑、难题与程序法回应[J]. 东方法学, 2021(3):187-200.
- [8] 郑海山. 大数据时代建构人工智能辅助量刑系统的路径探讨[J]. 湘江青年法学, 2018(1):68-87.
- [9] 马靖云. 智慧司法的难题及其破解[J]. 华东政法大学学报, 2019(4):110-117.
- [10] 陈静, 余建波, 李艳冰. 基于随机森林的用户流失预警研究[J]. 精密制造与自动化, 2021(2):21-24+51.
- [11] 张心向. 构成要件要素: 从文本概念到裁判类型[J]. 东方法学, 2020(1):58-67.
- [12] 谭红叶, 张博文, 张虎, 等. 面向法律文书的量刑预测方法研究[J]. 中文信息学报, 2020(3):107-114.
- [13] 张春云, 曲浩, 崔超然, 等. 基于过程监督的序列多任务法律判决预测方法[J]. 计算机科学, 2021(3):227-232.
- [14] 王文广, 陈运文, 蔡华, 等. 基于混合深度神经网络模型的法律文书智能化处理[J]. 清华大学学报(自然科学版), 2019(7):505-511.
- [15] 潘瑞东, 孔维健, 齐洁. 基于预训练模型与知识蒸馏的法律判决预测算法[J/OL]. 控制与决策, 2020(12):1-9[2021-07-09]. <https://doi.org/10.13195/>

- j. kzyjc. 2020. 0985.
- [16] 翁洋, 谷松原, 李静, 等. 面向大规模裁判文书结构化的文本分类算法[J]. 天津大学学报(自然科学与工程技术版), 2021(4): 418-425.
- [17] 刘宗林, 张梅山, 甄冉冉, 等. 融入罪名关键词的法律判决预测多任务学习模型[J]. 清华大学学报(自然科学版), 2019(7): 497-504.
- [18] 舒洪水. 司法大数据文本挖掘与量刑预测模型的研究[J]. 法学, 2020(7): 113-129.
- [19] 郑丽敏, 乔振铎, 田立军, 等. 基于BERT-LEAM模型的食品安全法规问题多标签分类[J]. 农业机械学报, 2021(7): 244-250+158.
- [20] 崔斌, 邹蕾, 徐明月. 基于BERT的诉讼案件违法事实要素自动抽取[J]. 科学技术与工程, 2021(9): 3669-3675.
- [21] 陈剑, 何涛, 闻英友, 等. 基于BERT模型的司法文书实体识别方法[J]. 东北大学学报(自然科学版), 2020(10): 1382-1387.

The Application Development and Technological Evolution of Artificial Intelligence in Judicial Sentencing

YU Miao LU Jiao GUAN Zhengxiang CHEN Liuya

Abstract: In recent years, artificial intelligence has enjoyed increasing theoretical discussion and practical application in the field of intelligent judiciary. Intelligent sentencing is an important link in AI-powered intelligent judiciary. Breakthrough has been made at home and abroad in the applying scenario and algorithmic design of practical cases. However, difficulties remain in key barriers, indefinite algorithm, and unclear outcome. This paper focuses on the issue of intelligent sentencing, closely examines the solution based on random forest algorithm, IP&MV algorithm, and BERT algorithm, and removes the barrier between algorithmic system and legal system. In future, intelligent sentencing system will maintain its subsidiary role in judicial trial, realize the full-time involvement of artificial intelligence in judicial practice, and defend fairness and justice of judicial trial.

Key words: intelligent sentencing; random forest algorithm; IP&MV algorithm; BERT algorithm