

基于多轮LLM和犯罪知识图谱的多被告人法律判决预测

王东升

引用本文

王东升. 基于多轮LLM和犯罪知识图谱的多被告人法律判决预测[J]. 计算机科学, 2025, 52(8): 308-316.

WANG Dongsheng. [Multi-defendant Legal Judgment Prediction with Multi-turn LLM and Criminal Knowledge Graph](#) [J]. Computer Science, 2025, 52(8): 308-316.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于大语言模型的移动应用隐私政策合规性检测方法](#)

Privacy Policy Compliance Detection Method for Mobile Application Based on Large LanguageModel
计算机科学, 2025, 52(8): 1-16. <https://doi.org/10.11896/jsjcx.250300156>

[多模态大语言模型的安全性研究综述](#)

Survey of Security Research on Multimodal Large Language Models
计算机科学, 2025, 52(7): 315-341. <https://doi.org/10.11896/jsjcx.241100141>

[基于置信度引导提示学习的多模态方面级情感分析](#)

Confidence-guided Prompt Learning for Multimodal Aspect-level Sentiment Analysis
计算机科学, 2025, 52(7): 241-247. <https://doi.org/10.11896/jsjcx.240600126>

[基于双重预训练的商品属性分类方法](#)

Commodity Attribute Classification Method Based on Dual Pre-training
计算机科学, 2025, 52(6A): 240500127-8. <https://doi.org/10.11896/jsjcx.240500127>

[基于大语言模型的审计领域命名实体识别算法研究](#)

Study on Named Entity Recognition Algorithms in Audit Domain Based on Large LanguageModels
计算机科学, 2025, 52(6A): 240700190-4. <https://doi.org/10.11896/jsjcx.240700190>

基于多轮 LLM 和犯罪知识图谱的多被告人法律判决预测

王东升

中国政法大学法治信息管理学院 北京 102249

摘要 一些研究利用先进的大模型(LLM)技术理解法律事实,预测被告人的罪名、刑期等判决结果。为进一步深入研究,选择了更为复杂的多被告人法律判决预测任务,它比单被告人预测更困难。具体地,将与 LLM 的交互由单轮升级为多轮,以此提高 LLM 对案件的理解能力。此外,构建了描述案件的两类犯罪知识图谱,其中犯罪关系知识图谱刻画了被告人之间的帮助关系,量刑情节知识图谱展示了案件的核心犯罪情节。通过犯罪知识图谱,设计了一个检索系统为 LLM 提供类案判决的参考。在多被告人法律判决预测实验中,所提方法的预测结果优于对比方法,这表明多轮 LLM 交互和犯罪知识图谱的设计是有效的。

关键词: 多被告人;法律判决预测;大语言模型;犯罪知识图谱

中图分类号 TP391.1

Multi-defendant Legal Judgment Prediction with Multi-turn LLM and Criminal Knowledge Graph

WANG Dongsheng

School of Information Management for Law, China University of Political Science and Law, Beijing 102249, China

Abstract Some studies use advanced Large Language Model(LLM) technologies to understand legal facts and predict the defendant's charges, prison term and other judgment results. For further in-depth research, this paper chooses the more complex task of predicting legal judgments for multiple defendants, which is more challenging than predicting for a single defendant. Specifically, upgrading the interaction with LLM from a single-turn to multi-turn process to enhance LLM's understanding of criminal cases. In addition, two types of crime Knowledge Graphs(KGs) are constructed to describe the case. The criminal relationship knowledge graph depicts the relationships of assistance between the defendants, while the sentencing circumstance knowledge graph represents the core criminal details of the case. Through crime knowledge graphs, a retrieval system is designed to provide LLM with references for similar case judgments. In experiments on predicting legal judgments for multiple defendants, the prediction results of the proposed method are better than the comparison methods, which shows that the designs of multi-turn LLM interactions and crime knowledge graphs are effective.

Keywords Multi-defendant, Legal judgment prediction, Large language model, Criminal knowledge graph

1 引言

法律判决预测(Legal Judgment Prediction, LJP)是人工智能在法律领域的一项重要任务^[1-2],它包括罪名预测、法条预测和刑期预测 3 项子任务。对于案件中的一个被告人来说,他/她因触犯一个或多个法律条款而被指控与法律条款对应的一个或多个罪名,同时被判处长短不等的拘役或有期徒刑。目前已有大量的工作尝试解决其中一种或者全部 3 种法律判决子任务,它们都取得了不错的预测效果^[3-4]。这些工作大多仅围绕一个被告的单被告人案件的案件陈述,任务相对明确和简单。然而,现实生活中还

存在着大量的团伙犯罪案件,其对应的是多被告人法律判决任务。

图 1 中的案件包含两个被告人,他们都殴打了同一个受害人。由于各种事实混杂在一起,因此判决此类案件的一个难点是模型容易混淆不同被告人的犯罪行为。如果采用单被告人模型分别预测每个被告人的判决,则难以将被告人与其犯罪事实对齐,进而出现更多判决失误。此外,由于是团伙作案,被告人之间相互影响,因此他们的判决之间有一定关系但又存在差别。例如,图 1 中被告人 A 帮助了被告人 B 殴打被害人 A,所以尽管两人都是故意伤害罪,但被告人 A 的刑期更长。

到稿日期:2024-09-29 返修日期:2025-01-22

基金项目:中国政法大学科研创新项目(24KYGH013);中央高校基本科研业务费专项资金

This work was supported by the China University of Political Science and Law Research Innovation Project(24KYGH013) and Fundamental Research Funds for the Central Universities.

通信作者:王东升(wangdsh@cupl.edu.cn)

案件事实: 2015年5月18日17时55分许,被告人[被告A]、[被告B]在<LOC>体育场福X楼下,因被告人[被告B]怀疑被害人A之前偷过自己的钱,辱骂并殴打被害人,被害人逃离时,被告人[被告A]将其拉住,后被告人[被告B]脱下鞋子在被害人身上、胳膊上、头部击打,被告人[被告A]对被害人拳打脚踢,致被害人腹部闭合性损伤(小肠破裂)、急性弥漫性腹膜炎、慢支并双下肺感染、切口感染住院治疗。经鉴定,被害人A之损伤程度属重伤二级。
犯罪关系: (被告A,帮助,被告B)
量刑情节: (被告B,怀疑被告人A,偷钱),(被告B,殴打,被害人A),(被告A,殴打,被害人A),(被害人A,受伤情况,重伤二级)
案件判决: 被告A(故意伤害罪;234;4年),被告B(故意伤害罪;234;3年6个月)

图1 多被告人法律判决预测案例

Fig.1 Example of legal judgment prediction for multiple defendants

为了解决法律判决预测问题,研究人员先后提出了规则判断和机器学习的方法。例如,Sukanya等^[5]利用三层编码器解决法律文本的长依赖问题,Zhao等^[6]采用图神经网络处理罪名混淆问题。随着预训练语言模型和大语言模型的快速发展与崛起,面向法律判决预测任务出现了大量更好的解决方案。例如,Xiao等^[7]构建了预训练语言模型 Lawformer,Zhou等^[8]采用预训练和微调技术在大规模中文法律数据集上构建了 LawGPT。然而,这些工作普遍面向单被告人案件,未考虑到团伙作案对应的多被告人犯罪情况。最近的一些研究开始关注多被告人案件。Pan等^[9]利用局部注意力让模型关注被告人名字周围的描述,采用全局注意力让模型关注单个被告人的罪名,从而解决多被告人犯罪事实难以区分的问题。Lyu等^[10]提出了层次推理网络,直接生成每个被告人的判决预测结果。但这些方法尚未探索类案在多被告人判决预测中的重要作用,而且对案件中多被告人及犯罪事实的区分和关联不够明显。

本文面向多被告人法律判决预测,提出了一套全新的解决方案。首先,将单轮 LLM 交互升级为多轮 LLM 交互,以增强 LLM 对任务的理解和推理能力。在第 1 次和第 2 次交互时,LLM 负责识别案件中多个被告人之间的帮助关系,以及包含案件重要信息的量刑情节。其次,参考法官判案时会参考先例的做法,设计了一种基于犯罪知识图谱的类案检索系统。具体来说,基于犯罪关系知识图谱和量刑情节知识图谱,该系统计算两个案件中任意两个被告人的相似度,然后对两个案件做整体的相似度评估。最后,在与 LLM 的第 3 次交互中,LLM 根据类案的犯罪事实及判决结果,预测当前案件对应的罪名、法条和刑期。在实验中,将所提模型与其他方法对比以验证其有效性。据调研,本文是首个采用多轮 LLM 与犯罪知识图谱解决多被告人判决预测问题的。总的来讲,本文的主要贡献如下:

1) 提出了一种新颖的多被告人法律判决预测方法,它通过与 LLM 多次交互完成犯罪关系知识图谱和量刑情节知识图谱的构建,以及罪名、法条和刑期的预测。

2) 利用两种类型的犯罪知识图谱,设计了一个面向多被告人的类案检索系统。该系统可以为 LLM 提供类案法律判决的参考。

3) 将所提模型与单被告人代表模型和多被告人代表模型进行综合对比,证明了该模型在 3 项判决子任务上的

效果更好。同时,通过消融实验表明了 LLM 多轮交互和犯罪知识图谱的重要性。

2 相关工作

2.1 法律判决预测

法律判决预测是司法服务领域的一项核心工作^[11],其判决结果决定了对被告人的惩罚力度,影响了法律的权威性和公正性。目前,罪名预测、法条预测和刑期预测是法律判决预测的 3 项预测内容。早期的研究基于规则进行判决,如 Kort^[12]使用公式从已判决案件的事实要素中得出数值,然后使用这些数值事实要素来预测其他案件的判决。这种方法的可扩展性差,代价高。随着机器学习的发展,研究人员开始将 BERT 和图注意力网络等模型应用于法律判决预测^[13-14]。Sukanya等^[5]设计了一种基于分层注意力模型的三层编码器,以解决法律案件冗长性质造成的长期依赖问题。Zhao等^[6]利用具有多图融合机制的图神经网络来解决有限数据条件下的罪名混淆问题。随着预训练模型的兴起,一些研究将预训练好的模型应用于法律判决预测。Chalkidis等^[15]在法律语料库上预训练了 BERT 模型,并针对下游任务对模型进行了微调。Xiao等^[7]发布了一个基于 Longformer^[16]的预训练语言模型 Lawformer,该模型擅长处理长篇法律文本,可用于类似案件检索、法律问答和法律判决预测等任务。以上相关工作虽然取得了很大进展,但都局限于一个案件一个被告人,忽视了大量团伙作案的案件。受到 Pan等^[9]研究的启发,本文重点关注多被告人案件的法律判决预测,这更加复杂且更有挑战性。

2.2 大语言模型

大语言模型是经过预训练、微调和人类反馈强化学习后的大规模神经网络^[17],它的自然语言理解能力、生成能力和逻辑推理能力非常强悍,因此被广泛应用于信息抽取^[18]、机器翻译^[19]、问答^[20]等任务。在司法领域,研究人员构建了许多法律大模型,用它们理解法律案件,分析案情,完成司法考试、舆情监测等工作。Zhou等^[8]利用大规模的中国法律数据集对 ChatGLM 和 Chinese-LLaMA 等中文大语言模型进行预训练和微调,从而构建了一个名为 LawGPT 的法律知识增强大语言模型。Huang等^[21]从中国司法考试中收集了一套法律咨询数据和客观问题,并对预训练的 LLaMA 基础模型进行指令微调,进而构建律师 LLaMA 大模型。此外,Wu等^[22]先借助领域模型找到先例,再使用 LLM 理解当前案件并给出判决。Shui等^[23]将 LLM 与信息检索系统相结合,以发现类似的案件及其判决,帮助 LLM 回忆特定的法律领域知识。Deng等^[24]通过采用基于人类司法推理框架的多任务微调,进一步提高了现有 LLM 的性能。然而,以上研究工作都是与 LLM 进行单轮交互,没有充分发挥 LLM 的潜力。与它们不同的是,本文受到思维链^[25]的启发,研究多轮 LLM 交互在法律判决预测任务中的应用。

2.3 犯罪知识图谱

知识图谱将实体、属性和实体之间的关系以三元组的形式组织到一起,这种显性知识易于理解,且能够方便地在各种

应用中使用^[26-27]。自知识图谱被提出后,研究人员在不同领域构建了大量的知识图谱^[28-29]。在法律领域,犯罪知识图谱可用于表述案件详情,刻画罪名相互关系等任务。例如,Zhang等^[30]通过异构知识图谱深度融合法律案件中的实体和事件,进而挖掘出法律事件背后的复杂法律知识。Correia等^[31]对巴西法律文件语料库中的嵌套法律实体进行标注,进而推动法律裁决应用程序的开发。Liu等^[32]通过构建案例三元组,刻画案件之间的关系和频率,有效增强了低频案件的预测效果。尽管之前的研究将犯罪知识图谱应用到了司法任务中,但它们构建三元组知识的效率偏低。与现有研究不同,本文利用通用大语言模型构建了犯罪关系和量刑情节两种类型的犯罪知识图谱,可高效地抽取案件中的重要信息,进而提高法律判决预测的准确性。

3 多被告人法律判决预测

本章首先对法律判决预测问题进行形式化,并提出解决此问题的整体框架;然后介绍利用 LLM 构建犯罪知识图谱,并对检索相似案例方法进行说明;最后基于犯罪知识图谱和先例案件,预测多被告人案件的法律判决结果。

3.1 问题形式化

在多被告人法律判决预测问题中,给定一段描述被告人信息和案件事实的法律判决书,表示为 $s = (\omega_1, \omega_2, \dots, \omega_n)$, 要求预测以下 3 项内容。

1) 罪名: 罪名是对罪犯犯罪行为的高度概括, 它反映了一种罪与另一种罪的本质区别。对于案件 s 中的每一个被告人 d , 他/她所触犯的罪名可以是一个或多个, 表示为 $a = (a_1, a_2, \dots, a_n)$ 。

2) 法条: 法条是法律的条文规定。法官判案时必须根据案件事实找到与之对应的法条, 据此判断被告人所犯罪名及对应的惩罚措施。对于案件 s 中的每一个被告人, 他/她所关联的法条可以是一个或多个, 表示为 $l = (l_1, l_2, \dots, l_n)$ 。

3) 刑期: 刑期是罪犯被关押的时长, 分为拘役和有期徒刑两类。拘役的时间比较短, 适用于较轻的罪行; 有期徒刑的时间相对较长, 适用于较重的罪行。每个被告人对应一个刑期, 使用 t 来表示。

综上所述, 法律判决任务是对于案件 s 中的每一个被告人 d , 给出恰当的法律判决 j , 具体表示为 $j = (a, l, t)$ 。

3.2 整体框架概述

本文利用多轮 LLM 交互和犯罪知识图谱解决多被告人判决预测问题, 该方法的整体框架如图 2 所示。框架的右上部分是以表格形式展示的先例库, 不同的 ID 表示不同的先例。每个先例由犯罪知识图谱和判决结果两部分构成, 犯罪知识图谱借助于 LLM 构建。框架的右下部分为基于图的类案检索, 通过本文设计的检索算法从犯罪知识图谱中查找某一个案例的相似案例, 并返回相关度排序后的结果。框架的左边是多轮 LLM 交互的步骤, 用户与 LLM 以对话的形式完成案件的判决预测。首先, LLM 根据给定的案件事实和用户指令, 分析出罪犯关系和量刑情节; 然后, 用户利用这些结果查找先例, LLM 在参考先例的基础上做出最终的罪名、法条和刑期预测。

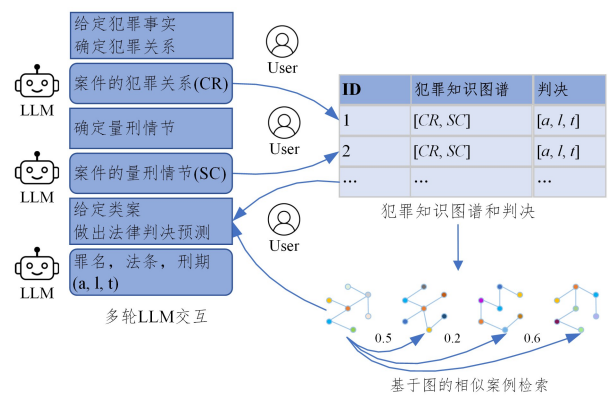


图 2 本文法律判决预测方法的总体框架

Fig. 2 Overall framework of the proposed legal judgment prediction method

3.3 犯罪知识图谱构建

在多轮 LLM 交互中, 本文利用 LLM 强大的信息抽取能力抽取三元组并构建犯罪知识图谱。一方面, 能够加深 LLM 对给定案件的理解, 为后续的预测任务做好准备。另一方面, 能够获取给定案件的关键信息, 后续利用这些信息查找相似案例会更加便利。本文设计了两类犯罪知识图谱: 犯罪关系知识图谱和量刑情节知识图谱。为获得高质量的知识图谱, 选择综合能力比较好的 LLM 面向案件进行信息抽取, 同时在设计 LLM 提示词时要求 LLM 从当前案件的犯罪事实中抽取案件对应的知识图谱, 避免添加其他信息。

3.3.1 犯罪关系知识图谱

在团伙犯罪案件事实描述中, 存在多个被告人, 他们之间相互帮助, 从而完成更大的犯罪活动。Lyu 等^[10]发现挖掘多被告人之间的帮助关系能够辅助判决预测任务, 因此将其作为一个额外的推理任务。受此启发, 本文从案件事实中抽取犯罪关系知识图谱, 从而构建被告人之间的帮助关系图。具体来说, 案件 s 中的多个被告人表示为 $D = (d_1, d_2, \dots, d_n)$, 通过 LLM 判断他们之间是否存在帮助关系。这个过程表示为:

$$KG_{CR} = \{(d_i, r, d_j)\}_{1 \leq i, j \leq n, i \neq j} = LLM(s, D) \quad (1)$$

其中, (d_i, r, d_j) 表示知识图谱中的三元组, d_i 和 d_j 表示两个不同的被告, r 表示帮助关系。需要注意的是, 有时一个案件中的多个被告人之间不存在互相帮助的情况, 这时 KG_{CR} 是一个空集。

图 3 展示了本文调用 LLM 时设计的提示词以及一个案例。最左边的案件事实被嵌入到提示词的 $\langle fact \rangle$ 中, 进而形成 LLM 的主要输入。此外, 为了让 LLM 的输出更加规范统一, 在提示词中增加了关于格式的要求。LLM 的输出经过解析处理后, 多个被告人之间的帮助关系形成了犯罪关系知识图谱。考虑到被告人之间会互相帮助, 本文没有在图谱中指定方向。需要注意的是, 每一个案件对应一个犯罪知识图谱, 如果多个被告人之间没有任何帮助关系, 则该案件对应的犯罪知识图谱是空集。

3.3.2 量刑情节知识图谱

法律案件具有文本较长的特征, 这容易让模型在处理后续内容时忘记较早的内容, 或者无法抓住重点内容。Lyu

等^[10]利用序列到序列模型生成与被告人的量刑情节,进而辅助法律判决任务。受此启发,本文从法律案件中寻找影响判决结果的重要因素。不同的是,本文基于案件本身抽取三元组事实,构建基于量刑情节的知识图谱。这不仅将案件由非结构化转为结构化,而且识别了案件中最重要的情节。具体地,对于一个多被告人案件 s ,使用一个 LLM 自动抽取若干三元组,这一过程表示为:

$$KG_{SC} = \{(head_i, relation_i, tail_i)\}_{1 \leq i \leq n} = LLM(s) \quad (2)$$

其中, KG_{SC} 是量刑情节知识图谱, $(head_i, relation_i, tail_i)$ 是第 i 个三元组。为缩小知识图谱的规模并凸显被告人的重要性,限制头实体必须是某个被告人。因此,这里的 $head_i$ 是某个被告人编号,例如被告人 A。对尾实体没有任何限制。

图 4 展示了量刑情节知识图谱构建的一个案例。最左边的案例与图 3 中的案例相同,该案例也是作为提示词的一部分,替换掉提示词中间的 $\langle fact \rangle$ 。一个 LLM 基于提示词中的案件内容和用户命令,抽取以被告人为头实体的三元组集合。此外,为了让 LLM 的输出规范统一,在提示词中添加了额外的指令。需要注意的是,一个 LLM 也可能会犯错,因此有时抽取的三元组的头实体并不是某一个被告人编号,这种三元组会被丢弃掉。对于每一个案件,所有符合条件的三元组共同构成了该案件对应的量刑情节知识图谱。

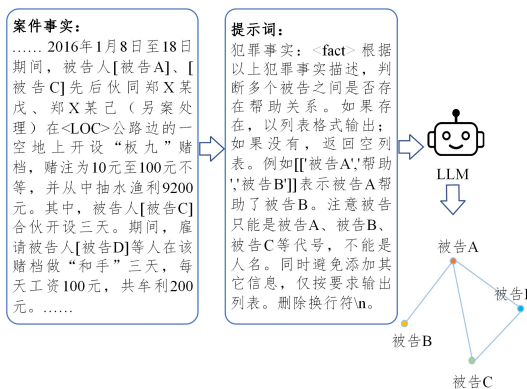


图 3 犯罪关系知识图谱构建案例

Fig. 3 Case on construction of criminal relationships knowledge graph

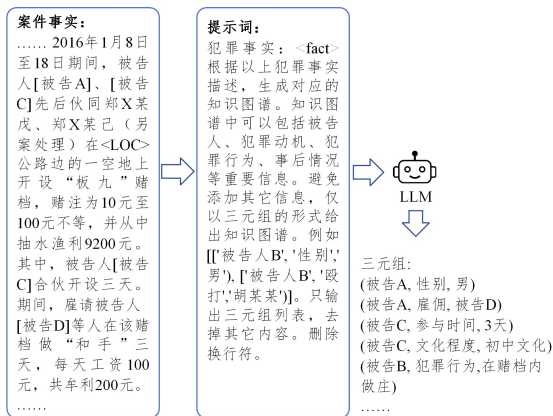


图 4 量刑情节知识图谱构建案例

Fig. 4 Case on construction of sentencing circumstances knowledge graph

经过实验与人工验证, LLM 的输出符合提示词的要求。同时,分析图 3 和图 4 构建的知识图谱可知, LLM 构建的犯罪关系知识图谱和量刑情节知识图谱与案件事实一致,能够用于相似案件的检索。

3.4 相似案例检索

最高人民法院发布的《关于完善统一法律适用标准工作机制的意见》旨在推进类案同判,保证司法的统一和公正。受此意见启发,为寻找一个案件的相似案件,本节设计了基于图的相似案例检索算法,并基于量刑情节知识图谱和犯罪关系知识图谱计算两个不同案例的相似度。

对于量刑情节知识图谱,每个三元组的头实体均为被告人编号,因此本文提出基于被告人的相似度计算方法。具体来说,图 5(a) 中的案件 case1 和案件 case2 均包含多个被告人,表示为 d_i^1 和 d_j^2 ($1 \leq i, j \leq n$)。对于案件 case1 中的每个被告人,计算其与案件 case2 中每个被告人的相似度,记为 $sim(d_i^1, d_j^2)$ 。案件 case1 和案件 case2 的相似度为上述所有相似度求的平均。为了评估两个被告人之间的相似性,将每个被告人的相关信息从量刑情节知识图谱中抽取出来。例如,图 5(b) 中的子图 $G_{d_i^1}$ 和子图 $G_{d_j^2}$ 分别对应案件 case1 中被告人 d_i^1 的信息和案件 case2 中被告人 d_j^2 的信息。这两个子图的中心都是被告人,周围的节点是被告人的性别、学历、犯罪行为等重要信息。为计算两个子图的相似性,借鉴 Jaccard 相似系数的思想,并结合文本模糊匹配算法,设计出如下相似度计算式:

$$sim(d_i^1, d_j^2) = sim(G_{d_i^1}, G_{d_j^2}) = \frac{N_{G_{d_i^1} \cap G_{d_j^2}}}{|G_{d_i^1}| + |G_{d_j^2}|} \quad (3)$$

其中, $N_{G_{d_i^1}}$ 表示子图 $G_{d_i^1}$ 中与子图 $G_{d_j^2}$ 中任一节点相似的节点的总数, $N_{G_{d_j^2}}$ 表示子图 $G_{d_j^2}$ 中与子图 $G_{d_i^1}$ 中任一节点相似的节点的总数, $|\cdot|$ 表示子图的节点总数。

此外,为了计算两个节点的相似度,应用 $textdistance^{1)}$ 中的基于 token 的重叠系数方法,其计算式为:

$$sim(G_{d_i^1}^x, G_{d_j^2}^y) = \frac{|G_{d_i^1}^x \cap G_{d_j^2}^y|}{\min(|G_{d_i^1}^x|, |G_{d_j^2}^y|)} \quad (4)$$

其中, $G_{d_i^1}^x$ 和 $G_{d_j^2}^y$ 分别为子图 $G_{d_i^1}$ 和子图 $G_{d_j^2}$ 中围绕被告节点的节点, $1 \leq x, y \leq n$ 。

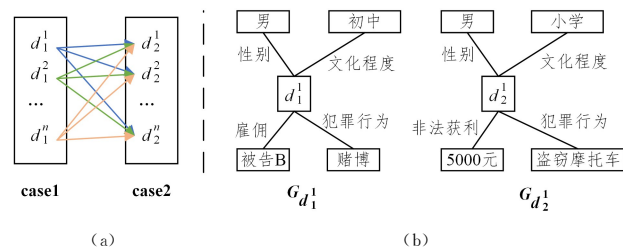


图 5 基于图的相似案例检索实例

Fig. 5 Case on graph-based similar case retrieval

考虑到关系与节点一样,也具有重要的语义知识,将关系文本与节点文本进行拼接,以提高节点相似度计算的准确性。

对于犯罪关系知识图谱,每个三元组的头实体和尾实体均为被告人编号,所有三元组组成的知识图谱蕴含了犯罪团

¹⁾ <https://pypi.org/project/textdistance/>

伙的组织关系。研究发现同类案件的犯罪团伙组织关系是相似的,因此采用 Jaccard 相似系数计算两个犯罪关系知识图谱的相似度,以辅助检索相似案件。以案件 case1 和案件 case2 为例,具体计算式如下:

$$\text{sim}(G_{cr}^1, G_{cr}^2) = \frac{|G_{cr}^1 \cap G_{cr}^2|}{|G_{cr}^1 \cup G_{cr}^2|} \quad (5)$$

其中, G_{cr}^1 和 G_{cr}^2 分别是 LLM 为案件 case1 和案件 case2 生成的犯罪关系知识图谱。如果两个知识图谱完全相同,则相似度为 1;如果完全不同,则相似度为 0。

为了整合由量刑情节知识图谱计算的相似度和由犯罪关系知识图谱计算的相似度,基于一个权重将它们相加:

$$\text{sim}(\text{case1}, \text{case2}) = \alpha \frac{\sum_{i=1}^n \sum_{j=1}^n \text{sim}(d_i^1, d_j^2)}{n} + (1 - \alpha) \text{sim}(G_{cr}^1, G_{cr}^2) \quad (6)$$

其中, α 是一个自定义的权重,用来调整两种犯罪关系知识图谱的重要性。

基于以上计算式,算法 1 和算法 2 梳理了相似案例检索的所有步骤。对于测试集的每一个案例,均通过这两个算法计算它的相似案例。在算法 1 中,首先根据式(1)和式(2)获得训练数据和测试案件 case1 的犯罪关系知识图谱和量刑情节知识图谱(第 1—6 行),然后计算 case1 与训练数据中每个案件 case2 的相似度(第 7—16 行)。基于量刑情节知识图谱计算 case1 和 case2 的相似度时,算法 1 根据算法 2 计算 case1 与 case2 之间任意被告人组合的相似度,并取平均值(第 8—13 行)。同时,算法 1 根据式(5)计算 case1 和 case2 在犯罪关系知识图谱中直接的相似度(第 14 行)。式(6)利用上述两类知识图谱的相似度计算 case1 和 case2 之间的相似度(第 15 行)。算法 1 通过综合案件的犯罪关系知识图谱相似度和量刑情节知识图谱相似度实现了相似案例的检索。

算法 1 相似案例检索算法

输入:测试集中的一个案件 case1、整个训练集 $\text{Data}_{\text{train}}$

输出:排序后的相似案例 similar_cases

```

1. for case c in  $\text{Data}_{\text{train}}$  do
2.   根据式(1)获得 c 的犯罪关系知识图谱
3.   根据式(2)获得 c 的量刑情节知识图谱
4. end for
5. 根据式(1)获得 case1 的犯罪关系知识图谱
6. 根据式(2)获得 case1 的量刑情节知识图谱
7. for a case case2 in  $\text{Data}_{\text{train}}$  do
8.   for a defendant  $d_1^i$  in case1 do
9.     for a defendant  $d_2^j$  in case2 do
10.      根据算法 2 计算 case1 中第 i 个被告人与 case2 中第 j 个被告人的相似度  $\text{sim}(d_1^i, d_2^j)$ 
11.    end for
12.  end for
13.  根据  $\frac{\sum_{i=1}^n \sum_{j=1}^n \text{sim}(d_1^i, d_2^j)}{n}$  计算所有被告人的平均相似度

```

```

14.  根据式(5)计算 case1 和 case2 对应的犯罪关系知识图谱之间的相似度  $\text{sim}(G_{cr}^1, G_{cr}^2)$ 
15.  根据式(6)计算 case1 和 case2 之间的相似度  $\text{sim}(\text{case1}, \text{case2})$ 
16. end for
17. 根据所有的相似度  $\text{sim}(\text{case1}, \text{case2})$  对  $\text{Data}_{\text{train}}$  排序,排序后为 similar_cases
18. return similar_cases

```

算法 2 两个被告人之间的相似度计算算法

输入:案件 case1 中 d_1^i 对应的子图 $G_{d_1^i}$ 、案件 case2 中 d_2^j 对应的子图 $G_{d_2^j}$

输出:子图 $G_{d_1^i}$ 和 $G_{d_2^j}$ 之间的相似度 sim_{ij}

```

1. 将子图  $G_{d_1^i}$  中的被告人节点  $d_1^i$  删除
2. 将子图  $G_{d_2^j}$  中的被告人节点  $d_2^j$  删除
3. for node  $G_{d_1^i}^x$  in  $G_{d_1^i}$  do
4.   for node  $G_{d_2^j}^y$  in  $G_{d_2^j}$  do
5.     根据式(4)计算  $G_{d_1^i}^x$  和  $G_{d_2^j}^y$  之间的相似度  $\text{sim}(G_{d_1^i}^x, G_{d_2^j}^y)$ 
6.     if  $\text{sim}(G_{d_1^i}^x, G_{d_2^j}^y) > \lambda$  then
7.        $N_{G_{d_1^i}} \leftarrow N_{G_{d_1^i}} + 1$ 
8.        $N_{G_{d_2^j}} \leftarrow N_{G_{d_2^j}} + 1$ 
9.     end if
10.  end for
11. end for
12. 根据式(3)计算子图  $G_{d_1^i}$  和子图  $G_{d_2^j}$  之间的相似度  $\text{sim}_{ij}$ 
13. return  $\text{sim}_{ij}$ 

```

3.5 多被告人法律判决预测

在多被告人法律判决预测中,本文首次采用多轮 LLM 交互的方法,这与之前的相关工作都不相同。在图 2 中,用户在多轮交互中首先让 LLM 根据图 3 和图 4 的提示词生成一个测试案件的犯罪关系知识图谱和量刑情节知识图谱。然后,基于 3.4 节的算法 1 和算法 2 寻找测试案件在训练集中的相似案件,它们也被称为类案。需要说明的是,训练集中的案件的两类犯罪知识图谱已经提前构建,可以在多轮 LLM 交互中随时使用。最后,基于先例案件和指令,LLM 预测该测试案件的罪名、法条和刑期。具体的指令为:

“分析以下几个相似案例及其判决结果:〈precedents〉当前案件的事实描述为:〈fact〉,给出当前案件的判决结果,以字典格式输出,例如{‘[被告 A]’: {‘accusations’: [‘开设赌场罪’], ‘laws’: [303], ‘term’: ‘有期徒刑七个月’}}。字典中每个被告只能出现一次。案件中所有的被告都需要给定其判决结果,同时避免添加其他信息,仅按要求输出字典。删除换行符。”

其中〈precedents〉为类案,〈fact〉为案件事实。

4 实验

4.1 数据集

单个被告人的公开犯罪数据集非常多,但多被告人的公开犯罪数据集比较少。本文采用 Lyu 等^[10]构建的一个多被告人案件数据集 MultiLJP,该犯罪数据集来源于中国裁判文书网¹⁾,由法律专业人员精心标注,且进行了匿名化和脱敏处

¹⁾ <https://wenshu.court.gov.cn/>

理。表 1 列出了该数据集的详细信息。

表 1 MultiLJP 数据集统计

Table 1 Statistics of MultiLJP dataset

统计项	训练集	测试集
案件数量	18968	2370
案件平均长度	3049.12	2890.96
单案件被告最大数	37	41
单案件被告最小数	2	2
单案件被告平均数	3.38	3.41
单案件罪名最大数	8	3
单案件罪名最小数	1	1
单案件罪名平均数	1.06	1.06
单案件法条最大数	8	3
单案件法条最小数	1	1
单案件法条平均数	1.06	1.06
罪名种类数	23	23
法条种类数	22	22
刑期种类数	11	11

4.2 实验设置

本文实验在 ubuntu 20.04 服务器上完成,代码运行环境是 python 3.11。在多次实验对比后,将式(6)中的权重 α 设置为 0.9,以取得最好的法律判决预测结果。此 α 值表明量刑情节知识图谱对计算相似案件的影响更大,而犯罪关系知识图谱对计算相似案件的影响较小。算法 2 中的 λ 为 0.8,即计算两个被告人的相似性时,若他们所对应的知识图谱间两个节点的相似度大于 0.8,则算法会认为两个节点相似。 λ 值小于 1,能够避免两个被告人相似性为 0。

对于图 2 总体框架中的 LLM,选用 OpenAI 开发的 gpt-4o-mini 模型。该模型具有较好的推理性能,且其 API 价格比较便宜,便于在大量数据集上做更多的实验。同时,为进一步降低成本,采用批处理 API,它虽然不是即时的交互,但是对生成的结果没有任何影响。此外,原始数据集中的验证数据集没有在本实验中使用,一是因为本文方法不是端到端模型,不需要验证集调整模型参数;二是因为训练集已经包含了足够多的数据用于类案检索。在评估指标方面,采用宏精确率、宏召回率和宏 F1,以便和其他多被告人预测研究进行比较。在评估罪名预测结果时,与 Lyu 等^[10]对罪名的处理相同,本文也对走私、贩卖、运输和制造毒品这些罪名进行了合并,具体做法是将这些罪名进行统一编码。

4.3 对比模型

在与其他方法进行比较时,本文选择了两大类模型:用于单被告人犯罪预测的模型和用于多被告人犯罪预测的模型。为了适应多被告人的场景,面向单被告人的模型会被适当改造。具体来说,将一个被告的名字和犯罪文本拼接到一起作为单被告人模型的输入,使得该模型也可以预测多个被告人。下面是本文选择的可以被改造的单被告人犯罪预测模型。

1)MPBFN。Yang 等^[33]提出了一种多视角双反馈网络(MPBFN),用于解决法律判决预测中多个子任务之间依赖关系没有高效利用的问题,同时该网络加入了注意力机制以便

为相似案例描述预测不同的刑期。

2)LADAN。为解决法律案件容易被误判的问题,Xu 等^[34]设计了可对法律文档进行提炼的注意力网络(LADAN),该网络通过图神经网络学习法律条款的细微区别,并利用它们对案件进行判决。

3)NeurJudge。考虑到判决环境对法律判决的重要影响,Yue 等^[35]通过子任务的预测结果抽取案件事实中的关键环境因素,并通过一个环境感知的神经网络 NeurJudge 完成罪名、法条和刑期的预测。

以下是本文选择的可以直接用于多被告人犯罪预测的模型。

1)BERT。为进一步提供 BERT 在各种 NLP 任务上的性能,Cui 等^[36]首次将全词掩码加入到预训练的中文 BERT 模型中。

2)Lawformer。Xiao 等^[7]发现在法律领域应用的主流预训练语言模型存在 tokens 过长的问题,因此他们研发了专门用于中文法律长文本的 Lawformer 模型,此模型适用于判决预测、法律问答等多种任务。

3)mT5。Text-to-Text Transfer Transformer(T5)模型在英语自然语言处理任务上取得了非常好的效果,Xue 等^[37]将其推广到 101 种语言,形成 mT5,并在 benchmarks 上对 mT5 进行了验证。

4)MAMD。为解决多被告人犯罪事实识别困难的问题,Pan 等^[9]采用局部注意力让模型关注被告人名字周围的描述,同时采用全局注意力让模型关注单个被告人的罪名预测结果。

5)HRN。针对单被告人的法律判决预测模型难以适应多被告人的场景,Lyu 等^[10]提出了层次推理网络,该网络能先后生成犯罪关系、量刑情节和判决预测结果。

4.4 实验结果和分析

本节首先将所提模型与其他模型进行了法律判决预测的对比实验,其次探究了多轮交互和犯罪知识图谱的重要作用,然后对选择最佳类案数量进行了多次实验,最后通过一个案例研究对法律判决效果进行了更详细的说明。

4.4.1 整体实验结果

本文在 MultiLJP 数据集上进行大量的对比实验,实验的评估指标是宏精确率、宏召回率和宏 F1。表 2 列出了对比模型和本文模型在 3 个子任务上的法律判决预测结果。分析可知,单被告人模型不能适应多被告人的场景,因为其没有考虑到多被告人之间的关系,同时预测某个被告人犯罪情况时,会受到其他被告人犯罪事实的影响。对比 3 个子任务后发现,罪名和法条的预测比刑期的预测相对容易一些,因为前两个是多标签分类问题,类别数量是固定的,最后一个是多分类问题,类别数量不固定,因此刑期预测任务更难一些。从宏 F1 值的视角看,相比其他模型,本文模型的整体效果最好,说明图 2 所示框架是有效的。此外,HRN 是一种具有竞争力的模型,其召回率较高,预测覆盖较为全面,但在精确率上有所牺牲,因此整体 F1 值低于本文模型。

表 2 基线模型和本文模型的法律判断预测结果

Table 2 Legal judgment prediction results of the baseline models and the proposed model

模型	罪名			法条			刑期		
	精确率	召回率	F1	精确率	召回率	F1	精确率	召回率	F1
MPBFN	32.79	33.20	31.59	34.73	34.22	34.35	21.59	20.91	20.70
LADAN	20.68	32.42	24.74	38.09	22.49	26.64	24.05	23.43	23.16
NeurJudge	34.19	25.36	27.55	41.72	36.96	38.15	27.56	25.63	25.95
BERT	36.80	20.10	25.14	34.19	29.68	30.70	23.95	22.68	21.55
Lawformer	38.97	29.12	32.76	36.18	35.33	34.00	22.66	20.68	18.30
mT5	63.33	49.94	52.86	74.28	53.65	58.84	39.13	24.23	23.04
MAMD	33.00	34.15	31.60	—	—	—	—	—	—
HRN	71.80	71.83	70.70	69.87	70.95	69.20	41.33	40.20	40.62
Ours	82.61	67.96	72.34	77.27	65.74	69.47	50.00	34.28	40.67

4.4.2 多轮交互的重要作用

为了验证多轮 LLM 交互的重要作用,在图 2 原始模型的基础上设计了一些对比实验。首先将多轮交互改为单轮交互,即用户仅向 LLM 提问一次。然而这里还涉及到类案的去留问题,为此将单轮交互进一步细分为无类案和有类案两类。实验结果如表 3 所列。与单轮交互相比,多轮交互对判决结果有明显的提升,这一点在罪名和法条方面体现得较明显。例如,与不带类案的单轮交互设置相比,多轮交互设计能让 F1 值提升 12.91%。对于刑期预测而言,多轮交互的帮助

作用有限,这表明该子任务的难度更高。

与无类案单轮交互相比,有类案单轮交互在 3 个子任务上的 F1 值分别有 6.29%,4%和 2.2%的提升。这说明类案为 LLM 提供了有价值的信息,LLM 可以参考类案的判决结果,进而做出更准确的判决。此外,不同设置实验的精确度几乎稳定不变,例如罪名预测任务的精确度均为 82.61%,这是因为本文采用的宏精确度对实验结果的变化不敏感。与此相反的是,宏召回率会随着预测结果的微小变化而发生改变。

表 3 单轮交互与多轮交互的判决结果对比

Table 3 Comparison of judgment results between single-turn interaction and multi-turn interaction

设置	罪名			法条			刑期		
	精确率	召回率	F1	精确率	召回率	F1	精确率	召回率	F1
单轮无类案	82.61	50.12	59.43	72.73	50.30	57.89	50	31.23	38.45
单轮有类案	82.61	58.44	65.72	77.27	55.53	61.89	50	34.24	40.65
多轮	82.61	67.96	72.34	77.27	65.74	69.47	50	34.28	40.67

4.4.3 犯罪知识图谱的影响

表 3 的对比实验证明了类案对法律判决预测有很大影响,同时,算法 1 和算法 2 利用两种犯罪知识图谱寻找类案,因此需要探究犯罪知识图谱对判决结果的影响。首先分别去掉犯罪关系知识图谱和量刑情节知识图谱,即在式(6)中仅使用一种相似度,然后测试模型的预测效果。此外,如果把两种犯罪知识图谱都去除,则不再有相似案例,多轮交互也变为单轮交互,这等同于表 3 中的不带类案的单轮交互。

从表 4 的实验结果可以看出,去掉犯罪关系知识图谱或量刑情节知识图谱都会让 F1 值下降,这说明两种类型的犯

罪知识图谱都能够改善相似案例的排序结果,从而为 LLM 提供更准确的类案以及类案的判决。与不使用犯罪关系知识图谱相比,不使用量刑犯罪知识图谱时 3 个子任务的 F1 值更低,说明量刑犯罪知识图谱包含更多语义信息,对类案检索帮助更大。在 3 个子任务中,罪名预测和法条预测在没有犯罪知识图谱协助时,其 F1 值下降更明显,说明 LLM 在做这两个子任务时非常依赖类案的判决结果。这种参考类案的判案方式与现实世界中法官的判案逻辑是一致的。此外,与表 3 一样,表 4 中的宏精确度也对实验结果的变化不敏感,所以精确度几乎稳定不变。

表 4 犯罪知识图谱对判决结果的影响

Table 4 Impact of criminal knowledge graphs on judgment results

设置	罪名			法条			刑期		
	精确率	召回率	F1	精确率	召回率	F1	精确率	召回率	F1
w/o CR	82.61	56.01	64.60	77.27	52.99	60.75	50	32.01	39.03
w/o SC	82.61	55.77	64.47	77.27	51.61	59.04	50	31.81	38.88
w/o ALL	82.61	50.12	59.43	72.73	50.30	57.89	50	31.23	38.45
ALL	82.61	67.96	72.34	77.27	65.74	69.47	50	34.28	40.67

4.4.4 相似案例数量的影响

本文还探究了不同相似案例数量对判决预测结果的影响。考虑到 LLM 对输入长度的限制,实验采用的类案数量为 0~4,结果如图 6 所示。对于 3 个子任务,当类案数量为 2 时,它们的 F1 值达到最高。这表明 1 个类案提供的参考判

决结果太少,需要增加类案的数量,以便为 LLM 提供更多判决依据。然而,类案数量超过 2 个时,不同的判决结果给 LLM 造成了一定程度的困惑,使 LLM 难以做出正确的判决。过多的类案中可能存在相似度较低的案件,给最终 LLM 的判决预测带来负面影响。这说明过多的类案并不能更好地

帮助 LLM 做判决预测任务,因此只需要提供少量高质量的相似案例即可。

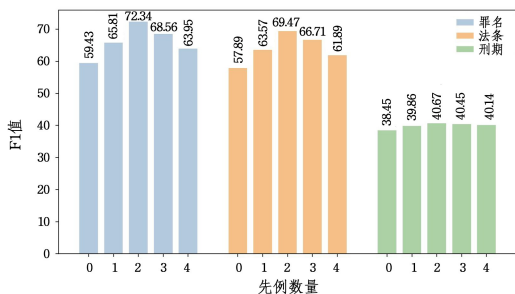


图 6 当 LLM 参考不同数量的先例时 3 个子任务的 F1 值比较

Fig. 6 Comparison of F1 scores for three subtasks when LLM references different number of precedents

4.4.5 案例分析

本节借助一个案例来说明本文模型相比其他模型所拥有的优势。在表 5 中,HRN 和本文模型都对一个刑事犯罪案件进行了法律判决预测。从中可以发现,HRN 难以区分十分相似的寻衅滋事罪和故意伤害罪,在本案例中受到“殴打”等描述的影响,错误地将本案分类为寻衅滋事罪。而本文模型基于当前案例的相似案例判定该案例是故意伤害罪,原因是案件中存在故意非法损害他人身体健康的行为。同时,HRN 和本文模型对被告 A 和 B 的刑期预测都不准,说明刑期预测任务是非常有挑战性的,但本文模型预测的刑期和实际的刑期更接近。此外,类案中的被害人为轻伤二级,而本案的被害人为重伤二级,因此本案被告人的刑期比类案被告人的刑期更长。

表 5 法律判决预测案例分析

Table 5 Case study of legal judgment prediction

案件事实:2015年5月18日17时55分许,被告人[被告A]、[被告B]在(LOC)体育场福X楼下,因被告人[被告B]怀疑被害人A之前偷过自己的钱,辱骂并殴打被害人,被害人预离开时,被告人[被告A]将其拉住,后被告人[被告B]脱下鞋子在被害人身上、胳膊上、头部击打,被告人[被告A]对被害人拳打脚踢,致被害人腹部闭合性损伤(小肠破裂)、急性弥漫性腹膜炎、慢支并双下肺感染、切口感染住院治疗。经鉴定,被害人A之损伤程度属重伤二级。
案件判决:被告A(故意伤害罪;234;4年),被告B(故意伤害罪;234;3年6个月)
HRN的判决:被告A(寻衅滋事罪;293;1年),被告B(寻衅滋事罪;293;1年)
类案:2014年3月31日零时许,被害人谭某红等人在位于(LOC)的拉飞歌酒吧有限公司皇宫俱乐部消费期间,因怀疑被李某平非礼而欲对李某平进行殴打,经俱乐部工作人员劝阻,李某平离开了现场。后谭某红等人认为俱乐部工作人员有意放纵李某平继而在俱乐部正门后与被告人[被告A]等人发生争执。并用手卡住[被告A]脖子,[被告A]遂指使被告人[被告B]等保安人员推搡和用拳脚对谭某红、张某河等人实施殴打,致使谭某红鼻骨、鼻中隔骨折。经鉴定,被害人谭某红的伤情属轻伤二级。
类案的判决:被告A(故意伤害罪;234;6个月),被告B(故意伤害罪;234;6个月)
本文模型的判决:被告A(故意伤害罪;234;3年),被告B(故意伤害罪;234;2年)

结束语 本文提出了融合多轮 LLM 和犯罪知识图谱的统一框架。该方法首先借助 LLM 建立案件的犯罪关系知识图谱和量刑情节知识图谱,然后通过类案检索算法计算案件的类案,最后利用 LLM 预测罪名、法条和刑期。大量对比实验证明,本文模型的预测效果优于多种单被告人和多被告人犯罪预测模型。在消融实验中,验证了多轮 LLM 交互和犯罪知识图谱的重要作用。此外,还探究了类案数量的最佳选择,并分析了一个具体的案例。

参考文献

[1] TONG S X, YUAN J L, ZHANG P L, et al. Legal Judgment Prediction via graph boosting with constraints[J]. Information Processing and Management, 2024, 61(2):103663.

[2] ZHANG C Y, QU H, CUI C R, et al. Process Supervision Based Sequence Multi-task Method for Legal Judgement Prediction[J]. Computer Science, 2021, 48(3):227-232.

[3] GAN L L, LI B K, KUANG K, et al. Exploiting Contrastive Learning and Numerical Evidence for Confusing Legal Judgment Prediction[C]// Conference on Empirical Methods in Natural Language Processing, 2023:12174-12185.

[4] FENG G Y, QIN Y B, HUANG R Z, et al. Criminal Action Graph: A semantic representation model of judgement documents for legal charge prediction[J]. Information Processing and Management, 2024, 60(5):103421.

[5] SUKANYA G, JAYARAJU P. Modified Hierarchical-Attention Network model for legal judgment predictions[J]. Data &

Knowledge Engineering, 2023, 147:102203.

[6] ZHAO Q H, GAO T H, GUO N. LA-MGFM: A legal judgment prediction method via sememe-enhanced graph neural networks and multi-graph fusion mechanism[J]. Information Processing and Management, 2023, 60(5):103455.

[7] XIAO C J, HU X Y, LIU Z Y, et al. Lawformer: A pre-trained language model for Chinese legal long documents[J]. AI Open, 2021, 2:79-84.

[8] ZHOU Z, SHI J X, SONG P X, et al. LawGPT: A Chinese Legal Knowledge-Enhanced Large Language Model[J]. arXiv: 2406.04614, 2024.

[9] PAN S P, LU T, GU N, et al. Charge Prediction for Multi-defendant Cases with Multi-scale Attention[C]// Computer Supported Cooperative Work and Social Computing, 2019:766-777.

[10] LYU Y G, HAO J T, WANG Z H, et al. Multi-Defendant Legal Judgment Prediction via Hierarchical Reasoning[C]// Findings of the Association for Computational Linguistics, 2023:2198-2209.

[11] QIN W C, CAO Z L, YU W J, et al. Explicitly Integrating Judgment Prediction with Legal Document Retrieval: A Law-Guided Generative Approach[C]// The 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2024:2210-2220.

[12] KORT F. Predicting Supreme Court Decisions Mathematically: A Quantitative Analysis of the “Right to Counsel” Cases[J]. The American Political Science Review, 1957, 51(1):1-12.

- [13] FENG Y, LI C Y, NG V. Legal Judgment Prediction via Event Extraction with Constraints[C]// Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. 2022;648-664.
- [14] LIU Y F, WU Y Q, ZHANG Y T, et al. ML-LJP: Multi-Law Aware Legal Judgment Prediction[C]// The 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2023;1023-1034.
- [15] CHALKIDIS I, FERGADIOTIS M, MALAKASIOTIS P, et al. LEGAL-BERT: The Muppets straight out of Law School[J]. arXiv:2010.02559, 2020.
- [16] BELTAGY L, PETERS M, COHAN A. Longformer: The Long-Document Transformer[J]. arXiv:2004.05150, 2020.
- [17] OUYANG L, WU J, JIANG X, et al. Training language models to follow instructions with human feedback[C]// Annual Conference on Neural Information Processing Systems 2022. 2022; 27730-27744.
- [18] SUN Q, HUANG K, YANG X C, et al. Consistency Guided Knowledge Retrieval and Denoising in LLMs for Zero-shot Document-level Relation Triplet Extraction[C]// The ACM on Web Conference. 2024;4407-4416.
- [19] JI B J, DUAN X Y, QIU Z Y, et al. Submodular-based In-context Example Selection for LLMs-based Machine Translation[C]// The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation. 2024; 15398-15409.
- [20] AGRAWAL G, PAL K, DENG Y L, et al. CyberQ: Generating Questions and Answers for Cybersecurity Education Using Knowledge Graph-Augmented LLMs[C]// Thirty-Eighth AAAI Conference on Artificial Intelligence. 2024;23164-23172.
- [21] HUANG Q Z, TAO M X, AN Z W, et al. Lawyer LLaMA Technical Report[J]. arXiv:2305.15062, 2023.
- [22] WU Y Q, ZHOU S Y, LIU Y F, et al. Precedent-Enhanced Legal Judgment Prediction with LLM and Domain-Model Collaboration[C]// The 2023 Conference on Empirical Methods in Natural Language Processing. 2023;12060-12075.
- [23] SHUI R H, CAO Y X, WANG X, et al. A Comprehensive Evaluation of Large Language Models on Legal Judgment Prediction[C]// Findings of the Association for Computational Linguistics: EMNLP 2023. 2023;7337-7348.
- [24] DENG C L, MAO K L, ZHANG Y Y, et al. Enabling Discriminative Reasoning in LLMs for Legal Judgment Prediction[J]. arXiv:2407.01964, 2024.
- [25] WEI J, WANG X Z, SCHUURMANS D, et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models[C]// Annual Conference on Neural Information Processing Systems. 2022;24824-24837.
- [26] ZHANG Z P, WANG A Q, ZHANG Y, et al. Relation pruning and discriminative sampling over knowledge graph for long-tail recommendation[J]. Information Sciences, 2024, 679; 120871.
- [27] LIU H Z, ZHU Y, WU Z H. Knowledge Graph-Based Behavior Denoising and Preference Learning for Sequential Recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(6); 2490-2503.
- [28] HU Y L, ZOU F T, HAN J J, et al. LLM-TIKG: Threat intelligence knowledge graph construction utilizing large language model[J]. Computers & Security, 2024, 145; 103999.
- [29] FAN Z L, CHEN C C. CuPe-KG: Cultural perspective-based knowledge graph construction of tourism resources via pre-trained language models[J]. Information Processing and Management, 2024, 61(2); 103646.
- [30] ZHANG Y X, ZHAI S L, MENG Y, et al. Event is more valuable than you think: Improving the Similar Legal Case Retrieval via event knowledge[J]. Information Processing and Management, 2024, 61(4); 103729.
- [31] CORREIA F, ALMEIDA A, et al. Fine-grained legal entity annotation: A case study on the Brazilian Supreme Court[J]. Information Processing and Management, 2022, 59(1); 102794.
- [32] LIU D G, DU W H, LI L, et al. Augmenting Legal Judgment Prediction with Contrastive Case Relations[C]// The 29th International Conference on Computational Linguistics. 2022; 2658-2667.
- [33] YANG W M, JIA W J, ZHOU X J, et al. Legal Judgment Prediction via Multi-Perspective Bi-Feedback Network[C]// The Twenty-Eighth International Joint Conference on Artificial Intelligence. 2019;4085-4091.
- [34] XU N, WANG P H, CHEN L, et al. Distinguish Confusing Law Articles for Legal Judgment Prediction[C]// The 58th Annual Meeting of the Association for Computational Linguistics. 2020; 3086-3095.
- [35] YUE L N, LIU Q, JIN B B, et al. NeurJudge: A Circumstance-aware Neural Framework for Legal Judgment Prediction[C]// The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021;973-982.
- [36] CUI Y M, CHE W X, LIU T, et al. Pre-Training With Whole Word Masking for Chinese BERT[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing, 2021, 29; 3504-3514.
- [37] XUE L T, CONSTANT N, ROBERTS A, et al. mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer[C]// The 2021 Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies. 2021;483-498.



WANG Dongsheng, born in 1992, Ph.D., lecturer. His main research interests include natural language processing and knowledge graph.