北大元法智能系统技术原理报告

北京大学法律人工智能实验室

一、问题和背景

随着人工智能时代的到来，基于机器学习和自然语言处理的智能系统被越来越多地应用于法律领域，以提供更便捷、更优质的法律服务。[1]然而，现有的法律智能系统更侧重于案件检索、合同审查和法律判决预测等法律细分领域。这些领域的服务对象更多是法律职业人士，并不直接面向社会公众。对于普通民众在生活中遇到的婚姻、借贷、租赁等法律问题，更多需要以法律咨询为主要形式的法律服务。而现有的法律智能系统，主要依靠传统的专家系统方法来构架法律咨询服务。这样的法律智能系统难以应对普通民众口语化、生活化的提问，构建成本也较高，知识的容量和可扩展性有限，除了在参考意义上提供相关知识以外，事实上难以通过做出终局性的逻辑推理结论以解决真实咨询场景中的法律问题。在日常生活中，普通民众仍然主要依靠人工的律师服务来获取法律援助和法律咨询。

大语言模型（Large Language Model）的出现为人工智能驱动的法律咨询提供了启示。GPT-4 被证明可以通过美国律师资格考试，这表明大语言模型可以在一定程度上掌握法律知识和能力。[2]而且，大语言模型的文本生成能力与对话生成能力与法律咨询的形式天然契合。但是，通用领域的 GPT、LLaMA、文心一言等大语言模型，存在缺乏法律知识、模型幻觉等问题，难以切中问题要害而直接应用于法律场景中。[3]近来，有研究者在开源大模型的基础上使用中文法律语料进行微调，提高了大模型在法律人工智能领域的能力。但是，对于真实人类世界的法律咨询而言，这些微调的法律大模型仍然存在不足。首先，微调模型没有在海量法律数据上进行充分预训练，没有掌握充足的法律知识。其次，这些微调模型对法律问题的分析和推理，往往是基于概率式的，难以严格符合法律规定和法律思维推理方式。最后，现实人类世界的法律咨询对话，并不总是用户提问、律师回答的结构。律师在理解当事人的法律问题后，也会主动向用户提问，以厘清当事人未说明或者未说明清楚的事实，进而据此做出解决问题的结论和方案。而这些关键的工作，在现有法律智能系统的研究中，仍是未被充分探索的问题。

为了解决这些悬而未决的问题，我们提出了将法律推理引擎与大语言模型相结合的方法，并根据该方法自主研发了能够进行法律问答咨询的智能系统——北大元法智能系统。相对其

他的微调法律大模型和通用领域大模型，北大元法智能系统基于首创的规则推理引擎技术和全栈自研的中国法律大模型，能够解决现有法律智能系统和大模型在真实法律服务场景的缺陷。在可预见的未来，元法智能系统不仅能应用于公共法律服务，还能够广泛应用于到立法、执法、司法、企业法务等法律场景，从而实现通用的法律智能技术，构建通用的法律智能系统，推动法律和社会治理的智能化，实现智能社会的良法善治。

二、技术原理

（一） 自研法律大模型

元法智能系统的法律知识来源于自研法律大模型。自研法律大模型是在法律领域数据上从头进行充分预训练的大语言模型。自研法律大模型是元法智能系统的一个重要组成部分，作为知识基础，决定了元法智能系统中法律知识的边界、要素和构造，以及静态的法律知识要素之间的关联关系。自研法律大模型的参数量超过了 100 亿，在海量通用语料和法律语料上进行了充分的预训练。模型的法律语料包含上亿篇权威法律文本，包括法律、司法解释、规范性文件、裁判文书、合同文本、法学资料、法律咨询数据等，确保了对法律术语、概念和理论的全面覆盖，使模型具备了法律知识的记忆、理解和运用等能力基础。

以法条背诵为例，法条背诵能够反映模型的法律知识记忆能力，能否记住了必要的法律概念、术语和法条。这对于其他通用领域大模型和微调法律模型而言，是一个困难的问题。对于通用大模型而言，其训练语料多为公开领域的百科、书籍、论文，其中并不包含法律法规、裁判文书，法律类语料的占比并不高。例如，对于开源模型 LLaMA，其预训练语料主要来自公开的数据集、维基百科、公开的论文库与代码库。[4]这就导致通用领域大模型在法律领域表现不佳。我们进一步对 GPT-4 在法律咨询中的表现进行了人工评测，使用 1000 条公共法律服务问题进行测试，结果显示 GPT-4 引用法规的准确率仅有 51%。

微调法律模型的表现更为不佳。在中文法律大模型评测指标上，现有微调法律模型背诵法条的准确率不足 15%。[5]微调法律模型的范式是较为流行的法律大模型研发范式，其核心方法是以开源的通用大模型作为基座，使用法律数据进行二次预训练和微调。在国内，已经有诸多研究者开源了微调的法律模型，例如 Lawyer LLaMa、ChatLaw、智海-录问、夫子-明察。但是，这些微调法律模型在法律知识记忆、理解和应用这三个层面，在相关的法律大模型评测指标上与 GPT 相比，仍是 GPT 的表现更佳。[6]这说明了基于开源大模型进行微调

的范式并不适合法律领域。

自研法律大模型并非沿用上述微调法律模型的范式，而是从头进行预训练，在预训练阶段使用的法律数据远超过 GPT 等通用模型。自研法律大模型在预训练阶段实现了对上万部法规的内化，极大提升了法律知识的记忆能力，从而避免在法律咨询中引用错误的法规，避免大模型幻觉现象的出现。

（二）基于一阶逻辑的法律推理引擎

法律推理引擎是北大元法智能系统的关键部分，决定了知识要素在解决问题的过程中的动态关联关系，构建了模型运用法律知识的思维规则、路径和工程机制，从而确定了本系统的基本属性以及与其他微调法律模型和通用大模型之间在属性上的本质区别。大语言模型虽然能解决幻觉问题，但是仍无法保证智能法律系统符合法律规则和法律推理思维逻辑。而且，大语言模型的决策过程是黑盒，不具备可解释性。因此，元法智能系统为了能够应用于真实的法律服务场景，将自研法律大语言模型和基于一阶逻辑的法律推理技术进行结合，使得法律智能系统的推理符合法律人思维逻辑，具有了完全可解释性。

一阶逻辑（First Order Logic）能够将基于自然语言表述的法律推理进行形式化，转换为一阶逻辑符号形式的形式化推理，模拟真实的法律咨询的法律推理。系统能将《中华人民共和国民法典》等法律的条文表示为一阶逻辑表达式，捕捉了法条之间的逻辑结构和层级关系，进一步将多个法律规则组合成类似于树状的推理结构，并基于符号化的推理算法，构建法律推理引擎。

结合一阶逻辑推理引擎和大语言模型，元法智能系统可以实现准确的多步法律推理过程，并让用户参与互动式的多轮对话，从而获取法律推理过程中的缺失信息。在咨询过程中，系统会根据提取的事实动态地遍历推理树，根据推理树提出相关的后续问题，以收集更多事实，最后得出结论，这模拟了人类律师向客户反复询问和推理的过程。这样，法律智能系统就能像人类法律专家一样，不再通过大模型的概率生成，而是通过可解释的逐步推导得出结论。

元法智能系统进一步将上述两个路径进行融合，以法律推理引擎的过程和结果作为训练数据，将其加入自研法律大模型的训练过程，指导法律大模型的生成。法律大模型与法律推理引擎结合的混合推理方法，使元法智能系统更加接近现实世界中的法律咨询场景，实现了更自然的人机交互，能够应用于真实的法律咨询问答服务场景。

三、结语

北大元法智能系统是基于北京大学法学院和北京大学人工智能研究院的学术资源和条件，根据北京大学整体的工作规划，由北京大学法律人工智能实验室研究团队实施研发的智能法律工程技术成果，也是北京大学重要的“人工智能+”交叉学科学术研究成果。本研究得到了北京大学武汉人工智能研究院国家智能社会治理实验综合基地开放课题“东湖高新区智能公共法律咨询服务系统”（ZX2222M）的资助，也得到了多级政府部门、法律实务界、学术界、相关产业领域的关注和支持。面向未来，北京大学法律人工智能实验室将以北大元法智能系统为基础，继续完善和深入研发通用法律智能技术，赋能社会治理的智能化，助力智能时代中国式现代化的法治建设。

---

[1] Surden, Harry. "Artificial intelligence and law: An overview." Georgia State University Law Review 35 (2019): 19-22.

[2] Katz, Daniel Martin, et al. "Gpt-4 passes the bar exam." Available at SSRN 4389233 (2023).

[3] Ji, Ziwei, et al. "Survey of hallucination in natural language generation." ACM Computing Surveys 55.12 (2023): 1-38.

[4] Touvron, Hugo, et al. "Llama: Open and efficient foundation language models." arXiv preprint arXiv:2302.13971 (2023).

[5] Guha, Neel, et al. "Legalbench: Prototyping a collaborative benchmark for legal reasoning." arXiv preprint arXiv:2209.06120 (2022).

[6] Fei, Zhiwei, et al. "LawBench: Benchmarking Legal Knowledge of Large Language Models." arXiv preprint arXiv:2309.16289 (2023).