




# XieZhi(獬豸): Chinese Law Large Language Model

Hongcheng Liu\*, Yuhao Wang\*, Yusheng Liao, Yutong Meng, Yu Wang   
Cooperative Medianet Innovation Center, Shanghai Jiao Tong University  
{hongcheng\_liu, colane, liao20160907, yuwangsJTU}@sjtu.edu.cn

## Abstract

Our primary vision is to ensure that everyone can access professional and reliable answers instantly when confronted with legal issues. Most large language models (LLMs) currently fall short in delivering proficient legal responses and authoritative legal references. In this paper, we introduce XieZhi, an LLM tailored for Chinese legal consultation. During the supervised fine-tuning stage, we enhanced its capabilities by incorporating distilled real legal consultation dialogues and situational Q&A from ChatGPT. Additionally, we refined the accuracy of legal references through iterative retrieval. Experimental results show that XieZhi excels at executing the legal consultation, significantly outperforming the original LLMs. This project can be available in <https://github.com/LiuHC0428/LAW-GPT>.

## 1 Introduction

Our vision is to make professional and reliable legal answers immediately accessible to everyone facing legal issues. We believe that professional legal services will only become habitual when they are as readily accessible as search engines were two decades ago and courier services a decade ago. Our goal is to integrate legal services into everyday life, thereby contributing to the development of a legally aware society. However, most large language models, like GPT-4, typically generate only general responses and lack the capability to provide specialized legal advice or corresponding Chinese legal references.

To realize this vision, we introduce XieZhi, a large language model designed for Chinese legal consultation. This model serves two primary functions: legal consultation and legal reference provision. For legal consultation, we have developed a new training dataset that includes 144k legal dialogues, refined by ChatGPT to offer professional advice. Additionally, we employ iterative retrieval methods to enhance the accuracy of legal reference searches, recognizing that generated responses are better suited for this purpose. We have assessed the model's performance in everyday scenarios, and the results affirm its effectiveness in providing legal consultation.

## 2 Methodology

### 2.1 Supervised Fine-tuning

In the first stage, we utilize both distilled and real-world data to equip our model for general legal consultations. We have developed a new law consultation dataset through ChatGPT distillation, focusing on two main areas: real-world legal dialogues and situational Q&A with legal references.

**Real-world legal dialogue** To achieve practical legal advice, it's crucial for the model to learn from real-world legal dialogues. We source our original dialogue data from the CrimeK-



Yu Wang is the corresponding author. The first two authors contributed to this work equally



Figure 1: The demonstration of the supervised fine-tuning. The two datasets are obtained from ChatGPT distillation.

gAssistant dataset<sup>1</sup>, which encompasses a variety of common legal issues, including scenarios such as consumer fraud and public disturbances. We refine this dataset by eliminating cases with unclear semantics or duplications and then employ ChatGPT to distill the remaining data for more nuanced responses.

**Situational Q&A** To bolster the model’s proficiency in providing legal references, we create situational Q&A using ChatGPT. Specifically, we generate questions and answers based on the core 9,000 legal provisions found in the legal manual of the People’s Republic of China. This method ensures that the dataset not only covers a wide range of situations but also incorporates specific legal references, significantly enhancing the model’s capability to deliver accurate legal advice.

## 2.2 Iterative Retrieval

After the supervised fine-tuning, the model is proficient in general legal consultation, albeit without providing legal references. Yet, the inclusion of legal bases is crucial in consultations and for safeguarding rights. Retrieval methods are effective in sourcing the correct laws from a knowledge base according to the query. Thus, we utilize retrieval techniques during the inference phase.

Specifically, we use the situational Q&A, which acts as a knowledge base containing the grounding legal statutes, to facilitate this process. We leverage the text2vec-base-chinese<sup>2</sup> model to encode legal statutes and user inquiries into sentence vectors. By calculating cosine similarity, we identify the most closely related legal provisions and input it into the model. However, understanding the legal provisions pertinent to a user’s query often requires inference. It has been observed that the semantic similarity between the original user queries and the legal statutes is not always high. Recognizing that LLMs excel at inference, we have adopted an iterative retrieval strategy to bolster the precision of our retrieval results. As depicted in Figure 2, our approach begins with using the original user query to elicit a response through the model. Subsequently, this generated response is utilized as a fresh query to fetch closely relevant legal provisions. The model generates the final response by

<sup>1</sup><https://github.com/liuhuanyong/CrimeKgAssistant>

<sup>2</sup><https://github.com/shibing624/text2vec>

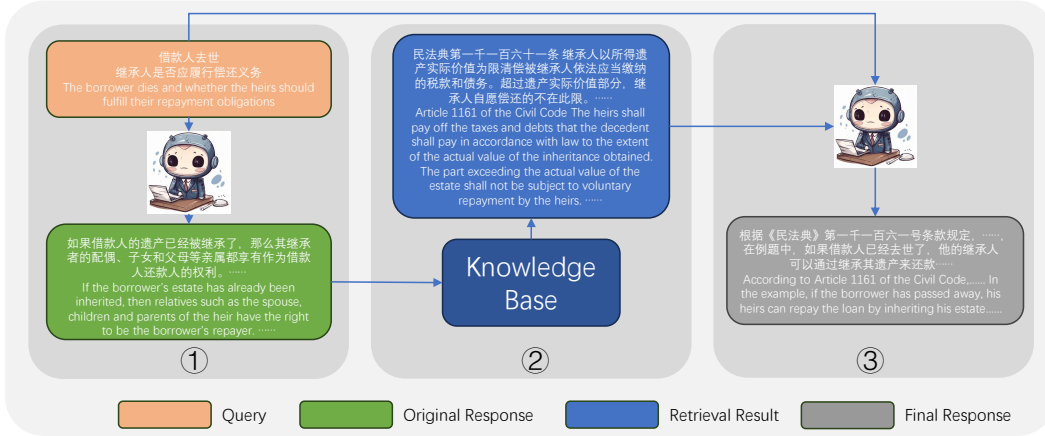


Figure 2: The demonstration of the inference stage. The first step aims at preliminary results generated. The second step aims at iterative retrieval. The third step aims at the final response generated.

synthesizing the retrieved legal terms with the original user query, ensuring that the legal guidance offered is both robust and precise.

### 3 Experiments

#### 3.1 Training Details

Our model is implemented in Pytorch using the Accelerate and Peft packages with Chatglm-6b as the base architecture. We employ the model parallelism technique to train the ChatGLM-6b using the LoRA method, with 16-bit precision, across a minimum of 4 NVIDIA GeForce RTX 3090 GPUs. In the supervised fine-tuning stage, we set the learning rate, training epoch, batch size, LoRA rank, and LoRA alpha as 5e-6, 1, 4, 8, and 16.

#### 3.2 Case Study

To demonstrate the capabilities of our law consultation model, we present two real-life examples<sup>3</sup>, as illustrated in Figure 3 and Figure 4. These examples highlight that XieZhi, equipped with retrieval capabilities, can generate the most accurate responses and provide the corresponding legal texts. Moreover, even without the retrieval feature, XieZhi still performs better than ChatGLM-6b, which validates the effectiveness of our Supervised Fine-Tuning (SFT) datasets.

### 4 Discussion

In this report, we introduce XieZhi, a Large Language Model designed for Chinese legal consultation. To enhance XieZhi's proficiency in legal advisory services, we have developed a new training dataset through ChatGPT distillation and incorporated iterative retrieval mechanisms during the inference stage.

However, the effectiveness of iterative retrieval is closely linked to the quality of the underlying knowledge base. Therefore, we are committed to continually developing a high-quality legal case knowledge base to support and improve future iterations of our law-oriented large language models.

<sup>3</sup>More cases can be found on <https://github.com/LiuHC0428/LAW-GPT>.

ChatGLM-6b	<p>问题:如何通过法律手段应对民间借贷纠纷?</p> <p>Q: How to deal with private lending disputes through legal means?</p> <p>民间借贷是一种常见的融资方式,但也存在风险。如果您遇到了民间借贷纠纷,以下是一些可以通过法律手段来应对的建议:</p> <ol style="list-style-type: none"> <li>1.收集证据:在处理任何纠纷之前,您需要收集尽可能多的证据来支持您的主张和证明对方存在过错。这些证据可能包括合同、借条、还款记录等文件。</li> <li>2.明确诉讼请求:根据您的实际情况和要求,明确自己的诉讼请求和诉讼费用的负担等内容。</li> <li>3.选择适当的律师:在面对复杂的民间借贷纠纷时,建议您选择一个专业的律师来帮助您进行处理。律师会根据案件的具体情况提供相应的建议和支持,并帮助您维护权益。</li> <li>4.提交起诉状:按照法律规定,您需要在指定的时间内向法院提交起诉状的副本和相关证据材料。</li> <li>5.等待判决结果:在法院受理后,会进行调解或者判决。如果双方无法达成一致意见,则法院将作出判决,并根据判决结果执行。</li> </ol> <p>需要注意的是,民间借贷的利率是有规定限制的,超过了规定的利率范畴,可能会被认定为非法高利借贷行为。因此,在使用民间借贷的过程中,需要注意利息的计算方式和标准,避免被不法分子利用。</p> <p>Private lending is a common way to raise funds, but there are also risks. If you're in a private lending dispute, here are some tips on how you can deal with it legally:</p> <ol style="list-style-type: none"> <li>1. Gather evidence: Before dealing with any dispute, you need to gather as much evidence as possible to support your claims and prove that the other party is at fault. This evidence may include documents such as contracts, IOUs, repayment records, etc.</li> <li>2. Clarify litigation claims: According to your actual situation and requirements, clarify your own litigation claims and the burden of litigation costs.</li> <li>3. Choose the right lawyer: When facing complex private lending disputes, it is recommended that you choose a professional lawyer to help you deal with them. A lawyer will provide advice and support on a case-by-case basis and help you defend your rights.</li> <li>4. Submit the complaint: According to the law, you need to submit a copy of the complaint and relevant evidentiary materials to the court within the specified time.</li> <li>5. Wait for the verdict: After the court accepts the case, mediation or judgment will be conducted. If the parties cannot reach an agreement, the court will make a judgment and enforce it based on the judgment.</li> </ol> <p>It should be noted that the interest rate of private lending is subject to prescribed restrictions, and if it exceeds the prescribed interest rate range, it may be regarded as illegal usurious lending. Therefore, in the process of using private lending, it is necessary to pay attention to the calculation method and standard of interest to avoid being used by criminals.</p>
	<p>遇到借贷纠纷,可以通过以下方式解决:</p> <ol style="list-style-type: none"> <li>1.协商:如果双方愿意协商的,可以就借款数额、还款期限和利息等进行谈判;如果无法达成一致意见或者对方拒绝履行协议,可以考虑起诉到法院要求强制执行;</li> <li>2.调解:如果双方同意进行调解的,可以向当地人民法院申请调解书来进行调节;需要注意的是调解不产生法律效力;</li> <li>3.仲裁:如果有证据证明存在借贷关系的,可以直接向当地的仲裁委员会申请仲裁;但是仲裁对裁决不服的话只能去法院诉讼解决了;</li> <li>4.法律援助中心寻求帮助。如果自己没有证据或判决不公的情况下,可以到当地的法律援助中心寻求援助;</li> <li>5.报警处理。如果对对方涉嫌刑事犯罪被逮捕了,可以先报案,由公安机关侦查后移送检察院审查起诉后再提起民事诉讼(民事赔偿部分);</li> </ol> <p>In the event of a loan dispute, it can be resolved in the following ways:</p> <ol style="list-style-type: none"> <li>1. Negotiation. If both parties are willing to negotiate, they can negotiate the amount of the loan, the repayment period and the interest; If an agreement cannot be reached or the other party refuses to perform the agreement, it may consider suing the court for enforcement;</li> <li>2. Mediation. If both parties agree to mediation, they can apply to the local people's court for mediation; It is important to note that mediation does not have legal effect;</li> <li>3. Arbitration. If there is evidence to prove the existence of a loan relationship, the applicant may directly apply to the local arbitration commission for arbitration; However, if the arbitrator is not satisfied with the award, he can only go to the court to settle the dispute;</li> <li>4. Ask for help from the Legal Aid Center. If you have no evidence or the verdict is unfair, you can go to your local legal aid center for assistance;</li> <li>5. Alarm handling. If the other party is arrested on suspicion of a criminal offense, the case may be reported first, and the public security organ will investigate and transfer it to the procuratorate for review and prosecution before filing a civil lawsuit (civil compensation part);</li> </ol> <p>根据《民法典》第468条的规定,对于非因合同产生的债权债务关系,适用于有关该债权债务关系的法律规定,而没有特定的规定时则适用于本编通则的有关规定,但是根据其性质不能适用的除外。因此,对于民间借贷纠纷,应该依据《民法典》第457条的规定,即借款人应当按照约定的利率支付利息;未按照的约定支付利息的,借款人可以随时请求借款人偿还本金并赔偿损失。此外,如果借款人到期不还款或者违反与出借人之间的借款协议中关于还款金额、还款日期等内容的,出借人可以通过向法院提起诉讼来维护自己的权益。在起诉前,出借人应当采取必要的证据措施,如收集借款人的身份信息、聊天记录、转账记录等信息以证明双方之间的关系和借款事实的真实性。同时,出借人也需要考虑自己的合法权益是否受到侵害,并根据实际情况决定是否委托律师进行维权。</p> <p>According to Article 1161 of the Civil Code, the heir shall pay off the taxes and debts payable by the decedent in accordance with the law to the extent of the actual value of the inheritance. The part exceeding the actual value of the estate shall not be subject to voluntary repayment by the heirs. Therefore, if the actual value of the borrower's estate exceeds the amount of debt that he should pay, then the heirs are not obligated to repay the additional debt. However, if the heirs voluntarily repay, it is not within this limit.</p> <p>In the example, if the borrower has passed away, his heirs can repay the loan by inheriting his estate. If the heir inherits all or most of the property, then he is obliged to pay all outstanding debts. He may not be able to take on additional debts if he does not inherit enough property, but that doesn't mean he shouldn't do his best to pay off his debts.</p>

Figure 3: The case study of the XieZhi with and without retrieval in law consultation.

	<p>问题:宠物狗被撞死，可以索要精神赔偿吗？</p> <p>Q: If a pet dog is hit and killed, can I claim moral compensation?</p>
ChatGLM-6b	<p>如果您的宠物狗因为交通事故而死亡，您可能有权要求对方支付精神赔偿。根据中国的《消费者权益保护法》和相关法律法规的规定，消费者因购买、使用商品或者接受服务受到人身、财产损害的，享有生命权、健康权和财产安全权等权益。</p> <p>因此，如果您的宠物狗因为交通事故而遭受精神和情感上的痛苦，您有权要求对方为您的狗提供经济上的帮助，以缓解您的痛苦。不过需要注意的是，这种精神赔偿通常是由法院根据具体情况判决的，因此在主张精神赔偿前，建议您先咨询专业的律师或法律援助机构，了解相关权利和法律规定。</p> <p>If your pet dog dies as a result of a traffic accident, you may be entitled to claim moral compensation. According to the Law of the People's Republic of China on the Protection of Rights and Interests of Consumers and relevant laws and regulations, consumers who suffer personal or property damage due to the purchase or use of goods or services shall enjoy the rights and interests of life, health and property safety.</p> <p>Therefore, if your pet dog is suffering mentally and emotionally as a result of a traffic accident, you have the right to ask the other person to provide financial assistance to your dog to ease your pain. However, it should be noted that this kind of moral compensation is usually awarded by the court on a case-by-case basis, so it is recommended that you consult a professional lawyer or legal aid agency to understand the relevant rights and legal provisions before claiming moral compensation.</p>
XieZhi	<p>针对你的问题，我进行分析如下：撞人的狗狗主人是否要承担责任取决于双方之间的责任划分。如果确定对方有过错并且造成了小狗的死亡，那么根据法律规定，有权要求其进行相应赔偿。例如经济赔偿金、精神损害抚慰金等。但是，具体的赔偿数额需要根据实际情况来计算。此外，也需要注意是否存在其他相关的证据支持索赔的请求。同时，对于动物的饲养人或管理人而言，如果没有尽到相应的管理义务导致动物受到伤害或者死亡的，也需要承担相应的赔偿责任。</p> <p>In response to your question, I analyze as follows: Whether the owner of the dog who hit the person is liable depends on the division of responsibility between the two parties. If it is determined that the other party is at fault and caused the death of the puppy, then according to the law, you have the right to claim compensation from the other party. For example, financial compensation, solatium for moral damages, etc. However, the specific amount of compensation needs to be calculated on a case-by-case basis. In addition, it is also important to note whether there is other relevant evidence to support the claim. At the same time, for the keeper or manager of the animal, if the animal is injured or killed if the corresponding management obligation is not fulfilled, it also needs to bear the corresponding liability for compensation.</p>
XieZhi with Retrieval	<p>根据《民法典》第1250条规定，因第三人的过错致使动物造成他人损害的，被侵权人可以向动物饲养人或者管理人请求赔偿，也可以向第三人请求赔偿。如果动物饲养人或管理人已经赔偿了被侵权人的经济损失，被侵权人还可以通过向第三人追偿来寻求进一步的经济赔偿。</p> <p>然而，对于宠物狗这种具有人格属性的动物，根据《民法典》第1247条规定，禁止饲养的烈性犬等危险动物造成他人损害的，动物饲养人或者管理人应当承担侵权责任。这意味着如果宠物狗被撞死了，受害人不能直接向动物饲养人或管理人要求经济赔偿，而是需要先通过诉讼程序主张自己的权利，认定是由第三人的行为导致的损失，然后由第三人承担相应的赔偿责任。</p> <p>According to Article 1250 of the Civil Code, if an animal causes damage to others due to the fault of a third party, the infringed party may claim compensation from the animal keeper or manager, or from a third party. If the animal keeper or manager has compensated the infringed party for economic losses, the infringed party can also seek further financial compensation by recovering from a third party.</p> <p>However, for pet dogs, which are animals with personality attributes, according to Article 1247 of the Civil Code, if a dangerous animal such as a fierce dog that is prohibited from being raised causes damage to others, the animal keeper or manager shall bear tort liability. This means that if the pet dog is hit and killed, the victim cannot directly claim financial compensation from the animal keeper or manager, but needs to first assert his rights through the litigation process, determine that the loss is caused by the act of the third party, and then the third party will bear the corresponding liability for compensation.</p>

Figure 4: The case study of the XieZhi with and without retrieval in law consultation.