

复旦大学计算机科学技术学院  
2023-2024 学年第一学期期末论文课程评分表

课程名称： 计算思维

课程代码： COMP119004.01

开课院系： 计算机科学技术学院

学生姓名： 王鲲溟 学号： 22307100011 专业： 经济管理试验班

论文名称： 生成式人工智能在法律咨询服务行业应用研究

(以上由学生填写)

成绩： 90

论文评语（教师填写）：

论文探究了生成式语言模型在法律咨询方面应用的可能性。通过数据收集、分析、实验设计，探索了人工智能生成文本和专业律师的建议之间的异同，对人工智能生成文本进行了较为细致的分析，提出模型虽然在关键词提取等方面表现较好，但和专业律师相比，存在回答不全面、忽视情境细节、案例特殊性等问题。文章最后总结了目前模型的局限性和缺陷，并提出了可能的改进意见。选题较好，结合前沿研究热点，也体现出了自己的实践和思考。符合复旦大学优秀通识核心课程论文的标准，推荐参与校优秀论文的评选。

任课教师签名：

日 期：

## 生成式人工智能在法律咨询服务行业应用研究

**【摘要】**随着人工智能的迅速发展，生成式人工智能在多个行业中的应用日益广泛，尤其在法律咨询服务行业中展现出巨大的潜力。本研究旨在探索生成式人工智能，尤其是 **GPT** 等模型在法律咨询服务中的应用效果。通过对政治、经济、社会和技术环境的分析，提出了**运用生成式人工智能优化法律服务行业的可能性**。研究方法包括实验设计、数据收集和数据分析，主要关注于模型与律师在法律咨询服务方面的性能比较。实验结果显示，虽然生成式人工智能在关键词提取、相似度计算等方面表现出色，但与专业律师相比，仍存在一定差距。对生成式人工智能在法律咨询服务行业的应用前景提供了深入的见解，并为未来相关研究提出了建议。

**【关键词】**：生成式人工智能；法律咨询服务；GPT 模型。

# 目 录

1 绪论 .....	6
1.1 背景介绍 .....	6
1.2 法律咨询服务行业现状 .....	6
1.3 论文目的 .....	8
2 文献综述 .....	8
2.1 生成式人工智能基本原理 .....	8
2.2 关键词提取算法 .....	9
2.3 分词算法 .....	10
2.4 词向量化方法 .....	11
2.5 生成式人工智能的行业应用 .....	11
2.6 法律咨询服务行业的现状和发展趋势 .....	12
3 研究方法 .....	12
3.1 实验设计 .....	12
3.1.1 研究目的和假设 .....	12
3.1.2 实验流程 .....	12
3.1.3 实验指标 .....	12
3.2 数据收集 .....	13
3.2.1 数据类型 .....	13
3.2.2 数据收集方法 .....	13
3.3 数据分析 .....	13
3.3.1 数据预处理 .....	13
3.3.2 数据分析 .....	13
4 研究结果 .....	14
4.1 实验结果 .....	14
4.2 结果分析 .....	21
5 讨论 .....	24
5.1 生成式人工智能的行业价值 .....	24

5.2 生成式人工智能面临的挑战和发展方向 .....	24
5.3 对法律咨询服务行业的影响和启示 .....	25
6 结论和建议 .....	25
6.1 发现和结论 .....	25
6.2 对未来研究的建议 .....	25
7 参考文献 .....	25
参考文献 .....	25
附 录 .....	27

# 1 绪论

## 1.1 背景介绍

随着科技的迅速发展，生成式人工智能已经为多个行业带来了革命性的变革，并且在服务行业也显示出巨大潜力。法律咨询服务行业特点突出，如信息处理量大、案例分析复杂、法律解释需准确等。法律咨询服务在全球范围内需求旺盛，但受限于律师服务价格昂贵及对隐私信息的担忧，普通民众在面临法律问题往往无法获得及时、有效的解答。本研究旨在利用人工智能技术，结合 GPT 系统进行法律咨询服务的研究，以降低成本、提高咨询效率，并确保用户隐私。

## 1.2 法律咨询服务行业现状

### 一、政治环境

长期以来，我国一直在加快推进法律服务业的发展，为本行业的发展营造了良好的环境。根据司法部《全国公共法律服务体系建设规划（2021-2025年）》，其中提到：“大力发展公共法律服务领域科技创新支撑技术，推进智慧法律服务重点实验室建设，重点突破精准普法宣传、法律援助智能保障……提高法律服务智能化水平。”<sup>1</sup>这为法律服务行业的科技赋能提供了坚实的政治基础。

如今，将科技创新与社会、经济建设相结合已经成为政治生活的重要议题。在党的二十大报告中，习近平总书记指出：推动战略性新兴产业融合集群发展，构建新一代信息技术、人工智能等一批新的增长引擎，构建优质高效的服务业新体系。随着大众创业，万众创新趋势的演进，在法律服务方面的科技创新也必将成为新时代的亮点。<sup>2</sup>这也为法律服务行业的科技创新提供了动力。

### 二、经济环境

表 1.1 2018-2022 年全国律师人数和办案数据

年份	律师人数（万）	增速（%）	律师办理各类法律事务（万）	增速（%）
2022	65.16	13.36	1274.4	-2.61

<sup>1</sup> 中华人民共和国司法部.全国公共法律服务体系建设规划（2021-2025年）[EB/OL].<https://www.gov.cn/zhengce/zhengceku/2022-01/25/5670385/files/cca208b72837425992c31b25e548e6ad.doc>,2021:14-15.

<sup>2</sup> 中华人民共和国中央人民政府.习近平：高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告[EB/OL].[https://www.gov.cn/xinwen/2022-10/25/content\\_5721685.htm](https://www.gov.cn/xinwen/2022-10/25/content_5721685.htm),2022-10-25.

2021	57.48	10.11	1308.5	17.41
2020	52.2	10.36	1114.5	-0.4
2019	47.3	11.82	1119	4.78
2018	42.3	15.89	1068	/

数据来源：司法部

根据司法部 2018-2022 年度律师、基层法律服务工作统计分析<sup>12345</sup>，最近五年我国律师人数保持增长，每年增速均超过 10%。但是律师人数增长时，律师办理各类法律事务数量的增长速度却相对较慢，2022 年还出现了负增长。总体上看，律师办理案件的增速远低于律师人数的增速，法律服务市场需要新的增长点。

**表 1.2 2018-2022 年律师办理诉讼案件代理情况**

年份	律师办理诉讼案件 量（万）	全国法院收案 量（万）	诉讼案件律师代理 率（%）	增速 （%）
2022	824.4	3372.3	24.45	0.95
2021	811.6	3351.6	24.22	12.91
2020	660.9	3308.5	21.45	5.35
2019	610.8	约 3000	20.36	2.47
2018	497.8	2505	19.87	/

数据来源：司法部

根据数据，最近五年诉讼案件律师代理率保持增长，说明律师服务整体市场仍在快速增长，并且增速仍有加快的趋势。

综合来看，律师办理案件增速有所增加，但是律师人数增速远高于办案增速。由此可见虽然法律服务市场整体正在快速增长，但律师行业缺乏开拓市场

<sup>1</sup> 中华人民共和国司法部.2022 年度律师、基层法律服务工作统计分析 [EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202306/t20230614\\_480740.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202306/t20230614_480740.html),2023-06-14.

<sup>2</sup> 中华人民共和国司法部.2021 年度律师、基层法律服务工作统计分析 [EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202208/t20220815\\_461680.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202208/t20220815_461680.html),2022-08-15.

<sup>3</sup> 中华人民共和国司法部.2020 年度律师、基层法律服务工作统计分析 [EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202106/t20210611\\_427394.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202106/t20210611_427394.html),2021-06-11.

<sup>4</sup> 中华人民共和国司法部.2020 年度律师、基层法律服务工作统计分析 [EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202006/t20200622\\_350049.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202006/t20200622_350049.html),2020-06-22.

<sup>5</sup> 中华人民共和国司法部.2020 年度律师、基层法律服务工作统计分析 [EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/201903/t20190307\\_350047.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/201903/t20190307_350047.html),2019-03-07.

的方法。

### 三、社会环境

表 1.3 2018-2022 年全国律师民事诉讼业务占比情况

年份	民事诉讼代理（万）	占比
2022	697.5	84.61
2021	660.1	81.35
2020	532.7	80.60
2019	479.2	78.45
2018	396.9	79.72

数据来源：司法部

根据数据，最近五年律师办理民事诉讼案件最多，近三年占比均超过 80%，且占比持续提高，这说明目前民事诉讼方面拥有巨大的发展潜力，这也为法律咨询服务的发展提供了契机。

### 四、技术环境

人工智能的迅速发展将深刻改变人类社会生活、改变世界。近年来，中国人工智能产业发展取得了重大进展，一些领域的技术创新能力位居世界前列。中国大力发展以人工智能为代表的数字技术，此举将为中国经济增长注入新动力，加快企业数字化和智能化升级。

#### 1.3 论文目的

根据《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》，我国人工智能技术快速发展、数据和算力资源日益丰富、应用场景不断拓展，为开展人工智能场景创新奠定了坚实基础。但仍存在对场景创新认识不到位，重大场景系统设计不足，场景机会开放程度不够，场景创新生态不完善等问题，需要加强对人工智能场景创新工作的统筹指导<sup>1</sup>，因此本研究尝试结合人工智能技术整合资源能力强、信息处理效率高、客观程度高的优势，将其应用到法律服务领域。

## 2 文献综述

### 2.1 生成式人工智能基本原理

<sup>1</sup> 科技部等六部门.关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见 [EB/OL].[https://www.gov.cn/zhengce/zhengceku/2022-08/12/content\\_5705154.htm](https://www.gov.cn/zhengce/zhengceku/2022-08/12/content_5705154.htm),2022-07-29



随着计算速度提升、大数据的涌现和算法改进，“深度学习”在自然语言处理方面取得了优秀表现，而生成式人工智能，即大型语言模型（LLM）是基于大量数据进行预训练的超大型深度学习模型。底层转换器是一组神经网络，这些神经网络由具有自注意力功能的编码器和解码器组成。编码器和解码器从一系列文本中提取含义，并理解其中的单词和短语之间的关系。本节以 GPT（Generative Pre-trained Transformer）<sup>1</sup> 的原理为例进行简要介绍。

### 一、输入编码

GPT 需要对输入的文本进行编码，除了对词语本身的信息进行编码外，还加入了位置编码和段编码。因为 GPT 的注意力机制没有顺序概念，所以需要位置信息也加入对输入的编码中。即字符编码与位置编码运算得到最终编码。

### 二、Transformer Decoder

GPT 抛弃了常见的循环结构 RNN，而是使用了 Transformer 的解码器部分，即遮蔽自注意层和全连接层。遮蔽自注意层是将每一个输入字符向量通过线性变换得到 Q(query)K(key)V(value)三个向量。其中 V 是这个字符实际的向量表示，要与对应的权重相乘。K 相当于这个字符的标签，要与当前字符的 Q 向量点乘。而 Q 用于与其它字符点乘，用于获得该字符与上下文的相关性。全连接层则是一个将每个神经元与前一层的所有神经元都连接起来的神经网络层，全连接层能够接收来自前一层的所有特征，并将它们线性组合以生成新的特征表示。实践证明，Transformer 的性能优于 RNN。

### 三、输出

GPT 最终输出每个词对应的融合全文语义信息的向量表示。

同时，大模型在使用前需要进行预训练以达到使用的目的。GPT 使用的预训练方式为 LM（Language Model），即根据文本预测下一个词。语料来源是互联网上有质量的文本。除了 LM 外还有 BI-LM，MLM，NSP 等训练方式。

## 2.2 关键词提取算法

在实验过程中使用的关键词提取算法为 KeyBERT 算法<sup>2</sup>。KeyBERT 的优势在于通过利用 BERT 的强大语义表示能力，能够更准确地捕捉文本中的语义信息，从而提高关键词提取的质量。也因如此，在无监督关键词提取工具

---

<sup>1</sup> Budzianowski P, Vulić I. Hello, it's gpt-2-how can i help you? towards the use of pretrained language models for task-oriented dialogue systems. [Z] Proceedings of the 3rd Workshop on Neural Generation and Translation, 2019.

<sup>2</sup> MaartenGr. KeyBERT. [Z] <https://github.com/MaartenGr/KeyBERT>, 2023-10-5

KeyBERT 表现十分优秀。现简单介绍 KeyBERT 算法如下：

一、文本嵌入：首先，KeyBERT 使用预训练好的 BERT（Bidirectional Encoder Representations from Transformers）模型对输入的文本进行嵌入，得到文档的向量表示。其中，BERT 也是一种大型语言模型，与 GPT 的区别在于 BERT 使用的是 Transformer 的编码器部分，并且 BERT 使用 MLM+NSP 的训练方式。MLM（Masked Language Model），即遮蔽语言模型，其训练方式为遮挡任意词，让机器反推回遮挡的词是什么。NSP（Next Sentence Prediction）的训练方式为判断一个句子是否是另一个句子的下一个句子。

二、关键词候选生成：对于每个文档，选择一组关键词的候选集。这些候选词由文档中的词汇或短语分词得到。

三、余弦相似度计算：对于每个关键词的候选，计算其与文档向量之间的余弦相似度。余弦相似度度量了两个向量在多维空间中的夹角，从而表示它们之间的语义相似性。

四、关键词排序：根据计算得到的余弦相似度，对关键词的候选集进行排序，选择与文档向量最相似的候选词作为最终的关键词。注意到有些关键词可能并没有太多实际意义，因此在这一步需要人工剔除无效的关键词。

### 2.3 分词算法

由于 KeyBERT 在关键词提取时无法做到对中文的分词，因此在使用 KeyBERT 进行关键词提取之前要首先进行分词。本研究使用 Python 的第三方库 jieba<sup>1</sup>进行分词，现简要介绍 jieba 分词算法如下：

一、生成前缀词典：将在统计词典中出现的每一个词的每一个前缀提取出来，统计词频，如果某个前缀词在统计词典中没有出现，词频统计为 0，如果这个前缀词已经统计过，则不再重复。

二、分割子句：列举所有在中文词典中可能会出现的字符，将连续出现的这些字符作为一个子句进行划分，这些字符之外的其他符号，我们便可以认为是中文标点符号，并把他们作为子句划分标志。

三、构造 DAG：将有向图存储为字典形式。

四、动态规划找到最大路径：每一个词出现的概率等于该词在前缀里的词频除以所有词的词频之和。如果词频为 0 或是不存在，当做词频为 1 来处理。

---

<sup>1</sup> fxsjy.jieba.[Z]<https://pypi.org/project/jieba/>,2020-1-20

最后使用动态规划算法算出概率最大的路径。

五、HMM 算法对于未登录词的识别：HMM 是一个生成式模型。用 HMM 模型进行分词，主要是将分词问题视为一个序列标注（sequence labeling）问题，其中，句子为观测序列，分词结果为状态序列。首先通过语料训练出 HMM 相关的模型，然后利用 Viterbi 算法进行求解，最终得到最优的状态序列，然后再根据状态序列，输出分词结果。

六、正则表达式辅助分词：为了处理日期、数字、英文单词或者其他符合某个特定规则的词语，需要使用正则表达式把连续的、不在词典中的非汉字字符提取出来。

## 2.4 词向量化方法

由于最终得到的关键词几乎不同，但在意义上可能相近，因此需要将最终得到的关键词再次进行向量化，以用于进行比较。在这里使用的算法是 Word2Vec 算法，现对算法简单介绍如下：

Word2Vec 是一种根据词生成向量的算法，为了发现词之间的相关性，通常需要足够多的维度向量<sup>1</sup>才能达到。其通常需要进行训练，主流训练方法有 CBOW 和 Skip-gram 两种。CBOW(Continuous Bag-of-Words Model)，即词袋模型，是通过上下文来预测当前值。即遮住一句话中的一个词，让机器推断被遮住的词是什么。而 Skip-gram (Continuous Skip-gram Model)则相反，是用当前词来预测上下文。即给出一个词，让机器推断整句话是什么。

本实验将使用 Word2Vec 的预训练模型 Chinese-Word-Vectors<sup>23</sup>，并且使用的是经由知乎网站爬取的数据进行的预训练模型，能够提供中文词对应的词向量的同时提供更符合网络问答的词向量结果。

## 2.5 生成式人工智能的行业应用

生成式人工智能涵盖了各种任务领域，包括图像生成（文本转图像）、文本生成（文本到文本）、语音生成（文本转语音）、音乐生成（文本到音乐）、代码生成（文本到代码）、视频生成（文本到视频）和科学内容生成（文本到

---

<sup>1</sup> Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space.[Z]In Proceedings of Workshop at ICLR,2013.

<sup>2</sup> Shen Li.Chinese-Word-Vectors.[Z]<https://github.com/Embedding/Chinese-Word-Vectors/>,2023-10-30

<sup>3</sup> Shen Li, Zhe Zhao, Renfen Hu, Wensi Li, Tao Liu, Xiaoyong Du, Analogical Reasoning on Chinese Morphological and Semantic Relations[C]ACL,2018.

科学），而在文本生成方面的应用尤为显著。目前的大语言模型已经能够做到文本翻译、搜索、提供情感陪伴等多种功能。

## 2.6 法律咨询服务行业的现状和发展趋势

**表 2.1 2019-2021 中国法律服务行业市场规模**

年份	市场规模（万元）
2019	842
2020	1178
2021	1503

数据来源：前瞻产业研究院

根据前瞻产业研究院的数据<sup>1</sup>，2019-2021 年，中国法律服务行业市场规模逐年上升。中国法律服务行业在 2020 年总体市场规模已经达到 1178 亿，2021 年超过 1500 亿元。目前法律服务市场规模逐渐扩张，预计将在未来几年稳步增长，并且伴随“互联网+”的浪潮，在法律服务领域人工智能等技术的应用将会逐步落地和完善。

## 3 研究方法

### 3.1 实验设计

#### 3.1.1 研究目的和假设

本研究旨在探索生成式人工智能在法律咨询服务行业的应用效果。研究假设为：生成式人工智能在法律咨询服务行业拥有着和律师相同的专业水平。

#### 3.1.2 实验流程

一、数据收集：在社交平台收集法律咨询服务的数据并加以整理。

二、实验执行：使用生成式人工智能进行相同的法律咨询服务，并记录得到的结果。

三、量化结果：将律师的法律咨询服务结果和实验结果进行量化，从而定量的统计人工智能生成结果的质量。

#### 3.1.3 实验指标

一、广度指标：计算人工智能生成结果对律师的相对广度，这里广度

<sup>1</sup> 前瞻产业研究院.2023-2028 年中国法律服务行业市场需求与投资战略规划分析报告 [EB/OL].<https://bg.qianzhan.com/report/detail/c8aebb3ce98e4dd3.html>

指提供信息的丰富程度。

二、深度指标：计算人工智能生成结果对律师的相对深度，这里深度指提供信息的准确程度。

## 3.2 数据收集

### 3.2.1 数据类型

本实验收集的数据均为文本形式，其中分为两部分：一部分为实验部分，即用于生成式人工智能对话生成结果的部分；另一部分为对照部分，即律师作出的回答文本。

### 3.2.2 数据收集方法

本实验的数据收集方法有两种：一是通过 Python 爬虫<sup>1</sup>爬取知乎等社交媒体平台上相应的板块，得到相应的数据；二是人工收集小红书等社交 app 上的相应内容，并进行整理和分类。

## 3.3 数据分析

### 3.3.1 数据预处理

一、数据清洗：对收集得到的数据进行人工清洗，将非专业律师回答、无效回答、无效提问等数据进行清洗，从而得到用于进行实验的数据库。

二、数据整合：将两种收集方式得到的数据进行整合，以便于进行统一的分析。在这之后，对人工智能进行相同的提问，收集人工智能的回答结果。在由于数据量较大，本实验会使用调用 API 得到人工智能回答的方式对问题进行批量处理。

三、数据量化：将律师的回答结果和人工智能的回答结果首先使用 jieba 第三方库将文本进行分词，再使用 KeyBert 关键词提取算法进行关键词提取。

### 3.3.2 数据分析

本实验使用以下统计方法确定不同回答之间的相似度：

余弦相似度：即通过计算两个向量的夹角余弦值来评估他们的相似度。在本实验中关键词会被转换成向量以进行计算。注意到不同回答得到的关键词向量可能会存在不同的维度，因此在比较之前需要将所有不同的维度合并成一个统一的维度集合，并为每个向量在缺失的维度上填充 0。

---

<sup>1</sup> 马哥 python 说.【2023 知乎爬虫】我用 Python 爬虫爬了 2300 多条知乎评论.[Z]<https://zhuannlan.zhihu.com/p/556843537.2023-08-04>

## 4 研究结果

### 4.1 实验结果

一、数据收集结果：经过数据收集后共得到知乎问答数 47 个，小红书问答数 20 个。经人工筛选后知乎有效问答数剩余 30 个。

图 4.1 知乎法律咨询问答示例



图 4.2 小红书法律咨询问答示例



再重逢



1

你好 请问短期挂名公司法定代表人，不参与任何公司的运作，后期变更了法定代表人公司除了任何问题，之前的法定代表人会承担连带责任吗。 6天前 四川 回复

作者赞过



律师|王铁平 作者



公司法定代表人变更登记后出的问题，与原法定代表人履职行为没有因果关系的，以变更之日起，原法人对以前签发的公司文件负责，现法人对公司以后的负责。

5天前 江西 回复

二、使用人工智能生成结果：对于每个问题使用 GPT 生成相应的结果。

图 4.3 文心一言生成结果示例

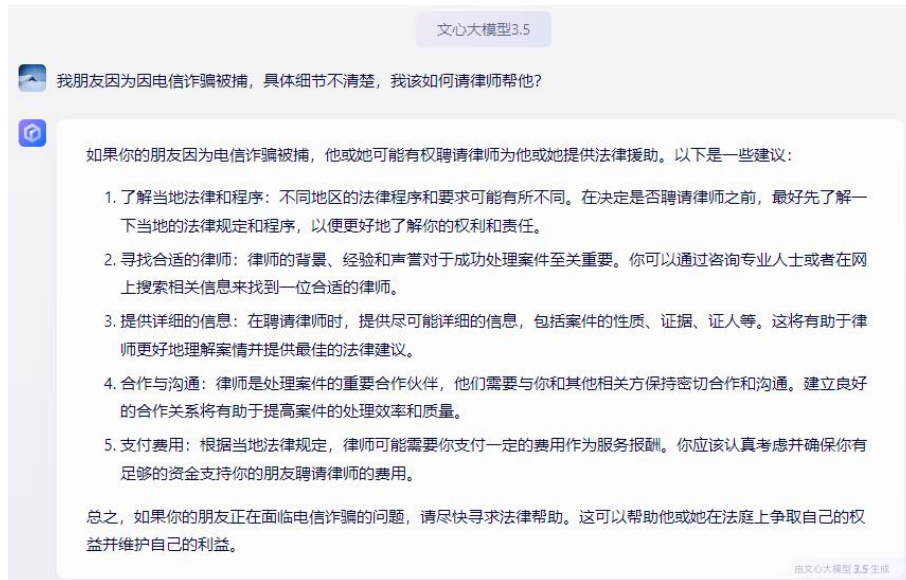


图 4.4 moonshot 生成结果示例

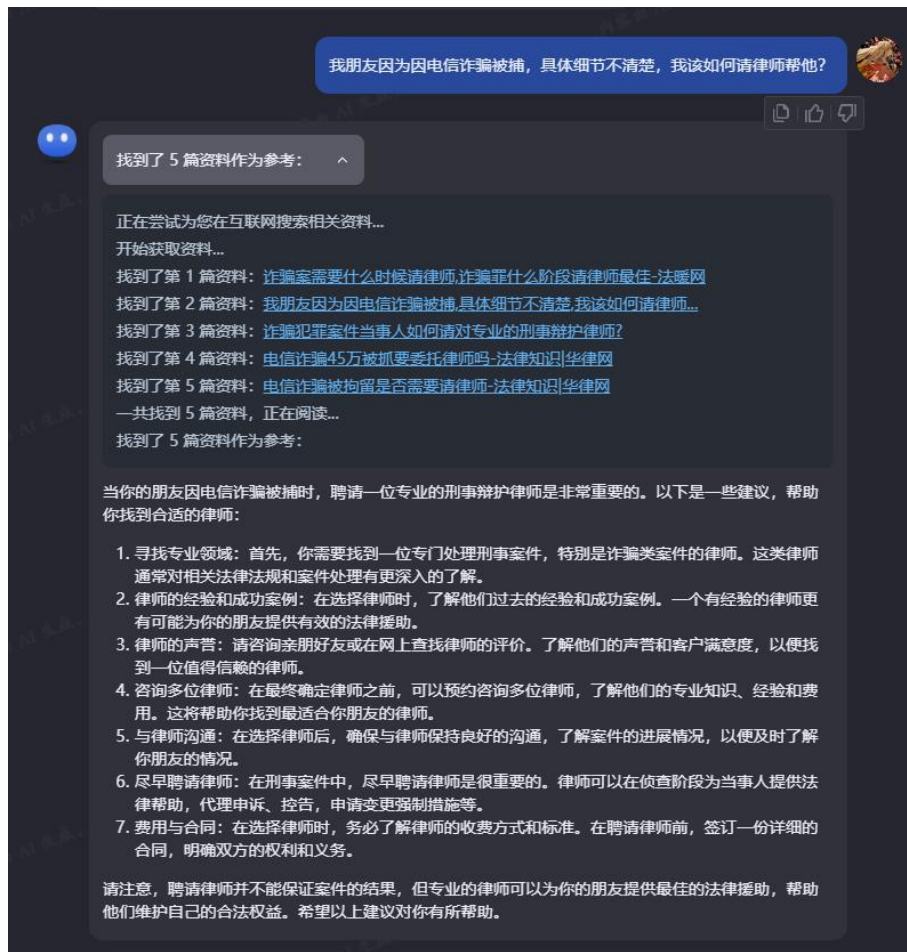
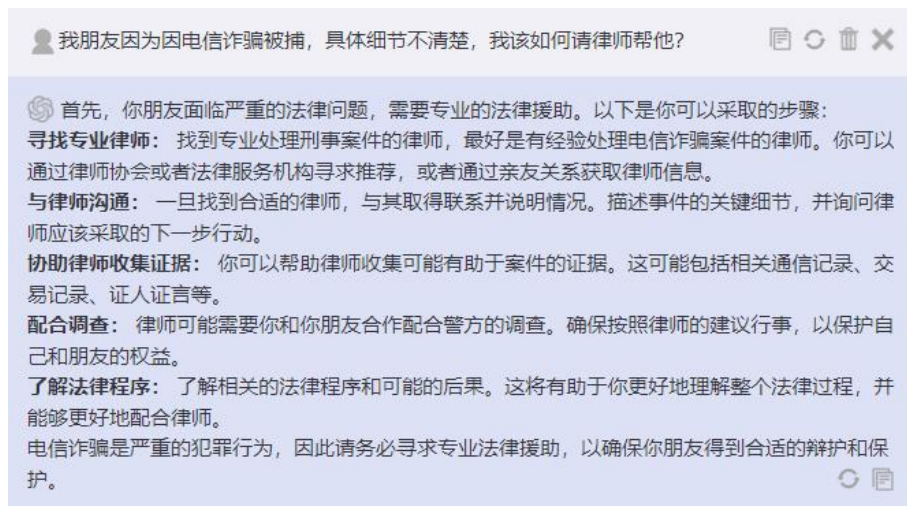


图 4.5 ChatGPT 生成结果示例



三、文本分词：将人工智能的生成结果和律师回答的文本结果进行分词处理，并且最终仅保留分词得到的文本形式。

四、数据量化：将人工智能的生成结果和律师回答结果进行关键词量化。并且仅提取出来每个回答的关键词文本，将关键词中无关紧要的词语删去。

图 4.6 生成结果描述性统计分析图



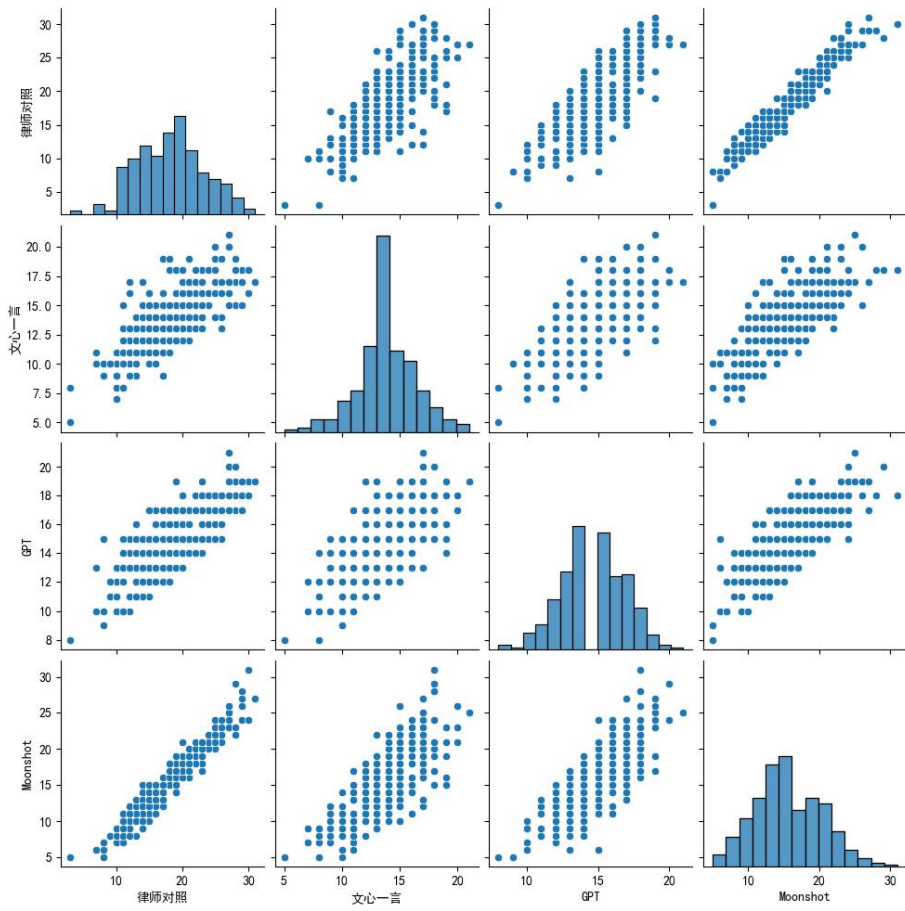


图 4.7 生成结果相关性分析图

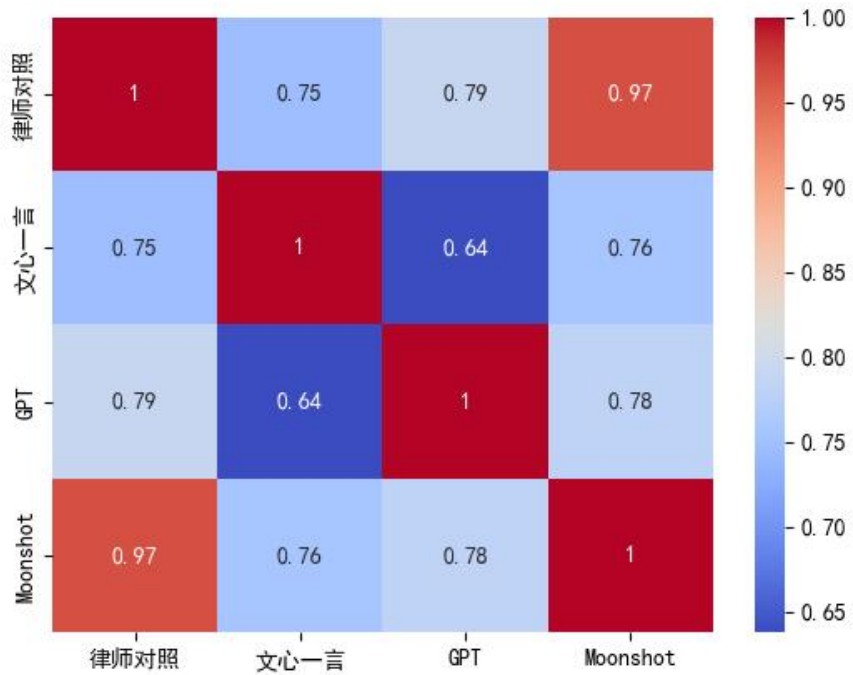


图 4.8 生成结果箱线图

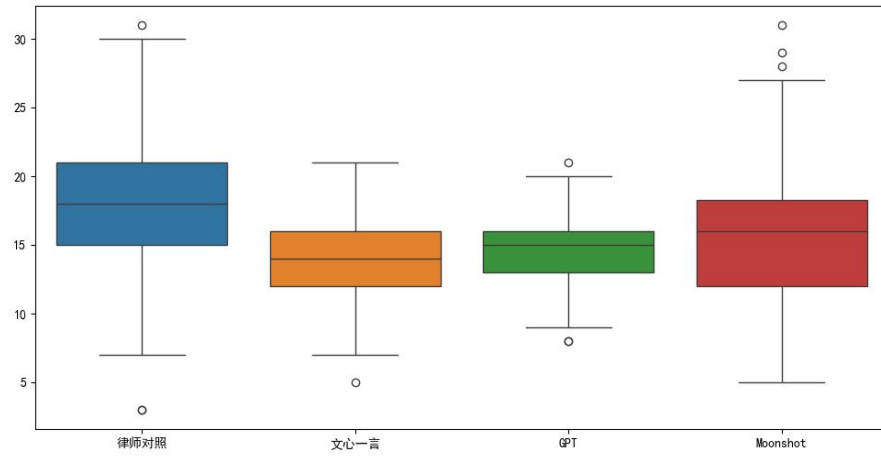


图 4.9 生成结果直方图

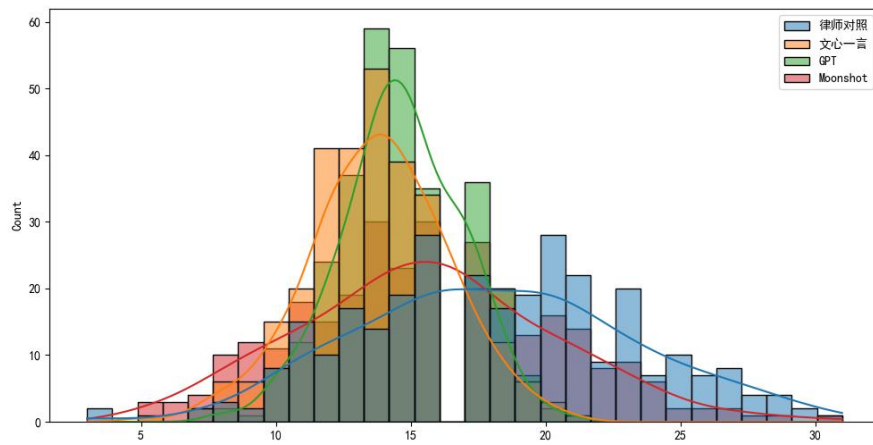


图 4.10 生成结果去除无效关键词后描述性统计分析图

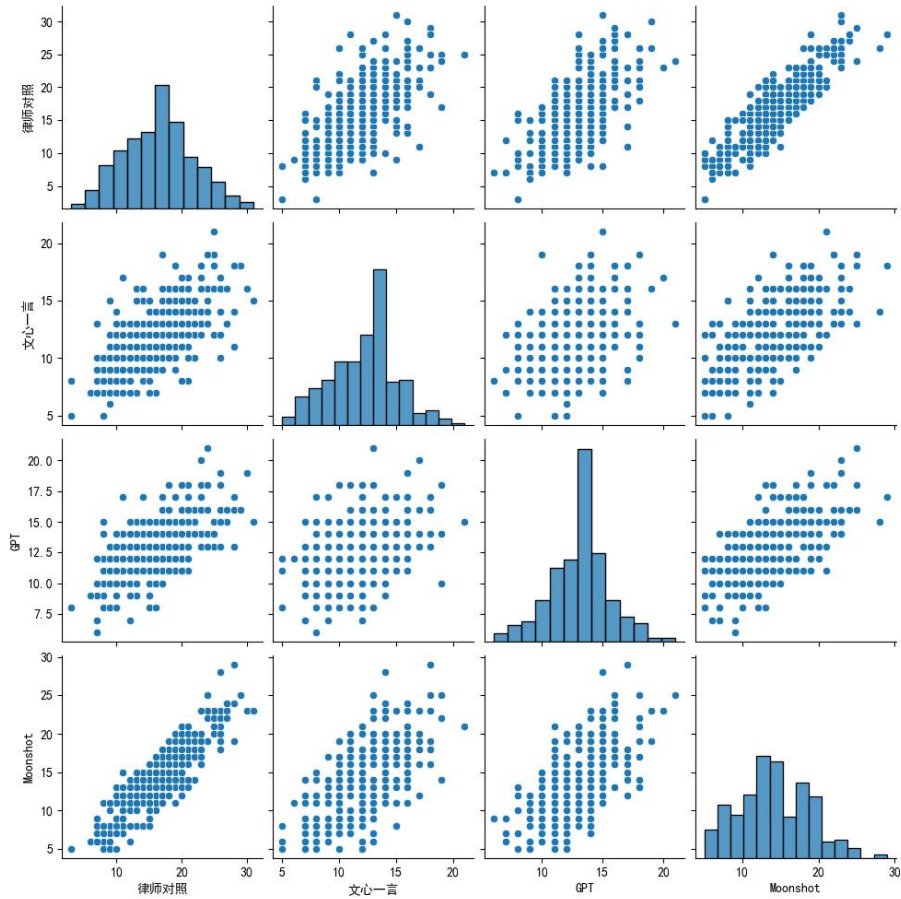


图 4.11 生成结果去除无效关键词后相关性分析图

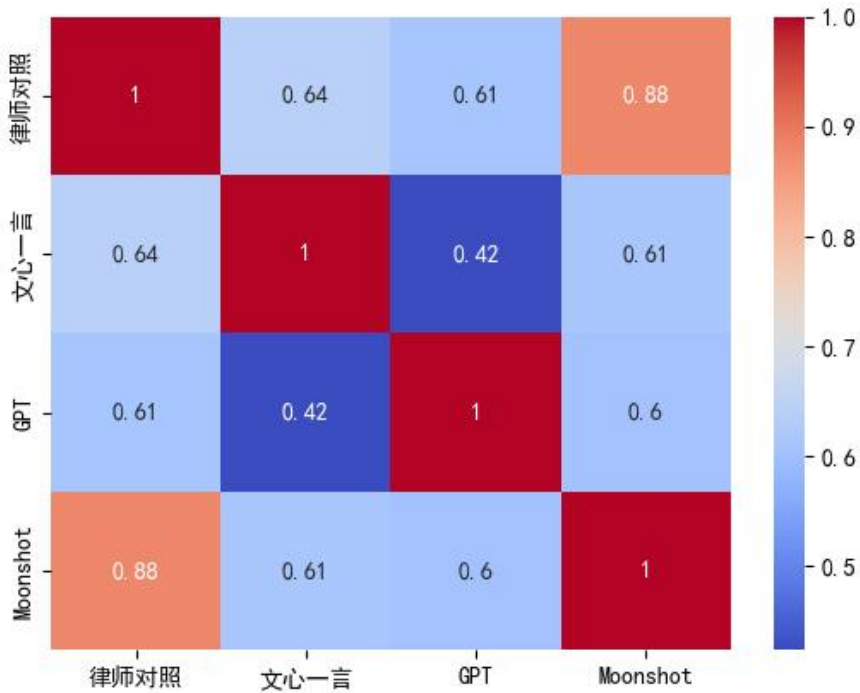


图 4.12 生成结果去除无效关键词后箱线图

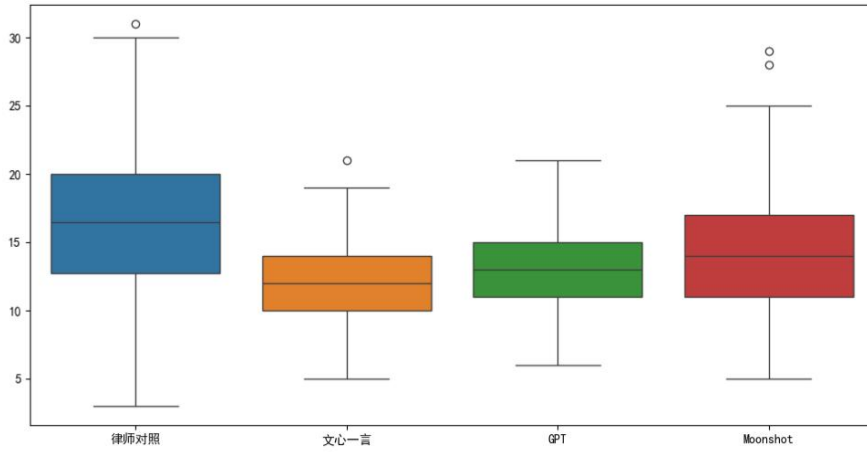
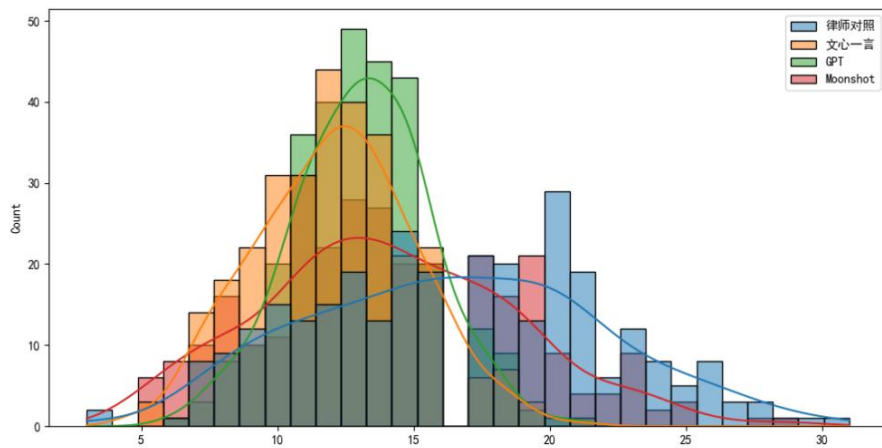


图 4.13 生成结果去除无效关键词后直方图



五、关键词向量化：将每组关键词转化成向量，并使用余弦相似度比较不同向量组之间的相似程度。

六、统计与分析：将最终得到的结果进行统计，并进行可视化处理。

图 4.14 GPT 与律师回答相似度分布图

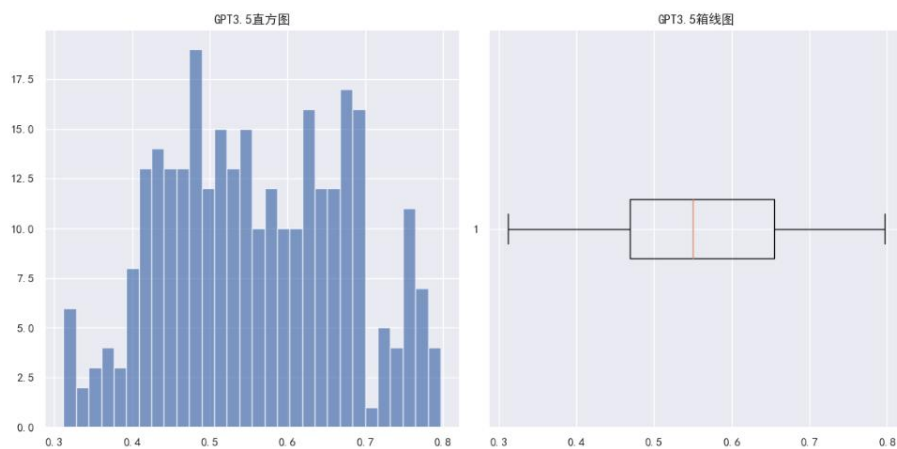


图 4.15 文心一言与律师回答相似度分布图

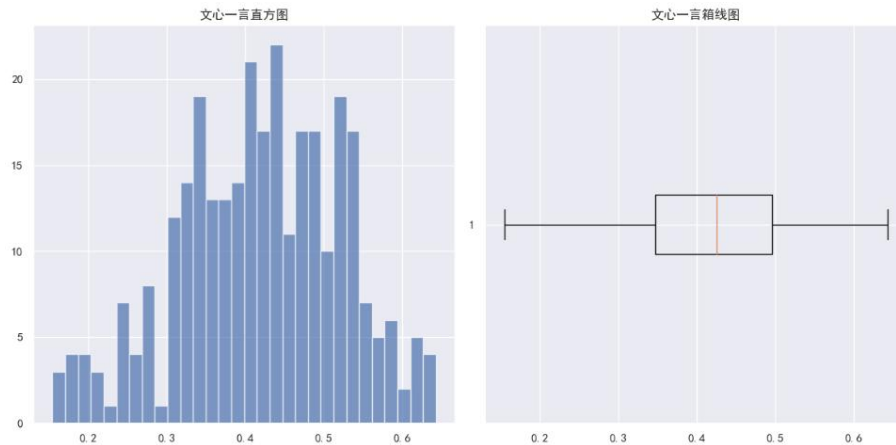
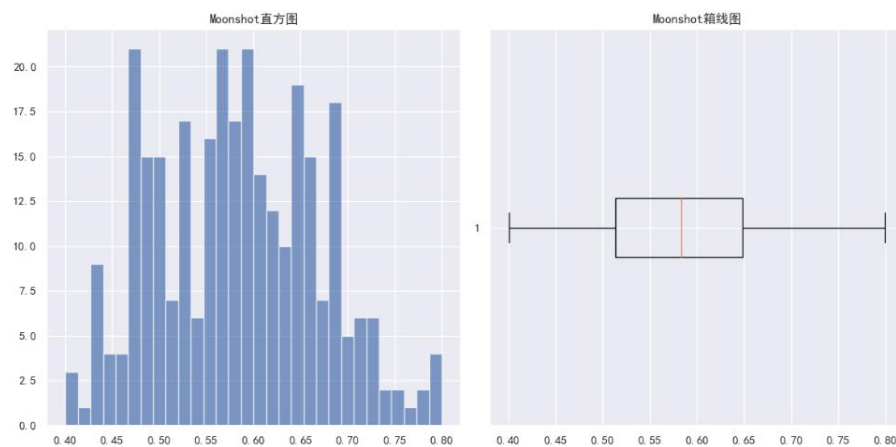


图 4.16 文心一言与律师回答相似度分布图



## 4.2 结果分析

根据得到的实验结果，我们可以进行如下总结：

一、在关键词提取中：

### 1.基本统计数据分析

**平均值：**律师回答的平均值最高（16.48），表明律师的回答在关键词数量上普遍高于其他模型。Moonshot 的平均值（14.09）是三个模型中最接近律师的，其次是 GPT3.5（13.11），文心一言的平均值最低（12.12）。

**标准差：**律师回答的标准差最大（5.36），这说明律师回答的变异性最大。Moonshot 的标准差（4.64）也相对较大，而 GPT3.5（2.50）和文心一言（2.91）的标准差较小，说明这两个模型的回答更加一致。

**极值：**律师回答的极值范围最广（3 到 31），表明律师的回答在所考虑的度量上变化最大。这反映了律师在回答问题时的个性化和情境适应性。相比之下，其他模型的极值范围较窄，则说明它们在处理问题时的方法更为一致。

**偏度系数：**所有类型的回答的偏度系数均为正值，这种现象符合我们的说

话习惯：在大多数情况下并不会说的很少，但是在有些情况下会说的很多，因此导致偏度系数为正值。

综合看来，相比于律师的回答，人工智能的回答通常更加简练，但可能内容并不齐全。注意到无论是去除无效关键词前后，律师回答的关键词分布都要比人工智能的回答均值要高，而律师回答的无效关键词相对于人工智能也更多，这可能是因为在回答时更倾向于用日常对话的方式进行回答，而人工智能则更倾向于使用简明扼要的重点陈述来进行回答。

而在人工智能的横向对比中，能够发现 Moonshot 的各项统计指标最与律师的统计指标相似，其次是 GPT3.5，最后是文心一言。原因可能是 Moonshot 具有联网搜索功能，因此更有机会能够作出类似律师的回答。而 GPT 的表现优于文心一言的原因可能是模型本身带来的差异。

## 2.相关性分析：

**相关系数：** Moonshot 与律师回答的关键词数量相关性最高（0.88），说明 Moonshot 的回答模式与律师最为接近。而 GPT3.5（0.61）和文心一言（0.64）则相关性较低。Moonshot 与律师回答的高相关性表明它在理解和模拟律师的逻辑和风格方面更为有效。

**均方误差和 R2 分数：** GPT3.5 的 MSE 最低（4.77），R2 分数最高（0.22），表明 GPT3.5 在预测律师回答方面的准确性和解释力最强，其预测与律师回答的实际值更为接近。Moonshot 的 R2 分数虽然最高（0.73），显示它在模拟律师回答的整体趋势方面表现出色，但其 MSE 相对较高（6.11），表明它在某些情况下与律师回答的差异较大。

所有模型与律师回答的拟合中，峰度系数均为负值，表示分布趋于均匀。其原因可能是模型的回答趋向结构相同，而不同律师的回答各有特色，因此导致模型与不同律师的回答在拟合的过程中出现较大的差异，因而导致峰度系数较低。

## 二、相似度分布分析

### 平均相似度：

Moonshot 的平均相似度最高（0.58），其次是 GPT3.5（0.55），文心一言最低（0.42）。在相似度计算方面，GPT3.5 和 Moonshot 的效果不相上下，可以看出 GPT3.5 除无法搜索外，模型本身的能力依然优秀，而文心一言对比两

个模型则略有差距，均值表现欠佳的同时偏度系数也为负值。

**中位数：**中位数结果与平均值类似，Moonshot (0.55) 和 GPT3.5 (0.57) 的中位数相似度较高，文心一言 (0.42) 较低。

**标准差：**GPT3.5 的标准差最大 (0.11)，Moonshot (0.10) 和文心一言 (0.09) 的标准差较小。GPT3.5 的较大标准差说明它在不同问题上的表现波动较大，而 Moonshot 和文心一言的较小标准差则说明它们在不同问题上的表现更为一致。

### 三、综合分析

Moonshot 通常与律师回答的相似度最高，可以看出它在模拟律师回答方面表现最佳。它的高相关性和相似度表明，Moonshot 可能更好地捕捉了律师回答的特点。

GPT3.5 在均方误差和 R2 分数方面表现最佳，表明它在预测律师回答的准确性上优于其他模型。然而，它的相似度略低于 Moonshot，说明它在某些情况下与律师回答的差异较大。

文心一言在所有指标上普遍低于其他两个模型，可以得到它在模拟律师回答方面的性能最弱。

### 四、原因分析

关于不同模型之间出现差异的原因，大体可以总结为以下几点：

**模型训练和优化：**不同模型的训练数据集、优化目标和算法导致它们在模拟律师回答方面的差异。例如，Moonshot 能够搜索相关的内容进行学习，这解释了其高度的相关性和相似度。

**回答风格和内容的差异：**不同模型可能在回答问题时采用不同的风格和内 容。例如，GPT3.5 和 Moonshot 可能更倾向于提供详尽的、条理清晰的回答，而其他模型可能更注重简洁和直接回答问题的核心。

**模型的泛化能力：**模型在处理不同类型问题时的泛化能力也可能影响其与律师回答的相似度。例如 GPT3.5 处理不同问题时可能会给出差距较大的回答，这会导致其整体的标准差较大。

**情境适应性：**律师在回答问题时会考虑更多的情境因素，如法律环境的具体细节和案例的特殊性。而模型难以有效地模拟这种情境适应性，这导致其与律师回答的相似度普遍不高。

## 五、结论

综上所述，Moonshot 在模拟律师回答方面整体表现最佳，尤其是在相似度和相关性方面。GPT3.5 在预测准确性方面表现较好，但在一致性上略逊于 Moonshot。文心一言在所有指标上表现较差，可能需要进一步的优化和训练以提高其性能。

## 5 讨论

### 5.1 生成式人工智能的行业价值

一、市场需求：目前对于法律咨询服务的市场需求持续增长，根据国家统计局数据显示，近年来法律服务的市场总量不断增加。这说明现阶段对于法律咨询有着充足的市场需求，并且人工智能领域尚未出现，市场空间广阔。

二、技术优势：相比于律师咨询的法律服务，人工智能拥有着处理迅速，回答客观，保护隐私的优点。人工智能法律咨询服务成本远低于律师咨询服务成本，在效果相同的情况下，有着不可替代的优势。

### 5.2 生成式人工智能面临的挑战和发展方向

挑战：

一、数据安全问题：生成式人工智能依靠大量数据进行训练，这些数据来源可能会侵犯到个人隐私，从而影响人们对人工智能的信任。并且生成式人工智能依赖大量数据进行训练。这些数据如果存在偏见，会导致人工智能产生有偏差的输出。这是我们所不愿意看到的。

二、安全问题：随着生成式人工智能技术的发展，如果监管力度不够，生成式人工智能很有可能用于制作虚假内容，用于欺诈等手段，因此生成式人工智能的监管也是关键。

三、可控性问题：生成式人工智能的输出结果有可能会出乎预料，可能会生成不准确、不恰当的内容，因此，保证其稳定性和可控性也是一大挑战。

发展方向：

一、个性化和定制化：可以针对不同人群和不同应用场景，开发相应的生成式人工智能，以应对不同的使用场合和使用需求。

二、多模态融合：结合文本、图像、声音等多种模式的数据，如支持语音对话、视频理解、图片生成，从而提高生成式人工智能的理解能力和生成能力，



以扩大生成式人工智能的应用范围。

### 5.3 对法律咨询服务行业的影响和启示

影响：

一、降低成本。通过自动化处理标准化的法律程序和咨询，生成式人工智能可以降低法律服务的成本，使得一些以往可能负担不起律师费用的个人或小型企业能够获得法律帮助。这能够大大提升社会福祉。

二、提高效率。生成式人工智能可以快速处理大量的法律文档和案例，这是律师所无法做到的。它能够带来两点好处：一是为律师提供快速的研究和参考工具，二是使法律咨询服务更容易被普通公众访问。

三、提高服务质量。结合了生成式人工智能的法律咨询服务可以有多种应用方式，如搜索类似历史情况、进行操作指引等。

启示：在技术快速发展的现在，各行各业都在追求与最新技术的融合，法律也应该接受新技术，应用新技术，以确保法律行业始终与时俱进。

## 6 结论和建议

### 6.1 发现和结论

经过本实验的结果，可以将发现整理为以下两点：

一、人工智能模型与律师回答仍存在一定差距，人工智能的回答和律师的回答在用词方面和相似度方面均需提升。

二、人工智能模型的搜索功能对于其能力有一定加成。具备了搜索功能的人工智能通常具有更强的能力。

### 6.2 对未来研究的建议

建议未来研究尝试使用法律咨询文本对人工智能大模型进行预训练，从而使人工智能具有类似律师的回答能力。同时建议未来研究尝试将搜索功能或插件加入到人工智能，使其能够以网络资源为参考进行回答，从而达到更好的效果。

## 7 参考文献

### 参考文献

- [1] 中华人民共和国司法部.全国公共法律服务体系建设规划（2021-2025 年）  
[EB/OL].[2021-12-30].<https://www.gov.cn/zhengce/zhengceku/2022-01/25/5670385/files/cca208b72837425992c31b25e548e6ad.doc>.

- [2] 中华人民共和国中央人民政府.习近平：高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告  
[EB/OL].[2022-10-25].[https://www.gov.cn/xinwen/2022-10/25/content\\_5721685.htm](https://www.gov.cn/xinwen/2022-10/25/content_5721685.htm)
- [3] 中华人民共和国司法部.2022 年度律师、基层法律服务工作统计分析  
[EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202306/t20230614\\_480740.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202306/t20230614_480740.html),2023-06-14.
- [4] 中华人民共和国司法部.2021 年度律师、基层法律服务工作统计分析  
[EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202208/t20220815\\_461680.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202208/t20220815_461680.html),2022-08-15.
- [5] 中华人民共和国司法部.2020 年度律师、基层法律服务工作统计分析  
[EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202106/t20210611\\_427394.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202106/t20210611_427394.html),2021-06-11.
- [6] 中华人民共和国司法部.2020 年度律师、基层法律服务工作统计分析  
[EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202006/t20200622\\_350049.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/202006/t20200622_350049.html),2020-06-22.
- [7] 中华人民共和国司法部.2020 年度律师、基层法律服务工作统计分析  
[EB/OL].[https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/201903/t20190307\\_350047.html](https://www.moj.gov.cn/pub/sfbgw/zwxxgk/fdzdgknr/fdzdgknrtjxx/201903/t20190307_350047.html),2019-03-07.
- [8] 科技部等六部门.关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见[EB/OL].[https://www.gov.cn/zhengce/zhengceku/2022-08/12/content\\_5705154.htm](https://www.gov.cn/zhengce/zhengceku/2022-08/12/content_5705154.htm),2022-07-29
- [9] Budzianowski P, Vulić I.Hello, it's gpt-2-how can i help you? towards the use of pretrained language models for task-oriented dialogue systems.[Z]Proceedings of the 3rd Workshop on Neural Generation and Translation,2019.
- [10] MaartenGr.KeyBERT.[Z]<https://github.com/MaartenGr/KeyBERT>,2023-10-5.fxsjy.jieba.[Z]<https://pypi.org/project/jieba/>,2020-1-20.
- [11] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space.[C]In Proceedings of Workshop at ICLR,2013.
- [12] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality.[C]In Proceedings of NIPS, 2013.
- [13] Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. Linguistic Regularities in Continuous Space Word Representations.[C]In Proceedings of NAACL HLT, 2013.
- [14] Shen Li.Chinese-Word-Vectors.[Z]<https://github.com/Embedding/Chinese-Word-Vectors/>,2023-10-30.
- [15] Shen Li, Zhe Zhao, Renfen Hu, Wensi Li, Tao Liu, Xiaoyong Du, Analogical Reasoning on Chinese Morphological and Semantic Relations[C]ACL,2018.

- [16] 前瞻产业研究院.2023-2028 年中国法律服务行业市场需求与投资战略规划分析报告 [EB/OL].<https://bg.qianzhan.com/report/detail/c8aebb3ce98e4dd3.html>,2022.
- [17] 马哥 python 说.【2023 知乎爬虫】我用 Python 爬虫爬了 2300 多条知乎评论.[Z]<https://zhuanlan.zhihu.com/p/556843537>.2023-08-04.
- [18] Yuanyuan Qiu, Hongzheng Li, Shen Li, Yingdi Jiang, Renfen Hu, Lijiao Yang. Revisiting Correlations between Intrinsic and Extrinsic Evaluations of Word Embeddings.[C]Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Springer, Cham, 2018. 209-221. (CCL & NLP-NABD 2018 Best Paper)

## 附录

提取关键词向量使用的 Python 代码:

```
# -*- coding: utf-8 -*-
import os

import jieba
from keybert import KeyBERT

def extract_keywords_chinese(text):
    # 使用 jieba 进行中文分词
    seg_list = jieba.cut(text, cut_all=False)
    text = " ".join(seg_list) # 将分词结果用空格连接, 形成新的字符串
    print(text)
    # 创建 KeyBERT 对象
    kw_model = KeyBERT(model='bert-base-chinese')

    # 提取关键词及其分数
    keywords = kw_model.extract_keywords(text, keyphrase_ngram_range=(1, 1),
stop_words=None, top_n=20)

    # 选择分数高于阈值的关键词
    selected_keywords = [[word, score] for word, score in keywords if score > 0.4]

    return selected_keywords

# 关键词向量的主函数
def main(text):
    keywords = extract_keywords_chinese(text)

    # 打印关键词
    for keyword in keywords:
        print(keyword)
    return keywords
```

```

# 获取文本
result = []
texts = os.read("./{every_text_name}.txt") # lawyer/wenxin/GPT/Moonshot
for text in texts:
    result.append(main(text))
print result

```

计算余弦相似度使用的 Python 代码:

```

import os

import numpy as np
from sklearn.metrics.pairwise import cosine_similarity
from gensim.models import KeyedVectors

# 加载预训练的词向量模型（例如，使用 Gensim 加载 Word2Vec 模型）
model = KeyedVectors.load_word2vec_format("D:\sgns.zhihu.txt", binary=False)

# 过滤函数，移除无关紧要的词
def filter_keywords(keywords, stop_words):
    return [word for word, score in keywords if word not in stop_words]

# 计算两个词的相似度
def word_similarity(word1, word2):
    if word1 in model.key_to_index and word2 in model.key_to_index:
        return cosine_similarity([model[word1]], [model[word2]])[0][0]
    return 0

# 比较两组关键词的相似度
def compare_keyword_vectors(vec1, vec2):
    similarities = []
    for word1 in vec1:
        for word2 in vec2:
            sim = word_similarity(word1, word2)
            similarities.append(sim)
    return np.mean(similarities)

# 比较两个关键词向量的余弦相似度的主函数
def main(vec1, vec2, result):
    similarity = compare_keyword_vectors(vec1, vec2)
    result.append(similarity)
    print("关键词向量的相似度:", similarity)

# 读取数据并进行比较

```

```

result = []
vec1_list = os.read(r".\vec1.txt").split()
vec2_list = os.read(r".\vec2.txt").split()
for i in range(len(vec1_list)):
    compare_keyword_vectors(vec1_list[i], vec2_list[i], result)
print result

```

部分统计数据:

去除无效关键词前关键词量分析:

	律师对照	文心一言	GPT	Moonshot
count	300.000000	300.000000	300.000000	300.000000
mean	18.110000	13.793333	14.660000	15.600000
std	5.128245	2.593526	2.218288	4.716725
min	3.000000	5.000000	8.000000	5.000000
25%	15.000000	12.000000	13.000000	12.000000
50%	18.000000	14.000000	15.000000	16.000000
75%	21.000000	16.000000	16.000000	18.250000
max	31.000000	21.000000	21.000000	31.000000

	律师对照	文心一言	GPT	Moonshot
律师对照	1.000000	0.745283	0.790327	0.965410
文心一言	0.745283	1.000000	0.637669	0.755731
GPT	0.790327	0.637669	1.000000	0.781281
Moonshot	0.965410	0.755731	0.781281	1.000000

Mean Squared Error for 文心一言: 3.1378513461518485

R2 Score for 文心一言: 0.35390843936475336

Mean Squared Error for GPT: 2.046810409750731

R2 Score for GPT: 0.6040772943365413

Mean Squared Error for Moonshot: 2.523297910862292

R2 Score for Moonshot: 0.8929848678301653

去除无效关键词后关键词量分析:

	律师对照	文心一言	GPT	Moonshot
count	300.000000	300.000000	300.000000	300.000000
mean	16.483333	12.123333	13.113333	14.086667
std	5.361487	2.905672	2.504943	4.641584
min	3.000000	5.000000	6.000000	5.000000
25%	12.750000	10.000000	11.000000	11.000000
50%	16.500000	12.000000	13.000000	14.000000
75%	20.000000	14.000000	15.000000	17.000000
max	31.000000	21.000000	21.000000	29.000000

	律师对照	文心一言	GPT	Moonshot
律师对照	1.000000	0.640638	0.611252	0.879662
文心一言	0.640638	1.000000	0.424029	0.614442

GPT            0.611252   0.424029   1.000000   0.604656  
Moonshot   0.879662   0.614442   0.604656   1.000000

Mean Squared Error for 文心一言: 7.32110825729145

R2 Score for 文心一言: 0.023707596449503043

Mean Squared Error for GPT: 4.768999466596605

R2 Score for GPT: 0.2218818854356519

Mean Squared Error for Moonshot: 6.105491712772007

R2 Score for Moonshot: 0.7344155369021359

关键词相似度分析:

GPT3.5 与律师回答相似度分布

Mean: 0.5533189889720057

Median: 0.5571134601085044

Standard Deviation: 0.11412146822716553

Min: 0.3010368311818521

Max: 0.7884678368282604

文心一言与律师回答相似度分布

Mean: 0.42147289808545707

Median: 0.42291998685199866

Standard Deviation: 0.10665747881810692

Min: 0.1543927673305402

Max: 0.6442277994985475

Moonshot 与律师回答相似度分布

Mean: 0.5827853614857839

Median: 0.5771417150001528

Standard Deviation: 0.09203284663167866

Min: 0.4060388562106841

Max: 0.7990614122301192