1、研究现状（*附主要的参考文献目录*）

1.1 问题的提出

在交通肇事案件中，各种证据可能是错综复杂的，例如在2019年的余金平案件中，被告人余金平有酒后驾驶、逃逸、自首、得到被害人家属谅解等多种情节，而检察院给出的量刑建议是缓刑与法院一审给出的有期徒刑两年和二审给出的三年徒刑的量刑相去甚远。同时在2015年的一项研究显示量刑建议被法院采纳的几率仅为57.9%[1]，由此可见在证据对于量刑结果的影响上，给出量刑建议的检察院与做出判决的法院的观点并不相同。而知识图谱可以直观地展现出案件要素、要素属性及案件要素间的关系，分析各类证据对于量刑结果的影响，并可以在给出证据的情况下对于量刑结果进行预测，提高量刑建议的采纳率。

1.2 国内外研究现状

知识图谱在交通事故的分析与预测中已有相关应用，其中王翠,胡昊天等人利用事理图谱技术构建道路交通领域知识库，揭示道路交通违法行为事件之间的逻辑关系，他们以开源数据为基础，面向道路交通违法行为构建事件语料库，通过事件抽取、关系抽取等步骤，构建交通违法事理图谱，在此基础上，实现了交通违法问答系统[2]。季青原, 吴建平等人基于交通事故知识图谱,使用GraphSAGE模型对交通事件进行表示学习,随后进行聚类分析,并删除交通事故随机噪声。然后使用多分类任务的全连接深度神经网络,利用事故的向量表示对事故标签进行预测。同时生成特定标签事故主要成因表,用以支持下游的事故预警以及防控等具体工作[3]。

此外在其他司法量刑领域，知识图谱也有相关的应用，例如在检察官办案过程中对盗窃案件性质把握不准确, 对量刑建议给出缺乏经验, 导致给出的量刑建议准确性不足等问题，乔钢柱、冯婷婷等人对盗窃案件法律文书理论和知识体系进行整理和分析, 对其中的隐式关系、深层关系进行挖掘、推理, 通过搭建本体模型, 提出了基于本体的盗窃案件法律文书知识图谱构建方法, 并且设计自定义推理规则, 实现了盗窃案件法律文书知识图谱在相似量刑类案推送测试功能[4]。王治政,王雷等人将量刑预测任务重新定义为基于知识图谱的链路预测任务，提出多视角的知识图谱嵌入方法，预测案件量刑。作者针对贩卖毒品类案件设计知识图谱本体模式，用于指导案情描述中关键要素的提取。然后使用图嵌入技术，从案件要素构成的辅助图中学习要素的初始表示，并融合知识图谱的结构化特征，对案件要素进行增强表示[5]。

1.3 研究的意义

在交通肇事量刑领域，针对被告人是否逃逸、得到被害人谅解等情节给出的量刑建议被法院接纳率较低，说明给出量刑建议的一方对于相关证据与交通事故量刑间的关系与法院方有所区别，本课题意在通过分析过往的交通肇事判决书并构建知识图谱，将判决书中信息结构化，同时可以通过输入相关的证据来预测该案件的量刑结果，对量刑建议的给出做出指导。

2、项目的研究内容、研究目标以及拟解决的关键问题。

2.1研究内容：

1. 基于判决书的关系抽取：对于中国裁判文书网上关于交通肇事案件的判决书进行关系抽取，得到相应的三元组。
2. 基于三元组的知识融合：将关系抽取得到的三元组进行融合得到结构化的图数据库。
3. 基于jena的量刑推理：构建推理规则，通过类似案件的判决结果推测出当前案件的量刑。

2.2 研究目标

使用知识图谱将交通肇事案件中的信息结构化存储，并能据此通过输入交通肇事案件的信息给出相应的量刑预测。

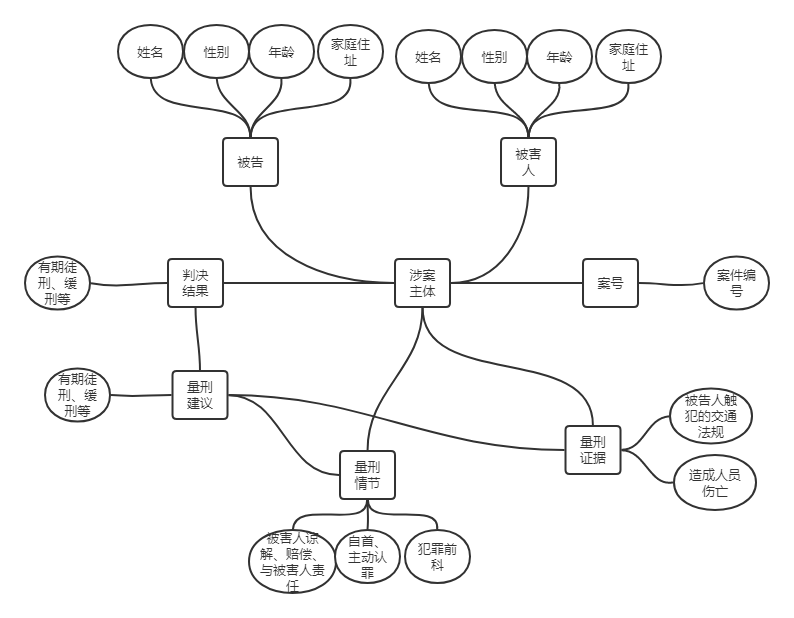
2.3 拟解决的关键问题

量刑建议被法院接受的概率较低，拟采用知识图谱，通过分析过往的判决书给出量刑预测，提高量刑建议的接纳率。

3、拟采取的研究方案及可行性分析

3.1 拟采取的研究方案

3.1.1交通肇事案件知识图谱本体的构建



3.1.2三元组提取

实体抽取：在课题背景下，主要识别的对象为人、量刑证据、量刑情节、机构、量刑结果。拟采用工具jieba完成既定目标，将判决书中常见的词语例如“逃逸”、“有期徒刑一年”等建立为一个词典，指导jieba进行分词并对实体进行标注。

关系识别：考虑到本课题中实体之间的关系较为明确，故采用远监督关系抽取的方法应该会有不错的效果。首先定义实体之间的主要关系如下表所示。

主要实体间关系：

|  |  |  |
| --- | --- | --- |
| 头实体 | 尾实体 | 关系 |
| 机构 | 判决结果 | 建议获刑 |
| 机构 | 判决结果 | 实际获刑 |
| 人 | 量刑证据 | 产生 |
| 人 | 量刑情节 | 产生 |
| 人 | 量刑情节 | 产生 |

PCNN模型训练：

1：将分词结果中的每一个词转换为词向量，并计算该词与两个实体之间的距离，构成卷积神经网络的输入。句向量标识为 其中s为句子单词数量，d为词向量加两个位置向量的维度。

2：对于长度为s的句子，首尾填充w-1长度（w为卷积长度）使用n个卷积核进行卷积操作并得到n个卷积结果。

3：分段最大池化，按照实体所在位置将卷积层输出划分为3部分进入池化层并得到输出经过非线性函数输出。

4：将输出转换为类别分数（W1为转换矩阵）

3.1.3图数据库生成

由三元组生成实体csv（entity.csv）和关系csv（relationship.csv）文件

entity.csv(ID、name、label)

relationship.csv(START\_ID,name,END\_ID,TYPE)

使用neo4j将csv文件转为图数据库。

3.1.4量刑预测

将提出出的三元组转换为RDF形式，使用jena推理引擎，并定义规则有相同的量刑证据和量刑情节的判决结果也会相同完成推理过程。

1. 丁琪,陈颖,廖思谦.量刑证据的司法适用——以交通肇事罪判决书为样本的分析[J].河南警察学院学报,2015,24(02):96-100.
2. 王翠,胡昊天,邓三鸿.面向道路交通违法行为的事理图谱构建技术[J].交通信息与安全,2022,40(01):36-44.
3. 季青原, 吴建平, 徐甲,等. 一种基于知识图谱的交通事故分析与防控方法及系统:, CN112732905A[P]. 2021.
4. 乔钢柱,冯婷婷,张国晨.基于知识图谱的盗窃案件法律文书智能推理研究[J].计算机系统应用,2019,28(07):206-213.
5. 王治政,王雷,李帅驰,孙媛媛,陈彦光,许策,王刚,林鸿飞.基于多视角知识图谱嵌入的量刑预测[J].模式识别与人工智能,2021,34(07):655-665.DOI:10.16451.

基于知识图谱的交通肇事案件量刑推荐技术

相关研究：2.1交通肇事案件量刑技术相关研究进展及问题分析

范例：XXX2021年在XX提出了 针对XXX场景 解决了XXX问题 达到了XXX效果 问题 存在不足

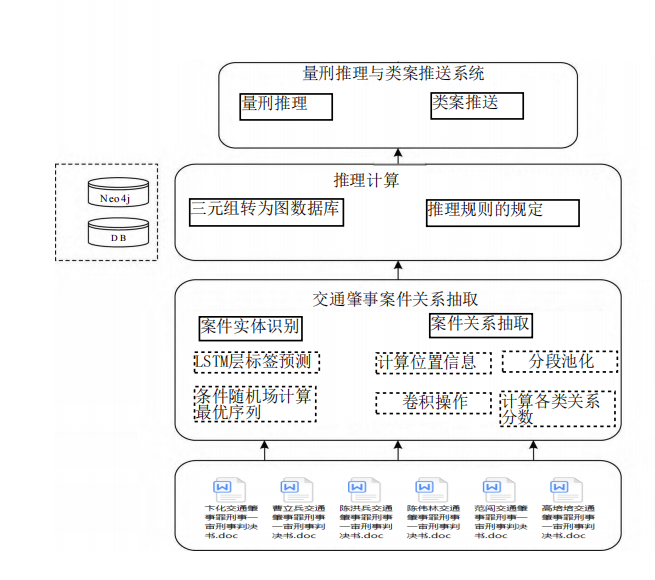
2.1.2进展和问题小结

说明现在研究 没有用知识图谱 知识图谱先进

2.2 基于知识的司法量刑推荐技术相关研究进展及问题分析

2.2.1相关研究：没有用到交通肇事案件；用到了司法 知识图谱识别方面推荐算法方面不足

2.2.2进展和问题小结

3总体设计

研究思路、关键技术

4