

彭科钦

19520482118 | keqin.peng@mail.bnu.edu.cn | [个人主页](#)

概要

彭科钦，2022 年 7 月毕业于北京师范大学人工智能学院并取得理学学士学位。

主要研究兴趣包括自然语言处理方法的应用和改进，关注文档级别信息抽取和挖掘。包括命名实体识别、实体关系抽取、事件抽取、文本摘要等。

教育经历

北京师范大学

本科, 人工智能学院

2018.09-2022.07

工学学士学位

获得京师奖学金二等奖

相关课程成绩: 数据库系统原理: (98/100), Python 编程: (94/100), 数据挖掘: (88/100), 软件工程: (86/100)

科研实习经历

清华大学人工智能产业研究院

NLP 实习生 (唯一)

2022.07-2023.04

研究方向: 从文本自动提取医疗实验程序并优化

负责内容:

- 制作医疗实验文本数据集, 包括收据收集、格式化、清洗、整理、标注、校对等;
- 对现有方法进行技术验证, 包括命名实体识别、实体关系抽取、语义角色标注、知识图谱, 搭建流水线;
- 提出新任务: 从自然语言文本中提取关键步骤流程图;
- 重新设计端到端模型已应对上述任务, 解决错误积累和错误传播问题;
- 与医疗实验室对接, 完成新技术部署;
- 总结方法, 完成论文。

北京师范大学

第二作者

2022.03-2022.07

研究方向: 文档级关系抽取方法改进

负责内容:

- 在已有数据集 (DocRED, CDR, GDR, etc.) 基础上, 完成现有模型 (CNN, Transformer, etc.) 技术验证;
- 设计并搭建新模型, 对现有模型进行优化 (如减少冗余数据输入、像素化处理实体关系等), 以提高文档级别关系抽取任务的准确性;
- 总结方法, 完成论文。

北京师范大学

毕业设计, 第一作者

2020.10-2022.02

研究方向: 多文本自动摘要生成方法改进

负责内容:

- 在已有数据集 (WikiSum) 基础上, 完成已有模型 (CNN, RNN, LSTM, Single Transformer) 技术验证;

- 设计两个 Transformer 层，对多文档内容进行自动总结摘要。突破了传统方法对文本尺寸的限制，同时保留多篇文档中对单一论点的论据和证伪，保留事件结构。

发表论文

U-Net with depthwise separable convolutions and big kernels for document-level relation extraction :

颜荣恩, 彭科钦, 党德鹏

2023 11th Transactions of the Association for Computational Linguistics. [Reviewing]

Document-level relation extraction with less redundant image-style feature :

颜荣恩, 彭科钦, 党德鹏

2023 11th Transactions of the Association for Computational Linguistics. [Reviewing]

基于事件结构的多文档摘要方法研究:

彭科钦, 王学松

北京师范大学毕业论文

项目经历

数据库开发: 科研任务发布平台

2021.03- 2021.05

- 有兴趣于跨学科的科研任务的学生苦于找不到集中的科研招聘信息，老师也面临着招不到足够多交叉专业学生的问题。我们开发科研任务发布平台，帮助匹配需要跨学科科研合作的师生。
- 平台收录和展示老师发布的科研任务信息，包括任务描述、截止日期、需要的人数以及相关专业。同时也将符合专业的申请信息推送给老师以供判断和筛选。
- 数据库开发主要负责建立数据库、日常维护与监控、根据项目需要，进行数据库结构更改、跟踪、优化等操作。

活动经历

美国大学生数学建模大赛-编程手:

2021.01-2021.03

- 与队员讨论不同的数学模型
- 将数学模型转换为可执行代码

河渠希望小学-支教成员、数学教研组组长

2019.01-2019.02

- 组织研讨数学相关支教事项，授课并且组织考试

相关技能

编程语言

C, C++, PostgreSQL, Latex, Matlab

办公软件

Microsoft Office Specialist(Expert), Overleaf, tableau

英语能力

四级 552:, 六级:518, 托福:97, 考研英语: 81