

杜晋华

唐杰教授 大语言模型方向四年级博士 2027-2028 毕业



绩点: 3.95/4.00

性别: 男

院系: 计算机系

排名: 1/329

年龄: 25

面貌: 中共党员

职务: 计算机系研究生党支部书记、计研五三班级带班助理(辅导员)



学历技能

Sep. 2022 - Jun. 2027

计算机科学与技术系工学博士

清华大学

- 专业主修: 高级机器学习 (A)、大数据分析 (A)、计算语言学 (A)、信息检索的前沿研究 (A)、网络计算与区块链技术 (A)、数智安全与标准化 (A)、数据挖掘原理与算法、知识工程、人工智能原理等
- 研究方向: 大模型、自然语言处理、机器学习、大数据分析
- 主要成果: 论文 13 篇: ML 方向 CCF-A 1 篇, NLP 方向 CCF-A 1 篇, CCF C-2 篇, LLM 方向 CCF A-2 篇, SCI 一区-2 篇, 三区-1 篇, 完整训练 5 个大模型: GLM、LogicGLM、MalayGLM、AiMed、公安大模型

Sep. 2018 - Jun. 2022

计算机科学与技术系工学学士

中国地质大学(北京)

- 专业主修: C++程序设计 (100)、科学与工程计算 (100)、高性能计算 (99)、人工智能 (97.9)、操作系统原理 (97)、数据结构 (94)、软件工程 (优秀) 等
- 研究方向: 机器学习、数值分析、机器人 (多智能体)
- 主要成果: 论文 6 篇: 机器学习与数值计算/建模方向 3 篇论文 (2 篇 EI, 1 篇核心), 机器学习与智能控制、机器人方向 3 篇论文 (1 篇 SCI, 1 篇 EI, 1 篇核心)
- 英语水平: IELTS(7.0)、CET6(525)、CET4(580)
- 相关技能: 掌握 Python、C/C++、熟悉 JAVA、Lua、Go 等编程语言, 熟悉使用 MATLAB 进行数值分析与计算

竞赛获奖

评奖评优

- 国家奖学金、国家励志奖学金、清华大学优秀研究生干部、金奖支队队长、硕博论坛卓越奖(第一名)
- 全区公安机关外聘专家、北京市优秀本科毕业生、北京市优秀本科毕业论文、北地先锋“十佳学生”

国家级竞赛

- 一等奖: 美国大学生数学建模、青创北京 2024 年“挑战杯”首都大学生创业计划竞赛主赛道金奖
- 二等奖: 全国高校计算机能力挑战赛、Robocup 机器人世界杯中国赛

省部级竞赛

- 一等奖: 蓝桥杯全国软件和信息技术专业人才大赛、移动应用创新赛
- 其他有代表性共计 50 余项目竞赛获奖, 参见抬头靠右二维码或个人主页 https://dujh22.github.io/Dujinhua_wiki/

科研经历

主研项目(课题负责)

面向基础模型深度复杂推理的关键技术研究

博士理论课题

- 基础模型复杂推理研究 子课题/科技部重点研发计划(清华大学、天津大学、智谱 AI)
 - 论文《Glm-4.5: Agentic, reasoning, and coding (arc) foundation models》Arxiv
 - 论文《A Survey of Post-Training Scaling in Large Language Models》CCF-A:ACL2025
 - 论文《Enhancing Test-Time Learning of LLM Agents via RL over Memory》CCF-A:ICML2026
 - 论文《Toward Evolvable Evaluation for Logical Reasoning》(在审) CCF-A:NeurIPS
 - 论文《A Self-Evolving Closed-Loop Learning Framework for Logical Reasoning》CCF-A:AAAI 在投

主研项目(课题负责)

面向医疗大数据分析利用的自然语言处理技术研究

博士应用课题

- 支持卫生健康数据要素的医疗知识图谱 子课题
 - 论文《中文电子病历命名实体识别的研究与进展》CCF-A:电子学报
 - 论文《KerNER: A Novel Named Entity Recognition Method Based on Knowledge Enhancement and Remote Supervision》CCF-C:TrustCom
 - 论文《KLDP: A Data Profiling Technique Based on Knowledge Graph and LLM》CCF-C:TrustCom
- 面向数据和知识双轮驱动的智慧医疗大模型研究 AiMed 子课题
 - 报告《ChatGPT 创造了 AI 狂潮》微软邀约
 - 论文《AiMed: Artificial Intelligence large language model for chinese Medicine》EI:MedAI/软著
 - 论文《Towards Artificial Intelligence for Science:A Case Study of Using ChatGPT》SCI 一区:big data Informatics and Health
 - 论文《Large Language Models Driven Reliable Clinical Decision-Making》SCI 一区:Engineering
 - 论文《Toward a Large Language Model-Driven Medical Knowledge Retrieval and QA System: Framework Design and Evaluation》
- 负责选择课题、课题调研、文献综述、理论创新、相关实验和论文书写

参研项目(技术)	区块链链上链下数据可信交互关键技术研究	纵向-科技部
- 大模型与区块链 子课题 - 网站《Open Data Entry 数据开放入口》 公共服务 - 论文《Is ChatGPT All That MedNLP Needs? A Systematic Evaluation of Knowledge Discovery Capabilities across Biomedical NLP Tasks》 Nature 在投 - 论文《OpenMonet: Open Model Orchestration Network》 在研 - 负责专业技术交叉研究、相关理论创新与应用落地、论文书写		

参研项目(技术)	老年肾功能减退临床综合评估管理及早期预警体系建立	纵向-科技部
- 数智涌现的新一代智慧医疗 子课题 - 论文《Research of Client Selection Algorithm in Cross-device Federated Learning》 CCF-A:软件学报 - 医疗数据标注系统、中文分词处理系统 相关服务 - 负责医工学科交叉研究、相关场景创新与应用落地、论文书写		

创业项目(项目负责)	基于强化学习的足球机器人教练系统研发	国家级
- 基于 ROS 的多级足球机器人系统的开发 国家级 - 论文《基于 Ubuntu 下 ros 变种 rocos 等操作系统架构的研究》 科技核心 - 论文《System Composition and Optimization of Small Soccer Robot》 外文 EI - 论文《基于改进人工势场和新型网格的路径规划》 北大核心 - 负责项目整体研究思路梳理和研究内容分工，完成核心 AI 算法开发和系统搭建		

合作项目(组长)		
- 论文《基于主成分分析对禁瓶令效果的研究》 国家刊物 - 论文《Global epidemic classification based on K-nearest neighbor algorithm》 外文 EI - 论文《Mid-term and long-term prediction of carbon emissions in Jiangsu Province based on PCA-STIRPAT improved GA-BP》 外文 EI		
负责 AI 算法开发与实验，文献查阅和调研		

外包项目(项目负责)		
- AiMed 医学知识大模型应用服务系统 软件著作权 - 风力发电机气动平衡检测系统、智慧型半人工酒店系统 软件著作权		
控制项目进度，负责建立系统框架，核心程序开发，方案展示及答辩		

实习经历

Mar. 2023 - Now		
中国北京智谱华章科技股份有限公司—AI 院	AI 研发岗	2023.03-Now
Sep. 2021 - Mar. 2023		
中国湖南网数科技有限公司—网络大数据研究中心	AI 研发岗	2021.09-2023.03
Sep. 2018 - Jun. 2022		
中国科学院信息工程研究所	助理研究岗	2021.06-2021.09
中国阿里云原生团队	算法岗	2021.03-2021.06
中国地质大学（北京）信息技术创新中心	竞赛主管	2018.09-2021.03

学工经历

Sep. 2022 - Present	
唐杰教授《高级机器学习》课程助教，第一作者负责书籍《大语言模型》的编写和修改，设立对应网站	
担任清华大学研团实践部定岗组长，构建优化“同行实践平台”微信小程序，持续服务全校硕博研究生	
在清华大学新闻网上以第一通讯员身份发表文章	
担任计算机系研团实践部书记，作为队长负责联络宣传带队全部的七系联合产业行活动，包括北京字节跳动、腾讯、美团、九坤投资、网易有道、杭州 13 家科技企业走访，主稿 23 篇推送	
担任带班助理，获得 3 次集体荣誉：笃行班团、卓越班团（全校仅 5 个）、清华大学甲级团支部	
获得 10 次个人荣誉：专业二等奖学金、社会工作二等奖学金、社会实践奖学金、清华大学优秀共青团员、校研团组长任职经历证明、计算机系研团书记任职经历证明、团支部书记任职聘书、新生骨干优秀学员	
担任校学生会体育部骨干，长期坚持跑步和健身，获得马拉松 10km 奖牌，半马奖牌，健美小组第五名	

个人陈述

认真负责，严谨细致，逻辑思维和学习能力强；拥有良好的人际沟通能力且善于团队合作。希望经过博士阶段系统的科研训练，在未来做出真正造福世界、改变世界，助力全人类向善、向好、向幸福的事。