

张宁豫

✉ zhangningyu@zju.edu.cn · ☎ (+86) 136-4571-7011 · 🌐 github · 🏠 homepage

📖 研究方向 (H-INDEX 9, CITATION 228)

知识图谱, 自然语言处理, 机器学习

🎓 教育背景

浙江大学 2012 – 2017

博士 计算机科学与技术, 导师: 陈华钧

浙江大学 2008 – 2012

学士 软件工程

👨‍💻 工作经历

阿里巴巴 (高级算法工程师) 2019 年 2 月 – 至今

1. 阿里概念图谱构建 (AliCG) 2019 年 2 月-2019 年 9 月

从零到一自动化构建细粒度和动态化的大规模概念图谱, 显著提升搜索业务效果, 产出论文 1 篇

- 提出自举对齐和弱监督融合一致性算法自动化挖掘海量细粒度概念, 基于用户行为实时更新图谱, 挖掘超过 400 万的细粒度概念和 2000 万的概念三元组
- 部署概念图谱服务, 提升 query 意图识别 1%, 搜索满意度 3%
- 阿里概念图谱获得了多个媒体如将门创投、AICug 等广泛报告, 投稿 KDD2020 论文 1 篇

2. 文本信息抽取 2019 年 2 月-2020 年 3 月

研究和开发文本信息抽取算法, 赋能通用和垂直搜索、问答和推荐, 产出论文 4 篇

- 基于极少 (2000 条) 样本、词典和大量无标注样本, 提出弱监督融合一致性算法, 在医疗领域实体识别任务上的准确率、召回率均超过 0.9, 垂搜满意度提升 3%
- 设计并组织开发基于远程监督和主动学习的通用信息抽取框架, 实现实体、关系、事件、情感、意图等信息的端到端抽取, 赋能文本内容理解, 提升知识图谱的实体覆盖率 5%
- 投稿 ACL2020 论文 1 篇, 发表 WWW2020、WSDM2020、NAACL2019 论文各 1 篇

3. 融合知识的自然语言处理 2019 年 9 月-2020 年 3 月

研究融合知识的预训练等自然语言任务, 显著提升医疗搜索业务效果, 产出论文 2 篇, benchmark 1 个

- 提出医疗领域融合知识的预训练算法, 基于医疗短语、实体的 mask 实现知识注入, 显著提升文本不一致判别准确率 10%, 搜索意图识别准确率 2%, 搜索满意度 3%
- 发表 AAAI2020、WSDM 2020 Health Day 论文各 1 篇, 投稿 SIGIR2020 论文一篇, 产出中文医疗 benchmark 1 个

之江实验室 (助理研究员) 2018 年 1 月 – 2019 年 2 月

1. 文本关系抽取 2018 年 1 月 - 2019 年 2 月

研究文本关系抽取算法, 产出论文 1 篇

- 提出基于胶囊网络的多关系抽取模型, 在 NYT 数据集上多关系抽取的准确率提升 2%, 发表 EMNLP2018 论文 1 篇
- 提出基于对抗生成网络的低资源知识图谱补全算法, 显著提升低资源关系抽取、知识图谱链接预测效果

2. 你说我画

2018 年 1 月 – 2019 年 2 月

指导学生优化文本生成图片模型，产出原型系统 1 个

- 基于 AttnGAN 和中文语料改进文本生成图片的模型
- 基于词的 Concept 信息和跨语言对抗训练增强表示学习
- 在中文语料取得 SOTA 效果, 英文语料较 AttnGAN Inception Score 提升 1%

3. 通用 NLP

2018 年 5 月 – 2018 年 9 月

研究文本匹配和文本分类算法，参加比赛和提供技术支持

- 基于 Siemase Network, ABCNN, MatchPyramid Network 训练文本匹配模型, 使用基于字符向量和词向量的联合表征, 取得蚂蚁 ATEC2018 复赛 top50
- 为第一届之江杯人工智能大赛提供算法支持
- 为之江印度 Hackathon 文本分类比赛提供数据及算法支持

👤 读博期间项目

知识图谱补全

2017 年 1 月 - 2018 年 1 月

研究知识图谱补全算法，产出论文 1 篇

- 提出基于地理空间位置知识约束的图谱补全算法，显著提升关系预测准确率，发表论文 1 篇

基于迁移学习的城市计算应用

2015 年 1 月 - 2017 年 1 月

研究基于迁移学习的城市计算，产出论文 2 篇，专利 1 个

- 基于迁移学习模型，解决目标城市数据稀疏性问题，并在真实应用场景如最优商铺选址、雾霾分析、内涝分析等取得较好效果
- 基于领域对抗训练改进算法，显著提升迁移学习效果

👤 论文

- **Relation Adversarial Network for Low Resource Knowledge Graph Completion**
Ningyu Zhang, Shumin Deng, Zhanlin Sun, Jiaoyan Chen, Wei Zhang, Huajun Chen
WWW 2020 (CCF A)
- **Long-tail Relation Extraction via Knowledge Graph Embeddings and Graph Convolution Networks**
Ningyu Zhang, Shumin Deng, Zhanlin Sun, Guanying Wang, Xi Chen, Wei Zhang, Huajun Chen
NAACL 2019. (NLP top conference)
- **One-Pass Extraction of Overlapping Triplets and Low-Resource Relations via Hierarchical Bidirectional Capsule network**
Ningyu Zhang, Shumin Deng, Jiaoyan Chen, Wei Zhang, Huajun Chen
ACL 2020 under review (CCF A)
- **AliCG: Fine-grained and Evolvable Conceptual Graph Construction for Semantic Search at Alibaba**
Ningyu Zhang, Qianghuai Jia, Shumin Deng, Gang Huang, Zhao Wang, Liang Dong, Wei Zhang, Huajun Chen
KDD 2020 under review (CCF A)
- **Entity Retrieval with Implicit Structure via Contextual Pretraining and Knowledge Graph Embedding**
Ningyu Zhang, Shumin Deng, Wen Zhang, Jiaoyan Chen, Wei Zhang, Huajun Chen
SIGIR 2020 under review (CCF A)
- **Meta-Learning with Dynamic-Memory-Based Prototypical Network for Few-Shot Event Detection**
Shumin Deng, Ningyu Zhang, Jiaojian Kang, Yichi Zhang, Wei Zhang, Huajun Chen
WSDM 2020 (CCF B)
- **Conceptualized Representation Learning for Chinese Biomedical Text Mining**
Ningyu Zhang, Qianghuai Jia, Kangping Yin, Liang Dong, Feng Gao, Nengwei Hua
WSDM 2020 Health Day (CCF B)

- **When Low Resource NLP Meets Unsupervised Language Model: Meta-pretraining Then Meta-learning for Few-shot Text Classification**
Shumin Deng, **Ningyu Zhang**, Zhanlin Sun, Jiaoyan Chen, Huajun Chen
AAAI 2020 (CCF A)
- **Context-aware Deep Model for Entity Recommendation in Search Engine at Alibaba**
Qianghuai Jia, **Ningyu Zhang**, Nengwei Hua
CIKM 2019 workshop (CCF B)
- **Knowledge-Driven Prediction and Explanation for Stock Fluctuation via Temporal Convolution Network**
Shumin Deng, **Ningyu Zhang**, Jiaoyan Chen, Jeff Pan, Huajun Chen, Wei Zhang
WWW 2019 workshop (CCF A)
- **Attention-Based Capsule Networks with Dynamic Routing for Relation Extraction**
Ningyu Zhang, Shumin Deng, Zhanling Sun, Xi Chen, Wei Zhang, Huajun Chen
EMNLP 2018 (CCF B)
- **Structured Knowledge Base as Prior Knowledge to Improve Urban Data Analysis**
Ningyu Zhang, Shumin Deng, Huajun Chen, Xi Chen, Jiaoyan Chen, Xiaoqian Li, Yiyi Zhang
ISPRS International Journal of Geo-Information (2018). (IF=2.022)
- **Location Based Link Prediction for Knowledge Graph**
Ningyu Zhang, Xi Chen, Jiaoyan Chen, Shumin Deng, Wei Ruan, Chunming Wu, Huajun Chen
Journal of Chinese Information Processing (2018). (Core Journals)
- **Semantic framework of internet of things for smart cities: case studies**
Ningyu Zhang, Huajun Chen, Xi Chen, Jiaoyan Chen (2016)
Sensors. (IF=3.302)
- **Transfer Learning for Urban Computing: A Case Study for Optimal Retail Store Placement**
Ningyu Zhang, Huajun Chen, Xi Chen
KDD 2015 workshop. (CCF A)
- **Big smog meets web science: smog disaster analysis based on social media and device data on the web**
Jiaoyan Chen, Huajun Chen, Guozhou Zheng, Jeff Z Pan, Honghan Wu, **Ningyu Zhang**
WWW 2014 (CCF A)

👤 专利

- 国内授权专利 4 个
- 国内受理专利 6 个

⚙️ IT 技能

- 编程语言: Python, Java, C++, PHP
- 大数据处理框架: Spark、Hadoop
- 机器学习框架: Tensorflow, Keras, Pytorch, Scikit-learn

♡ 获奖情况

浙江大学优秀博士生	2016
浙江大学优秀硕士生	2014

📄 其他

- 个人主页: <https://zxlzr.github.io/>
- Github: <https://github.com/zxlzr>