

方林涛 / LINTAO FANG

📍 443 Huangshan Road, USTC, Hefei, 230000 📞 +86 18756073629 ✉ flt@mail.ustc.edu.cn

🌐 linkedin.com/in/lintaofang/ 🐙 github.com/rat-racer

求职目标: 寻求算法工程师, 推荐算法工程师或数据挖掘工程师职位

EDUCATION

中国科学技术大学	Hefei, Anhui, China
计算机科学与技术 / 硕士 机器学习、数据挖掘、图神经网络	Sept. 2019 – July 2022
中国科学技术大学	Hefei, Anhui, China
计算机科学与技术 / 本科 优秀新生、国家励志奖学金	Aug. 2015 – July 2019

WORKING EXPERIENCE

中国科学技术大学 - 大数据分析与应用实验室	Hefei, China
硕士研究生、导师: 徐童副教授	Sept. 2019 – July 2022
<ul style="list-style-type: none">- 研究方向为数据挖掘, 主要从事于 GNN 图神经网络, 网络表征等机器学习相关领域的理论和应用研究。- 应用网络表示学习在专利挖掘, 公司竞争关系挖掘等领域, 挖掘新的问题并基于 GNN 提出新的解决方案。- 针对特点场景与问题, 选择合适的数据集和特征表示方法, 设计算法框架模型, 实现并优化模型效果。- 研究成果已被录用论文 3 篇, 目前在投 1 篇 (CCF 2B + 2C), 已申请专利 1 项。	
华为云 - Cloud BU 知识图谱技术团队	Hangzhou, China
算法实习生、导师: 郑毅博士	Jan. 2021 – May 2021
<ul style="list-style-type: none">- 阅读调研知识图谱表示学习相关的最新学术研究工作小组并作分享, 预研并实现 KG Embedding 算法。- 模型基于 GraphSAGE 和 RGCN, 解决目前知识图谱嵌入方法对图谱更新敏感以及缺乏利用实体属性问题。- 支持图谱 mini-batch 训练方式, 图谱发生更新 (增删改) 时, 实现对节点的动态更新, 无需重新训练。- 期间继续推进科研任务, 设计并实现算法模型, 完成基于图表征技术的技术流动预测的科研项目。	

PUBLICATIONS

Patent2Vec: Multi-view Representation Learning on Patent-Graphs for Patent Classification (World Wide Web)
Deep Technology Transfer Prediction via Multi-aspect Dynamic Network Embedding (ICDM 21 Under Review)
Ontological Structure Aware Knowledge Transfer for Incremental Knowledge Graph Embedding (IJCNN 2020)
Author Contributed Representation for Scholarly Network (APWeb-WAIM'2020)

NOTABLE PROJECTS

Patent2Vec - 基于多视图网络表示学习的专利嵌入与自动分类	USTC, Hefei
硕士研究生、导师: 徐童副教授	June 2020 – Oct. 2020
<ul style="list-style-type: none">- 根据专利领域知识构建多视图网络, 使用 Graph SAGE 和 Metapath2vec 实现单视图的表征。- 提出了多视图表示学习模块, 其中包括多视图 Enhance 和 Align 部分, 最终得到提炼后的 Embedding。- 较基准模型 (Deepwalk, node2vec, GCN 等) 提高了 6% 的 Precision@1 和 4% 的 NDCG@5 score。- 论文已被 WWW 录用, 目前跟进的技术流动预测论文已经投递于数据挖掘领域顶会 ICDM。	
基于 Spark 和 Tensorflow 的电影推荐系统	Java, Jetty, Redis, Tensorflow, Keras, Spark, JavaScript
<ul style="list-style-type: none">- 使用 Spark 进行数据预处理, 生产训练样本和特征, 样本供于模型训练, 特征存入 Redis 供线上推断使用。- 使用 Tensorflow 和 Spark MLlib 实现 CF, Embedding MLP, Wide&Deep, NeuralCF, DeepFM, 双塔模型等。- 使用 Tensorflow 训练并导出模型, 用 Tensorflow Serving 载入模型并提供服务, 实现用户界面的猜你喜欢模块。- 基于物品序列和 Spark 实现 Item2vec, 训练获得物品 Embedding, 实现电影界面的相似推荐模块。	
基于 Lucene 的中文信息搜索引擎	Java, Lucene, ICTCLAS, HTML/CSS, JSP, Tomcat
<ul style="list-style-type: none">- 基于 ICTCLAS 的中文分词器与 Lucene 实现简易搜索引擎, 语料来自于数十万条新浪新闻。- 实现搜索结果分页, 搜索历史记录与关键词高亮功能, 并进行前端美化。	
老司机 App - 基于因子分解机 FM 的汽车内容推荐系统	EigenTech, Hangzhou
本科暑期实习项目、导师: 曹斌博士	June 2018 – Sept. 2018
<ul style="list-style-type: none">- 实现基于 FM 的推荐系统, 完成客户需求, 相应的推荐场景为老司机汽车 App 的帖子推荐。- 从阿里云 OSS 下载日志, 进行数据预处理、特征构建, 训练并保存更新的 FM 预测模型。- 基于 Tensorflow Serving 进行服务部署, 相较于基准模型, 离线测试的 TopK 点击率指标提升了 30%。	

EXPERIENCE AND AWARD

KDDCUP 2019 上下文感知交通方式推荐 Top 2% 排名 45/1696	2019
SIGIR 2020 大宗商品股价预测, Top 6% 排名 30/457	2020