

# 杜利强

(+86) 186-0019-8951 · keris.du@gmail.com · 开发工程师 · <http://duliqiang.me/>

## 个人简介

目前在友信金服担任开发工程师，负责反欺诈方面的工作。在此之前，在新浪微博担任搜索算法工程师负责数据开发及排序模型迭代工作，在诺禾致源担任生物信息工程师负责开发短序列比对程序。此外，我是一个人工智能爱好者，在业余时间完成了优达机器学习工程师纳米学位，同时跟踪阅读相关书籍和论文。

## 教育背景

中国农业大学, 计算机科学与技术, 硕士研究生	2014/9 - 2017/1
中国科学技术大学, 电子科学与技术, 理学学士	2008/9 - 2012/6

## 技术能力

- 编程语言: C/C++, Python, Java, JavaScript, SQL, Shell
- 数据库: MySQL, MongoDB, Redis, JanusGraph
- 框架: Hadoop, Spark, TensorFlow, PyTorch
- 模型: LR, Svm, GBDT, xgboost, CNN, RNN, LSTM
- 人工智能: Machine Learning, Deep Learning, Reinforcement Learning
- 开发工具: Visual Studio Code, Vim, SVN, Git
- 操作系统: Linux/macOS

## 工作经历

北京友信科技有限公司, 开发工程师	2018/10-至今
<ul style="list-style-type: none"><li>• 基于用户关系网络进行公共服务电话挖掘，用户 tagging，统计特征挖掘，关联模式挖掘以及失联修复；</li><li>• 用户行为埋点数据分析；</li><li>• 迭代文本分类模型，对基于用户信息的搜索结果进行分类；</li><li>• 优化黑名单系统，提高了系统的可扩展性、可维护性、可读性和鲁棒性；</li><li>• 开发了一种埋点收集系统，可方便进行计数和耗时统计，如异常数、错误数统计，接口耗时统计等。</li></ul>	

新浪微博, 搜索算法工程师	2017/1-2018/9
<ul style="list-style-type: none"><li>• 负责微博搜索部门各业务线数据的接入、属性计算、分发和存储。</li><li>• 迭代排序模型改进站外文章的展现效果。</li><li>• 开发微服务应用于热门搜索和综合搜索，包括赞服务、转发服务和热门评论服务。</li></ul>	

北京诺禾致源生物信息技术有限公司, 生物信息开发工程师	2012/7-2013/10
<ul style="list-style-type: none"><li>• 开发短序列比对程序以检测 DNA 中的插入或删除。</li><li>• 利用 cuda 技术加速现有的比对程序。</li></ul>	

## 项目

用户行为埋点数据分析   北京友信科技有限公司	2019/12-至今
基于 APP 行为数据挖掘特征，构建逻辑回归模型，模型 auc=0.8。	

用户关系网络   北京友信科技有限公司	2018/10-至今
基于图数据库构建的用户关系网络，包含了 10 亿多个节点和 100 亿多条边。在关系网项目中我的工作主要包括： <ul style="list-style-type: none"><li>• 公共服务电话挖掘，用户 tagging，统计特征挖掘，关联模式挖掘以及失联修复；</li><li>• 进行性能调优，带来了十分高效的数据注入效率，达数千条边每秒；</li><li>• 开发关系网络可视化平台，支持用户关系直接交互。</li></ul>	

## 黑名单系统 | 北京友信科技有限公司

2018/10-至今

系统由若干个表构成，保存了满足特定规则的交易涉及的用户信息，如银行卡、注册手机号码、紧急联系人，设备号等。在每次进件时会调用黑名单 API，检查进件包含的信息是否命中黑名单，若命中，进件将被拒绝。其中我的工作包括：

- 开发新 API 服务，维护旧 API 服务；
- 系统重构，提高了系统的可扩展性、可维护性、可读性和鲁棒性；
- 日报和周报生成及推送。

## 网络筛查系统 | 北京友信科技有限公司

2018/10-至今

基于用户信息从百度搜索、贴吧、论坛等获取相关的搜索结果，然后对返回的文本结果进行分类。其中我的工作主要包括：

- 迭代优化文本分类模型；
- 日报和周报生成及推送。

## 优达机器学习工程师纳米学位

2018/10-2019/2

<https://graduation.udacity.com/confirm/G3M94KR7>

为了对机器学习的有一个全面的了解，我在业务时间完成了优达机器学习工程师纳米学位，系统性的学习了机器学习、深度学习和强化学习，并完成了相关项目：

- 预测波士顿房价；
- 创建客户细分；
- 狗品种分类；
- 基于 deep-Q 网络实现模拟器的飞控。

## 数据中间件 | 新浪微博

2017/1-2018/9

主要功能包括：数据获取，数据属性计算，数据分发和数据持久化。涉及微博搜索的各个业务线，包括热门搜索，综合搜索，微博故事和话题推荐。

## 文章搜索 | 新浪微博

2017/1-2018/9

通过包含长微博以及站点外文章以丰富微博搜索的结果。其中我的工作主要包括：

- 迭代排序模型，改进站外文章展现效果，提高召回率和准确率；
- 维护和改进爬虫系统，确保内容的实时性。

## 出版物

- 
- Liqiang Du, Hong Chen, Shuli Mei, and Qing Wang. 2016. Real-time human action recognition using individual body part locations and local joints structure. In *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry - Volume 1* (VRCAI '16), Vol. 1. ACM, New York, NY, USA, 293-298. DOI: <https://doi.org/10.1145/3013971.3013974>