

## 姓名：王 帅

性别：男

出生年月：1989.01

居住地：浙江省杭州市

邮箱：wangliaofan1988@163.com

电话：18501252287

### 工作经历

- 2024.05-至今 负责人 算法交易创业
- 2024.02-2024.05 架构师 百度（北京）（安全部业务安全风险负责人）
- 2020.05-2023.07 算法专家 字节跳动
- 2019.02-2020.05 技术专家 阿里巴巴（钉钉安全部业务安全风险负责人）
- 2015.03-2019.02 算法专家 蘑菇街（风控算法负责人）
- 2013.04-2015.02 项目经理 中国空间技术研究院 503 系统部

### 教育经历

- 2010.09-2013.04 保送工学硕士 华中科技大学多谱信息处理国家级重点实验室  
(图像识别与人工智能研究所) (Top 10%)
- 2006.09-2010.07 工学学士 哈尔滨工业大学航天学院 (Top 5%)

### 技术背景

- 英语：CET4(603), CET6(570), TOEFL-iBT(95), 能熟练使用英语口语与书面英语；
- 编程：熟悉 Python, Scala, Java, 学习使用过 C、C++、C#；
- 机器学习：熟悉线性及 Logistic 回归分类、SVM、混合高斯模型、DNN、CNN, LSTM, Auto Encoder 等统计机器学习/深度学习算法，熟悉 PyTorch 等深度学习框架；
- 大数据工具：熟悉 SPARK, Prediction IO, HIVE, HADOOP, HBASE, TACHYON, ZEPPELIN, CAFFE, KERAS 等利器。

### 业界分享

- 2017.07.21, 51CTO 全球创新技术峰会特邀讲师，“智+应用”专场分享《机器学习在蘑菇街风控业务中的技术实践》，链接：  
<http://wot.51cto.com/act/2017/innovation/page/agent>。
- 2017.09.09, VSRC 城市沙龙上海站特邀嘉宾，分享《机器学习在蘑菇街业务安全中的应用实践》，链接：[https://mp.weixin.qq.com/s/7t5zMuAscs\\_l8f1poMrjVA](https://mp.weixin.qq.com/s/7t5zMuAscs_l8f1poMrjVA)。

### 专业积累

- 《机器学习互联网业务安全实践》，第一作者，电子工业出版社，2019.08 月出版；
- GITHUB: <https://github.com/xuanyuansen>
- “A foundational framework designed for intelligent satellite-based applications incorporating data mining techniques”，第一作者，被 GLAC2014 录取，随院代表团代表单位赴法国巴黎教科文组织总部参会，并作现场口头英文汇报与海报

展示;

- “Optimizing the number of clusters with global complexity criterion”, 第一作者, 被 ICCEE2012 录取;
- “Simulation of uneven blurred images impacted by the aero-optical effect of high-speed flow field”, 第一学生作者, Proc. SPIE 8002, 800209 (2011), EI 检索。

## 项目经历

### ➤ 2024.02-2024.05 百度安全部架构师

作为安全部的架构师, 负责整合百度安全部内部的业务安全风险能力, 包括活动风控、设备风控、流量风控、网页安全、文本监测挖掘等等, 再面向百度内部各个业务线赋能, 提升业务线与黑灰产对抗的能力, 保证业务安全风险收敛。

### ➤ 2020.05-2023.07 字节跳动算法专家

#### 网页内容安全风险 (基于浏览器内核)

**2021.06-2023.07** 基于自研浏览器内核获取的非用户隐私行为数据, 包括网页浏览跳转关系等等, 叠加网页自身的文本图片等数据, 通过 TextGcn 等机器学习算法识别黄赌毒, 挂马、钓鱼等不安全链接和场景, 提升用户的安全浏览体验, 期间创新性地提出了基于网站文本和网页跳转关系构建异构图的风险识别算法, 并且从底层重构了算法的实现, 提升算法效率, 优化内存空间使用, 有效解决了多语言非法站点识别的问题, 提升垃圾网页召回率 **20%** 以上, 并基于此技术方案沉淀技术专利一项, 已经通过国家专利局审批。

#### 飞书业务安全风险

**2020.05-2021.06** 针对飞书面临的安全问题, 从账户视角全面梳理飞书安全现状, 包括 IM、日历、会议等子业务场景, 明确整体的安全需求以及治理路径; 同时为了解决飞书账户安全相关的问题, 与集团安全团队沟通与协作, 分析问题、拆解需求、沉淀和制定安全解决方案, 同时带领算法同学完成数据挖掘和机器学习建模, 识别不同场景下的问题用户、租户以及其作弊路径, 解决飞书账户侧的安全问题, 例如邀请短信被滥用、散播违禁信息、垃圾注册登录等等。

### ➤ 2019.02-2020.05 钉钉安全部业务安全算法及产品负责人 (4 人团队)

#### ToB 业务安全风险方案及实施

**2019.02-2020.05** 针对钉钉业务安全风险尚缺乏体系化产品、数据沉淀无闭环和风险运营与管理缺乏系统和产品层面完成支撑的情况, 全面梳理了旧系统存在的问题, 包括数据依赖分散、名单关键词管理耦合较高和基础原子能力不足等等问题, 针对上述问题, 并且在同中台能力联动的基础上, 提出了钉钉业务安全系统 (DTASS) 的方案, 包括业务安全中心, 数据中心, 审核处罚中心, 能力中心和监控中心, 并带领工程开发同学完成了数据中心和审核处罚中心的实现, 打通了整个业务安全风险系统的数据流转, 解决了风险可运营、可管理和可追溯等问题。

### ➤ 2017.03-2019.02 蘑菇街风控算法团队负责人 (8 人团队)

#### 算法团队组建与管理

- ✧ **算法团队组建**, 通过社招快速组建风控算法团队, 组件团队后建立了规范的算法论文分享及模型评审制度, 同时根据公司业务需求进行算法策略调研与研发方向规划, 过程中指导团队算法同学改进所负责业务的算法, 有效保证了风控业务的稳定及后续工

---

作的开展，并支撑公司的时尚业务，例如图片质量评价、时尚度评价等等；

- ✧ **算法研发方向**，根据全站风控业务的性质，以及用户的行为链路将风控业务做如下整合：用户风险控制（**用户风控**），内容风险控制（**内容风控**），交易风险控制（**交易风控**）和金融风险控制（**金融风控**），实现结对编程，解决了业务中单点人力问题；
- ✧ **团队内部人才培养**，招聘面试 50 人以上，辅导实习生及新同学，先后带过 10 人以上；
- ✧ **积极拥抱开源生态**，引入有价值的组件，不重复造轮子，在组内引入 Prediction IO, Zeppelin, Tachyon, Keras, Theano, Tensorflow 等工具，提升算法服务业务的能力。

➤ **2015.03-2019.02 蘑菇街电商风控算法负责人（从 0 到 1 搭建了电商风控算法架构）**  
**项目一：电商风控**

**第一阶段，统计学习模型初建**，引入了统计机器学习算法模型大幅提升了系统的召回，从 5% 到 20% 左右，过程中挖掘引入了浏览，转化，搜索等强业务相关的行为特征，并提出了分时特征的新颖思路（扩充了特征的维数，提高了模型的效果）。

**第二阶段，统计学习模型实时化**，独立调研，选用了 GITHUB 上高星项目 Prediction IO 进行二次开发，独立搭建了实时计算的架构，后续又对该系统进行改进来降低 RT 及解决单点问题等，包括存储从 HBASE 换成 REDIS，通过 ES 索引组建 PIO 集群等。

**第三阶段，深度学习模型及图模型引入**，经过数月左右的探索，利用特征主体在 APP 内的行为打点信息，建立了 LSTM 融合模型，再次提高反作弊系统的召回 20% 以上；与此同时，为了打击精刷及团伙作弊，研究了图挖掘算法在反作弊中的应用，经过调研采用了 Louvain、Fraudar 等图挖掘算法来识别作弊团伙，并改进了部分算法的开源实现。

**第四阶段，深度学习模型实时化**，经过前三个阶段的研发，交易风控系统已经较为完善，能够满足公司整体业务需求，同时为反套利，优惠券防刷，虚假注册等提供大量数据支撑，是电商风控最核心的业务。为了进一步提升风控策略的时效性，基于 tornado web server 构建了深度学习模型的实时化服务架构。

**项目二：金融信贷风控**

**2016.08~2018.09**，作为信贷风控算法负责人，负责蘑菇街信用支付产品白付美的风险识别算法相关工作，主要负责信用评分模型的构建，分别是贷前评分模型、贷中评分模型和贷后评分模型，对应的是贷前准入模型、贷中监控模型和催收预警模型，结合风控工作经验，在信贷建模的工作过程中快速学习消费信贷评分建模的基本定义和方法，除了使用传统的评分卡以外，积极引入机器学习模型，并结合电商交易风控的特征工程，一个月时间内使 M0 逾期率下降 20% 以上，后续 M1 逾期率、M2 逾期率均有较大幅度下降。作为算法项目负责人，前期负责算法的实现，后期负责算法工程化及模型的部署、上线与优化，模型上线后通过积累用户行为等数据，不断优化贷前准入、贷中监控和催收预警模型。

**项目三：内容风控等**

**2016.07~2018.12**，指导相关算法同学在 UGC/IM 反垃圾业务中进行新策略调研，使用了 word2vec、LSTM、CNN 等算法并进行融合，提升了算法在业务中的表现，消息垃圾率降低 30% 以上，并且成功复用 Prediction IO 集群框架，实现了独立的算法服务。

**2015.11~2017.12**，为了配合推荐部门的图墙排序升级策略，从无到有建立了点击反作弊策略体系，过程中借鉴交易风控的研发思路，采用了规则和模型相结合的策略体系，

---

并根据用户的点击轨迹建立了全站商品的点击图，直接在该有向图上挖掘点击最异常的商品。

### 风控工作感悟

- 工程与策略兼备，理论过关，数据敏感，开放心态，自驱负责