

范叶亮

✉ leo-van@hotmail.com · ☎ (+86) 186-3159-3756 · 🌐 <https://leovan.me>

🎓 教育背景

河北工业大学 硕士 (保送) 信息管理 GPA: 85/100 Top 2% 2012 – 2015

- 二等奖学金 (Top 5%), 三好学生 (Top 5%), 优秀学生干部 (Top 5%)
- 河北省普通高等学校优秀毕业研究生 (Top 2%)
- “创青春” 全国大学生创业大赛河北省一等奖

河北工业大学 学士 工商管理 GPA: 87/100 Top 2% 2008 – 2012

- 一等奖学金 (Top 2%), 三好学生 (Top 5%), 优秀学生干部 (Top 5%)
- 经济管理学院学生会主席

👤 工作经历

京东金融 2015 年 03 月 – 至今

👤 算法工程师

Daat (金融知识图谱项目) 项目负责人 2018 年 04 月 – 至今

- 京东金融知识图谱技术白皮书: 调研整个知识图谱技术栈, 包括各细分领域的相关模型, 数据, 评测和开源工具, 汇总整理最终调研报告。
- 自动化敏感信息识别: 对于入库到数据仓库中的数据进行自动敏感信息识别, 辅助数据加密策略实施。根据数据的元信息 (例如: 表名, 表注释, 字段名, 字段注释等) 和值信息 (即字段存储的数据值), 利用 Wide & Deep 网络构建识别模型。提取传统特征构建 Wide 网络, 针对文本特征, 利用 Char Embedding + CNN 构建 Deep 网络, 模型测试数据的 F1-Score 为 95%+。

全视之眼 (中文地址分析项目) 项目负责人 2015 年 03 月 – 2018 年 04 月

- 地址分析算法引擎: 设计开发地址分词, 分类, 完整度, POI 识别, 相似度等算法 (准确率 90%+)。
- 地址画像系统: 依托地址分析算法引擎, 设计地址 POI 画像相关指标体系, 存储和服务框架。对内, 服务于线下支付等场景下的营销和运营, 平均提升用户转化率 30%+; 对外, 接入京东征信和京东金融稻田等平台, 提供精确查询, 模糊查询和经纬度范围查询等多种查询方式, 实现金融科技的对外输出。
- 授信和反欺诈模型: 依据地址画像构建授信策略, 辅筛选和授信用户超过 1000 万人; 依托地址分析算法和地址画像, 构建中介识别等反欺诈模型, 识别套现订单金额约 20 万元/日。
- 全视之眼项目斩获 2017 年度京东集团京芽杯创新大赛 “创新种子奖”, 排名 20/378。
- 企业地址画像系统: 构建企业、地址和用户关系, 设计企业地址画像相关指标体系, 服务于京东蓝鲸征信系统 (<https://icredit.jd.com/>)。
- 农村金融小站选址系统: 根据地址画像系统, 结合农村金融相关业务, 构建基于地址画像的农村金融小站选址模型和系统。

用户行为分析项目 算法设计和开发 2017 年 10 月 – 2017 年 12 月

- 设计开发了一种 Behavior2Vec 用户行为表示方法, 基于层次聚类和深度搜索, 构建了一个用于识别用户异常行为的混合模型。在校园白条中介激活和账号盗用问题上, 相比传统的 Bag of Words 和 N-GRAM 等表示方法, 在保证准确率 80%+ 的前提下, 提高异常用户识别数量 3+ 倍。

动产融资项目 算法设计和开发

2015 年 03 月 - 2015 年 10 月

- 设计开发一种基于 Bass 扩散模型, 优化的时间序列相似度和聚类的混合产品生命周期识别模型, 尾货识别准确率 95%+, 辅助贷款方商品质押准入决策和质押率的制定。
- 设计开发不同数据源的商品信息融合模型, 相同商品识别准确率 90%+, 配合 ElasticSearch 开发商品模糊匹配查询服务, 为商品质押提供精准的价格等相关信息。

方正国际软件有限公司

2013 年 11 月 - 2014 年 02 月

👤 算法工程师, 实习

智慧公安 (山海易绘 EzMap) 算法设计和开发

2013 年 11 月 - 2014 年 02 月

- 设计开发基于 EMD 和 SVR 的时间序列预测算法, 用于犯罪分析和智慧城市中的时序预测。
- 设计开发基于 k-Prototypes 的串并案算法设计, 用于犯罪分析中的串并案分析。
- 开发公安元数据导入工具和基于 XML 的警用地理信息系统数据库升级工具。

🔧 专业技能

- 开发语言: R: ★★★★★ Python: ★★★☆☆ Java: ★★★★★ Lisp: ★★☆☆☆ SQL: ★★★★★
- 平台框架: Linux: ★★★★★ TensorFlow: ★★★★★ Qt: ★★☆☆☆
- 语言能力: CET-6: 518, 熟练运用英语进行口语交流, 阅读资料, 撰写文档

📄 研究成果

- Li Jie, **Fan Yeliang***, et al. (2013) An Improved Forecasting Algorithm for Spare Parts of Short Life Cycle Products Based on EMD-SVM. In *Information Science and Cloud Computing Companion (ISCC-C), 2013 International Conference on*, (pp. 722-727). IEEE. 2013
- 一种中文地址分词方法及系统 (专利, CN105159949A, 第一发明人) 2015
- 一种产品库存预测方法及装置 (专利, CN106056239A, 第一发明人) 2016
- 一种产品生命周期的识别方法和装置 (专利, CN106408217A, 第一发明人) 2017
- 一种计算地址相似度的方法和装置 (专利, CN107239442A, 第一发明人) 2017
- Zhou Feng, Li Huafei, Zhan Lizhang, **Fan Yeliang**, et al. (2018) A Novel Ensemble Strategy Combining Gradient Boosted Decision Trees and Factorization Machine Based Neural Network for Clicks Prediction. In *Big Data and Artificial Intelligence, 2018 International Conference on*. 2018
- 一种基于信息抽取的数据仓库宽表自动化构建方法 (专利, 申请中, 第一发明人) 2018
- 一种敏感信息的智能识别方法 (专利, 申请中, 第一发明人) 2018

📁 开源项目

- Github: <https://github.com/leovan>
- R 语言数据科学导论: 一份以 R 语言为基础的数据科学入门教程, 🌟 Stars: 6。
- Sci-Hub EVA: 一个跨平台的 Sci-Hub 界面化应用, 🌟 Stars: 226。

👤 自我评价

- 适应能力, 学习能力强, 喜欢尝试和挑战
- 有敏锐的洞察力, 深刻的分析力和高度的责任心
- 平等的换位思考, 舒服的与人沟通, 具有团队合作精神