

个人信息

- 手机：18573101656 / Email：xbtlin@gmail.com
- 林轩 / 男 / 1990
- 硕士 / 国防科技大学计算机科学与技术（2012.09 ~ 2014.12）
- 本科 / 东北大学计算机科学与技术（2008.09 ~ 2012.06）
- 工作年限：2年
- 期望职位：算法/数据挖掘工程师



工作经历

美团 算法开发岗（2015年7月 ~ 至今）

餐饮前台品类体系

从0到1建设餐饮前台品类体系。我独立完成餐饮前台品类建设的所有工作。第1步是利用后台品类和前台品类映射关系为一部分商家打上前台品类。第2步是利用1产生的标注数据，用模型训练的方法为剩下的商家打上前台品类。第2步问题可以抽象为文本多分类问题，自己独立开发的模型在多分类(14类)问题正确率达到95%，覆盖率90%。将商家的三级品类覆盖率从48%提高到80%，该品类体系已被点评APP、广告、搜索部门使用。用到朴素贝叶斯原理、Spark等。

餐饮商家画像

餐饮商家画像。我负责特色类标签的生产，整个餐饮商家画像数据格式定义和对外提供。生产过程是运营给出策略，我用Spark从源Hive中读取数据，根据策略进行处理后写回目标Hive表，我定义了餐饮商家画像统一的数据格式。并提供给点评、美团APP、广告等其他部门。

商家不接待项目

根据商家的线上数据预测不接待商家，维护美团形象，挽回公司损失。我参与数据分析，特征提取流程搭建，算法选型，特征工程等工作。此项目的难点在于线下不接待行为反映到线上的信息有限。我通过选取不接待商家对应的强特征，模型AUC从0.66提高到0.84。用到的技术有GBDT、Hive、Spark、awk、Java、Maven。

天天平价项目

该项目监控团购同款商品美团价格和竞对价格，保证美团价格有竞争力。我参与数据解析、同方案匹配，负责方案对比、审核工作台和智能审核开发。此项目的难点是如何将千万级数量的美团方案和竞对方案一一匹配。之前采用规则匹配的正确率是70%，我们采用决策树算法将正确率提高到93%。此项目第2个技术点是将

匹配的方案对和处理结果存档，该方案对再次进来时，结合当前线上数据自动处理。第3个技术点是智能审核，即通过抓取网页数据并模拟人工审核的过程，提高审核速度，解放人力。我参与技术点1，负责2、3。用到的技术有Hive、Spark、Weka、Java、Spring、Mybatis、Maven。

买单项目

建立一套系统，及时发现新美大缺失门店和买单价格劣势，通知相应销售人员，并接受销售和总部的反馈。独立完成后台开发的全部工作，包括需求评审、数据和流程设计、API设计、系统开发、测试、前后端联调。

酒泉卫星发射中心 软件开发岗（2014年11月～2015年6月）

军用信息系统项目

实现故障在线报修模块，提高报修效率。独立完成该模块的前后端设计、开发和测试。

其他相关项目

分布式容错存储系统 (2013年8月～2014年10月)

国家863计划项目“云计算关键技术与系统”的一部分。基于Hadoop的文件系统HDFS开发出低冗余高可靠性的分布式容错存储系统RDFS。我是唯一的设计与开发人员。提出多节点修复算法GRC和分组数据互换修复技术GDER，并实现，已发表论文。

论文 / 林轩,王意洁,裴晓强,许方亮,符永铨. GRC：一种适用于多节点失效的高容错低修复成本纠删码[C]. 第二届CCF大数据学术会议. (已录用) 并推荐到计算机研究与发展(增刊). (已录用)

学生网上作业交付及文档相似性检测系统（2012年2月～2012年6月）

网站有用户注册、登录、作业上传、作业下载和文档相似性检测功能。我独立完成全部设计与实现工作。用向量空间模型(VSM)作为文档相似性检度量，用Lucene组件分词，用Hash表去重形成词典，用Java实现业务逻辑。抄袭阈值通过训练多组样本得到。

相关经历

- 国际大学生数学建模竞赛 (MCM/ICM) 国际二等奖，负责算法和MATLAB实现
- 全国大学生数学建模竞赛国家二等奖，负责算法和MATLAB实现
- 中国机器人大赛暨RoboCup公开赛冠军，负责机器人足球算法和实现

技能清单

以下为我熟练使用的技能

- 熟悉的机器学习算法：线性回归/朴素贝叶斯/决策树/SVM等
- 大数据处理：Spark Java API/Hadoop
- 主语言：Java 脚本语言：Python/Bash
- 框架：Spring/Mybatis
- 依赖管理：Maven 版本管理：Git
- 数据库相关：MySQL/Hive