

# 崔 晨阳

☎ (+86) 188 1146 4201  
✉ ccyjava@gmail.com

## 教育背景

- 2012.09 – 2015.07 硕士在读, 北京大学, 信息科学技术学院, 计算机科学与技术系.
- 2008.09 – 2012.07 工学学士, 安徽大学, 计算机科学与技术学院, 计算机科学与技术系.

## 项目及科研经历

- 2013 实习, *Hulu*北京软件开发公司 软件开发实习生.  
用LDA对精准广告投放系统的效果优化
- 该系统是通过用户的访问行为, 来预测用户的年龄, 性别等信息, 以此来投放广告。我们以注册用户 (含有年龄, 性别等信息) 为训练集, 对匿名用户进行预测。
  - 我的工作是用logistic回归实现模型, 并把用户的行为特征, 用LDA 的方法训练出主题, 以此进行特征降维并训练模型, 最终AUC提高了2%
- 挖掘视频涉及的情节-用自然语言处理的方法
- 通过对视频介绍文本, 评论内容进行文本挖掘, 获取视频设计的情节。
  - 我们用关键词抽取的方法抽取出一系列关键词, 统计重要度, 定义出一系列的情节, 然后给每个show和每个情节打分, 标识其相关度。
- 2010 科研, 实验室项目, 基于Hadoop平台的SVD算法研究.
- 负责搭建Hadoop平台, 并在平台上编写相似矩阵求解程序
  - 以第二作者身份在ICCSNT发表一篇论文
- 2013 科研, 实验室项目, 代码注释推荐系统.
- 通过对开源软件库的挖掘, 分析待标注代码的结构, 用协同过滤和关键词生成的方法推荐代码注释
  - 负责整个项目的调研设计开发, 目前该项目正在进行中
- 2012 个人项目, 基于HMM的无指导分词系统.
- 2012 个人项目, 基于向量空间模型的搜索引擎.

## 荣誉

- 2011 铜奖, 36届ACM/ICPC国际大学生程序设计竞赛亚洲区赛 (大连) 铜奖.

## 语言技能

英语 CET6 能无障碍阅读计算机英文材料, 活跃于stackoverflow等国外论坛

## 专业技能

- 掌握Java, Python, 熟悉C, Shell脚本, 常见的web技术, 常见的数据库, Hadoop, Linux
- 了解并非常感兴趣自然语言处理、机器学习与信息检索相关知识技术