

徐彩旭

✉ csxucaixu@gmail.com · ☎ (+86) 188-9697-0393

🔑 基本信息

性 别: 男

出生年月: 1993.03.12

籍 贯: 江苏盐城

政治面貌: 中共党员

学 历: 硕士在读

研究方向: 大数据研究与应用

🎓 教育背景

苏州大学, 江苏, 苏州

2015-2018

在读硕士 计算机技术

专业排名前 15%

东南大学成贤学院, 江苏, 南京

2011-2015

学士 计算机科学与技术

专业排名前 3%

👨‍💻 实习/项目经历

百度金融黑名单系统

2017.06-至今

项目属性: 暑假实习项目 (百度金融部)

项目角色: 项目贡献者

主要工具: Python2.7, Linux, MySql, Hadoop, MapReduce, bs4

利用 Hadoop、MapReduce 统计百度内部与金融相关的数据; 爬取外部失信数据; 内部数据结合外部数据提高黑名单系统的 Precision、Recall 评估指标。

- 数据的统计;
- 数据采集、数据处理及数据管理;
- 黑名单系统 Precision 和 Recall 的评估;

基于位置社交网络的用户关系强度预测

2016.01-2017.02

项目属性: 机器学习项目

项目角色: 独立研发完成

主要工具: Python3, Linux, C++, MySql, Scikit-Learn, XGBoost

利用斯坦福大学开源 SNAP 数据 (近亿条签到数据, 几百万用户) 构建分类模型, 模型预测任意用户对关系强度。在召回率相同的情况下, 精确度比最好的模型提高 10%。

- 平台搭建 (Linux 系统 +MySql+Python3)
- 数据处理与管理
- 代码编写 (多维度抽取特征、构建多视角分类器模型、模型调测)

轨迹相似度计算系统

2016.03-2016.08

项目属性: 比赛项目 (中国软件杯)

项目角色: 项目主要负责人

主要工具: Java, Python3, JS, UML, Linux

利用全国近三千万轨迹数据搭建相似度计算平台, 平台可以快速响应请求: 计算任意两条轨迹的相似度; 任意画出一条轨迹, 从数据库中快速检索出与当前轨迹最相似的轨迹, 并以地图显示的形式反馈给用户。

- 数据的聚类处理与存储;

- 该项目采用 BS 模式, 前台采用 JSP+CSS+DIV;
- 后台 Java 实现, 构建基于聚类簇的数据结构 (以响应快速读取操作); 定义相似度并设计高效算法;

公司信誉度评估分类系统

2015.06-2015.11

项目属性: 实习项目 (苏州汇誉通数据科技有限公司)

项目角色: 项目贡献者

主要工具: Python3, Linux, Scikit-Learn

到各大信息公开网站采集公司的异常记录、公司基本属性等数据; 根据数据构造特征; 对部分公司进行聚类标记, 构建随机森林回归预测模型, 模型输出公司所属类别。

- 数据的采集与清洗;
- 数据存储与管理;
- 特征工程与模型调测;

IT 技能

- 编程语言: Python、C++、C、Shell、Java;
- 数据库: 熟悉 MySQL、Oracle;

语言水平

- 中文 (普通话): 母语
- 英语: 听、说良好, 读、写流利, CET-4

社会活动

计算机科学与技术学院学工助理, 东南大学成贤学院	2013
院 DV 工作室技术部副部长	2012
院科学与技术协会副部长	2012

★ 获奖情况

基于时空数据的用户社交关系强度预测软件 (计算机软件著作权)	2016.05
苏州大学三等奖学金	2015-2016
东南大学成贤学院优秀毕业生	2015.06
东南大学成贤学院三好学生	2013-2014
东南大学成贤学院综合一等奖学金	2013-2014
全国计算机二级 C 语言能力等级证书	2013
江苏省高等学校计算机二级 VFP 能力等级证书	2013
院级二等奖学金	2012-2013
院级三等奖学金	2012-2013
院级三等奖学金	2011-2012

☆ 学科竞赛

DataCastle: 交通线路通达时间预测, 26 名	2017.01
第五届“中国软件杯”大学生软件设计大赛决赛三等奖	2016.08
东南大学第五届程序设计竞赛二等奖	2014.05
院级网页设计制作大赛一等奖	2011.12