**Mr. Wang, Fengwei**

Room 402, East Building of Cao Guangbiao Building, Yuquan Campus, Zhejiang University, Hangzhou, Zhejiang Province, 310027, China.

Phone: (+86) 15967135691 E-mail: [cf.wfwei@gmail.com](mailto:cf.wfwei@gmail.com) GitHub: [github.com/wfwei](https://github.com/wfwei)

**Education**

2011-Present **Master of Science candidate** in Computer Science and Technology, **Zhejiang University** (ZJU) (GPA: 3.68/4)

2006-2011 **Bachelor of Engineering** in Information Management And Information System, **East China University Of Science And Technology**(ECUST)(GPA: 3.39/4)

**Basic Abilities**

Programming Proficient with commonly used data structure and algorithms; Proficient with Java, familiar with C(++), Python and linux scripts.

Professional Solid foundation of Data Mining and Machine Learning knowledge

English CET4（586）CET6（564）

**Professional Experience**

2013.3-2013.6 **Chief Programmer** Snapshot Crawler. It is a web crawler which can snapshot a whole website. It works well with Chinese character encodings and has excellent fault tolerance. I also add a URL sampling module which can filter structurally similar URLs, which use similarity hashing skills to improve efficiency

2012.8-Present **Chief Programmer** News Classification and Aggregation Application. It regularly crawls long media news and weibo(like twitter) news, then classifies them into different topics. Currently I have completed on the topic of 'American Drama' with a website and a chrome extension. I will continue to work on this amazing work and make it my graduation project.

2013.5-2013.7 **Chief Programmer** “You Pull I Push”Recommendation System. Simply drop local files to the search box, and get personalized recommendations instantly, this effectively solved the famous cold-start problem. We use collaborative filtering and tag based methods to recommend movies, books, music and etc.

2012.1-2013.6  **Chief Programmer** Information Accessibility Project. According to Website Accessibility Standards, Design automatic check methods and online check platform which is based on J2EE.

**Scholarships & Honors**

2011-2013 Grand Prize of Youdao Problem , Second Place of Baidu ML Competition, Best Innovation Award of Insigma IN Plan

2007-2011 National Scholarship & merit student , Excellent graduate of Shanghai

**Self-Evaluation**

I'm a coder who believes that technology changes life, and I love creative thinking and make a difference; I'm also a sports lover, I especially like swimming, mounting and tai-chi

**王峰伟**

浙江省杭州市浙江大学玉泉校区曹光彪大楼东楼402室 310027

电话:15967135691 邮件:[cf.wfwei@gmail.com](mailto:cf.wfwei@gmail.com) GitHub: github.com/wfwei

**教育背景**

浙江大学硕士 计算机科学与技术(平均绩点：3.68/4 ) 2011-2014

主要课程 机器学习导论，数据挖掘，机器学习，计算理论等

华东理工学士 信息管理与信息系统(平均绩点：3.39/4) 2007-2011

主要课程 数据结构和算法，数据库概念与技术，C(Java)程序设计，运筹学等

**基本技能**

编程基础 熟悉常用数据结构和算法；熟悉Java，熟练使用Python，C，以及Linux常用命令

专业技能 拥有机器学习和数据挖掘的扎实基础，熟悉爬虫，分词器等工具

英语水平 四级-586，六级-564，具备较好的听说读写能力

**项目经历**

主要开发者 网站镜像抽样爬虫 2013.4-2013.6

修改开源爬虫crawler4j的源码，提供镜像网站的功能，同时改善对中文编码和不规范网点的容错能力；根据url结构相似性，进行网页抽样爬取，同时使用相似性哈希加快速度。该爬虫已在实验室无障碍平台上使用

项目负责人 新闻分类和聚合器 2012.6-2013.9

对微博短消息和新闻的分类聚合研究：该应用通过不断爬取新闻和微博，对其分类，去重，整合成主题相关的类别；已完成对美剧相关的微博，新闻分类聚合工作，并作为i看美剧的新闻模块上线；目前在继续研究微博的主题分类研究，作为自己的毕设课题。

主要开发者 “你拉我推”搜索推荐应用 2013.5-2013.7

颠覆人们对搜索框的认识，用户将本地文件拖拽到搜索框，就能得到个性化的推荐结果，有效解决冷启动问题，同时生成可理解性的推荐理由。推荐内容包括电影，书籍，音乐，文档等，算法包括协同过滤，标签推荐等

主要开发者 网站无障碍检测和改造 2012.4-2013.6

根据网站无障碍标准，设计自动化检测方法及在线监测平台，设计网站镜像爬虫抓去网

站，根据网页拓扑关系以网页DOM结构，使用层次聚类的方法进行网页抽样。

**奖励情况**

有道难题特等奖，百度机器知我心二等奖，浙大网新“IN 计划”最佳创新奖 2012-2013

国家奖学金，校优秀学生，上海市优秀毕业生 2007-2011

**个人评价**

我是一个坚信技术改变生活的码农，爱思考，爱创新，爱动手实践；我也是一个热爱运动的人，爱游泳，爱爬山，爱太极。