吴俊

电话：156-5270-3181   E-mail：wujunhust@163.com

教育背景

2012.9-2015.7	硕士	中国科学院计算技术研究所	计算机应用技术（免试推荐）	**Top 15%**
2008.9-2012.7	学士	华中科技大学	电子信息工程（GPA:90 排名：6/210）	**Top 3%**

实习经历

2014.6-2014.7	百度 LBS 研发部 数据清洗方向	百度实习项目	核心设计/开发人员

◆ 项目介绍：百度地图的数据来源很多，有些来源并不可靠，如果进入后续处理流程会对处理结果进行干扰，因此需要对来源进行准入拦截，将不合法的数据不进入后续的流程。

◆ 个人职责：

- **负责接手拦截策略**，现有策略先对数据进行继承校验，以及黑名单过滤，对名称，坐标，地址，城市，电话进行单字段交验以及对坐标，地址，城市进行多字段校验，对校验通过的数据进行后续处理。
- **负责升级拦截策略**，针对现有的策略继承校验可能存在的隐患，以及黑名单过滤编写的不够健壮的问题，以增加白名单的方式进行升级，对部分单字段以及多字段校验根据调研进行升级。
- **负责编写数据脚本**，编写 awk 脚本，通过获取不同文件的字段，将多个文件的数据整合到一个文件中，作为分析依据。

◆ 研究成果：升级策略有效提升了继承校验的合理性，将继承校验的错误从 20%下降到 10%，其他策略的升级有效提高代码的健壮性。

2013.3-2013.7	百度推广 无线广告主推广站数据统计	百度实习项目	核心设计/开发人员

◆ 项目介绍：百度推广广告主数量日益增多，随着无线站点的普及，很多广告主只有 PC 站，缺少相应的无线站，需要提供 PC 站转换为无线站的框架，以及对无线站访问及转换情况进行统计并展现，实现即时获取日常消费信息。

◆ 个人职责：

- **负责抓取线上日志信息**，下载线上日志处理文件，判断文件合法性有效性完整性。
- **负责统计更新数据脚本设计**，将广告主名片的 pv，uv 及转化量进行统计，将统计结果更新到线上数据库，实现每日定时统计。
- **负责改进脚本**，对脚本运行时间较慢的问题，将日志拆成 90 个子文件并行处理，及增加数据库索引对的方法，提高运行速率。
- **负责更新优惠券物料算法设计**，针对无线优惠券的需求添加了一个优惠券数据表，用于存放优惠券物料信息，编写更新物料脚本定时在线上抓取最新物料信息，比较线上数据库信息，更新物料。

◆ 研究成果：脚本已经上线用于对无线名片广告主的数据统计，每天定时统计用户消费转换情况，运行时间从 3000 秒下降到 37 秒。

项目经历

2012.9-2012.12	人脸确认系统实现	机器学习项目	核心设计/开发人员

◆ 项目介绍：目前人脸检索在受控环境下容易受光线，及脸部旋转影响，为克服上述问题，我们提取图像 LBP 特征，运用 Adaboost 方法进行特征选择，实现对给定的数量有限的人脸图像集进行训练，确认被搜索引擎返回的图像中是否存在我们训练的人脸图像。

◆ 个人职责：

- **调研特征提取算法**，考虑图像的特点，选用 LBP 特征进行特征选择，有效克服光线明暗及旋转的影响。
- **负责特征选择算法设计**，对提取到的 LBP 特征，利用 Adaboost 方法特征选择，得到一组弱分类器组成强分类器用于分类。
- **负责改进特征选择算法**，在传统的 Adaboost 算法基础上，改进算法，将计算分类器错误率时间从 $O(N)$，下降到 $O(1)$。

◆ 研究成果：实现人脸确认系统，准确率 85%以上，训练一个弱分类器的时间从 40 秒下降到 0.5 秒。

获奖情况

◆ **国家励志奖学金**（前 **5%**，2009、2011）、**校学习优秀奖学金**（前 **10%**，2009、2010）

◆ **校三好学生**（前 **10%**，2009、2014）、**校优秀毕业生**（2012）、**优秀团员**（2009）

个人技能

◆ 熟悉基本数据结构和算法，逻辑思维能力较强。

◆ 熟悉 c/c++开发，熟悉 PHP，shell 脚本语言，了解 mysql，Verilog。

◆ 软考中级软件设计师，计算机等级考试二级 C 语言，三级网络，四级网络工程师

社会实践

2008.9-2012.6	**学习委员**，班级获优良学风班（2009/2011）称号。校记者团外联部**副部长**。
2009.10-2011.3	华中科技大学导游**志愿者**，曾参加小学支教活动和华中大开放日等公益活动。

自我评价

◆ 稳重，踏实，能够潜心下来做事情。

◆ 在大学培养了自己对新的知识，在最短的时间学会并且学好的能力。

◆ 高中班主任曾这样评价我：一个有理想远大的人，一个勤奋好问的人，一个谦虚有潜力的人。

Wu Jun

Mobil: (+86)156-5270-3181   E-mail: wujunhust@163.com

Education

2012.9-2015.7	Master	Institute Of Computing Technology Chinese Academy Of Sciences	Computer Science	**Top15%**
2008.9-2012.7	Bachelor	Huazhong University of Science and Technology(GPA:90)	Electronics Information	**Top3%**

Internship Experience

2014.6-2014.8   Data clean group   **Baidu LBS Internship Project**   **Designer/Developer**

◆ Project Description：Baidu map data has many sources , some sources are not reliable.If they go into the subsequent processing ,it will have a bad influence on the results, so we need get access to sources ,intercept the illegal data.

◆ Personal Responsibilities：

- **Responsible for taking up interception strategy,** existing policies on data will do inherited check and blacklist filtering, and name, coordinates, address, city, phone single field inspection .And a pair of coordinates, address, city, multi-field validation.The verification data pass the polices can go through subsequent processing.

- **Responsible for improving interception strategy,**the existing policy has some problem on inheritance check for possible hazards, as well as not robust enough .In order to solve that ,we add white list approach to upgrade.What's more,   according to research ,we improve part of single and multi-field validation field upgrade.

- **Responsible for Write data script,** written awk script, by acquiring field different files,that will integrate data from multiple files into one file, as the analysis basis.

◆ Result：Upgrade strategy has a good effect on enhance the rationality of the inheritance check, nherit the validation error decreased from 20% to 10%, other upgrade strategies to improve the robustness of code effectively.

2013.3-2013.7   Advertisers consumption statistics   **Baidu Promotion Internship Project**   **Designer/Developer**

◆ Project Description：With the growing number of advertisers in Baidu Promotion and the great increasing of mobile equipment,many advertisers only have PC webs,lack of mobile web.what we need to do is provide a tool for advertisers to convert PC webs into mobile ones,what's more, we also need to make it clear and statistic that how the mobile ones function so that we can gather the detailed consumption data.

◆ Personal Responsibilities：

- **Responsible for online data capture,**download the processed files,and test their legality,validity and completeness.

- **Responsible for update statistic script design,**statist advertisers' pv,uv and consumption information and update the database.

- **Responsible for improve statistics script,** as script runs tediously,we split the log file into 90 sub files and add a index pair in database.

- **Responsible for update discount material algorithm design,**I add a discount table into the database for storing the discount material data to satisfy the command of discount material,and write a script to update the discount information every hour.

◆ Result：The scripts have been used for daily statistics and the time of the script running reduce from 3000(s) to 37(s).

Project Experience

2012.9-2012.12 Face Recognition System   **Machine Learning Project**   **Designer/Developer**

◆ Project Description：Currently,face retrieval in a controlled environment is susceptible to light and face rotation.In order to overcome these problems, we extract images LBP features, using Adaboost feature selection to get some weak classifiers combined a strong classifier.With the help of this system we can determine whether the face is our training people easily.

◆ Personal Responsibilities：

- **Responsible for Feature extraction algorithm research.** Considering the characteristics of the image,we choose LBP feather to overcome the influence of the lightness and rotation.

- **Responsible for feature selection algorithm design.**Using Adaboost feature selection algorithm get some weak classifiers to combine a strong classifier for classification.

- **Responsible for improve adaboost algorithm.** Improving the algorithm reduces the time of computing error rate from O(N) to O(1).

◆ Result：Complete face recognition system,and correct rate has been above 85% with the time training reduces from 40(s) to 0.5(s)..

Awards

- **National Scholarship for Encouragement (top 5%,** 2009, 2011),   Scholarship for outstanding study of HUST(**top 10%,** 2009, 2010).
- **Excellent Student Cadre (top 10%,** 2009, 2014),   Excellent member of China's Communist Youth League(2009)

Personal skills

- Familiar with the basic data structures, algorithms. Having a strong sense of logic mind.
- Familiar of **C/C++ ,PHP.**Familiar with **shell,** know **mysql** .
- Software design certification,computer 2(C),3(NET),4(NET) certification.

Self-evaluation

- **Staid,pragmatic,**being able to **concentrate** on doing something.
- A strong ability of **learning,**having a strong sense of **learning new** knowledge.
- Teacher once made commence on me like this :an **ambitious,diligent,inquisitive** person with **great potential.**