Wei Li

Female | (+86)131-2005-5598 | liwei.silvia@gmail.com

EDUCATION

Tsinghua University, Master, Software Engineering, Distributed Machine Learning Area   08.2012-Present

Huazhong University of Science and Technology, Bachelor, Software Engineering   09.2008-06.2012

National Scholarship, University Best Bachelor Thesis Award, University Excellent Student Leader

INTERNSHIPS

Google (Beijing, China)   07.2014-10.2014

Software Engineering Intern, Geo Imagery Group

➢ Speed up the auto geo-tagging online system for UGC Photos introducing caching mechanism in 3 weeks.

➢ Perform async data processing and backend service communication via Producer Graphs on server side.

➢ Reduce timeline loading from minutes to instant. Optimize pipeline and speed up 80% of total requests.

Hulu (Beijing, China)   11.2013-06.2014

Software Development Intern, Recommendation Team

➢ Build a new metric collection framework that parses JSON texts into Hadoop jobs based on Scalding(scala).

➢ Improve decision tree model of Shows You Watch (SYW). Write Hadoop jobs to analyze user behavior logs.

➢ Reduce code lines for monitoring an algorithm from hundreds to tens. Included as SYW patent author.

Microsoft (Beijing, China)   03.2013-08.2013

Project Manager Intern, Windows Plan & PC Ecosystem Department

➢ Windows and IE readiness & adoption. Track compatibility bugs in database and cooperate with partners.

Baidu (Beijing, China)   10.2011-01.2012

Research & Development Intern, NLP Department

➢ Research and implement Coordinate Descent and Cutting Plane Algorithm for linear SVM classifier library.

➢ Refine Machine Learning Research Tools: data preprocessing, task evaluation, feature selection.

RESEARCH & DEVELOPMENT EXPERIENCE

Startup project "Taobook" App, Campus P2P library and floating bottles platform   11.2013-01.2014

➢ Project management and lightweight implementation of user matching, book categorization and retrieval.

EI Compendex indexed paper: G. Shen, W. Li. A Message Passing Graph Match Algorithm Based on a Graphical Model. Active Media Technology 8th International Conference. Proceedings. 2012.

CONTESTS

Advertising Strategy based on Rewards-point & Tagging for Online Video Mobile Client   11.2013

My Sohu Product Optimization Contest, 1st Prize

E-commerce Data Collection and Implementation based on Recommender System   08.2010

The 1st Chinese University Students Service Outsourcing Innovation Application Contest, 2nd Prize

SKILLS SUMMERY

Fast learning ability working with Hadoop/Scalding, Spark, Petuum, Application Framework, open source utils. Sense of requirements and data developed through real production experience and data mining background. Strong team spirit and communication skills demonstrated through well-written design docs and diverse roles including project management and presentation in internship, startup and contests.

李 唯

女 | 131-2005-5598| liwei.silvia@gmail.com | http://liweithu.me

教育经历

清华大学   2012.08–至今
软件学院软件工程专业，硕士，机器学习及分布式计算方向

华中科技大学   2008.09–2012.06
软件学院软件工程专业，学士，模式识别方向
国家奖学金（5/220），校三好学生、优秀毕业论文、优秀毕业生、科技创新优秀个人、优秀学生干部

实习经历

Google   2014.07–2014.10
实习软件工程师，地理图像组
➢ 在 Google+图片地理位置自动标注在线服务中加入缓存机制，三周内完成高质量设计文档及功能实现。
➢ 设计缓存格式，使用 Producer Graph 和后端服务交互进行异步数据处理，用 EasyMock 编写单元测试。
➢ 将时间线的加载时间从十到几百秒降至瞬间，提升用户体验；优化图片标注流水线，提速 80%的查询。

Hulu   2013.11–2014.06
实习开发工程师，个性化及推荐组
➢ 调研分布式计算框架 Scalding (Scala)，实现将 JSON 文本解析反射生成 Hadoop 任务的指标统计新框架。
➢ 分析用户行为数据，统计验证为 Shows You Watch 模型引入新特征时，对用户观看模式的猜想。
➢ 将监测新算法所需新增代码量从几百降至 0-20 行，撰写系列技术文档；被纳为 SYW 专利参与者。

Microsoft   2013.03–2013.08
实习项目经理，Windows 规划及合作伙伴部
➢ 与合作伙伴沟通协调解决 Windows 8 & Blue 平台下的网络游戏及网页兼容性问题，跟踪管理缺陷。

百度   2011.10–2012.01
实习研发工程师，自然语言处理部
➢ 机器学习算法库：调研并实现新算法 Coordinate Descent 和 Cutting Plane，提供线性 SVM 的快速优化。
➢ 机器学习调研配套工具：优化数据预处理、分类/回归任务效果评价、分类任务特征选择工具。

研究项目

创业项目"淘书达人"安卓应用，面向高校学生的 P2P 图书馆及漂流瓶交流平台   2013.11–2014.01
➢ 进行敏捷项目管理，产品设计，以轻量级方式实现图书分类、图书检索、用户配对等后台部分。
基于概率图模型的图相似度研究，EI 核心论文 G. Shen, W. Li. A Message Passing Graph Match Algorithm Based on a Generative Graphical Model. Active Media Technology 8th International Conference. 2012.

比赛经历

基于积分互动和主动标注的视频移动端推荐及广告投放策略，"My sohu"产品优化大赛，一等奖   2013.11
基于推荐系统的电子商务数据收集解决方案，首届中国大学生服务外包创新应用大赛，二等奖   2010.08

技能特点

快速学习能力。能够快速调研 Hadoop(Scalding)/Spark/Application Framework 等大规模系统的运作并产出。
对数据和用户需求敏感。有文本挖掘和产品设计的背景，可增加分析解决实际问题思路的多样性。
团队合作和沟通能力。乐于在实习、创业、比赛中负责项目管理、演讲展示、文档撰写，主动肯干。