

个人简历

个人信息

姓名: 吴德超
性别: 男
出生日期: 1992.2
学校及专业: 山东理工大学计算机科学与技术
学历: 硕士
电子邮件: dechaowu@163.com
电话: 15711057804



期望工作

期望工作性质: 实习
期望职位: 数据挖掘
期望工作地点: 北京
期望月薪: 面议

专业技能

1. 使用 Ubuntu 进行开发, 对于 Linux 常用命令较为熟悉, 擅长使用 Shell/Python 编程快速处理问题。
2. 熟练使用 Python 编程语言, 有 Java、Scala、Python 项目开发和维护经验。
3. 熟悉常用的机器学习算法, 比如决策树、线性回归、逻辑回归、随机森林、GBDT 和聚类等。
4. 熟悉 MySQL、redis、mongodb、HBase 等数据库。
5. 熟悉 Hadoop2 中 hdfs 存取数据以及 MapReduce 框架中数据 shuffle 和排序过程, 另有短期的 MapReduce 程序开发经验。
6. 熟悉 Spark Core、SQL、Mllib, 有 Spark SQL 和 Mllib 方面的项目开发经验。
7. 熟悉 Scrapy 爬虫框架, 有一定的爬虫项目经验。

工作经历

2017.02 — 2017.11	北京领众科技有限公司 大数据工程师 (实习)
	一. 福彩大数据平台项目 1. 编写 MapReduce 程序对系统日志进行初步清洗。 2. 基于彩民购彩数据对彩民兴趣进行分析, 进而得到每个彩民的画像。 3. 基于彩民购彩数据建立模型对彩民行为进行沉迷度判定, 最大程度防止用户沉迷。

	<p>4. 根据历史一段时间的彩票销售金额建立模型对未来彩票销售额进行预测。</p> <p>二 . QST 实训项目</p> <ol style="list-style-type: none"> 1. 担任数据结构讲师。讲解链表、搜索、排序、图论、动态规划等算法。 2. 担任 Python 讲师。讲解 Python 面向对象、爬虫框架 Scrapy 以及机器学习库 scikit-learn 等。 3. 担任 Hadoop 讲师。讲解 Hadoop 架构以及基础的 mapreduce 编程。 4. 担任 Spark 讲师。讲解 spark core 、 sql 、 streaming、mllib 等。
2016.05 — 2016.12	<p>北京巧达科技有限公司</p> <p>数据研发工程师（实习）</p> <ol style="list-style-type: none"> 1. 对海量简历文件做数据处理，包括：内容解析、字段填充、数据排重、数据传输等。 2. 基于推荐算法实现智能简历推荐和推送，挑选出 HR 感兴趣的简历定期推送数。

项目经历（近期）

2017.10~2017.11	<p>QST 青软实训</p> <p>大数据讲师</p> <p>该项目是公司与 QST 青软实训合作，对大学在校本科生进行 IT 技能培训，我在该项目中担任数据结构、Python、Hadoop 与 Spark 讲师，通过该教学任务我对大数据技术知识做了系统的整理，使我的知识体系不断清晰，另外也锻炼了自身的表达能力。</p>
2017.02 - 2017.9	<p>福彩大数据管理平台开发</p> <p>智慧福彩</p> <p>核心业务是分析各地彩民购彩以及彩票销售情况等，平台供福彩内部人员使用，我负责的部分为彩民画像、彩民防沉迷、销售预测。彩民画像采用半监督学习思想，先聚类然后分析得到标签最后进行类别判定；彩民防沉迷模型算法为随机森林，对用户的购彩行为进行预测，最大可能识别出沉迷购彩的用户人群；销售预测则是对全国各省份城市的购彩金额进行预测，训练模型的算法是 ARIMA 时间序列化。以上各功能均是基于 Spark 进行开发。</p>
2016.08 - 2016.12	<p>UU 推荐</p> <p>推荐算法研究</p>

	核心业务为智能简历推荐，根据 HR 的搜索记录或者职位的描述提取职位特征，利用协同过滤算法在众多简历寻找匹配度最高的简历推荐给 HR；我的职责是简历内容解析、字段分析、推荐算法设计等。
2016.04 - 2016.05	文本分类器（NLP） 代码实现 在机器学习算法的学习过程中写的一个小项目，基于朴素贝叶斯算法对海量文本做内容分类。 github 地址： https://github.com/TrustMe5/Naive-Bayes-TextClassifier 。
2016.01 - 2016.03	参加 kaggle 与阿里云天池大数据竞赛 模型构造与调参 Kaggle 和天池大数据主要以 Data Mining 的比赛为主，性质是数据建模与分析预测，选择合适的数据模型解决具体问题。
2016.01 - 2016.02	山东理工大学在线打字系统 独立开发，全权负责 开发这个打字系统的初衷是为了提升我们 ACM 实验室集训队员的打字能力，系统分为练习和比赛两部分，该系统记录队员的日常在线打字情况，并定期组织比赛对队员进行测试。所用技术为：flask、bootstrap
2015.11 - 2016.01	山东理工大学 ACM 在线评测系统开发 数据抓取 学校老版 ACM 代码评测系统漏洞较多，维护代价较大，由导师带领指导开发新版评测系统。我在团队中的职责是 ACM 题面的抓取，所用技术为开源框架 Scrapy。

教育背景

2015.9 -- 至今	山东理工大学	计算机技术（硕士）
2011.9 -- 2015.6	济宁医学院	计算机科学与技术（本科）

个人评价

本人性格开朗、为人诚恳、乐观向上、兴趣广泛、拥有较强的学习和适应能力；工作态度认真负责，积极进取，具有较好的团队协作和沟通能力。

在校表现

硕士阶段：

- 1.成绩优秀，连续两年获得学院三等奖学金。

2.参加 ACM 算法集训和竞赛，参加 Kaggle 大数据竞赛，锻炼了数据挖掘和编程能力，培养了事先计划、做事严谨、事后总结的习惯。

3.撰写并发表《基于 Hadoop 的分布式聚类算法研究》论文，具备研究与创新能力。

4.担任辅导员助理，协助辅导员工作，提高了沟通、组织协调能力。

本科阶段：

1.连续四年获得学院学业奖学金。

2.担任学校社团外联部干事，培养了较强的沟通、组织协调、团结协作能力。

3.长期担任辅导班兼职老师，工作认真负责，培养了较强的沟通表达能力。