

# 个人简历

李 永 博

求职意向：Python数据分析师



Li Yong Bo

女 25岁 未婚  
汉族 本科 应届毕业生  
18822827151 (微信)  
1536645035@qq.com

## 教育背景

2014 – 2018 韩山师范学院 数学与应用数学 本科(统招)

## 个人技能

语言技能：大学英语四级

办公技能：熟练使用Office办公套件

专业技能：

- 1、熟悉数据分析流程：需求分析、数据获取、数据预处理、探索性分析、变量选择、模型开发、模型评估等；
- 2、掌握SQL语言，能在MySQL环境下对数据进行增删改查，了解MongDB和Redis的基本操作；
- 3、熟悉Python基本语法，会用Numpy、Pandas、Matplotlib进行数据清洗、数据分析和可视化；
- 4、了解机器学习的常见算法，如线性回归、决策树、朴素贝叶斯等，能够利用scikit-learn进行模型的调用和调整；
- 5、熟练使用Excel，能够用Excel进行数据的处理、统计分析和辅助决策操作；
- 6、了解前端知识，能够利用Requests、BeautifulSoup、Selenium以及Scrapy框架进行简单的爬虫；
- 7、熟练掌握Linux命令。

## 工作经历



## 项目经验

### 项目名称：小说分类系统

#### 项目描述：

根据客户所提供的文本数据,基于已知的小说标签,对数据集(9万篇小说)进行模型训练,构建标签系统,用于小说分类,便于统计管理。

#### 项目职责:

- 1.处理类别文件,根据类别得到小说标签;
- 2.数据预处理:调用 jieba 分词进行分词,过滤停用词,去掉非中文字符以及单个字;将分词后的数据集进行 Bunch 化操作;
- 3.构建特征工程:用 TF-IDF 算法计算特征分词的权重,对文章进行量化处理,用 N-gram 对词的顺序处理;
- 4.数据集拆分:调用 train\_test\_split 进行数据集拆分;
- 5.模型训练:用多项式朴素贝叶斯算法训练分类器,并进行拉普拉斯平滑处理;
- 6.模型评估:调用 classification\_report 评分函数在验证集上对分类模型进行评分,f1-score在 98% 以上。

### 项目名称：小说推荐系统

#### 项目描述：

通过统计用户喜好和购买历史等标准,为用户推荐符合用户需求的小说,搭建个性化推荐系统。

#### 项目职责:

- 1.数据采集:获取用户基本信息(性别、年龄等)和用户行为信息(浏览历史、收藏、购买历史等)以及小说信息(类别、作者、介绍等)
- 2.数据清洗:用 Numpy、Pandas 对数据中的缺失值和异常值进行处理;对数据进行探索性分析,发现数据规律;
- 3.数据转换:对非数字型数据(如性别、类别等)进行LabelEncoder(标签分类)或Onehot(独热编码)操作,以便于对数据进行更好的操作
- 4、构建特征工程:构建用户与用户之间的相似度矩阵UserCF(基于用户的协同过滤)、构建小说与小说之间的相似度矩阵ItemCF(基于小说的协同过滤)



5、数据建模:对比各个模型(RF、LR、XGBoost 等),选择表现较好的XGBoost 模型进行训练;

6、模型评估：利用GridSearchCV（网格搜索交叉验证）对模型的参数进行调优，以求达到最好的推荐效果（最优可以达到75%）

### 项目名称：数据分析职位分析

#### 项目描述：

对爬取的数据进行数据清洗并进行可视化分析，为公司招聘数据分析岗位提供参考意见。

#### 项目职责：

- 1.将从MongDB数据库中取出数据并转换成DataFrame类型，对数据结构等进行简单分析;
- 2.用 Numpy、Pandas 对数据进行基本处理，包括去重、缺失值和异常值处理等;
- 3.调用jieba分词对职位描述信息进行分词，过滤停用词，并利用正则表达式对其剔除常规信息;
- 4.对数据进行可行性分析，如城市需求分布、行业需求分布、工作年限薪资分布、教育背景薪酬分布等，并通过Matplotlib、Pandas等工具可视化显示出来;
- 5.调用wordcloud库生成词云图，分析职位需求和技能要求。



## 个人评价

- ❖ 从事数据分析2年，有优秀的数学思维和敏锐的洞察力，做事有条理，工作认真负责；
- ❖ 自主学习能力较强，喜欢利用各种资源来提升自己的专业能力，能持续学习数据分析的各种知识；
- ❖ 性格积极乐观，自信稳重，具有良好的团队合作精神和较强的沟通能力。