

简历

姓名：陈功 性别：男
年龄：26 岁 学历：本科
联系方式/微信：17623350266
电子邮件：allenchen0710@163.com



教育经历

2010/09-2014/07

重庆大学 机电一体化 本科

课程：C 语言程序设计、 线性代数、 概率论与数理统计等

证书：CET-4、全国计算机等级二级

技能描述

- 1.熟练掌握 Python 语言，能够熟练在 ubuntu、windows 下进行 python 开发。
- 2.熟练掌握 numpy、pandas、matplotlib、seaborn 等基本数据分析工具，能够运用工具对数据进行清洗，处理、分析及可视化。
- 3.熟练掌握 SQL 语言，掌握 MySQL，MongoDB，redis 等数据库常用操作。
- 4.熟练掌握机器学习数据的预处理，以及主成分分析 PCA。
- 5.了解机器学习原理及基本算法，线性回归、逻辑回归、朴素贝叶斯、决策树、随机森林、XGboost、adaboost、K 近邻、SVM、聚类等，能够对各种算法进行运用及算法融合。
- 6.了解神经网络、CNN、RNN 等深度学习算法。
- 7..熟练掌握机器学习数据的预处理，以及主成分分析 PCA。
- 8.能够运用 scikit-learn，对各种算法进行使用，运用 jieba 库对中文进行分词处理。
- 9.掌握爬虫的基本原理，熟悉爬虫库的基本使用，urllib，requests、bs4、selenium，

对爬虫框架 scrapy 有熟练地运用，能够进行分布式数据抓取，解析处理入库等操作。

10.熟练掌握 linux 操作系统的常规操作，熟悉 shell 编程，掌握 git 版本控制系统。

工作经历

2015 年 9 月-2018 年 12 月 百成俊网络科技有限公司 数据分析师

- 1、负责公司项目的数据清洗，处理，分析。
- 2、运用可视化工具为项目经理提供分析数据的依据。
- 3、调节参数模型，调试出最合适的机器学习模型。

2014 年 7 月-2015 年 8 月 雅马哈发动机研发 BD 设计

负责公司摩托车 BD 设计研发。

项目经验

游戏用户留存预测 2018/4-2018/9

项目描述：使用游戏用户在游戏中的数据，分析用户次日的流失情况，对于预测将要流失的客户，采取挽留措施。

项目职责：

- 1.由游戏公司提供他们挖掘的数据（等级、金币、钻石、体力、战力、游戏动作等），并运用 numpy、pandas、matplotlib 对数据进行可视化分析。
- 2.排除干扰信息。如：游戏动作少于一定次数的玩家，注册 7 日以内的玩家，剔除冷门动作。
- 3.根据对数据的分析情况采用 pandas 等工具对数据进行清洗，缺失值处理、异常值修改。
- 4.通过对比经典学习算法，综合考虑效果之后运用 XGboost 对数据进行分类。
- 5.对测试集计算结果，准确率达到 81%左右。对于流失玩家的预测精确率和召回率为 79%，82%，对于非流失玩家的预测精确率和召回率为 82%，80%。

信用评分模型 2017/1-2017/6

项目描述：根据用户以往的信用数据，对用户的还款行为进行分析，帮助金融企业决定是否贷款做出决策。

项目职责：

- 1.数据预处理，包括数据清洗、缺失值处理、异常值处理。
- 2.探索性数据分析，使用直方图、箱形图获取样本分布的大概情况，为进一步的操作打基础。
- 3.变量选择，筛选出对违约状态影响最显著的指标，主要有单变量特征选择方法和基于机器学习模型的方法。
- 4.模型开发，该步骤主要包括变量分段、变量的 WOE（证据权重）变换和逻辑回归估算三部分。
- 5.模型评估，该步骤主要是评估模型的区分能力、预测能力、稳定性。
- 6.使用 Logistic 模型跑数据，并对生成的模型进行评估。
- 7.运用 ROC 曲线评估，结果的 AUC 值符合要求。

商城推荐系统 2016/3-2016/8

项目描述：基于隐语义模型，维护一个用户、item 预测表，将预测的结果由高到低选取最高几项通过推荐栏显示的方式推荐给用户。

项目职责：

- 1.在 matplotlib 中通过箱行图，直方图，热力图等可视化工具分析用户和商品数据。
- 2.使用 numpy、pandas 对数据进行数据清洗、缺失值处理、异常值修改。
- 3.经过分析提取正反馈、负反馈数据。
- 4.利用梯度下降法优化出预测表中各个参数。
- 5.利用参数求出用户 item 的所有预测结果，使用结果推荐给用户。

自我评价

具有良好的逻辑分析能力和沟通能力

具有阳光、积极的生活工作态度

工作认真负责、严谨，精益求精，具有团队协作精神，学习能力强