

肖智元

意向岗位: 数据分析师

电话: 18527850256

邮箱: 1786032122@qq.com

性别: 男

年龄: 21



教育经历

2020.09 - 2024.06

西华师范大学

人工智能 | 本科

科学计算与数据分析、数据库原理及应用、线性代数、概率论与数理统计、数据结构、离散数学、操作系统。



相关技能

- 数据库: 熟悉MySQL
- 数据处理和分析: 熟悉Numpy、Pandas、Matplotlib
- 网络数据爬取: 熟悉Requests、Selenium、BeautifulSoup4、Xpath、正则表达式
- 数据清洗和分析工具: 熟悉Excel (基本函数和Power Query)
- 可视化工具: 熟悉PowerBI
- 机器学习算法: 了解Logistic Regression、支持向量机、K最近邻分类、高斯朴素贝叶斯、分类决策树、随机森林分类



项目经历

2021.10 - 2022.07

ELE英语学习网站用户数据分析

用户行为分析

背景:

对校内ELE英语学习网站项目产生的用户信息进行分类分析, 用以监督、帮助大家英语学习

相关工具: SQL、python、django

任务: 对不同的用户制作通用的学习看板, 可视化用户的学习进度

行动:

- 数据提取: 使用SQL和Python对不同用户进行分组, 并收集他们每天的学习数据, 以了解他们的学习进度。
- 数据建模: 根据艾宾浩斯遗忘曲线的原理, 利用用户的学习数据建立了一套学习方案, 帮助用户优化记忆效果。
- 数据可视化: 针对不同用户, 我们制作了学习进度和学习看板。通过直观的图表展示, 用户可以清楚地了解自己的学习进度, 并且可以及时调整学习计划。

成果:

学习看板可以提高学习效率, 通过定期复习, 用户的记忆效果明显提升, 并能够提供个性化学习建议

背景：数学建模校赛《三胎政策对人口状况、经济及其影响的研究》

相关分析工具以及模型：matlab，python，numpy，sklearn，Leslie模型

任务描述：

- 2035年时中国的人口结构情况。
- 分析三胎政策对国家经济增长情况的影响。

项目过程：

- **分析问题选择模型**：选择了Leslie模型作为研究人口结构和经济增长的分析工具，因为它具有较高的准确性和可靠性。
- **收集数据与数据清洗**：选择《中国人口统计年鉴》上数据，进行数据清洗和补充缺失值，确保了数据的准确性和完整性。
- **数据分析**：基于Leslie模型，建立离散模型，将人口划分为多个年龄段，预测每年的人口结构和经济增长情况。

项目结果与分析：Leslie模型和第六次人口普查数据，三胎政策有望缓解劳动力短缺问题，并提供更多的劳动力资源，稳定和促进经济增长。

数学建模与编程

背景：

在2022年的数学建模省赛B题中，我作为团队的成员，对问题三四进行建模分析预测

相关分析工具与模型：python、numpy、pandas、sklearn、SPSS、决策树、Pearson 相关性

任务：

- 问题三：未知类别玻璃文物的化学成分进行分析，鉴别其所属类型，并对分类结果的敏感性进行分析。
- 问题四：针对不同类别的玻璃文物样品，分析其化学成分之间的关联关系，并比较不同类别之间的化学成分关联关系的差异性。

数据处理：对数据进行处理得到纹饰、类型、颜色、表面风化、以及二氧化硅(SiO₂)、氧化钠(Na₂O)等十四种化学成分

问题三行动：

- **分析问题选择模型**：根据文献阅读和调研以及相关算法的分析，使用有监督的决策树。
- **数据分析**：决策树算法通过有监督学习对已知类别数据进行训练。未知类别数据代入模型进行预测分类。

结果：该模型经过数据维度规约和剪枝操作，在预测分类时表现稳定且准确。

问题四行动：

- **分析问题选择模型**：数据为连续变量，因此选择 Pearson相关性分析多变量间的关系，使用偏相关分析更进一步分析两两变量之间的相关关系并判断化学成分关联关系的差异性的差异性。
- **数据分析**：对数据进行决策树分析。计算多变量与两变量的相关关系

结果：分析玻璃文物样品的化学成分关联性，比较不同类别间的差异。



博客

CDSN：<https://blog.csdn.net/sjyrpb>

项目主页：<https://sjyrpb.github.io>



在校经历

数学建模四川省省一等奖

蓝桥杯B组省二等奖

计算机设计大赛校二等奖



自我评价

我有数量统计技能、编程技术能力以及问题解决能力。注重细致入微和准确性，在数据处理过程中力求精确性。具备建模和预测能力，能够使用统计方法和机器学习技术构建模型并进行预测。还具备良好的沟通能力，能够将复杂的数据分析结果以简明扼要的方式传达给非技术人员。