



崔鹏

📞 13075522108 • @ 12110802@mail.sustech.edu.cn

## Profiles

🔗 [Cui-Peng-624](#)

## Skills

### 编程语言

◆ ◆ ◆ ◆ ◆

Python, R, Java, SAS, C/C++

### 工具与技术

◆ ◆ ◆ ◆ ◆

数据分析, 机器学习, 爬虫技术, 大模型API集成, 大数据处理

### 深度学习框架

◆ ◆ ◆ ◆ ◆

PyTorch

## Languages

### 全国大学英语四级考试

总分: 538

◆ ◆ ◆ ◆ ◆

### 全国大学英语六级考试

总分: 476

◆ ◆ ◆ ◆ ◆

### TOEFL

总分: 90

◆ ◆ ◆ ◆ ◆

## Awards

新生奖学金三等奖

2021-2022校级奖学金三等奖

## Summary

南方科技大学数据科学与大数据专业大四学生, 具备扎实的数据分析与深度学习技能。精通Python, 熟悉R、Java、SAS, 并通过多个项目积累了丰富的实战经验。自主开发创新算法DirectionalKMeans, 构建了Telegram-Academic-Bot, 展现了我在AI与数据分析方面的创新能力。期待在人工智能与生成式AI领域继续深造。

## Education

### 南方科技大学

数据科学与人工智能

3.6/4 (87/100)

2021.9 - .

数据科学与大数据专业

## Projects

### 数据分析项目

参与多个数据分析项目, 包括悉尼房价、FIFA足球运动员和网球比赛等数据的分析, 使用Python、R语言进行数据清理、可视化和建模, 提升了数据分析和处理复杂数据集的能力。

### Telegram-Academic-Bot

🔗 <https://github.com/Cui-Peng-624/Telegram-Academic-Bot>

一个基于Telegram、OpenAI API和Pinecone API等API, 文本分类模型和爬虫的自动化工具, 支持快速查询论文、推荐论文及与PDF文档交互, 运用RAG技术实现问答功能。

### DirectionalKMeans算法

🔗 <https://github.com/Cui-Peng-624/DirectionalKMeans-Experimental>

自主设计并实现DirectionalKMeans算法, 解决了KMeans在指定方向上分类的局限性。此算法通过删除无用数据点并结合投票机制改进了分类效果。

### 大模型训练可视化工具

🔗 <https://github.com/Cui-Peng-624/TrainingVisualizer>

自行开发了一个大模型训练过程的可视化工具, 方便展示和监控深度学习模型的训练进度与性能指标。该工具极大提高了实验调试效率。

### SC-Position-Encoding

🔗 <https://github.com/Cui-Peng-624/SC-Position-Encoding>

探索并对比了球坐标与Transformer中位置编码的效果差异, 将球坐标应用于卷积神经网络 (CNN), 评估其在不同任务中的表现。此项目展示了对卷积和Transformer架构的深入理解和创新能力。