

scikit-learn简介

韩雁泽

华清远见开源实验室
yyzlab.com.cn



目录

CONTENTS |

- 01.** scikit-learn简介
- 02.** scikit-learn的主要功能
- 03.** scikit-learn的优缺点

01 scikit-learn简介





工具介绍 - 开源，可商业使用 - BSD许可证

华清远见
HQYJ.COM

简单高效

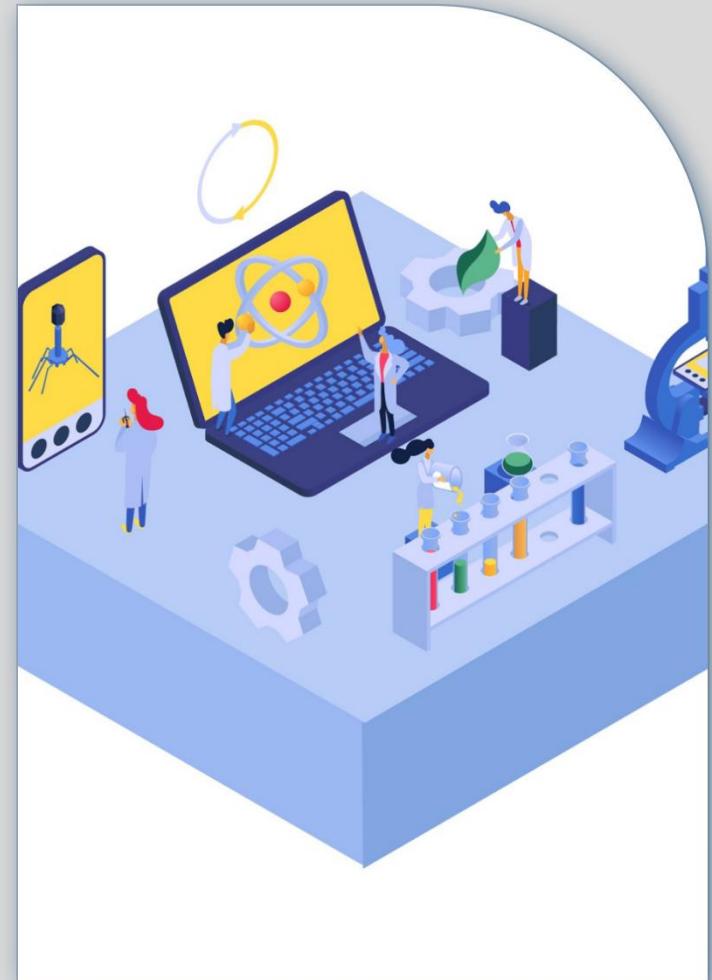
scikit-learn提供了简单高效的算法和工具，方便用户快速进行数据分析和机器学习。

模块化设计

scikit-learn采用模块化设计，使得用户可以根据需要自由组合不同的算法和工具。

丰富多样的算法

scikit-learn提供了丰富多样的机器学习算法，包括分类、回归、聚类、降维等，满足用户不同需求。





适用场景



01

数据挖掘

scikit-learn适用于数据挖掘和数据分析任务，能够高效地处理大规模数据集。

02

机器学习

scikit-learn提供了丰富的机器学习算法，包括分类、回归、聚类等，适用于各种机器学习任务。



01

安装scikit-learn

可以使用pip或conda等包管理工具进行安装，例如：
pip install -U scikit-learn

02

基本使用方法

导入必要的模块和函数，例如：
from sklearn.linear_model
import LinearRegression



机器学习的定义

01

支持向量机

scikit-learn支持向量机算法，用于分类和回归分析。

03

随机森林

scikit-learn随机森林算法，通过集成学习提高分类精度。

02

决策树

scikit-learn决策树算法，适用于分类和回归问题。

02 scikit-learn的主要功能





数据预处理

华清远见
HQYJ.COM

数据清洗

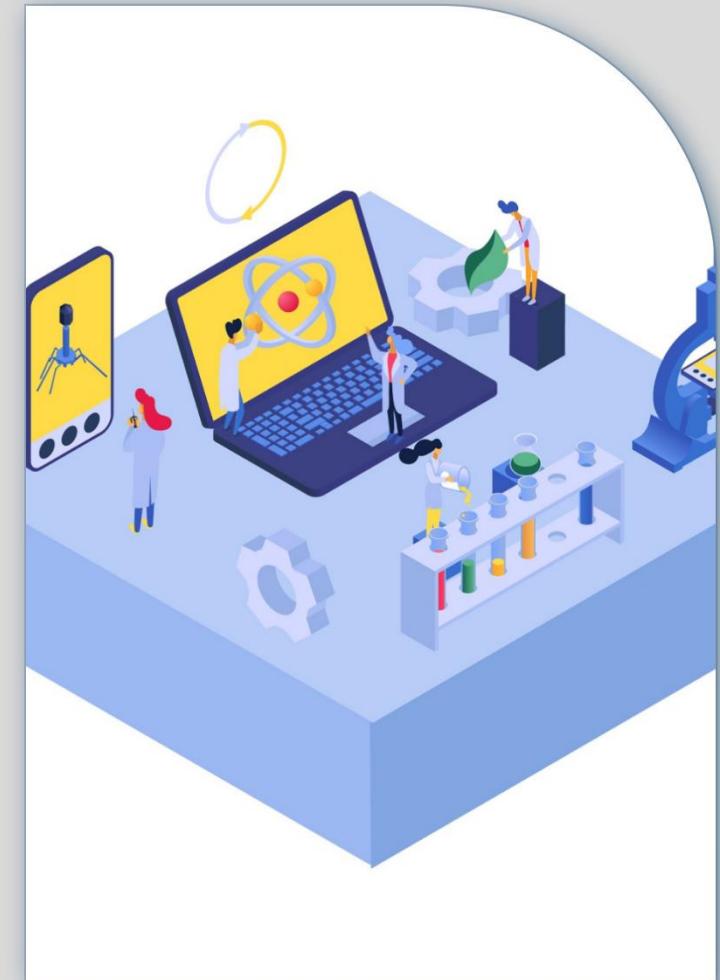
scikit-learn提供了数据清洗的功能，包括缺失值处理和异常值检测等。

特征选择

scikit-learn提供了特征选择的功能，可以帮助用户选择最重要的特征，提高模型的准确率。

特征转换

scikit-learn提供了特征转换的功能，可以将原始特征转换为更符合模型需求的特征。





分类、回归算法

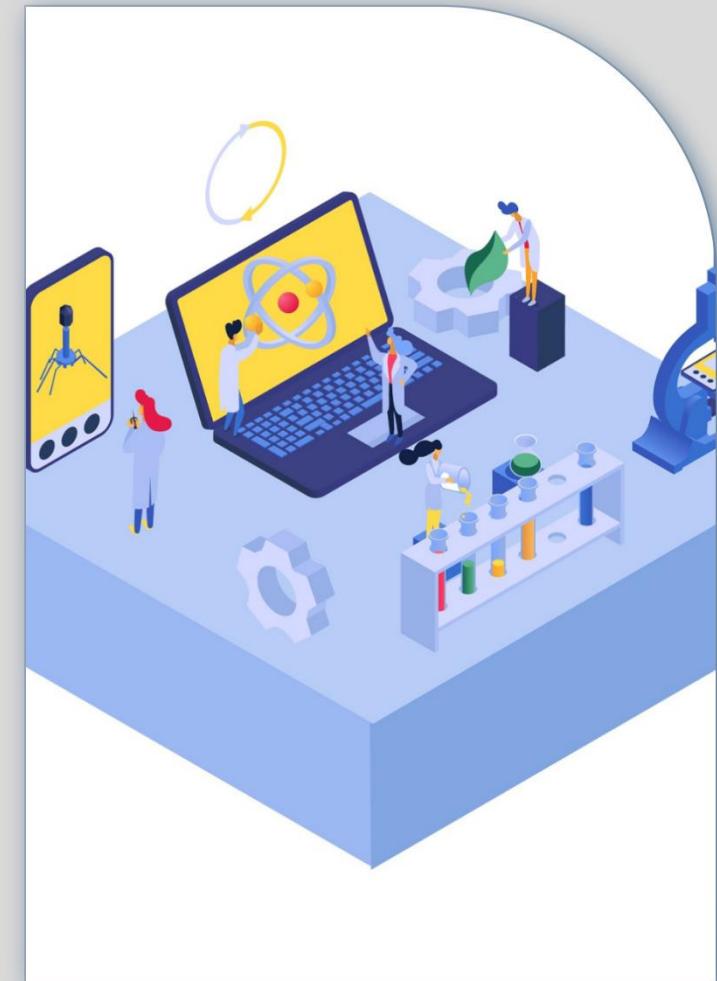
各种回归算法、支持向量机、决策树、随机森林等。

聚类算法

K-means、谱聚类、DBSCAN等。

降维技术

PCA（主成分分析）、LDA（线性判别分析）等。



03 scikit-learn的优 缺点





优点



01

功能强大

scikit-learn提供了丰富的机器学习算法和工具，可以满足各种不同的需求。

02

易于使用

scikit-learn的API设计简洁明了，易于理解和使用，降低了机器学习的门槛。

03

社区支持

scikit-learn拥有庞大的用户和开发者社区，为使用者提供了丰富的资源和支持。



缺点



01

不易理解

对于初学者来说，scikit-learn的文档和API可能难以理解，特别是英文官网，入手难度高，中文网站API往往滞后。

科技赋能·智引未来



技术领先
品质保障



超多干货
实时更新



海量视频
贴身学习

了解更多信息请访问：

集团官网：www.hqyj.com

创客学院网址（在线课程）：www.makeru.com.cn

研发中心网址（研发产品）：www.fsdev.com.cn

元宇宙实验中心网址（在线实验）：www.yyzlab.com.cn