



# 基于Azure的Python机器学习

平安金融壹账通大数据研究院

微软MVP 王大伟



# 目录

## CONTENTS

- >> Azure与Python
- >> 如何用Azure完成机器学习
- >> Azure与自动机器学习
- >> Azure的相关学习资料





# Azure与Python

# 日渐流行的Python



TIOBE给出的排行榜是具有权威性质的，是判断语言流行趋势的指标。

Sep 2019	Sep 2018	Change	Programming Language	Ratings	Change
1	1		Java	16.661%	-0.78%
2	2		C	15.205%	-0.24%
3	3		Python	9.874%	+2.22%
4	4		C++	5.635%	-1.76%
5	6	▲	C#	3.399%	+0.10%
6	5	▼	Visual Basic .NET	3.291%	-2.02%
7	8	▲	JavaScript	2.128%	-0.00%
8	9	▲	SQL	1.944%	-0.12%
9	7	▼	PHP	1.863%	-0.91%
10	10		Objective-C	1.840%	+0.33%

TIOBE排行榜的网址是：<https://tiobe.com/tiobe-index/>



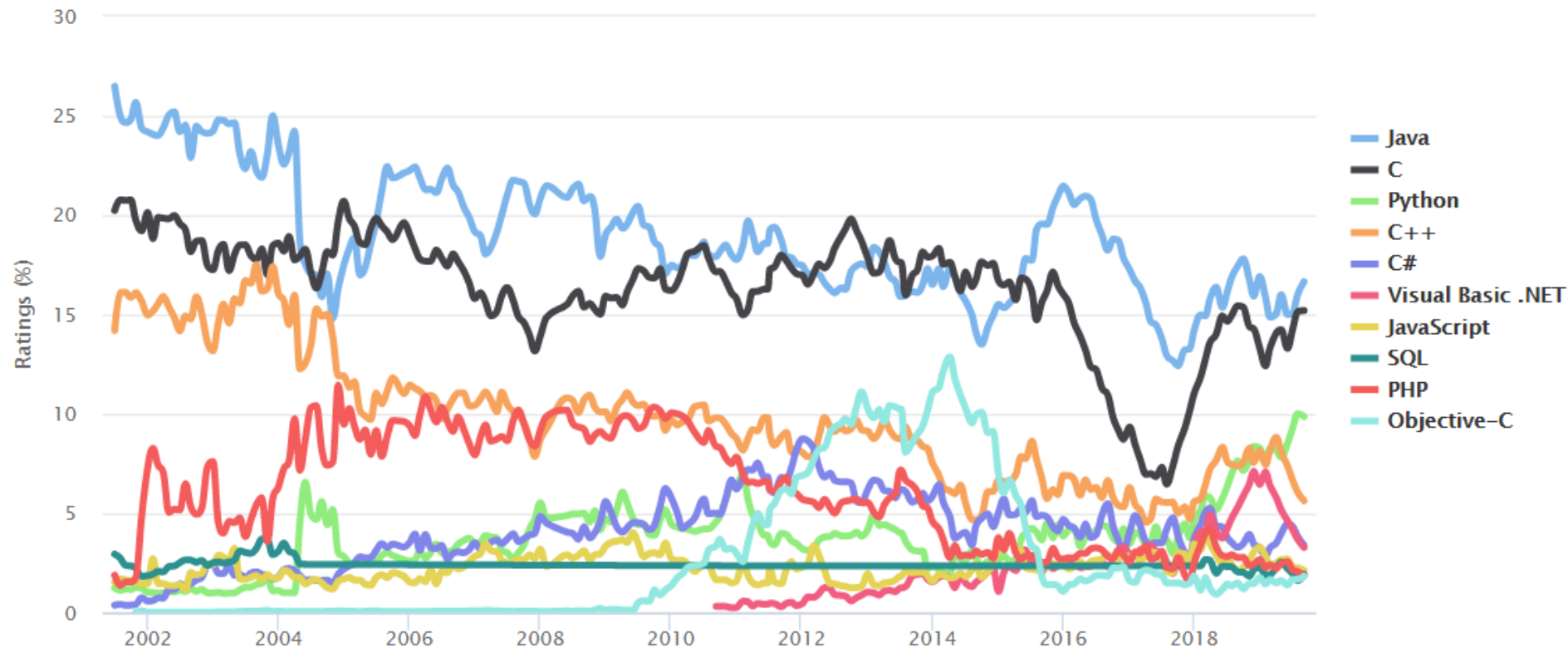
# 日渐流行的Python

PyConChina  
2019



## TIOBE Programming Community Index

Source: [www.tiobe.com](http://www.tiobe.com)



# 日渐流行的Python

## Very Long Term History

To see the bigger picture, please find below the positions of the top 10 programming languages of many years back. Please note that these are *average* positions for a period of 12 months.

Programming Language	2019	2014	2009	2004	1999	1994	1989
Java	1	2	1	1	9	-	-
C	2	1	2	2	1	1	1
Python	3	7	5	6	23	21	-
C++	4	4	3	3	2	2	2
Visual Basic .NET	5	9	-	-	-	-	-
C#	6	5	6	7	18	-	-
JavaScript	7	8	8	8	16	-	-
PHP	8	6	4	5	-	-	-
SQL	9	-	-	89	-	-	-
Objective-C	10	3	29	37	-	-	-



# 日渐流行的Python



Year	Winner
2018	🏆 Python
2017	🏆 C
2016	🏆 Go
2015	🏆 Java
2014	🏆 JavaScript
2013	🏆 Transact-SQL
2012	🏆 Objective-C
2011	🏆 Objective-C
2010	🏆 Python
2009	🏆 Go
2008	🏆 C
2007	🏆 Python

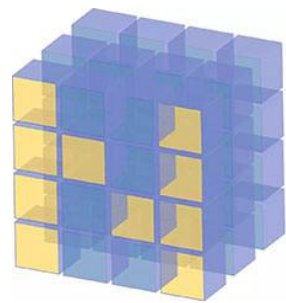


# 为什么用Python完成机器学习



Python的优势：易学习、大量不断更新的各领域库、尤其适合完成机器学习相关任务。

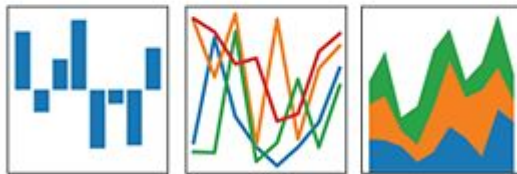
Python机器学习相关库介绍：



NumPy

pandas

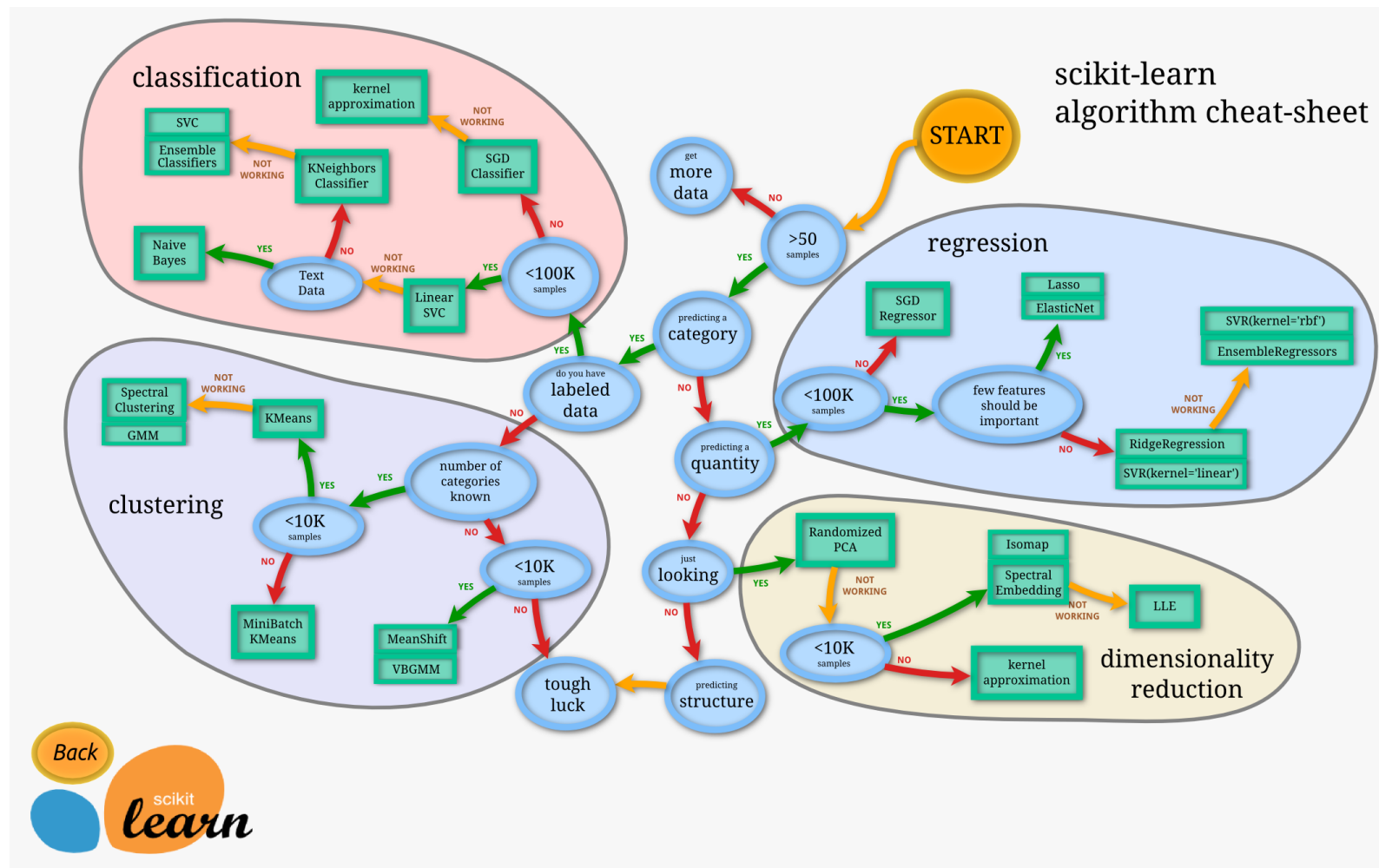
$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



matplotlib  
Version 3.1.1







# Azure是什么？

Azure 是一个不断扩展的云计算服务集合。通过 Azure，公司和组织可以加快发展步伐，提高工作效率，节省运营成本。



Azure网站：<https://azure.microsoft.com/zh-cn/>



# Azure行业解决方案

Azure 解决方案提供相关产品、服务和第三方应用程序，满足各种需求，帮助成功应对商业挑战。



## 政府

构建安全的解决方案，帮助政府机构更好地服务和保护自己的公民。



## 金融服务

重定义客户体验、提高员工能力、优化风险管理并转变产品使命。



## Retail

通过基于数据的客户见解，提供出色的个性化无缝体验。



## 制造

获取可操作见解，快速响应客户反馈和市场趋势。



## 医疗健康与生命科学

制定解决方案、提高医疗保健业绩和效率，同时利用技术创新。



## 能源

开发可用于优化数字操作、支持工作人员、实现客户效果和创造可持续未来的安全解决方案。





# 如何用Azure完成机器学习

# Azure机器学习



进入Azure服务页面：<https://portal.azure.com/#home>

Microsoft Azure

搜索资源、服务和文档(G+I)

**Azure 服务** [全部查看\(超过 100 个\)](#) > [创建资源](#) >

虚拟机 应用程序服务 存储帐户 SQL 数据库 Azure Database for PostgreSQL Azure Cosmos DB Kubernetes 服务 函数应用

**Microsoft Learn**  
通过 Microsoft 的免费在线培训了解 Azure

**Azure Monitor**  
监视应用和基础结构

**安全中心**  
保护应用和基础结构的安全

**成本管理**  
免费分析和优化云支出

**最新资讯** [查看你的所有最新资讯](#) > [查看你的所有资源](#) >

名称	类型	上次查看
<a href="#">docs-ws</a>	机器学习服务工作区	6 天前
<a href="#">docs-ws1</a>	机器学习服务工作区	7 天前
<a href="#">docsws4386490351</a>	存储帐户	7 天前
<a href="#">Free Trial</a>	订阅	7 天前
<a href="#">docs-aml</a>	资源组	3 周前

**实用链接**

[技术文档](#)

[Azure 服务](#)

[近期 Azure 更新](#)

[Azure 迁移工具](#)

[查找 Azure 专家](#)

**Azure 移动应用**

Download on the App Store

GET IT ON Google Play



# Azure机器学习



通过点击“所有服务”，我们可以看到Azure可提供的服务非常多  
找到我们本次需要的“AI + 机器学习”

所有服务

🔍 搜索AI + 机器学习

全部内容

常规

计算

网络

存储

Web

移动

容器

数据库

分析

AI + 机器学习

物联网

混合现实

AI + 机器学习 (8)

💡 Batch AI

★

🧠 机器学习服务工作区

★

🧬 Genomics 帐户

★



在“机器学习服务工作区”中可以看到已有的服务

主页 > 机器学习服务工作区

## 机器学习服务工作区

默认目录

+ 添加

≡ 编辑列

↻ 刷新

|

◆ 分配标记

🗑 删除

订阅: Free Trial

按名称筛选...

所有资源组

所有位置

2 个项

<input type="checkbox"/> 名称 ↑↓	资源组 ↑↓	位置 ↑↓
<input type="checkbox"/>  docs-ws	docs-aml	Southeast Asia
<input type="checkbox"/>  docs-ws1	docs-aml	Southeast Asia



选择“添加”按钮，填写相关信息

主页 > 新建 > 机器学习服务工作区

## 机器学习服务工作区

创建

---

\* 主要    标记    \* 审阅

\* 工作区名称

docs-ws ✓

订阅

Free Trial ▼


资源组

(新项) docs-aml ▼

新建

位置

东南亚 ▼



为方便起见，这些资源会自动添加到工作区(如果区域支持的话): [Azure 存储](#)、[Azure Application Insights](#) 和 [Azure Key Vault](#)。








等待服务部署完成

主页 > Microsoft.MachineLearningServices - 概述


 **Microsoft.MachineLearningServices - 概述**  
部署


概述


输入


输出


模板


 删除

 取消

 重新部署

 刷新

 **部署完成**

 部署名称: Microsoft.MachineLearningServices  
订阅: [Free Trial](#)  
资源组: [docs-aml](#)

▼ 部署详细信息 (下载)

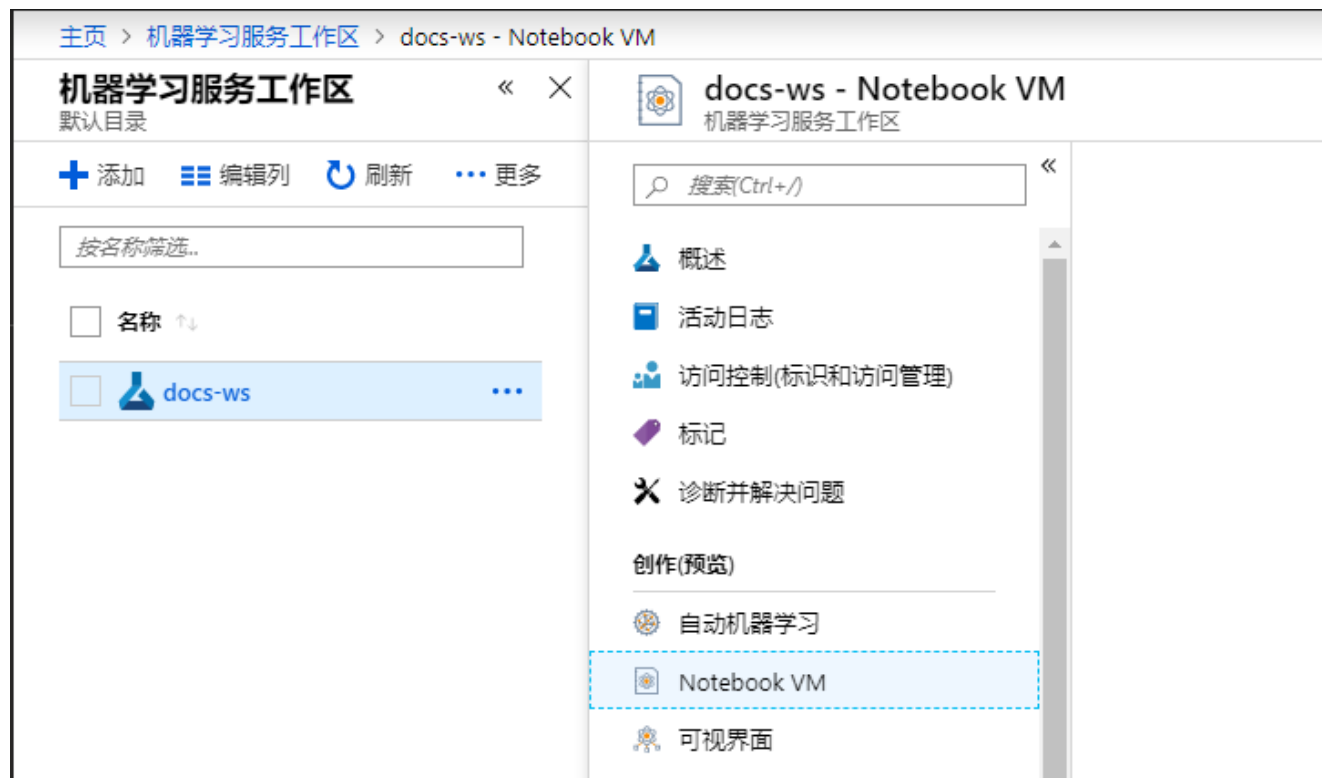
^ 后续步骤

转到资源



单击此处添加标题

将部署的服务内容的Notebook VM打开





## 新建一个笔记本

### 新笔记本 VM

名称 \*

虚拟机大小 \*

[选择虚拟机大小](#)

创建

取消





点击新建的笔记本的 Jupyter部分， 可以打开Jupyter

## 笔记本 VM (预览)

+ 新建   刷新   开始   停止   重新启动   删除

名称	状态	URI
azure-ml	 正在运行	<a href="#">Jupyter</a> <a href="#">JupyterLab</a>





点击新建的笔记本的 Jupyter部分， 可以打开Jupyter



Files

Running

Clusters

Select items to perform actions on them.



0



/



live.com#wangdawei



# Azure机器学习

视频演示Azure机器学习的demo





# Azure与自动机器学习

# 什么是自动机器学习？



机器学习的一般步骤包括：问题定义、数据收集、特征工程、模型选择、模型评估、模型应用。

而算法工程师的工作一般是从特征工程开始。

自动机器学习的自动体现在：自动特征工程、自动模型选择、自动超参数优化等。





# 自动机器学习解决了什么问题？



如何有效选择对应于  
特定数据集的模型至  
关重要。

自动模型选择

自动特征工程

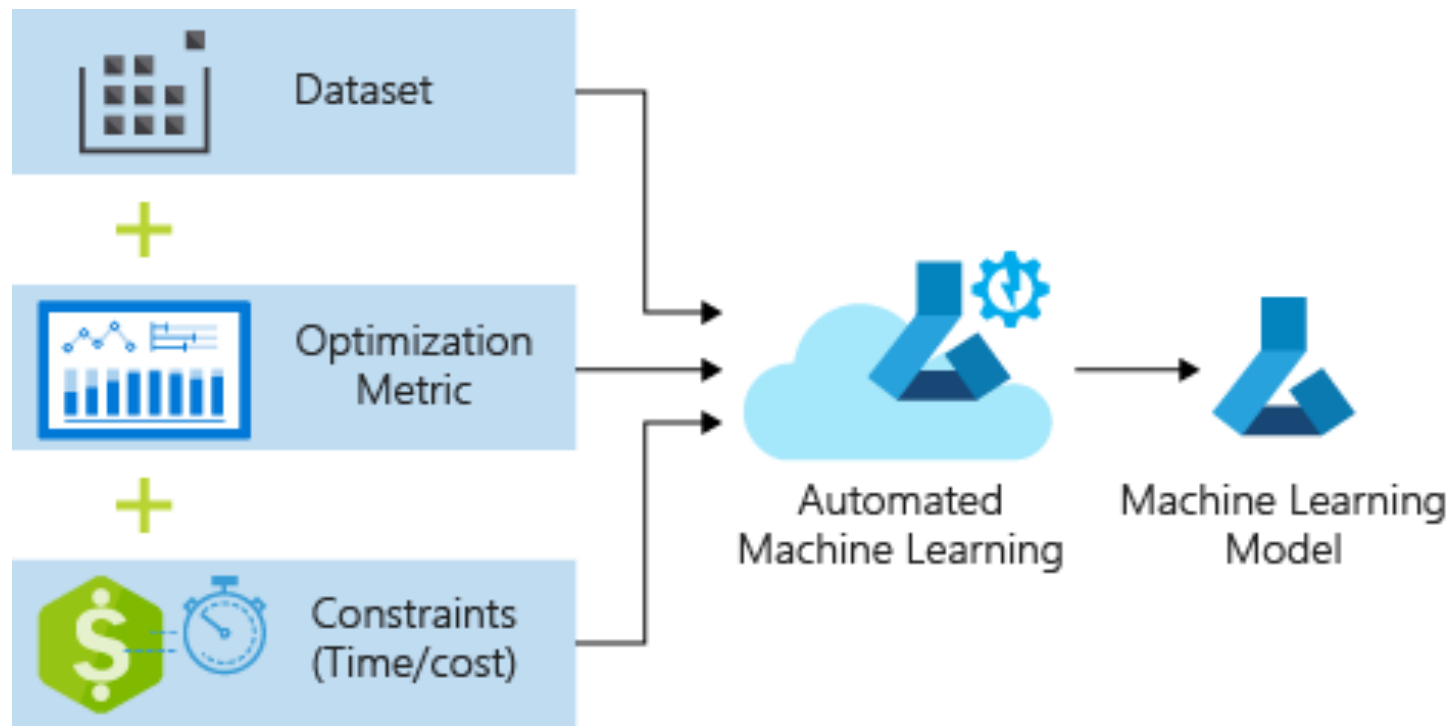
自动超参数优  
化

手动特征工程效率低、  
可移植性差、受到创  
造力的限制。

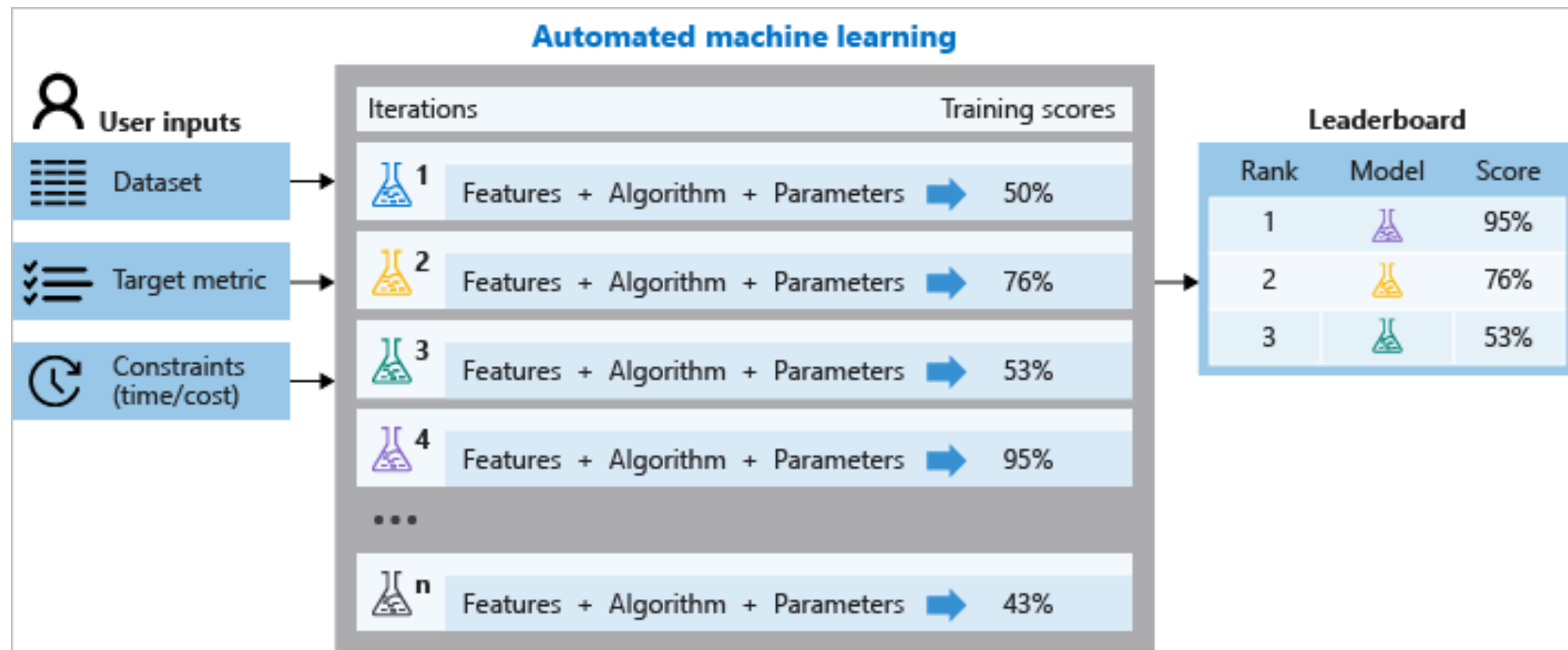
大量超参数如何自动  
优化？



## 利用Azure自动机器学习



# 什么是自动机器学习？



# Azure机器学习

视频演示Azure自动机器学习的demo





# Azure的相关学习资料

# Azure & Python 机器学习的相关学习资料

What is the Azure Machine Learning SDK for Python?

<https://docs.microsoft.com/zh-cn/python/api/overview/azure/ml/intro?view=azure-ml-py>

各种库、模块、方法的作用与参数：

## azureml-core package

### Packages

core	<p>Contains core packages, modules and classes for Azure Machine Learning.</p> <p>Main areas include managing compute targets, creating/managing workspaces and experiments, and submitting/accessing model runs and run output/logging.</p>
data	<p>Contains modules for data related functionality.</p> <p>It includes Datastore, which stores connection information to an Azure storage service, DataPath, which points to a location within a datastore, and Dataset, which points to a DataPath and contains transformations applied to the data.</p>

Azure Machine Learning SDK for Python (Latest)

按标题筛选

SDK overview

Install/update the SDK

Tutorial: train & deploy models

Get support

API Reference

API Reference

> azureml-core

> azureml-accel-models

> azureml-dataprep

> azureml-explain-model

> azureml-opendatasets

> azureml-pipeline-core

> azureml-pipeline-steps

> azureml-telemetry

> azureml-tensorboard

> azureml-train-automl

> azureml-train-core

> azureml-widgets

> azureml-contrib-datadrift

> azureml-contrib-notebook

PyConChina  
2019





# THANK YOU

wechat :

lezi10121643

请备注姓名 - 行业



Python爱好者社区公众号