VUL·AI 华云安

创新 卓越 奋斗 专业

人工智能应用经验分享

北京华云安信息技术有限公司

汇报人：安全研究院 郝伟

日　期：2022/01/11 18:00

主要目标

提高公司主要成员对人工智能价值的认识，具体包括：

管理人员：正确理解人工智能在各方面决策中的影响力；

销售人员：更好地宣传公司产品中人工智能的能力水平；

产品人员：更好地将人工智能应用到产品设计中；

技术人员：理解在开发中是否需要以及如何使用人工智能技术；

其他人员：对其岗位可能有相应的帮助作用。

背景

- **人工智能日趋强大**

  随着谷歌的深度学习带动的新一轮的人工智能研究热潮，越来越多的研究人员和工程技术人员的加入，人工智能的能力不断突破，带给我们越来越多的惊喜。

- **人工智能是大势所趋**

  人工智能、基因工程、纳米科学 并称为21世纪三大尖端技术，是各行各业都在探寻人工智能与本行业的切入点，加速人工智能助力行业的发展是大势所趋。

- **人工智能比较复杂**

  人工智能基于大量复杂的理论知识，一般人难以理解，因此常规理解大都基于字面意思或实际应用等表面内容，导致不同人群对人工智能理解深度有限，理解偏差巨大。

对人工智能认识的一些偏差

- 常见误区
  - 范围过大：人工智能技术能够应用于所有的场景；
  - 能力过强：所有难以解决问题只要使用人工智能都能解决；
  - 易于实现：建个模型或调个接口就能实现人工智能。
- 具体观点
  - 人工智能比自动化性能更好；
  - 人工智能一定要使用Python；
  - 深度学习才是人工智能；
  - 机器学习不如深度学习等。

VUL·AI 华云安

主要内容

01 什么是人工智能

02 什么是机器学习

03 机器学习应用原理

04 常见问题

VUL·AI
华云安

01 | 什么是人工智能

以下能力属于人工智能的范畴吗？

1. CPU或显卡的运算能力；
2. 内存和硬件的存储能力；
3. Word、Excel、PowerPoint的文字处理能力；
4. 科学软件的数学计算、公式指导、解方程等能力；
5. 仿真软件的力学、结构学、工程学等模拟能力；
6. 网络应用中的爬虫、口令爆破、指纹识别能力；
7. 人脸识别、文字识别等识别类软件的识别能力。

人工智能的本质

AI, Aritificail Intelligence

人工 智能？

智能的内涵

记忆
MEMORY

证明
PROVE

想象
IMAGINATION

理解
UNDERSTANDING

判断
JUDGEMENT

计算
CALCULATION

情感
EMOTION

语言
LANGUAGE

控制
CONTROL

推理
REASONING

归纳
CONCLUDING

演绎
DEDUCTING

以下能力属于人工智能的范畴吗？

1. CPU或显卡的运算能力；
2. 内存和硬件的存储能力；
3. Word、Excel、PowerPoint的文字处理能力；
4. 科学软件的数学计算、公式指导、解方程能力；
5. 仿真软件的力学、结构学、工程学等模拟能力；
6. 网络应用中的爬虫、口令爆破、指纹识别能力；
7. 人脸识别、文字识别等识别类软件的识别能力。

都是人工智能？

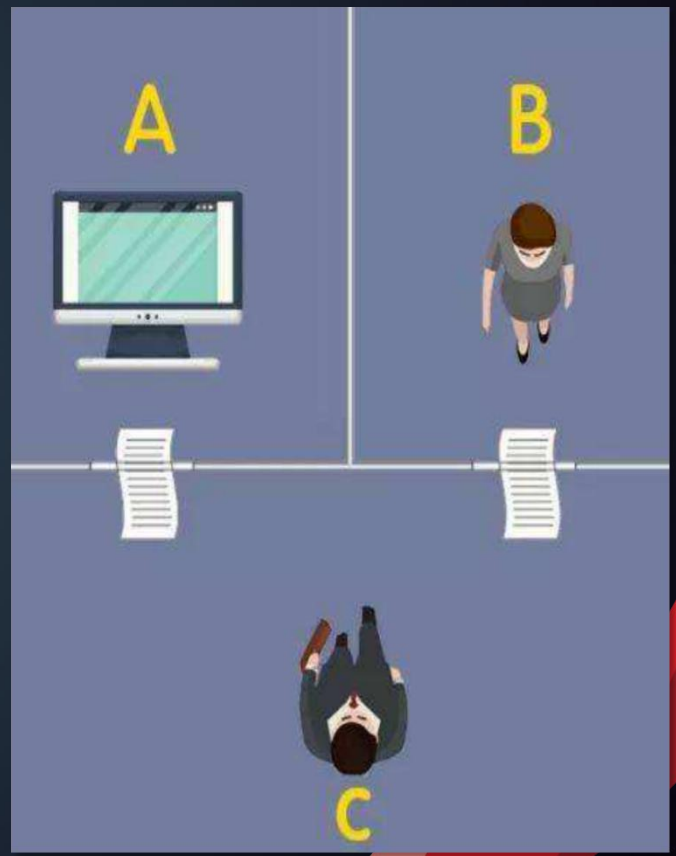
人工智能判定标准：图灵测试

测试目标 设计智能软件与人交流，使人类相信它是人类。

规则判定

- 设定1：将被测试**机器A**与**被测试人类B**完全隔开；
- 设定2：**测试者C**通过通信装置向A和B提问，A和B负责回答；
- 设定3：通信装备不包含任何可用于判断是人还是机器的额外信息；
- 判　定：让多名C参与测试，如果C判断A和B谁是机器人的误判率超过30%，那么这台机器就通过了测试，被认为具有人类智能。

注：俄罗斯人弗拉基米尔·维西罗夫2014年开发的人工智能软件尤金·古斯特曼，成功让人相信它是一个13岁的男孩，首次通过图灵测试的计算机，是人工智能发展的一个里程碑事件。



图灵测试原理示意图

图灵测试相关知识

- 图灵测试，来自于图灵在1950年发表的论文：计算机器与智能。
- 艾伦·麦席森·图灵(1912-1954), 现代计算机科学之父，人工智能之父。主要贡献为图灵机。图灵机是一种抽象的数据模型，定义了现代计算机的核心工作原理。
- 图灵奖，美国计算机协会(ACM, Association for Computing Machinery)1966年为纪念图灵设置了图灵奖。每年仅有1项成果会获奖，是计算机领域的顶级奖项，被誉为计算机界的诺贝尔奖。本世纪获奖内容：2000 姚期智, 通讯复杂性和伪随机数生成理论, 2015 ECDH, 2002 RSA，2018年 深度学习。

人工智能范围定义

- **广义的人工智能**

对人思维过程的模拟，涉及学科包括数学、神经生理学、心理学、计算机科学、信息论、控制论、不定性论、仿生学、社会结构学等，主要实现方式包括结构模拟，制造出“类人脑”的仿生系统，如中科大的类脑项目[1]；或是能模拟，从其功能过程进行模拟，通过不同的方式实现类似功能，如各类仿生系统。

- **狭义的人工智能**

一般指机器学习，即通过建立一定的算法模型，对数据样本集对模型进行训练，从而使模型具备实现具有一定人类智慧能力。在计算机领域，**人工智能与机器学习内涵基本相同。**

[1] 类脑智能技术及应用国家工程实验室, http://leinao.ustc.edu.cn/25858/list.htm

VUL·AI
华云安

02 | 什么是机器学习

机器学习

注：以下分类颗粒大小不相同

K-邻近算法 | 决策树 | 回归算法 | SVM | 关联规则学习

CNN RNN DNN
深度学习
神经网络

贝叶斯分类 | CBR | 逻辑回归 | 降维算法

推荐算法 | 强化学习 | 聚类算法 | 降维算法

机器学习

通常构建的数学模型对输入样本数据进行学习，从而形成并不断提高其对未知样本的处理能力的人工智能方法。

生物仿生学

机械力学

传感技术

...

人工智能

由神经生理学、计算机科学、信息论、控制论、...、仿生学、社会结构学等学科构建的模拟人类行为的综合学科

机器学习定义与示例

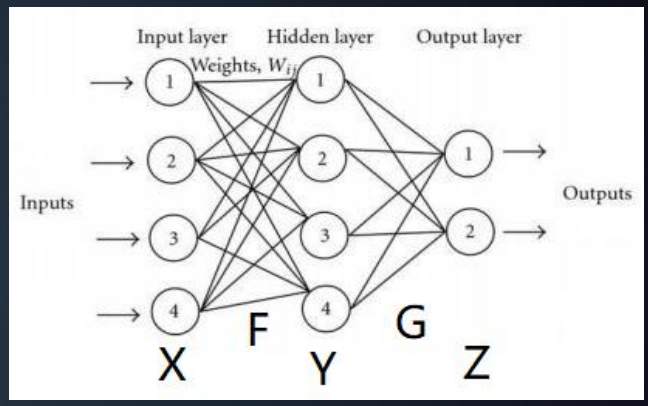
业内定义：给定任务**T**由模型M执行，执行结果的好坏可以用性能指标**P**衡量，执行任务可以获得经验**E**，**若E的更新迭代可以提升M的性能指标P**，那么模型M就是一种具有机器学习能力的模型。

场景	T	M	E	P	解释
人脸识别	从图像中识别出人脸信息	人脸识别模型	面部参数	人脸识别正确率	通执行识别人脸的任务，积累人脸识别信息提高对人脸的识别准确率。
文字翻译	将语言A转换为语言B	语言翻译模型	语言转换权重	翻译准确程度	通过执行文字翻译任务，积累翻译信息，提高翻译识别率。
语音识别	语音信号转为文字	语音识别模型	两种信息的对应关系	文字提取准确率	通过积累人脸识别信息E提高识别率。
对比人类学习（非机器学习）	学习某项知识	人脑思维	对知识的理解	掌握度	通过学习某个知识，提高对知识的理解，从而提高对此知识的掌握程度。

重要特性：在达到模型性能瓶颈前，模型的性能会随着使用的次数变得越来越好。

常见概念

- **监督学习类型**
  - 监督学习: 有标准答案，可以根据答案来学习，如分类问题;
  - 无监督学习: 无标准答案，可以通过制定判断标准来学习，如识别问题。
- **强化学习**
  以奖励的引导学习：按某条件执行，结果正确此条件的权值增加，反之下降。
- **神经网络**
  本质上是多层的函数映射关系，通过多次映射能够实现较为复杂的功能。



- 如 G(F(x)) 就是2层神经网络，第1层 F: X→Y, 第2层G: Y→Z，其中X是输入，Z是输出，Y是隐藏层。
- **深度学习**
  层数较高的神经网络，之所以最后会出现深度学习，本质上是计算机性能的提高，具备了使用更多的层的网络条件，从而显著提高了人工智能的处理能力。

技术路线

技术上游
难度：高

技术下游
难度：低

理论基础

线性代数
微积分
概率论
数值分析
数理统计
...

神经网络　深度学习
强化学习　表示学习
决策树　集群分类
卷积　集成学习
联接主义　KNN
RNN　...

相关学科

AI核心算法

TensorFlow CUDA
CNTK　Singa
Caffe2　**Pytorch**
Spark MLlib　Keras
H2O Apache Mahout
Azure ML Studio
...

底层技术框架

工程实现

语音识别　图像识别
自动驾驶　语义理解
智能家居　无人机
自然语言处理
知识图谱　网络安全
游戏娱乐　语言翻译
...

百度EasyDL, OpenCV，ImageMagic, GraphicsMagick

Jieba, HanLP, NLTK Bert, Ansj, Spark NLP

指纹识别，验证码识别，路径爆破，漏洞智扫，攻防智能决策

中层应用场景开发

银行/微信/支付宝人脸识别
讯飞输入法
小爱同学
指纹解锁
人脸门禁
AI电话应答
搜索推荐
...

上层产品封装

应用领域

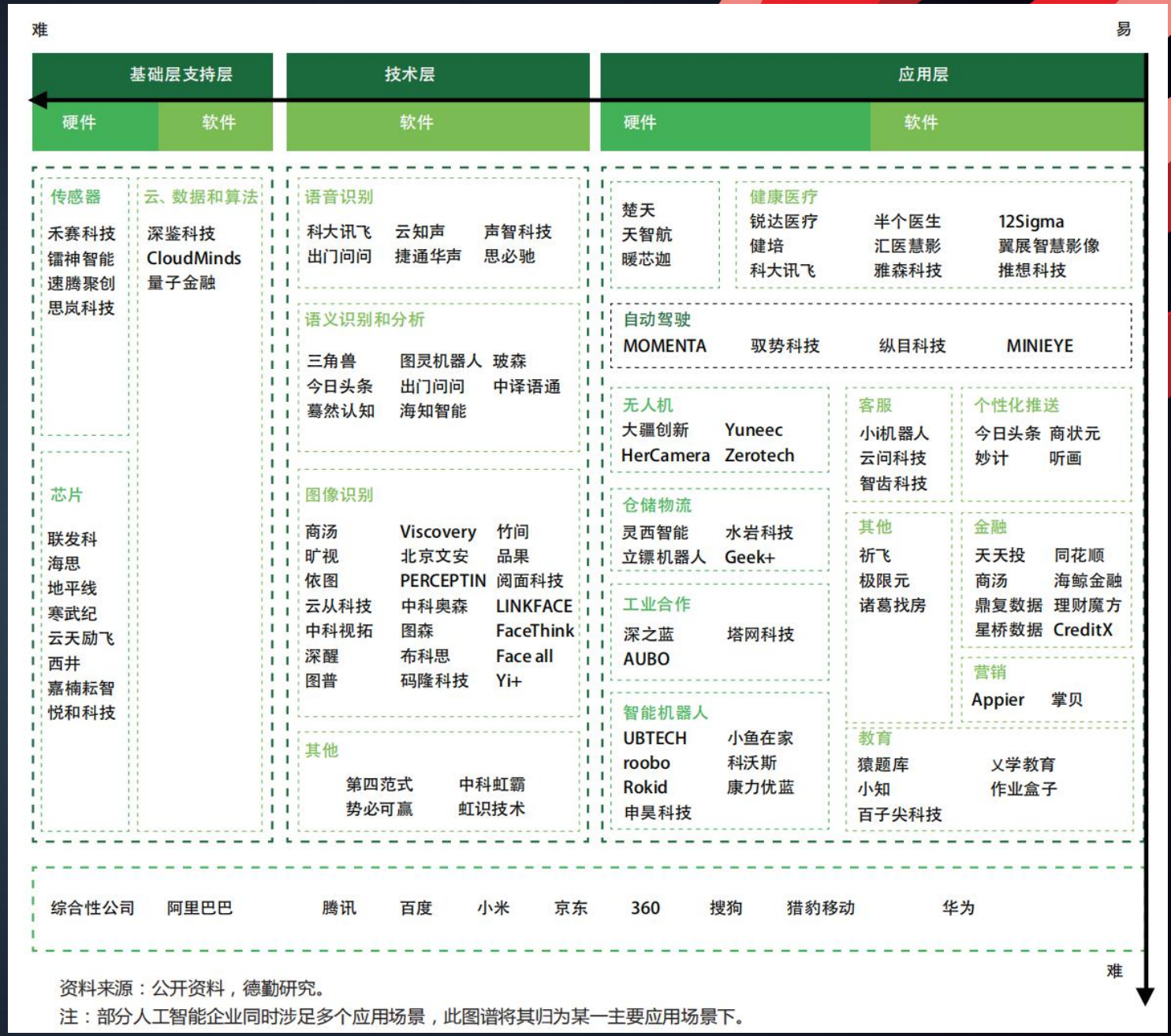
右图为主要应用领域的业内主要知名公司。

主要分为以下几类：

基础支持：底层算法技术；

核心技术：包括语音、图像和语义等算法模型；

应用层面：与各个行业相结合的具体应用场景。



资料来源：公开资料，德勤研究。

注：部分人工智能企业同时涉足多个应用场景，此图谱将其归为某一主要应用场景下。

商汤科技

中国“AI四小龙”——商汤、旷视、云从、依图。

公司简介：商汤科技成立于2014年，由香港中文大学教授汤晓鸥创立，自主研发原创深度学习平台Parrots，主要业务是为不同应用场景提供计算机视觉和深度学习领域的核心算法。

研发团队：截至2021年6月，共有**40**位教授，**250**多位博士，**3593**名科学家和工程师，研发人员占比七成。核心团队由两部分组成：一部分是来自麻省理工学院、香港中文大学、清华大学、北京大学的博士、硕士等；另一部分则是来自微软、谷歌、联想、百度等相关领域的从业者。

团队成就：截至2021年6月，已生成的人工智能模型总数为**22196**个，在各项全球竞赛中已获得**70**多项冠军，发表了**600**多篇顶级学术论文，拥有**8123**项知识产权组合，包括中国**4169**项，海外**3954**项。

融资情况：2021年12月30日，商汤科技正式在港交所挂牌上市，开盘价3.91港元，总市值**1301**亿港元。截止 2022/01/11 16:00，股价为 7.35元，总市值 2446亿。

网络安全领域

网络攻防

防火墙

设备指纹识别

密码爆破

流量攻击检测

内存取证

恶意样本识别

密码爆破

资产探测

恶法域名检测

指纹识别

恶意软件识别

弱口令分析

病毒检测

域名动态生成

智能漏扫

系统异常检测

推荐阅读：A Survey on Machine Learning Techniques for Cyber Security in the Last Decade, 2020.12, IEEE Access.
https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2730527/Hameed%252C%2BIbr.pdf

VUL·AI
华云安

03 | 机器学习应用方法

问题：根据人体尺寸确认性别

基于规则的处理方法

1. 数据定义：选择身高体重作为输入数据；

2. 判断规则：身高在160左右，体重在50公斤判定为女性，其他为男性。

主要问题：不同国家的人使用判断偏差大，

解决方法：需要对身高和体重的范围进行人工配置。

问题：根据人体尺寸确认性别

机器学习方法

1. 选择身高体重作为输入，
2. 使用 身高：单位cm，# 第1维

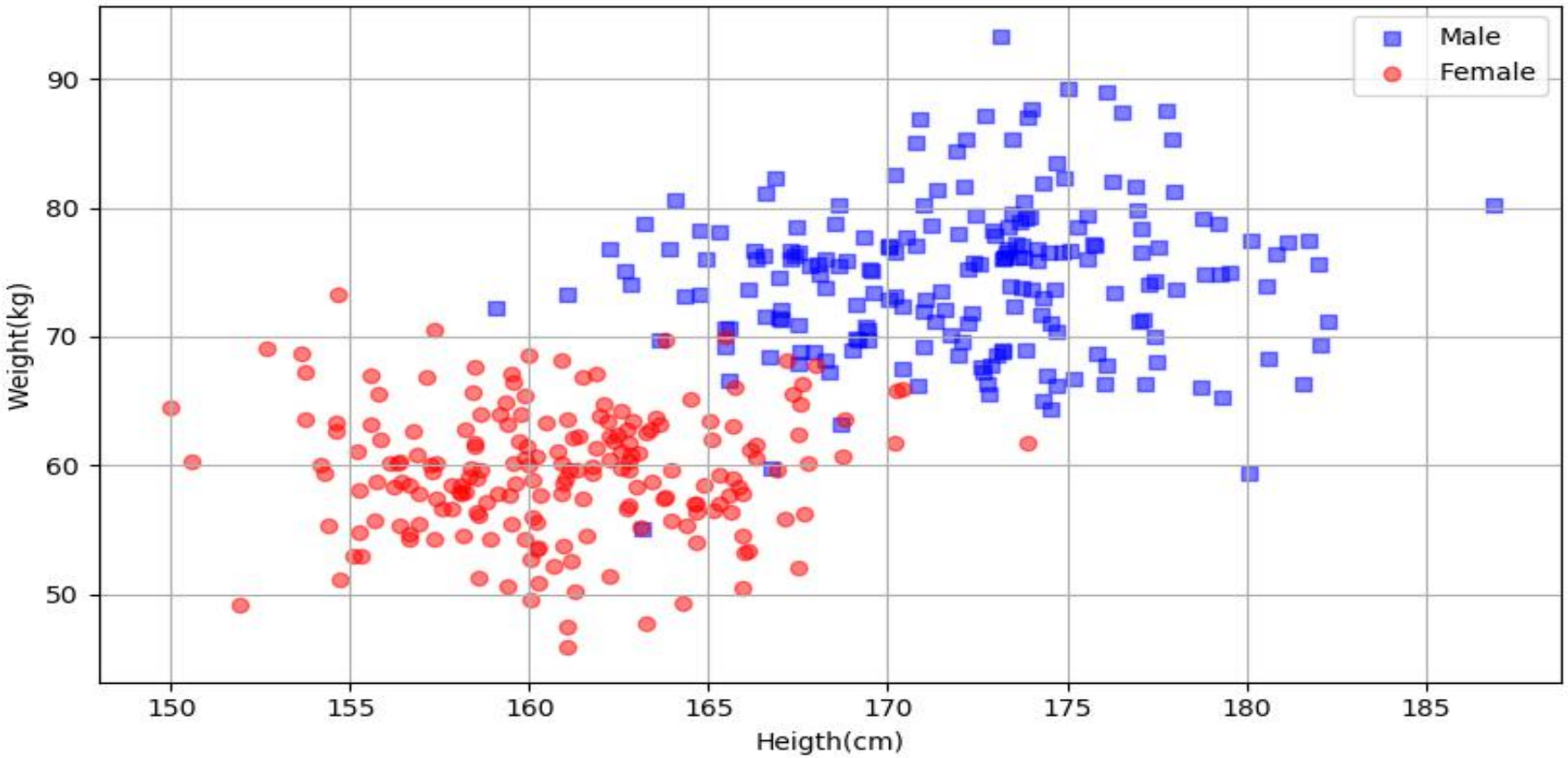
   体重：单位kg，# 第2维

   性别：0=女，1=男。

注：其中只有身高和体重两个特征，因此是二维数据。

建立样本集，如：(174,96;1), (181,87;1), (165,110;0), (158,104;0), ...

问题：根据人体尺寸确认性别



问题：根据人体尺寸确认性别

升维: 二维→五维

身高： 单位cm,

胸围： 单位cm,

腰围： 单位cm,

臀围： 单位cm,

体重： 单位kg,

性别： 0=女， 1=男。

示例样本：

(174,78,75,78,96;1),

(165,86;66,80,52;0)

降维: 五维→四维

身高： 单位cm,

胸围： 单位cm,

腰围： 单位cm,

体重： 单位kg,

性别： 0=女， 1=男。

示例样本：

(174,78,75,96;1),

(165,86,66,52;0)

降维: 四维→三维

身高： 单位cm,

胸腰比： 百分比,

体重： 单位kg,

性别： 0=女， 1=男。

示例样本：

(174,1.04,96;1),

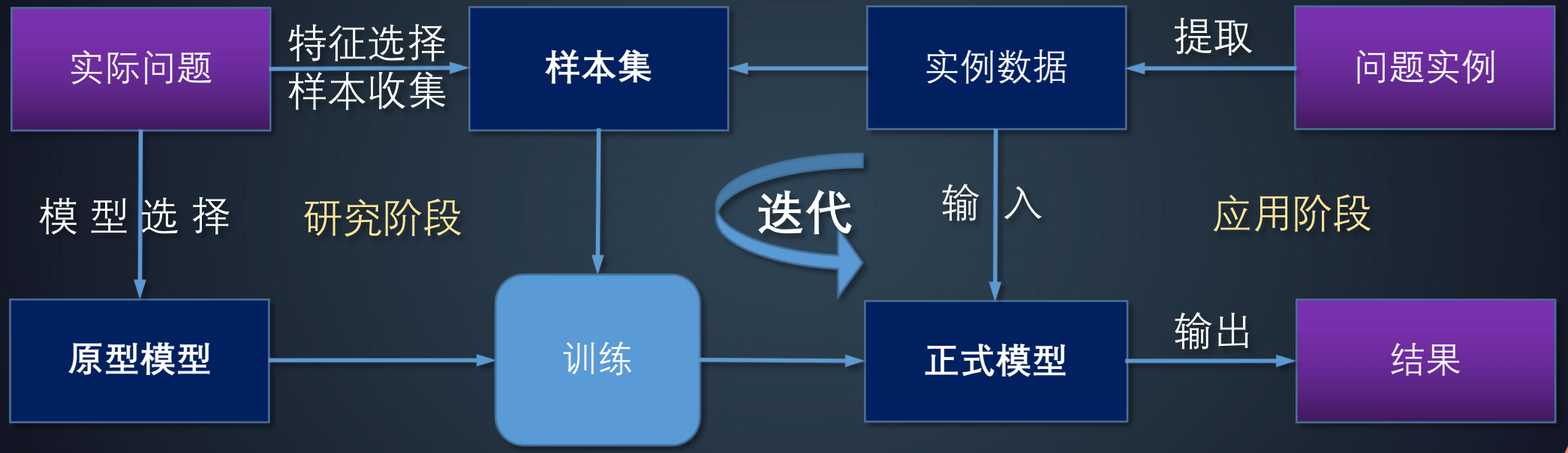
(165,1.33,52;0)

机器学习模型建立流程

1. 问题分析：对问题进行分析，理解问题的本质及主要难点；
2. **特征选择**：选择合适的特征，如选择漏洞影响力、利用难度、权限要求等评估漏洞可利用性；
3. 模型选择：选择能够有效表示问题的模型，如KNN、SVM、回归、决策树等；
4. **样本准备**：根据模型的要求准备数据，获得方式可以下载、购买、自建等；
5. 模型训练：使用准备好的样本对模型进行训练，使模型初步具备应用能力；
6. 实战应用：通过部署和接口调用等形式， 在实战中使用模型，此时模型的能力还有待提高；
7. 迭代优化：通过实战收集的数据，进一步提高模型的性能。

注：标红项为机器学习应用难点。

机器学习应用过程



机器学习一般处理流程示意图

核心价值

1. 机器学习模型

理论模型与实战应用差别很大，需要长期进行模型的研发和改进，从模型**数量**和**质量**两方面提高。

2. 模型学习样本

样本是模型的核心基础，必需以多种方式不断积累不同模型的样本数量。

3. 丰富经验的专业开发人员

需要有一定数据的高水平从业人员，通过学习、交流、试错、实战不断提高人工智能应用水平。

VUL·AI
华云安

04 常见问题

人类学习VS机器学习

	人类学习	机器学习
主体	大脑	计算机
输入	教材与习题	数值样本
方式	教学与复习	样本训练
主体能力	智力水平	计算性能
如何学习	学习方法	算法模型
学习效果	测试问答	性能指标
学习收获	知识经验	模型权值
能力变化	可以通过不断学习提高学习能力	可以通过新数据的迭代提高模型水平

自动化 VS 智能化

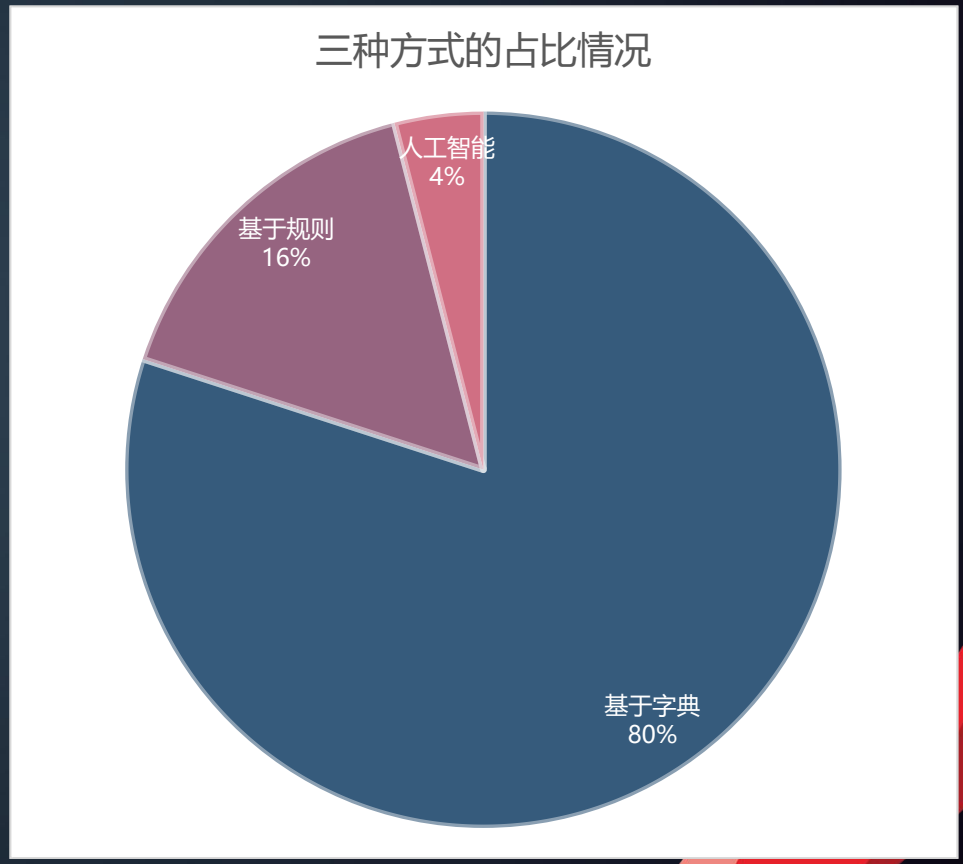
	自动化	智能化
实现方式	基于运行规则	基于智能模型
应变能力	无	有
实现难度	低	高
排错难度	低	高
处理能力	中	强
性能要求	低	高
稳定性	高	低
适应性	人工通过参数手工调节	可以根据新数据自适应

案例：口令爆破--猜出目标账号的密码

Level 1: 基于字典的查表法 (80%: 大部分情况)

Level 2: 基于密码规则 (16%: 剩余20%的80%)

Level 3: 基于人工智能的推断法 （4%: 剩余情况）



注： 占比数据为实际的粗略的估计

VUL·AI
华云安

Q&A 欢迎提问