



# Python程序设计（混合式）



杨尚东

南京邮电大学计算机学院，数据科学与工程系

[shangdongyang.github.io](https://shangdongyang.github.io)

2023/9/8

# 目录

- **开场白**
- **Python简介**
- **人工智能简介**
- **Python开发环境安装**

# 开场白

## 口 上课情况

- ✓ 上课时间：周五下午，6-7节，仙林校区，教3-313
- ✓ 答疑时间：周五下午，8-9节，计算机学科楼，418室
- ✓ 教材信息：陈景强等. Python编程及人工智能应用
- ✓ 课程总评：平时成绩（40%）+期末考试（60%，闭卷）
- ✓ 课程交流：课程邮箱（用于接收作业）[njupt\\_python@163.com](mailto:njupt_python@163.com)
- ✓ QQ群：920825955



[shangdongyang.github.io/course/python](http://shangdongyang.github.io/course/python)

混合式48学时=24理论+8实验+16自学

# 开场白

## □ 课程内容

- ✓ 第1章：人工智能与Python概述
- ✓ 第2章：Python语言基础 **(实验一)**
- ✓ 第3章：线性回归及Python实现
- ✓ 第4章：逻辑斯蒂分类及Python实现 **(实验二)**
- ✓ 第5章：最大熵模型及Python实现 **(自学4学时)**
- ✓ 第6章：K-近邻分类与K-均值聚类 **(实验三)**
- ✓ 第7章：朴素贝叶斯分类及Python实现 **(自学4学时)**
- ✓ 第8章：决策树及Python实现 **(自学4学时)**
- ✓ 第9章：神经网络及Python实现 **(实验四)**
- ✓ 第10章：在图形识别的应用案例 **(自学4学时)**



Python基础 → 人工智能基础 → 小实验 → ? ? ?

# 开场白

口 课后作业：中国大学MOOC

✓ 浏览器网址访问：

<https://www.icourse163.org/spoc/course/NJUPT-1464077167>

✓ 昵称格式：**NJUPT学号姓名**，如NJUPTB20030925徐震

✓ 认证为南邮学生才可访问

✓ **一共6次章节作业**

✓ MOOC成绩由两部分组成：

1. 单元测验

2. 编程作业：**需要进行互评才有成绩**

✓ 注意每次作业提交的截止日期，过后不能再补

**实践、认识、再实践、再认识.....**

**毛主席《实践论》**

# 开场白

## □ 4次实验安排

- ✓ 实验一：Python语言基础实验

 第7周周五8-9节，计算机学科楼地点待定

- ✓ 实验二：线性回归和逻辑斯蒂分类实验

 第11周周五8-9节，计算机学科楼地点待定

- ✓ 实验三：K-近邻分类与K-均值聚类实验

 第12周周五8-9节，计算机学科楼地点待定

- ✓ 实验四：神经网络实验

 第14周周五8-9节，计算机学科楼地点待定

## □ 实验要求

- ✓ 提前预习、设计代码、**独立完成实验**（编码、调试、运行成功、检查）
- ✓ **每次实验一份电子档报告**，不交纸质
- ✓ 实验成绩根据实验**准备、运行、报告**等综合打分

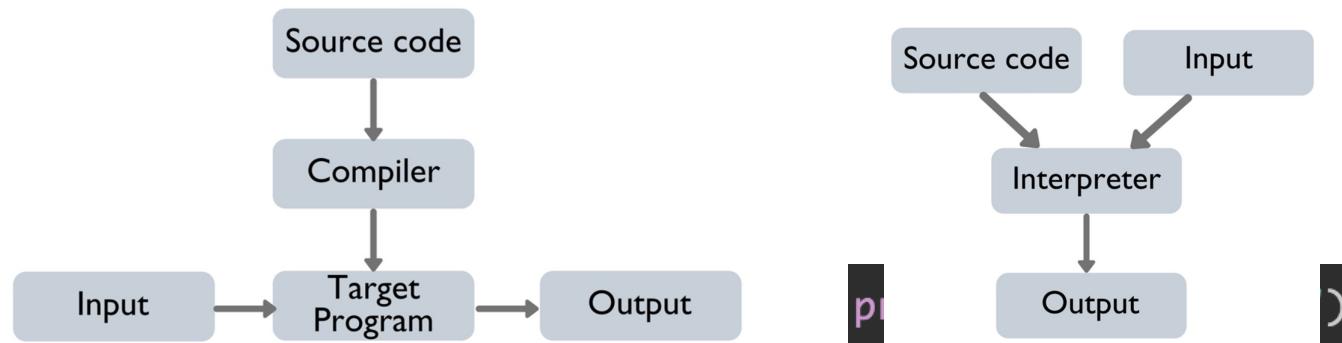
# 目录

- 开场白
- Python简介
- 人工智能简介
- Python开发环境安装

# Python简介

## □ Python (<https://www.python.org/>)

- ✓ Python诞生于1989年的圣诞节前后
- ✓ 吉多·范罗苏姆 ([Guido van Rossum](#))
- ✓ BBC的show: “**Monty Python's Flying Circus**”



- ✓ Python0.9 (1991, Single) ...> Python1.0 (1994, Mailing List)
- ✓ Python2.0 (2000, SourceForge) ...> Python3.0 (2008, Unicode)

 **Dropbox**

 **facebook**

 **豆瓣**  **douban**

 **YouTube**

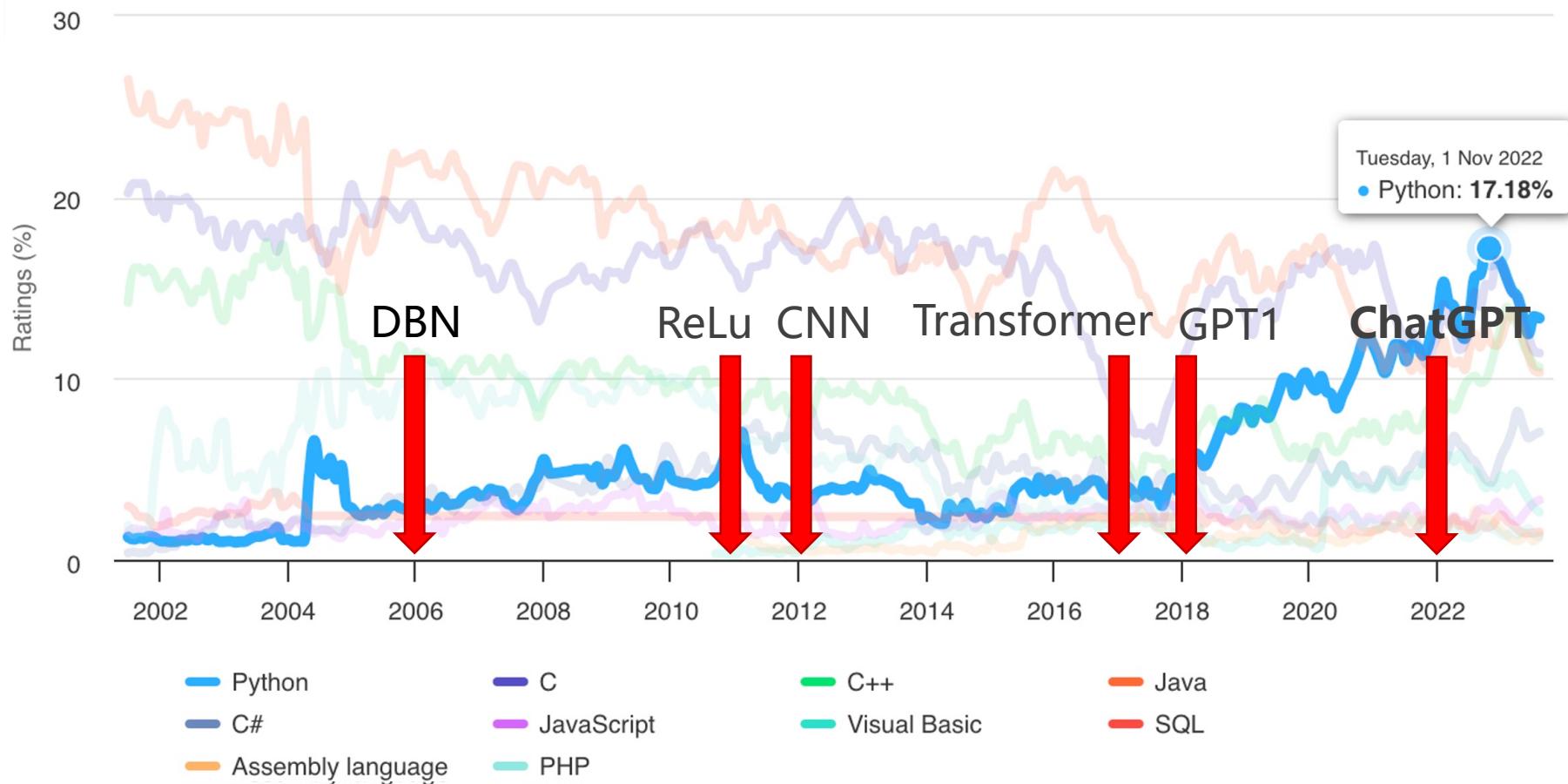
 **Instagram**

 **知乎**  
有问题上知乎

# Python简介

## TIOBE Programming Community Index

Source: www.tiobe.com



# Python简介

## □ Python

- ✓ 简便、跨平台
- ✓ 强大的社区
- ✓ 丰富的库与框架
- ✓ 多用途

## □ 人工智能

- ✓ 数字化社会（第三次工业革命）
- ✓ 各个行业都需要人工智能来实现更自动、智能的应用
- ✓ 探索智慧的原理和极限

## □ Python与人工智能

- ✓ 数据科学：numpy, scipy, cvxopt...
- ✓ 机器学习：sklearn, xgboost...
- ✓ 深度学习：tensorflow, pytorch...

# Python简介

## □ Python语言比C语言容易学

- ✓ 本校将Python语言作为文科专业的入门编程语言，而C语言一般只给工科专业开设
- ✓ 美国的常春藤联盟基本都是以Python为入门
- ✓ 在学习Python的过程中，可以忽略数组越界、内存回收、动态内存分配等细节，而更关注计算机科学算法的本质

## □ 有C语言基础，学习Python语言更加容易，因此本门课通过7个学时讲解Python语言，其它课时讲解人工智能算法的Python实现

## □ 编程语言的基本思想都是相通的

- ✓ 建议可以用“边用边学”的方式来学习

# Python简介

## □ Python语言是一种“解释型”语言

- ✓ Python需要一个解释器

✓ 在运行的时候将程序翻译成机器语言。解释型语言的程序不需要在运行前编译，在运行程序的时候才翻译，专门的解释器负责在每个语句执行的时候解释程序代码。这样解释型语言每执行一次就要翻译一次，效率比较低

- ✓ C语言是“编译型”语言

- ✓ 两个不兼容的版本：Python2.x、Python3.x

## □ Python

- ✓ 动态类型、脚本式语言、引用计数（内存管理）

## □ Python语言被称为“人工智能语言”

- ✓ NumPy、SciPy、Matplotlib

- ✓ Scikit-learn、PyTorch、Tensorflow

- ✓ NLTK、OpenCV、SpeechRecognition



# 目录

- 开场白
- Python简介
- 人工智能简介
- Python开发环境安装

# 人工智能简介

- 人工智能（Artificial Intelligence），简称为AI
- 是一门研究、开发用于**模拟、延伸和扩展**人类智能的理论、方法、技术及应用系统的技术科学，已成为计算机科学的一个重要分支。
- 根据维基百科的定义，人工智能所谓的“智能”，指的是机器的智能，区别于人和动物展示的自然智能。
- 人工智能是以人类行为标准进行猜想、假设、研究、实验而造就出来的一种类人类智能。

# 人工智能简介

口二十世纪五六十年代，逐渐孕育出人工智能概念，三位“人工智能之父”：

- ✓ 1950年，**马文·明斯基** (Marvin Minsky)，建造世界上第一台神经网络计算机
- ✓ 1950年，**阿兰·图灵** (Alan Turing)，提出“图灵测试”
  - …> 指测试者与被测试者（一个人和一台机器）隔开的情况下，通过一些装置（如键盘）向被测试者随意提问。进行多次测试后，如果机器让平均每个参与者做出超过30%的误判，那么这台机器就通过了测试，并被认为具有人类智能
- ✓ 1956年，**约翰·麦卡锡** (John McCarthy)，提出“人工智能”一词



2006年达特茅斯会议50年纪念，当事人重聚达特茅斯（左起：摩尔、麦卡锡、明斯基、赛弗里奇、所罗门诺夫）

# 人工智能简介

- 人工智能的第一次高峰：在1956年的达特茅斯会议之后，人工智能迎来了第一段发展高峰期，计算机被广泛应用于数学和自然语言领域，用来解决代数、几何和英语问题
- 人工智能的第一次低谷：在20世纪70年代，人工智能进入了一段痛苦而艰难岁月。
  - ✗ 计算机性能不足
  - ✗ 问题的复杂性
  - ✗ 数据量不足

# 人工智能简介

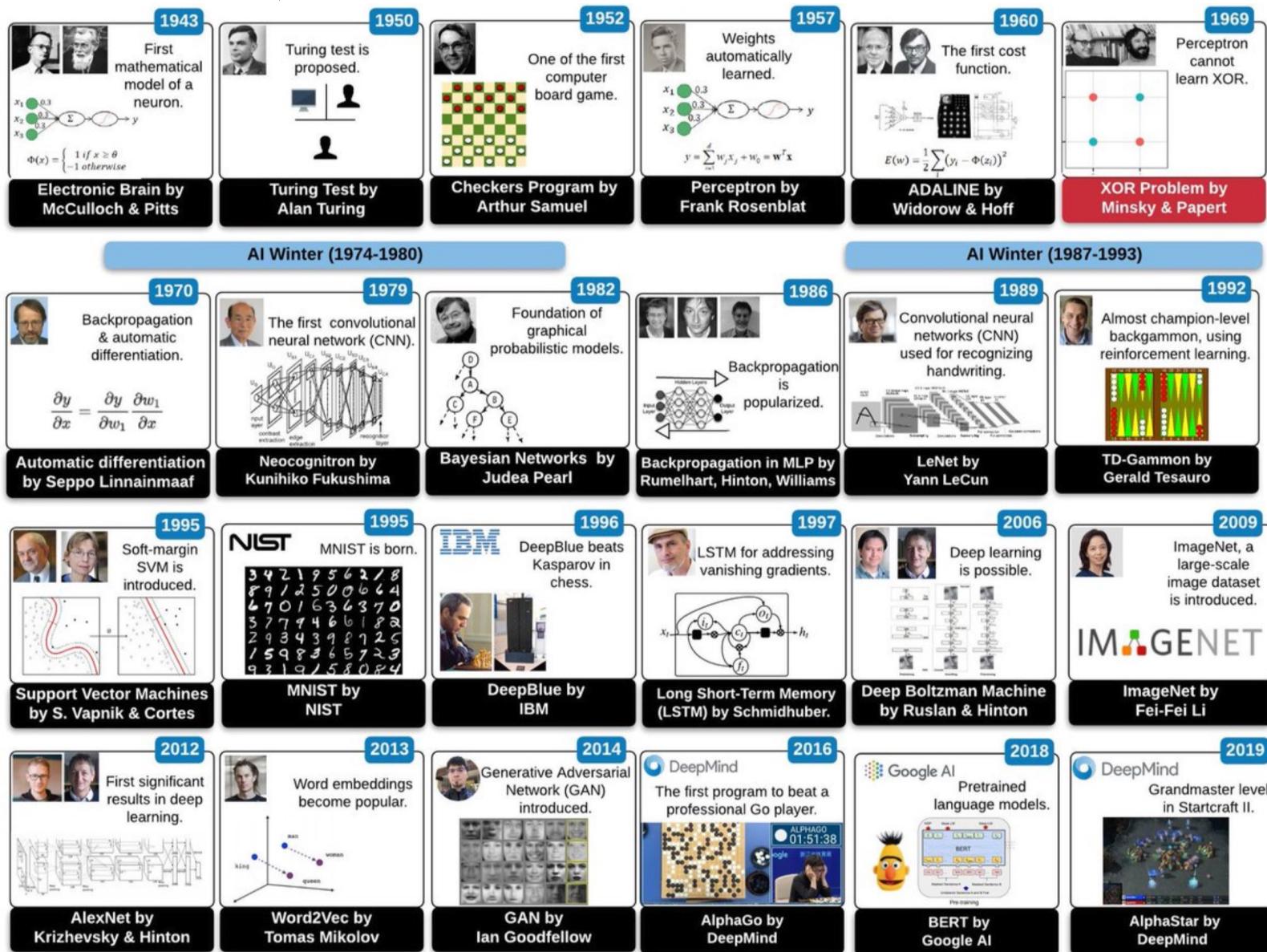
- 口 人工智能的第二次高峰：1980年，卡内基梅隆大学为数字设备公司设计了一套名为XCON的“专家系统”。这是一种采用人工智能程序的系统，可以简单理解为“知识库+推理机”的组合，具有完整的专业知识和经验。
- 口 人工智能的第二次低谷：在维持了仅仅7年之后，这个曾经轰动一时的人工智能系统就被宣告结束。到1987年，苹果和IBM公司生产的台式机性能均超过了Symbolics等厂商生产的通用计算机。从此，专家系统风光不再。

# 人工智能简介

口 人工智能的第三次高峰：2006年Hinton等人提出的深度学习技术并获得成功，人工智能再一次掀起了全球的热潮。

- ✓ 2011年，IBM开发的人工智能程序“沃森”（Watson Jeopardy）参加一档智力问答节目并战胜了两位人类冠军
- ✓ 2015年，在ImageNet图像识别竞赛中，基于深度学习的人工智能算法在准确率方面第一次超越了人类肉眼，人工智能实现了飞跃性的发展
- ✓ 2016年，AlphaGo战胜人类围棋冠军
- ✓ 2017年，AlphaGoZero（第四代AlphaGo）战胜了在人类高手看来不可企及的第三个版本“大师”
- ✓ 2022年，OpenAI开放了ChatGPT

# 人工智能简介



# 人工智能简介

## 口 人工智能的三大学派

✓ 符号主义 (Symbolism) :

…> 一种基于逻辑推理的智能模拟方法

…> 又称逻辑主义、心理学派或计算机学派。

…> 启发式算法→专家系统→知识工程，知识图谱

✓ 连接主义 (Connectionism) :

…> 又称仿生学派或生理学派

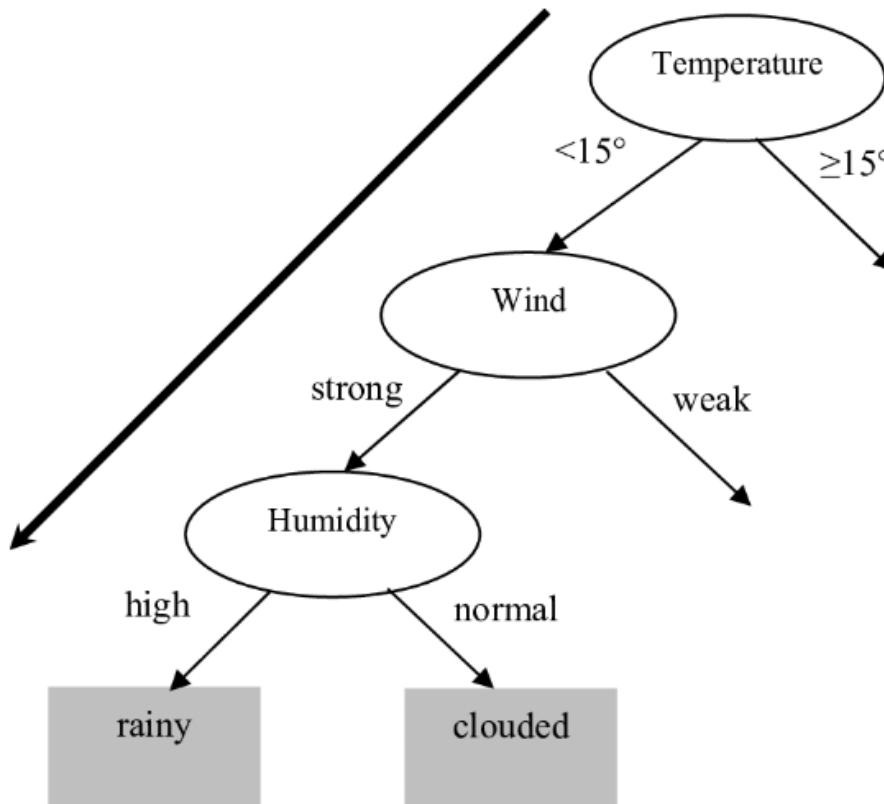
…> 是一种基于神经网络及网络间的连接机制与学习算法的智能模拟方法，比如深度学习

✓ 行为主义 (Actionism) :

…> 强化学习：“感知——行动”的行为智能模拟方法

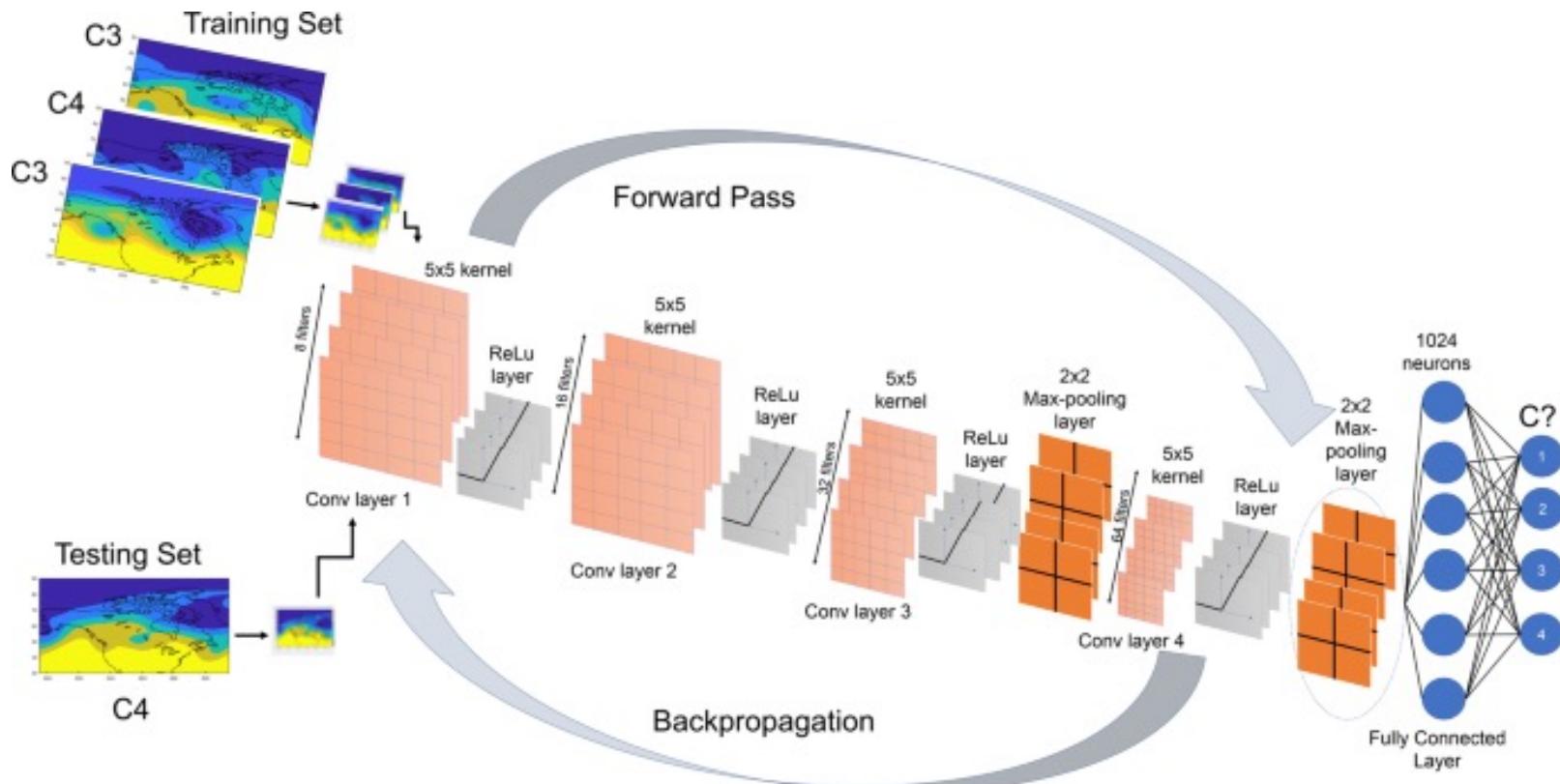
# 人工智能简介

□ 符号主义 (Symbolism) : 又称逻辑主义、计算机学派, 认为认知就是通过对有意义的表示符号进行推导计算, 并将学习视为逆向演绎, 主张用显式的公理和逻辑体系搭建人工智能系统。如用决策树模型输入业务特征预测天气:



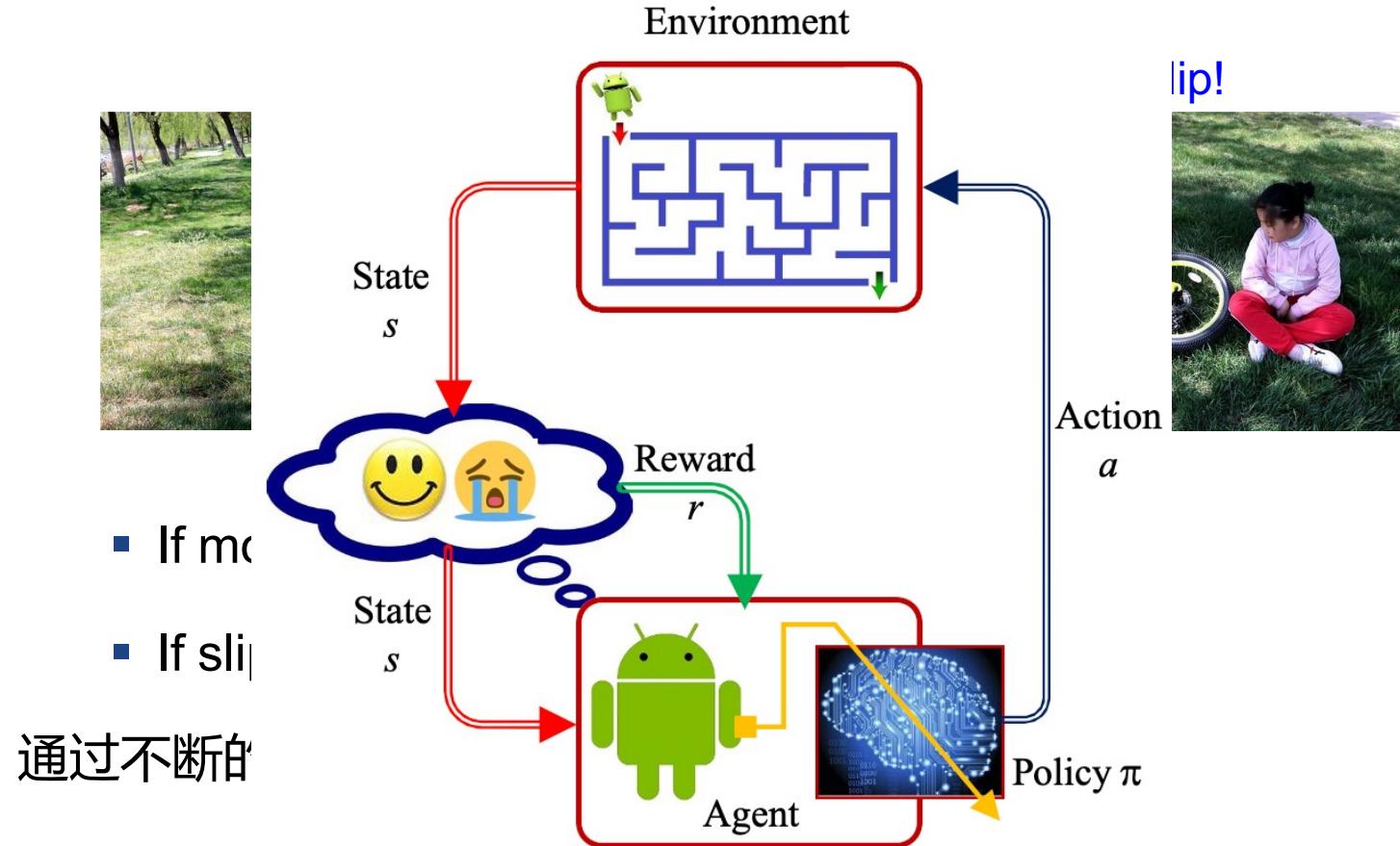
# 人工智能简介

□ 连接主义 (Connectionism) : 又叫仿生学派，笃信大脑的逆向工程，主张是利用数学模型来研究人类认知的方法，用神经元的连接机制实现人工智能。如用神经网络模型输入雷达图像数据预测天气：



# 人工智能简介

□ 行为主义 (Actionism) : 以控制论及感知-动作型控制系统原理模拟行为以复现人类智能



# 人工智能简介

## □ 强人工智能：

- ✓ 具备与人类同等或者超越人类智慧，具有心智和意识，能根据自己意图开展行动
- ✓ 类人的人工智能，即机器的思考和推理就像人的思维一样
- ✓ 非类人的人工智能，即机器产生了和人完全不一样的知觉和意识，使用和人完全不一样的推理方式
- ✓ 图灵测试

## □ 弱人工智能：面向特定的问题或任务

- ✓ 计算机视觉
- ✓ 自然语言处理
- ✓ 语音识别等

# 人工智能简介

## □ 图灵测试

### □ 混淆视听的测试

- ✓ 「你家养了猫？」
- ✓ 「哪里买的毯子？」
- ✓ 「你今天心情怎么样？」
- ✓ 「 $9566 \times 55523$  等于几？」

### □ 「温诺格拉德」测试

“镇上的议员们拒绝给愤怒的游行者提供游行许可——因为他们担心会发生暴力行为”——是谁在担心暴力行为？

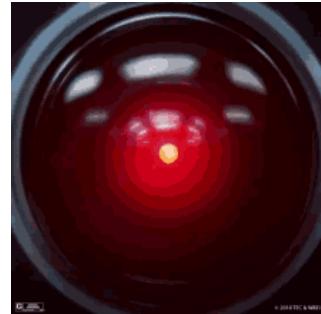
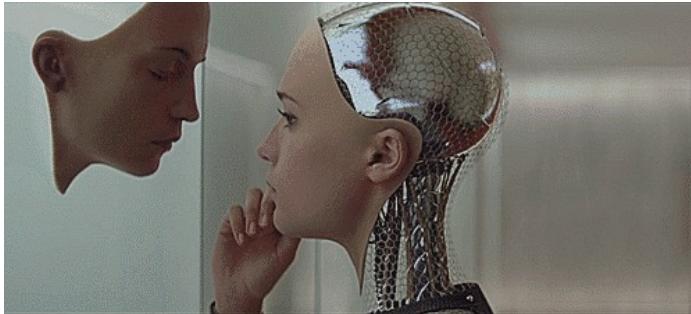
我：猫坐在毯子上，因为它很温暖。什么很温暖？

Siri：有趣的问题。

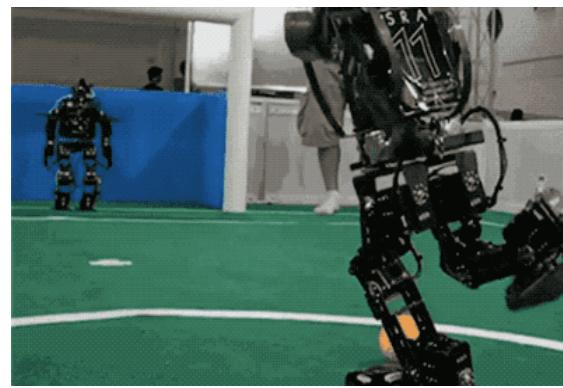
B、愤怒的游行者

# 人工智能简介

## □ 我们以为的人工智能



## □ 实际中的人工智能



# 人工智能简介

- 南京大学**周志华**教授认为强人工智能是**不可实现的**
  - ✓ “从技术上来说，主流人工智能学界的努力从来就不是朝向强人工智能，现有技术的发展也不会自动地使强人工智能成为可能；即便想研究强人工智能，也不知道路在何方；即便强人工智能是可能的，也不应该去研究它。”
- 北京大学**黄铁军**教授则认为强人工智能是**可实现的**
  - ✓ “强人工智能是可以做出来的，而且可以在有生之年做出来，30年之内是可以做出来的；把最难的问题放在最后，不要把最难的问题放在第一步，如果放在第一步就是哲学家，就是争论，上来就开始给你很多的一套一套的，但是没用，有用是你去实践，去做。”
- “我们不能既承认人类智力的优越性，又否认可以通过人类智力实现人工智能”

# 人工智能简介

## 口 机器学习 (Machine Learning)

- ✓ 从**已知数据**中**学习**其中蕴含的规律或者规则并利用这些规律和规则推广到未来的新数据
- ✓ 按方法分类，机器学习可分为
  - 有监督学习：需要大量有标注数据
  - 无监督学习：无需标注，但是计算难度很大
  - 半监督学习：为小部分样本提供真实标注
- ✓ 按任务分，机器学习可分为
  - 回归（有监督）：预测的数据对象是连续值。教材第3章
  - 分类（有监督）：预测的数据对象是离散值。教材第4-7章
  - 聚类（无监督）：在数据中寻找隐藏的模式或分组。教材第6章
- ✓ 线性回归、逻辑斯蒂分类、K近邻分类、K均值聚类、贝叶斯分类、决策树

# 人工智能简介

## □ 深度学习 (Deep Learning) :

- ✓ 深度学习是一种实现机器学习的技术，并不是一种独立的学习方法
- ✓ “深度”的含义
  - …> 深度等于更大规模的网络
  - …> 深度等于更抽象的特征
- ✓ 深度学习模型是以人工神经网络为基础
  - …> 人工神经网络训练速度慢、容易过拟合、经常出现梯度消失以及在网络层次比较少的情况下效果并不比其他算法
  - …> 随着计算性能的提升和互联网时代数据量的爆炸式增长，深度学习效果大幅提升
  - …> 三大要素：硬件、数据、算法



深度学习“三巨头”，2018年图灵奖获得者

# 人工智能简介

## 口 人工智能、机器学习、深度学习的关系：

- ✓ 机器学习是一种实现人工智能的方法
- ✓ 深度学习是一种实现机器学习的技术
- ✓ 机器学习和深度学习不是人工智能的全部

## 口 三者关系



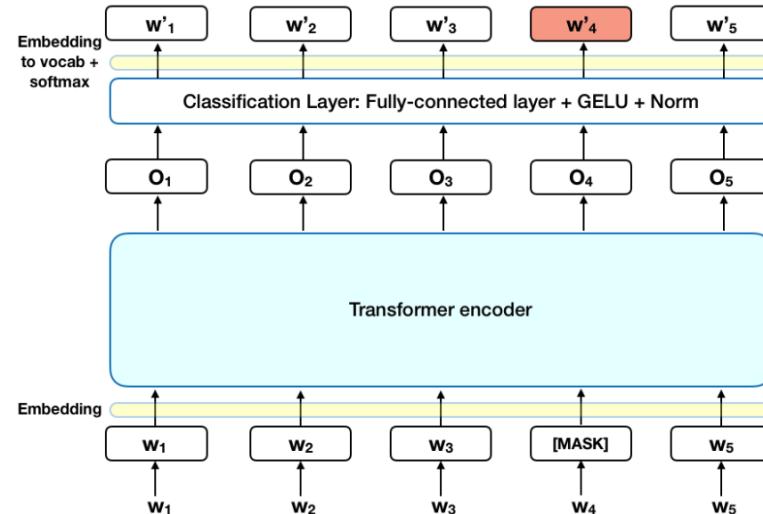
# 人工智能简介

- 人工智能还从各个方面模拟人的智能。
- 人最基本的智能包括语言、视觉、听觉等，让机器分别拥有这些智能，便产生了人工智能技术的应用分支领域
  - ✓ 自然语言处理
  - ✓ 计算机视觉
  - ✓ 语音识别
  - ✓ 推荐系统
  - ✓ .....

# 人工智能简介

## 口 自然语言处理

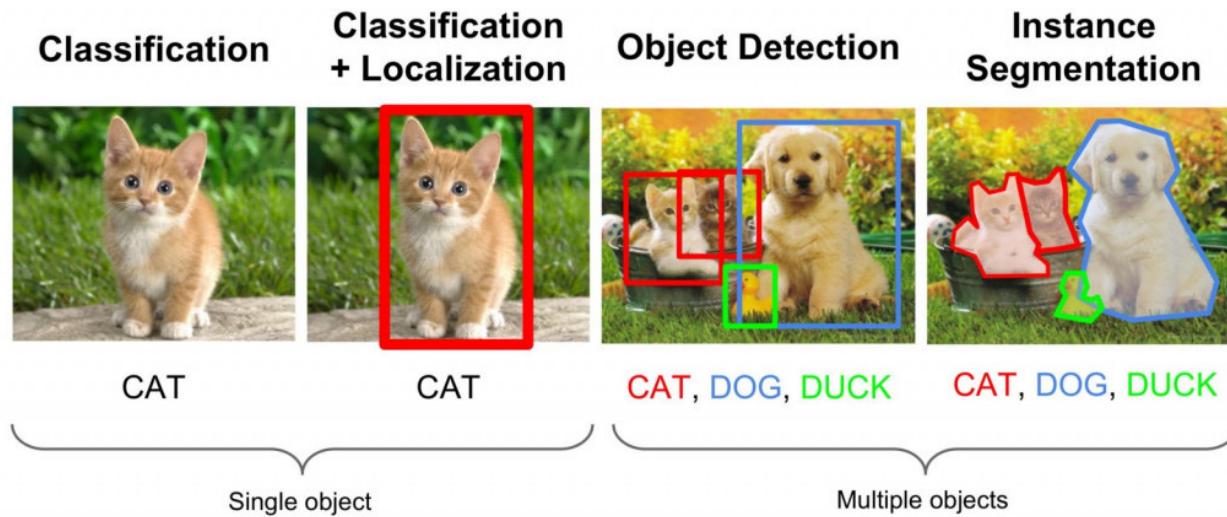
- ✓ Natural Language Processing, 简称NLP
- ✓ 利用计算机对自然语言进行智能化处理。
- ✓ 研究目标是让机器理解并生成人类语言，用自然语言与人类平等流畅地沟通交流，最终拥有“智能”
- ✓ 基础自然语言处理技术主要围绕语言不同层级展开
  - …> 包括音位、形态、词汇、句法、语义、语用、篇章
- ✓ 自然语言处理任务
  - …> 机器翻译、人机对话
  - …> 自动问答、文档摘要
- ✓ 目前，NLP的主流方法是基于深度学习的方法



# 人工智能简介

## □ 计算机视觉

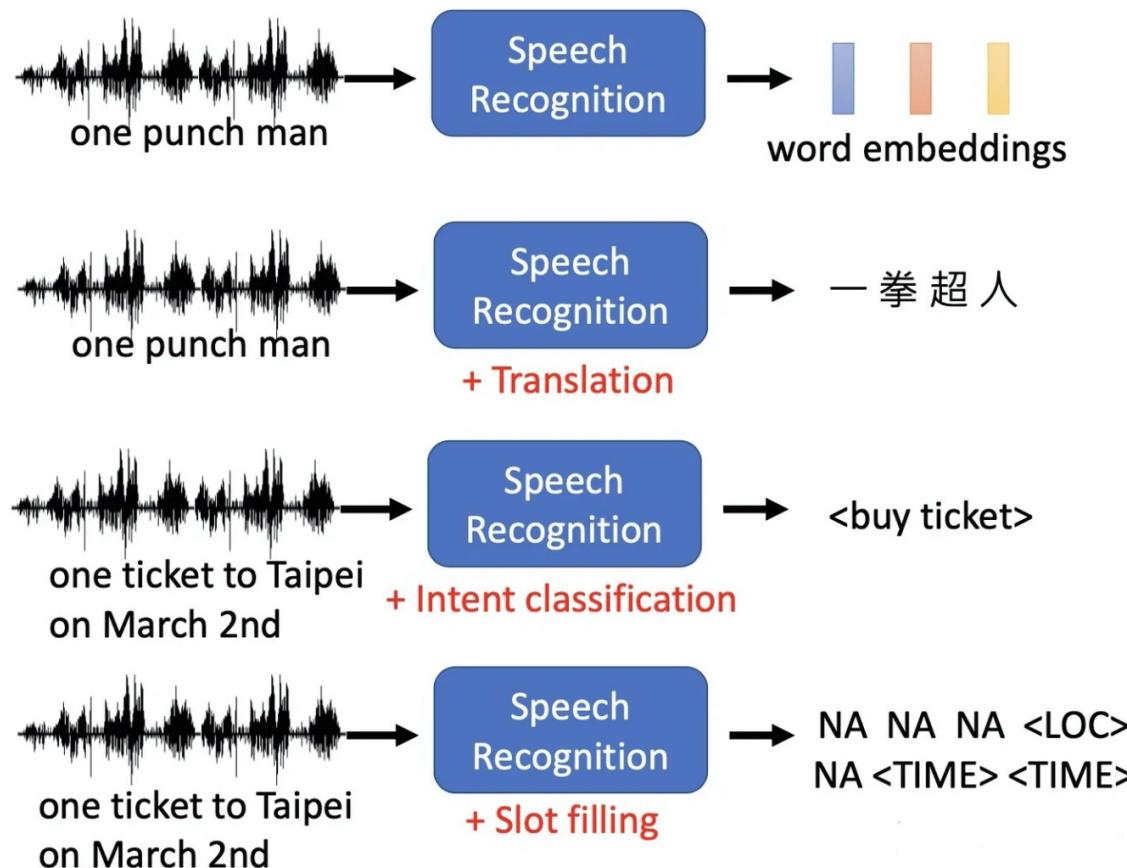
- ✓ Computer Vision, 简称CV。是人工智能领域中研究如何使机器“看”的科学
- ✓ 主要任务
  - 人脸识别、图片问答、图像识别、目标检测和跟踪等
- ✓ 当前计算机视觉的研究主要是基于深度学习技术
  - Fast RCNN、Faster RCNN、Yolo等



# 人工智能简介

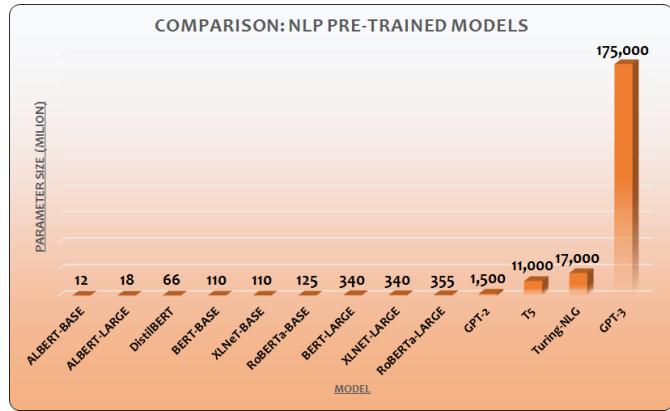
## □ 语音识别

□ Automatic Speech Recognition, 简称ASR, 把人类的语音信号转  
变为想要的文本或者命令

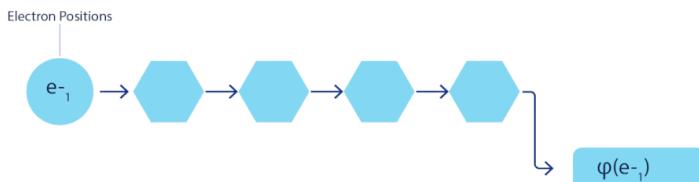


# 人工智能简介

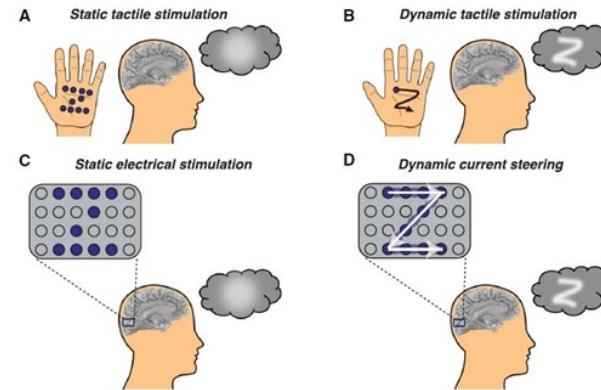
## 口 人工智能最新进展



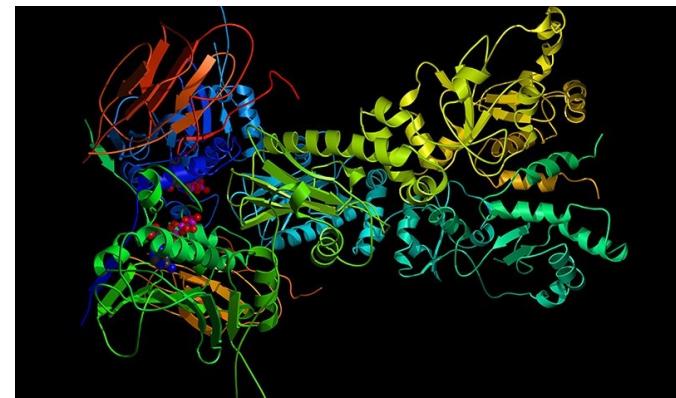
GPT-3: 1750亿参数，训练所用的数据量达到45TB，训练费用超过1200万美元。



FermiNet: 来近似计算薛定谔方程，在精度和准确性上都满足科研标准



美国贝勒医学院利用动态颅内电刺激新技术，用植入的微电极阵列构成视觉假体，在人类初级视皮层绘制W、S和Z等字母的形状，成功地能够让盲人“看见”了这些字母。



AlphaFold2: 其准确性可以与使用冷冻电子显微镜等实验技术解析的蛋白质3D结构相媲美

# 人工智能简介

## □ ChatGPT

✓ 李宏毅 (Hung-yi Lee)

✓ <https://speech.ee.ntu.edu.tw/~hylee/ml/2023-spring.php>

# 人工智能简介

## □ AGI

- ✓ 人们把具备与人类同等智慧、或超越人类的人工智能称为通用人工智能 (artificial general intelligence)

## 1. 强化学习

- ✓ 环境—智能体—反馈—学习

## 2. 自监督学习

- ✓ 自监督学习算法不依赖注释，而是通过暴露数据各部分之间的关系，从数据中生成标签，这一步骤被认为是实现人类智能的关键。
- ✓ Lecun：“我们作为人类学习的大部分知识和动物学习的大部分知识都是在自我监督的模式下，而不是强化的模式。基本上是通过观察世界，并与之进行一些互动，主要是以独立于测试的方式进行观察，这是我们不知道如何用机器复制的（学习）类型。”

## 3. 其它

# 人工智能简介

## 口 人工智能不足的地方

- ✓ 小样本学习、零样本学习
- ✓ 推理
- ✓ 常识

# 人工智能简介

口 人工智能正在改变多个关键行业，并且应用范围在迅速扩展，提供影响企业利润的实质好处

- ✓ 自动驾驶：特斯拉的自动驾驶汽车
- ✓ 聊天机器人：Siri、Cortana和Alexa
- ✓ 社交媒体：利用深度学习和自然语言处理来过滤令人反感的内容
- ✓ 艺术创造
- ✓ 市场营销：广告投放
- ✓ 金融业：依靠计算机和数据科学家来确定市场的未来模式
- ✓ 智慧农业：帮助农民找到更有效的方法来保护农作物免受杂草侵害
- ✓ 智慧医疗：跟踪人们在养老院、家庭护理的表现
- ✓ 游戏行业：DeepMind 基于AI开发的AlphaGo
- ✓ 太空探索：嫦娥五号通过人工智能自主决策

# 人工智能简介

技术及应用标准与规范

科研学术机构与各层次人才

通信与信息网络

数据存储设施

物联网  
与微型传  
感器

2020年中国人工智能产业图谱



# 人工智能简介

## 口 我国人工智能产业规模大

- ✓ 到2020年，中国人工智能产业规模超过1500亿元，带动相关产业规模超过1万亿元

## 口 需求企业薪资高

- ✓ 百度、京东、谷歌、讯飞、阿里巴巴、腾讯



# 目录

- 开场白
- Python简介
- 人工智能简介
- Python开发环境安装

# Python开发环境安装

## □ 安装Python

- ✓ 官方下载网址: <https://www.python.org>
- ✓ 演示下载和安装最新版Python

## □ 运行Python

- ✓ 交互模式
  - …> 打开命令行提示符界面，输入python指令后
  - …> 出现“>>>”提示符，输入语句print("hello world")
- ✓ 文件模式
  - …> 新建文本文件，改后者名为“.py”
  - …> 编辑文件内容为print("hello world")，通过右键菜单运行

# Python开发环境安装

## □ Pip指令

- ✓ 安装指令: pip install [-i 安装源] 包名称
- ✓ “-i 安装源” 可指定国内源以提升下载速度
- ✓ 清华大学安装源: <https://pypi.tuna.tsinghua.edu.cn/simple>
- ✓ 阿里云安装源: <http://mirrors.aliyun.com/pypi/simple>
- ✓ 豆瓣安装源: <http://pypi.douban.com/simple>

## □ 本教材需要用到的包:

- ✓ NumPy、Matplotlib、Scikit-learn、Pytorch、Tensorflow
- ✓ 使用pip指令——安装演示

# Python开发环境安装

## □ 显示安装包信息

- ✓ pip show 包名称
- ✓ 如: pip show numpy

## □ 卸载包

- ✓ pip uninstall 包名称
- ✓ pip uninstall numpy

# Python开发环境安装

## □ 集成开发环境

- ✓ 简称IDE，一般包括代码编辑器、编译器、调试器和图形用户界面等工具，能够提升开发效率

## □ Python常用IDE

- ✓ IDLE: Python自带，演示print("hello world")
- ✓ JupyterLab: 基于Web的Python集成开发环境
  - ...> 安装指令: pip install jupyterlab
  - ...> 运行指令: jupyter lab
- ✓ PyCharm:
  - ...> 下载: <https://www.jetbrains.com/pycharm-edu/>
- ✓ Vscode+命令行 **(推荐)**
  - ...> 下载: <https://code.visualstudio.com/download>

# Python开发环境安装

- 安装Python人工智能开发环境需要先安装Python，再通过pip指令下载安装相应的库，最好再安装一款好用的集成开发环境
- Anaconda是一个Python的集成库，是一个开源的Python发行版本，包含了conda、Python等180多个科学包及其依赖项 **(推荐)**
- 下载地址： <https://www.anaconda.com/>
- 清华源：<https://mirrors.tuna.tsinghua.edu.cn/>
  - ✓ ==> 选择 anaconda
  - ✓ ==> 选择 archive
  - ✓ ==> 下拉至下方最新版本点击下载

# 总结

- 课程要求
- Python简介
- 人工智能简介
- Python环境安装
  
- 作业: Spoc第一章 (DDL: 9月16日00:00)
- 任务: 每个人在自己的电脑上安装Python环境
- ✓ 推荐anaconda + vscode的组合 