

# 导论

机器学习与人工智能（一）

陈一帅

[yschen@bjtu.edu.cn](mailto:yschen@bjtu.edu.cn)

北京交通大学电子信息工程学院

# 内容

- 定义
- 类型
- 回顾
- 现状

# 人工智能

AI: Artificial Intelligence

使计算机像人那样聪明

# 机器学习

Machine Learning

# 机器学习

- Arthur Samuel, 1959

“Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed”

“让机器像人一样学习，不需要清楚的编程”

# 机器学习

- Tom Mitchell, 1998

“Well posed Learning Problem: A computer program is said to learn from experience  $E$  with respect to some task  $T$  and some performance measure  $P$ , if its performance on  $T$ , as measured by  $P$ , improves with experience  $E$ .”

“一个适定的学习问题是：一个计算机程序，为了任务  $T$ ，从经验  $E$  中学习，改进它的性能  $P$ 。它在任务  $T$  上的表现（由  $P$  测量）随经验  $E$  而改善”

# 实现方法

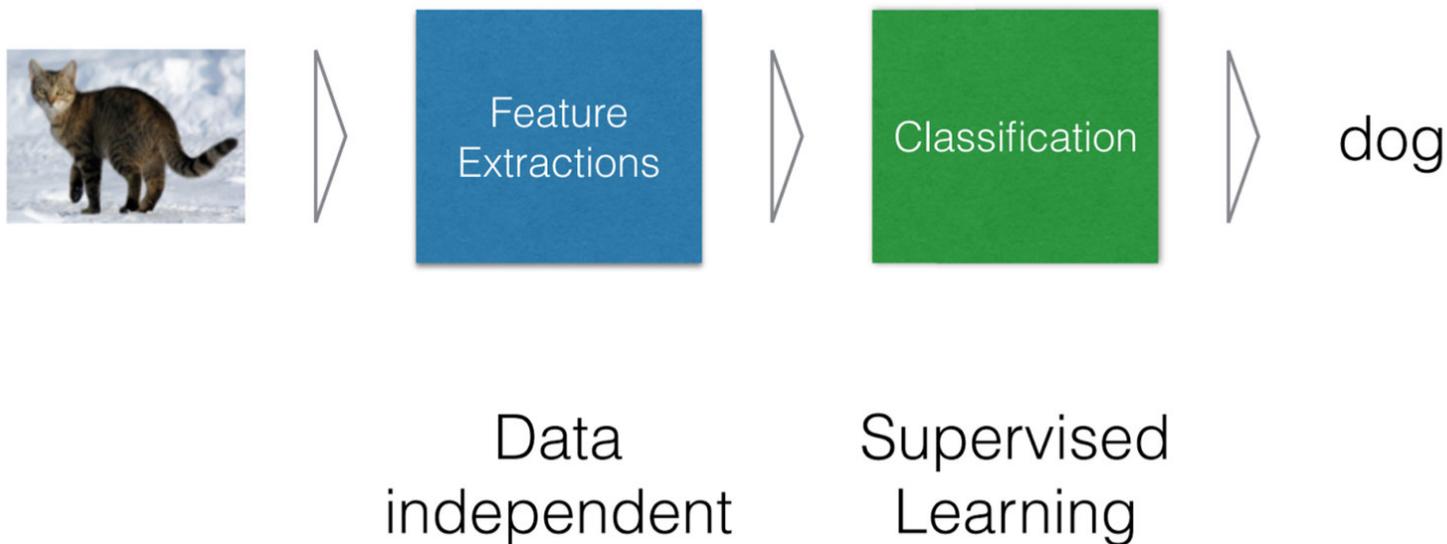
- 定义规则
  - 猫 = “有胡须的四足动物”
- 分解为一系列子问题
  - 找腿，而且要有四只
  - 找胡子
- 分解为更细节的技术问题
  - 找边缘
  - 将猫从背景中提取出来：背景分离

# 机器学习方法

- 不需给出“什么是猫”的清晰的规则
  - 给计算机很多猫的图片，也给很多不是猫的图片（狗、汽车、...）的图片
  - 计算机从这些数据中，自己学出来什么是“猫”
- 效果
  - 比人的识别准确度都要高

# 机器学习方法

- 特征提取
- 机器学习



# 七种机器学习任务

# 1) 分类

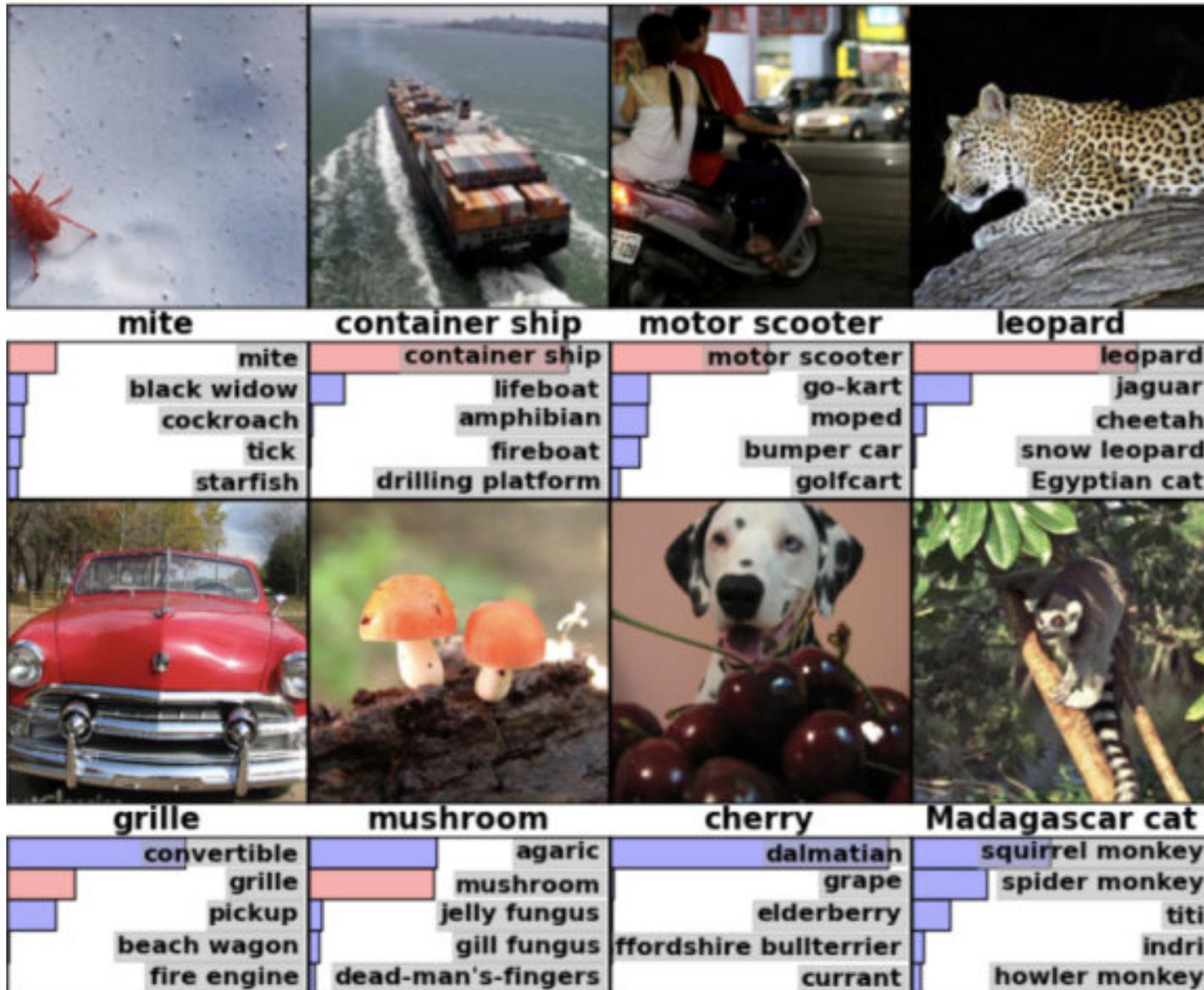
Classification

学习“类别”

# 分类： 图片分类 (ImageNet)

																							
<b>lens cap</b>	<b>abacus</b>	<b>slug</b>	<b>hen</b>																				
<table border="1"> <tr><td>reflex camera</td></tr> <tr><td>Polaroid camera</td></tr> <tr><td>pencil sharpener</td></tr> <tr><td>switch</td></tr> <tr><td>combination lock</td></tr> </table>	reflex camera	Polaroid camera	pencil sharpener	switch	combination lock	<table border="1"> <tr><td>abacus</td></tr> <tr><td>typewriter keyboard</td></tr> <tr><td>space bar</td></tr> <tr><td>computer keyboard</td></tr> <tr><td>accordion</td></tr> </table>	abacus	typewriter keyboard	space bar	computer keyboard	accordion	<table border="1"> <tr><td>slug</td></tr> <tr><td>zucchini</td></tr> <tr><td>ground beetle</td></tr> <tr><td>common newt</td></tr> <tr><td>water snake</td></tr> </table>	slug	zucchini	ground beetle	common newt	water snake	<table border="1"> <tr><td>hen</td></tr> <tr><td>cock</td></tr> <tr><td>cocker spaniel</td></tr> <tr><td>partridge</td></tr> <tr><td>English setter</td></tr> </table>	hen	cock	cocker spaniel	partridge	English setter
reflex camera																							
Polaroid camera																							
pencil sharpener																							
switch																							
combination lock																							
abacus																							
typewriter keyboard																							
space bar																							
computer keyboard																							
accordion																							
slug																							
zucchini																							
ground beetle																							
common newt																							
water snake																							
hen																							
cock																							
cocker spaniel																							
partridge																							
English setter																							
																							
<b>tiger</b>	<b>chambered nautilus</b>	<b>tape player</b>	<b>planetarium</b>																				
<table border="1"> <tr><td>tiger</td></tr> <tr><td>tiger cat</td></tr> <tr><td>tabby</td></tr> <tr><td>boxer</td></tr> <tr><td>Saint Bernard</td></tr> </table>	tiger	tiger cat	tabby	boxer	Saint Bernard	<table border="1"> <tr><td>lampshade</td></tr> <tr><td>throne</td></tr> <tr><td>goblet</td></tr> <tr><td>table lamp</td></tr> <tr><td>hamper</td></tr> </table>	lampshade	throne	goblet	table lamp	hamper	<table border="1"> <tr><td>cellular telephone</td></tr> <tr><td>slot</td></tr> <tr><td>reflex camera</td></tr> <tr><td>dial telephone</td></tr> <tr><td>iPod</td></tr> </table>	cellular telephone	slot	reflex camera	dial telephone	iPod	<table border="1"> <tr><td>planetarium</td></tr> <tr><td>dome</td></tr> <tr><td>mosque</td></tr> <tr><td>radio telescope</td></tr> <tr><td>steel arch bridge</td></tr> </table>	planetarium	dome	mosque	radio telescope	steel arch bridge
tiger																							
tiger cat																							
tabby																							
boxer																							
Saint Bernard																							
lampshade																							
throne																							
goblet																							
table lamp																							
hamper																							
cellular telephone																							
slot																							
reflex camera																							
dial telephone																							
iPod																							
planetarium																							
dome																							
mosque																							
radio telescope																							
steel arch bridge																							

# 分类： 图片分类 (ImageNet)



# 例：垃圾邮件分类

- 邮件特征
  - 发件时间
  - 发件人
  - 敏感词
- 结果
  - 是否垃圾邮件

# 分类性能评估

- True Negative: 真阴
- False Positive: 假阳
- False Negative: 假阴
- True Positive: 真阳

negative class	TN	FP
positive class	FN	TP
	predicted negative	predicted positive

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

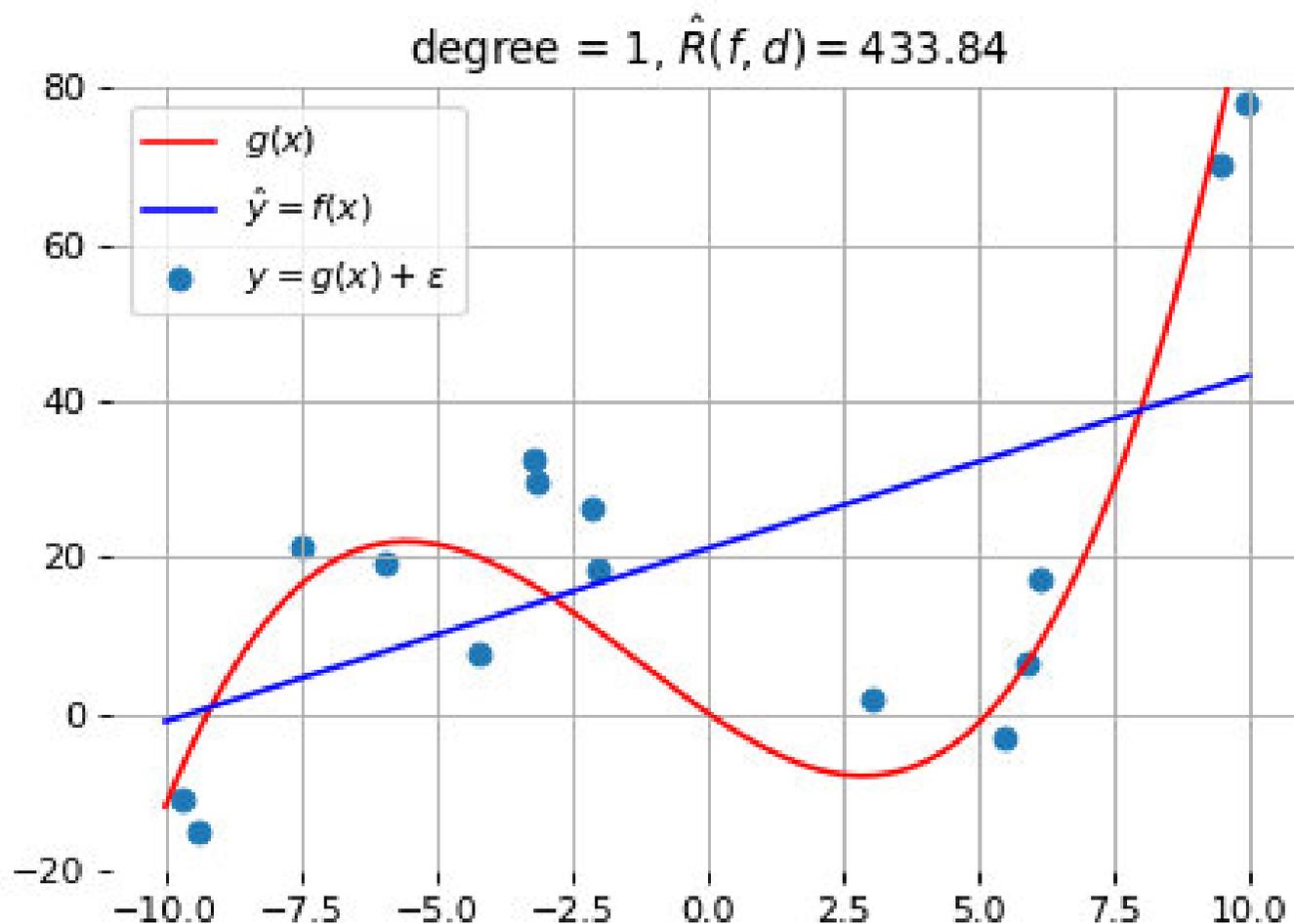
## 2) 回归

Regression

学习“数值”

# 线性回归与非线性回归

直线：线性回归；曲线：非线性回归



# 股市预测

茅台股价和平均目标价

## Drink to That

Moutai may be the first stock in China to reach 1,000 yuan



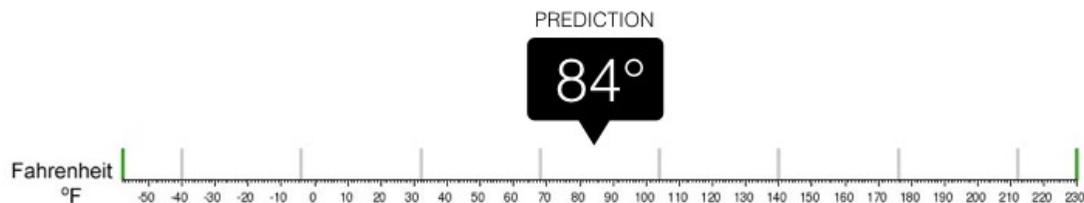
预测涨跌，分类；预测股价，回归

# 练习：预测温度，分类回归？



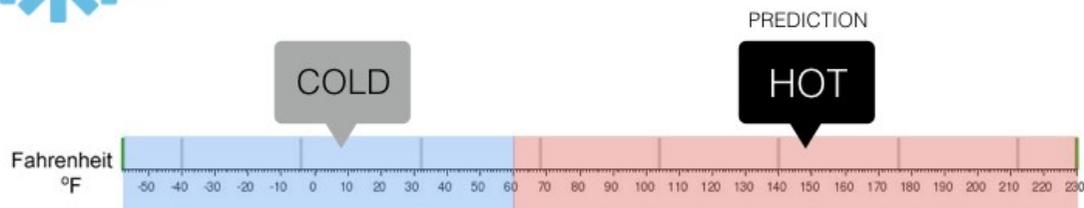
## Regression

What is the temperature going to be tomorrow?



## Classification

Will it be Cold or Hot tomorrow?



# 3) 排序

Ranking

# 搜索排序



百度为您找到相关结果约48,900,000个

搜索工具

投资有风险，选择需谨慎。

## [【人人贷】投资理财\\_人人贷\\_AAA信用\\_行业合规排名前三](#)



人人贷投资,投资理财,8年专业品牌,银行资金存管,AAA信用认证,靠谱省心!投资理财就上人人贷,新手专享1288元红包,行业合规排名前三,限量发布!

www.renrendai.com 2018-08 ▾ V2 - 评价 - 广告

## [【京东金融】安全的理财产品-京东理财\\_理财产品丰富\\_多重保险](#)



安全的理财产品-京东理财,门槛低,收益可观,选取优质产品,更有专业团队为您保驾护航!安全的理财产品-就选京东理财,互联网专业投资平台,优质金融服务.

jr.jd.com 2018-08 ▾ V3 - 评价 - 广告

## [【钱多多】投资理财-交易透明合规\\_靠谱理财平台](#)



投资理财-资金银行存管,上海股交所挂牌平台,8月返现送不停!投资理财参考回报12% 新手另享加息3.6%\_5年运营平台 让您投资无..

新手25天标 资金银行存管 8月返现送不停! 历史年化12%

www.d.com.cn 2018-08 ▾ V3 - 评价 - 广告

## [【京东金融】各银行理财产品收益率对比-京东理财\\_理财产品丰富..](#)

各银行理财产品收益率对比-京东理财,门槛低,收益可观,选取优质产品,更有专业团队为您保驾护航!各银行理财产品收益率对比-就选京东理财,互联网专业投资平台,优质金融服务.

jr.jd.com 2018-08 ▾ V3 - 评价 - 广告

大鱼号/原粹番

# 搜索排序



找到约 157,000,000 条结果 (用时 0.49 秒)

## 普通人如何合理的理财投资，有哪些书可以学习阅读？ - 知乎

<https://www.zhihu.com/question/22818974> ▼

2015年10月12日 - 实际上，**投资理财**并不像专家们鼓吹得那么难，保持合理的心态，寻找合理的逻辑，学习相关的知识，你是有可能会比身边的人赢得更高的收益率的。投资是一种技能， ...

如何进行**投资理财**? 35 个帖子 2018年5月6日

家庭应该如何**投资理财**? 15 个帖子 2016年6月22日

如何系统的学习**投资理财**的知识? 8 个帖子 2016年1月5日

大学生，怎么开始**投资理财**入门? 14 个帖子 2016年1月1日

[www.zhihu.com](#)站内的其它相关信息

## 富翁投资经：十大理财好习惯有钱一辈子负债也是一种资产\_高端休闲\_ ...

[people.pedaily.cn](#) › 悦·生活 › 高端休闲 › 投资 ▼

2016年4月27日 - 富翁**投资**经：十大**理财**好习惯有钱一辈子负债也是一种资产，储蓄和**投资**高效并行观点：不储蓄，绝对成不了富豪；储蓄不是美德，而是手段；努力工作 ...

## 文学城：投资理财 (tzlc) - 文学城论坛

<https://bbs.wenxuecity.com/tzlc/> ▼

准确及时的国内外投资市场动态信息，以及理财讨论，理财规划，家庭理财，生活理财常识，个人理财技巧。

**投资理财**视野涵盖基金、银行、保险、外汇、期货、债券等不同

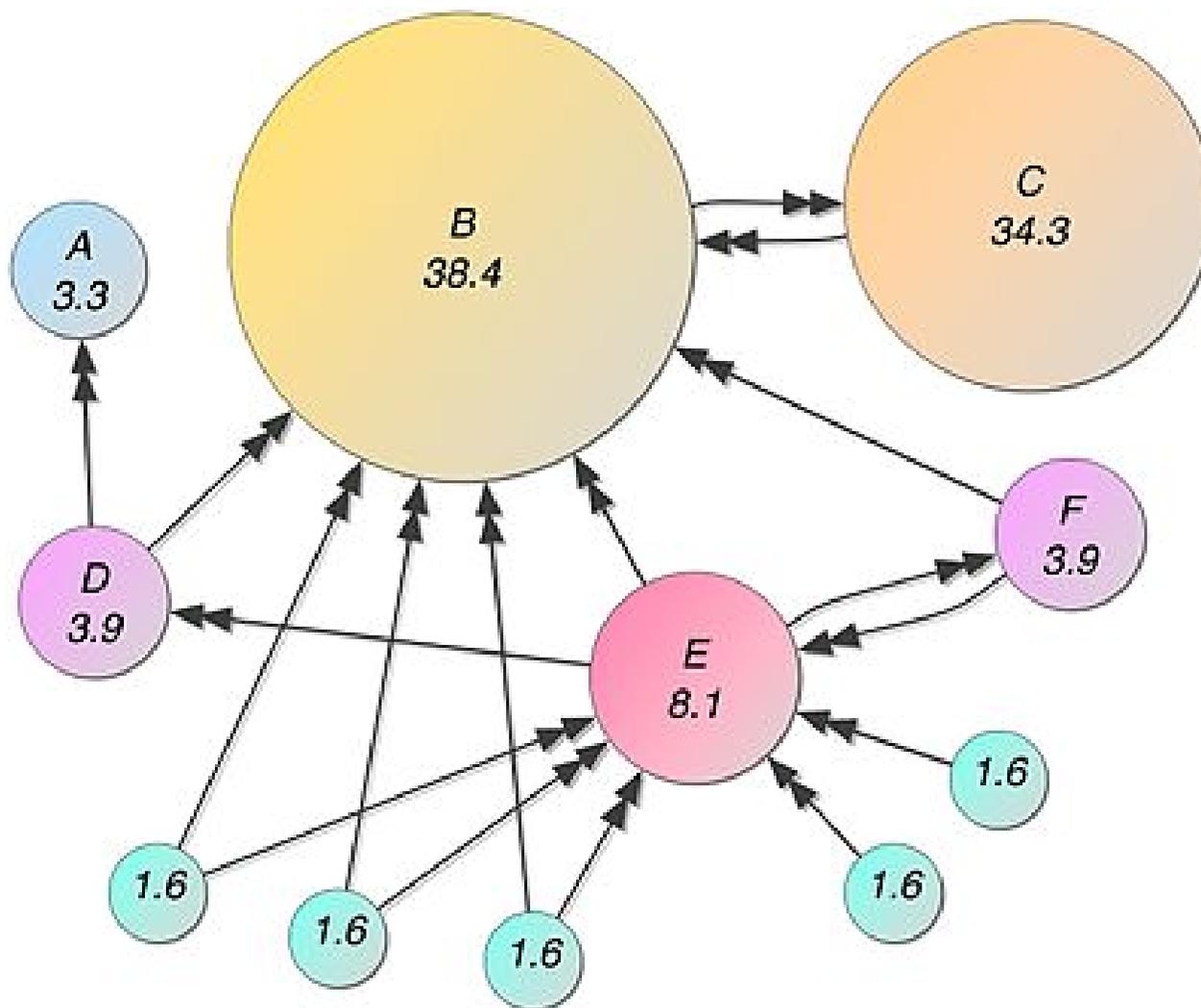
大鱼号/原创

# 排序考虑因素（谷歌2016年）

- 超链接：29%
- 定期制作、原创、“思想领袖”的内容：23%
- 元页面、标题、标签中的关键字：8%
- 手机和平板上的响应速度：8%
- 有针对转换优化的着陆页：8%
- 干净的代码：6%
- 网站访问速度：5%
- 社会网络中的热度：4%
- 网址年龄：4%

# PageRank: 基于超链接排序

随机访问各种超链接，来到一个页面的概率



## 4) 推荐

# 推荐： 淘宝

淘宝网  
Taobao.com

The screenshot displays the Taobao.com homepage with the following elements:

- Navigation Bar:** Includes links for 宝贝 (Items), 天猫 (Tmall), and 店铺 (Stores). A search bar is located on the right with a search button labeled 搜索 (Search).
- Category Bar:** Lists various categories such as 首页 (Home), 女装 (Women's Clothing), 聚划算 (Group Buy), 电器城 (Appliance City), 超市 (Supermarket), 一淘网 (1tao.com), Hitao (Hitao), 旅行 (Travel), 云手机 (Cloud Phone), 特色中国 (Specialty China), and 消费者保障 (Consumer Protection).
- Left Sidebar (淘宝服务 - Taobao Services):**
  - 购物 (Shopping):** 司法拍卖 (Judicial Auction), 淘金币 (Taobao Gold Coins), 天天特价 (Daily Special Price), 跳蚤市场 (Flea Market), 全球购 (Global Buy), 试用 (Trial), 清仓 (Clearance).
  - 生活 (Life):** 彩票 (Lottery), 电子书 (E-books), 水电煤 (Utilities), 保险 (Insurance), 外卖 (Food Delivery), 理财 (Finance), 电影 (Movies), 生活服务 (Life Services).
  - 互动 (Interaction):** 随便逛逛 (Browse), 淘女郎 (Taobao Girl), U站 (U-Station), 网友晒家 (Share Home), 淘宝淘快 (Taobao Fast), 懂懂买 (Buy Smart).
  - 工具 (Tools):** 阿里旺旺 (Alibaba旺旺), 支付宝 (Alipay), 浏览器 (Browser).
  - 其它 (Others):** 品牌特卖 (Brand Special Sale), 品牌馆 (Brand Pavilion), 品牌团购 (Brand Group Buy), 天猫特卖 (Tmall Special Sale), 设计品 (Design Products), 特价商品 (Special Price Goods).
- Main Banner:** A large green banner for "新物种" (New Species) featuring a giraffe illustration. Text includes "史无前例" (Unprecedented), "14.97米" (14.97m), and "9月9日凶猛开卖!" (Fierce sale on September 9th!).
- Right Sidebar (便民服务 - Convenience Services):**
  - 公告 (Announcements), 规则 (Rules), 论坛 (Forum), 安全中心 (Security Center), 公益 (Public Welfare).
  - 快捷新人, 人人有礼 (Fast Newcomer, Everyone has a gift).
- Bottom Section:** Promotional banners for "天猫Tmall.com 精选" (Tmall.com Selection) and "天猫俱乐部" (Tmall Club) with a "秋季网购50元优惠券" (Autumn online shopping 50 yuan coupon). Below these are product categories: 低至7折 (Low to 70% off), 爆款狂欢 (Hot sale), 秋冬新品 (Autumn/Winter New Arrivals), and 5折起售 (50% off). Brands like UNIQLO, 南极人 (NANPOLAR), and Mobil Garden are visible.

# 推荐：头条



# 推荐：合作过滤

- Collaborative Filtering
- 基于用户
  - 找到兴趣相同的人，他喜欢的，推荐给你
- 基于物品
  - 找到用户喜欢分布相近的东西，推荐给你

# 合作过滤：基于用户

					
A					
B					
C					
D					
E					

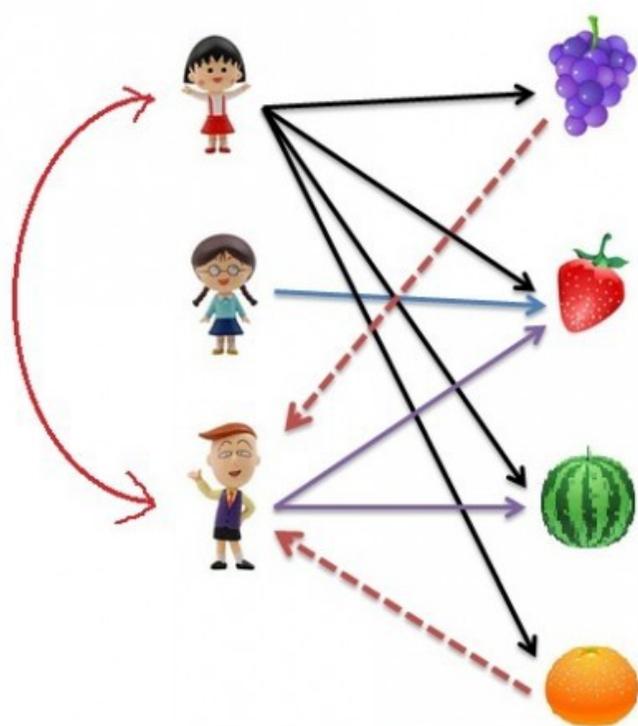
谁和用户E的兴趣相同？谁相反？

# 合作过滤：基于物品

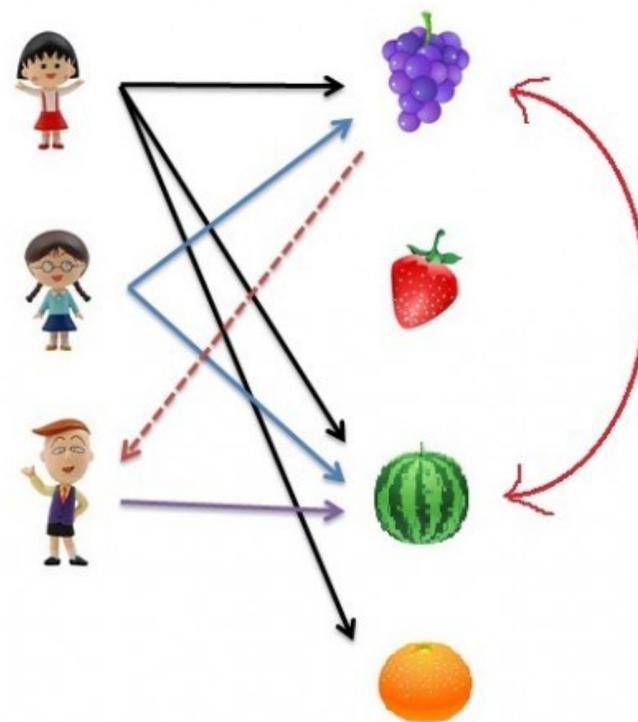
					
A					
B					
C					
D					
E					

谁和“视频”的用户分布相同？谁相反？

# 推荐：合作过滤

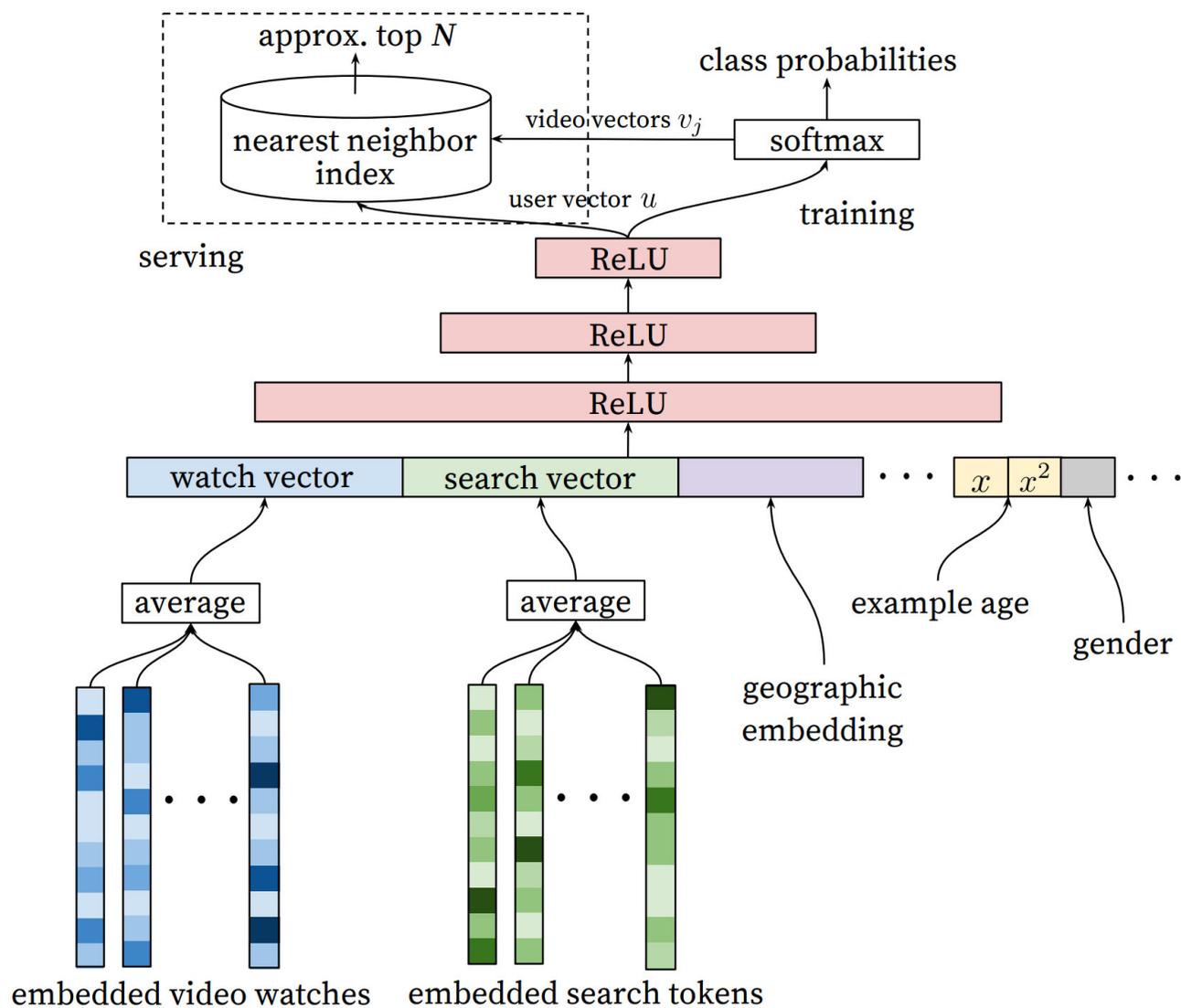


User-based filtering



Item-based filtering

# 推荐 (深度学习)

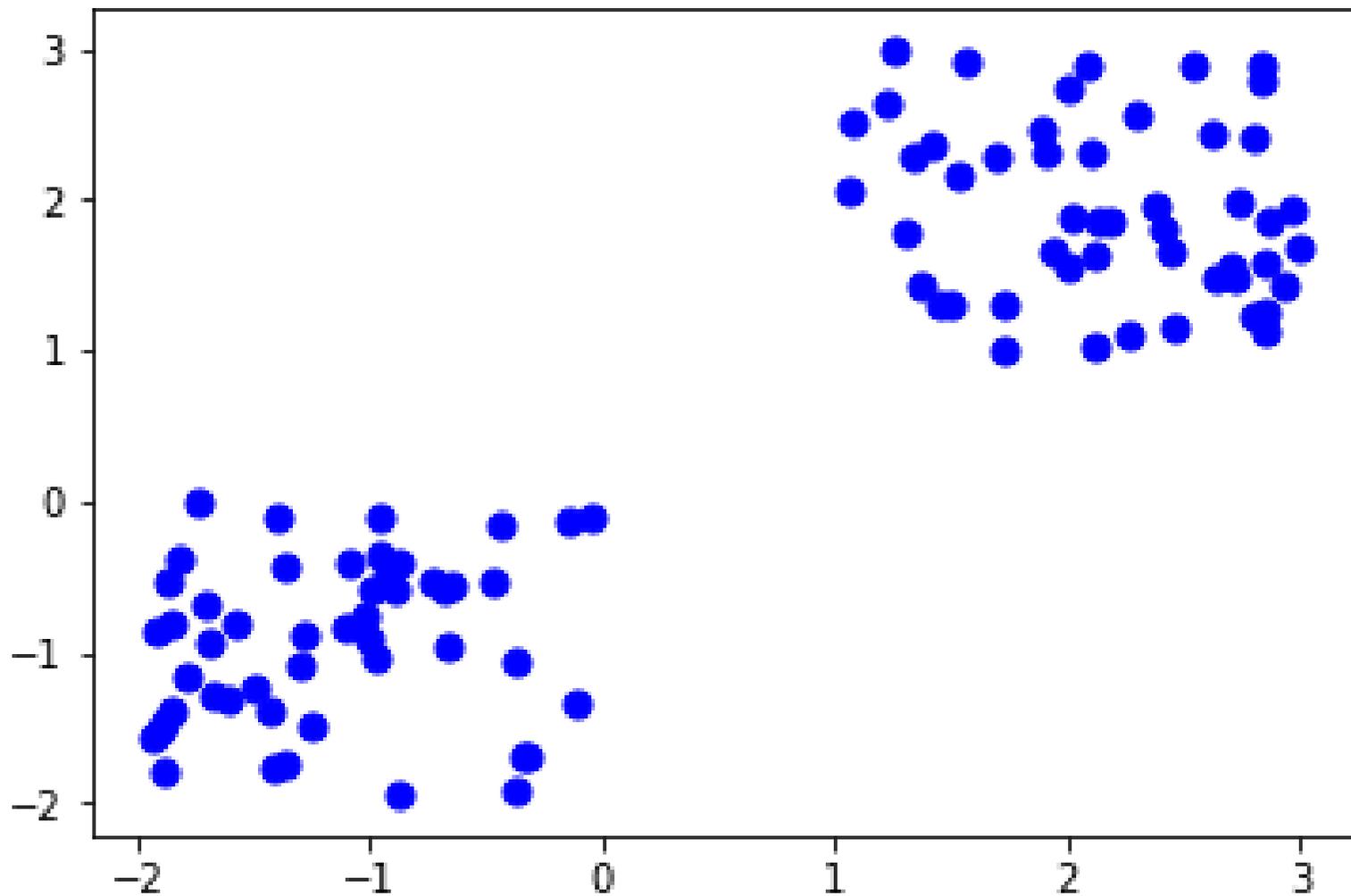


# 5) 聚类

物以类聚

# 聚类

将聚在一起的样本，作为一类

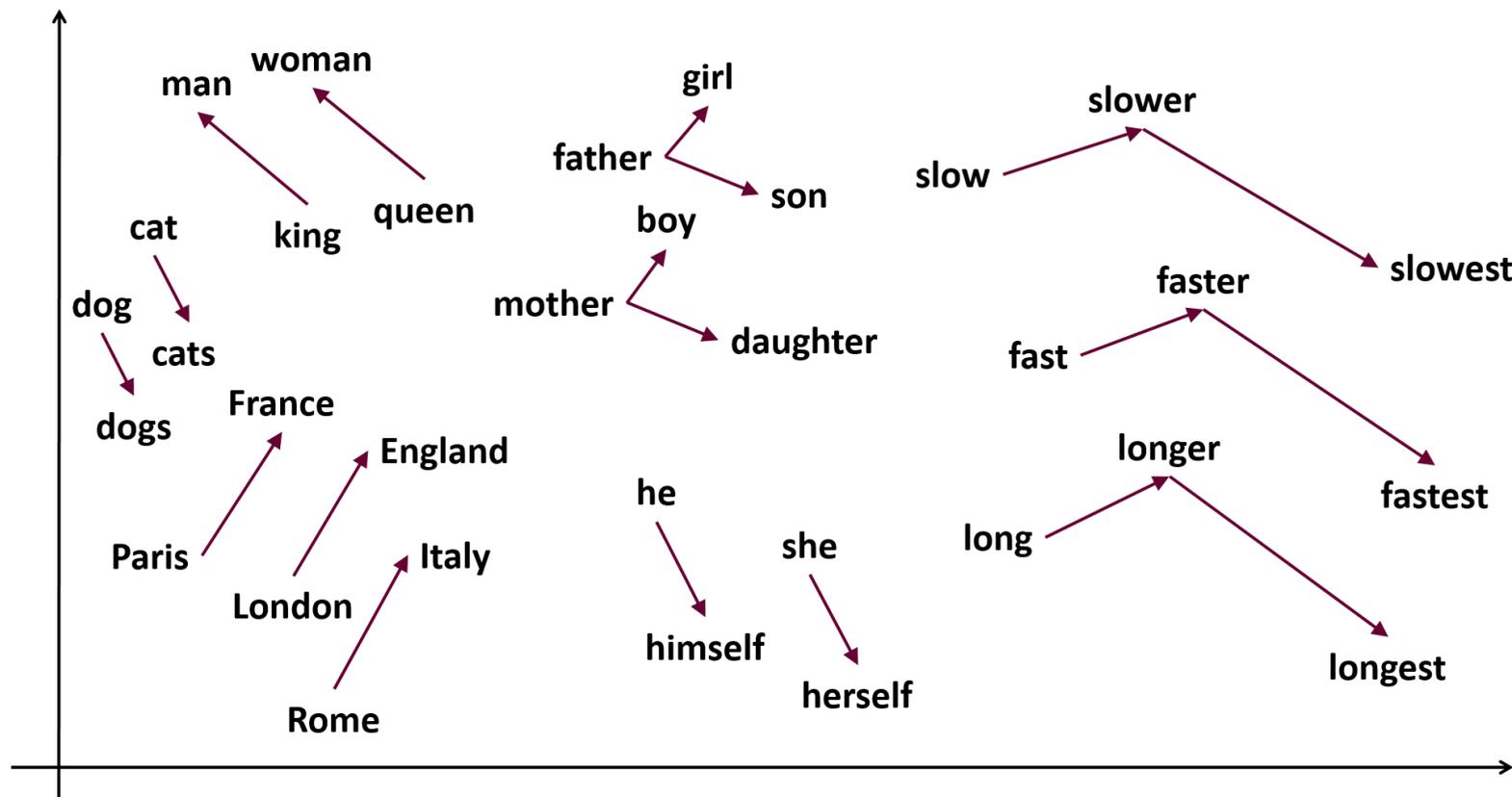


## 6) 表征

将复杂数据，用数学向量表示

# 单词表征

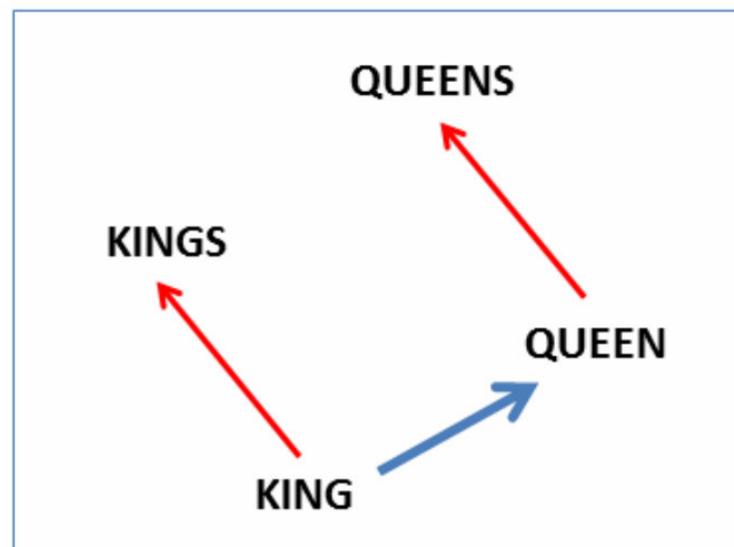
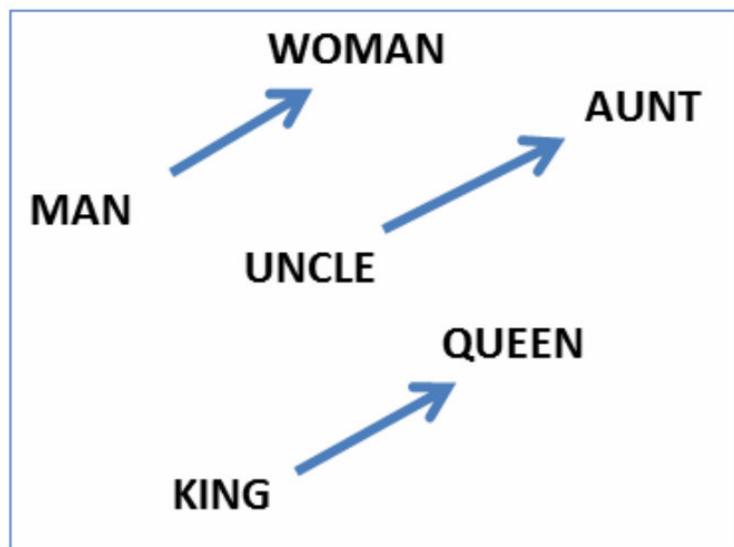
单词位置表示了它们的意义



# 单词表征

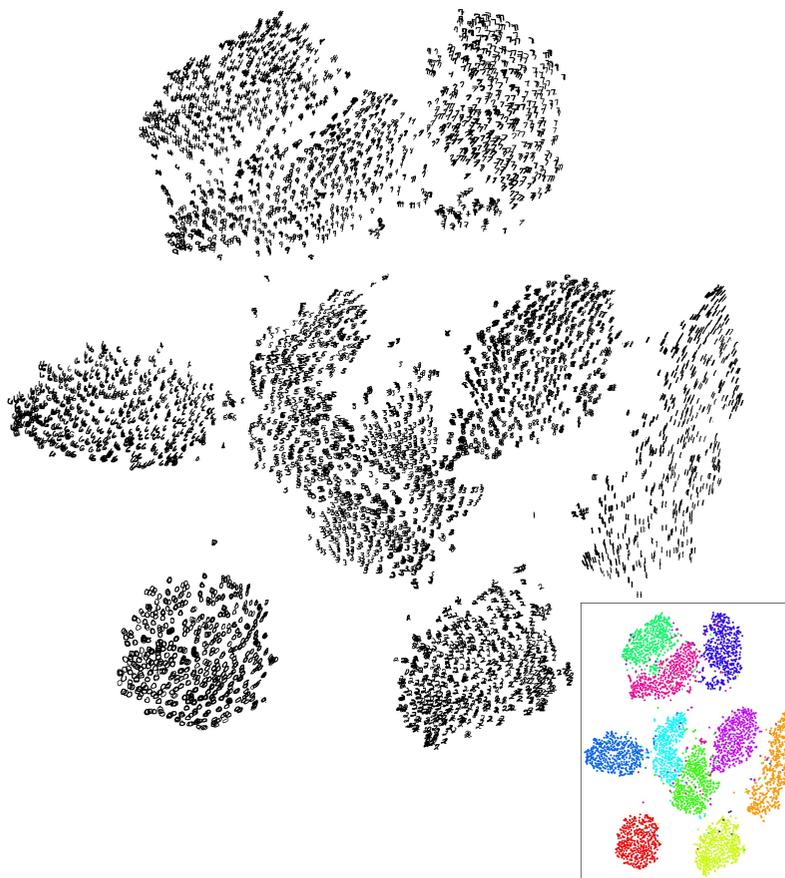
两个单词间的距离向量表示了它们语义上的关系

- 性别关系
- 单复数关系



# 图像表征

图片位置表示了它们的内容



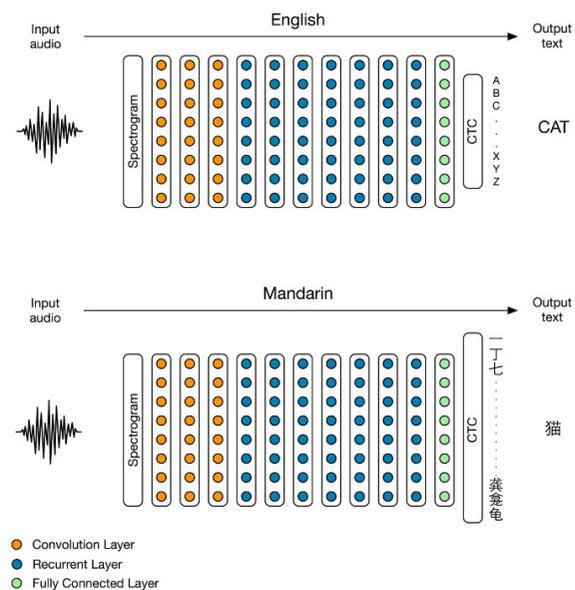
MNIST 0-9数字图片，表征后聚成10个簇

# 7) 结构化预测

利用结构化信息

# 语音识别

- 下一个词和前面的词有很强的关系
- 因此，可以利用前面的单词，预测后面的单词



[Baidu 2014]

# 同声传译

669 人在观看：[【抗韩中年人】157期.1400法强8秒一个大 ...](#) 立即围观 >

人工智能技术：语音翻译和合成

去bilibili观看

分享

扫一扫 手机看



00:00 / 09:04



360P



进入bilibili,一起发弹幕吐槽!

去吐槽

# 小结：七种机器学习任务

1. 分类
2. 回归
3. 排序
4. 推荐
5. 聚类
6. 表征
7. 结构化预测

# 人工智能

Artificial Intelligence

使计算机像人那样聪明

# 智能

- 理解：文字、视觉
- 交流：文字、语音
- 计划：设计、规划
- 学习：知识、模仿

# 图灵测试 (1)

让人认为是在和人对话

## Turing Test - Conversation

:

**Q: Please write me a sonnet on the subject of the Forth Bridge.**

**A : Count me out on this one. I never could write poetry.**

**Q: Add 34957 to 70764.**

**A: (Pause about 30 seconds and then give as answer) 105621.**

**Q: Do you play chess?**

**A: Yes.**

**Q: I have K at my K1, and no other pieces. You have only K at K6 and R at R1. It is your move. What do you play?**

**A: (After a pause of 15 seconds) R-R8 mate.**

# 图灵测试 (1)

- 问：请以“福斯桥”(Forth Bridge)为主题写一首十四行诗。
- 答：我不会回答这道题，我从来没有写过诗。
- 问：34 957 加70 764 等于多少？
- 答：(在经过大概30 秒之后给出答案)105 621。
- 问：你会下象棋吗？
- 答：会的。
- 问：我在K1 处有棋子K，而且没有其他棋子。你在K6 处有棋子K，在R1 处有棋子R。现在轮到你了，你要下哪一步棋？
- 答：(在经过15 秒之后)棋子R 走到R8 处，将军。

# 图灵测试 (2)

## Turing Test - Conversation

:

**Q: In the first line of your sonnet which reads "Shall I compare thee to a summer's day," would not "a spring day" do as well or better?**

**A: It wouldn't scan.**

**Q: How about "a winter's day," That would scan all right.**

**A: Yes, but nobody wants to be compared to a winter's day.**

**Q: Would you say Mr. Pickwick reminded you of Christmas?**

**A: In a way.**

**Q: Yet Christmas is a winter's day, and I do not think Mr. Pickwick would mind the comparison.**

**A: I don't think you're serious. By a winter's day one means a typical winter's day, rather than a special one like Christmas.**

*...and so on, .*

## 图灵测试 (2)

- 问：14行诗的首行是“你如同夏日”，你不觉得“春日”更好吗？
- 答：它不合韵。
- 问：“冬日”如何？它可是完全和韵的。
- 答：它确是合韵的，但没有人愿意被比作“冬日”。
- 问：你不是说过匹克威克先生让你想起圣诞节吗？
- 答：是的。
- 问：圣诞节是冬天的一个日子，我想匹克威克先生对这个比喻不会介意吧。
- 答：我认为您不够严谨，“冬日”指的是一般冬天的日子，而不是某个特别的日子，如圣诞节。

# 技术

- 机器学习
- 搜索
- 推理
  - 符号推理
  - 逻辑推理
- 统计

# 人工智能60年沉浮

1957年至今

# 1956-1974，第一波

- 1956年，达特茅斯会议，提出“人工智能”这一名词
- 1957年，感知机模型
- 进展曾经不错
- Marvin Minsky 1967年预言：AI在一代人内就可以被解决

# ELIZA对话机器人

Welcome to

```
EEEEEE LL      IIII ZZZZZZZ  AAAAA
EE      LL      II     ZZ     AA  AA
EEEEEE LL      II     ZZZ    AAAAAAA
EE      LL      II     ZZ     AA  AA
EEEEEE LLLLLL IIII ZZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.

The original program was described by Joseph Weizenbaum in 1966.

This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?

YOU: Men are all alike.

ELIZA: What is the connection, do you suppose ?

YOU: They're always bugging us about something or other.

ELIZA: Can you think of a specific example ?

YOU: Well, my boyfriend made me come here.

ELIZA: Is it important to you that your boyfriend made you come here ?

YOU: He says I'm depressed much of the time.

ELIZA: I am sorry to hear that you are depressed.

YOU: It's true. I am unhappy.

ELIZA: Can you explain what made you unhappy ?

YOU: █

# ELIZA对话机器人

一位女士对ELIZA进行早期测试时的部分交谈内容：

- 机：有什么麻烦吗？
- 人：天下的男人都很相似。
- 机：怎么相似的？
- 人：他们总是会用这样那样的事情烦扰我们。
- 机：你能举个实例吗？
- 人：嗯，我男朋友让我产生这样的想法。
- 机：你男朋友让你这么想的啊。
- 人：他说我大部分时间都情绪低落。
- 机：听到你说情绪低落，我很难过。
- 人：是这样的，我不快乐。
- 机：你能解释一下什么让你不快乐吗？

# ELIZA对话机器人

## ELIZA- Dialog

:

**Ramu.** I have a terrible headache

**Eliza.** May be you should consult a doctor of medicine.

**Ramu.** Okay I'll do that then

**Eliza.** I am glad you agree

**Ramu.** I want to eat an icecream

**Eliza.** Why do you wish to eat an icecream..

**Ramu.** My friends are intellectuals.

**Eliza.** Intellectuals... You said you have a headache

**Ramu.** Lets discuss religion

**Eliza.** You haven't mentioned your family

**Ramu.** I want to improve myself.

**Eliza.** That is interesting. Perhaps your terrible headache has something to do with your wanting an icecream

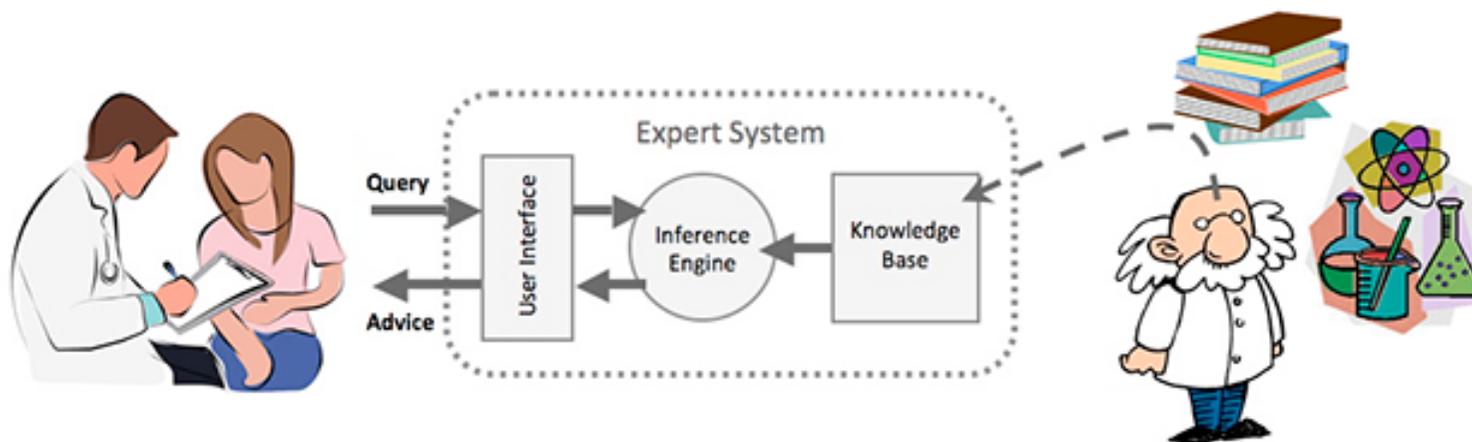
# 第一个冬天：1974-1980

- 感知机模型被发现有缺陷
- 人工智能逐渐被忽略
- 创业、资助消失



# 1980-1987: 专家系统

- 记录专家的行为与决策
- 提取规则，建立专家系统



# 第二个冬天：1987-1993

- 性能未达预期
- 陷入低谷



# 1993-至今

- 机器学习
- 深度学习
- 计算机视觉
- 阿尔法狗
- 最近又有衰退迹象

*Turing Award Won by 3  
Pioneers in Artificial Intelligence*



# 进展

# 人机对话



Hey, Wynton Marsalis is playing this weekend. Do you have a preference between Saturday and Sunday?

-S



[Google Inbox Smart Reply]



[Amazon Echo / Alexa]

利用文字的相关性

# 折衣服 (2010)

2707 人在观看：[起小点TOP10 VOL395 上忍劫鬼斩破疾风](#) 立即围观 >

人工智能：叠毛巾机器人

去bilibili观看

分享

播放器初始化...[完成]

加载用户配置...[完成]

加载视频地址...[完成]

加载视频内容...



00:00 / 00:00



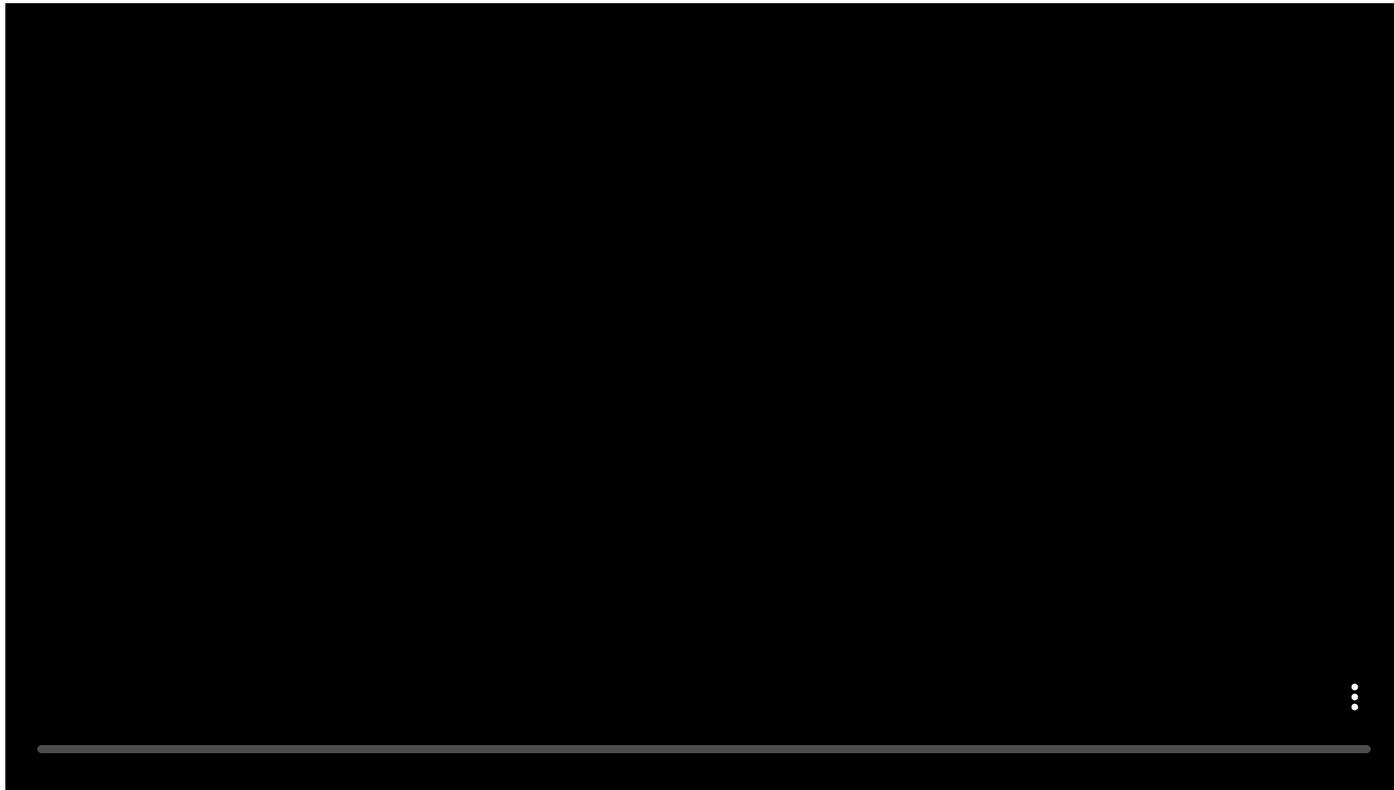
360P



进入bilibili,一起发弹幕吐槽!

去吐槽

# 开车 (2016)



# 分捡物体 (2018)

1131 人在观看： [【WNS中字】180424 惊爆眼球！韩男团公...](#) 立即围观 >

人工智能：机器人自动流水线物品分类

去bilibili观看

分享

扫一扫 手机看



00:00 / 00:33



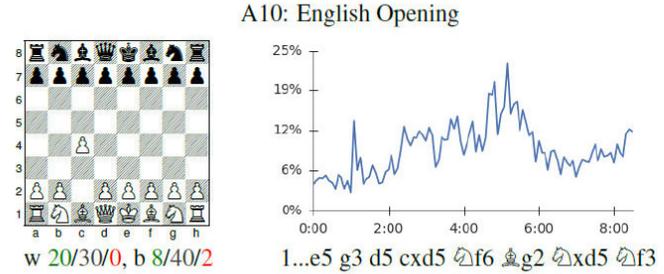
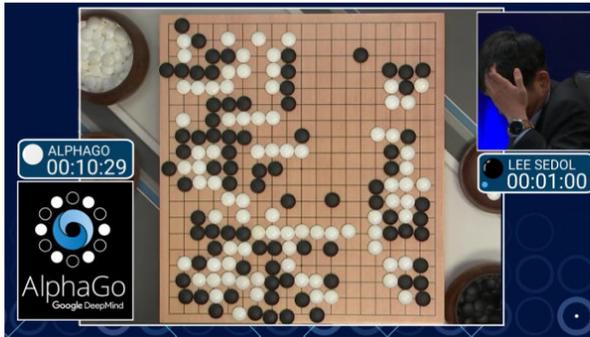
360P



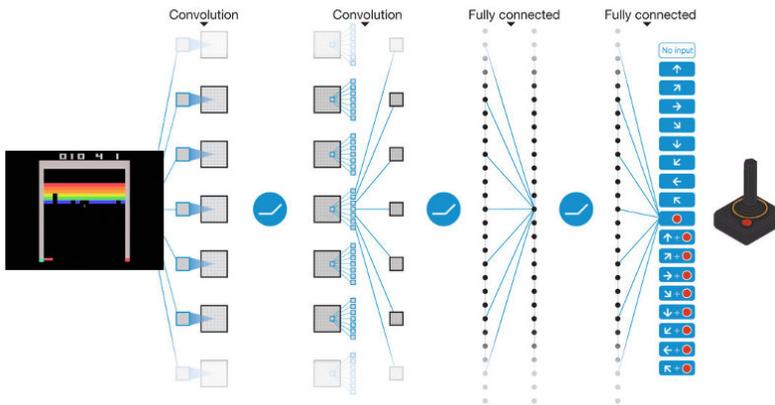
进入bilibili,一起发弹幕吐槽!

去吐槽

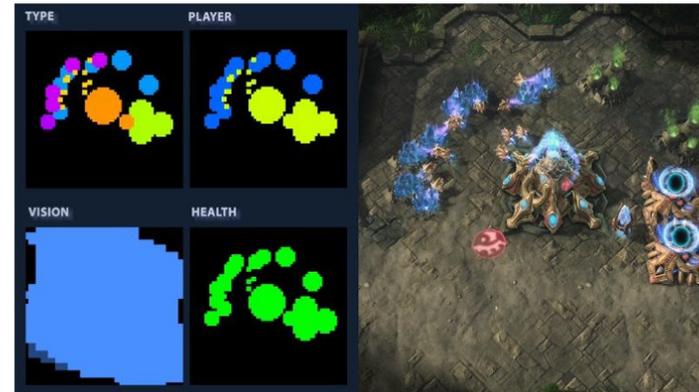
# 游戏



[Deepmind AlphaGo / Zero 2017]



[Atari Games - DeepMind 2016]



[Starcraft 2 for AI research]

# Atari

1631 人在观看：[【星际老男孩】4月24号olimo周赛+赌王大赛](#) [立即围观 >](#)

人工智能：Atari游戏机器人玩家

去bilibili观看

分享

扫一扫 手机看



00:00 / 01:42



360P

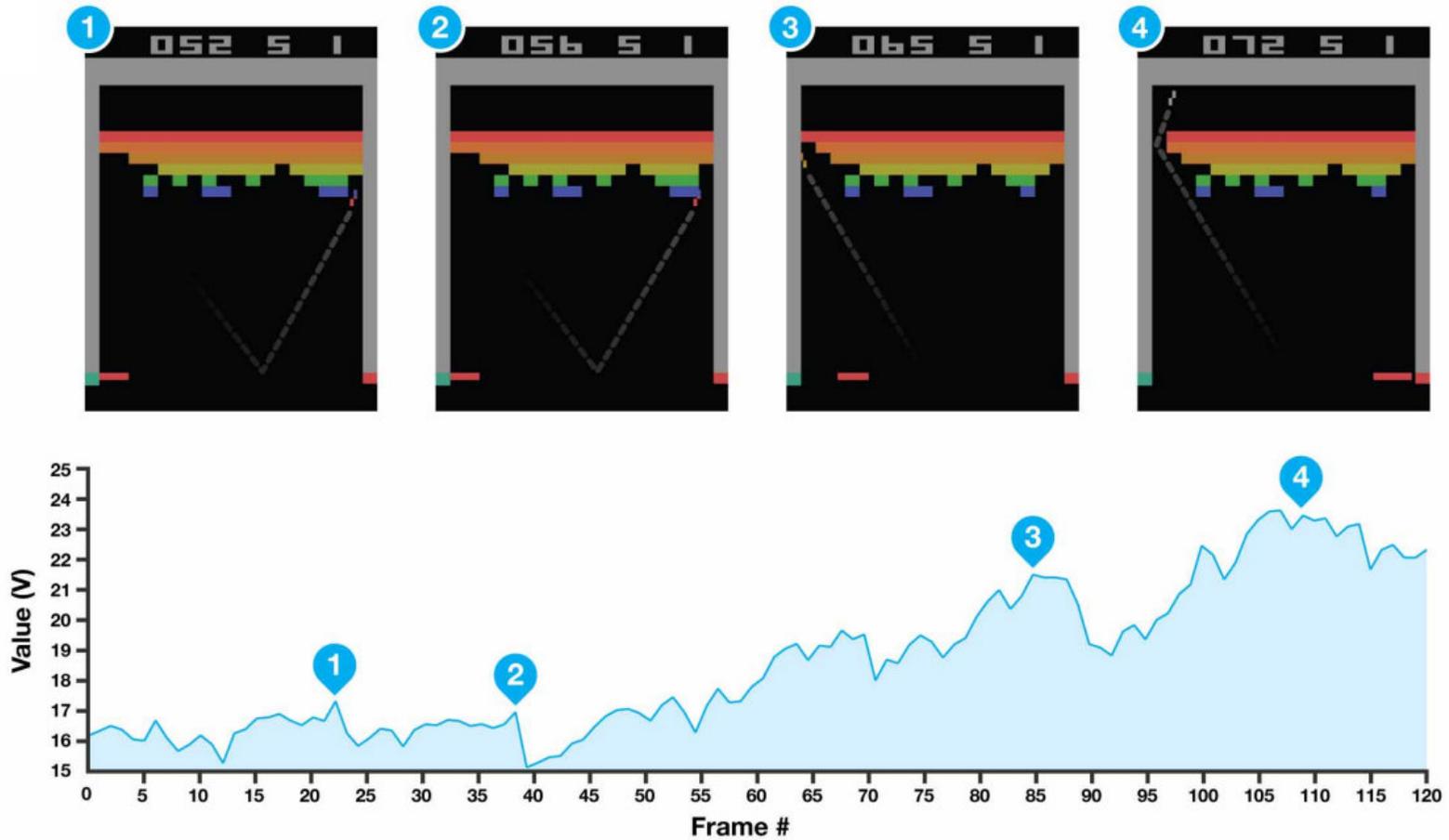


进入bilibili,一起发弹幕吐槽!

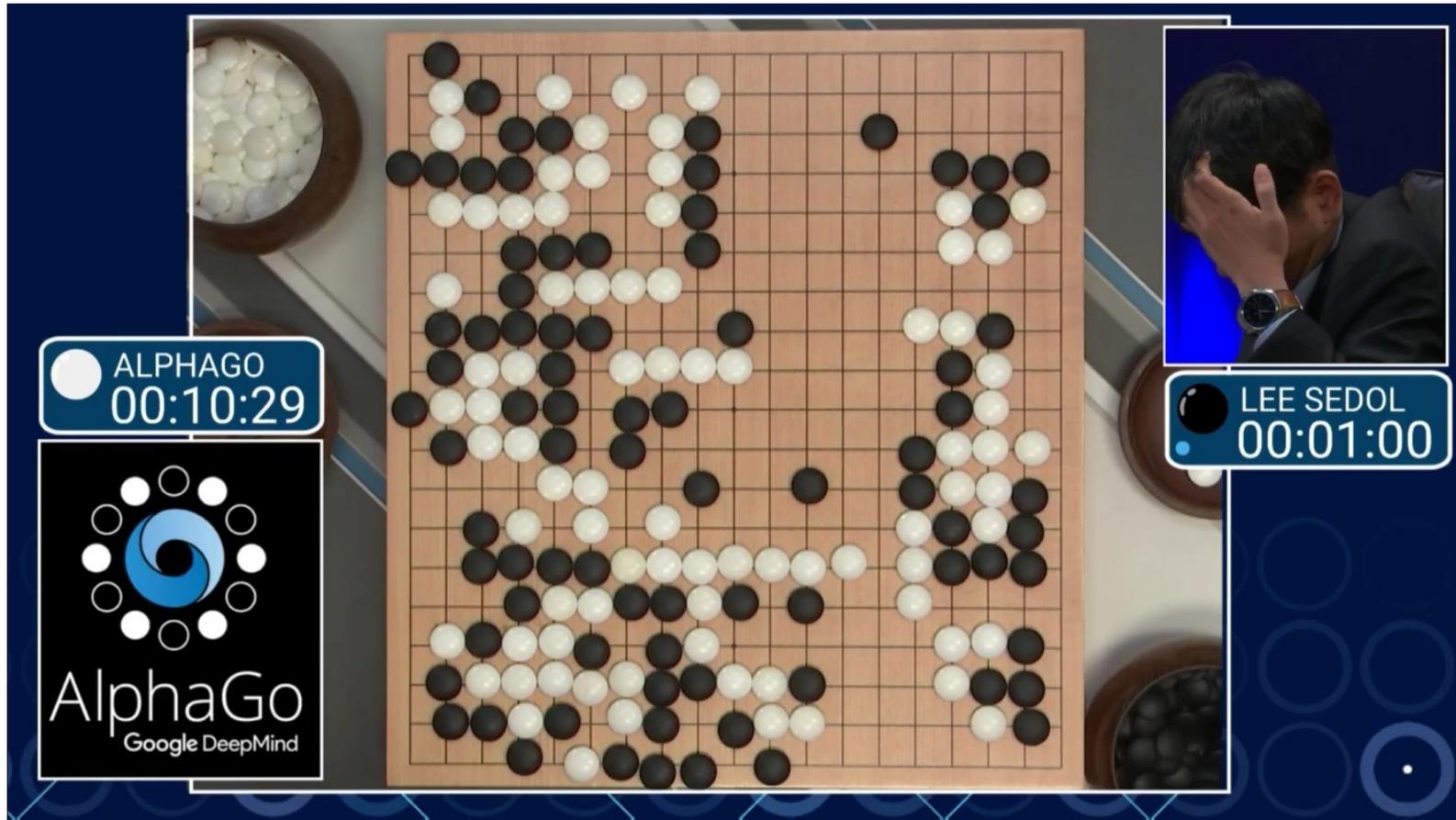
去吐槽

# Atari

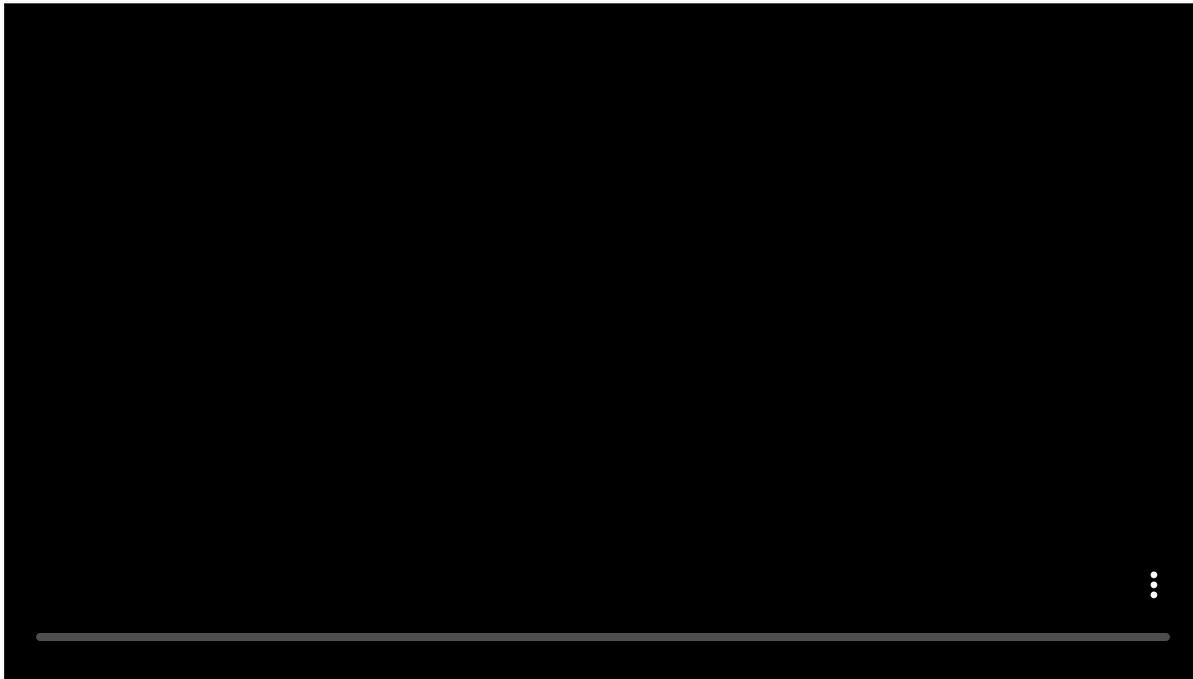
## 增强学习值函数



# AlphaGo



# AlphaGo



# 小测验

- 什么是人工智能
- 预测明天的股票是涨还是跌，是分类还是回归问题？
- 预测明天的气温多少度，是分类还是回归问题？
- 人工智能从诞生起就一帆风顺，这一说法对吗？