



数据挖掘与信息检索实验室

Data Mining & Information Retrieval Laboratory

VALSE Tutorial



## 因果发现与因果性学习

蔡瑞初

数据挖掘与信息检索实验室

广东工业大学

[cairuichu@gmail.com](mailto:cairuichu@gmail.com)

# Outline

---



- ☑ **背景知识**

- ☑ **因果关系发现**

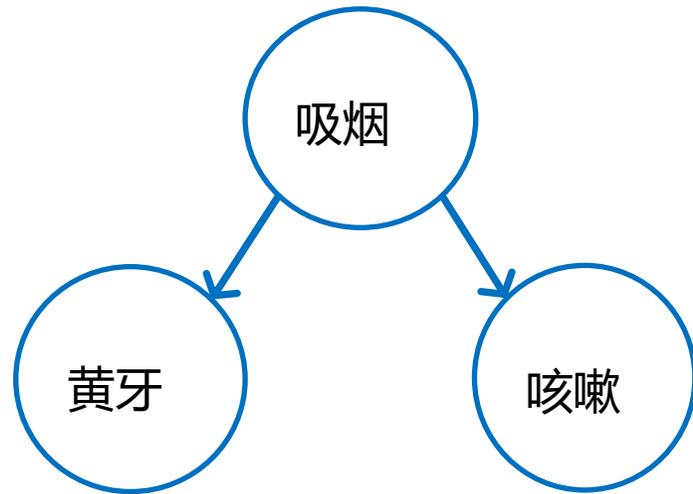
- ☑ **因果性学习**

# 背景知识

---

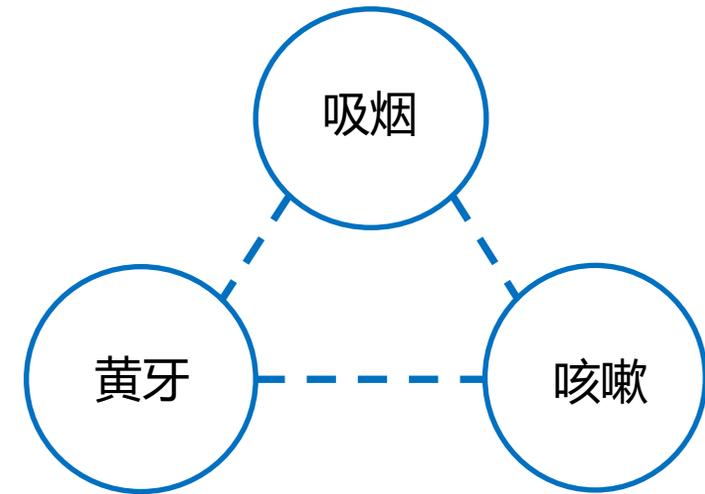
# 什么是因果关系?

## ☑ 关联 vs. 因果



因果

不对称, 简洁  
能够指导干预



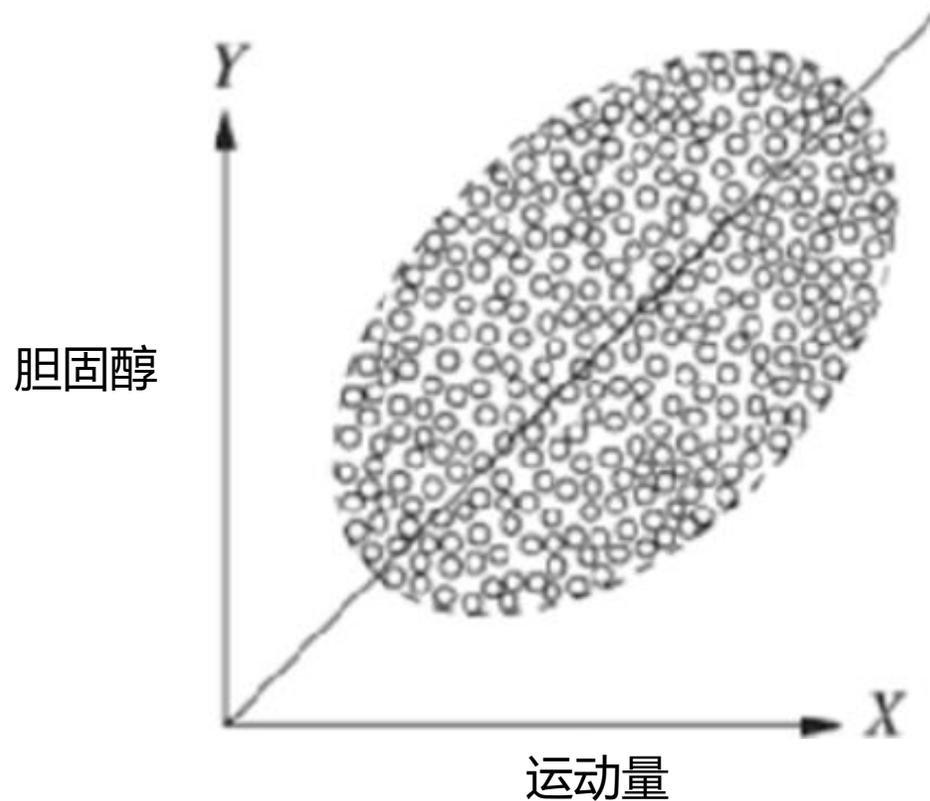
关联

对称, 冗余  
可能给出错误的干预建议

# 为什么需要因果关系？从统计的角度

## ☑ 隐混淆(辛普森悖论)

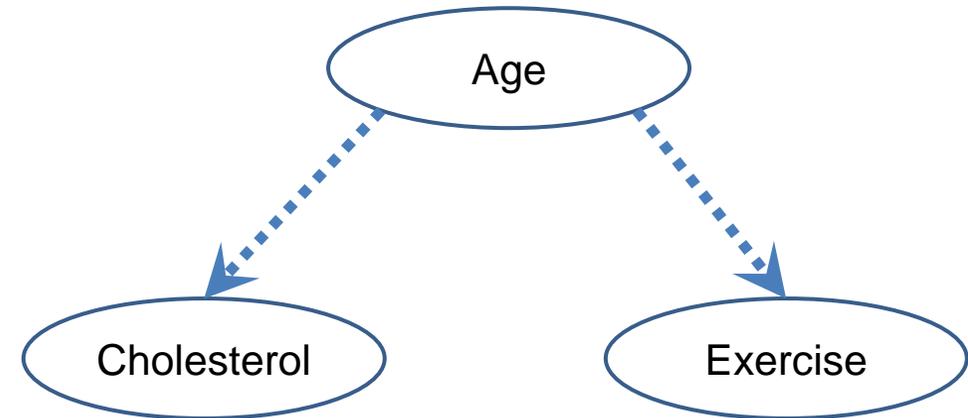
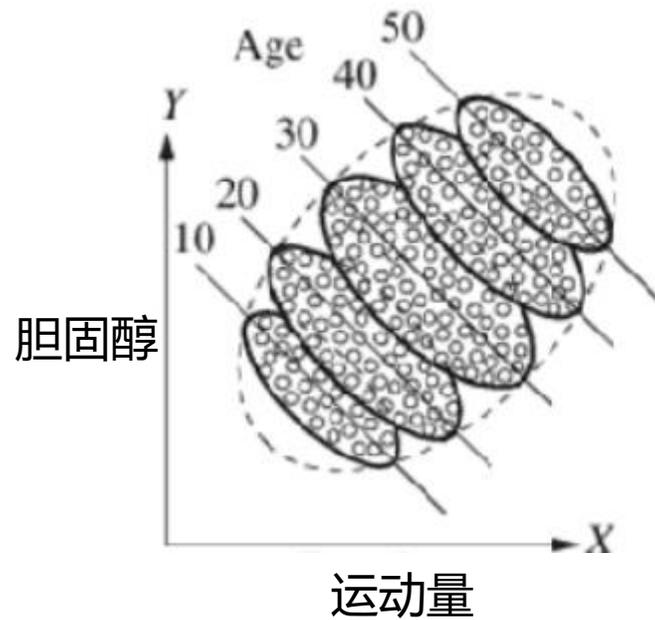
- 胆固醇研究：“胆固醇-运动量”正相关？



# 为什么需要因果关系？从统计的角度

## ☑ 隐混淆（辛普森悖论）

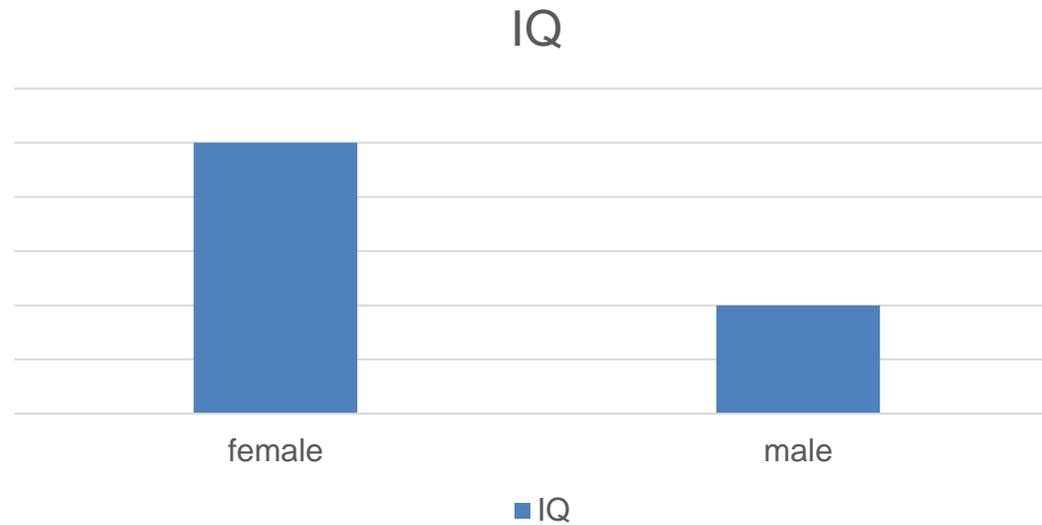
- 胆固醇研究：年龄是混淆因子（共因）！



# 为什么需要因果关系？从统计的角度

## ☑ 选择偏差

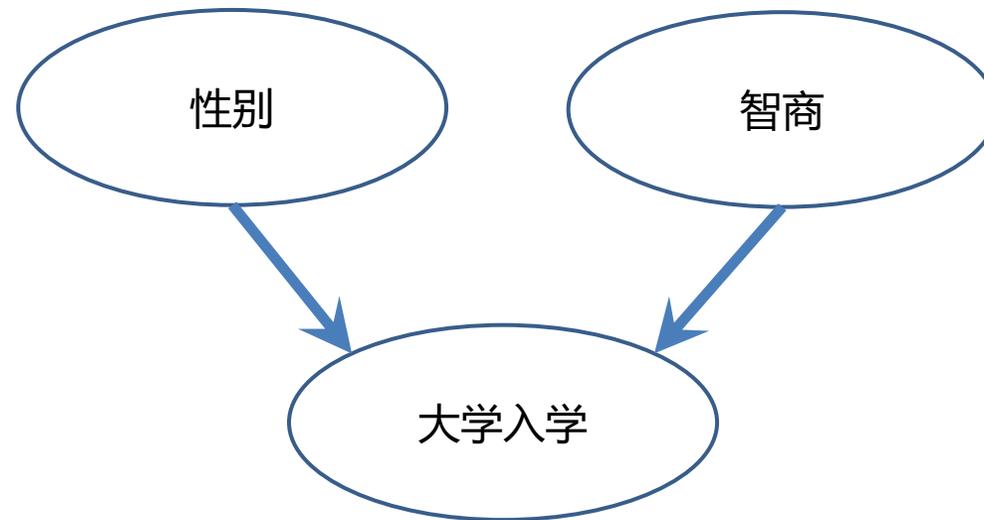
- 1950年美国大学生IQ测试：女性比男性智商高很多？



# 为什么需要因果关系？从统计的角度

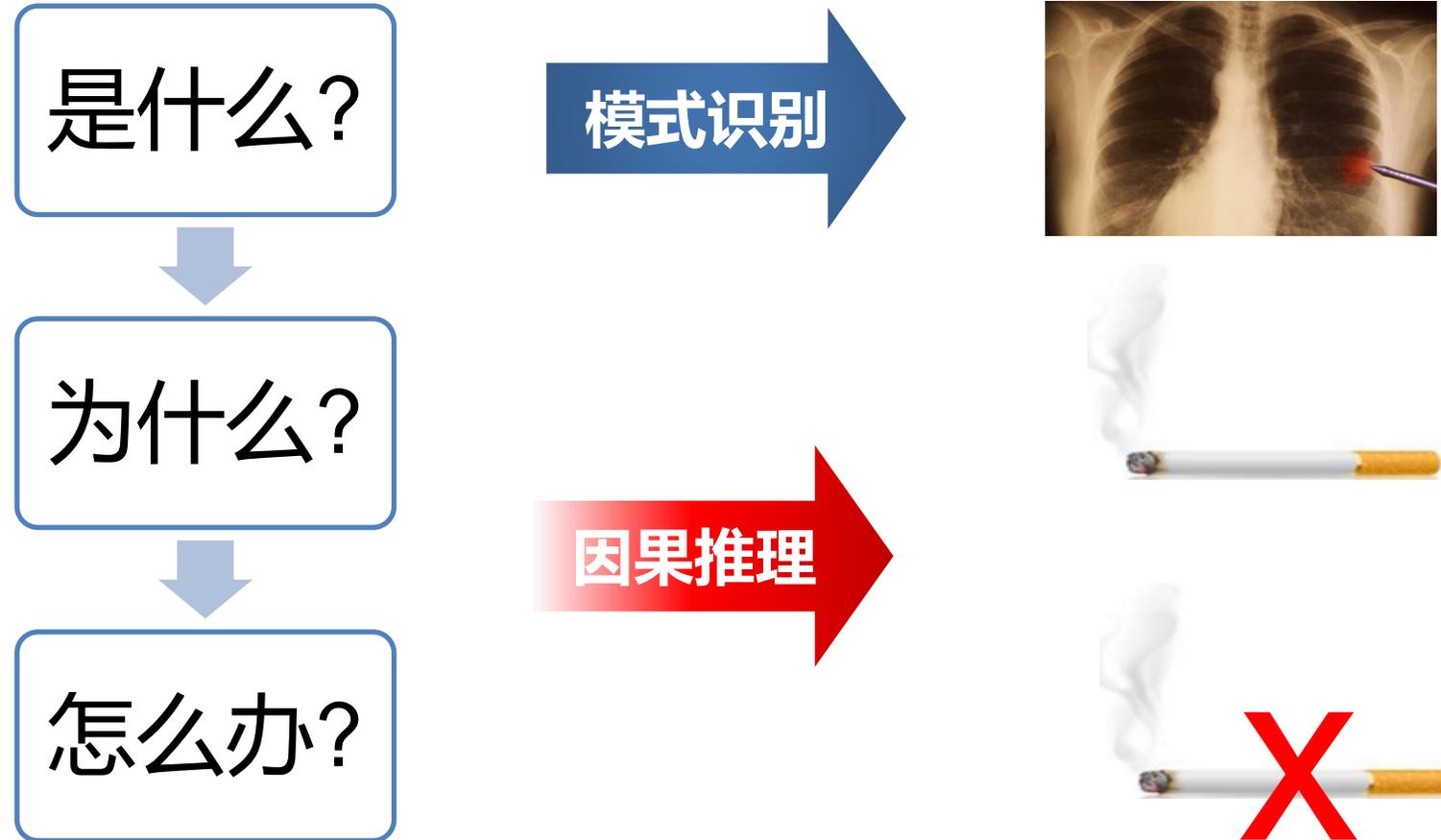
## ☑ 选择偏差

- 1950年美国大学生IQ测试：在校大学生这一条件，导致性别和智商发生伪关联！



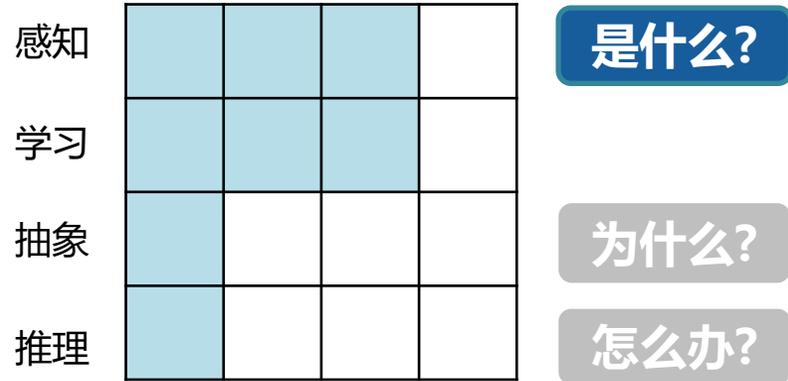
# 为什么需要因果关系？从人工智能的角度

☑ 人工智能三大任务



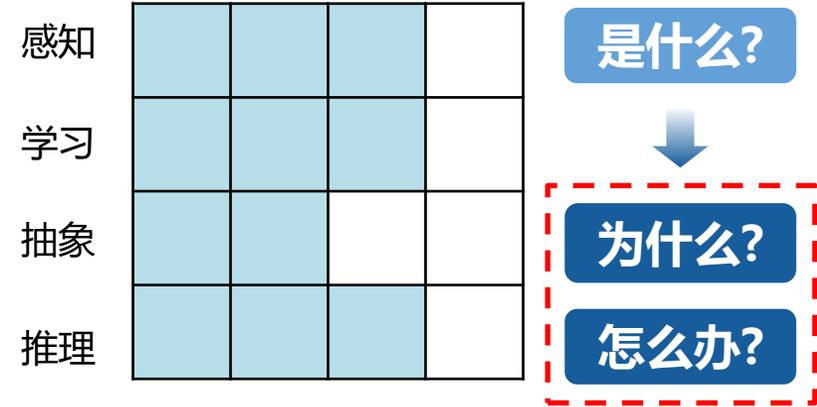
# 为什么需要因果关系？从人工智能的角度

## 深度学习为代表的人工智能



因果推理

## 新一代人工智能



因果关系是实现真正意义人工智能的必经之路。

—— Judea Pearl  
图灵奖得主、因果奠基人



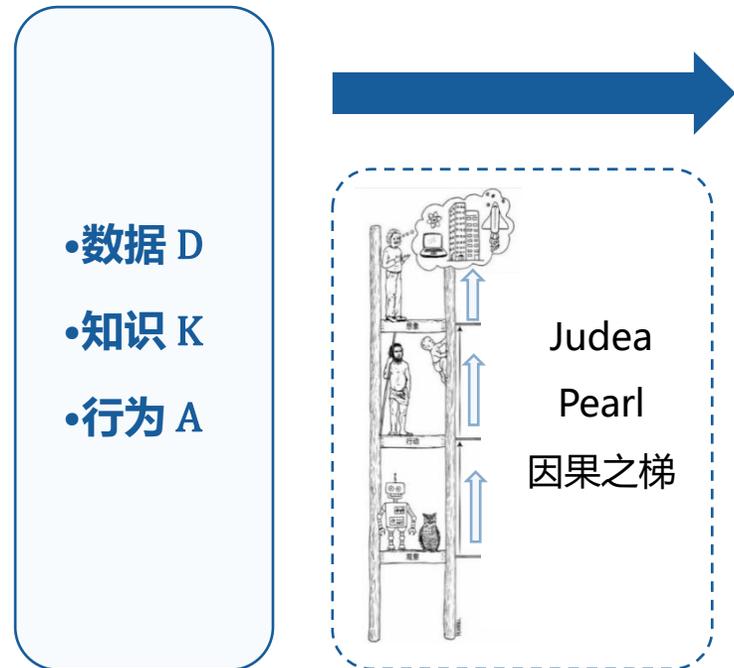
新一代人工智能发展的关键要素是对因果关系的理解。

—— Yoshua Bengio  
图灵奖得主、深度学习奠基人

# 为什么需要因果关系？从人工智能的角度

## 1. 因果发现

为什么：因果知识从哪里来？

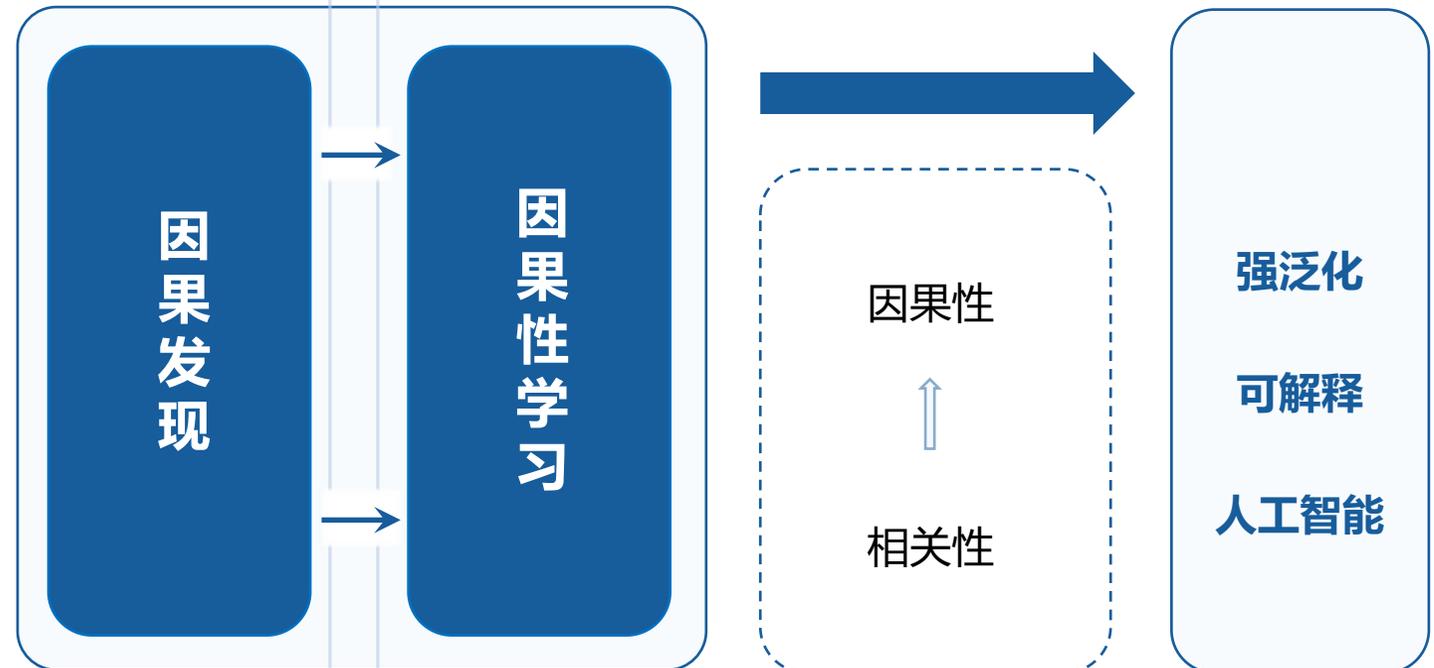


Pearl J, Mackenzie D. The book of why: the new science of **cause and effect**[M]. Basic books, 2018.



## 2. 因果性学习

怎么办：因果知识怎么利用？



Zhang K, Schölkopf B, Spirtes P, Glymour C. **Learning causality and causality-related learning**: some recent progress[J]. National science review, 2018, 5(1): 26-29.