第7章 数据产品开发及数据科学中的人文与管理

编写思路

学习方法、要求及建议

疑难知识点的解读

1.本章定位与内容简介
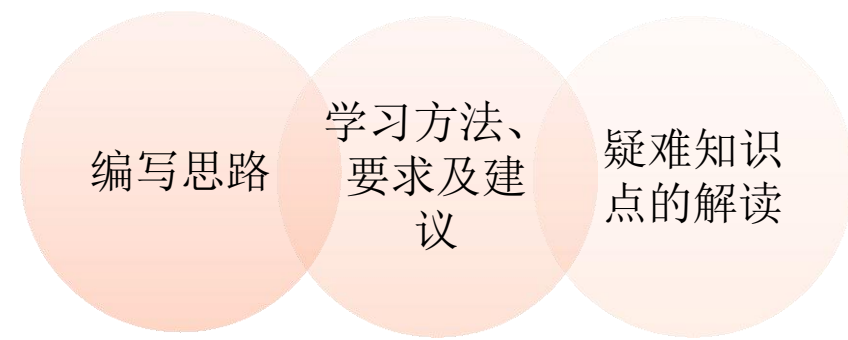

统计学

机器学习

领域知识

基础理论

数据加工

数据计算

数据管理

数据分析

数据产品
开发

数据可视化
与故事化

⋯

人文与管理

7.1数据产品开发及数据科学的人文与管理属性

7.2 数据产品及开发

7.3 数据科学的项目管理

7.4 数据能力

7.5 数据治理

7.6 数据安全

7.7 数据偏见

7.8 数据伦理与道德

7.9 继续学习本章知识

习题

2.本章学习提示及要求

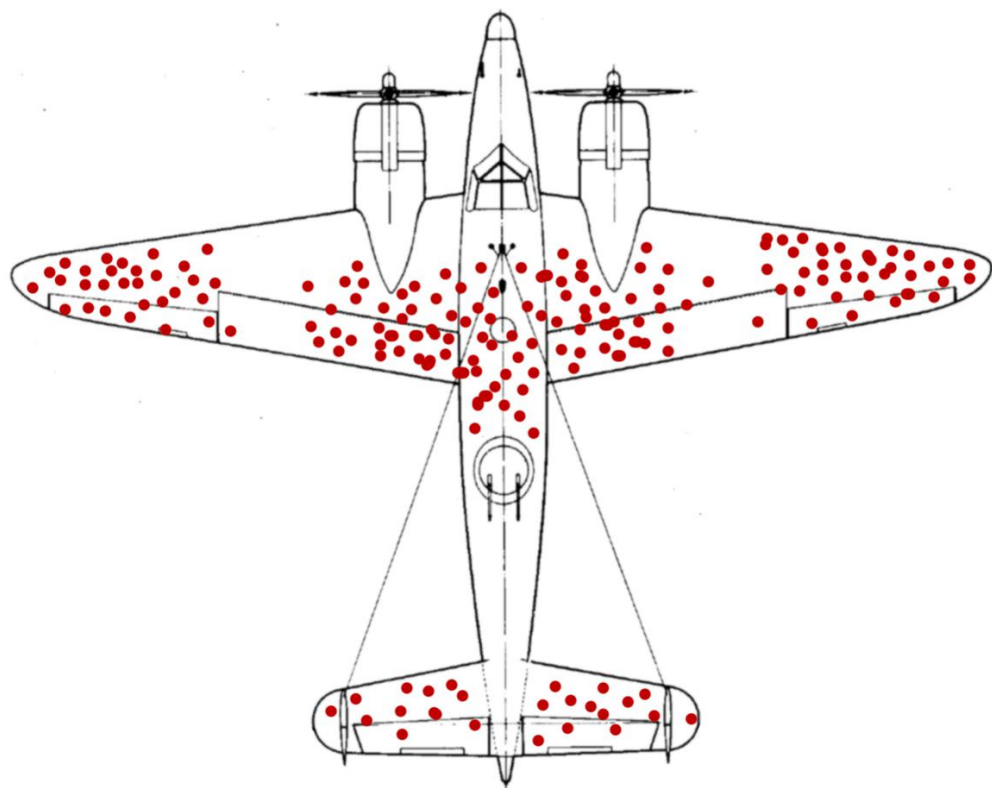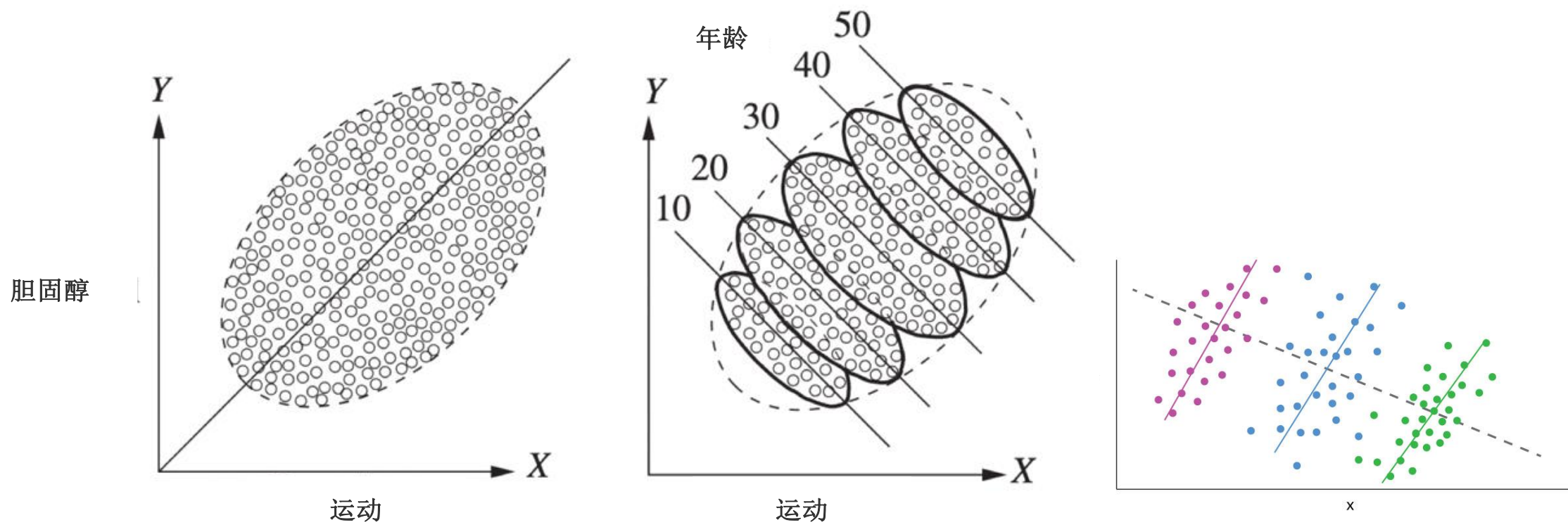| 了解	理解	掌握	熟练掌握
• 数据科学中的人文与管理问题的重要地位 • 数据伦理、数据道德的重要意义	• 数据科学的项目管理方法 • 数据安全及数据保障方法 • 数据产品的概念	• 数据治理的含义及治理方法 • 数据偏见及其避免方法	• 数据能力的评估与改进方法

3. 幸存者偏差

幸存者偏差（**Survivorship Bias**）
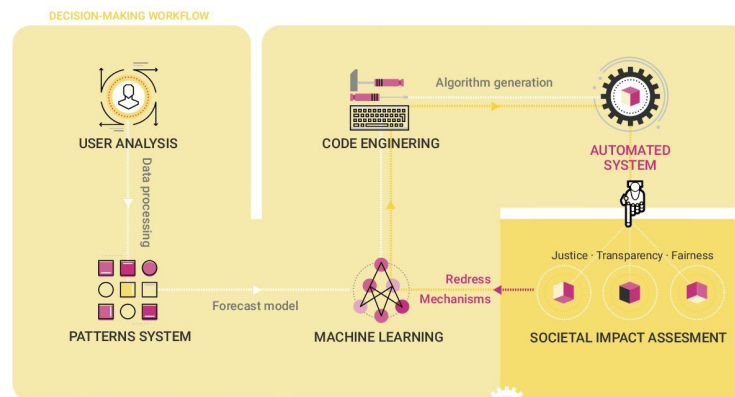

Abraham Wald

4. 辛普森悖论

(来源：PEARL ET AL., 2016)

5.算法歧视

ALGORITHMIC DISCRIMINATION

With the rise of big data, complex algorithms have been created for decision-making purposes and with the idea that they will take efficient, rapid and better decisions. As these algorithms are adaptive and created to learn from humans, they also learn human bias commonly related to race, gender, age or nationality, for instance. The problem of algorithmic discrimination is very complex. It is a technological problem reflecting a societal problem, and there is no easy fix. It affects many aspects of our daily lives, whether we notice it or not. And it definitely shows how technological development has to go hand-to-hand with a societal impact assessment.

DECISION-MAKING WORKFLOW

USER ANALYSIS   CODE ENGINERING   AUTOMATED SYSTEM
Data processing   Algorithm generation
                                            Justice · Transparency · Fairness
PATTERNS SYSTEM   Forecast model   MACHINE LEARNING   Redress Mechanisms   SOCIETAL IMPACT ASSESMENT

CRITICAL TOPICS: OPACITY, UNFAIRNESS, ABUSE OF POWER, DISCRIMINATION...
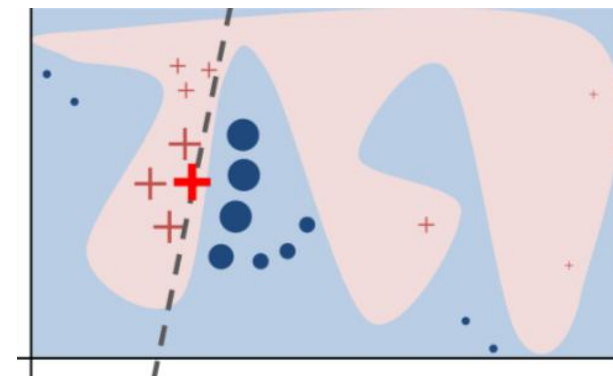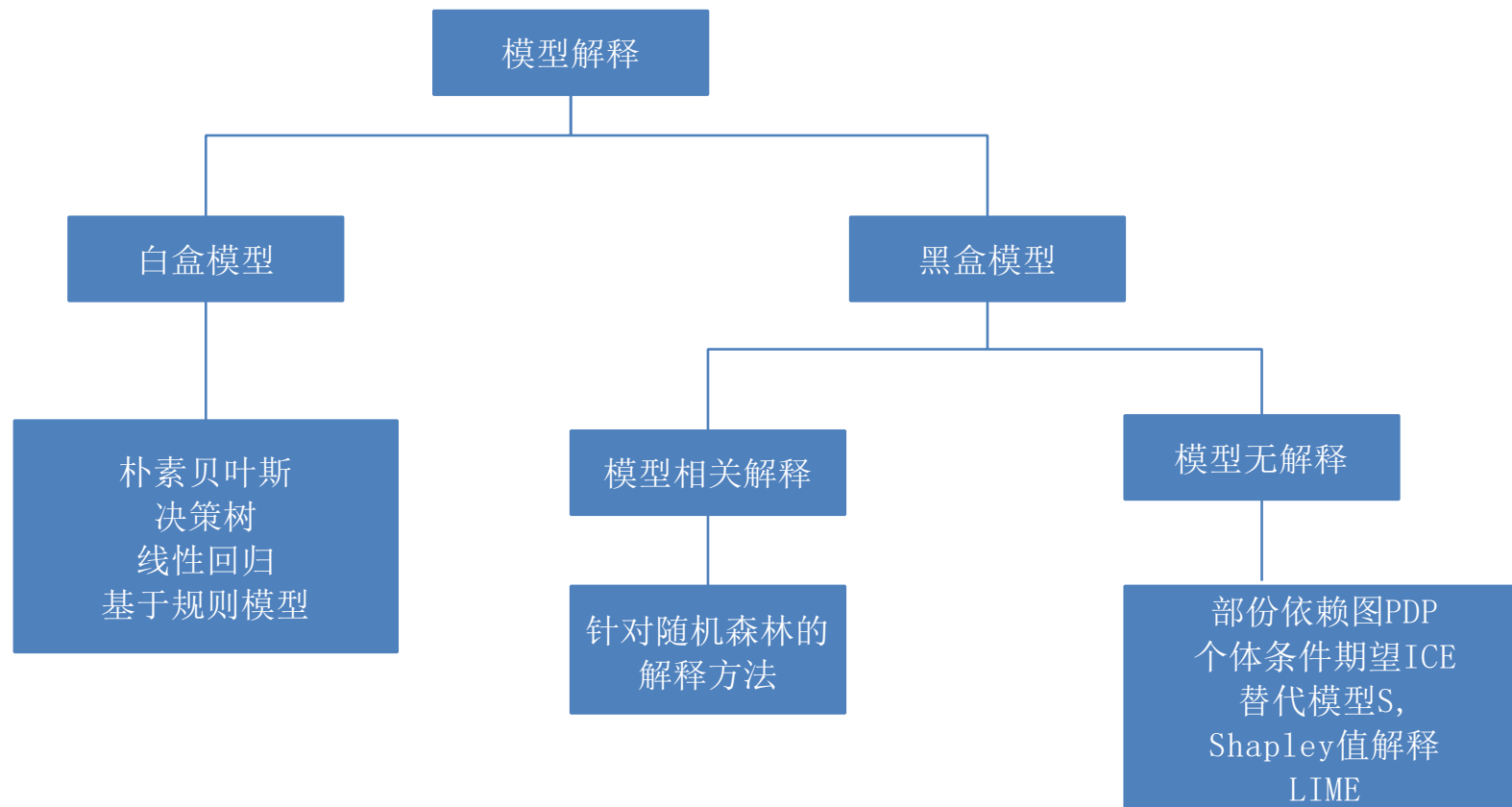
RACISM

SEXISM

Algorithms, as well as the learning process that decision-making algorithms follow, should be subjected to audits, in order to eliminate human bias.

Algorithm audits should be promoted and become mandatory. Governments should lead by example, informing citizens on how decisions are being made and explaining with transparency the whole algorithmic process. This should also become the case in the private sector.
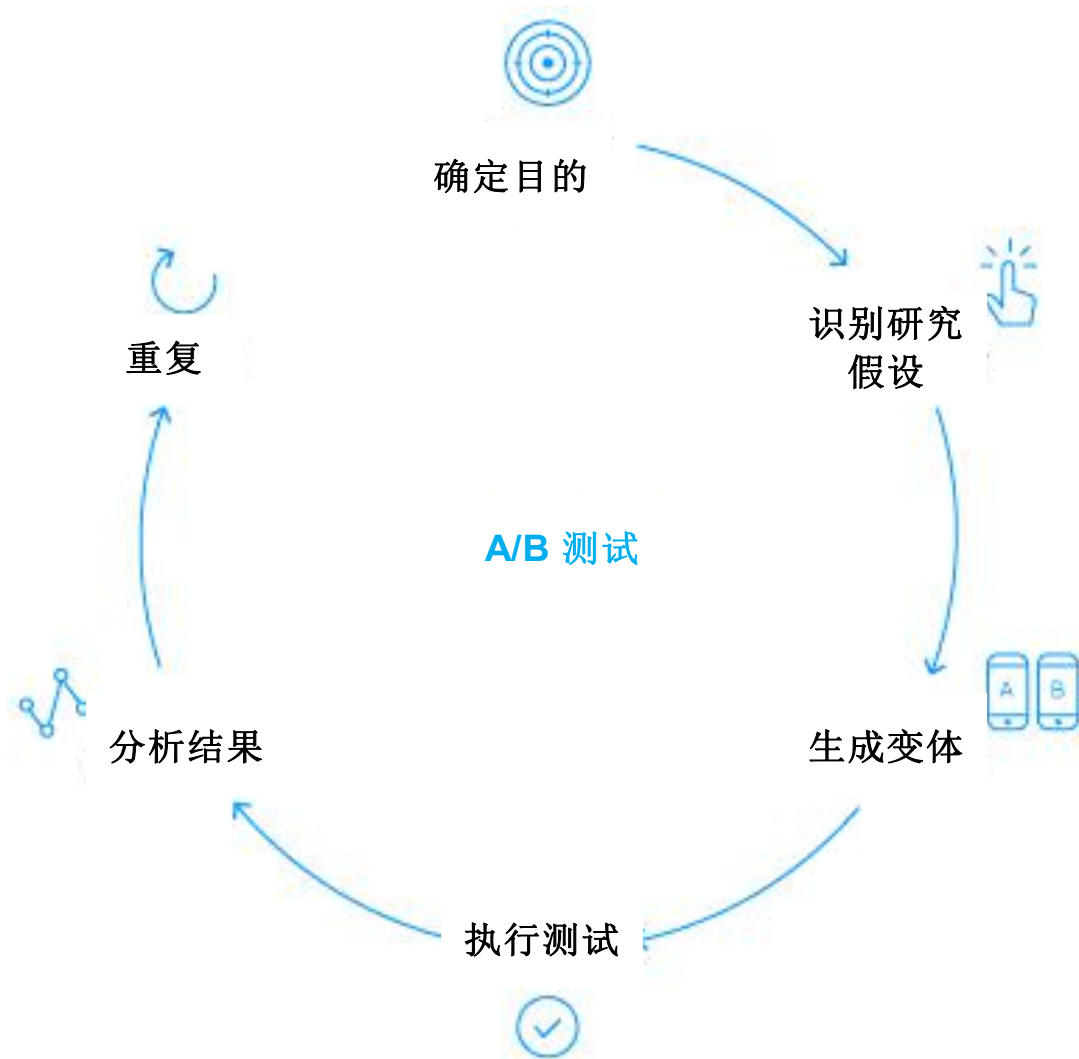
Societal impact assessment should be part of the design process of every decision-making algorithm, especially if these algorithms are part of public structures.

Individuals or groups could be the target of algorithmic discrimination, based on their characteristics — such as race, age, gender, sexual orientation or religion, among others. Simple features, such as names, are enough to allow algorithms make unfair decisions that could prevent individuals from getting a loan or being selected for a job.

" Through the use of algorithms, computers can process data, provide solutions to problems, and even make decisions for us "

eticas foundation

谋公司，2014年，自动筛选简历算法，5万简历+500个模型，男性

6.AI与机器学习的可解释性

模型解释

白盒模型

黑盒模型

朴素贝叶斯
决策树
线性回归
基于规则模型

模型相关解释

模型无解释

针对随机森林的
解释方法

部份依赖图PDP
个体条件期望ICE
替代模型S,
Shapley值解释
LIME

7. A/B测试

确定目的

识别研究
假设

A/B 测试

重复

分析结果

生成变体

执行测试

Randomized controlled trial (RCT)
大样本随机双盲对照试验

8.如何继续学习本章知识

数据产品开发

- 基于数据加工和数据柔术的数据产品开发
- 基于大数据的传统产品的创新

数据的生产要素作用

- 数据不仅是资源，更是资产
- 法律法规、伦理道德、安全隐私、偏见、可解释性

小结

1.本章定位与内容简介

2.本章学习提示及要求

3.幸存者偏差

4.辛普森悖论

5.算法歧视

6.AI与机器学习的可解释性

7.A/B测试

8.如何继续学习本章知识