

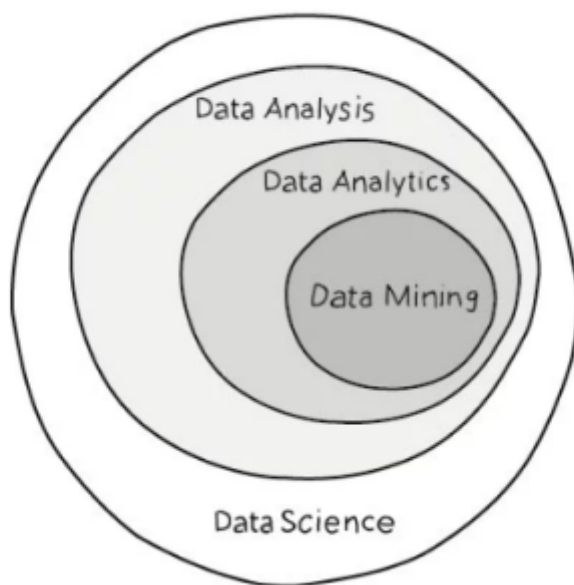
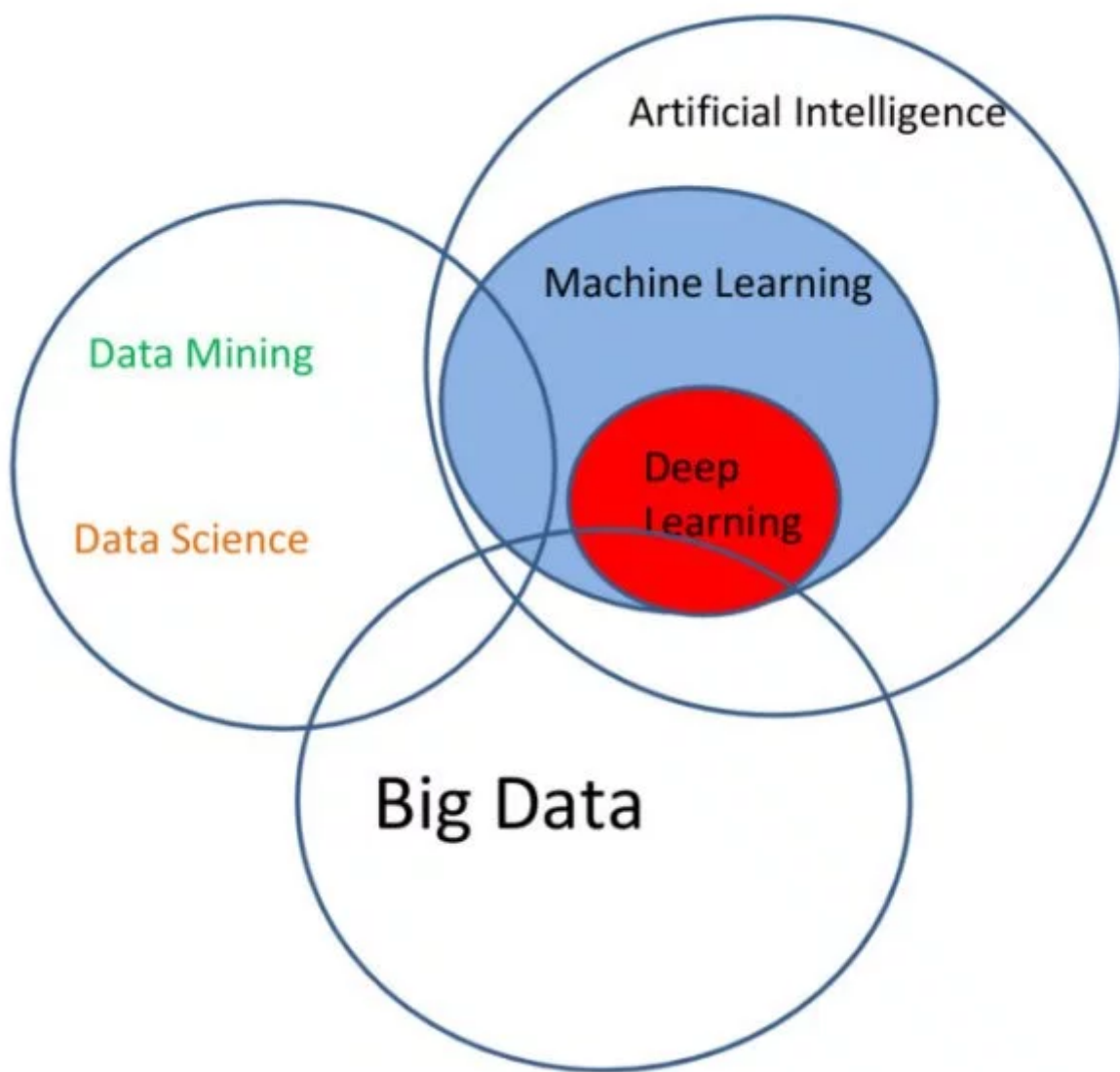
Lecture 1:Introduction

课程基本介绍

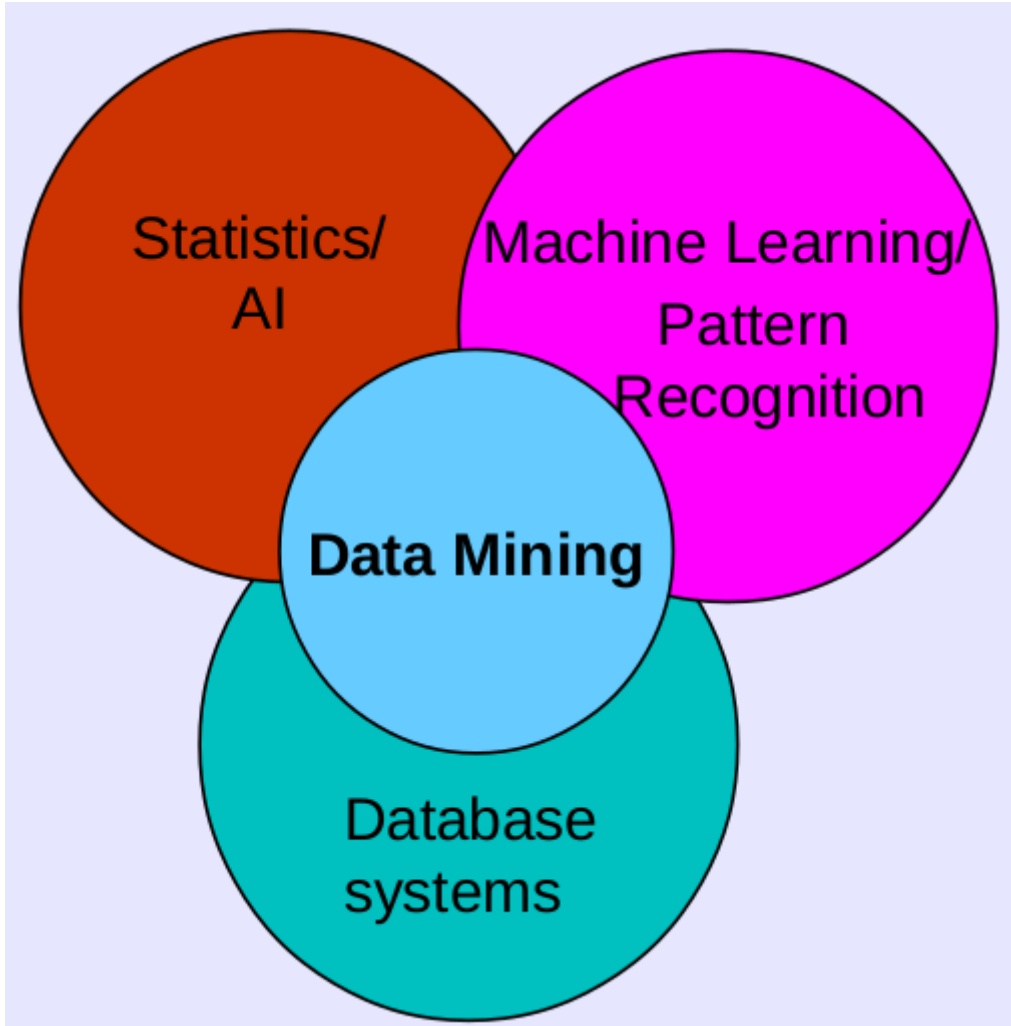
- CS422 探讨数据挖掘与机器学习和AI关系，更多放在与机器学习的关系上.
- 编程模块采用R语言.
- 学习几种基本算法,包括但不限于购物篮分析、最近邻居、决策树、频繁项目集、回归、分类、聚合.
- 目标：精通数据挖掘架构、词汇、技术方面，并可使用相应工具来构建模型并输出解释；可通过学习的算法部分实现评估使用合适的算法与工具实现输出.
- 包含课程讲座、编程作业、2-3次主题讨论会、考试(期中、期末).
- 需要一定的线性代数、统计学知识
- 准时交作业

Introduction:Placing data mining in context

- Data mining vs. machine learning vs. ...



- 从机器学习/AI中汲取目标识别、统计学、数据库系统
- 传统技术不适合原因
 - 数据量巨大
 - 数据高维度
 - 数据的异构、分布性特征



例子：

- 父母未知女孩怀孕收到避孕套广告
- Google邮箱通过扫描邮件实现广告推送
- 商场/实体通过用户会员卡形式推测用户购买行为
- Facebook广告问题
- 避免广告/隐私泄密，proton邮箱
- 各州通过法律保护用户隐私

Introduction:AI

- AI研究起源于Dartmouth College in 1956.
- 1960's:美国DoD捐款AI项目

- 1974年AI研究进展开始缓慢
- 1974-1980进入AI冬天
- 1987-1993进入第二AI冬天
- LISP:一种符号计算语言使用一阶逻辑
 - 亲属事实关系
 - 规则关系
 - 推理关系
- 现代AI
 - 1993-2011
 - 通过决策理论和经济学把AI带入到贝叶斯理论、马卡罗夫模型、优化、信息论和随机建模
 - 2011-
 - 深度学习:AI在计算机视觉、语音识别取得长足进步
 - 现代AI短板部分
 - 基于模式匹配,并非通过常识推理
 - 需要大数量集来理解使用哪种模型匹配
 - 人类大脑更敏捷,不需要大数据集
 - 灾难性的健忘
 - 对于人类简单的东西(视觉、语言)机器很难做对
 - 对于机器简单的东西(大数据量集匹配)人类很难最对
 - 例子:指示牌、指示牌上粘一个创口贴(AI模型是针对某个信息作出的训练,对于部分差异性信息无法准确识别)
 - 未来
 - 通用AI智能
 - 我们在下一个AI冬天的风口浪尖?
- AI:通用AI智能的挑战
 - 为了使AI确实有用,需要关注常识信息
 - 很多基于教会AI常识信息项目
 - 看到很多常识AI项目回归到一阶符合逻辑上很有趣
 -

```
(#$isa #BillClinton #UnitedStatesPresident)
```

"Bill Clinton belongs to the collection of U.S. presidents."

```
(#$genls #Tree-ThePlant #Plant)
```

"All trees are plants."

```
(#$capitalCity #France #Paris)
```

"Paris is the capital of France."

Sentences can also contain variables, strings starting with "?". These sentences are called "rules". One important rule asserted about the #isa predicate reads:

```
(#$implies  
  ($and  
    ($isa ?OBJ ?SUBSET)  
    ($genls ?SUBSET ?SUPERSET))  
  ($isa ?OBJ ?SUPERSET))
```

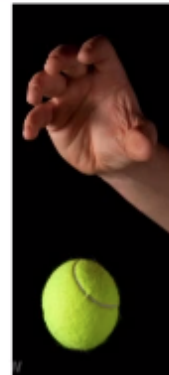
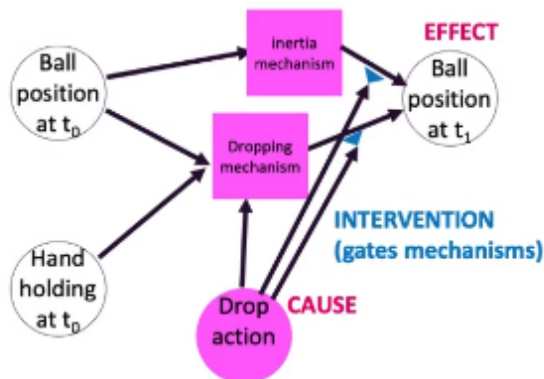
○ 因果关系

■ SPARSE DEPENDENCIES BETWEEN ABSTRACT VARIABLES

Also consistent with Baar's Global Workspace Theory (1997) of conscious processing.

Linguistic example:

"if I drop the ball, it will fall on the ground"



**An abstract outcome
can be predicted
accurately from very
few conditioning
abstract variables**



© CanStockPhoto.com - csp55707317

- 人会识别为因汽车爆胎停在路边的女士寻求帮助
- AI会识别为一个女士在路边招手
- 特斯拉问题，AI/车主 为车祸负责？