

三驾马车驱动人工智能

陈熙霖

中国科学院计算技术研究所



人工智能的再次复兴——新的平衡

人类的智慧与机器的智能

- 天资聪慧
- 阅历丰富
- 身强体壮
- 算法能力强(学习、泛化)
- 训练数据全
- 硬件跑得快



2017年Libratus , AlphaGo Zero
2016年AlphaGo
2011年Watson

AI复兴(1997-2006)
1997年深蓝夺冠

第二次高潮(1980-1987)
联结主义、五代机

第二次低谷(1987-1997)
五代机失败

黄金时期(1956-1967)
符号主义、启发式

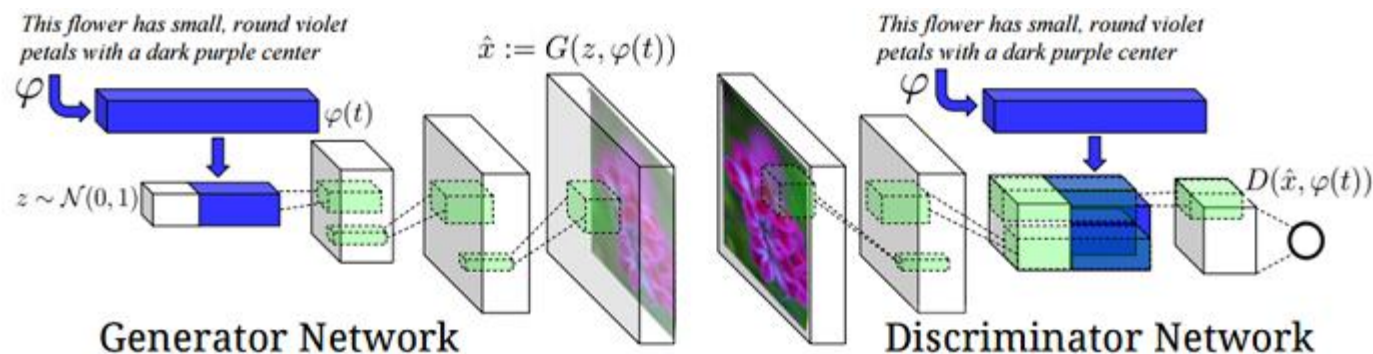
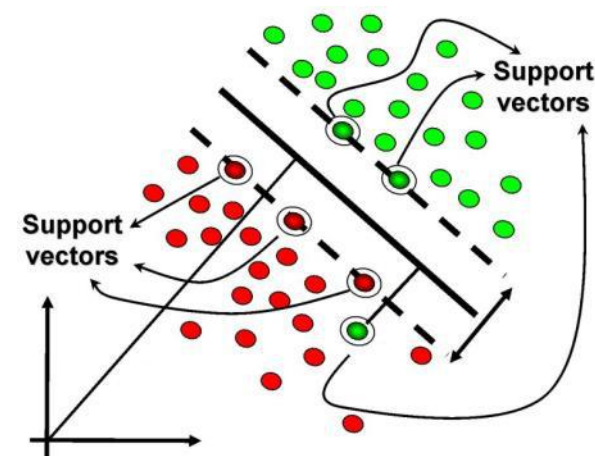
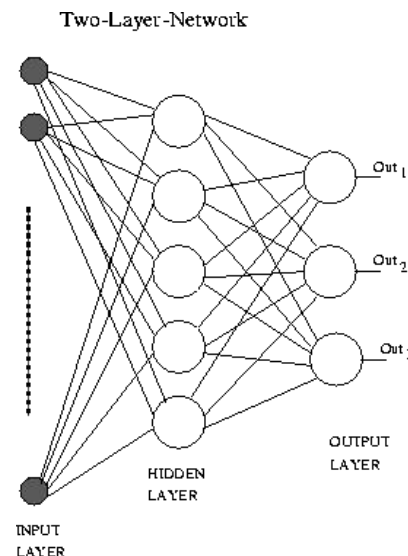
第一次低谷(1967-1980)
机器翻译的例子

1956年达特茅斯会议
开启AI

人工智能的九曲十八弯

背后的算法进步

- BP (Rumelhart et. al. 1986)
- SVM (Vapnik et. al. 1995)
- AdaBoost (Freund et. al. 1995)
- DNN (Hinton 2006)
- GAN (Goodfellow et. al. 2014)



算法与数据——AI从来没有拒绝过数据

- 数据使AI走出Toy World

- 鲁肃与吕蒙的故事

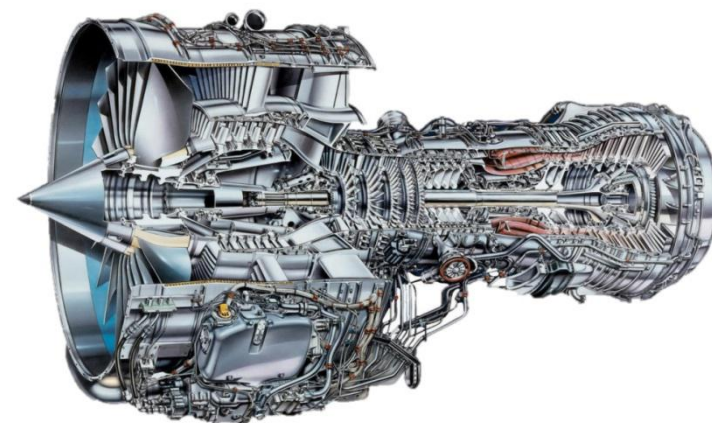
- 及鲁肃过寻阳，与蒙论议，大惊曰：“卿今者才略，非复吴下阿蒙！”
蒙曰：“士别三日，即更刮目相待，大兄何见事之晚乎！”

- 算法和数据——天资聪慧 vs. 见多识广

- 人工智能问题，维数空间大，受困于维数灾难——海量数据使得训练得以成为可能

- AI使数据发挥价值

- 从“垃圾 Scale”数据到发挥价值
 - 没有发动机，石油只能点灯



几个观点



- 采集的是数据，自生的也是数据
 - 左右互搏也是训练
 - 什么样的问题可以采用自生数据——数学家可以低头思考，天文学家必须仰望星空
 - 明确规则，封闭体系——各类博弈游戏
- 大规模训练数据使AI逃出实验室，强劲硬件平台使AI走向应用
 - 计算能力限制
 - 功耗约束
 - 从“算不了”、“算不起”到规模应用



未来——三驾马车在牵制与互动中前行

- 离开学习(数据)的神童(算法)不会成为智者(走不出实验室)
- 没有硬件支撑，AI/大数据也难以真正应用
- 缺乏分析的大数据只能是Large(垃圾) Scale的数据
- 没有新的应用需求，硬件发展也会迷失

谢谢！