**产生式系统的搜索**

**状态空间法(S,F,G)**(初始状态 操作符 目标状态)

如食人生番:动作op(b1,b2,b3)(人数,野人数,哪岸到哪岸)

状态(a1,a2,a3)(左岸人num,左岸野人num,船在左or右)

状态转换:

**搜索:**宽/深:后代放进open-后代是否目标-取点放进closed

**一致代价搜索**: Dijkstra算法。

**贪婪最优搜索**:仅利用启发函数作为代价。

**A\*算法**:

最优性:可采纳：树搜索的A\*，一致：图搜索的A\*

可采纳性：；一致性：

/深度优先:,

/一致代价,

/宽度优先:足够大，，

**计算复杂性**

**P类问题**:有多项式时间算法解决的**判定问题**(有)

**NP类问题**:对猜想存在多项式时间算法来验证的判定问题，非确定性多项式问题

**NP完全问题**:(NP中最难的问题)判定问题对所有其他判定问题有,则是NP完全的。

**NP难题**:完全是NP难的。至少和一样难。如TSP问题不是判定问题，但它是NP难题。

**约束满足问题和搜索**:

**约束满足问题**:变量X、值域D、约束C。

**回溯搜索：**一次赋值一个变量，合法时从该状态继续为下一变量赋值，没有合法值则返回父节点尝试新赋值/选择顺序：最少剩余值、最多约束项、最少约束值

**推理：**节点相容（变量在其值域中的所有取值均满足其一元约束）、边相容（变量在其值域中的所有取值均满足与该变量相关的二元约束）、k-相容、路径相容。

**搜索后推理：**在变量赋值后对它进行边相容检查，删除与该变量相关的未赋值变量不相容的值，冲突时回溯。

**局部搜索：**爬山法：从初始状态开始，不断向h(n)函数值增加的状态移动/随机梯度下降：一次用一个样例计算梯度

**博弈树搜索**:画与或图，我方或、MAX，对方与、MIN

**与或图**:注意弧线！有弧线是与，没弧线是或。

叶节点一定可解;非叶节点:若含有或(与)后继点,那么至少一个(全部)后继可解时,该非叶节点可解。

非叶节点无后继不可解/或(与)后继全不(有一个)可解。

**极大极小搜索**:我方取后继的MAX,对方取后继的MIN。

**剪枝**:用深搜。是的下界;是的上界。

:，中止该MIN层后续搜索；

:，中止该MAX层后续搜索。

效率:设深度为P,每节点有B个后继,生成端节点最小值。p为偶数:;p为奇数:

**蒙特卡洛树搜索**:权重(N是总访问次数)

、算法开始到现在,节点j被访问的次数和平均收益

流程:选择-拓展-模拟-反向传播

**谓词逻辑与归结原理:**命题是具有唯一真值的陈述句。

2.∧:合取(与) ∨:析取(或) →:蕴含 ↔:等价

3.为假当且仅当

4.永真/重言式:真；永假/矛盾式:假；可满足式:至少一个成真赋值；**非重言式的可满足式**:至少一个成真一个成假

5.吸收律:

,

**重要:**



,

6.原子公式:不含任何联结词的公式//子句:任何文字的析取式; 文字:原子或原子//合取**范式**:简单析取式构成的合取式//析取范式:简单合取式构成的析取式

**命题逻辑的推理规则**

7.附加:假言推理:

简化:;拒取式:

析取三段论:

假言三段论:

等价三段论:

构造性二难:

可做:前提引入/结论引入/置换规则比如换

**命题逻辑的归结方法**

8.子句:简单析取式,项是一个变量或者其否定。

子句集:合取范式中所有子句的集合, 换成“，”。

9.归结原理:如和都为真，那么为真。

10.归结式:归结

11.**归结（消解）法**:为了证明真,先转为证明该命题公式永假。提取其子句集，归结为空即可。

**谓词逻辑**12.个体词/谓词 任意/存在量词 约束出现/自由出现//函数是个体域到个体域的映射，不同于谓词

13.换名:辖域中约束出现的变量名换掉/替代:自由出现的个体变量名字可换掉

14.谓词公式永真称为逻辑有效/永真的，谓词公式永假成为不一致/不可满足。

15.**谓词演算公式**







注意：



16.**前束范式**:将量词均提到最左边

17.谓词推理:存在用常量替代,任意用变量or常量替代

**谓词逻辑归结原理**

18.skolem标准型:变前束范式;消去存在量词;略去全称

量词例子: |

19.**置换**，用代替

结合律,不满足交换律。

20.合一：,存在置换使。

21.**最一般合一**:是F的最一般合一置换,那F的任一

合一,都有置换有(从左比较对应项不同就换)

22.**归结过程**:归谬-化合取范式-子句集-归结-空子句。

**监督学习:线性回归**

**线性回归**:找函数族/找优化准则/找最优函数

,,

评价：，MSE（改平方）

**最小二乘**:,

**最大似然估计**:,

概率密度函数:

去找使似然函数值最大的参数值，结果同最小二乘法。

参数

，

**假设检验**:，

P小H1真，检验越合理。

确定系数:，越接近1拟合效果越好

**多项式回归**:,

**多元回归**:，

，

**监督学习:Logistic回归**

**L回归**:,

,

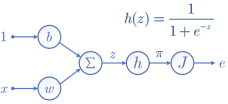
，

**对数似然**:

，需要

**经验损失**:，需要

交叉熵

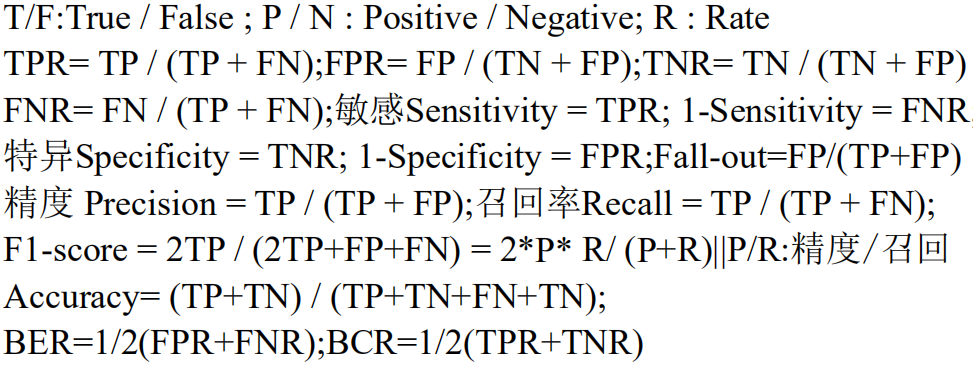


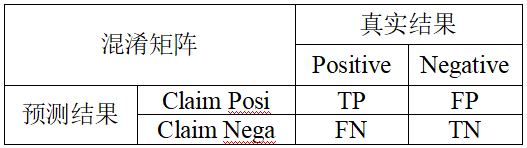
梯度下降找最小:



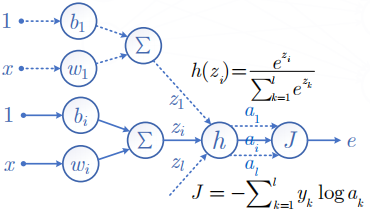
|**从输出算梯度**

**二分类的评价**





**softmax回归:**







**一元变多元**:

**前馈神经网络:输入单元—隐层单元—输出单元**

**线性单元**：

，

**Logistic单元**：



**Softmax**:





**双曲正切单元**：，

**整流线性单元(ReLU)**：

梯度消失问题略优。其他隐层单元：softplus

**输入单元—隐层单元(一般ReLU)—输出单元(按需选)**

**网络结构设计**:可表示区域的数量是深度的指数函数(d深度,每个隐层k单元m输入)：

**随机梯度下降**:一次用一个样例更新梯度(SGD)

**动量梯度下降**位移导是速度，速度导是受力，将位移类比参数,受力类比负梯度。参数更新公式:

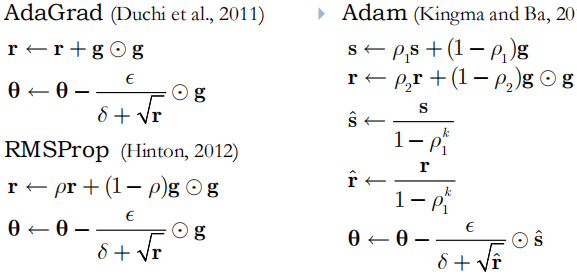
Nesterov动量：

**参数初始化策略**

破坏对称性:与同一个输入相连的具有相同激活函数的隐层单元参数不同；随机初始化:权重随机产生,偏置设为零

**学习速率调整方法**:学习速率反比于梯度累计平方和的根

线性衰减: //自适应衰减:



**正则化**:增加数据/提前终止/修改目标函数加入惩罚项

简化网络结构:随机扔掉一些节点/连接

**批次标准化**:SGD过程中对隐层输出标准化,可以用较大学习率/提升训练速度/收敛快/性能好(类似扔节点的正则化方法)/初始参数要求低

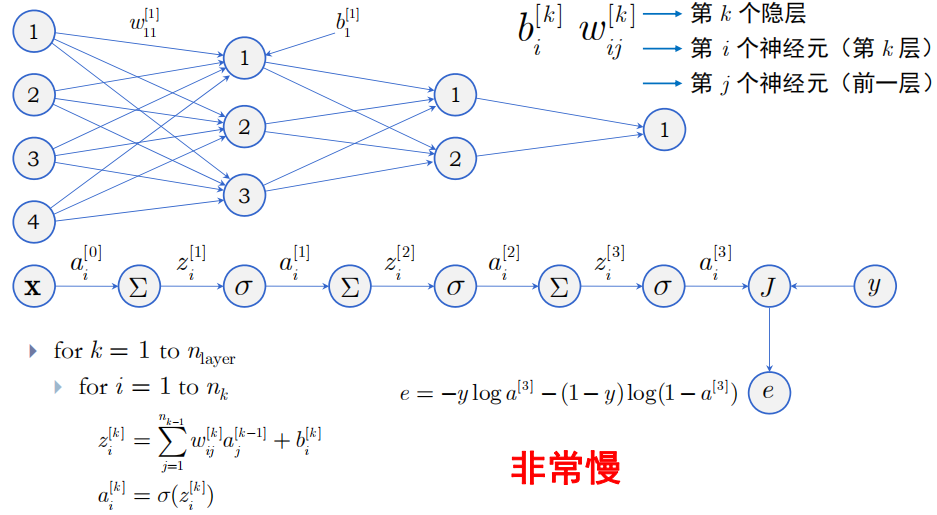


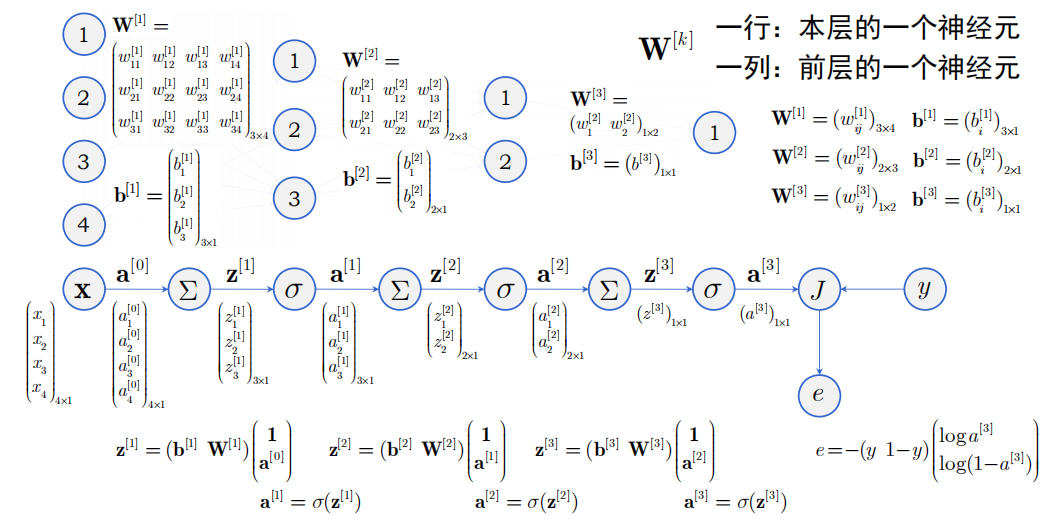
**卷积神经网络**

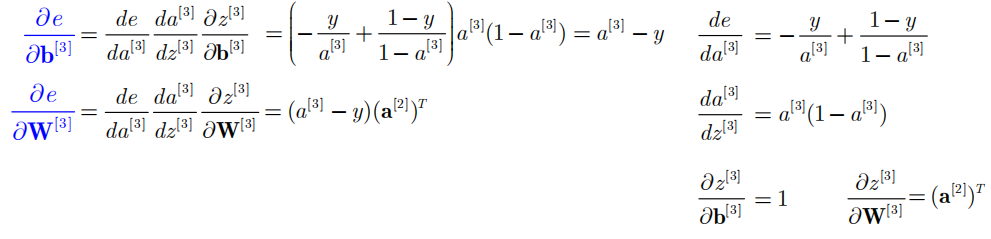
**卷积**:填充:允许卷积核超出边界(Samepadding)如边缘补0。

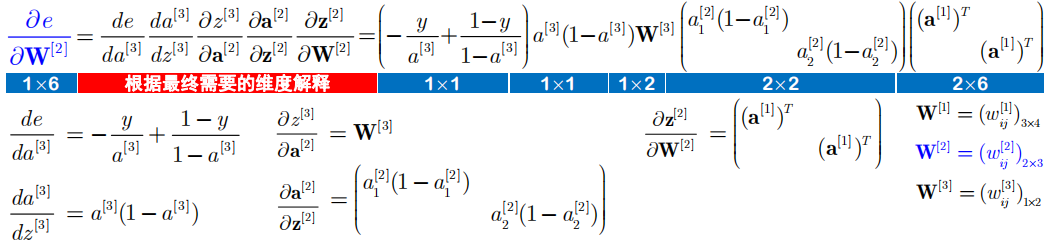
**稀疏连接**:可以用全连接的前馈网络实现,不参与卷积的点权重为0。节点N，卷积核m\*m，连接

**参数共享**:在不同位置卷积时像素点一样则结果一样(等变表示)，局部特征为制不重要，。









非共享卷积核/平铺卷积核:不共享参数/不完全共享参数

1\*1卷积核可以降维/多个卷积核升维/减少计算量

**池化**:用某一位置相邻输出的统计特征代替该位置的输出。

池化函数:最大池化/平均池化/随机池化

步长大于1的池化能够**降低输入规模**。

**卷积神经网络典型结构:**卷积层、池化层叠起来，再由全连接网络完成相关学习。卷积功能:特征提取—端到端学习。

**训练过程**:前向/后向,关键是池化和卷积单元的梯度计算。

数据的扩增:原图抽取小图、原图进行几何变换等。

图卷积模型:

**强化学习**

**马尔可夫过程：二元组** 状态空间

状态转移矩阵，

**马尔可夫回报过程：四元组**

状态转移产生回报。折现因子**;**期望状态回报,

累计回报:

状态价值函数:

贝尔曼期望方程:，

**马尔可夫决策过程:五元组**

行动导致状态转移产生回报行动空

间:；状态转移矩阵:其中；行动期望回报

**策略**:,s下选策略a的概率.

,

**状态价值**:

**行动价值**:由后续状态价值的加权和计算。



**贝..期望方程**:



矩阵形式:，

**动态规划**—策略评价(状态价值计算)、策略改进

，

最优状态价值:

最优行动价值:

最优a价值是同时刻最优q价值

最优策略(**贪心策略**):

策略迭代：已知策略pi，用动态规划计算出各个行动下的状态价值v（直至收敛，或者算一步改进一次），在某状态可行的所有策略中选择使得状态价值最大的那个（贪心策略），然后重复以上过程，得到一列状态、策略。

同步迭代:算完所有s更新一次策略。(不用算v至收敛)

异步迭代:算完一个s就更新一次策略。

**蒙特卡洛—**无P矩阵时状态价值预测、策略改进

已知观测片段，对某个状态，以其出现作为开始计算出这些幕各自的，对这些取平均值来估计v(s)。

首次访问/每次访问:作为起始的s是否第一次出现。

增量式蒙特卡洛:

定步长蒙特卡洛:

无模型时预测行动价值，

增量式:

定步长:

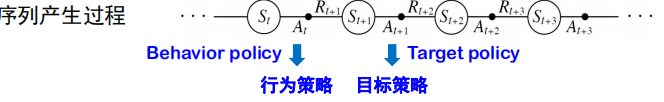
策略改进:

**贪心改进**:

**时序差分**—从部分序列学习

递推状态价值:

蒙从完整序列学习,适用带终止的决策过程/离线学习(必须获得完整片段来估计累计汇报)/非马更有效；时序从部分序列学习,可用于无终止状态的决策过程/在线学习,只需获取下一状态的及时回报/马环境更有效。

序列产生:

在线策略:目标策略和行为策略相同(SARSA)

离线策略:不同(Q-learning,E-SARSA)

当前状态,根据行为策略选择获得回报、下一状态,根据目标策略从产生,计算,迭代更新。不断循环以更新Q表。

**算法SARSA**:行为策略和目标策略都是



,

**算法Q-learning**:行为策略目标策略



**算法E-SARSA**:行为策略目标策略行动期望



**增量式价值近似**

**状态价值近似**:特征提取,函数族(线性函数),

优化准则,

优化方法SGD:

**行动价值近似**:特征提取,函数族(线性函数),

优化准则,

优化方法SGD:

**批量式价值近似(经验回放)**

**经验回放的SGD方法**:

交替进行采样和梯度下降至收敛

**状态价值的蒙特卡洛经验回放:**

积累经验:

从经验采样:，计算系数:

**状态价值的时序差分经验回放:**



**行动价值的时序差分经验回放:**



**深度强化学习**:用神经网络做价值近似,Deep Q-Network