**第八章 概述、人工神经网络**

1. 计算智能：概念性的东西考选择

（1）定义：关于在复杂和变化的环境中如何实现智能行为的自适应机制研究；综合了人工神经网络、进化计算、群智能和模糊系统。

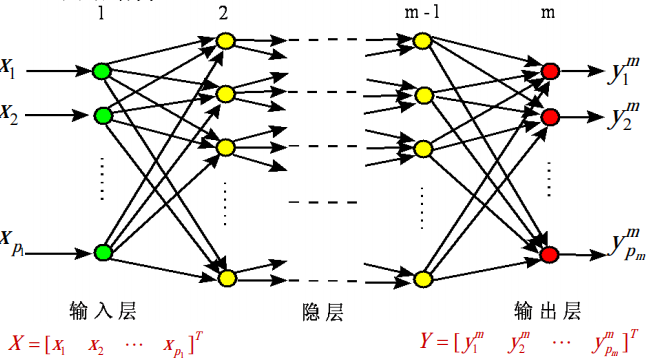
（2）包括：ppt 5-6

2. 人工神经网络ANN特征

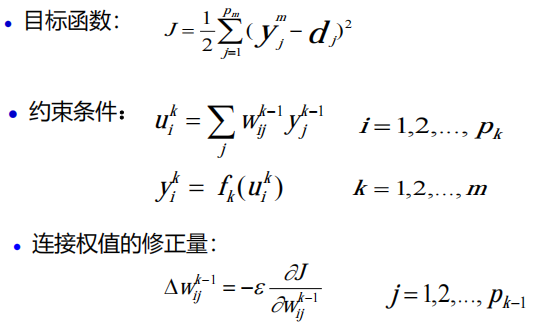
并行分布处理，非线性映射，通过训练进行学习，适应与集成，硬件实现

3. BP神经网络

（1）结构

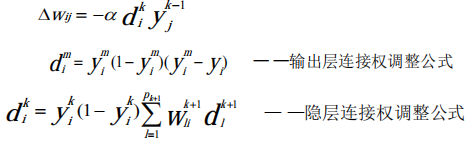


（2）基本思想



（3）为什么需要反向误差传播，误差如何反向传播（自己看看ppt，32-37）

从输出层方向计算到第一个隐层，按连接权值修正公式向减小误差方向调整网络的各个连接权值。



权重更新必须是从最后一层往前推算

4. Hopfield神经网络（主要是离散和连续，后面拓展不用看）（ssy注：我觉得这jb玩意不会考结构图啥的，甚至不会考，ppt47-56页）

（1）离散型：结合神经网络低线性的动态行为，单层的全反馈

（2）连续型：联想记忆的功能，达到稳态后输出一个和输入模式最为相近的模式

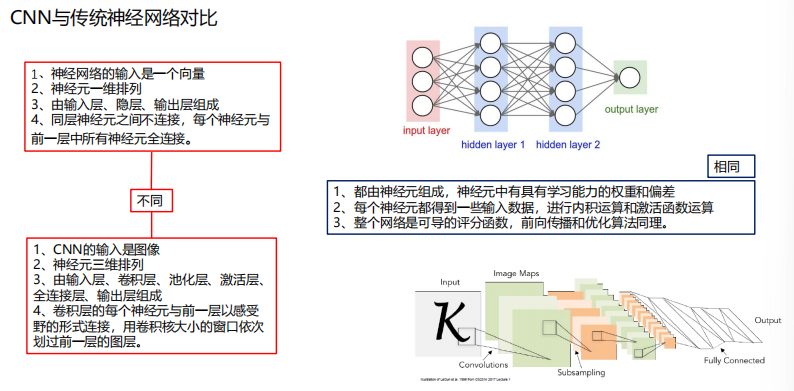
5. 车间调度问题不考

6. 卷积神经网络（从理解、特点出发）

（1）CNN是一个多层的神经网络，每层由多个二维平面组成，而每个平面由多个独立神经元组成。

C层为特征提取层（卷积层），S层是特征映射层（下采样层）。CNN中的每一个C层都紧跟着一个S层。

（2）CNN与ANN的区别



**（3）每个层的作用**★★★

a. 卷积层

• 特点：局部连接（区别于全连接），参数共享

• 优点：减少参数量，节约计算成本

b. 池化层

• 作用：减少数据量、减轻过拟合

• 特点：无参数运算，一旦选定池化滤波器尺寸、步长及最大/平均池化，这一层输入数据发生的改变就固定。

c. 激活层：非线性拟合作用，使网络模型拟合出的特征更复杂

常用的激活函数：f：激活函数



d. 全连接层

• 连接方式：该层神经元对前一层中所有数据是全部连接的。

• 最后一层全连接中，在神经元的输出端使用合适的损失函数，就能让单个神经元变成一个线性分类器。即最后一层全连接层的作用为分类器

CNN优缺点稍微理解

循环神经网络RNN：图理解一下，知道是神经序列上神经网络的展开；其问题：梯度爆炸

BPTT理解一下，持续性的反向误差传播（很难出题）

LSTM理解一下门控方式

transformer肯定不考

gans网络理解一下：学习是一种数据的分布，生成器、判别器能力的提升，优化方式怎么实现的

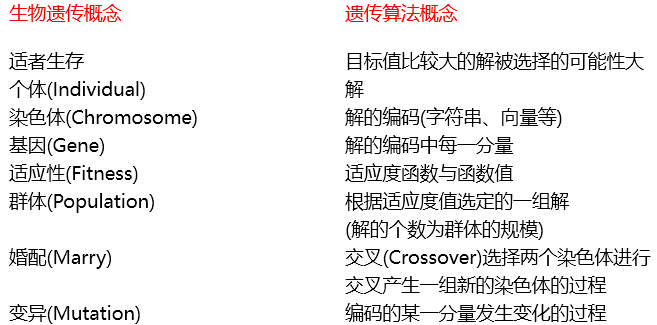
**第十章 进化计算，群智能算法**

1. 自然界中的进化：种群里的遗传性状在世代之间的变化

进化的实质：种群基因频率的改变

2. 遗传算法的机制：遗传算法是模仿生物遗传学和自然选择机理，通过人工方式构造的一类随机优化搜索算法，是对生物进化过程的一种数学仿真，是进化计算的一种最重要形式。

3. 生物遗传与遗传算法概念迁移★



4. 遗传算法的特点：一种全局优化概率算法

·无论是**线性的还是非线性**的，**离散的还是连续的**都可处理，可直接对结构对象进行操作

·利用**随机技术**指导对一个被编码的参数空间进行效率搜索；采用群体搜索策略，易于并行化

·仅用**适应度函数**来评估个体，并在此基础上进行遗传操作，使种群中个体之间进行信息交换。

·进化算子的各态历经性使得遗传算法能够非常有效地进行概率意义地**全局搜索**。

·对于各种特殊问题可以提供极大地灵活性来混合构造领域独立地**启发式**，从而保证算法地有效性。

5. 遗传算法的5个基本要素（三年之内不会重复）

1. 个体的参数编码形式；
2. 初始群体的设定；
3. 适应度函数的设计；
4. 控制参数的设定；
5. 遗传操作设计。

（1）编码：把问题的解的参数形式转换成基因链码的表示形式

位串编码：二进制编码（染色体也是二进制的），格雷编码，实数编码（旅行商、生产调度等）

（2）群体设定

①根据问题固有知识，把握最优解所占空间在整个问题空间中的分布范围，然后，在此分布范围内设定初始群体。

②随机产生一定数目的个体，从中挑选最好的个体加到初始群体中。不断迭代，直到初始群体中个体数目达到了预先确定的规模。

* 群体规模太小，遗传算法的优化性能不太好，易陷入局部最优解；群体规模太大，计算复杂。

（3）适应度函数的构造

适应度：个体在种群生存的优势程度度量，用于区分个体的“好与坏”

适应度函数的尺度变换：防止适应度计算的过程中适应度的值分布不合理或者难以体现个体的特性

①线性变换

②非线性变换

幂函数变换法

指数变换法

注意：适应度函数值必须非负。根据情况做适当的处理（如：加Cmax）

选择合适区间的尺度，解决欺骗问题（过早收敛：适应度超常；停滞现象：陷入局部峰值）

（4）参数的设定——选择

从当前群体中按照一定概率选出优良的个体，使它们有机会作为父代繁殖下一代子孙；个体适应度越高，其被选择的机会就越多。

**个体选择概率的分配方法：**

①适应度比例方法 或 蒙特卡罗法

完全由适应度大小来决定选择概率，各个个体被选择的概率和其适应度值成比例

②排序方法：按适应值大小从好到坏依次排列

分为：线性排序，非线性排序

**选择个体方法：**

①转盘赌选择：适应度值越好的个体被选择的概率越大；

按个体的选择概率产生一个轮盘，轮盘每个区的角度与个体的选择概率成比例

②锦标赛选择方法

每次从种群中取出一定数量个体（成为竞赛规模），然后选择其中最好的一个进入子代种群。重复该操作，直到新的种群规模达到原来的种群规模。

③和选择

：从规模为的群体中随机选取个体通过重组和变异生成个后代，再选取个最优的后代作为新一代种群；

：从个后代与其父体共中选取个最优的后代

④Boltzmann锦标赛选择

⑤最佳个体保存方法

（5）参数的设定——交叉（对两个/多个基因而言）cross，概率：

定义：在选择操作的基础上，根据一定的概率（称为交叉概率）进行交叉操作。

目的：能够在下一代产生新的个体，是遗传算法获取新的优良个体的最重要的手段。

但并不是所有被选择的个体都要进行交叉操作：过多：适应度好的被破坏；过少：搜索停滞不前

把两个父个体的部分结构进行替换重组，生成新个体。不同的个体编码方法可以有不同的算法。

**基本交叉算子：（二进制编码的）**

①一点交叉：随机设定一个交叉点，该点前或后的两部分结构进行互换

②二点交叉：随机设置两个交叉点，将两个交叉点之间的码串相互交换

③均匀交叉（一致交叉）：按照均匀概率抽取一些位，将被抽取位互换

**修正的交叉方法：**

①部分匹配交叉

②顺序交叉

**实数编码的交叉方法：**

①离散交叉：部分离散交叉：二进制的点式交叉；整体离散交叉：二进制的均匀交叉

②算术交叉：也分为部分和整体

（6）参数的设定——变异（对单个基因而言）mutate，概率：

变异概率是加大种群多样性的重要因素。

不是所有被选择个体都要进行变异操作：太小很难产生新个体，太大会使GA成为随机搜索

**整数编码的变异方法：**

①位点变异：随机挑选一个或多个基因座，并对这些基因座的基因值以变异概率作变动

②逆转变异：随机选择两点，将其基因值以逆序插入原位置

③插入变异：将码随机插入某两点中间

④互换变异：随机选取染色体的两个基因进行简单互换

⑤移动变异：随机选取一个基因，向左或者向右移动一个随机位数

⑥自适应变异：类似位点变异，变异概率随群体中个体的多样性程度而自适应调整

**实数编码的变异方法：**

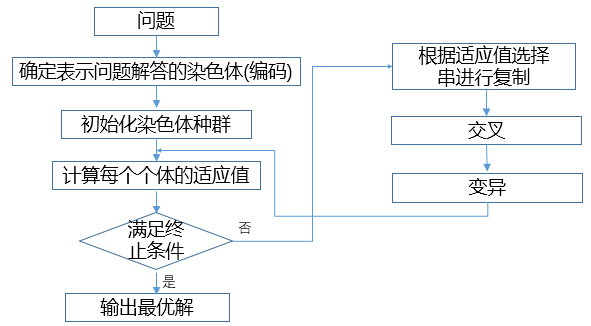
①均匀性变异

②正态性变异

③非一致性变异

④自适应变异

6. 遗传算法基本流程（看图说话）



7. 遗传算法的改进（了解即可，不考方法、公式，考的概率低）

（1）双倍体遗传算法

基本思想：采用**显性**和**隐性**两个染色体同时进行进化（基本遗传算法是单倍体）

优势：可以记忆以前有用的基因块，延长了有用基因块的寿命，提高了算法的收敛能力，在变异概率低的情况下能保持一定水平的多样性

（2）双种群遗传算法

基本思想：多种群同时进化，并在每一代运行结束后，交换种群之间优秀个体所携带的遗传信息，以打破种群内的平衡态，达到更高的平衡态

优势：有利于跳出局部最优

（3）自适应遗传算法（AGA）

当种群各个体适应度趋于一致或者趋于局部最优时，使和增加；

当群体适应度比较分散时，使和减少，以利于优良个体的生存。

对于适应度高于群体平均适应值的个体，选择较低的和，使得该解得以保护进入下一代；

对低于平均适应值的个体，选择较高的和值，使该解被淘汰。

8. 差分进化算法DE(Differential evolution)：基于**实数编码**的具有**保优思想**的**贪婪**遗传算法（了解基本思想、优缺点）

①参数编码：实数编码，**不必进行数制转换**，可直接在解的表现型上进行进化操作

②初始种群的产生：可以从给定边界约束内的值中随机选择

③适应度函数的设计：具体设计方法与遗传算法相似

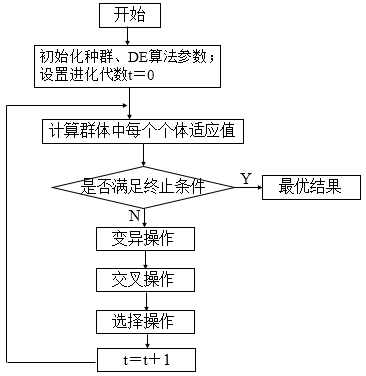
④差分操作设计：

·变异（F表示缩放比例因子，实常数）（选择可能出：哪种是差分操作中的变异？）

·交叉：二项式（bin）交叉、指数（exp）交叉

·选择：贪婪准则去选择，以产生下一代个体

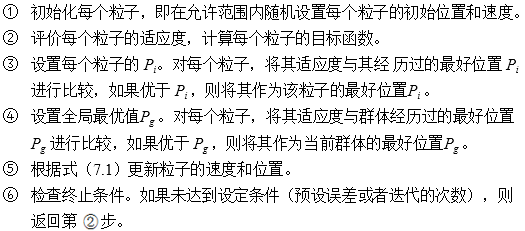
⑤控制参数设置：种群规模，缩放比例因子，交叉概率，最大迭代代数，终止条件



9. 粒子群优化算法PSO（去年考了）

**（1）原理**：将每个个体看作*n*维搜索空间中一个没有体积质量的粒子，在搜索空间中以一定的速度飞行，该速度决定粒子飞行的方向和距离。所有粒子还有一个由被优化的函数决定的适应值。

**（2）算法流程**

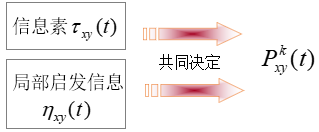


10. 蚁群算法ACO（最早解决TSP问题）★★★★★一定会考

（1）基本原理：蚂蚁在运动过程中，根据各条路径上的信息素决定转移方向。

（2）转移概率

表示在*t*时刻蚂蚁 *k* 选择从元素(城市) *x* 转移到元素(城市) *y* 的概率，也称为随机比例规则。

，

***α* 值越大**：该蚂蚁越倾向于选择其它蚂蚁经过的路径，该状态转移概率越接近于贪婪规则。

当 *α =* 0时：不再考虑信息素水平，算法就成为有多重起点的随机贪婪算法。

当 *β =* 0时：算法成为纯粹的正反馈的启发式算法。

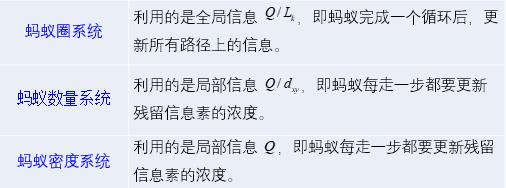
（3）信息素浓度消散规则

用参数1-*ρ*表示信息素消逝程度，蚂蚁完成一次循环，各路径上信息素浓度消散规则为：



（4）浓度更新规则

有三种蚁群的信息素浓度更新规则：（第一种最好，）



（5）**参数分析**★★★★重点看

**信息素启发因子*α***

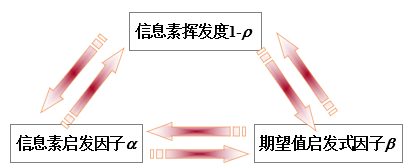
* 反映了蚁群在路径搜索中随机性因素作用的强度；
* *α*值越大，蚂蚁选择以前走过的路径的可能性越大，搜索的**随机性减弱**；
* 当*α* 过大时会使蚁群的搜索过早陷于局部最优。

**期望值启发式因子***β*

* 反映了蚁群在路径搜索中先验性、确定性因素作用的强度；
* *β*值越大，蚂蚁在某个局部点上选择**局部最短路径的可能性越大**；
* 虽然搜索的收敛速度得以加快，但蚁群在最优路径的搜索过程中随机性减弱，易于陷入局部最优。

**信息素挥发度1-***ρ*

* 当要处理的问题规模比较大时，会使那些从来未被搜索到的路径(可行解)上的信息量减小到接近于0，因而降低了算法的全局搜索能力；
* 而且当1-*ρ* 过大时，以前搜索过的路径被再次选择的可能性过大，也会影响到算法的**随机性能和全局搜索能力**；
* 反之，通过减小信息素挥发度 1-*ρ* 虽然可以提高算法的随机性能和全局搜索能力，但又会使算法的**收敛速度降低**。



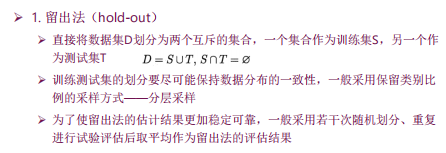
**第十一章 机器学习**

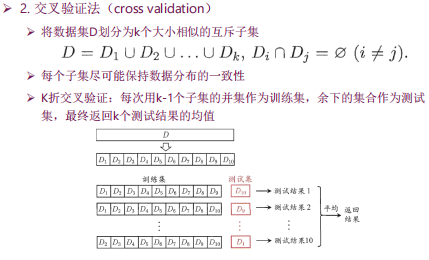
**1. 几种评估方法**★★

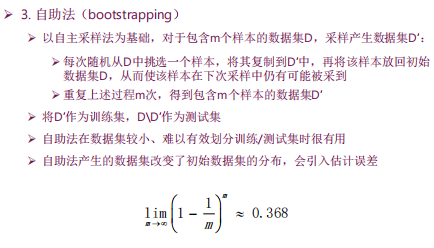
留出法：分层采样

交叉验证法(K折交叉验证)：每次用第k个子集作测试集，剩下的作训练集，最终返回k个测试结果的均值

自助法：在数据集较小、难以有效划分训练/测试集时很有用；会引入误差

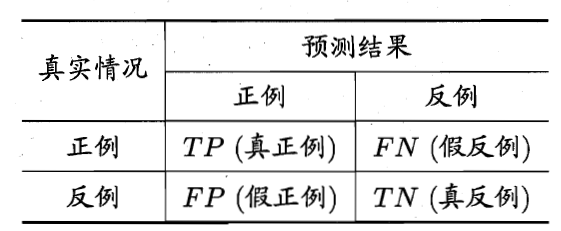


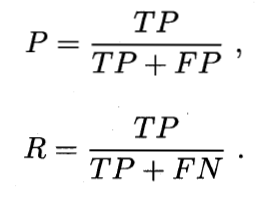
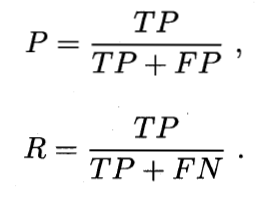




**2. 性能度量的几个指标**（AUC面积，ROC curve等）★★

（1）查准率 & 查全率：



查准率：查全率：

（2）ROC曲线：受试者工作特征，将测试样本进行排序，最可能是正例的排在最前面，最不可能是正例的排在最后面，设置分类阈值（截断点），将样本分为两部分（前一半做正例，后一半做反例）；排序本身质量的好坏，体现了学习器在不同任务下的期望泛化能力。（对角线：随机模型）

* + **横轴**：假正例率FPR false positive rate
  + **纵轴**：真正例率TPR true positive rate

若一个学习器的ROC曲线被另一个学习器的ROC曲线完全“包住”，则可断言后者的性能优于前者。

**AUC value**：area under ROC curve

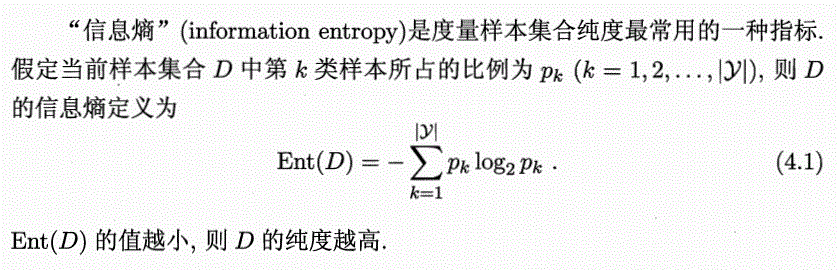
ROC曲线的特性：当测试集中的正负样本的分布变化的时候，ROC曲线能保持不变

3. 线性模型：理解，梯度下降，线性回归的意思，看一眼

4. 分类：理解一下

5. 决策树：基本流程，基本思想的理解

**6. 决策树中最优的属性，指标：信息熵-信息增益-基尼指数**（适合出小题）自己看一下

对于二分类任务：|y|=2

信息熵越小，纯度越高（熵是混乱度，越大越混乱）

7. 决策树预剪枝、后剪枝优缺点

①预剪枝的思想：在决策树生成过程中，对每个结点在划分前先进行估计，若当前结点的划分**不能**带来决策树泛化性能提升，则停止划分并将当前结点标记为叶结点。

（降低过拟合，减少时间开销；欠拟合风险增加）

②后剪枝的思想：先从训练集生成一棵完整的决策树，然后自底向上地对非叶结点进行考察，若将该结点对应的子树替换为叶结点**能**带来决策树泛化性能提升，则将该子树替换为叶结点。

（保留了更多分支，欠拟合风险小，泛化能力好；时间开销大）

8. 支持向量机：从概念角度出题，eg：支持向量的作用，映射时和函数引入的好处，软间隔之后怎么进行处理

9. 聚类：性能度量：有参考模型无参考模型，无参考模型的指标：想要性能好，DBI（无专家系统） RAND（有专家系统）等指数是大还是小，距离选择

10. Knn：K均值，伪代码理解