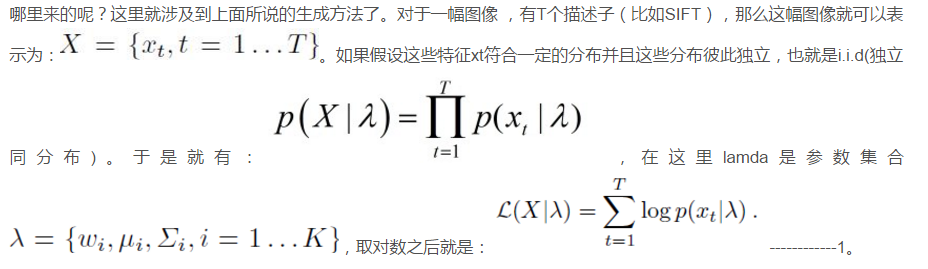
**机器学习各类算法适应场景**

https://blog.csdn.net/u010159842/article/details/52918815 Fisher Vector和GMM原理

GMM K个高斯分布的线性组合逼近分布模型

Fisher vector本质上是用似然函数的梯度vector来表达一幅图像，这个梯度的意义describes the direction in which parameters should be modified to best fit the data，说白了就是数据拟合中对参数调优的过程。



然后对公式1求偏导（对权重，均值，标准差求偏导），得到fisher vector。由于每一个特征是d维的，需要K个高斯分布的线性组合，一个Fisher vector的维数为（2\*d+1）\*K-1维。

*经过fisher vector的编码，大大提高了图像特征的维度，能够更好的用来描述图像。FisherVector相对于BOV的优势在于，BOV得到的是一个及其稀疏的向量，由于BOV只关注了关键词的数量信息，这是一个0阶的统计信息；FisherVector并不稀疏，同时，除了0阶信息，Fisher Vector还包含了1阶(期望)信息、2阶(方差信息)，因此FisherVector可以更加充分地表示一幅图片。*

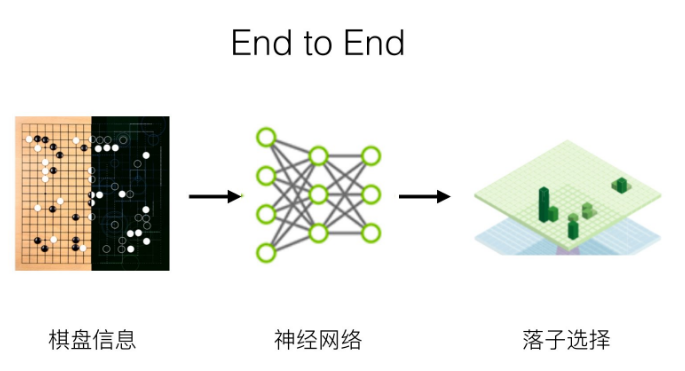
**SVM推导**

**LR推导**

**对alphago的理解**

围棋19\*19网格点，每个网格点三种状态。总的叫做状态空间复杂度。

监督学习 增强学习 人工设定



**XGBOOST原理** 包括优点，损失函数，最优化的方法等

**最小二乘与极大似然函数的关系？从概率统计的角度处理线性回归并在似然概率为高斯函数的假设下同最小二乘简历了联系**

L2正则的本质？限制解空间范围，缩小解空间，控制模型复杂度

怎么判断特征的重要性

**accuracy和precision、recall这些一样吗？AUC的解释**

roc曲线 roc曲线上每个点反映着对同一信号刺激的感受性。横轴fpr假正类率,纵轴tpr真正类率。

FPR=FP/(FP+TN) TPR=TP/(TP+FN) accuracy=(TP+TN)/(all)

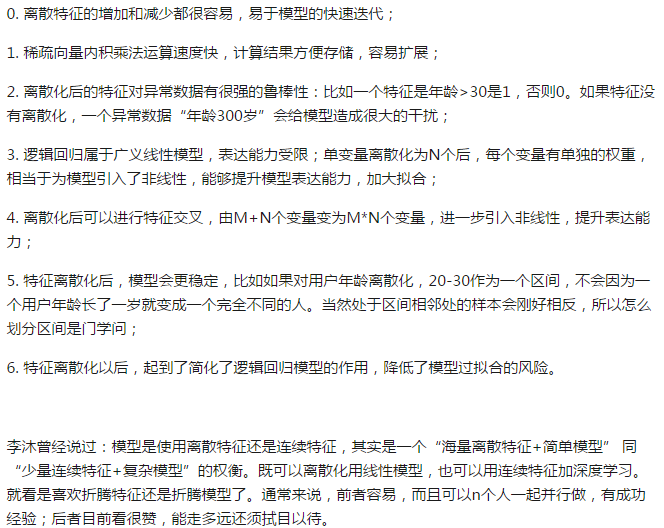
AUC(Area under Curve)：Roc曲线下的面积，介于0.1和1之间。Auc作为数值可以直观的评价分类器的好坏，值越大越好。

**首先AUC值是一个概率值，当你随机挑选一个正样本以及负样本，当前的分类**[**算法**](http://lib.csdn.net/base/datastructure)**根据计算得到的Score值将这个正样本排在负样本前面的概率就是AUC值，AUC值越大，当前分类算法越有可能将正样本排在负样本前面，从而能够更好地分类。**

**为什么使用ROC曲线**既然已经这么多评价标准，为什么还要使用ROC和AUC呢？因为ROC曲线有个很好的特性：当测试集中的正负样本的分布变化的时候，ROC曲线能够保持不变。在实际的数据集中经常会出现类不平衡（class imbalance）现象，即负样本比正样本多很多（或者相反），而且测试数据中的正负样本的分布也可能随着时间变化。

**keras底层用TensorFlow和theano时，代码有何不同？TensorFlow原理、流程图，session是啥？**

**为啥LR的输入特征一般是离散的而不是连续的**



了解各种优化算法不？梯度下降和随机梯度下降的区别？牛顿法和拟牛顿法的区别？为啥提出拟牛顿？因为牛顿法涉及海塞矩阵，它的逆矩阵求解很麻烦

 KNN的使用场景