* 在卷积神经网络计算中，已知输入特征层大小为32x32x64, 使用标准卷积计算，带偏置项，卷积核大小为3\*3，输出特征层数目为64，请问卷积层的参数个数为？

每个Filter有3\*3\*64个参数（64为# Channels）。总共有64个Filter（输出特征层数为64）。共3\*3\*64\*64个参数。

详细公式：<https://blog.csdn.net/gaishi_hero/article/details/81512404>

* 什么数据集不适合用深度学习？

1. 数据集太小，数据样本不足时，深度学习相对其它机器学习算法，没有明显优势。
2. 数据集没有局部相关特性，目前深度学习表现比较好的领域主要是图像／语音／自然语言处理等领域，这些领域的一个共性是局部相关性。图像中像素组成物体，语音信号中音位组合成单词，文本数据中单词组合成句子，这些特征元素的组合一旦被打乱，表示的含义同时也被改变。对于没有这样的局部相关性的数据集，不适于使用深度学习算法进行处理。举个例子：预测一个人的健康状况，相关的参数会有年龄、职业、收入、家庭状况等各种元素，将这些元素打乱，并不会影响相关的结果。（如果特征顺序打乱并不会影响结果，则不适合用深度学习）

* 在二分类问题中，当测试集的正例和负例数量不均衡时，以下评价方案哪个是相对不合理的（ A ）（假设precision=TP/(TP+FP),recall=TP/(TP+FN)。）

A: Accuracy:(TP+TN)/all

B: F-value:2\*recall\*precision/(recall+precision)

C: G-mean:sqrt(precision\*recall)

D: AUC:ROC曲线下面积

对于分类器，主要的评价指标有precision，recall，F-score，以及ROC曲线等。

在二分类问题中，我们主要关注的是测试集的正样本能否正确分类。当样本不均衡时，比如样本中负样本数量远远多于正样本，此时如果负样本能够全部正确分类，而正样本只能部分正确分类，那么(TP+TN)可以得到很高的值，也就是Accuracy是个较大的值，但是正样本并没有取得良好的分类效果。因此A选项是不合理的。在样本不均衡时，可以采用BCD选项方法来评价。

* 9、在其它条件不变的前提下，以下哪种做法容易引起机器学习中的过拟合问题（ D ）

A 增加训练集数量

B 减少神经网络隐藏层节点数

C 删除稀疏的特征

D SVM算法中使用高斯核/RBF核代替

机器学习中发生过拟合的主要原因有：

（1）使用过于复杂的模型；

（2）数据噪声较大；

（3）训练数据少。

由此对应的降低过拟合的方法有：

（1）简化模型假设，或者使用惩罚项限制模型复杂度；

（2）进行数据清洗，减少噪声；

（3）收集更多训练数据。

本题中，A对应于增加训练数据，B为简化模型假设，C为数据清洗（稀疏特征即很大部分特征值都为0）。D选项中，高斯核的使用增加了模型复杂度，容易引起过拟合。选择合适的核函数以及软边缘参数C就是训练SVM的重要因素。一般来讲，核函数越复杂，模型越偏向于过拟合；C越大模型越偏向于过拟合，反之则拟合不足。

* 假设你需要调整超参数来最小化代价函数（cost function），会使用下列哪项技术？ D

A.穷举搜索

B.随机搜索

C.Bayesian优化

D.都可以

* 关于朴素贝叶斯分类算法，描述正确的是：A

A.它假设属性之间相互独立

B.根据先验概率计算后验概率

C.对于给定的待分类项X={a1,a2,...,an}，求解在此项出现的条件下各个类别 yi 出现的概率，哪个P(yi|X)最大，就把此待分类项归属于哪个类别。

D.有最小错误率判断规则和最小风险判断规则

B应该加一句“根据贝叶斯公式”

C的解释：简单点回复吧，这是朴素贝叶斯，p(yi|X)可能为0，甚至可能X的一组特征在训练集中没有同时出现，此时P(yi|X)如何计算。所以想要求X出现的概率还是要根据后验来。

* 模式识别中，不属于马式距离较之于欧式距离的优点的是（ ）

欧氏距离是在N维空间中两个点的真实距离；马氏距离表示数据的协方差距离。

而欧式距离的特征是：平移不变性、旋转不变性。

马式距离的特征则是：平移不变性、旋转不变性、尺度不变性、不受量纲影响、考虑了模式分布。

<https://www.zhihu.com/question/20852004> 关于协方差的解释（两个变量的同向程度）

* 统计模式分类问题中，当先验概率未知时，可以使用(A)

1. 最小最大损失准则
2. 最小误判概率准则
3. 最小损失准则
4. N-P判决

链接：https://www.nowcoder.com/questionTerminal/34017bb5168b448abf3bb92231b70ea3

来源：牛客网

A. 考虑p(wi)变化的条件下，是风险最小

B. 最小误判概率准则， 就是判断p(w1|x)和p(w2|x)哪个大，x为特征向量，w1和w2为两分类，根据贝叶斯公式，需要用到先验知识

C. 最小损失准则，在B的基础之上，还要求出p(w1|x)和p(w2|x)的期望损失，因为B需要先验概率，所以C也需要先验概率

D. N-P判决，即限定一类错误率条件下使另一类错误率为最小的两类别决策，即在一类错误率固定的条件下，求另一类错误率的极小值的问题，直接计算p(x|w1)和p(x|w2)的比值，不需要用到贝叶斯公式\_

简单概括：所以可以理解为：先验概率未知，其实就是说不能用生成模型，只能用判别模型。

* 4、如果以特征向量的相关系数作为模式相似性测度，则影响聚类算法结果的主要因素有（B ）

A已知类别样本质量

B分类准则

C量纲

A选项，类别已知暂且不说，样本质量是不会影响聚类结果的。因为聚类的任务只是把数据按照相似性原则进行划分，不存在分类问题中由于训练集样本存在噪声数据，从而影响分类结果的情况。此外，在类别已知的情况下，直接按照样本的类别标签进行聚类就可以了，用不到复杂的聚类算法，所以也就不存在影响聚类算法结果这么一说了；

B选项，分类准则是指选取什么特征将该特征相似的数据聚为一类，这个会直接影响到样本聚类的结果；

C选项，两个特征向量的相关系数与其量纲无关，故不选。

* 6、以下( )不属于线性分类器最佳准则？

A感知准则函数

B贝叶斯分类

C支持向量机

DFisher准则

何为线性分类器？ <https://www.zhihu.com/question/30633734>

线性分类器有三大类：感知器准则函数、SVM、Fisher准则，而贝叶斯分类器不是线性分类器。

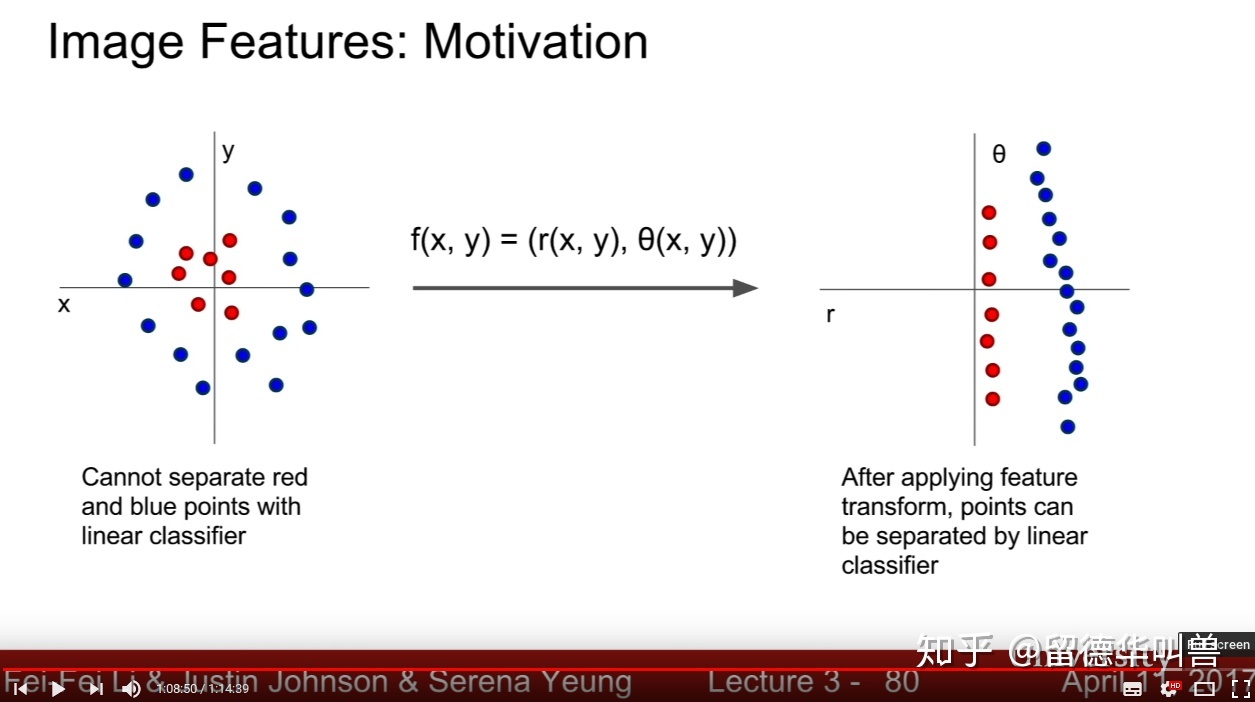
感知准则函数 ：准则函数以使错分类样本到分界面距离之和最小为原则。其优点是通过错分类样本提供的信息对分类器函数进行修正，这种准则是人工神经元网络多层感知器的基础。

支持向量机 ：基本思想是在两类线性可分条件下，所设计的分类器界面使两类之间的间隔为最大，它的基本出发点是使期望泛化风险尽可能小。（使用核函数可解决非线性问题）

Fisher 准则 ：更广泛的称呼是线性判别分析（LDA），将所有样本投影到一条远点出发的直线，使得同类样本距离尽可能小，不同类样本距离尽可能大，具体为最大化“广义瑞利商”。

根据两类样本一般类内密集，类间分离的特点，寻找线性分类器最佳的法线向量方向，使两类样本在该方向上的投影满足类内尽可能密集，类间尽可能分开。这种度量通过类内离散矩阵 Sw 和类间离散矩阵 Sb 实现。

使用线性分类器分类一组线性不可分数据：高维映射



* 

注意最大间隔就是垂直平分线

* 