## 深度学习损失函数的创新设计

我们的方法属于非参数判别方法，它倾向于由所提供样本数据直接求出在某一准则函数下的最优参数，这种方法必须由分类器设计者首先确定准则函数，并根据样本数据和该函数最优的原理求出函数的参数，因此关键是准则函数（损失函数Loss的设计）



1. **增加Fisher类别结构的损失函数**

采用fisher准则度量类别结构，使得CNN抽取特征时考虑这个结构准则。

Sw为类内距离和 Sb为类间距离，显然类内距离越小越好，类间距离越大越好。

这样损失函数变为：

1. **增加良好流形结构的损失函数**

采用流形结构的良好性，使得CNN抽取特征时考虑这个结构准则。

（1）

（2）

JM为流形重构误差，参考LLE线性重构和ISOMAP测地距离重构。公式（2）增加构造的基本向量的正交性，越正交越好。

这样损失函数变形为：

1. **增加特征空间的稀疏性的损失函数**

抽取的特征空间可能是高维的，但越稀疏越好。稀疏性(公式（1）)可以用来度量样本的多样性，越稀疏，样本的多样性越好。其次，可以用少量的特征来表述，与简单性原则一致。

(1)

(2)

**公式（2）表达样本之间的正交性，越小越好，表明样本之间的区分性越好。**

这个规则也可以用来做选择性集成，选择集成的分类器。

**【4】网络结构化简单性原则的损失函数**

目的是通过感知函数 构造网络

* 第一部分是**经验风险最小化**，即预测的误差要最小。只顾这一项容易导致过学习
* 第二部分是**表示复杂性感知，要稀疏化**，即简单样本要用简单的特征数量来表示，表示是神经网络抽取的的特征
* 第三部分是**结构复杂性感知**，要简单化

1. **基于特征选择原则的损失函数**

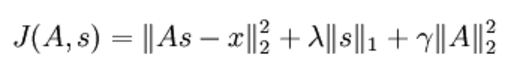
若选择出的特征能够很好的分类，那么特征抽取是有能力的，它提供了可供选择的特征。

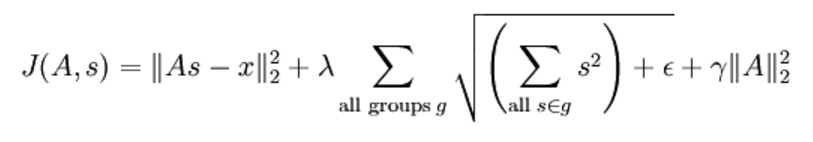
* 距离度量一般认为是差异性或者分离性的度量。对于两个特征f1和f2，如果特征f1引起的两类条件概率差异大于特征f2，则认为特征f1优于特征f2。
* 如果特征f1的信息增益大于特征f2的信息增益，则认为特征f1优于特征f2。
* 如果特征f1与类别C之间的相关性（Pearson correlation coefficient）大于特征f2与类别C之间的相关性，则认为特征f1优于特征f2。同样也可以计算得到属性与属性之间的相关度，属性与属性之间的相关性越低越好。
* 假定两个样本，若它们的特征值相同，且所属类别也相同，则认为它们是一致的：否则，则称它们不一致。
* 分类器错误率度量，它倾向于选择那些在分类器上表现较好的子集。

其中表示对选择出的特征的不恰当性度量，

**【5】稀疏性变形的损失函数：**

拓扑sparse coding主要是模仿人体大脑皮层中相邻的神经元对能提取出某一相近的特征





公式J(A,s)的第一项说明重构误差最小，第二项是要稀疏，第三项是基向量要最小。然后深度神经网络采用下列损失函数：

，

【6】多种损失函数的组合

将以上几种损失函数组合或部分损失函数组合，得到新的损失函数。