摘要：

本文主要研究了DNN(Deep neural networks )在多种语音识别的标准下胜于GMM的研究过程

I Introduction

1. GMMS之优点

make them suitable for modeling the probability distributions over vectors of input features that are associated with each state of an HMM.

huge amount of

research has gone into finding ways of constraining GMMs to

increase their evaluation speed and to optimize the tradeoff

between their flexibility and the amount of training data

required to avoid serious overfitting [6].

1. 2.GMMS之缺点

GMMS在非线性处对数据进行建模，是统计低效的

1. DNN值优点
2. DNN 初始作用
3. DNN超越CMM
4. DNN的两个阶段

II 训练深度神经网络

每个隐藏单元J，使用逻辑函数来将来自下一层的所有输入xj，

映射到上一层需要的标量yj

为了减小过度拟合，我们可以将较大的权重转换为他们的平方，或是在数据显示的正确性开始出错额时候停止训练

拥有多层隐藏层的深度神经网络实际上是很难优化的，从一个随机的开始点来进行梯度下降，并不是寻找一个好的权重级最好的方法（除非初始的权重就是经过精心选择的）

减小权重或是提前终止学习这两种方法可以减少过度拟合，但是只能通过减少建模力度

很大的训练集可以在保留建模力度的情况下减小过度拟合，但仅仅是建立在高昂的训练的计算代价之上

因此我们需要通过使用数据集中的信息来获取一个更好的方法，来建立多层非线性特征检测

III 生成预训练

这个想法是：一次学习一个层的特征，并利用这层这层特征状态的行为来训练下一层

预训练后，这些特征检测器形成的多级，就可以作为一个较好的起始点。在DNN中进行反向传播的时候，就可以微调这些来自预处理后的权重

RBM：限制玻尔兹曼机：其限制条件是在给定可见层或者隐层中的其中一层后，另一层的单元彼此独立

IV一个RBMs的有效学习程序

网络通过能量函数，为每个可见及不可见的向量的可能的部分分配一个概率

给予可见向量的概率是由所有可能的不可见向量之和而来

由于RBM中隐性单元之间连接性的缺失，获取一个无偏差的期望样本（<vi,hi>data）就显得较为容易

（取到<vi,hi>model）它能由这样的方法完成：从任意的可视单元的随机状态开始，用很长的时间执行交替的吉布斯采样

V 对真实数据进行建模

VI 使用堆栈RBMs创建深度信念网络

VII DNN与HMM的交互

VIII 预处理DNN的波形

IX 调整DBN-DNN，来优化交互信息

X 卷积DNNs下的语音分类及识别

XII DNNs和GMMs的区别总结

XIII 在很大的词汇量的语音数据集上比较DBN-DNNs与GMMs

XIV 在语音识别任务上的堆声音研究

XV 配电盘的语音识别任务

XVI 谷歌声音输入的语音识别

XVII yutube语音识别任务

XVIII 英文广播新闻语音识别任务

IXX DBN-DNN 模型在***LVCSR***任务上的主要结果

XX DNNs的识别加速

XXI DNN的交替预训练算法

XXII DNN的交替微调算法

XXIII 深度神经网络的其他应用

XXIV 使用DBN-DNNs向GMM-HMM系统提供输入特征

XXV 使用DNN对 基于检测的语音识别进行精度值估计

XXVI 总结以及未来的方向