**自编码器相关文献综述**

## 自编码器介绍[6]

### 传统自编码器（1986）(Auto-Encoders)

一般来说，传统自编码器主要包括编码阶段和解码阶段，且结构是对称的，即如果有多个隐层时，编码阶段的隐层数量与解码阶段相同。如图1所示：

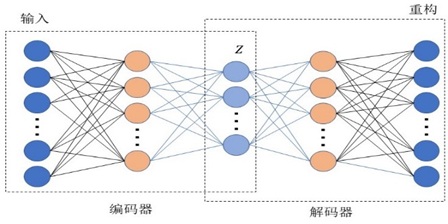


图1

传统自编码器的目的就是要在输出层重建输入数据。编码过程：,解码过程：. 当,称为绑定权重自编码器。

模型的损失函数为：

或者(交叉熵形式)：

我们发现自编码器没有涉及输入数据的标签信息。因此，传统自编码器被看作一种无监督学习模型。传统自编码器的隐层表达有三种不同的形式：压缩机构，稀疏结构和等堆结构。当输入层神经数量大于隐层时，称为压缩结构。反之，当小于时，称为稀疏结构。而如果两者相等时，称为等堆结构。

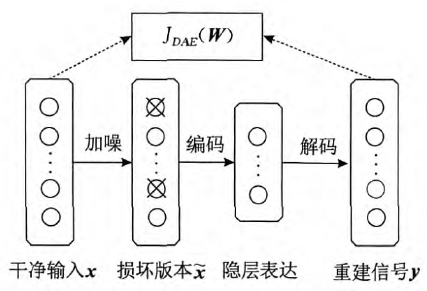
对于传统自编码器，还可以在损失函数增加一个权重衰减项来控制权重的减小程度，我们称之为传统正则自编码器。损失函数为

权重衰减项的增加能有效抑制静态噪声对目标和权重向量中不相关成分的影响，显著提升网络的泛化能力，可以有效避免过拟合现象的产生。

### 降噪自编码器（2008）(Denoising Auto-Encoders)

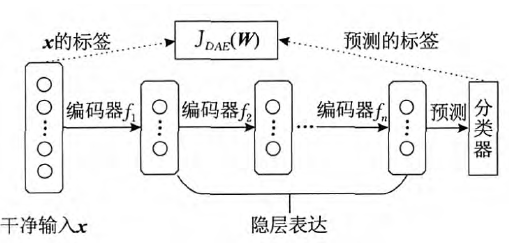
在自编码器中，我们真正关心的是隐层表达，那么到底什么是好的表达？首先它应该能够捕获输入信号的稳定结构，具有一定的鲁棒性，且对重建信号有用。因此，2008年Vincent提出了降噪自编码器。

降噪自编码器的研究目标是：隐层表达对被局部损坏的输入信号的鲁棒性。因此，降噪自编码器通过对干净输入信号人为加入一些噪声，使干净信号受到局部损坏，产生与它对应的一个损坏信号，然后将这个损坏信号送入传统自编码器，使其尽量重建一个与干净输入相同的输出。



降噪自编码器的目标就是最小化损失函数：

进一步将降噪自编码器进行堆叠，构造一个堆叠降噪自编码器并将其用于图像分类，结构如图：



边缘降噪自编码器的主要思想是在降噪自编码器的基础上，对噪声干扰进行边缘化处理，因此称为边缘降噪自编码器，针对线性和非线性问题，Chen分别提出了领域适应性边缘降噪自编码器和非线性表示边缘降噪自编码器。

### 边缘降噪自编码器（Marginalized Denoising Auto-Encoders）

**(1) 领域适应性边缘降噪自编码器（2012）**

领域适应性降噪自编码器是由一些单层降噪自编码器组合而成，因此，它的输入信号与降噪自编码器一样，是经过加噪处理的损坏信号。如果采用均方误差作为损失函数，那么其的目标为：

通常为了让模型具有普遍性，将实验过程重复多次，每次通过添加不同的噪声，产生m个不同的损坏信号作为输入，因此损失函数为：

领域适应性边缘降噪自编码器具有以下优点：具有比较强的特征学习能力；训练速度快；更少的中间参数，更快的模型选择以及基于逐层训练的凸性。但缺点在于智能用于线性表示。

### 稀疏自编码器（2011）(Sparse Auto-Encoders)

自编码器最初提出基于降维思想，但是当隐层节点比输入节点多时，自编码器就会失去自动学习样本特征的能力，此时就需要对隐层节点进行一定的约束。通过对隐层神经元的大部分输出进行抑制使网络达到一个稀疏的效果。

为了实现抑制效果，稀疏自编码器通过对隐层神经元输出的平均激活值进行约束，利用KL散度迫使其与一个给定的稀疏值相近，并将其作为惩罚项添加到损失函数中，因此，稀疏自编码器的损失函数可表示为：

代表所有训练样本在隐层神经元j上的平均激活值：,

如果通过稀疏表达可以完美重建输入，那么说明这些恶稀疏表达已经包含了输入的大部分主要特征，可以看作是对输入的一种简单表示，这样就在保证模型重建精度基础上，极大降低数据维度。

### k-稀疏自编码器（2013）(K-Sparse Auto-Encoders)

由上一节可以知道，稀疏自编码器可以使模型通过抑制隐层神经元达到一个很好的稀疏效果，但是无法指定哪些神经元处于激活状态，些神经元被抑制，使其无法获得一个准确的稀疏度．而k-稀疏自编码器针对这个问题，对稀疏自编码器进行了改进．两者之间最本质的区别在于ｋ-稀疏自编码器仅使用线性激活函数，同时在隐层只保留ｋ个最大激活值．仅使用线性激活函数是为了有效提升模型的训练速度，而通过指定ｋ的取值，可以避免使用ＫＬ散度，同时使模型获得一个准确的稀疏度．通过以上两点改进，ｋ－稀疏自编码器能很好的处理大数据量问题，这是稀疏自编码器无法做到的。

k-稀疏自编码器的基本思想为：（1）在神经网络前馈阶段，首先对输入进行线性变化，计算其激活值，然后利用排序算法或者Relu函数选取k个最大的激活值，k的设置可以看成一个规则化，防止使用过多的隐层单元去重建输入；（2）利用隐层的稀疏表示计算输出和误差，最后使用反向传播算法对网络参数进行优化。

### 收缩自编码器（2011）(Contractive Auto-Encoders)

衡量一个模型的效果如何可以通过两个标准来评判：（1）模型是否可以很好地重建输入信号；（2）模型对输入数据一定程度下的扰动是否有不变性。

对于第1点，大部分自编码器都能很好地完成，但是第2点很难达到很好效果。然而，对于分类任务而言，第2个标准更加重要。收缩自编码器的提出的主要目的是：抑制训练样本在所有方向的扰动，通过在传统自编码器的目标函数上增加一个惩罚项来达到局部空间收缩的效果。该惩罚项是关于输入的隐层表达的Jacobian矩阵的F范数，其目的是为了使特征空间在训练数据附近的映射达到收缩效果：

那么模型的损失函数为：

在功能上压缩自编码器和稀疏自编码器具有一定的相似性。在非线性情况下，收缩

自编码器的收缩性与稀疏自编码器的稀疏性非常类似，两者都鼓励稀疏表示，稀疏自编码器通过使其大部分隐层神经元受到抑制，隐层输出对应于激活函数的左饱和区域，而收缩自编码器通过将隐层神经元的输出推向它的饱和区域来达到收缩性。同样的，收缩自编码器的鲁棒性与降噪自编码器也是如出一辙，两者都对输入噪声具有鲁棒性，其主要区别在于作用对象不同，降噪自编码器是针对重建信号的鲁棒性，而收缩自编码器则是针对隐层表达的鲁棒性。

### 高阶收缩自编码器（2014）(Higher order Contractive Auto-Encoders)

将收缩自编码器对于输入微小变化的鲁棒性进一步提升，所以提出了高阶收缩自编码器。该模型对收缩自编码器的目标函数做了进一步改进，在其基础上又增加了一个二阶惩罚项，该惩罚项是关于输入的隐层表达的 Hessian矩阵的 F范数。

### 饱和自编码器（2013）(Saturating Auto-Encoders)

从一定角度上来说，饱和自编码器和使用非线性变换的收缩自编码器具有非常类似的地方，两者都鼓励隐层神经元的输出落在激活函数的饱和区域。不同之处在于，饱和自编码器是通过在传统自编码器的隐层神经元上引入一个正则项，使隐层神经元的激活函数至少包含一个零梯度区域，即饱和区域，而该正则项的目的就是尽量使激活值落在相应激活函数的饱和区域。

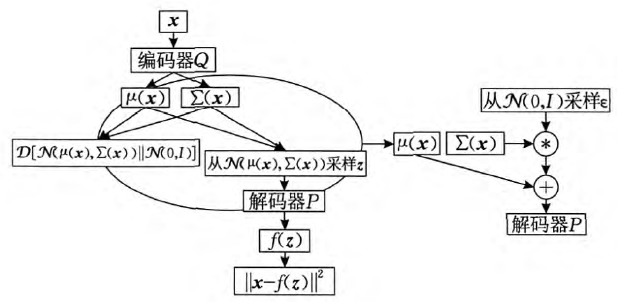
### 卷积自编码器（2011）(Convolutional Auto-Encoder)

卷积自编码器，顾名思义就是将传统自编码器的全连接替换为卷积网络。这样就可以处理二维图像或视频信号，来保留二维信号的空间信息。

**变分自编码器（2014）(Variational Auto-Encoder)**

在前述的自编码器中，都没定义隐层向量的分布，因此学习的隐层表示都是训练数据的压缩表示，具有不连续性和非完备性。因此从编码器的角度将，变分自编码器通过定义隐变量分布为高斯分布，并且在损失函数中用KL散度约束该分布近似于标准高斯分布。而产生连续紧致的分布。这样得到的隐向量就具有了一定含义。可以直接从隐空间中采样得到新的隐变量，得到一个原始数据集未出现过且有意义的数据。损失函数如下：

网络结构如下：

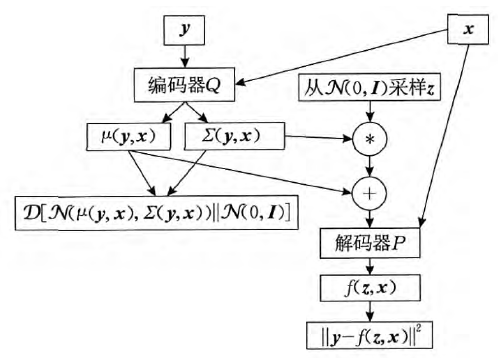


### 条件变分自编码器（2015）(Conditional Variational Auto-Encoders)

条件变分自编码器是指在原始输入(x)加入特定的标签或其他信息(y)，则在生成时，输出结果会产生与标签相关的数据。

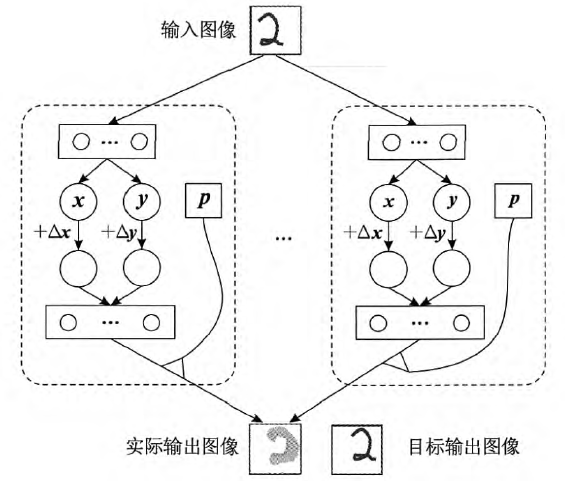
具体模型与损失函数和变分自编码器的差别不大。

如下图为模型结构：



### 变换自编码器（2011）（Transforming Auto-Encoders）

变化自编码器是由Hinton等人于2011年提出，他们认为神经网络应该使用一种被称为“胶囊”（capsule）的基础单元去执行一些非常复杂的内部计算，然后把计算机过封装在一个较小的向量里，该向量中包含了关于输入的大量信息。



其实胶囊与卷积神经网络中的局部单元非常相似，但是一个最重要的区别在于，卷积神经网络池化层通常只输出那个网络中最活跃的神经元的结果，而胶囊会将识别单元的所有输出封装到一个简单向量中并用于后续更高级别的活动．此外，由于识别单元计算出来的位置信息非常精确，因此胶囊可以更好地利用这些信息，甚至可以精确到图像的每一个

像素．

变换自编码器的最大优点在于，可以通过任何一种已知的方式去强制规定一个胶囊的输出，这些输出代表的是图像的任意性质．但是它也有一个非常严重的缺陷，那就是每个胶囊在同一时间只能代表可视化实体的一个实例．

### 区分自编码器（2015）(Discriminative Auto-Encoders)

2015年提出了一种新的自编码器——区分自编码器，它通过对隐层神经元增加一个Fishe区分准则，使隐层特征对几何结构变化具有可区分性和不敏感性．区分自编码器最关键的技术在于它对传统正则自编码器的损失函数所做出的改进，其损失函数具体可描述为:

和所表达的意义是的类内离散度和类间离散度。而区分自编码器的最终目标就是使类内离散度尽量小，同时类间离散度尽量大，很明显这对于分类任务非常有利。

### 大边缘自编码器（2017）（Large Margin Auto-Encoders）

受到区分自编码器最小化类内离散度和最大化类间离散度思想的启发，Liu等在2017年提出了大边缘自编码器，通过使属于不同类的样本在隐层空间具有更大的边缘距离来进一步提升自编码器的区分能力．

与区分自编码器的思想类似，大边缘自编码器也是对传统正则自编码器的损失函数进行了改进，通过对损失函数增加一个大边缘惩罚项，使带有不同类标签的样本在ｋ 最近邻中保持一个安全距离。因此，大边缘自编码器的损失函数可表示为：

### 张量自编码器（2014）(Tensor Auto-Encoder)

前面所介绍的自编码器模型都是基于向量空间的，因此无法学习大数据的特征，因为一个向量无法对大数据的高度非线性分布建模，尤其是异类数据。因此，2014 年 Zhang等人提出了张量自编码器的概念，同时通过堆叠多个张量自编码器去构建一个深度计算模型用于大数据的特征提取.

张量自编码器和传统自编码器非常类似，同样包括一个输入层，一个隐层和一个输出层．唯一的不同在于张量自编码器的每一层都用一个张量表示。

张量自编码器的编码过程可描述为：

解码过程可描述为：

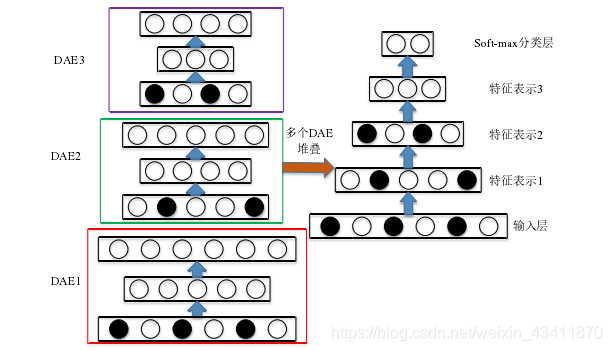
张量自编码器的编解码过程和传统自编码器基本完全一样。不同之处在于 ，是(Ｎ＋１)阶张量，和是N阶张量，而⊙为一个多点积过程，一个(Ｎ+１)阶和 Ｎ 阶张量的多点积操作结果为一个Ｎ 阶张量。

假 设 有 一 个 训 练 样 本 集 张量自编码器首先根据张量距离将每个训练样本从张量形式转换为向量形式。而为了从复杂数据中尽可能地获取未知分布，张量自编码器采用张量距离作为重建误差中的平均平方和误差项，因此张量自编码器的损失函数可表示为：

其中G是一个度量系数。

### 堆叠自编码器（stacked auto encoder，也称栈式自编码器或级联自编码器）

上述的自编码器都可以堆叠起来，形成堆叠自编码器（比如堆叠降噪自编码器，堆叠稀疏自编码器等），使得将原始数据以最小的信息损失降低到较小的维度。具体模型结构如下图：



上图的右边是一个分类模型，每一层的特征层首先由自编码器训练得到隐向量表示，这个过程叫预训练。第一个自编码器DAE1的输入为原始数据，训练得到编码向量，作为特征表示1；第二个自编码器DAE2的输入为DAE1的编码向量，训练得到特征表示2；DAE2的编码向量作为DAE3的输入，训练得到特征表示3. 然后连接一个softmax层，构建一个分类器。最后再进行分类任务的学习，调整模型参数及特征表示，这个过程叫做微调（fine-tuning）。

**【问题】堆叠自编码器比普通的分类网络优势在哪里？**

堆叠自编码器通过层层堆叠学习到一个较好的表示；但是没有证据说明在压缩到相同的维度，堆叠自编码器的效果一定比单个自编码器效果好；另外对于分类任务，在普通BP网络中，从原始数据提取底层特征，再到高层特征，在做分类任务。在训练过程中，反向传播算法将特征逐渐提取为与分类任务相关的特征，最终得到一个分类想过优异的模型；而堆叠自编码器先是通过无监督训练提取和原始数据相比损失尽可能小的特征。但是没有证据证明这个特征一定有利于分类，或者说比BP网络提取的特征更好。

如果把分类网络看成一个端到端的函数映射。这个函数将是一个复杂的多峰和多谷的函数，因此，给定不同的初始值，则函数会收敛到该初始值周围的局部最小点。考虑到堆叠自编码器的分类模型，它将训练自编码器的参数，作为后续训练分类器的初始参数；而BP网络的参数初始值是随机的，可以定义为高斯分布等其他。这样网络就是从不同的参数初始值训练分类模型，而且很可能会收敛到不同的局部最优值。没有证据可以说明前者的局部最优解更接近BP网络的局部最优解。

从训练难度，训练耗时的角度将，也没有证据证明前者比后者更省时，更省内存，训练更加简单。

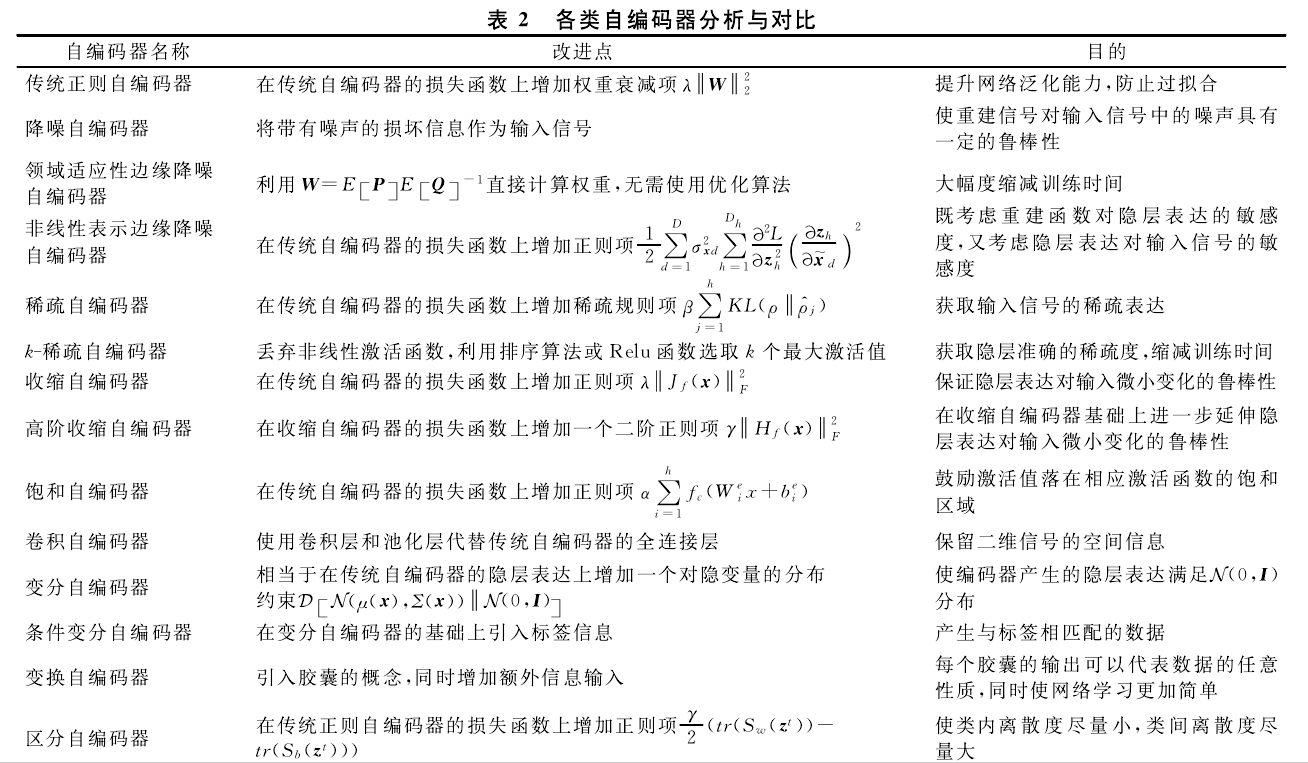
目前没有理论或具体的实验验证，大多数文章都是采用堆叠自编码器的方法去做分类，然后用实验结果去说明他们的方法更好，并没有更好的说明原因。

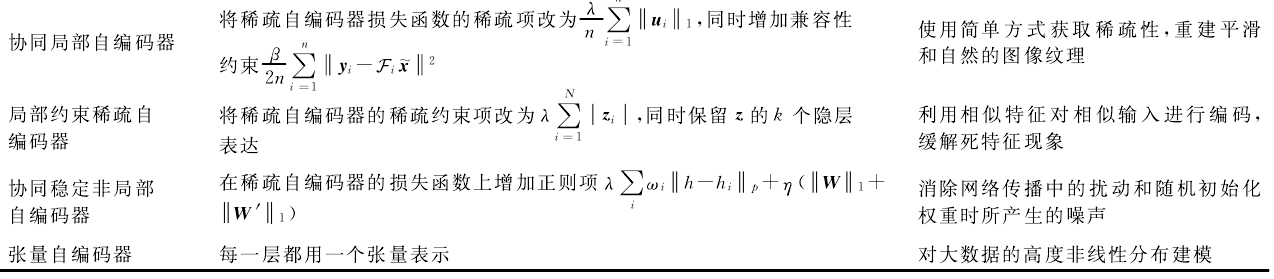
**通过查询，首先有一个基于实验的认知，在数据的标签较少的情况下，SAE模型可以在训练分类器模型之前就找到一个较好的网络初始值，这样后期通过较少的带标签样本以更少的训练迭代步数使网络达到局部最优。但是当标签数量足够时，无法说明哪个更好。我们可以想到的是，对于一个分类任务，一般使用的是有监督模型，数据是带有标签的。当数据量少时用SAE有利于得到更有的分类模型。**

**但是如果我们对于基本的分类模型多几次迭代，调整好超参数，是不是会得到差不多的效果？这也是未知的。**

## 自编码器对比与分析

通过对大量自编码器的研究发现，大部分自编码器的创新点都在于对基础的损失函数增加各种正则项约束，使模型达到不同的效果，比如稀疏性、鲁棒性、收缩性、兼容性等．同时，大部分自编码器进行约束的对象都是隐层表达，因为对于大部分研究者而言，自编码器最重要的性质还是在于其能自动学习有效特征，**因此大部分自编码器所做的改进都是针对隐层表达．而其中研究最多的是隐层表达对输入噪声的鲁棒性，包括边缘降噪自编码器、收缩自编码器、饱和自编码器、协同稳定非局部自编码器．除此之外，还有一些其他的创新方向，比如领域自适应边缘降噪自编码器和ｋ－稀疏自编码器，在满足各种约束的前提下，大幅度减少了训练时间．卷积自编码器是为了有效利用二维数据的空间相关信息，变换自编码器是为了使网络更加简单，信息高度集中，而张量自编码器则是为了能够实现对高维数据的建模．**下表对前面介绍的自编码器进行了一个归纳总结，特别指出了各类自编码器的改进点以及所要达到的目的。





## 应用场景

**应用1：数据分类**

根据对大量文献的研究，利用自编码器进行分类的方法只有一个类似的通用模型——堆叠自编码器，依据数据和任务需要，堆叠各种自编码器的编码部分用于特征提取，最后采用有监督学习对网络参数微调。

基于以上模型的研究方法最大的区别在于用的自编码器，损失函数以及送入自编码器网络的数据不同。典型的任务有图像分类，3D模型分类，无线信号分类，AMR音频分类等

**应用2：异常检测**

通过对文献的研究，我们发现基于自编码器的异常检测也有一个类似的框架，即先利用自编码器去学习正常数据的特征，建立正常的分布，然后在测试阶段，重建误差比较高的数据被认为是异常数据。因此正常数据的特征学习成为了其中非常关键的步骤。

**应用3：数据生成**

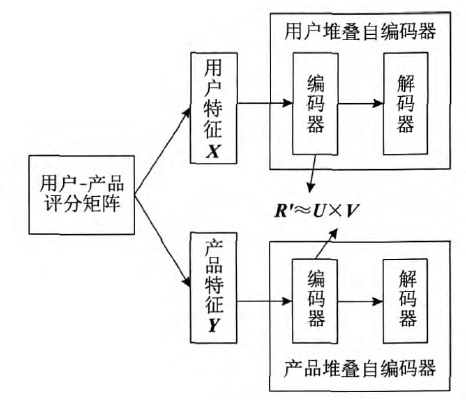
数据生成模型是指通过模型识别数据，进而产生出类似原始数据但有原始数据不同的，有意义的数据。比如理解图片，音频，视频，文本内容等。现在很火的应用有：人脸生成（指定风格，风格迁移），语音和音乐合成，提升图片画质，提升音频清晰度，机器翻译，文本生成，换脸，换表情，视频实时渲染转换，增强学习等。

在深度学习，目前比较流行的数据生成基础模型有变分编码器(VAE) 和生成式对抗网络（GAN）。其中基于VAE的解耦模型由于其隐变量的分布是可推断的，因此基于VAE的解耦模型的可解释性更强。目前利用VAE及其改进模型生成的数据类型有图像，音乐和自然语言等。

**应用4：产品推荐**

协同过滤技术的提出使推荐系统真正受到了大家的关注。协同过滤的主要理念是相似用户之间具有相似的偏好，主要思想是首先对用户-产品评分矩阵Ｒ进行矩阵因式分解，来学习用户和产品的特征表达，然后基于相似性计算结果进行产品推荐，其中矩阵因式分解是协同过滤方法中的核心步骤。此外还有一些改进的方法，通常是引入一些额外的支持信息，比如用户与产品属性之间的社会关系，作为规则化去丰富特征表达的先验知识。但是由于评分矩阵中的评分数据通常远远小于产品数量，因此会出现数据稀疏的问题，而一些新出现的用户或产品，由于缺乏评分数据，又存在冷启动问题。

因此为了克服这些困难，有些文献提出采用自编码器去自动学习用户和产品的特征表达。据分析，这些方法存在一个共同点，那就是都采用两个独立的堆叠自编码器网络去分别学习用户和产品的特征Ｕ 和Ｖ，然后通过Ｒ′≈Ｕ×Ｖ 获得一个新的评分矩阵用于后续工作，其基本框架如下图所示，而不同点在于其中采用的堆叠自编码器不同。



**其他应用**

除了以上几类研究比较多的应用领域，自编码器还被应用于图像重建、图像配准、人脸对齐、数据增强、目标检测、目标追踪、血管分割、数字水印、股票市场预测等。

通过上述描述，我们发现各种改进方法的提出可归纳为以下几个方向：(1) 将比较优秀的传统特征提取方法与各种自编码器相结合，共同用于特征的提取；(2) 结合各研究领域的先验知识，对自编码器的损失函数进行修改，一般都是通过增加能反映该先验 知识的惩罚项来达成；(3) 将各种自编码器与一些比较好的方法相结合，比如 ELM、k最近邻、GAN等，以提升模型的整体效果；(4) 将两个自编码器进行组合，并通过一些方法建立两个自编码器之间的联系。

## 关于自编码器的最新论文

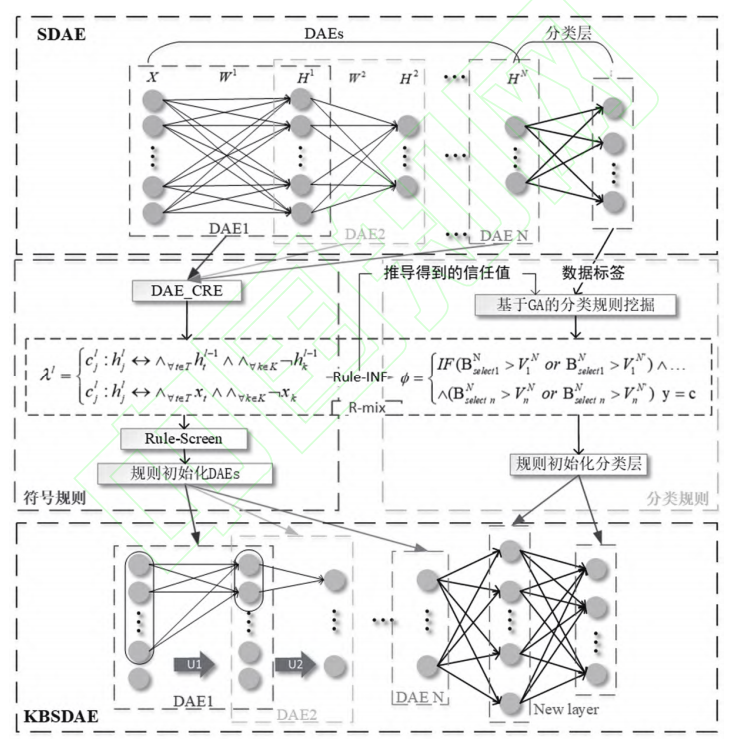
**1. 知识堆叠降噪自编码器[1]**

创新点：将知识，逻辑规则融入堆叠降噪自编码器中，使之具有可解释性和可视化特性。

【摘要】

知识堆叠降噪自编码器（Knowledge-based stacked denoising Autoencoder, KBSDAE）。**尝试以一种逻辑语言的方式有效解释网络结构及内在运作机理，同时确保逻辑规则可以进行深度推导。进一步通 过插入提取的规则到深度网络, 使 KBSDAE 不仅能自适应地构建深度网络模型并具有可解释和可视化特性, 而且有效地提高 了模式识别性能.** 大量的试验结果表明, 提取的规则不仅能够有效地表示深度网络, 还能够初始化网络结构以提高 KBSDAE 的特征学习性能, 模型可解释性与可视化, 可应用性更强.

模型结构如图：



模型特点：

**面对深度神经网络的”黑箱问题”, 本文提出了一套全新的知识表达规则系统, 尝试解释并强化深度神经网络.** 该系统可以**对 SDAE 网络进行简单表示并从 SDAE 中抽取和插入知识**. 通过这套系统,可以理解到网络内部的知识并建立性能更加强大的KBSDAE. 规则系统创新性的将符号类型和数值类型的规则进行有机结合, 使得这种混合规则具有较高的推导性能和可理解性, 在网络规模愈加复杂的当下这种规则形式不失为一条具有研究价值的路径.试验证明,混合规则系统可以有效表示网络并提取网络知识, 具有高推导精度和稳定性. 利用规则初始

化后的 KBSDAE 相较于传统 SDAE 具有更快的收敛速度, 更高的预测精度和数据灵敏度. 下**一步可以将规则与深度网络可视化相结合以提升对网络的解释能力. 可以尝试解释更复杂和庞大的网络, 如卷积神经网络.**

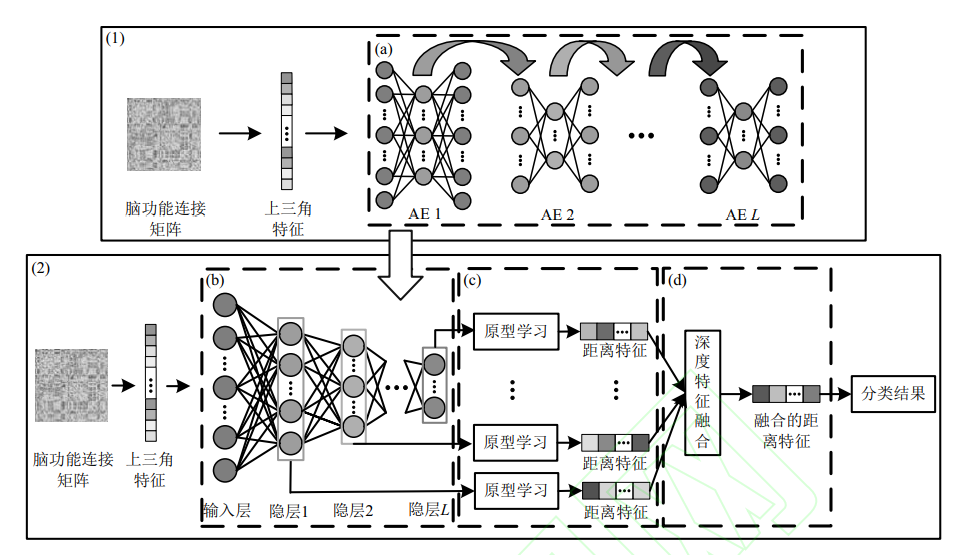
**2.基于原型学习与深度特征融合的脑功能连接分类方法研究[2]**

应用场景：脑功能连接分类方法

创新点：基于原型学习与深度特征融合，并结合栈式自编码器（即堆叠自编码器）提取特征

【摘要】

近年来, 基于深度学习的**脑功能连接分类方法**已成为一个研究热点. 为了进一步提高脑功能连接的分类准确率, 获得与疾病相关的鉴别性特征, 本文提出了一种基于原型学习与深度特征融合的脑功能连接分类方法. **该方法首先使用栈式自编码器从脑功能连接中提取从低层次到高层次的深度特征**; **然后利用原型学习在自编码器的各隐层中提取表示样本类别信息的距离特征; 最后采用深度特征融合策略将这些距离特征融合, 并将该融合特征用于脑功能连接的类别标签预测.** 在ABIDE数据集上的实验结果表明, 与其他同类方法相比, 该方法不仅具有较高的分类准确率, 而且能够更加准确地定位与疾病相关的脑区.



新的知识点：原型学习与深度特征融合

**3.基于 SAE 的睡眠呼吸中止综合症识别及预测算法[3]**

应用场景：睡眠呼吸中止综合症

任务：识别与预测

创新点：将SAE应用到新的场景OSA识别中

【摘要】

睡眠呼吸暂停综合症(OSA)是一种常见的睡眠中呼吸停止的睡眠障碍，在各年龄段人群中均拥有较高的患病比例。然而目前 OSA 的主要诊断仍依靠复杂而昂贵的多导睡眠监测(PSG)，严重制约了 OSA 的筛查效率。**针对这个问题，本文提出了一种基于栈式自编码网络(SAE)的鼾声 OSA 症状识别算法。该算法首先利用心电、鼾声数据进行 OSA 症状的特征抽出，两种特征在 SAE 的输入层或隐层结合后共同训练OSA 诊断模型。**该模型利用了深度神经网络发掘复杂特征及多模态特征相关性的能力，在自编码网络的隐层训练心电、鼾声特征的相关性，并利用这种相关性提高模型的识别性能。OSA 症状识别及发病预测实验中，采

用心电、鼾声同时输入的情况下，比传统算法获得了5%-7%的识别精度提升；在采用鼾声输入时，也获得了96.5%的识别精度。

**4.基于深度学习的作曲家分类问题[4]**

应用场景：作曲家分类

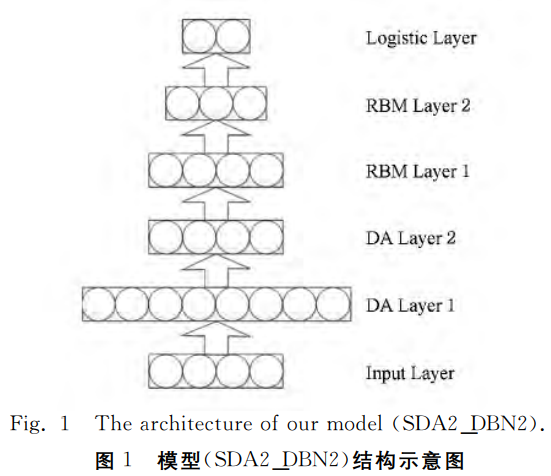
任务：分类

创新点：结合深度信念网络与堆叠降噪自编码器；应用到新场景作曲家分类

**【摘要】**

在音乐信息检索领域，作曲家分类是一个十分重要的问题，这一问题的目标是通过音频数据来识别相应的作曲家信息．传统的分类算法都是通过提取复杂的特征来进行分类的，而深层神经网络在特征学习上具有比较强的能力，因此提出用深层神经网络来解决这一问题．为了结合不同深层神经网络模型的优点，设计了一种混合模型，该模型基于深度置信网络和级联去噪自编码器，可以较好地解决作曲家分类问题．实验表明，该模型取得了７６．２６％的正确率，这一结果比单纯用某一种模型搭建的深层神经网络以及支持向量机要好．和图像数据类似，人脑在提取音乐特征也是分层的，每一层对信号的处理不一样，因此混合模型在解决作曲家分类问题上具有一定的优势

模型结构如图：



去噪自编码器擅长对数据进 行扩充，而受限玻尔兹曼机擅长对数据进行投影.

**5. 栈式自编码器特征表达能力研究[5[**

任务：研究栈式自编码器的表达能力

创新点：隐节点个数，隐层数对SAE的影响。创新点少

**【摘要】**

以深度学习中的自编码器为基础，堆积稀疏自编码形成深度网络，以贪心算法 逐层训练实现特征层的层抽象映射，监督训练 Softmax 分类器，然后使用 BP（反向传播） 算法优化权值，构建 SAE（栈式自编码器），并重点对其特征表达能力进行探究。 第一，以 单层自编码器作为特征表达的基础模块，探究自编码器中隐藏层节点数对特征表达能力 影响；第二，重点探究对于多层自编码器的理解，堆积自编码器是否是一个好的获取理想 特征表达的途径， 主要评估多层自编码器对于特征表达的准确性和稳定性影响。 **基于 MNIST（美国国家标准与技术研究所数据库）数据集的实验与 PCA（主成分分析）、LLE （局部线性嵌入）、BP 算法特征表达能力进行对比分析， 验证栈式自编码器特征表达能力 的有效性。**

## 引用

[1]刘国梁,余建波.知识堆叠降噪自编码器[J/OL].自动化学报:1-13[2021-06-17].https://doi.org/10.16383/j.aas.c190375.

[2] 梁玉泽,冀俊忠.基于原型学习与深度特征融合的脑功能连接分类方法研究[J/OL].自动化学报:1-12[2021-06-17].https://doi.org/10.16383/j.aas.c190747.

[3] 党鑫,柯登峰,韦然,张振.基于SAE的睡眠呼吸中止综合症识别及预测算法[J/OL].自动化学报:1-9[2021-06-17].http://kns.cnki.net/kcms/detail/11.2109.tp.20200430.1630.002.html.

[4]胡振,傅昆,张长水.基于深度学习的作曲家分类问题[J].计算机研究与发展,2014,51(09):1945-1954.

[5] 朱成.栈式自编码器特征表达能力研究[J].电信快报,2019(03):28-33.

[6]袁非牛,章琳,史劲亭,夏雪,李钢.自编码神经网络理论及应用综述[J].计算机学报,2019,42(01):203-230.