自编码器：一种数据压缩算法

性质：

数据相关的（人脸训练出来的不能压缩树木图片）

有损的

自动学习 压缩与解压缩通过神经网络实现。

自监督算法 不是一个无监督算法 输入与输出做损失函数的计算

条件：

编码器，解码器（一般是参数化方程）

损失函数——使用SGD最小化损失函数

应用：

数据去噪

进行可视化降维

过程：

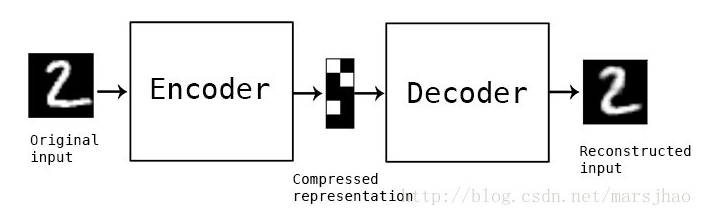
从输入到输出的不完全复制

编码器：h=f(x)

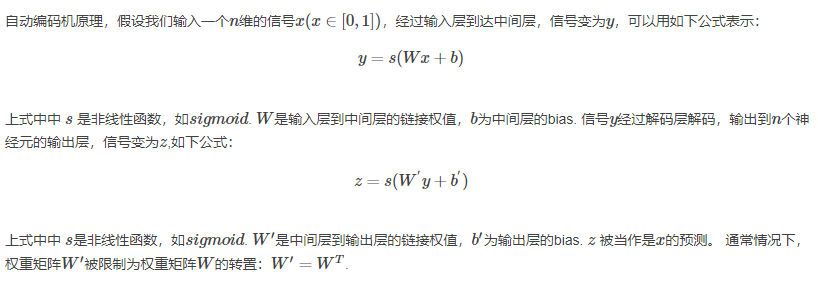
解码器：r = g(h)

增加约束条件，使得输出结果只能与输入近似，往往是考虑哪些部分被最优先赋值（主要特征）

因此数据的特征是编码器最感兴趣的地方。



原理：



损失函数：

可以用平方损失函数



如果输入的数据为位向量或者是位概率向量，可以用交叉熵（对数损失函数）



欠完备自编码器：

约束条件：要求h的维度比输入的x小

其他类型编码器详细见https://blog.csdn.net/marsjhao/article/details/73480859