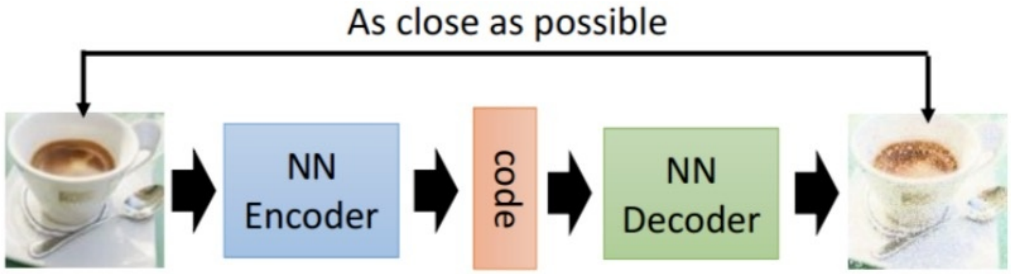
自动编码器

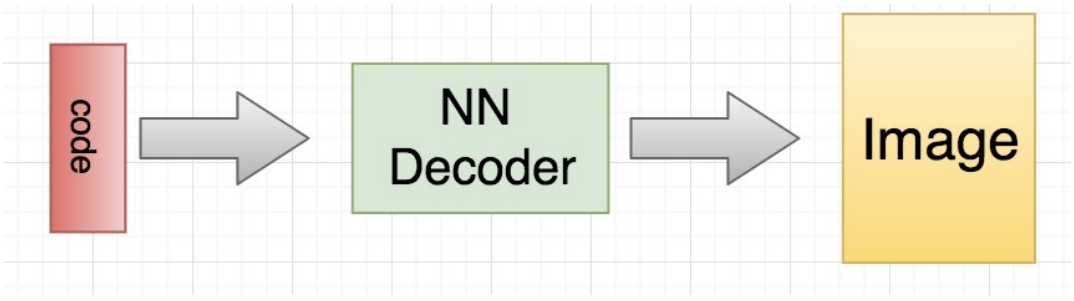
自动编码器最开始是作为一种数据压缩方法，同时还可以在卷积网络中进行逐层预训练，但是随后更多结构复杂的网络，比如 resnet 的出现使得我们能够训练任意深度的网络，自动编码器就不再使用在这个方面，下面我们讲一讲自动编码器的一个新的应用，这是随着生成对抗模型而出现的，就是使用自动编码器生成数据。

自动编码器的一般结构如下



由上面的图片，我们能够看到，第一部分是编码器(encoder)，第二部分是解码器(decoder)，编码器和解码器都可以是任意的模型，通常我们可以使用神经网络作为我们的编码器和解码器，输入的数据经过神经网络降维到一个编码，然后又通过另外一个神经网络解码得到一个与原始数据一模一样的生成数据，通过比较原始数据和生成数据，希望他们尽可能接近，所以最小化他们之间的差异来训练网络中编码器和解码器的参数。

当训练完成之后，我们如何生成数据呢？非常简单，我们只需要拿出解码器的部分，然后随机传入code，就可以通过解码器生成各种各样的数据



变分自动编码器

变分编码器是自动编码器的升级版本，其结构跟自动编码器是类似的，也由编码器和解码器构成。

回忆一下，自动编码器有个问题，就是并不能任意生成图片，因为我们没有办法自己去构造隐藏向量，需要通过一张图片输入编码我们才知道得到的隐含向量是什么，这时我们就可以通过变分自动编码器来解决这个问题。

其实原理特别简单，只需要在编码过程给它增加一些限制，迫使其生成的隐含向量能够粗略的遵循一个标准正态分布，这就是其与一般的自动编码器最大的不同。

这样我们生成一张新图片就很简单了，我们只需要给它一个标准正态分布的随机隐含向量，这样通过解码器就能够生成我们想要的图片，而不需要给它一张原始图片先编码。

一般来讲，我们通过 encoder 得到的隐含向量并不是一个标准的正态分布，为了衡量两种分布的相似程度，我们使用 KL divergence，利用其来表示隐含向量与标准正态分布之间差异的 loss，另外一个 loss 仍然使用生成图片与原图片的均方误差来表示。

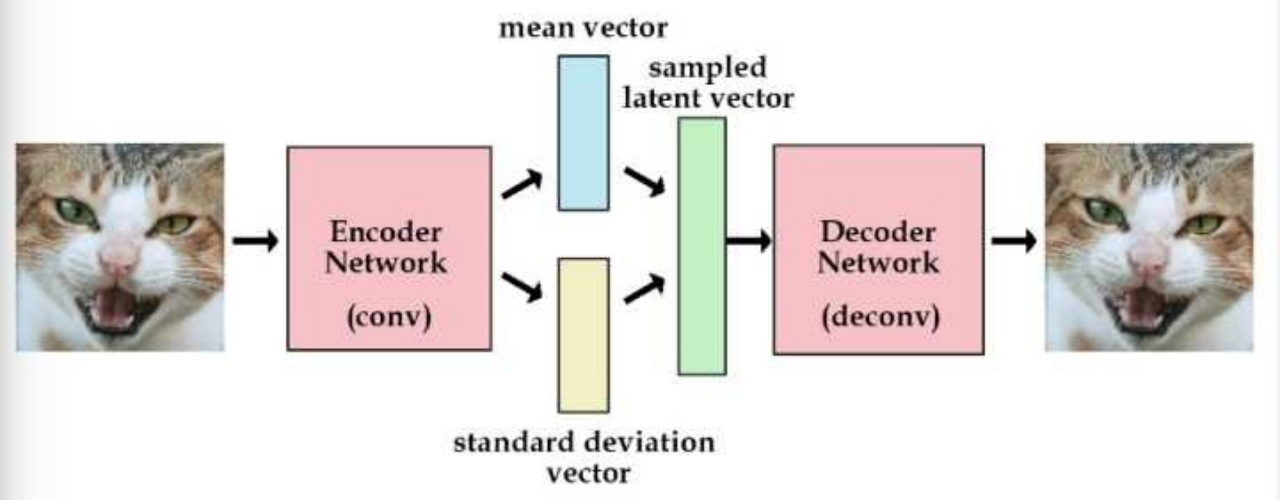
KL divergence 的公式如下

$$DKL(P||Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx \tag{1}$$

重参数

为了避免计算 KL divergence 中的积分，我们使用重参数的技巧，不是每次产生一个隐含向量，而是生成两个向量，一个表示均值，一个表示标准差，这里我们默认编码之后的隐含向量服从一个正态分布的之后，就可以用一个标准正态分布先乘上标准差再加上均值来合成这个正态分布，最后 loss 就是希望这个生成的正态分布能够符合一个标准正态分布，也就是希望均值为 0，方差为 1

所以标准的变分自动编码器如下



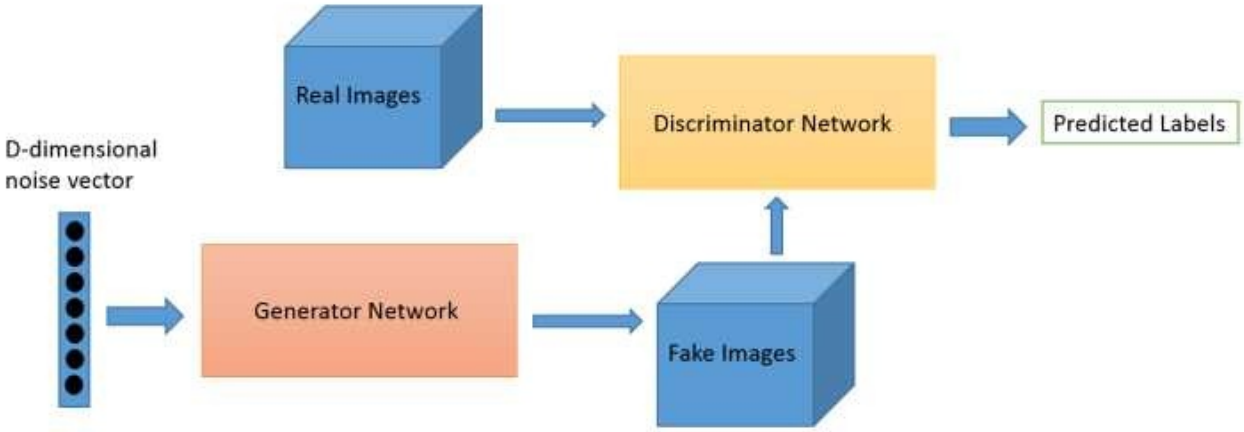
前面我们讲了自动编码器和变分自动编码器，不管是哪一个，都是通过计算生成图像和输入图像在每个像素点的误差来生成 loss，这一点是特别不好的，因为不同的像素点可能造成不同的视觉结果，但是可能他们的 loss 是相同的，所以通过单个像素点来得到 loss 是不准确的，这个时候我们需要一种全新的 loss 定义方式，就是通过对抗进行学习。

GANs

这种训练方式定义了一种全新的网络结构，就是生成对抗网络，也就是 GANs。这一部分，我们会形象地介绍生成对抗网络，以及用代码进行实现，而在书中会更加详细地介绍 GANs 的数学推导。

根据这个名字就可以知道这个网络是由两部分组成的，第一部分是生成，第二部分是对抗。简单来说，就是有一个生成网络和一个判别网络，通过训练让两个网络相互竞争，生成网络来生成假的数据，对抗网络通过判别器去判别真伪，最后希望生成器生成的数据能够以假乱真。

可以用这个图来简单的看一看这两个过程



Discriminator Network

首先我们来讲一下对抗过程，因为这个过程更加简单。

对抗过程简单来说就是一个判断真假的判别器，相当于一个二分类问题，我们输入一张真的图片希望判别器输出的结果是1，输入一张假的图片希望判别器输出的结果是0。这其实已经和原图片的 label 没有关系了，不管原图片到底是一个多少类别的图片，他们都统一称为真的图片，label 是 1 表示真实的；而生成的假的图片的 label 是 0 表示假的。

我们训练的过程就是希望这个判别器能够正确的判出真的图片和假的图片，这其实就是一个简单的二分类问题，对于这个问题可以用我们前面讲过的很多方法去处理，比如 logistic 回归，深层网络，卷积神经网络，循环神经网络都可以。

Generator Network

接着我们看看生成网络如何生成一张假的图片。首先给出一个简单的高维的正态分布的噪声向量，如上图所示的 D-dimensional noise vector，这个时候我们可以通过仿射变换，也就是 xw+b 将其映射到一个更高的维度，然后将他重新排列成一个矩形，这样看着更像一张图片，接着进行一些卷积、转置卷积、池化、激活函数等进行处理，最后得到了一个与我们输入图片大小一模一样的噪音矩阵，这就是我们所说的假的图片。

这个时候我们如何去训练这个生成器呢？这就需要通过对抗学习，增大判别器判别这个结果为真的概率，通过这个步骤不断调整生成器的参数，希望生成的图片越来越像真的，而在这一步中我们不会更新判别器的参数，因为如果判别器不断被优化，可能生成器无论生成什么样的图片都无法骗过判别器。

生成器的效果可以看看下面的图示


100
noise
Generator
Fake
images
Label = 1
Discriminator
Loss