

# 人工智能之深度学习

## 自编码器(AutoEncoder)

主讲人：刘老师(GerryLiu)

## 课程要求

- 课上课下 “九字” 真言
  - 认真听, **善摘录, 勤思考**
  - **多温故, 乐实践**, 再发散
- 四不原则
  - **不懒散惰性, 不迟到早退**
  - **不请假旷课, 不拖延作业**
- 一点注意事项
  - 违反 “四不原则” , 不推荐就业

## 课程内容

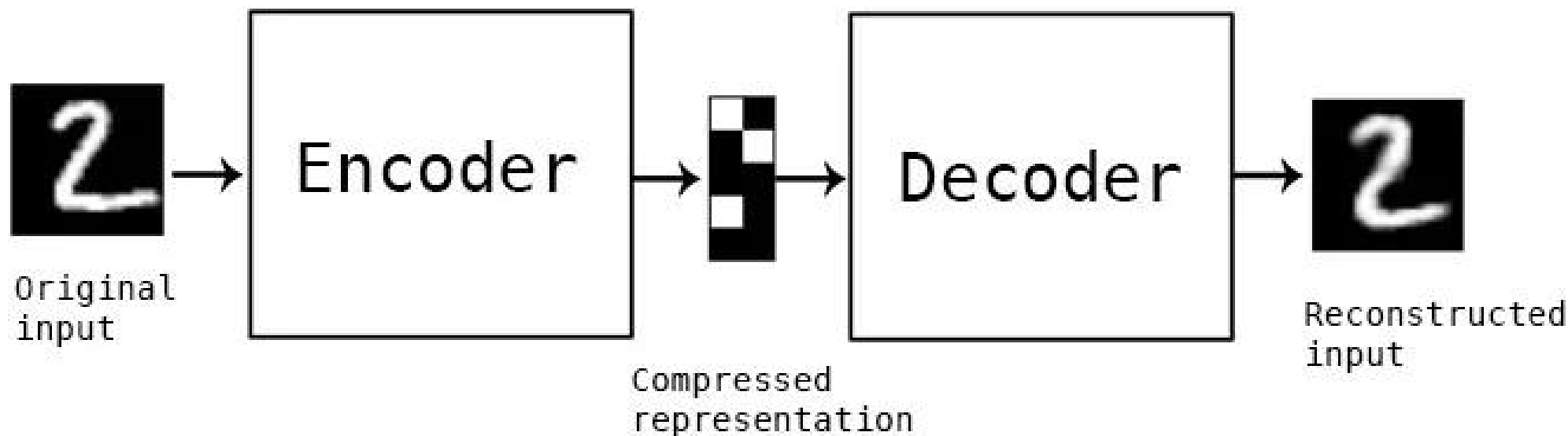
- 什么是自编码器(AutoEncoder, AE)
- 变分自编码器(Variational AutoEncoder, VAE)
- 案例代码(手写数字生成)

## 自编码器

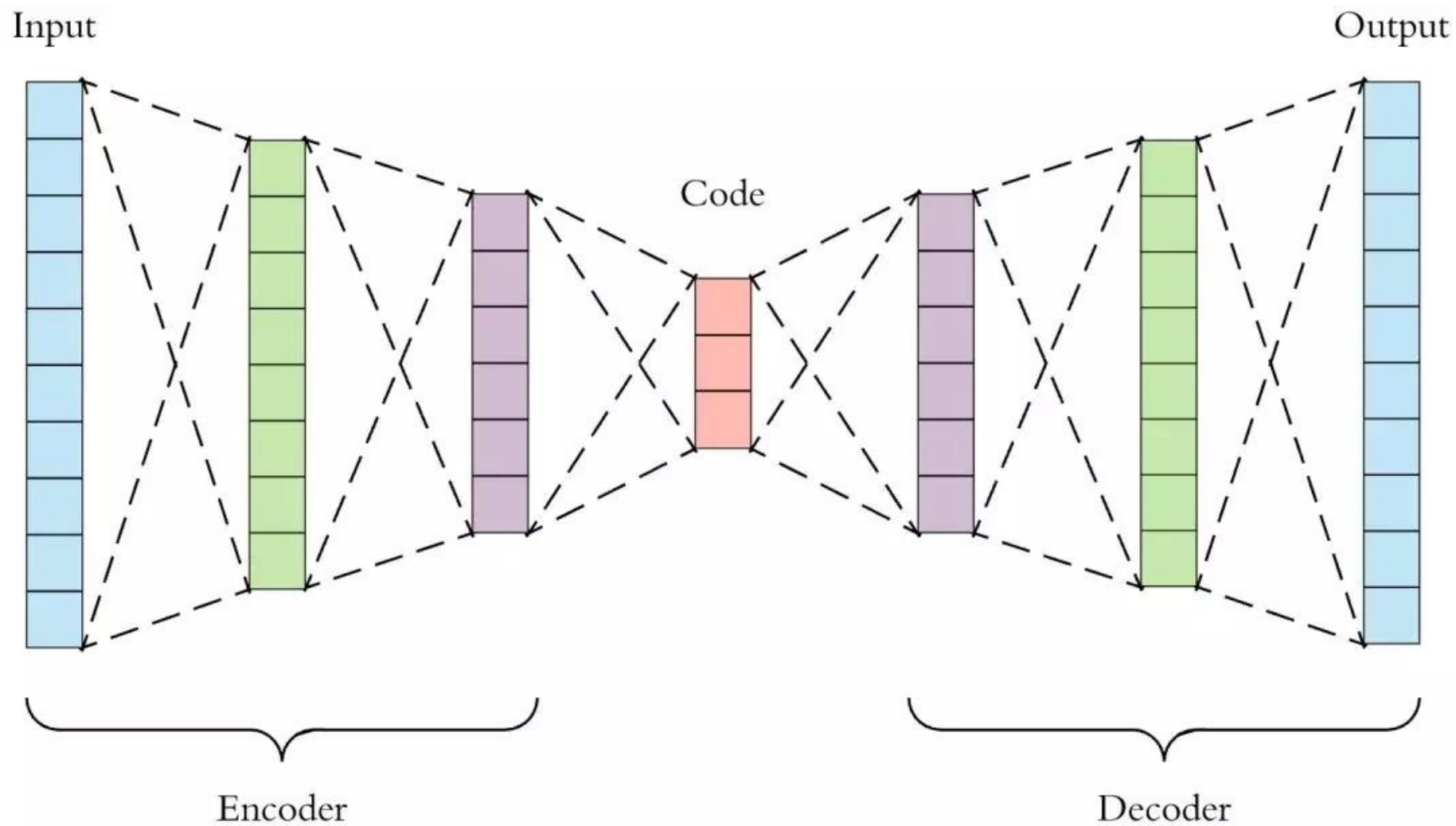
- 自编码器作为一种深度学习领域无监督的算法，本质上是一种数据压缩算法，和生成对抗网络一样，属于生成算法的一种。
- 自编码器(AutoEncoder, AE)就是一种利用反向传播使得输出值等于输入值的神经网络，它将输入压缩成潜在特征/高阶特征，然后将这种表征重构输出。主要包含以下三个特征：
  - 数据相关性。
  - 数据有损性。
  - 自动学习性。

## 自编码器

- 构建一个自编码器主要包括两部分：**编码器(Encoder)和解码器(Decoder)**。编码器将输入压缩为潜在空间特征，解码器将潜在空间特征重构输出。
- **自编码的核心价值是在于提取潜在的高阶空间特征信息。**主要应用是两个方向：**数据去噪**以及进行**可视化降维**。



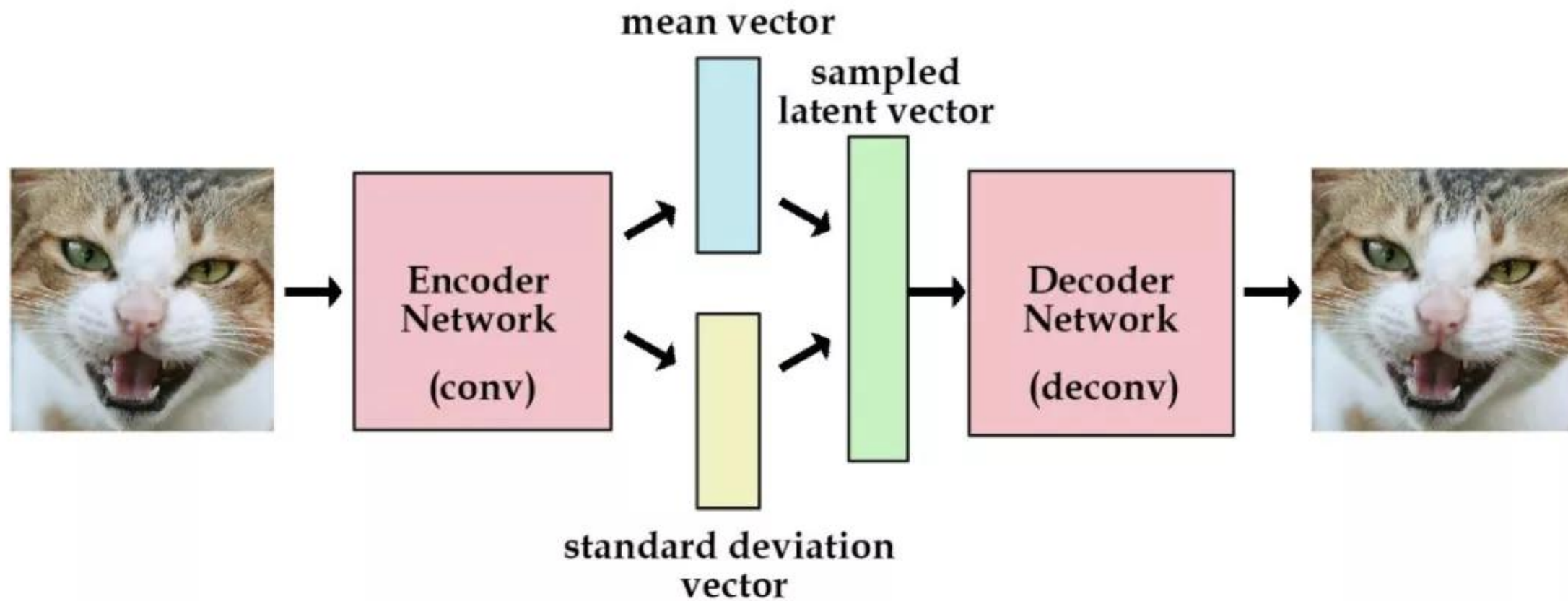
# 自编码器



## 变分自编码器

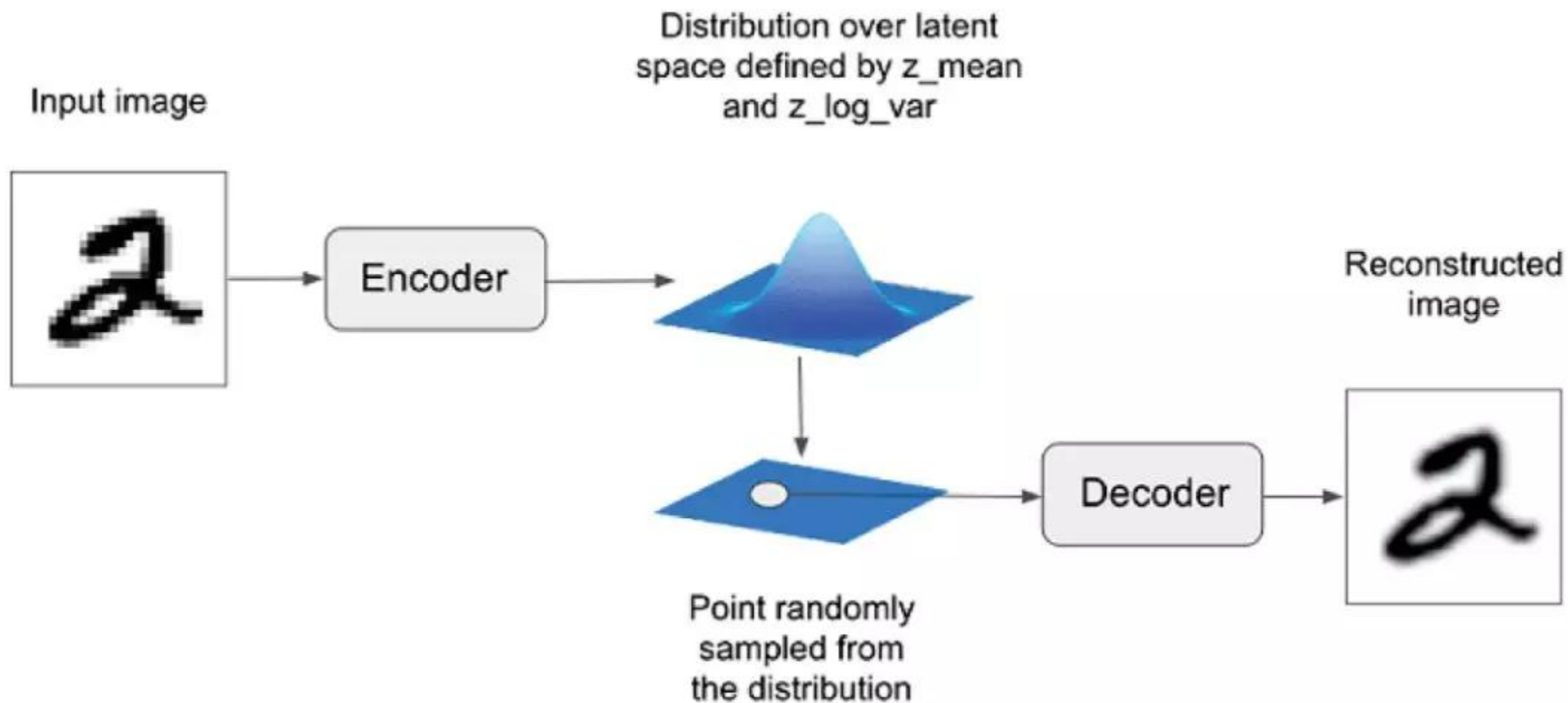
- 变分自编码器(Variational AutoEncoder, VAE)是常见的生成模型的一种，特别适合利用概念向量进行图像生成和编辑的任务。
- 经典的自编码器由于本身是一种有损的数据压缩算法，在进行图像重构的时候不会得到效果特别好的结果，而VAE则不是将输入图像压缩成潜在表征，而是将输入图像数据转换为最常见的两个统计分布参数：均值 $\mu$ 和标准差 $\sigma$ ，然后解码器使用这两个参数从分布中进行随机采样得到隐变量，对隐变量进行解码重构就可以得到输出值。

# 变分自编码器





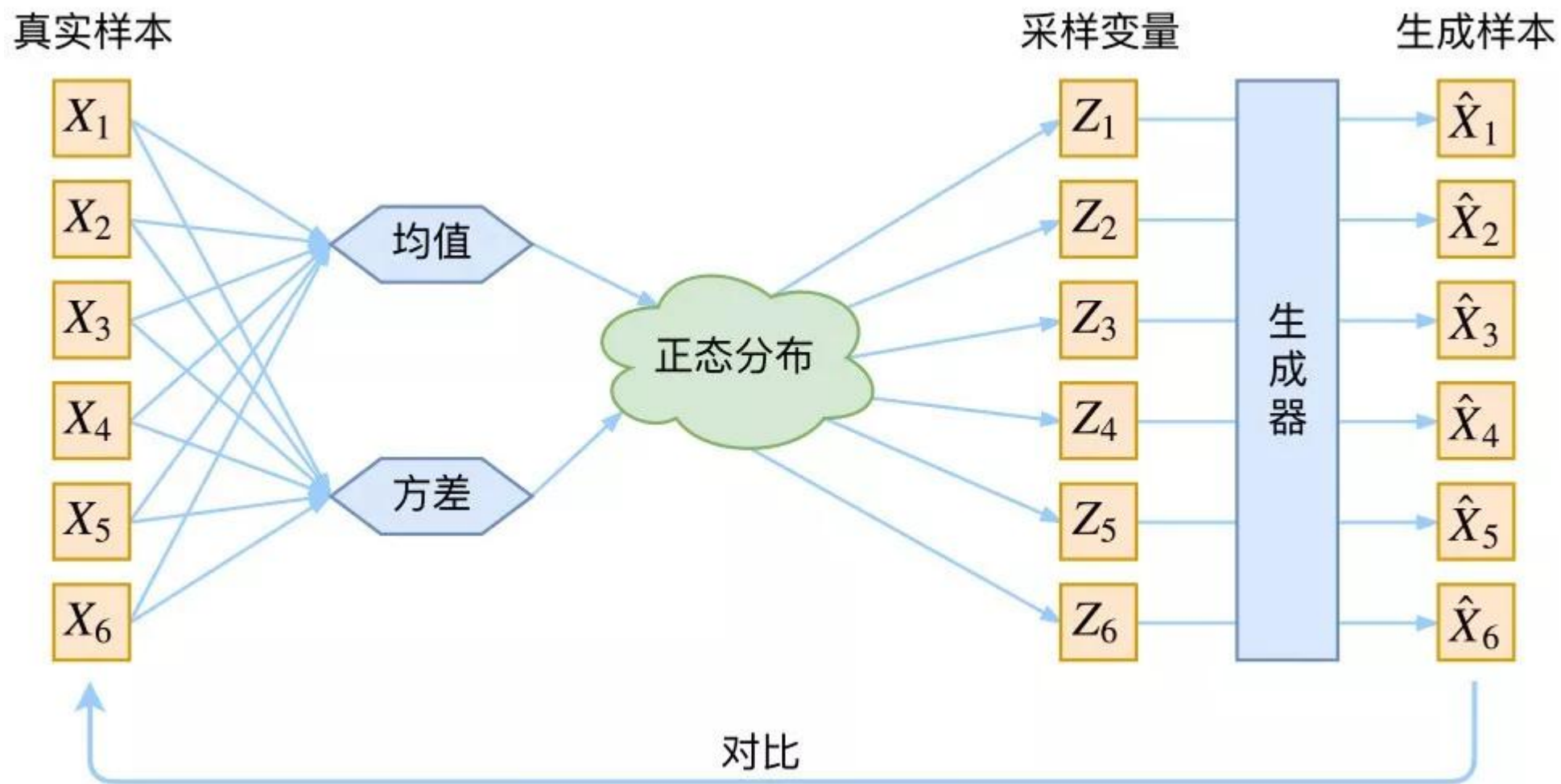
# 变分自编码器



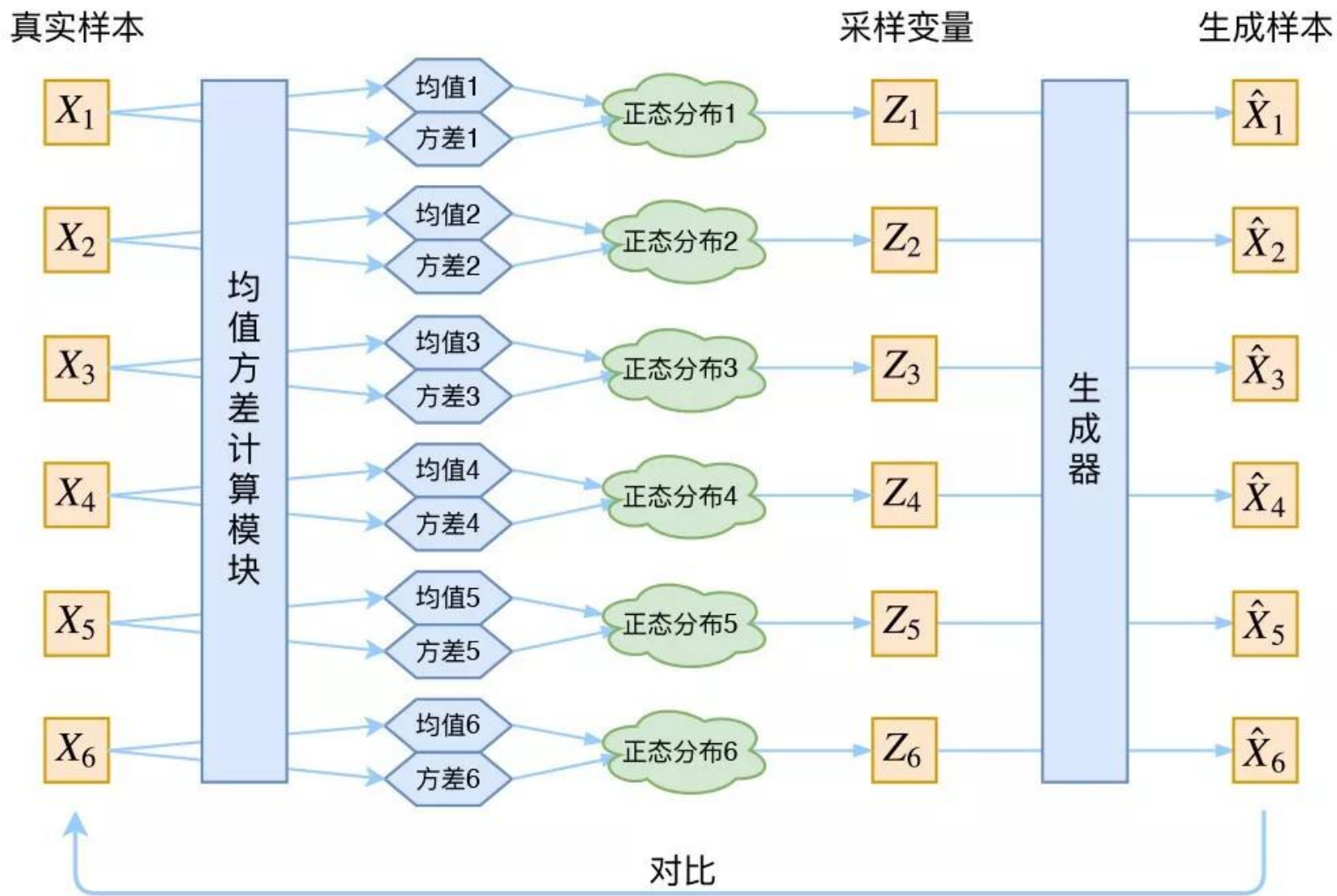
## 变分自编码器

- 首先编码器模块将输入图像转换为潜在空间上的两个参数/高阶特征：均值和方差，这两个参数可以定义潜在空间中的一个正态分布；
- 然后从这个正态分布中进行随机采样；
- 最后由编码器模块将潜在空间中的采样点映射回原始输入图像，从而达到最终的重构目的。

# 变分自编码器



# 变分自编码器



## 自编码器案例

- 手写数字生成VE模型
- 手写数字生成VAE模型

