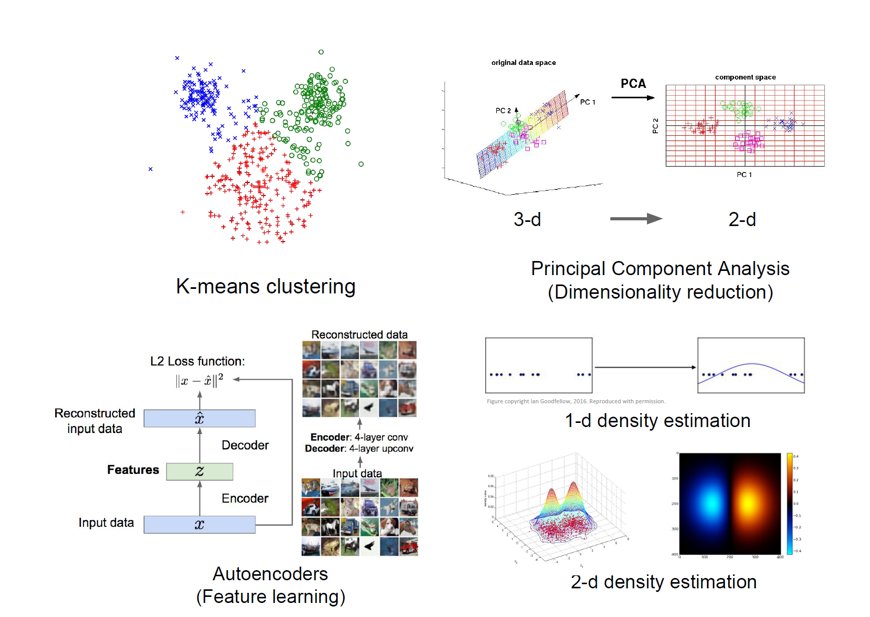
Cs231n 13강 요약

1. Unsupervised Learning

지금까지 배운 것은 supervised learning에 대해 주로 배웠다. Classification, regression, object detection, semantic segmentation, image captioning 등이다.

그렇다면 unsupervised learning은 data가 라벨 없이 주어져 있을 때, 데이터의 숨어있는 기본적인 구조를 학습하는 것을 목표로 학습하는 방법이다. 예시로는 clustering, dimensionality reduction, feature learning, density estimation 등이 있다.

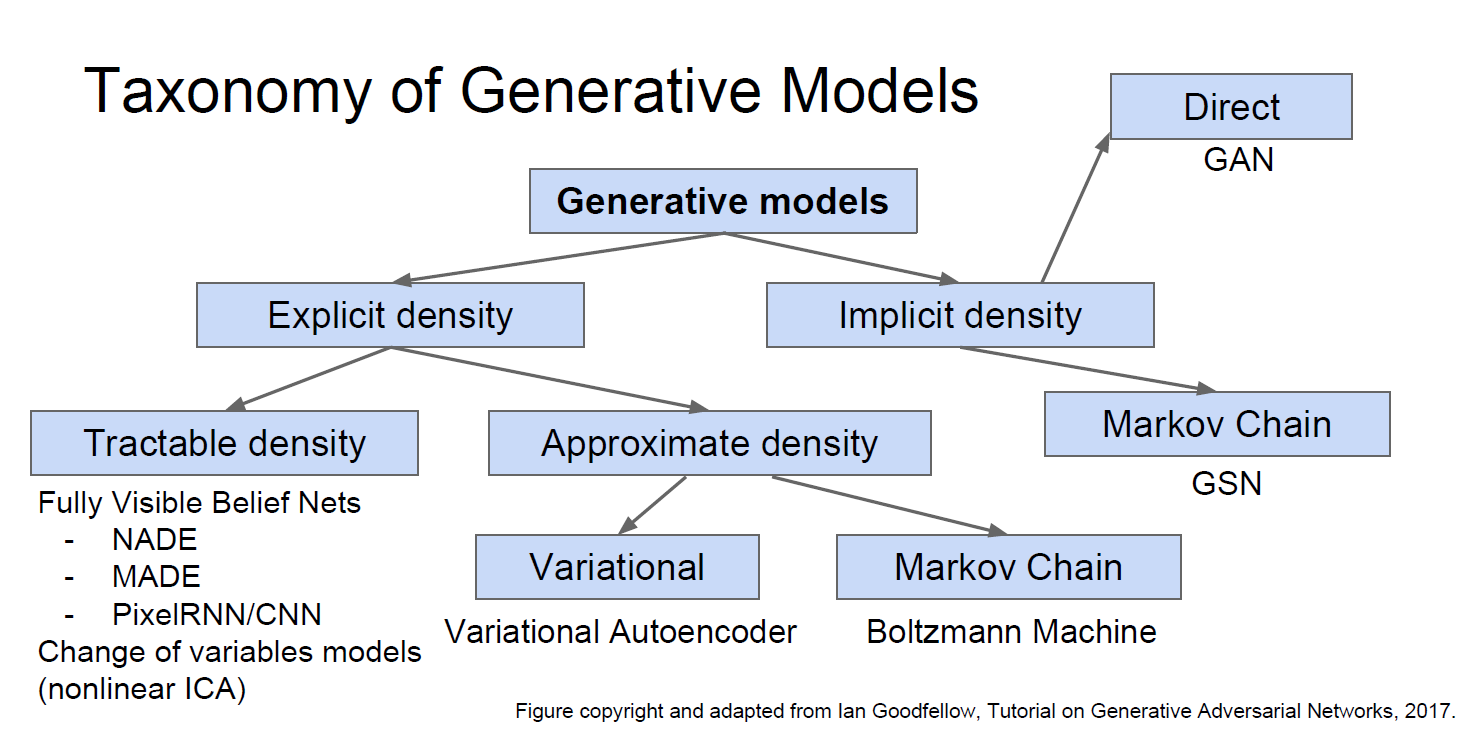


Unsupervised learning의 장점은 데이터를 많이 모을 수 있다는 점이다. 하지만 확실한 학습 결과로 분류 기준을 사람이 알 수 없다는 것이 단점이다.

1. Generative models

Generative model은 한국어로 생성 모델이라고 하며, data를 입력했을 때 동일한 분포에서 새로운 샘플들을 생성하는 것을 말한다. 생성 모델도 unsupervised learning에 속하는데, density estimation(분포 추정)을 하는 것이 핵심 과제이다. 분포 추정 방법은 크게 두 가지가 있다.

1. Explicit density estimation: 생성 모델을 명시적으로 나타내 주는 방법
2. Implicit density estimation: 생성 모델을 정의하지 않고 sample을 얻는 방법

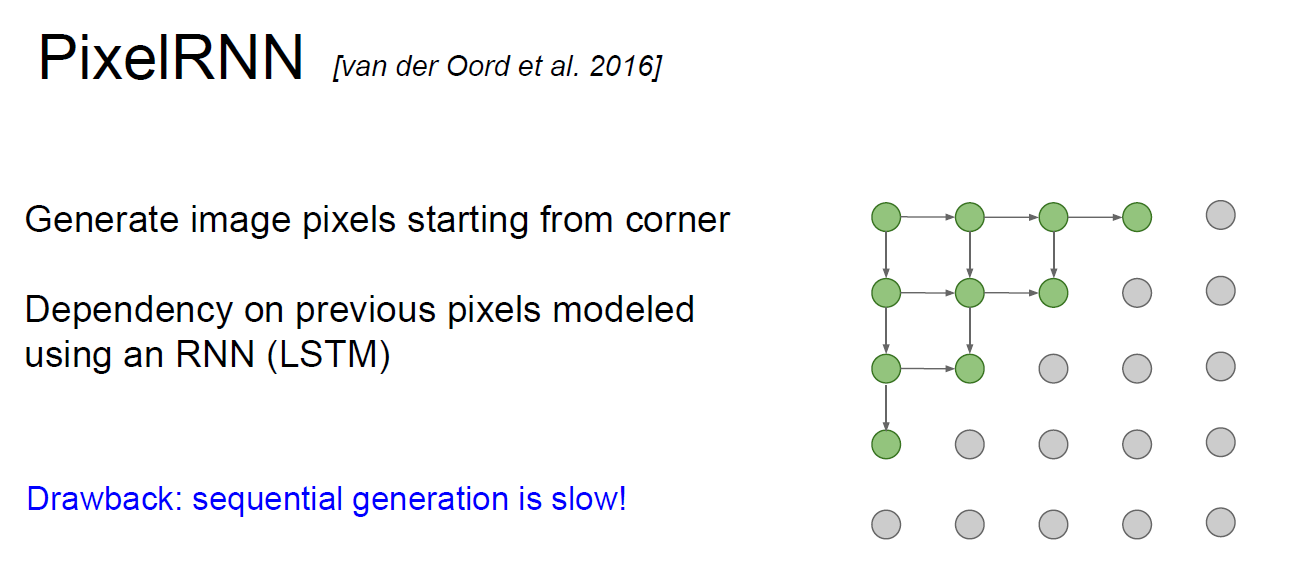


위 사진은 generative model의 종류를 나타낸다.

오늘 강의에서는 generative model 중 PixelRNN/CNN, Variational Autoencoder, GAN에 대해 설명한다.

1. PixelRNN/CNN

pixelRNN/CNN은 explicit density estimation 중 계산 가능한 density를 다루는 방법이다. Explicit density model을 만들기 위해, chain rule을 이용해 이미지를 1차원 분포들 간 곱의 형태로 분해한다. 이 분포들은 neural network를 이용하여 표현한다. 그렇다면 이전 pixel들을 어떻게 정의할까? 이전 pixel을 정의할 때 RNN을 사용하는 방법이 바로 PixelRNN이다.



인접한 픽셀들을 이용해 순차적으로 픽셀을 생성하고, LSTM을 사용해 모델링한다.

하지만 속도가 굉장히 느리다는 단점이 있다.

이를 보완하기 위해 나온 것이 바로 PixelCNN이다. RNN 대신 CNN을 사용해 픽셀을 생성하는데, 훨씬 빠르게 이미지를 생성할 수 있다.

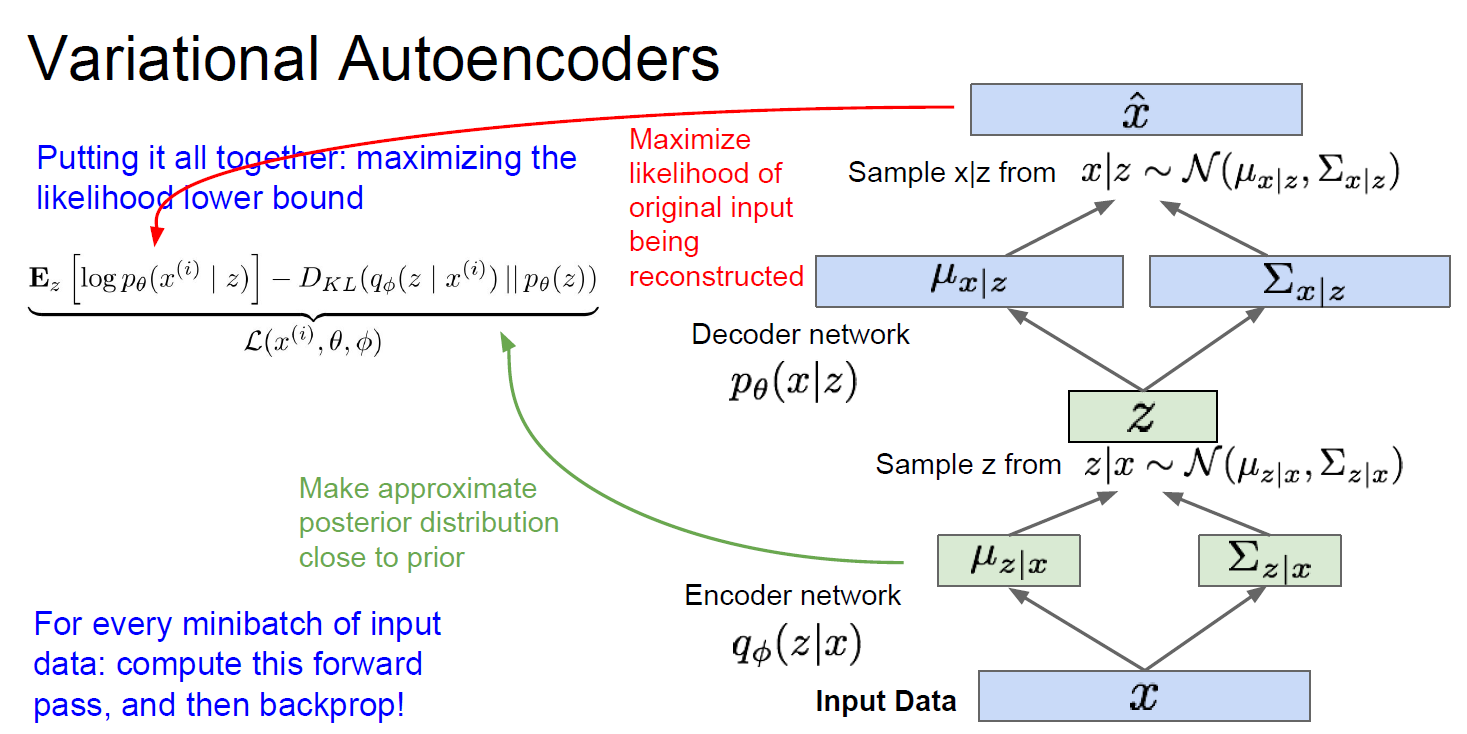
1. Variational Autoencoders

앞서 본 PixelCNN은 확률 모델이 계산 가능한 함수로 정의되었다. 하지만 VAE(Variational autoencoders)는 확률 모델이 계산 불가능한 함수로 정의된다. 따라서 lower bound를 구해서 계산 가능한 형태로 만들어 준다.

여기서 encoder란 input data x가 들어왔을 때, 그 특징 벡터 z를 추출하는 과정을 말한다. 예를 들면 sigmoid, FC layer, ReLU, CNN 등이 있다. 반대 개념인 decoder는 특징 벡터 z를 이용해 input data처럼 재구성하는 과정이다. 원본을 reconstruct한다고 생각하면 좋다.

따라서 VAE에서는 복원된 이미지와 원본 이미지간의 차이를 계산하기 위해 L2 loss function을 사용한다. 이때 decoder는 training 과정에서만 쓰이고, 실제로 test를 할 때에는 encoder만 가지고 진행한다.

아래 이미지는 training 과정이다.



이 방법은 서로 독립되어 있는 변수 z를 통해 여러 가지 이미지를 생성할 수 있지만 생성된 이미지가 blur한 특징을 갖는다는 점이 단점이다. 그리고 PixelRNN/CNN보다 이미지의 퀄리티가 떨어진다.

1. GAN(Generative Adversarial Networks)

GAN은 확률 분포를 계산하지 않고 단지 sample을 얻어내는 방법을 말한다. 즉 결과만 뽑아내는 것이다. GAN을 구현하기 위해서는 두 가지의 Network가 필요하다.

1. Generator Network: discriminator를 속여 가짜 이미지를 만드는 것이 목적
2. Discriminator Network: 진짜와 가짜 이미지를 구별하는 것이 목적

