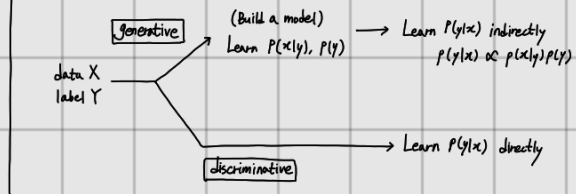


* generative model vs discriminative model

- 생성 모델 : posterior를 직접적으로 도출 / 데이터 범주 분류를 학습
데이터 X가 생성되는 과정도 두 개의 확률 모형 $p(x), p(x|y)$ 로 정의하여 $p(x|y)$ 를 직접적으로 도출
- 판별 모델 : posterior를 직접적으로 도출 / 결정 경계 (decision boundary)를 학습
데이터 X가 주어졌을 때 레이블 Y가 나타날 조건부 확률 $p(y|x)$ 를 직접적으로 변환
ex. 선형 SVM, 로지스틱 SVM



* 생성모델 Generative model

: 주어진 학습 데이터를 학습하여 학습 데이터의 분포를 따르는 유사한 데이터를 생성. Joint distribution, $p(v, h) \propto p(v|h)p(h)$

$$p(y=k|x) = \frac{p(x|y=k)p(y=k)}{p(x)} \quad (k: \text{클래스})$$

Supervised

- QDA 이차판별분석법 : 독립변수 X가 연속이고 학습률과 다항식 정제분포로 가정 $p(x|y=k) = \frac{1}{(2\pi)^{d/2} |\Sigma_k|^{1/2}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)}$, $p(y=k|x) = \frac{p(x|y=k)p(y=k)}{\sum_l p(x|y=l)p(y=l)}$
- LDA 선형판별분석법 : 각 Y클래스에 대한 독립변수 X의 고차항을 모두가 공통된 공분산 행렬을 가지는 다변수 정규분포로 가정
판별함수가 X에 대한 선형함수, 경계선 모양이 직선이 된다.

Unsupervised

통계적 생성 모델 : 밀도 추정 = 관측된 데이터의 분포로부터 전체 변수의 확률 분포를 추정

- KDE (Kernel Density Estimation) 커널 밀도 추정
- GMM (Gaussian Mixture Model) 혼합 가우시안 모델

deep learning Explicit density : 학습 데이터의 분포 기반

Tractable density : 학습 데이터의 분포를 직접적으로 구함.

$$p(x) = \prod_{i=1}^n p(x_i | x_{1:i-1})$$

Likelihood of image x Probability of i-th pixel value given all previous pixels
need to define

Chain rule 이용 pixelRNN

+ RNN (Recurrent Neural Network) 순환 신경망 : 은닉층이 개입될 때 입력층과 출력층이 개입될 수 있음. (순환 구조와 데이터 예측)

은닉층 $h_t = \tanh(W_h x_t + W_{hh} h_{t-1} + b)$ (W_h : 입력층 관련 가중치, W_{hh} : 은닉층 관련 가중치)

출력층 $y_t = f(W_y h_t + b)$

pixelCNN

+ CNN (Convolutional Neural Network) 합성곱 신경망 : 합성곱층(합성곱 연산 + 활성화 함수) + 풀링층(특정 영역 크기를 줄임)

Approximate density : 학습 데이터의 분포를 추정

Variational

Variational Autoencoder (VAE): 추론된 특징이 가우시안 확률 분포에 기반한 확률 분포 $p_\theta(x) = \int p_\theta(z) p_\theta(x|z) dz$, $p_\theta(x) = \frac{p_\theta(x|z)p_\theta(z)}{p_\theta(z)}$

+ Auto Encoder (AE) (인코더-디코더 구조. sequence to sequence 구조) : 은닉층의 구조가 Encoder(해석 방식 학습, 추론) + Decoder(리퍼스트링된 데이터로부터 저해독의 특징을 학습한 비지도 학습)

Markov Chain

Restricted Boltzmann Machine (RBM, 제한된 볼츠만 머신) : 방향성 X, 출력 ↔ 입력 $p(x, h) = \frac{1}{Z} e^{-E(x, h)}$, $E(x, h) = -h^T W x - c^T x - b^T h$

+ Deep Belief Network (DBN) : RBM 여러 층 $p(x, h', b', \dots, h') = \prod_{k=1}^{K-1} p(h^k | h^{k-1}) p(h^k, b^k)$

Implicit density : 학습 데이터의 분포를 몰라도 생성

Markov chain

GSN (Generative Stochastic Networks)

Direct

GAN (Generative Adversarial Networks) 적대적 생성 모델 : Generator ↔ Discriminator

object function
최적화 방법

$$\min_{\theta_g} \max_{\theta_d} [E_{x \sim p_{data}} \log D_\theta(x) + E_{z \sim p_\theta} \log (1 - D_\theta(z))] \quad (\text{Minimax Game})$$

- DCGAN (Deep Convolutional GAN) : Generator 부분에 CNN 적용 (세미 자동)

- CycleGAN : 두 개 GAN을 Cycle 형태로 구성. Unpaired한 이미지들 변환. 서로 다른 Generator, Discriminator.