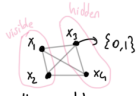
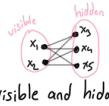
7- deep generative models

generative model = a model that can generate new data samples that resemble a given dataset
↳ k-means, hidden markov model, naive bayes, boltzman machines, GANs, variational autoencoders $p(x,y)$
discriminative model = model that focuses on learning the decision boundary between different classes in a dataset, conditional models $p(y|x)$
↳ better in classification



boltzman machine = stochastic fully connected recurrent neural network
↳ energy based model = energy func determines the probability distribution over all possible states → minimize
↳ visible units = connected to the external environment and represent data ↘ used in recommendation systems
↳ hidden units = used to discover features in the data
↳ training = based on maximum likelihood



restricted boltzman machine = there are no intra-layer connections between visible and hidden layers
↳ training can be performed using gradient descent of the maximum likelihood → expensive
↳ contrasive divergence = efficient way to approximate the gradient
① start with a training sample = start with input data (visible layer) and compute the probabilities of the hidden layer states
② gibbs sampling = using these probabilities, a sample of the hidden states is generated, then the visible states are reconstructed from this hidden layer, and the hidden states are resampled from this
③ update model parameters = using the difference in the product of probabilities between the original and reconstructed samples ⟶ approximates the gradient
④ iterative process = repeat for each example

gaussian bernoulli RBM = designed to handle continuous/real valued input data → others use binary data

deep belief networks (DBNs) = composed of stacking RBMs on top of each other
↳ greedy, layer by layer training
↳ able to learn distributed representations → [] = vertical + rectangle (previously unseen can be expressed better)

generative adversarial networks (GANs) =

↳ training = for D, min the probability of wrong decision
for G, max the probability of D making a mistake

deep convolutional generative adversarial networks (DCGANs) = improved GANs with NNs

autoencoders = input x → encoder network → representation h → decoder network → $\hat{x}$ minimize $\|x-\hat{x}\|^2$
↳ typically used for h feature extraction, after training decoder is trown out

variational encoders = x → encoder network → $p(z|x)$ $\xrightarrow{\text{draw sample}}$ z → decoder network → $p(x|z)$ $\xrightarrow{\text{draw sample}}\hat{x}$
↳ designed to model the input data in probabilistic manner, which helps in generating new points that are similar to inputs