UNIVERSITÀ PARTHENOPE
DiST
Dipartimento di Scienze e Tecnologie

# Machine Learning (part II)

# Deep Generative Models

Angelo Ciaramella

# Deep Belief Networks

- Deep belief networks (DBN)
  - first non-convolutional models to successfully admit training of deep architectures
  - Deep belief networks demonstrated that deep architectures can be successful
  - Generative models with several layers of latent variables
  - Latent variables are typically binary
  - Visible units may be binary or real
  - no intra-layer connections
    - Usually, every unit in each layer is connected to every unit in each neighboring layer
  - construct more sparsely connected DBNs
  - The connections between the top two layers are undirected
  - A DBN with only one hidden layer is a Restricted Boltzmann Machine (RBM)

# Boltzmann Machines

- BM
  - connectionist approach to learning arbitrary probability distributions over binary vectors

- Energy function

$$\mathbf{x} \in \{0,1\}^d$$

$$P(\boldsymbol{x}) = \frac{\exp\left(-E(\boldsymbol{x})\right)}{Z}$$

joint probability distribution over the observed variables

$$E(\boldsymbol{x}) = -\boldsymbol{x}^\top \boldsymbol{U} \boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x}$$

# Boltzmann Machines

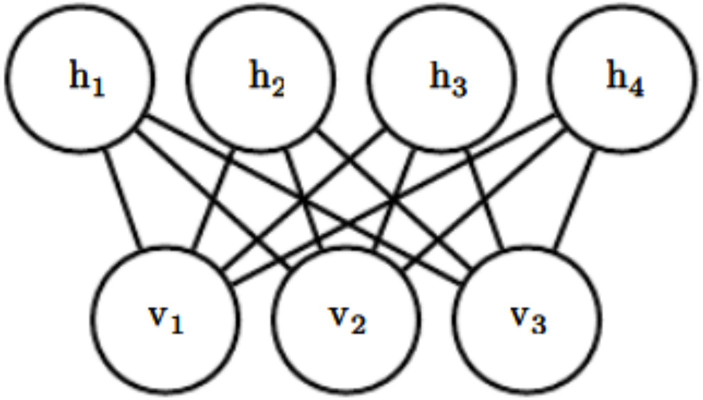
- Powerful when
  - not all the variables are observed
  - latent variables act similarly to hidden units in MLP
- Energy function

$$E(\boldsymbol{v}, \boldsymbol{h}) = -\boldsymbol{v}^\top \boldsymbol{R} \boldsymbol{v} - \boldsymbol{v}^\top \boldsymbol{W} \boldsymbol{h} - \boldsymbol{h}^\top \boldsymbol{S} \boldsymbol{h} - \boldsymbol{b}^\top \boldsymbol{v} - \boldsymbol{c}^\top \boldsymbol{h}$$

- Learning
  - intractable partition function

# Restricted Boltzmann Machines

- Restricted Boltzmann Machines (RBM)
  - Energy based model (named Harmonium, 1986)
  - undirected probabilistic graphical model
  - a layer of observable variables and a single layer of latent variables
  - may be stacked (one on top of the other) to form deeper models



# RBM

- Energy function

$$E(\boldsymbol{v},\boldsymbol{h}) = -\boldsymbol{b}^\top\boldsymbol{v} - \boldsymbol{c}^\top\boldsymbol{h} - \boldsymbol{v}^\top\boldsymbol{W}\boldsymbol{h}$$

$$P(\mathbf{v}=\boldsymbol{v},\mathbf{h}=\boldsymbol{h}) = \frac{1}{Z}\exp\left(-E(\boldsymbol{v},\boldsymbol{h})\right)$$

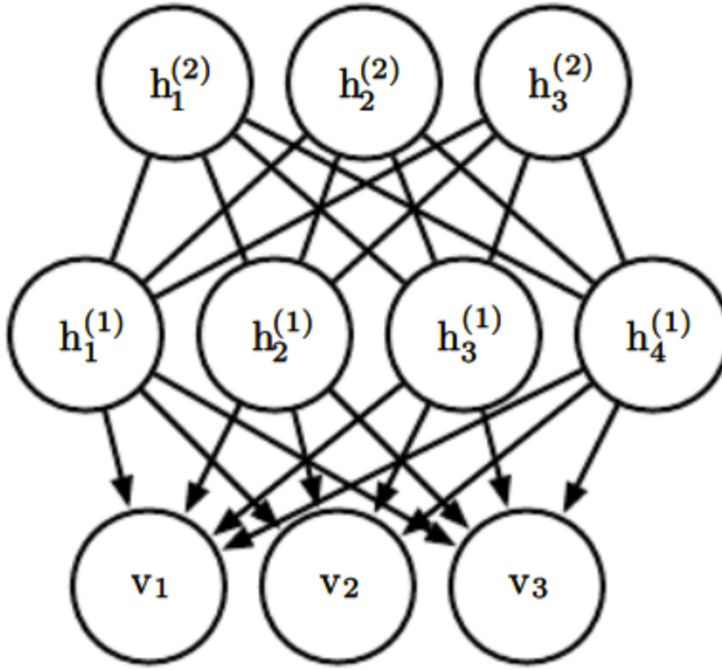
$$Z = \sum_{\boldsymbol{v}}\sum_{\boldsymbol{h}}\exp\left\{-E(\boldsymbol{v},\boldsymbol{h})\right\}$$

# Training RBM

- RBM training
  - training models that have intractable partition functions
    - CD
    - SML
    - Ratio matching
    - …

# DBNs

The connections between the top two layers are undirected.
The connections between all other layers are directed




A deep belief network is a hybrid graphical model involving both directed and undirected connections

# DBNs

- Probability distribution

$$P(\boldsymbol{h}^{(l)}, \boldsymbol{h}^{(l-1)}) \propto \exp\left(\boldsymbol{b}^{(l)^\top} \boldsymbol{h}^{(l)} + \boldsymbol{b}^{(l-1)\top} \boldsymbol{h}^{(l-1)} + \boldsymbol{h}^{(l-1)\top} \boldsymbol{W}^{(l)} \boldsymbol{h}^{(l)}\right),$$

$$P(h_i^{(k)} = 1 \mid \boldsymbol{h}^{(k+1)}) = \sigma\left(b_i^{(k)} + \boldsymbol{W}_{:,i}^{(k+1)\top} \boldsymbol{h}^{(k+1)}\right) \forall i, \forall k \in 1, \ldots, l-2,$$

$$P(v_i = 1 \mid \boldsymbol{h}^{(1)}) = \sigma\left(b_i^{(0)} + \boldsymbol{W}_{:,i}^{(1)\top} \boldsymbol{h}^{(1)}\right) \forall i.$$

$$\mathbf{v} \sim \mathcal{N}\left(\boldsymbol{v}; \boldsymbol{b}^{(0)} + \boldsymbol{W}^{(1)\top} \boldsymbol{h}^{(1)}, \boldsymbol{\beta}^{-1}\right)$$

Real-valued visible units

# Training DBN

- Learning
  - Training an RBM to maximize by CD or SML

$$\mathbb{E}_{\mathbf{v}\sim p_{\text{data}}} \log p(\boldsymbol{v})$$

  - Second RBM maximize

$$\mathbb{E}_{\mathbf{v}\sim p_{\text{data}}} \mathbb{E}_{\mathbf{h}^{(1)}\sim p^{(1)}(\boldsymbol{h}^{(1)}|\boldsymbol{v})} \log p^{(2)}(\boldsymbol{h}^{(1)})$$

- procedure can be repeated indefinitely on many layers

# DBF as generative model

- Improve classification models
  - weights from the DBN and use them to define an MLP

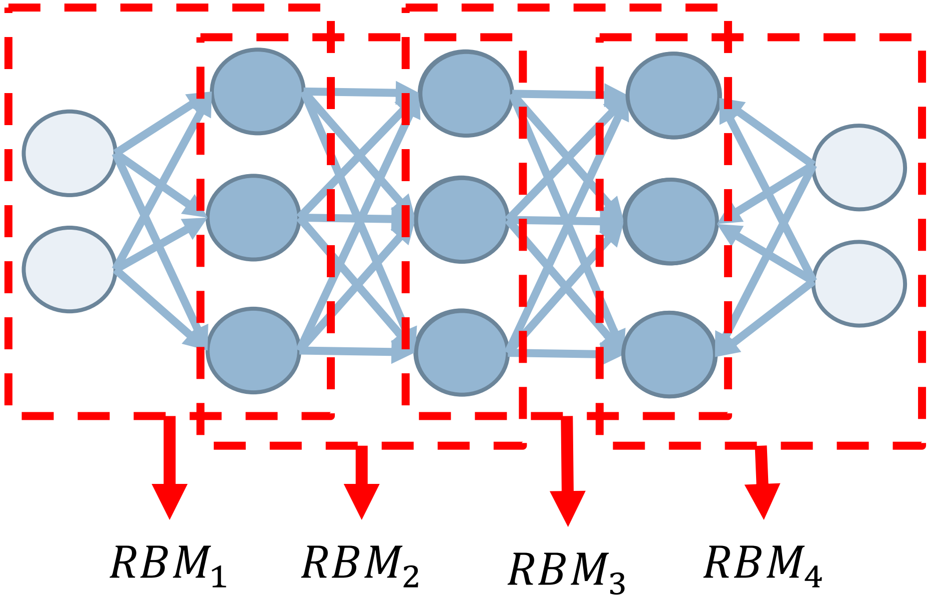
$$h^{(1)} = \sigma\left(b^{(1)} + v^{\top} W^{(1)}\right).$$

$$h^{(l)} = \sigma\left(b_i^{(l)} + h^{(l-1)\top} W^{(l)}\right) \forall l \in 2, \ldots, m,$$

  - Training MLP for classification tasks (discriminative fine-tuning)

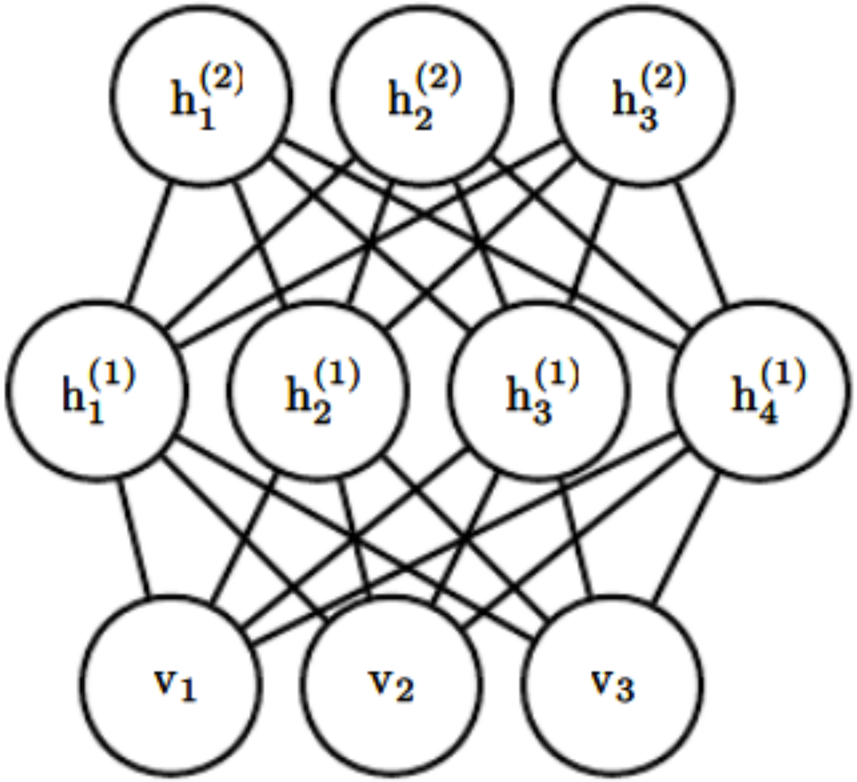
# Deep Belief Networks






**unsupervised, layer-wise, greedy pre-training**

# Deep Boltzmann Machines




entirely undirected model

# Deep Boltzmann Machines

- joint probability

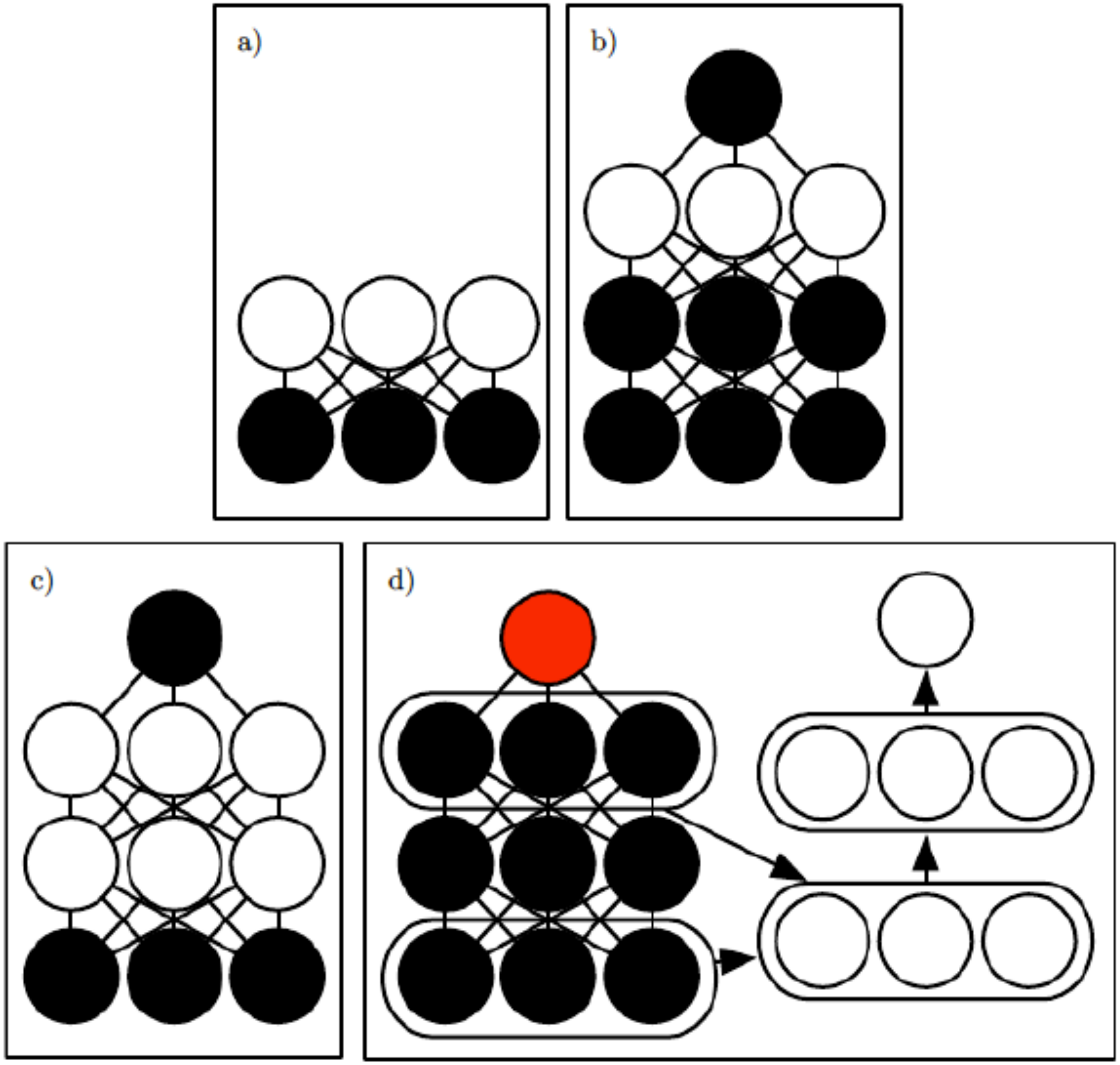
$$P\left(\boldsymbol{v}, \boldsymbol{h}^{(1)}, \boldsymbol{h}^{(2)}, \boldsymbol{h}^{(3)}\right) = \frac{1}{Z(\boldsymbol{\theta})} \exp\left(-E(\boldsymbol{v}, \boldsymbol{h}^{(1)}, \boldsymbol{h}^{(2)}, \boldsymbol{h}^{(3)}; \boldsymbol{\theta})\right)$$

- Energy function

$$E(\boldsymbol{v}, \boldsymbol{h}^{(1)}, \boldsymbol{h}^{(2)}, \boldsymbol{h}^{(3)}; \boldsymbol{\theta}) = -\boldsymbol{v}^{\top}\boldsymbol{W}^{(1)}\boldsymbol{h}^{(1)} - \boldsymbol{h}^{(1)\top}\boldsymbol{W}^{(2)}\boldsymbol{h}^{(2)} - \boldsymbol{h}^{(2)\top}\boldsymbol{W}^{(3)}\boldsymbol{h}^{(3)}$$

- Learning
  - challenge of an intractable partition function
  - challenge of an intractable posterior distribution

# Jointly Training DBM



# Generative Adversarial Networks

- GANs
  - game theoretic scenario
    - generator network must compete against an adversary

$$x = g(z; \theta^{(g)})$$

Generator network

$$d(x; \theta^{(d)})$$

Discriminator network

probability that x is a real training example rather than a fake sample drawn from the model

# Generative Adversarial Networks

- Zero-sum game

$$v(\theta^{(g)}, \theta^{(d)})$$

payoff of the discriminator

- to maximize its own payoff

$$g^* = \arg\min_g \max_d v(g, d)$$

$$v(\theta^{(g)}, \theta^{(d)}) = \mathbb{E}_{\mathbf{x}\sim p_{\text{data}}} \log d(\boldsymbol{x}) + \mathbb{E}_{\boldsymbol{x}\sim p_{\text{model}}} \log (1 - d(\boldsymbol{x}))$$

This drives the discriminator to attempt to learn to correctly classify samples as real or fake. Simultaneously, the generator attempts to fool the classifier into believing its samples are real.

# GANs



Images generated by GANs trained on the LSUN dataset