

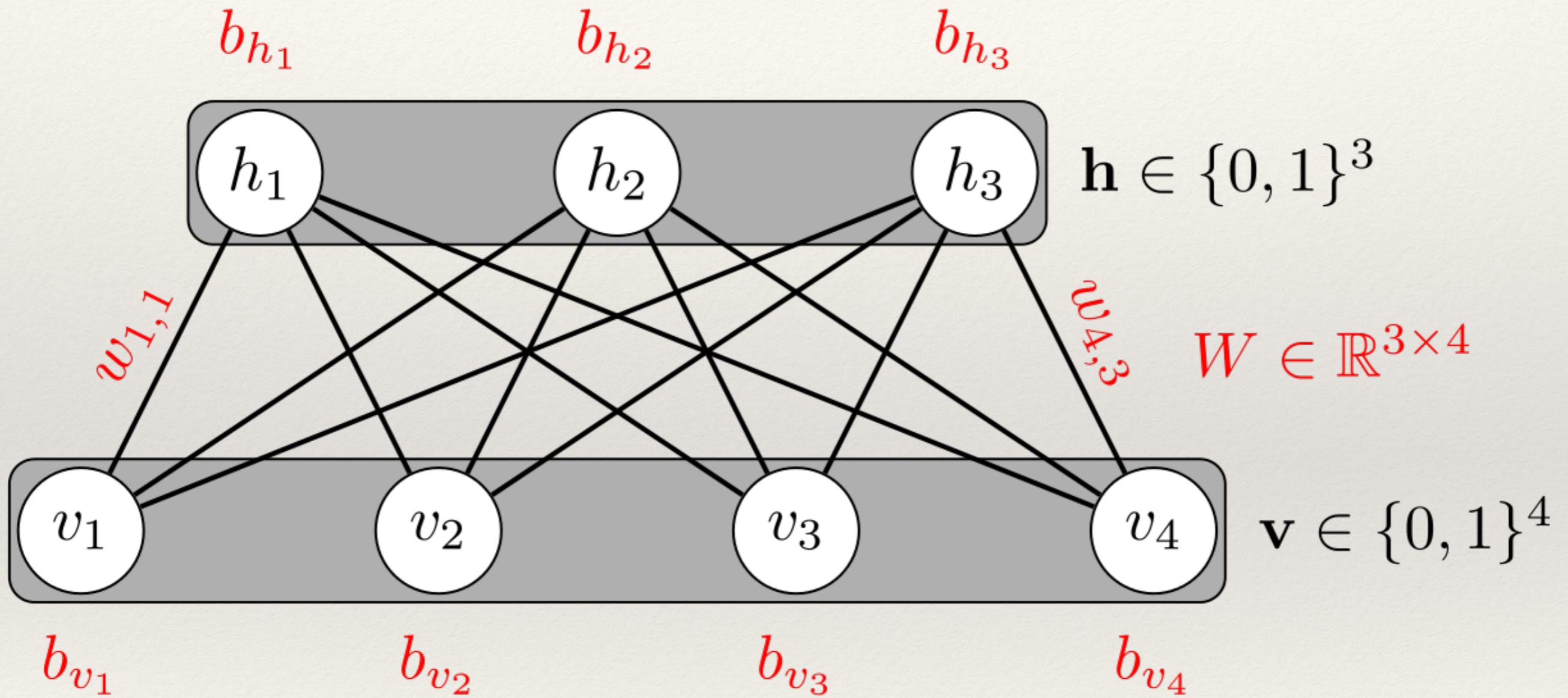
# *E9 205 Machine Learning for Signal Processing*

---

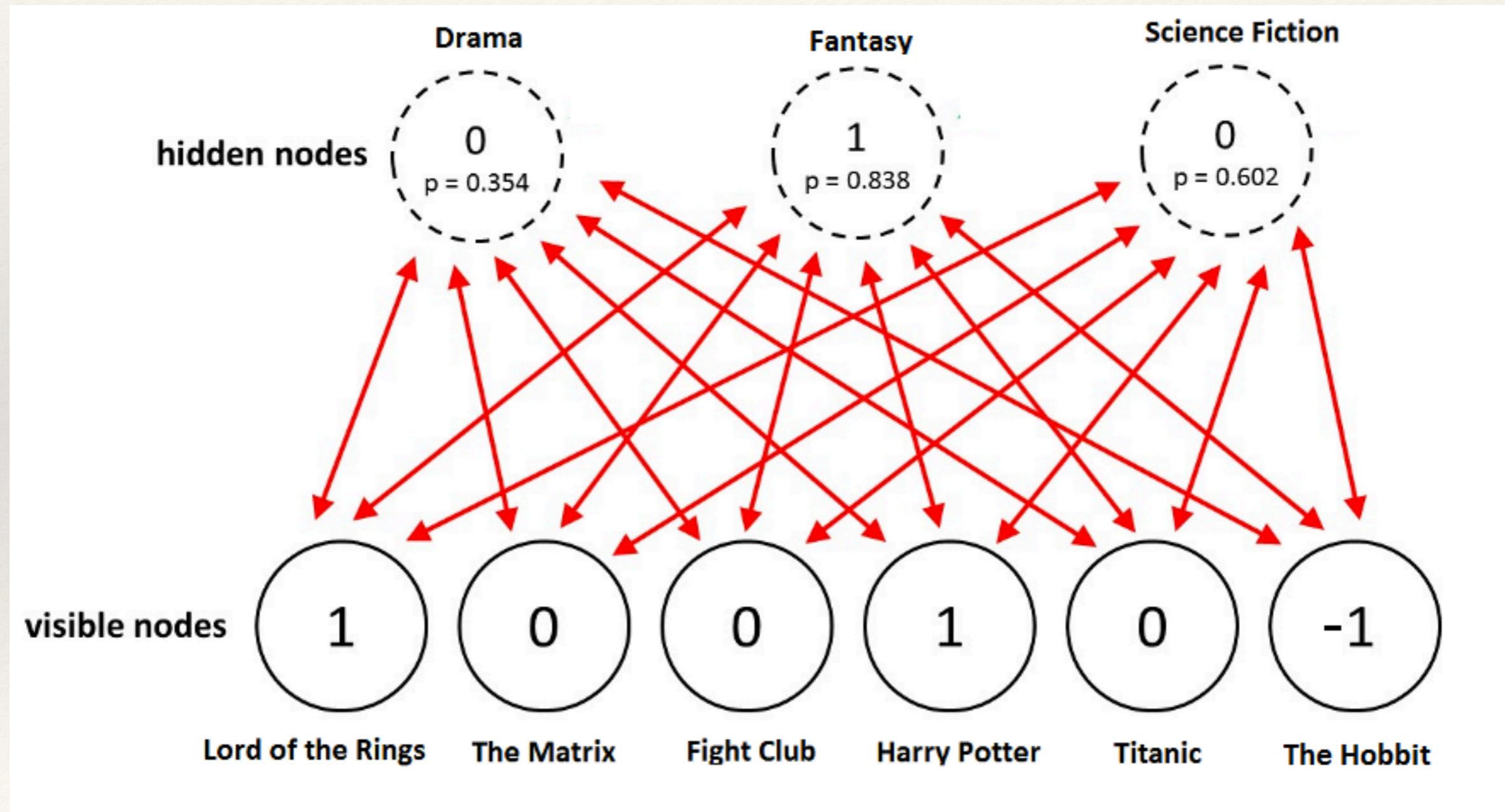
Deep Unsupervised Learning

12-11-2018

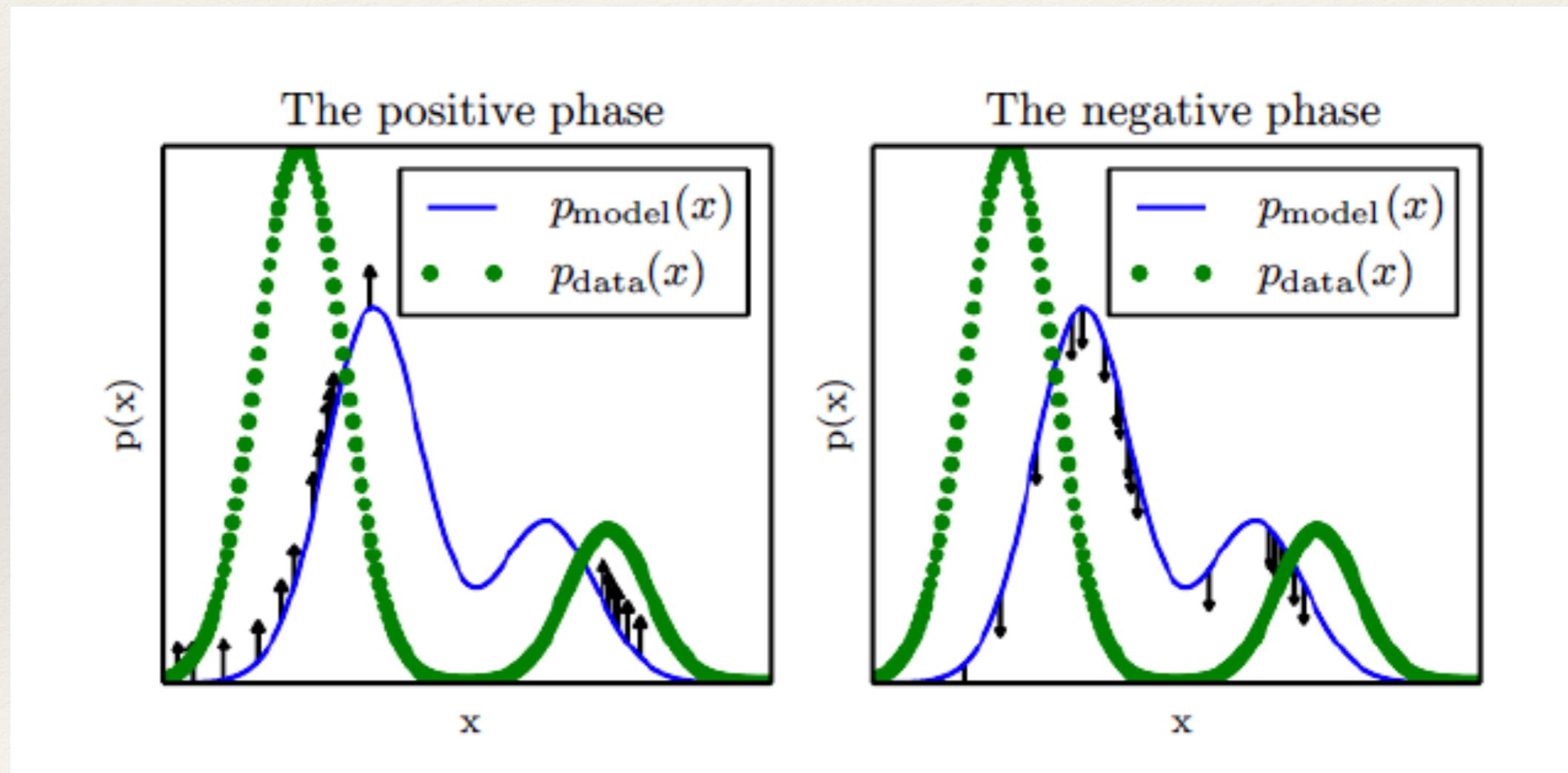
# Restricted Boltzmann Machines



# Restricted Boltzmann Machines



# Positive and Negative Phase



# Contrastive Divergence Algorithm

---

**Algorithm 18.2** The contrastive divergence algorithm, using gradient ascent as the optimization procedure.

---

Set  $\epsilon$ , the step size, to a small positive number.

Set  $k$ , the number of Gibbs steps, high enough to allow a Markov chain sampling from  $p(\mathbf{x}; \boldsymbol{\theta})$  to mix when initialized from  $p_{\text{data}}$ . Perhaps 1-20 to train an RBM on a small image patch.

**while** not converged **do**

    Sample a minibatch of  $m$  examples  $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}\}$  from the training set.

$\mathbf{g} \leftarrow \frac{1}{m} \sum_{i=1}^m \nabla_{\boldsymbol{\theta}} \log \tilde{p}(\mathbf{x}^{(i)}; \boldsymbol{\theta})$ .

**for**  $i = 1$  to  $m$  **do**

$\tilde{\mathbf{x}}^{(i)} \leftarrow \mathbf{x}^{(i)}$ .

**end for**

**for**  $i = 1$  to  $k$  **do**

**for**  $j = 1$  to  $m$  **do**

$\tilde{\mathbf{x}}^{(j)} \leftarrow \text{gibbs\_update}(\tilde{\mathbf{x}}^{(j)})$ .

**end for**

**end for**

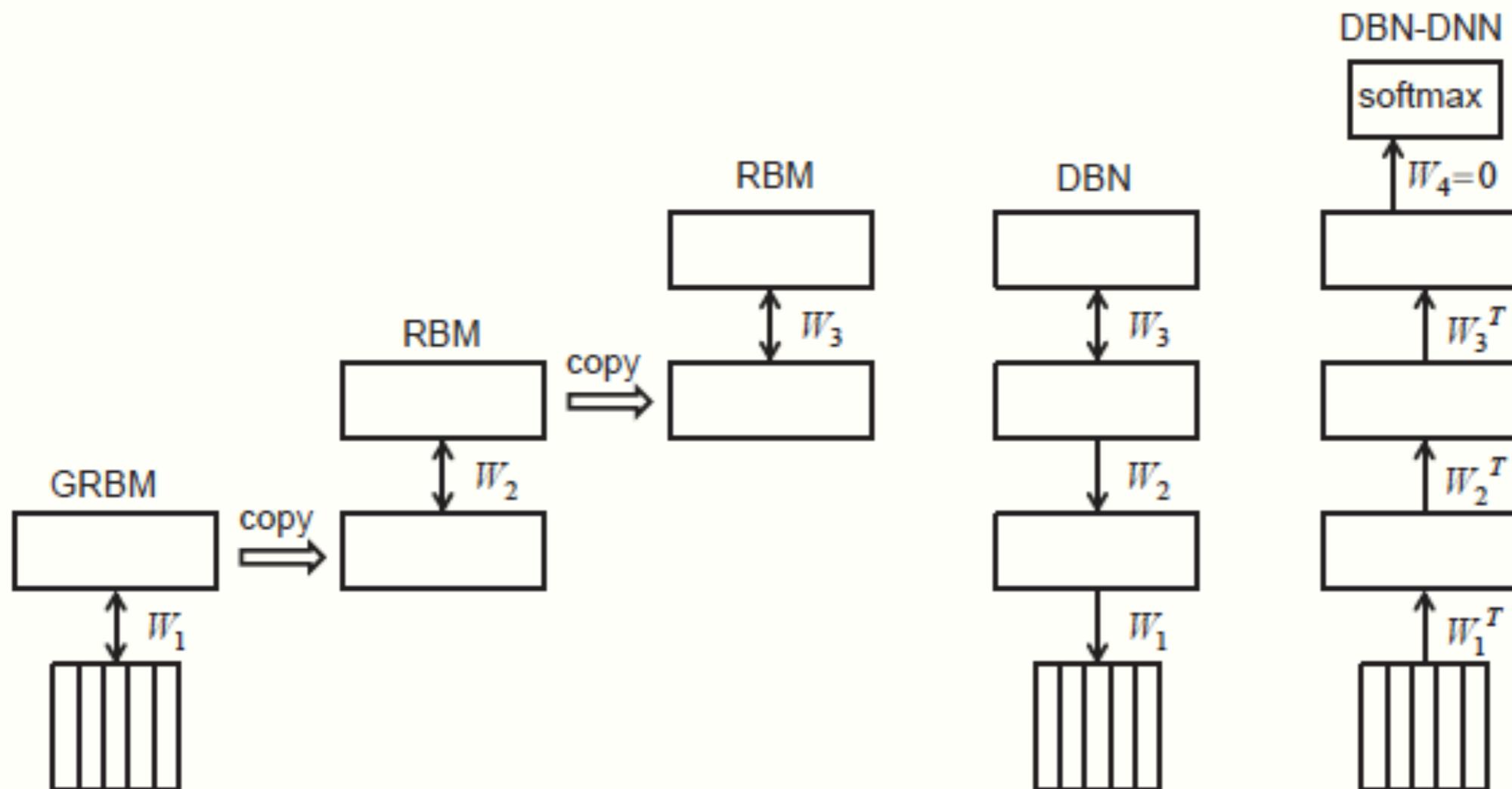
$\mathbf{g} \leftarrow \mathbf{g} - \frac{1}{m} \sum_{i=1}^m \nabla_{\boldsymbol{\theta}} \log \tilde{p}(\tilde{\mathbf{x}}^{(i)}; \boldsymbol{\theta})$ .

$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \epsilon \mathbf{g}$ .

**end while**

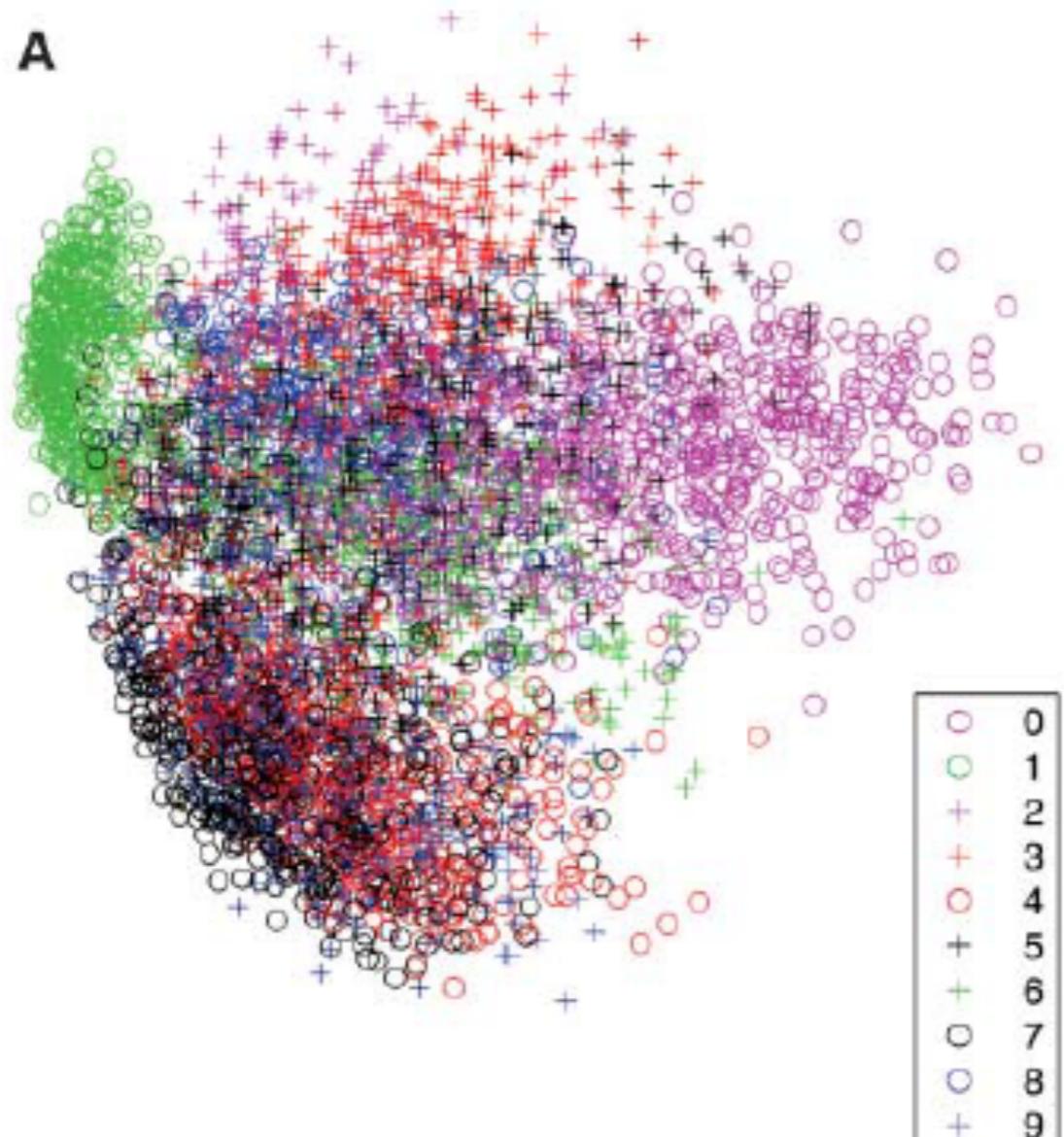
---

# DBN for DNN Initialization

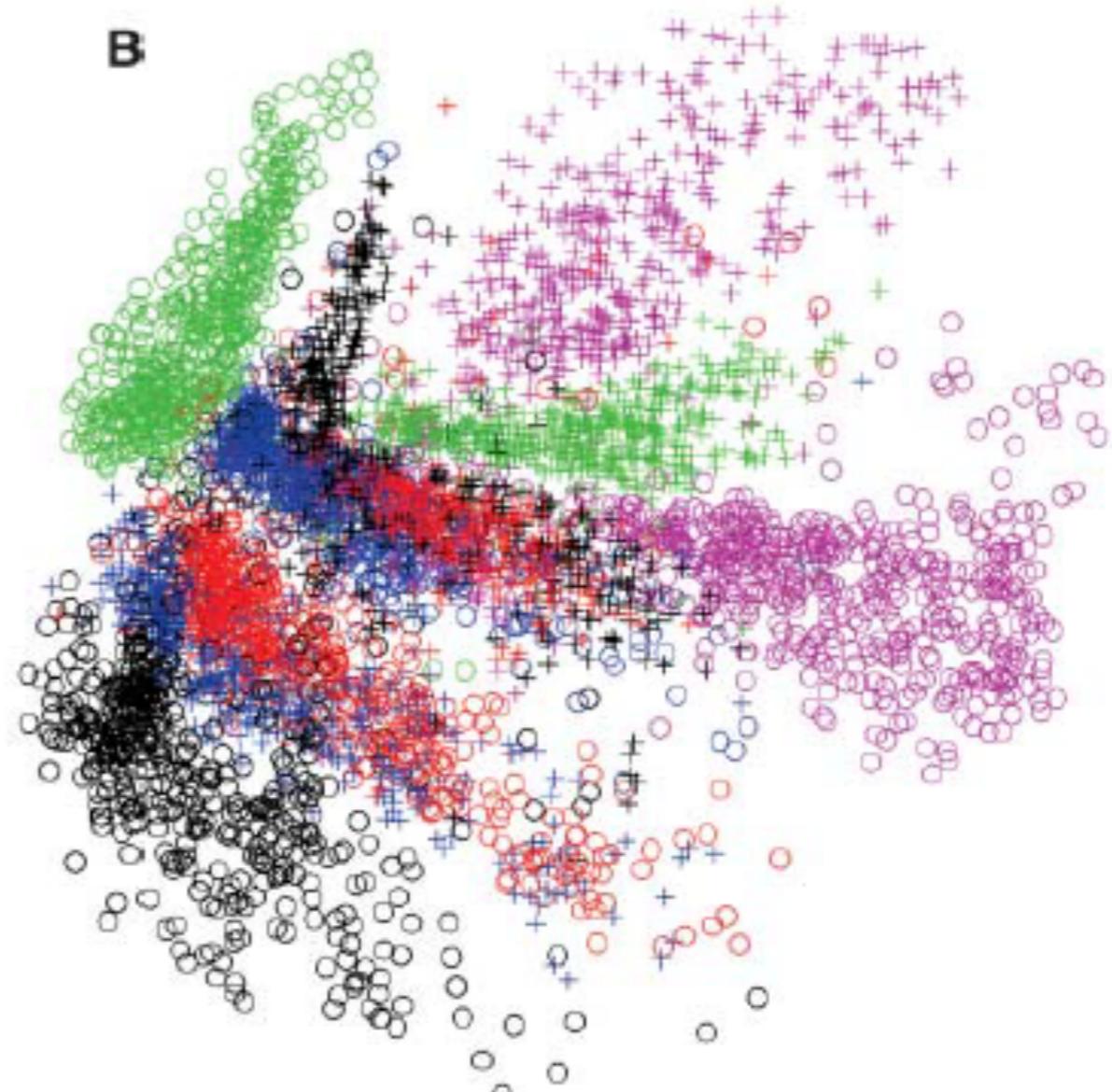


[Hinton, 2013]

# Restricted Boltzmann Machine



**PCA**



**RBM**

# Data Generation Using RBMs

