



南京航空航天大学

NANJING UNIVERSITY OF AERONAUTICS & ASTRONAUTICS

Deep Belief Networks (深度信念网络)



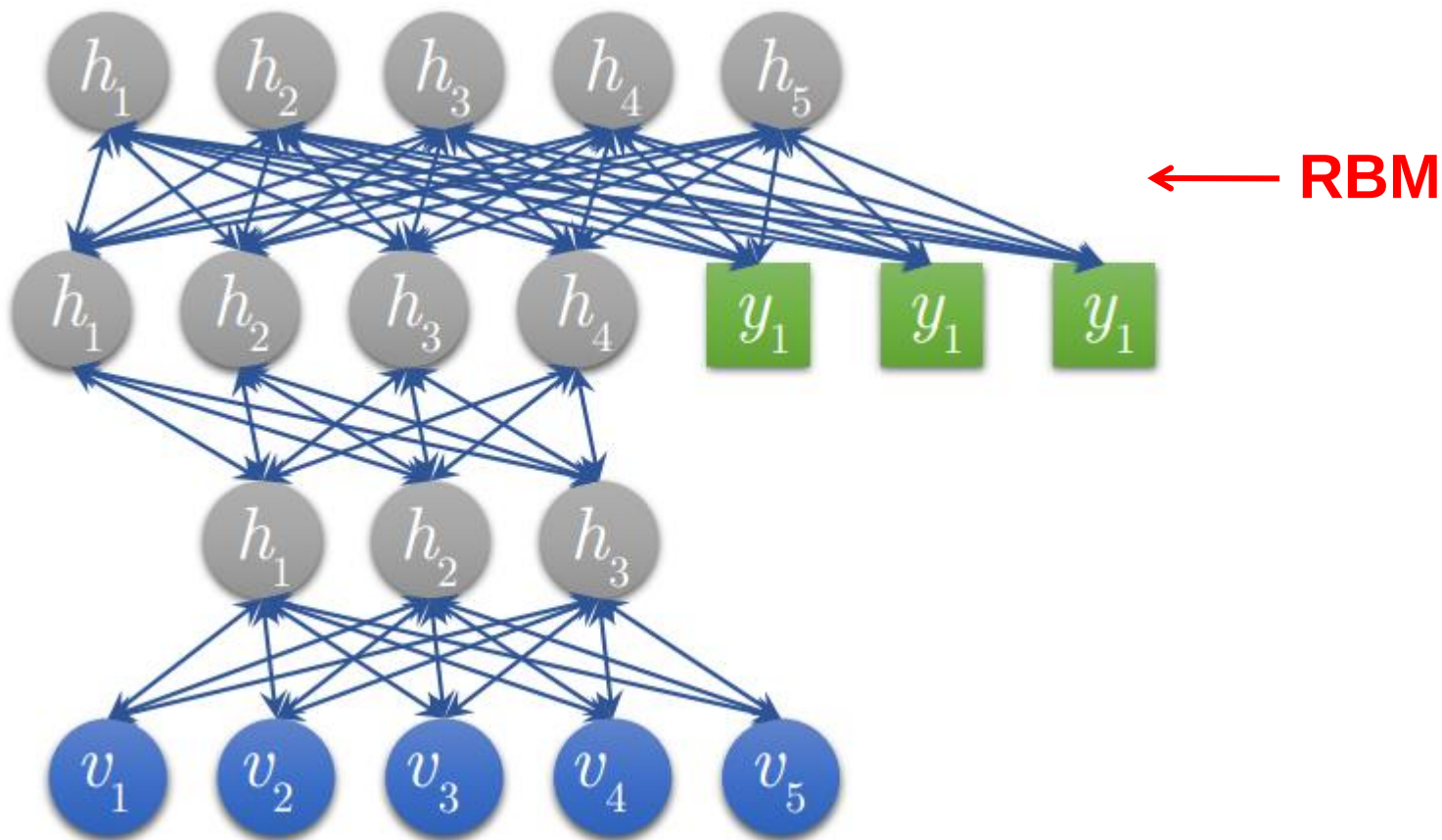
颜逸凡



DBN

南京航空航天大学

NANJING UNIVERSITY OF AERONAUTICS & ASTRONAUTICS

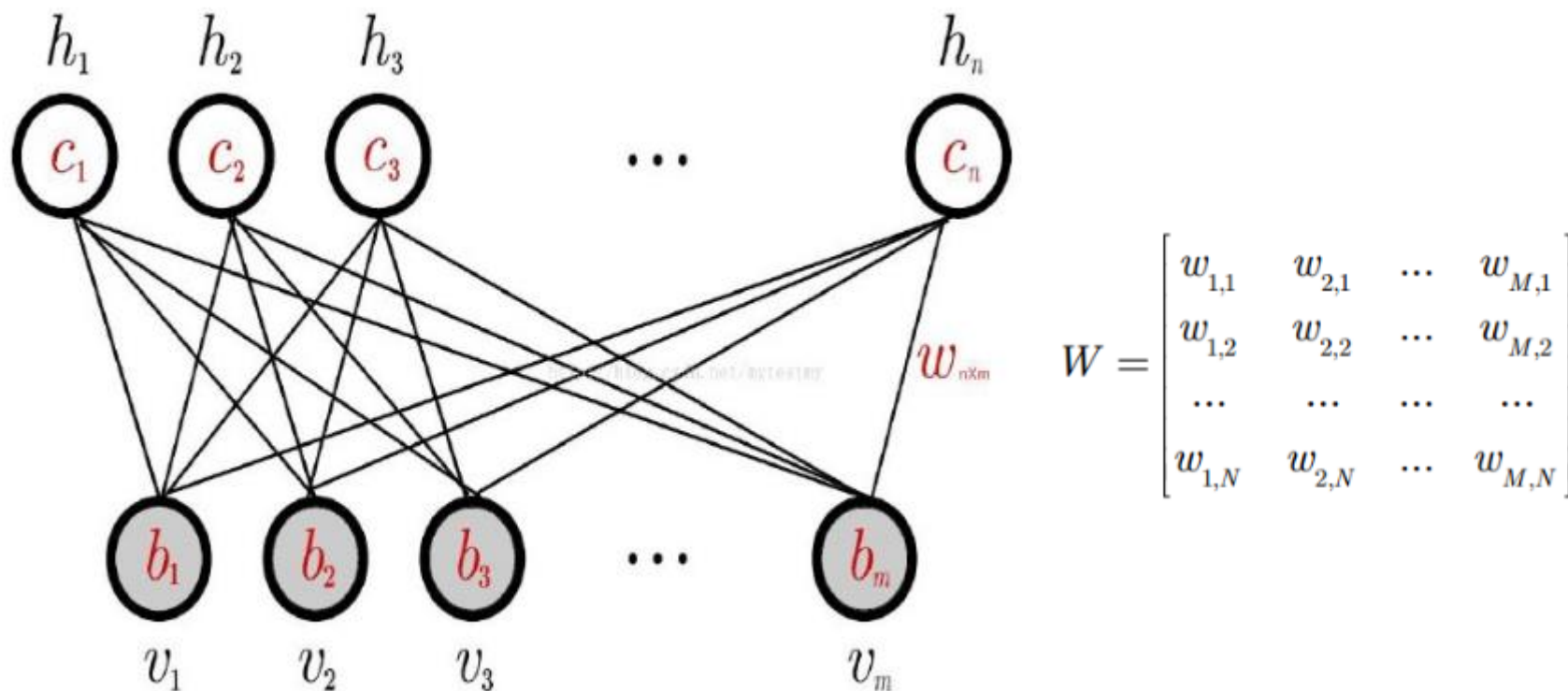




南京航空航天大学

NANJING UNIVERSITY OF AERONAUTICS & ASTRONAUTICS

RBM(受限玻尔兹曼机)





RBM的使用

南京航空航天大学

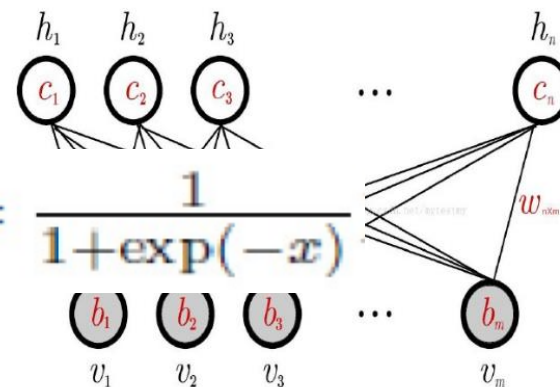
NANJING UNIVERSITY OF AERONAUTICS & ASTRONAUTICS

$$\mathbf{x} = (x_1, x_2, \dots, x_M)^T$$

$$p(h_j = 1 | v) = \sigma\left(\sum_{i=1}^m w_{ji} v_i + c_j\right)$$



$$\sigma(x) = \frac{1}{1 + \exp(-x)}$$



$$h_j = \begin{cases} 1, & P(h_j = 1) \geq u \\ 0, & P(h_j = 1) < u \end{cases}$$

$$u \sim U(0,1)$$

$$W = \begin{bmatrix} w_{1,1} & w_{2,1} & \dots & w_{M,1} \\ w_{1,2} & w_{2,2} & \dots & w_{M,2} \\ \dots & \dots & \dots & \dots \\ w_{1,N} & w_{2,N} & \dots & w_{M,N} \end{bmatrix}$$





南京航空航天大学

NANJING UNIVERSITY OF AERONAUTICS & ASTRONAUTICS

训练RBM

算法 1. 对比散度 (Contrastive Divergence)

对于 训练集中的每一条记录 $\mathbf{x}$ 有:

将 $\mathbf{x}$ 附给显层 $\mathbf{v}^{(0)}$, 计算它使隐层神经元被开启的概率

$$P(h_j^{(0)} = 1 | \mathbf{v}^{(0)}) = \sigma(\mathbf{W}_j \mathbf{v}^{(0)})$$

式中的上标用于区别不同向量, 下标用于区别同一向量中的不同维。

然后, 从计算出的概率分布中抽取一个样本

$$\mathbf{h}^{(0)} \sim P(\mathbf{h}^{(0)} | \mathbf{v}^{(0)})$$

用 $\mathbf{h}^{(0)}$ 重构显层

$$P(v_i^{(1)} = 1 | \mathbf{h}^{(0)}) = \sigma(\mathbf{W}_i^T \mathbf{h}^{(0)})$$

同样, 抽取出显层的一个样本

$$\mathbf{v}^{(1)} \sim P(\mathbf{v}^{(1)} | \mathbf{h}^{(0)})$$

再次用显层神经元 (重构之后的) 计算出隐层神经元被开启的概率

$$P(h_j^{(1)} = 1 | \mathbf{v}^{(1)}) = \sigma(\mathbf{W}_j \mathbf{v}^{(1)})$$

按下式更新权重:

$$\Delta w_{ij} \leftarrow \Delta w_{ij} + p(H_i = 1 | \mathbf{v}^{(0)}) \cdot v_j^{(0)} - p(H_i = 1 | \mathbf{v}^{(k)}) \cdot v_j^{(k)}$$

$$\Delta b_j \leftarrow \Delta b_j + v_j^{(0)} - v_j^{(k)}$$

$$\Delta c_i \leftarrow \Delta c_i + p(H_i = 1 | \mathbf{v}^{(0)}) - p(H_i = 1 | \mathbf{v}^{(k)})$$



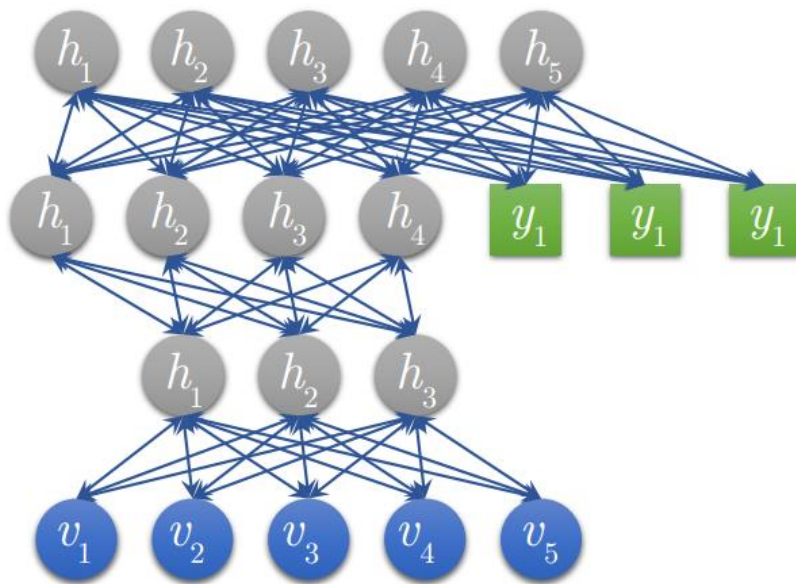


南京航空航天大学

NANJING UNIVERSITY OF AERONAUTICS & ASTRONAUTICS

DBN的训练过程

1. 首先充分训练第一个 RBM;
2. 固定第一个 RBM 的权重和偏移量, 然后使用其隐性神经元的状态, 作为第二个 RBM 的输入向量;
3. 充分训练第二个 RBM 后, 将第二个 RBM 堆叠在第一个 RBM 的上方;
4. 重复以上三个步骤任意多次;
5. 如果训练集中的数据有标签, 那么在 RBM 的显层中除了显性神经元, 还需要起进行训练:
 - a) 假设顶层 RBM 的显层有 500 个显性分成了 10 类;
 - b) 那么顶层 RBM 的显层有 510 个显性, 相应的标签神经元被打开设为 1, 而

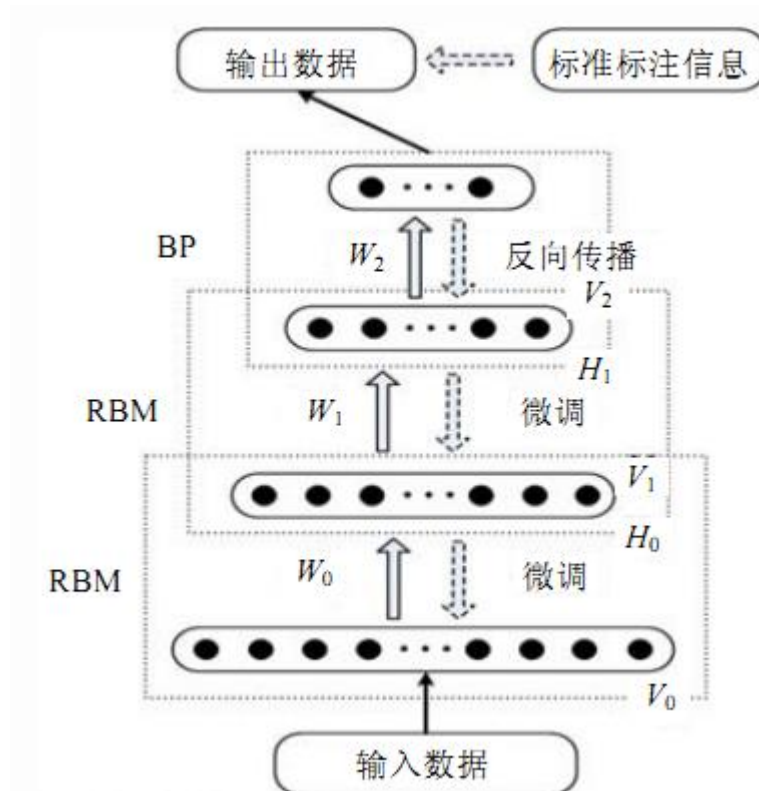




南京航空航天大学

NANJING UNIVERSITY OF AERONAUTICS & ASTRONAUTICS

经典的DBN网络结构





南京航空航天大学

NANJING UNIVERSITY OF AERONAUTICS & ASTRONAUTICS

Thank you

