Autoencoders

Monday, 19 February 2024 21:57

Supervised Learning:

Labeled data set for training:

$$\mathcal{D} = \{\bar{x}_n, \bar{t}_n\}_{n=1}^{N}$$

Labeled data $\mathcal{D}_L = \{\bar{x}_n, \bar{t}_n\}_{n=1}^{N_L}$

Unlabeled data $\mathcal{D}_U = \{\bar{x}_m\}_{m=1}^{N_U}$

$$N_U \gg N_L$$

Semi-supervised Learning:

Self-training method

Train the model using $\mathcal{D}_L$

Use trained model on unlabeled data $\mathcal{D}_U$

- Identify confidently labeled examples in $\mathcal{D}_U$ and add them to $\mathcal{D}_L$.
- Re-train the model using the updated $\mathcal{D}_L$.

Repeat until a condition is reached.

Two-phase training:

1. Pre-training using "unsupervised" learning with $\mathcal{D}_U$
2. Fine-tuning using supervised learning with $\mathcal{D}_L$

Deep Feedforward Neural Network (DFNN)

- More than 2 hidden layers.

Input		Desired output	Actual output
$\bar{x}_n$	$\longrightarrow$	$\bar{t}_n$	$\bar{y}_n$
d		K	K

Hetero-associative mapping

Auto-associative mapping

Input		Desired output	Actual output
$\bar{x}_n$	$\longrightarrow$	$\bar{x}_n$	$\bar{y}_n$

Principal Component Analysis (PCA) method for dimension reduction:

$$\bar{x}_n \longrightarrow \bar{a}_n$$

$$d \qquad\qquad l < d$$

$$\mathcal{D} = \{\bar{x}_n\}_{n=1}^{N}$$

Covariance Matrix

$$C = \frac{1}{N}\sum_{n=1}^{N}(\bar{x}_n - \bar{\mu})(\bar{x}_n - \bar{\mu})^t$$

$$C\bar{q} = \lambda\bar{q}$$

$\bar{q}_i$: Eigenvector $\quad i = 1, 2, \ldots, d$

λ_i : Eigenvalue

$$a_{ni} = (\bar{x}_n - \bar{\mu})^t \cdot \bar{q}_i \quad \leftarrow \text{Linear method}$$

$$\bar{a}_n = [a_{ni}]_{i=1}^{l}$$

$$\bar{z}_n = \bar{x}_n - \bar{\mu}$$

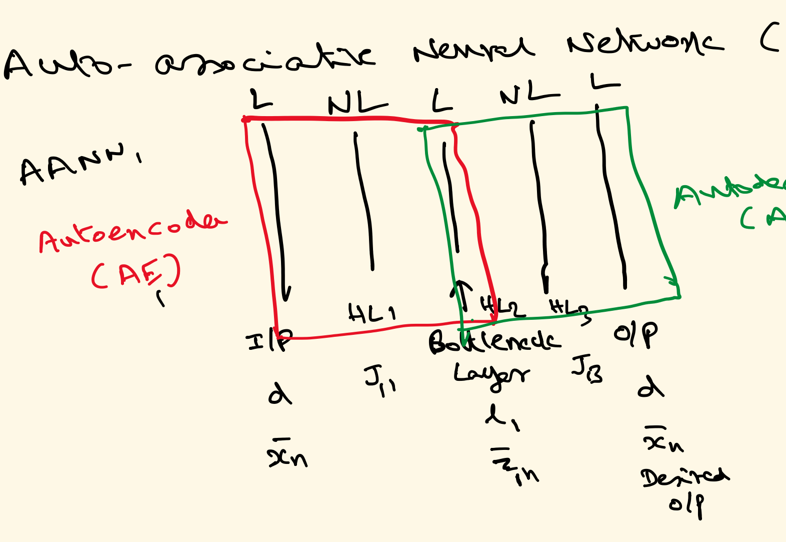
$$\hat{\bar{z}}_n = \sum_{i=1}^{l} a_{ni}\bar{q}_i$$

$$\bar{z}_n \longrightarrow \hat{\bar{z}}_n \simeq \bar{z}_n$$

Encoding	Decoding
$\bar{x}_n \longrightarrow \bar{a}_n$	$\bar{a}_n \longrightarrow \hat{\bar{x}}_n$
$d \qquad l$	$l \qquad d$

Auto-associative Neural Network (AANN)

AANN$_1$



Tanh in HL1 & HL3

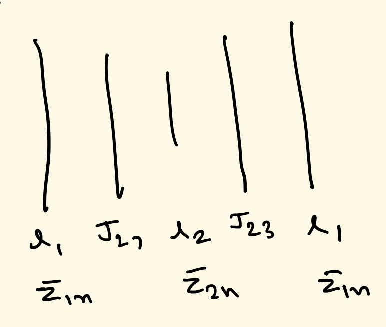
$$l_1 < d$$

Linear activation fn. $\phi(a) = a$

Training Dataset $\mathcal{D}_1 = \{\bar{x}_n\}_{n=1}^{N} \quad \bar{x}_n \in \mathbb{R}^d \quad \mathcal{D}_1 = \mathcal{D}_U$

$$\mathcal{D}_2 = \{\bar{z}_{1n}\}_{n=1}^{N} \quad \bar{z}_{1n} \in \mathbb{R}^{l_1}$$

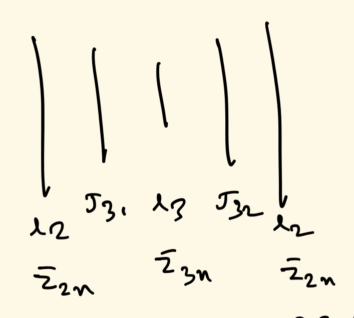
AANN$_2$



$$l_2 < l_1$$

$$\mathcal{D}_3 = \{\bar{z}_{2n}\}_{n=1}^{N} \quad \bar{z}_{2n} \in \mathbb{R}^{l_2}$$

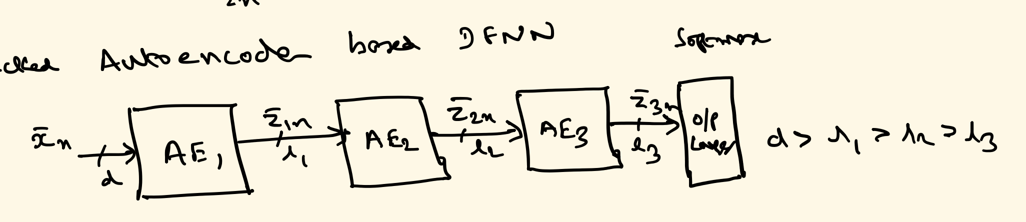
AANN$_3$



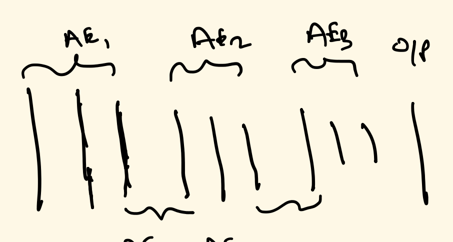
l from PCA

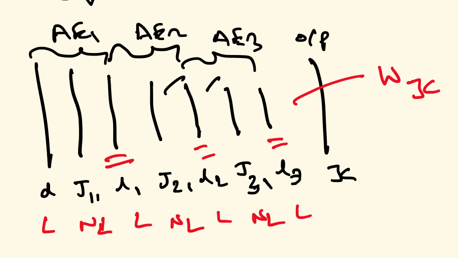
$l_3 < l$

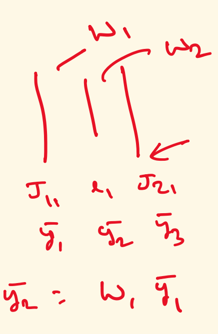
Stacked Autoencoder based DFNN



$$d > l_1 > l_2 > l_3$$







Merging

$\bar{y}_1$: output of HL$_1$ of AE$_1$

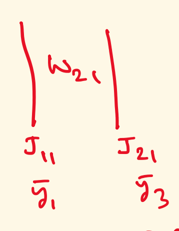
$$\bar{y}_2 = W_1\bar{y}_1$$

$$\bar{y}_3 = \bar{\Phi}(\bar{a}_3)$$

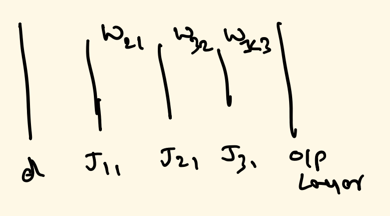
$$= \bar{\Phi}(W_2\bar{y}_2)$$

$$= \Phi(\underbrace{W_2 \cdot W_1}\bar{y}_1)$$

$$= \bar{\Phi}(W_{21}\bar{y}_1)$$



$$\bar{y}_3 = \bar{\Phi}(W_{21}\bar{y}_1)$$



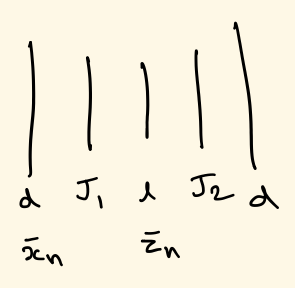
7-layer AANN

Stacked Autoencoder with 3 AEs

6 HLs after merging

Greedy layer-wise training

Contractive Autoencoder



$$x_{ni} + \Delta x_{ni} \longrightarrow z_{nk} + \Delta z_{nk}$$

$$\tilde{E} = \frac{1}{2}\sum_{n=1}^{|\mathcal{D}|}\|\bar{x}_n - \bar{y}_n\|^2 + \frac{\lambda}{2}\left\|\frac{d\bar{z}}{d\bar{x}}\right\|_F^2$$

M ; Jacobian matrix

Frobenius norm of matrix

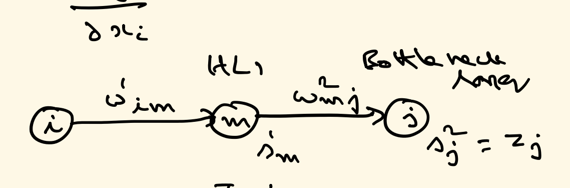
$$\|\bar{x}\|^2 = \bar{x}^t\bar{x} = \sum_{i=1}^{d} x_i^2$$

$$\underset{L\times K}{\|M\|^2} = \sum_{i=1}^{L}\sum_{j=1}^{K} m_{ij}^2$$

$$\underset{d\times l}{M} = \frac{d\bar{z}}{d\bar{x}} = [m_{ij}]_{i=1}^{d}\,{}_{j=1}^{l}$$

$$m_{ij} = \frac{\partial z_j}{\partial x_i}$$

$\dfrac{\partial z_j}{\partial x_i}$



$$\frac{\partial z_j}{\partial x_i} = \sum_{m=1}^{J} w_{im}\, w_{mj}\, (1 - s_m^{1\,2})$$

$$z_j = \sum_{m=1}^{J_1} w_{mj}\, s_m^1 + b_j$$

$$s_m^1 = \Phi\left(\sum_{i=1}^{d} w_{im}x_i + b_m\right)$$

$$\frac{\partial s_m^1}{\partial x_i} = w_{im}\underbrace{\frac{d\,\Phi(a_m^1)}{da_m^1}} = w_{im}\cdot(1 - s_m^{1\,2})$$

$$m_{ij} = \frac{\partial z_j}{\partial x_i} = \sum_{m=1}^{J_1} w_{mj}\, w_{im}\,(1 - s_m^{1\,2})$$

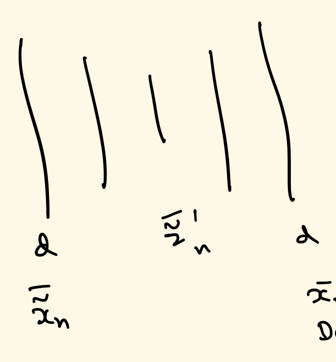
$$\|M\|_F^2 = \sum_{i=1}^{d}\sum_{j=1}^{l} m_{ij}^2$$

Image Denoising Autoencoder

Training dataset: $\mathcal{D} = \{\bar{x}_n, \tilde{\bar{x}}_n\}$

$\bar{x}_n$: Clean Image; $\tilde{\bar{x}}_n$: Noisy image

AANN$_1$



AANN$_1$	AANN$_2$	AANN$_3$
$\tilde{\bar{x}}_n \quad \bar{x}_n$	$\tilde{\bar{z}}_n^1 \quad \bar{z}_n^1$	$\tilde{\bar{z}}_n^2 \quad \bar{z}_n^2$

$\tilde{\bar{z}}_n^1$: Output of AE$_1$ when $\tilde{\bar{x}}_n$ is input to AANN$_1$ after training

$\bar{z}_n^1$: Output of AE$_1$ when $\bar{x}_n$ is input to AANN$_1$ after training