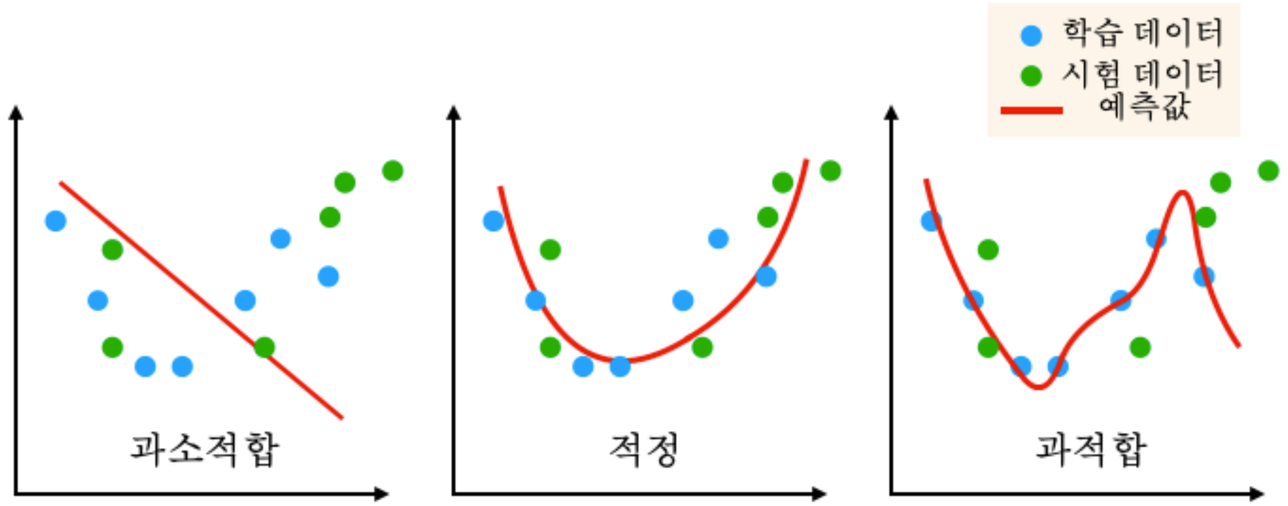
딥러닝 올인원

과적합 방지

21강

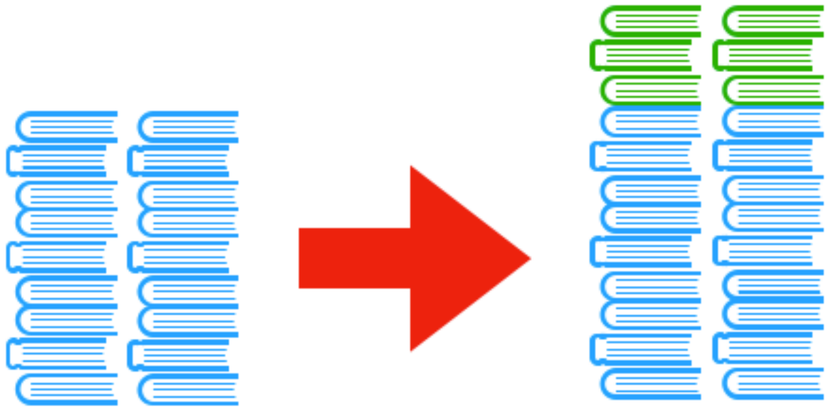
딥러닝호형

과적합(Overfitting)



과적합 해결방법

데이터 확보

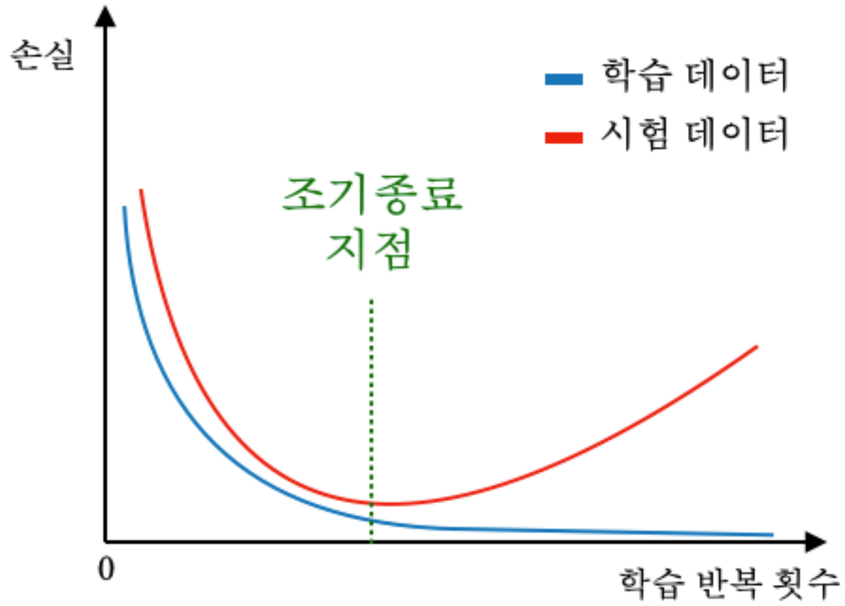


과적합 해결방법

데이터 증식(Data Augmentation)

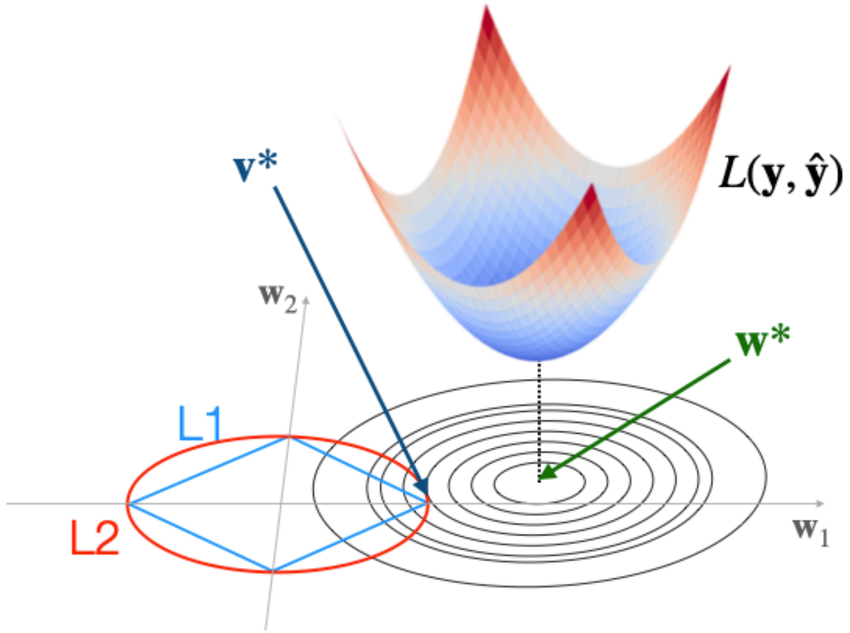
과적합 해결방법

조기 종료(Early Stopping)



과적합 해결방법

L1, L2 정규화(L1, L2 Regularization)

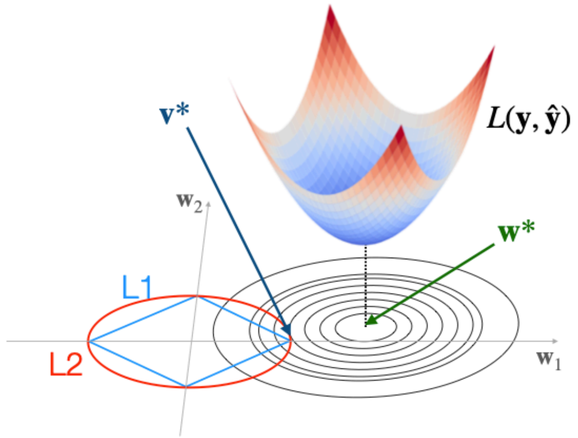


$$\tilde{L} = L(\mathbf{y}, \hat{\mathbf{y}}) + \lambda \Omega(\mathbf{w}), \lambda \geq 0$$

L1: $\Omega(\mathbf{w}) = \sum_{i=1}^{n} |w_i|$ L2: $\Omega(\mathbf{w}) = \sum_{i=1}^{n} w_i^2$

과적합 해결방법

L1, L2 정규화(L1, L2 Regularization)



$$\tilde{L} = L(\mathbf{y}, \hat{\mathbf{y}}) + \lambda \Omega(\mathbf{w}), \lambda \geq 0$$

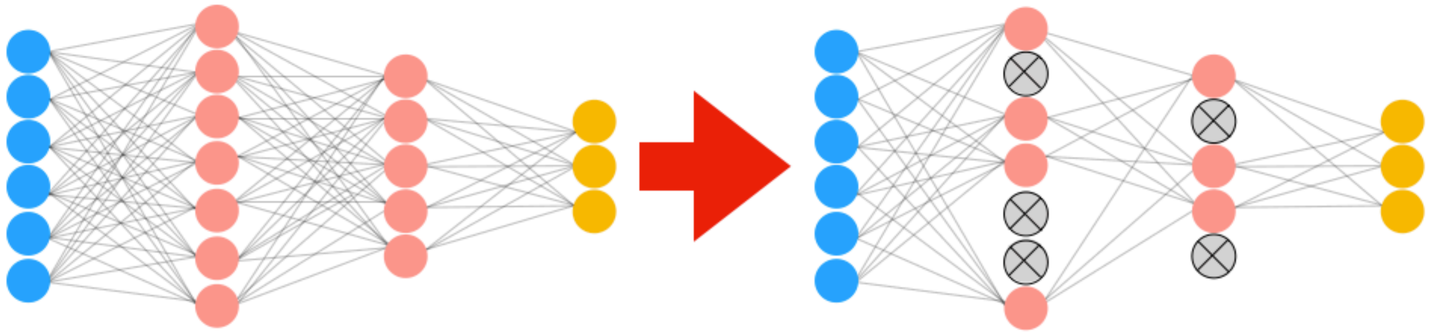
$$w \longleftarrow w - \mu \frac{\partial \tilde{L}}{\partial w} = w - \mu \left(\frac{\partial L}{\partial w} + \lambda \frac{\partial \Omega}{\partial w} \right)$$

L1: $$w \longleftarrow w - \mu \left(\frac{\partial L}{\partial w} + \lambda sign(w) \right), sign(w) = \begin{cases} 1 (\text{if } w \geq 0) \\ -1 (\text{if } w < 0) \end{cases}$$

L2: $$w \longleftarrow w - \mu \left(\frac{\partial L}{\partial w} + 2\lambda w \right) = (1 - 2\mu\lambda) w - \mu \frac{\partial L}{\partial w}$$

과적합 해결방법

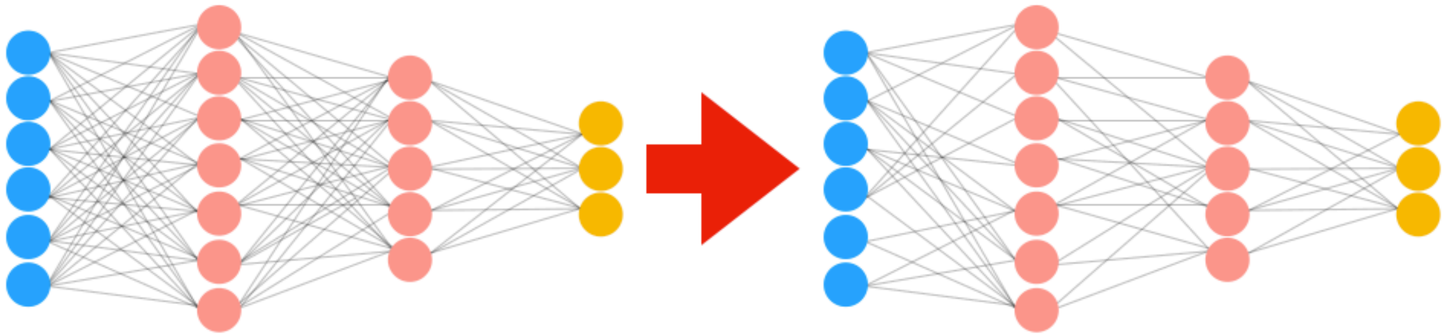
드롭아웃(Dropout)



Dropout 적용 전 Dropout 적용 후

과적합 해결방법

드롭커넥트(DropConnect)

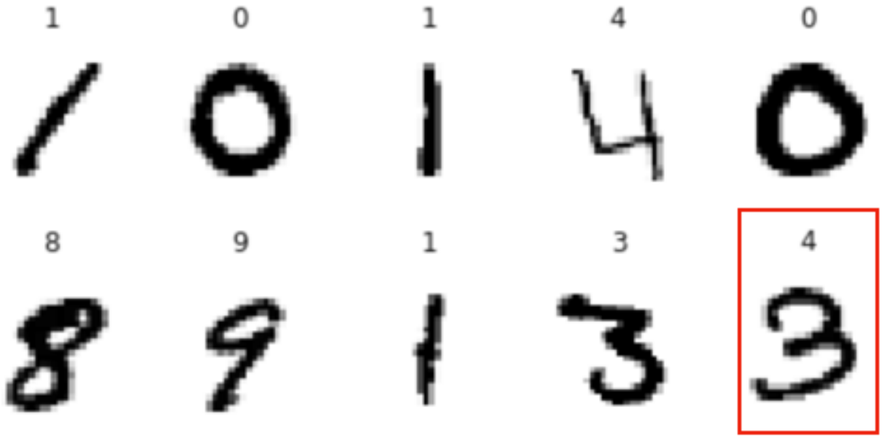


Drop Connect 적용 전

Drop Connect 적용 후

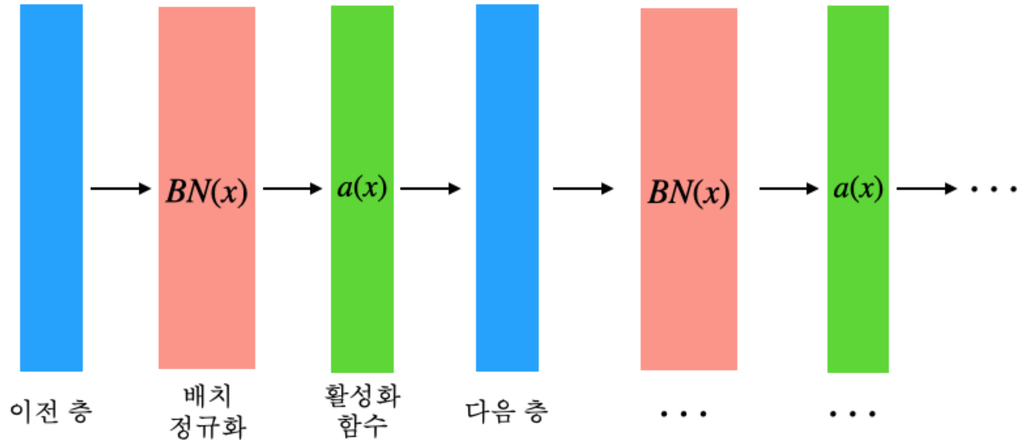
과적합 해결방법

교란라벨(DisturbLabel)



과적합 해결방법

배치 정규화(Batch Normalization)



학습 시: $\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{{\sigma_B}^2 + \varepsilon}}$, $\mu_B \leftarrow \frac{1}{n}\sum_{k=1}^{n} x_k$, ${\sigma_B}^2 \leftarrow \frac{1}{n}\sum_{k=1}^{n}(x_k - \mu_B)^2$

$$BN(x_i) = \gamma \hat{x}_i + \beta$$

평가 시: $\hat{x}_i \leftarrow \frac{x_i - \mu_{BN}}{\sqrt{{\sigma_{BN}}^2 + \varepsilon}}$, $\mu_{BN} \leftarrow \frac{1}{n}\sum_{k=1}^{n} \mu_B^k$, $\sigma_{BN}^2 \leftarrow \frac{1}{n}\sum_{k=1}^{n} {\sigma_B^k}^2$

$$BN(x_i) = \gamma \hat{x}_i + \beta$$

과적합 해결방법

라벨 스무딩(Label Smoothing)

$$y_{ls} = (1-\alpha)\, y + \frac{\alpha}{K}$$

(K 는 클래스 수, α는 스무딩 비율, y는 0 또는 1)

$\alpha = 0.1, K = 3 \qquad y_{ls} = 0.9y + 0.03 \qquad (1,0,0) \rightarrow (0.93,0.03,0.03)$

$\alpha = 0.3, K = 5 \qquad y_{ls} = 0.7y + 0.06 \qquad (0,0,1,0,0) \rightarrow (0.06,0.06,0.76,0.06,0.06)$

$\hat{y} = (0.1,0.1,0.76,0.01,0.03), y = (0,0,1,0,0)$

$\hat{y} = (0.1,0.1,0.76,0.01,0.03), y = (0.06,0.06,0.76,0.06,0.06)$

과적합 해결방법

노이즈 주입(Noise Injection)

- 입력값 자체나 레이어 중간의 임의의 노이즈를 주입

$$\tilde{x} = x + \alpha, \tilde{x} = \alpha x + \beta$$

과적합 해결방법

교차 검증(Cross Validation)