딥러닝 올인원

확률적 경사 하강법
10강

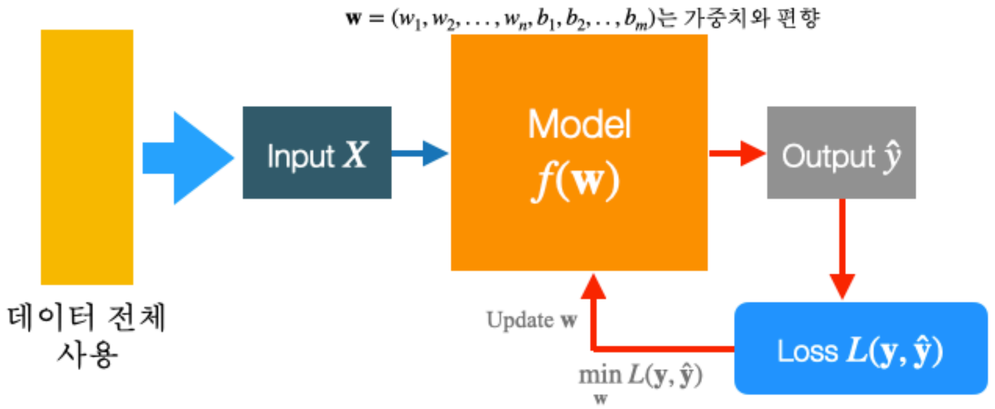
딥러닝호형

확률적 경사 하강법

SGD(Stochastic Gradient Descent)

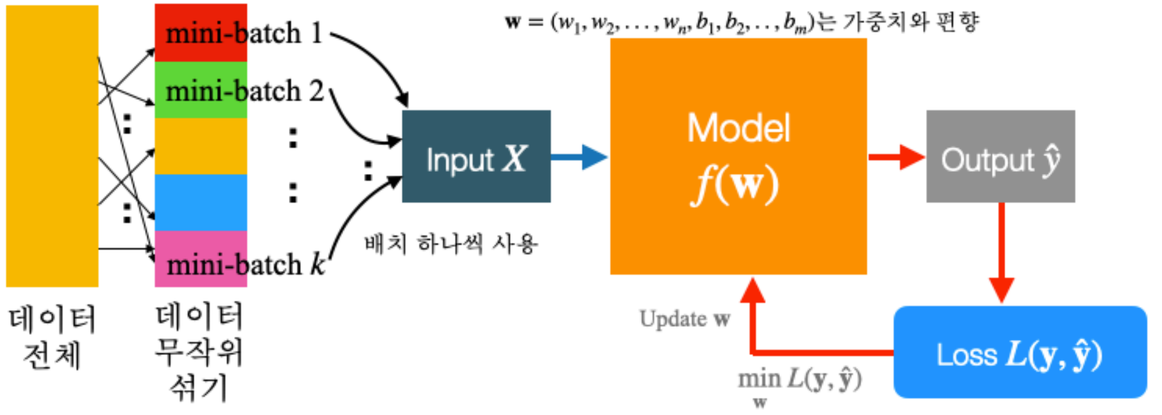
경사하강법

$$w \leftarrow w - \mu \nabla L(w)$$

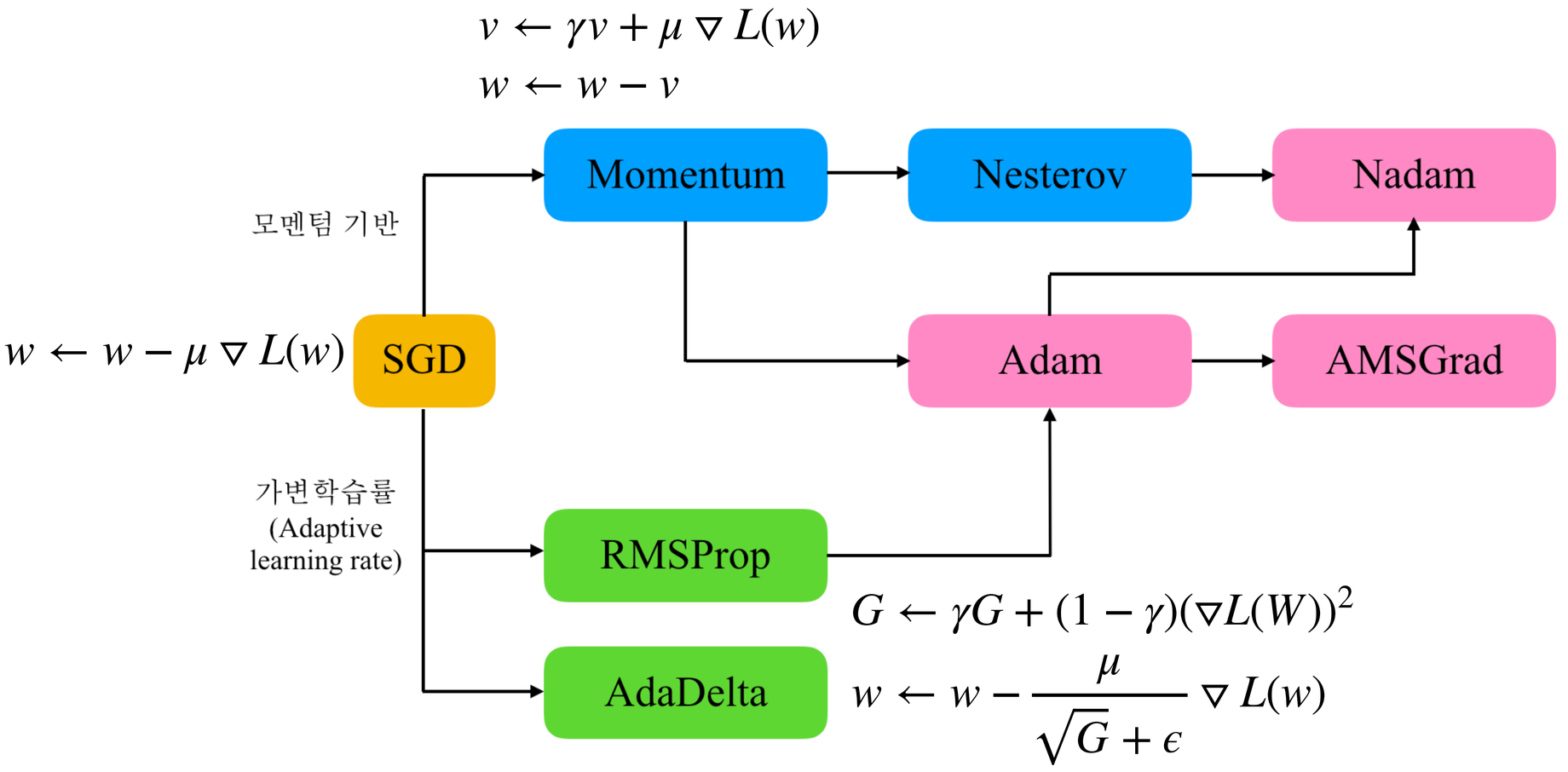


SGD

$$w \leftarrow w - \mu \nabla L(w)$$



여러 가지 최적화 방법



여러 가지 최적화 방법

ADAM(Adaptive Moment Estimation)

RMSProp와 Momentum 개념을 합침

$$m \leftarrow \beta_1 m + (1 - \beta_1) \nabla L(W)$$

$$v \leftarrow \beta_2 v + (1 - \beta_2)(\nabla L(W))^2$$

$$\hat{m} = \frac{m}{1 - \beta_1^t} \quad \hat{v} = \frac{v}{1 - \beta_2^t}$$

$$w \leftarrow w - \frac{\mu}{\sqrt{\hat{v}} + \epsilon} \hat{m}$$

여러 가지 최적화 방법

스케줄링 - learning rate을 조절하자!

- StepLR
- ExponentialLR
- Cosine Annealing 등등