PHY654

Machine Learning in Particle Physics

Saturday, 19 August, 2024

## Early Stopping:

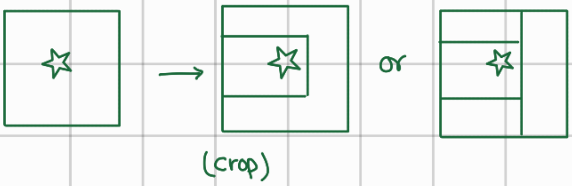
Regularization

(L2, drop-out)

niter = 180, 200


MNIST

Data Augmentation - I




## Solving ODE via NN

$$\frac{dg(x)}{dx} = g'(x) = -\gamma g(x) \;;\; g(0) = g_0$$

Assume: $\gamma = 2, \; g_0 = 10$

Analytical solution: $\boxed{g(x) = g_0 e^{-\gamma x}} = 10 e^{-2x}$

Trial Solution:

Neural network

$$g_t(x) = g_t(x, p) = h_1(x) + h_2(x, N(x,p)) = g_0 + x N(x,p)$$

$p \to w, b$

$$\boxed{g_t'(x) = -\gamma g_t(x)}$$

Cost function: $(LHS - RHS)^2 \to$ minimize

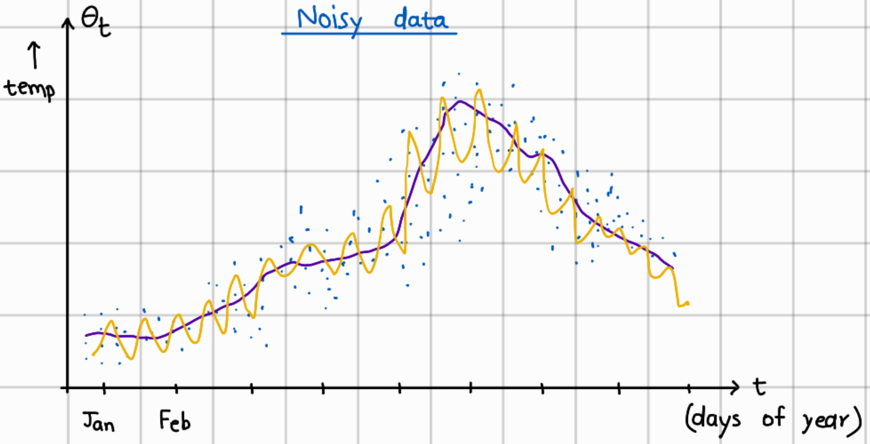
For $x_i \;;\; i = 1, 2, \dots N$

$$\min_{p} \left\{ \sum_{i=1}^{N} \left[ g_t'(x,p) - [-\gamma g_t(x,p)] \right]^2 \right\}$$

## Adam Optimization

[TensorFlow (keras)] → Next Class.

Exponentially weighted average




Initialize : $V_0 = 0$

day 1 — $V_1 = 0.9 V_0 + 0.1\ \theta_1$

day 2 — $V_2 = 0.9 V_1 + 0.1\ \theta_2$

...

for day t — $\boxed{V_t = \beta V_{t-1} + (1-\beta)\theta_t}$

$V_t$ approximately avg over $t = \frac{1}{1-\beta}$ days

$\beta = 0.9 \rightarrow t = 10$

$\beta = 0.98 \rightarrow t = 50$

→ peak will shift to the right

$$V_{100} = 0.9\ V_{99} + 0.1\ \theta_{100}$$

$$V_{99} = 0.9\ V_{98} + 0.1\ \theta_{99}$$

$$V_{98} = 0.9\ V_{97} + 0.1\ \theta_{98}$$

$$V_{100} = 0.1\ \theta_{100} + 0.9(0.1\theta_{99} + 0.9(0.1\ \theta_{98} + 0.9(0.1\ \theta_{97} + \cdots)))$$

$$= 0.1\ \theta_{100} + (0.1)(0.9)^1\theta_{99} + (0.1)(0.9)^2\theta_{98} + (0.1)(0.9)^3\theta_{97} + \cdots$$

→ weighted sum