

- Redefine the model to be  $w_2 * t_u^{** 2} + w_1 * t_u + b$ .

Optimizer은 SGD를 사용함을 가정한다.

a) What parts of the training loop, and so on, need to change to accommodate this redefinition?

화씨(Fahrenheit) = (섭씨(celsius) X 1.8) + 32 인 일차 공식은 이미 알려진 output 정보이다. 그래서 모델을 2차 계산식으로 변환 후, 학습 epoch를 진행하면  $w_2$ 의 값은 점점 작아져서 0이 될 것이다. 이때 learning rate이 크다면  $w_2$ 의 변화량이 oscillating할 수 있다. 따라서 이를 방지하고자 learning rate 인 lr을 줄이는 게 좋다. Input이 하나에서 둘로 증가하고, lr을 줄인 상태이므로  $w_2 = 0$ ,  $w_1$ 는 원하는 결과값을 얻기 위해서는 epoch를 더 증가해야 할 것이다.

b) What parts are agnostic to swapping out the model?

전체적인 계산방식인 forward와 backward, objective function과 activate function은 변하지 않는다. Test set과 validation set에 관한 Batch의 대중적인 비율도 유지하여 학습을 진행한다.

c) Is the resulting loss higher or lower after training?

충분한 학습을 하면 loss는 일차식 model일 때와 같은 값으로 수렴할 것이다. 그러나  $w_2 = 0$ 이됨을 학습하는 불필요한 계산과정이 추가되었으므로, 더 좋은 model이라고 할 수 없다.

d) Is the actual result better or worse?

결과적으로  $w_2 = 0$ 이됨을 학습하는 불필요한 계산과정이 추가되었으므로, 더 좋은 model이라고 할 수 없다.