machine learning is what?

looking for a function

types of task

regression: use data find scalar (The function)

Classification: Give options find correct one

Structure learning: generate structured result (image, article etc)

How to find a function

① write a function with unknow Parameter (guess

is with domain knowledge

def model: a function with unknown parameter

$y = b + wx_1$

feature

weight

bias

② def loss Loss of a function $L(b,w)$

✓ how good a set of value is

from train material $x_i \rightarrow f(x_i)$

$|\hat{y} - y_1| = e_1 \rightarrow |\hat{y} - y_i| = e_i$

$\hat{y}$ | label : the real number

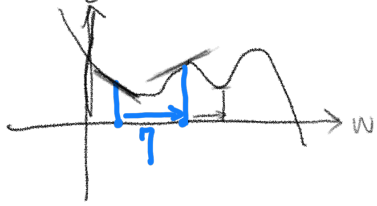
Loss $L = \frac{1}{N} \sum_n e_n$ $\begin{cases} e = |y - \hat{y}| & \text{mean absolute error} \\ e = (y - \hat{y})^2 & \text{mean square error} \end{cases}$

If y $\hat{y}$ are probability distributions $\Longrightarrow$ Cross entropy

③ optimization

$w^* \; b^* = \arg \min_{w,b} L$ ⓞ mute to only 1 variable

gradient decent ① random pick initial w^0



② Compute $\frac{\partial L}{\partial w}\Big|_{w=w_0}$

η is learn rate : the rate of update of data

↑ hyper parameter self define parameter

$w' \leftarrow w^0 - \eta \frac{\partial L}{\partial w}\Big|_{w=w_0}$

stop

set times hyperparameter

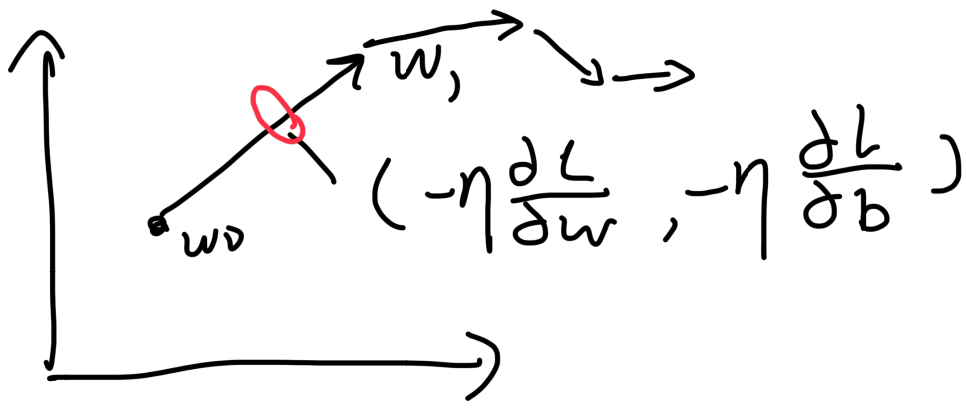
$\frac{\partial L}{\partial w} = 0 \rightarrow$ local extrema

gradient decent of more than 1 unknown p

① initial w^0 b^0

② $\left.\frac{\partial L}{\partial w}\right|_{w=w^0, b=b^0}$ $w^1 \leftarrow w^0 - \eta \left.\frac{\partial L}{\partial w}\right|_{w=w^0, b=b^0}$

$\left.\frac{\partial L}{\partial b}\right|_{w=w^0, b=b^0}$ $b^1 \leftarrow b^0 - \eta \left.\frac{\partial L}{\partial b}\right|_{w=w_0, b=b_0}$



Step ① ② ③ is called Training (with current data)

④ forecast

⑤ use your domain knowledge to modify