lattitude
90°
0
-90°
-30°C
0
40°C
aver temp

where can I ski?

⇒ cannot use a simple linear model for this kinda data...
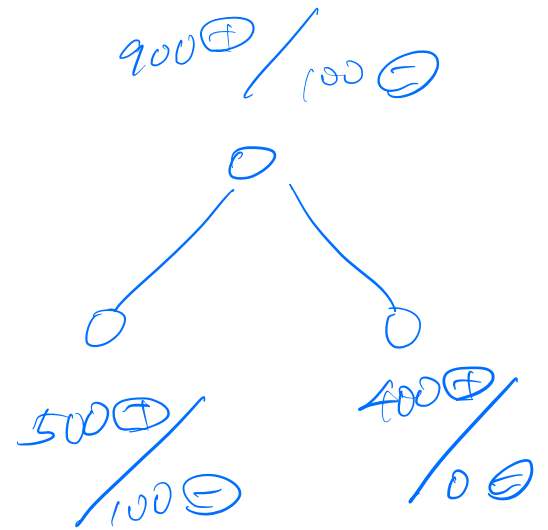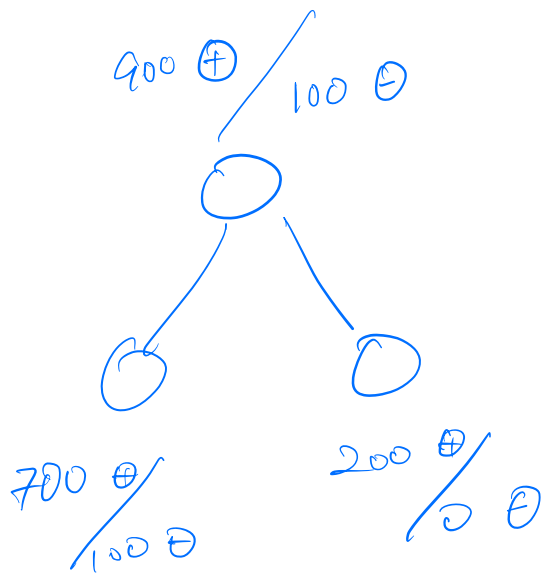
    ⌐→ decision tree

training not that slow (faster than neural network)
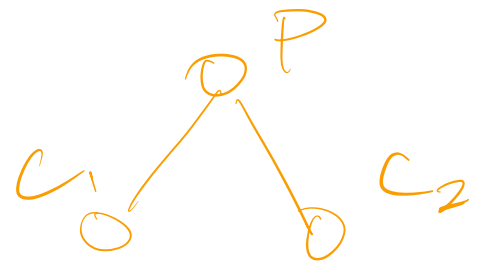
△ how do select features on node & value (page. 2)

900 ⊕ / 100 ⊖

900 ⊕ / 100 ⊖

700 ⊖ / 100 ⊖

200 ⊖ / 0 ⊖
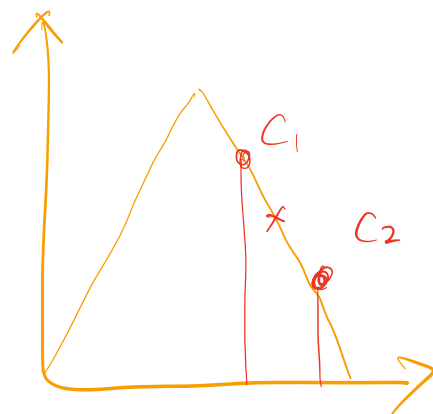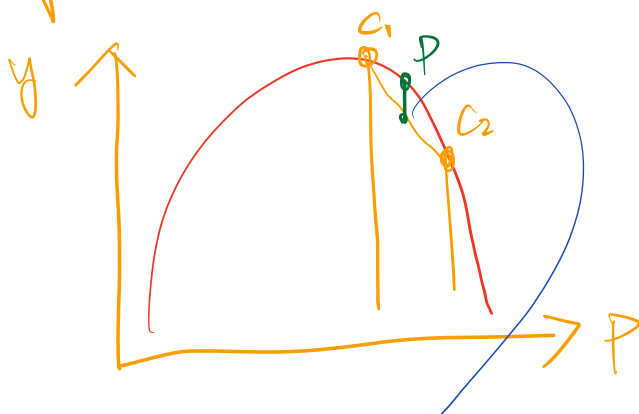
500 ⊕ / 100 ⊖

400 ⊕ / 0 ⊖

bad partition

⟹ What kind of metric to be designed for better partition

maximum entropy

$$H(s) = -\sum_{C=1}^{C} P(c) \log P(c)$$

P

C_1 C_2

Why we use entropy?

C_1 P C_2

y → P

C_1 C_2

1. information gain
2. Gini $1 - \Sigma p^2$ $\Big\}$ entropy low
to judge whether the data is a good split or not

We can aslo use regression tree

$f_1, f_2, \ldots f_T$

$\sum_{i=1}^{T} W_i \cdot f_i$: Bagging

P18 : usually : give probability value.

Difference between Boosting & Bagging

Boosting.

$f_1 \quad f_2 \quad f_3 \quad ---- \quad f_T$

fix the error from f_1

Adaboost
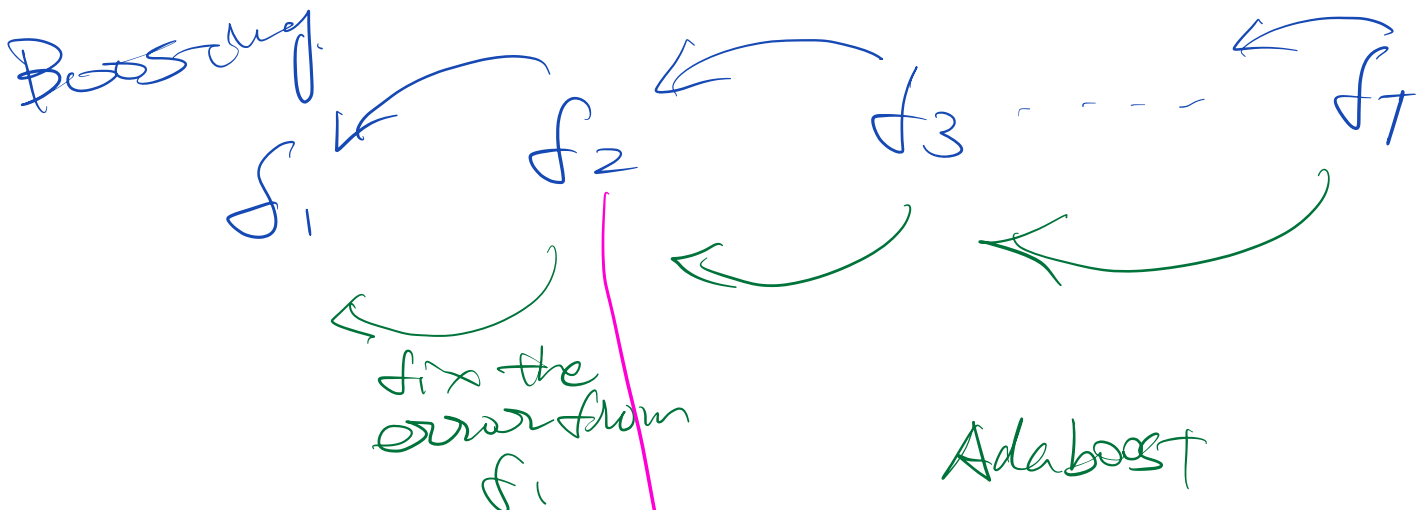
model to predict the error, cannot use independaty.