Supervised

Training (x_i, y_i)

x_i → Image

y_i → label

100K → {male, female} → Discrete classification

y_i is aesthetic rating [1, 5] → Continuous Regression

Unsupervised

Only know x_i

x_i → 50-d gene signature for 1000 pat.

Can we split into clusters discrete

50-d → 3-din

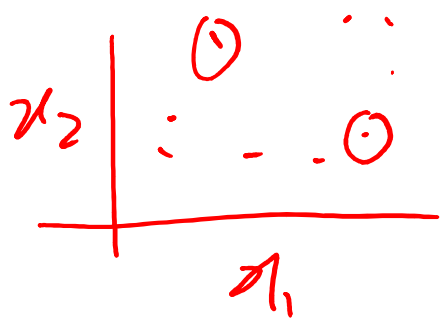
Sup ⟵ ⟶ Unsup

Semi - Supervised

Weakly Supervised

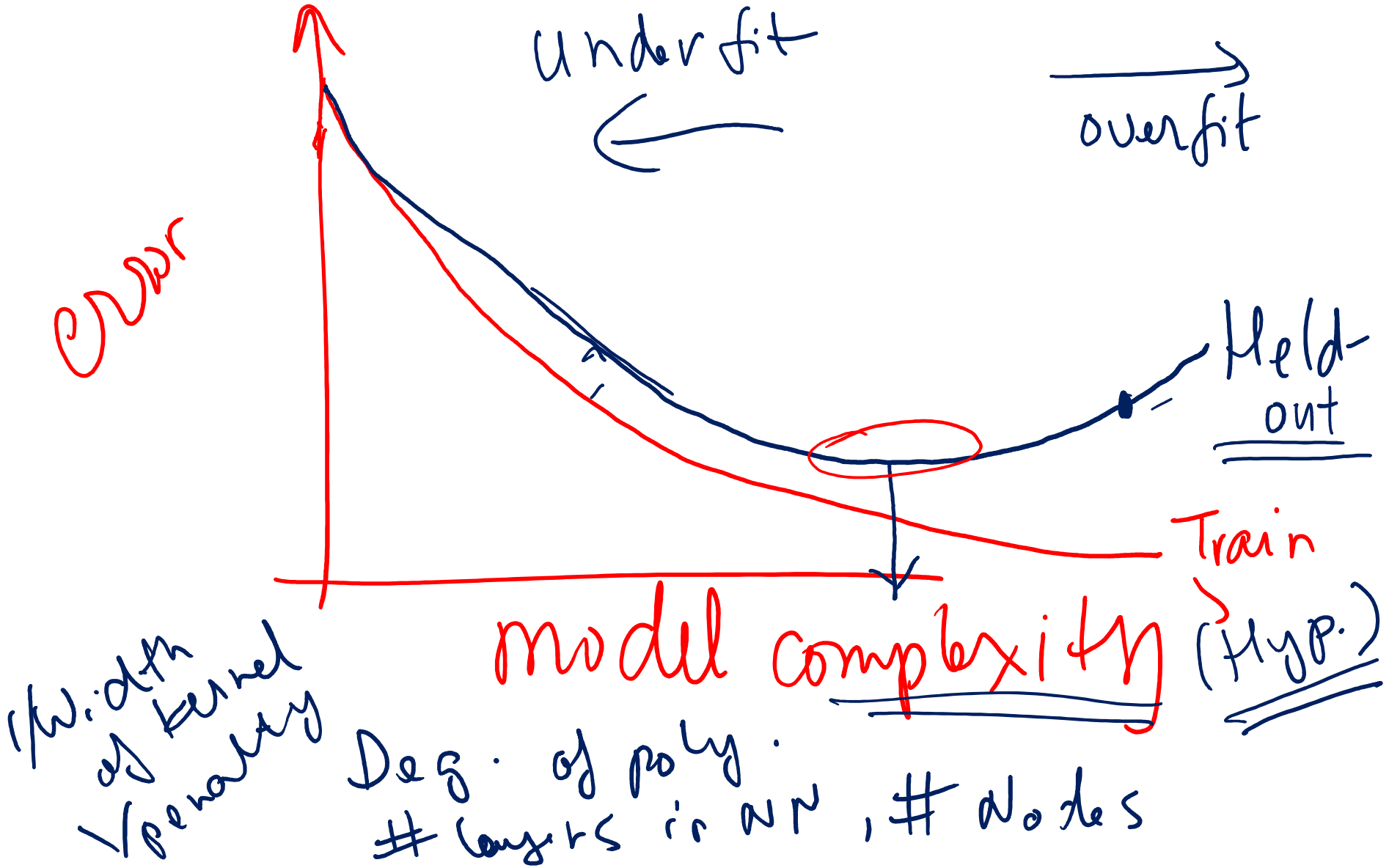
Self- Supervised

R. L. 30k



Some samples have known y_i

There are ups of that samples you something nbout


Underfit
overfit
Error
Held-out
Train
model complexity (Hyp.)
(Width of Kernel Penalty)
Deg. of poly.
layers in NN, # Nodes

Grid Search (Hyp. tuning)

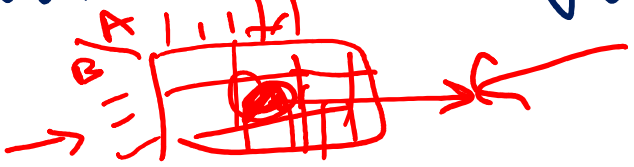
Prepare a set of hyp. values

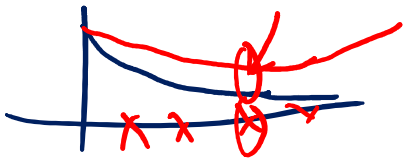
→ $\lambda = \{0.01, 0.1, 1, 10, 100\}$

For each setting in the set ↑

- 70% Train parameters on trg. set
- 30% Test performance on validation set

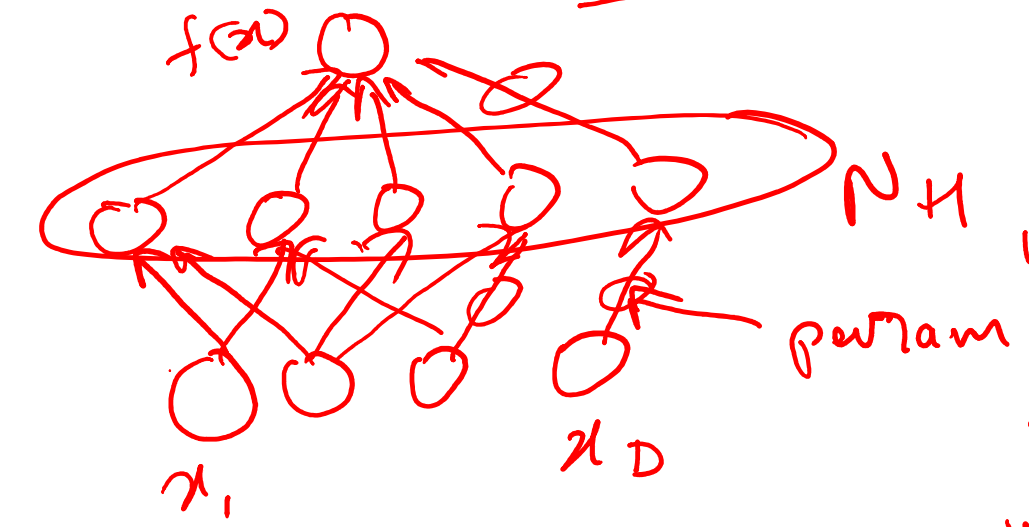
Pick the hyp. w/ lowest val. error

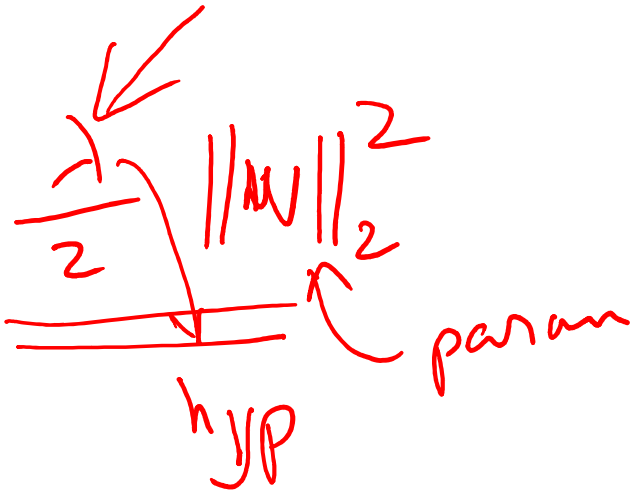




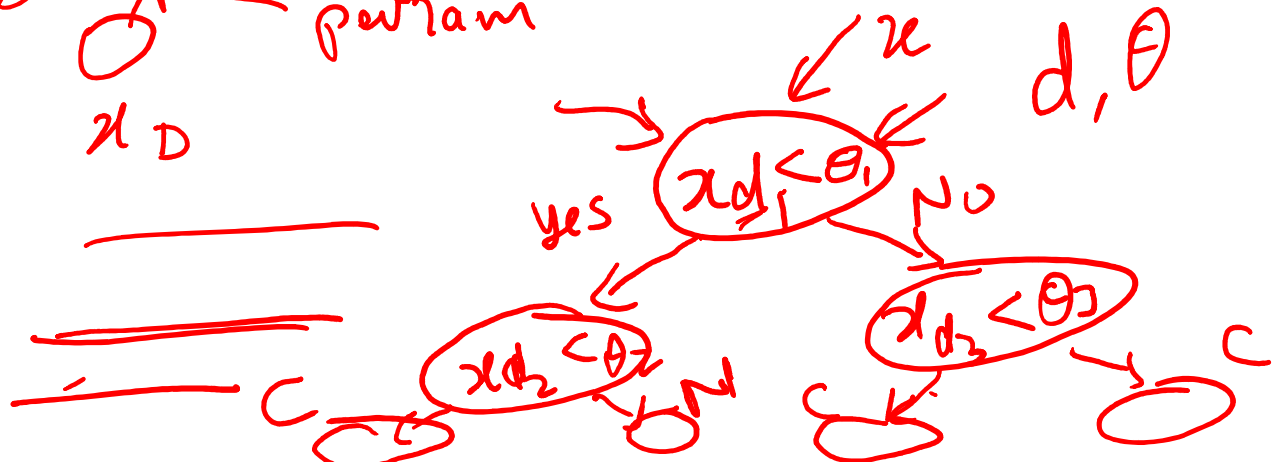
$$f(x) = v_0 \sum_{d=1}^{D} w_d x_d$$

$$\rightarrow \sum_{i=1}^{N} (y_i - f(x_i))^2 + \frac{1}{2} \|W\|_2^2$$





is hyper param.



50-d ⟶ {0, 1}

#tree depth → [R.F.]

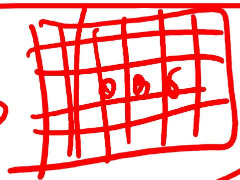
kernels, Spenalty, width → [S.V.M.]

#layers, #nodes → N.N.

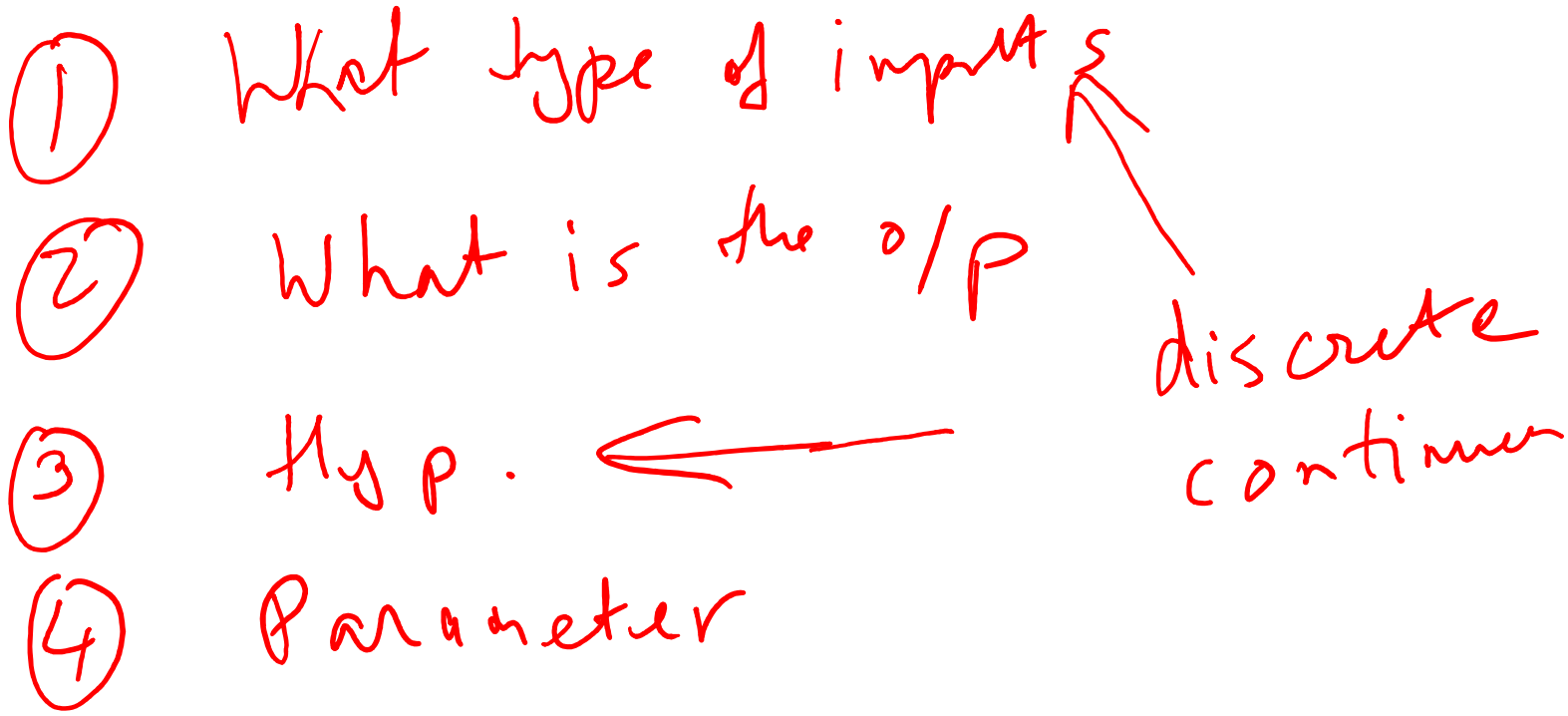
XGBOOST

which is the best?

Lots of data

Except D.L. →  → The cat jumps

Graphical


1 What type of inputs
2 What is the o/p
discrete
continuous
3 Hyp.
4 Parameter