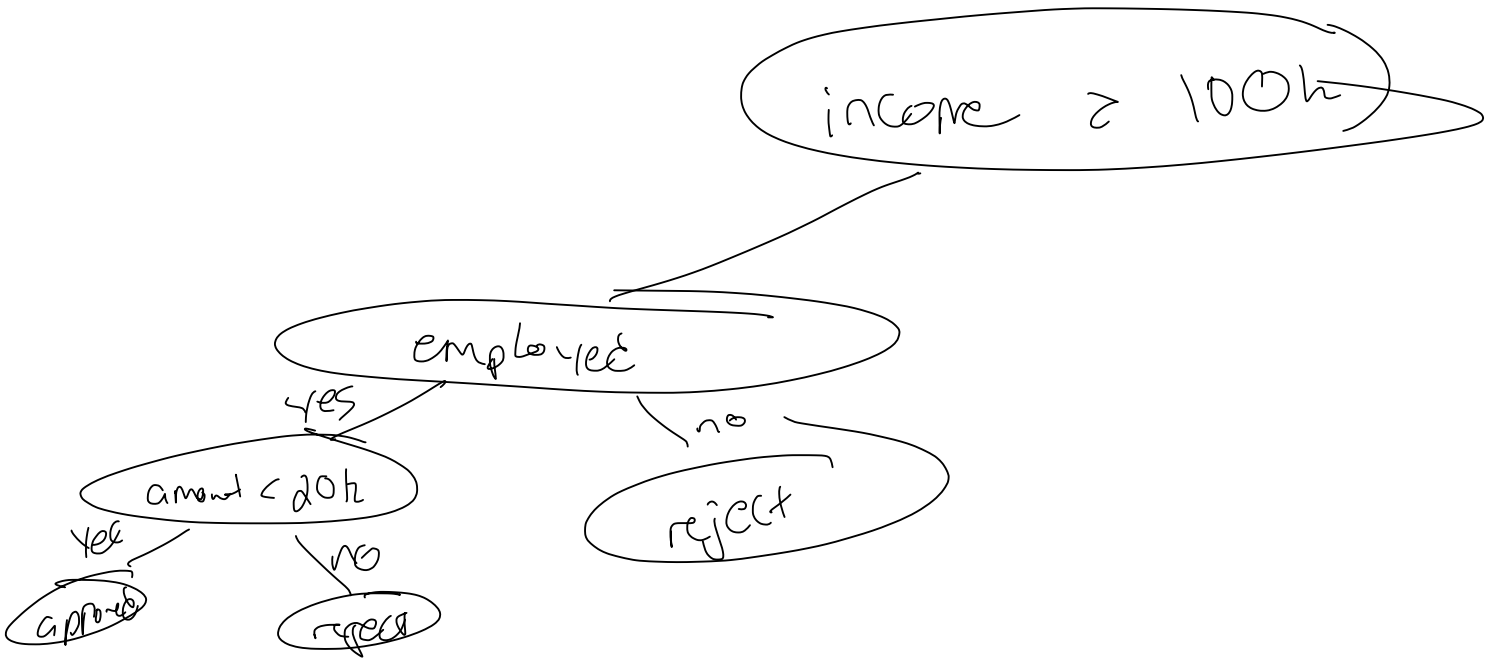
employment	income	requested amount	approved?
un employed	0	100k	rejected
employed	50k	10k	approved
ure employed	0	1k	approved
employed	100k	100k	rejected



Choosing what to split on:

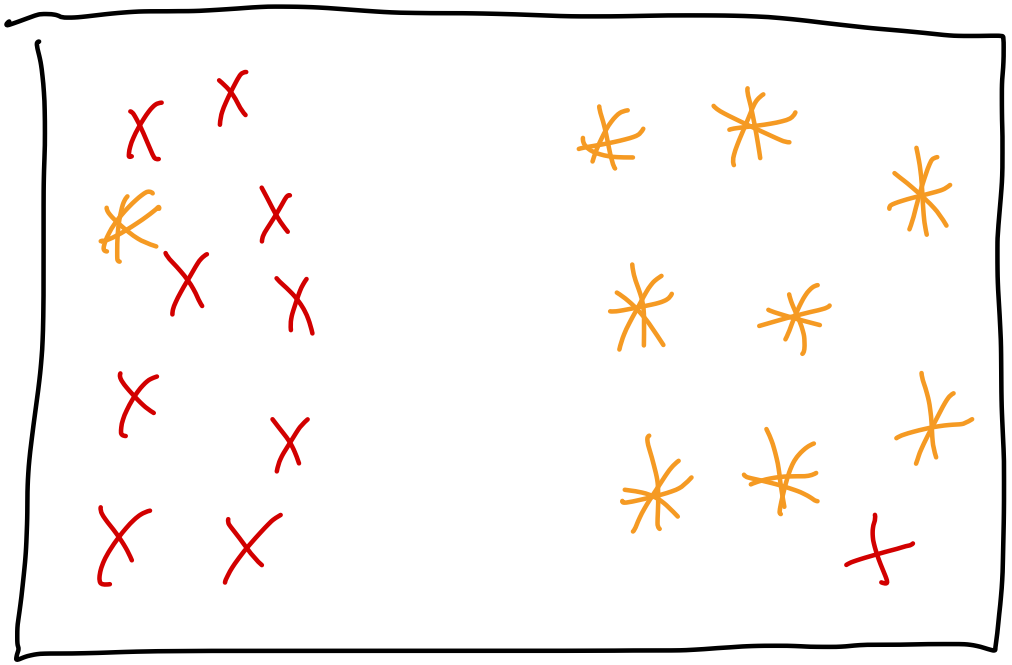
1) Consider all and choose best

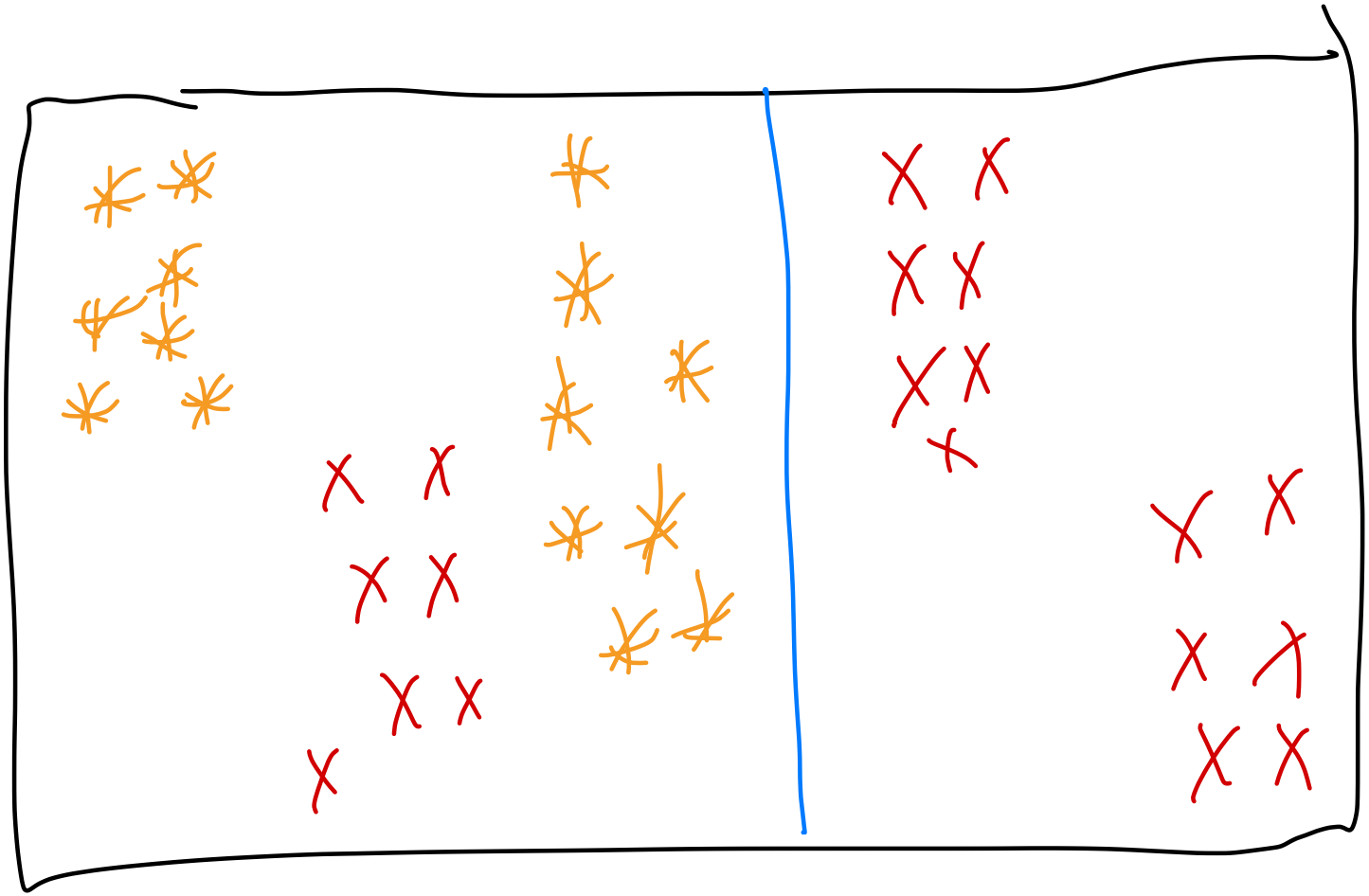
2) Choose one at random

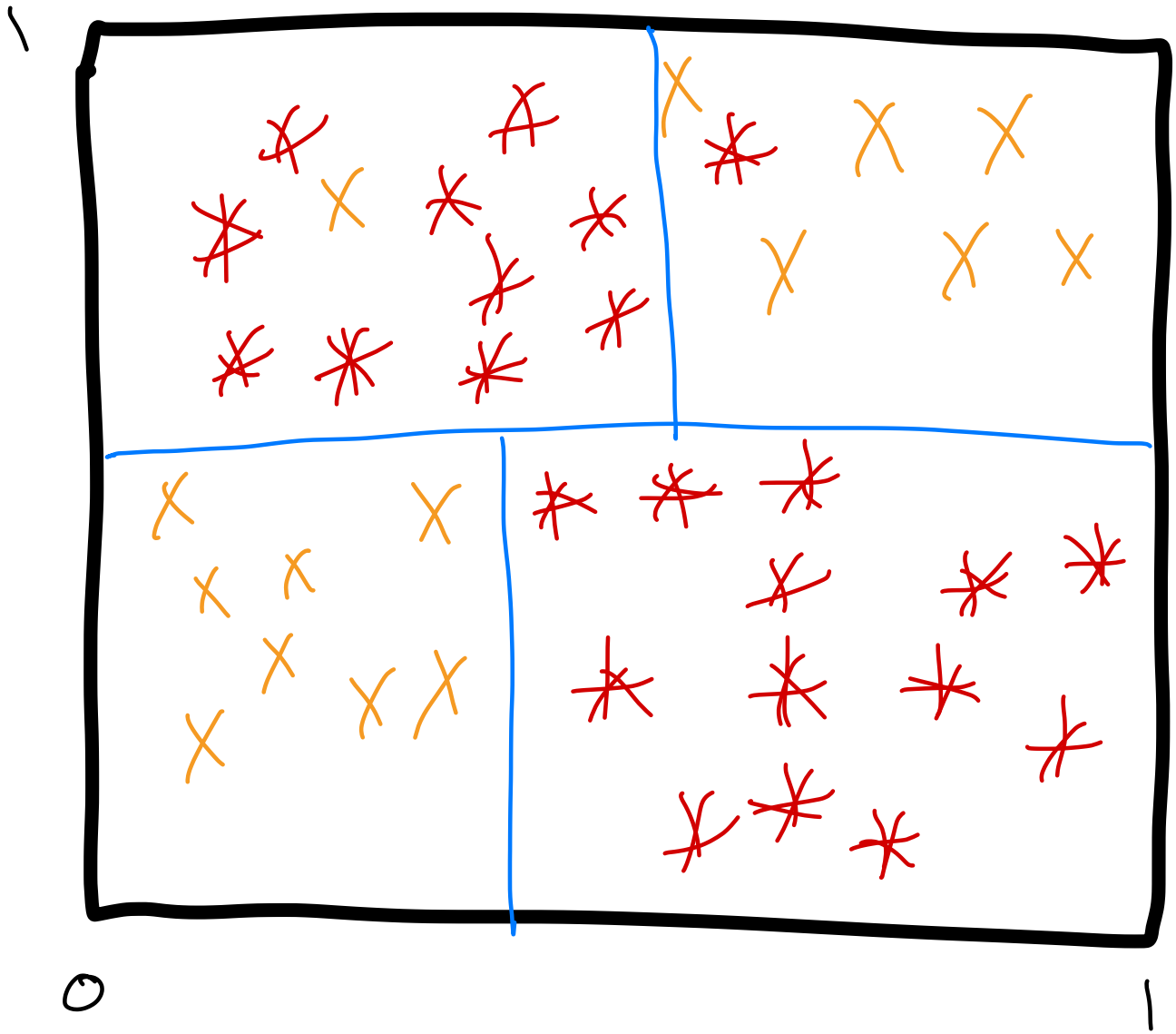
3) Randomly choose a subset of features
   → choose best from subset

Choosing threshold:

- want most informative






1
0
1

Entropy → means the purity of a collection

↳ on average, # of bits required to represent the class item

Formula

pool of items D

$N(D)$ = # of items in pool D

$N(i, D)$ = # of items with class i in D

Entropy, $H(D)$

$$H(D) = -\sum_i \left[\frac{N(i,D)}{N(D)} \log_a \frac{N(i,D)}{N(D)} \right]$$

Let P be all data at node

P_e be left pool

P_r be the right pool

$$\frac{N(P_e)}{N(P)} H(P_e) + \frac{N(P_r)}{N(P)} H(P_r)$$

↳ ~~#of~~ bits to classify if we split into P_l and P_r

information gain

$$I(P_e, P_r, P) = H(P) - \left[\frac{N(P_e)}{N(P)} H(P_e) + \frac{N(P_r)}{N(P)} H(P_r) \right]$$