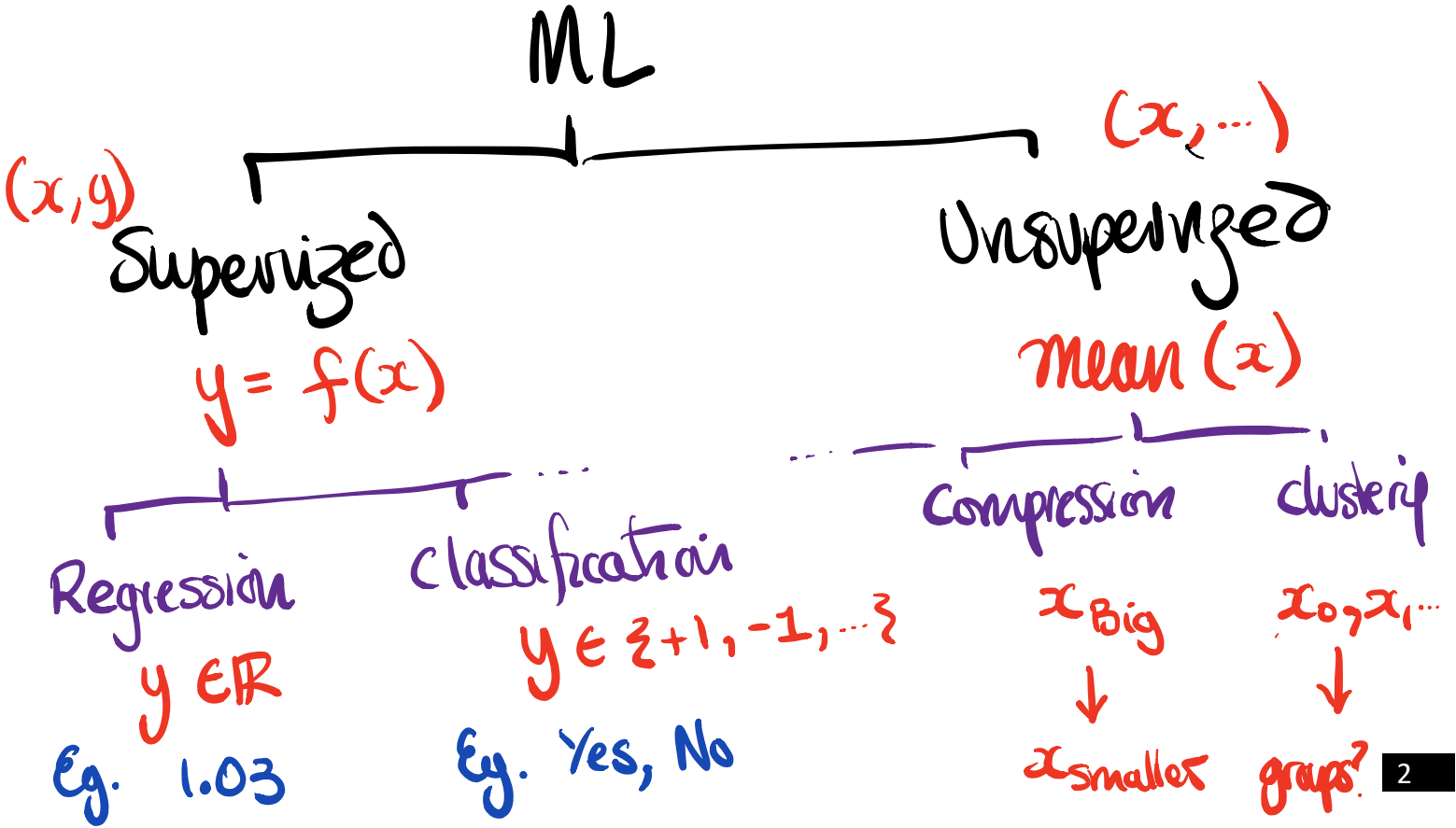
Introduction to Machine learning

How do Machines "learn"?

Review: Map of ML



Supervised ML

$\mathcal{D} = (x, y)$ ← Actual historical cases

FIND: $\hat{y} = \hat{f}(x)$ ← Estimate

$\hat{y} - y$ ← Error

USE:

Loss (Error) on $\mathcal{D}$ ← Hold (x, y) fixed!

x	y	$\hat{y}$	ℓ
3	2	2.1	0.1

} change $\hat{f}$ until ℓ small

Aside:

MODEL $\hat{y} = \hat{f}(x) = ax + b$ — a, b: constant; x: variable

$= \hat{f}(x; a, b)$ — x: variable; a, b: Parameters ≡ Constants

LOSS: Bespoke formula, Error per Point

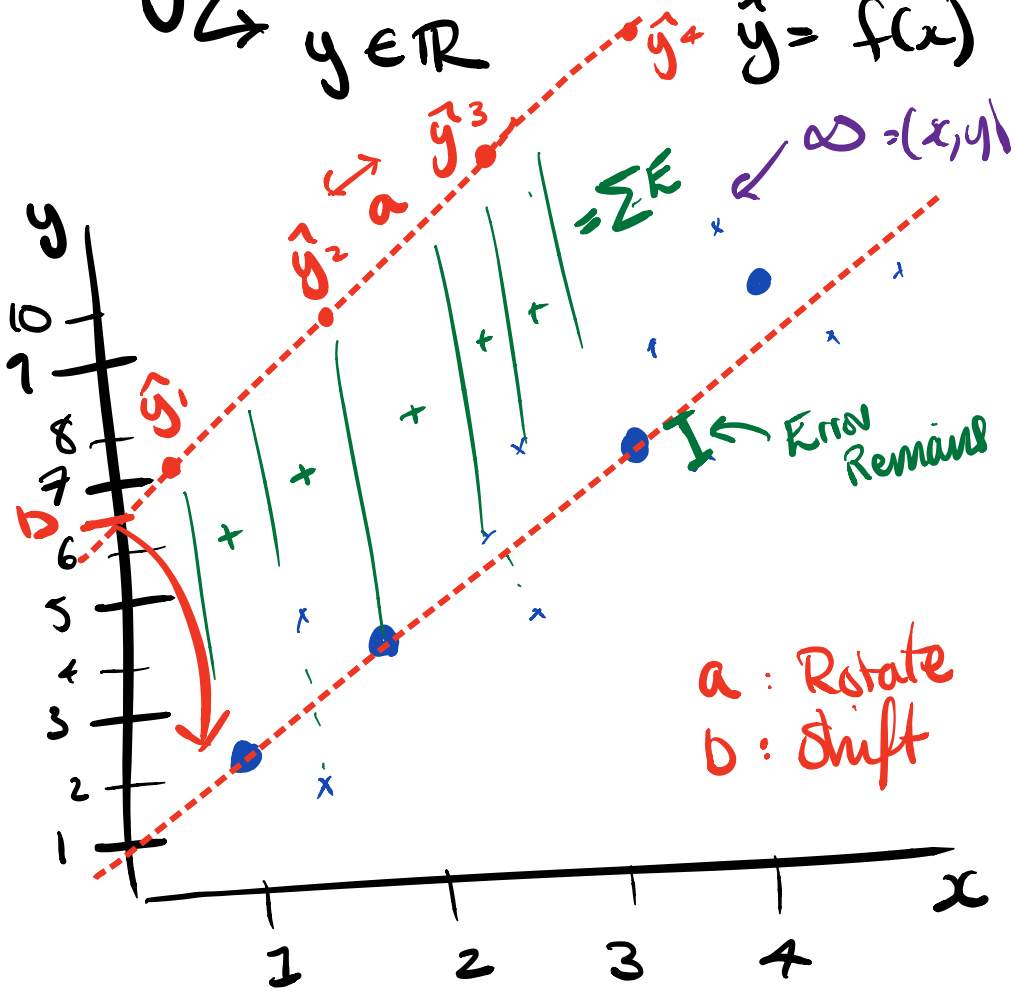
$$\ell(\hat{y}, y) = (\underbrace{\hat{y}}_{\hat{f}(x)} - y)^2 = \left[\underbrace{(ax + b)}_{\hat{f}} - y\right]^2$$

y: fixed from ∞

$\ell(a, b; x, y)$

Regression : linear

↳ $y \in \mathbb{R}$ $\hat{y} = f(x)$



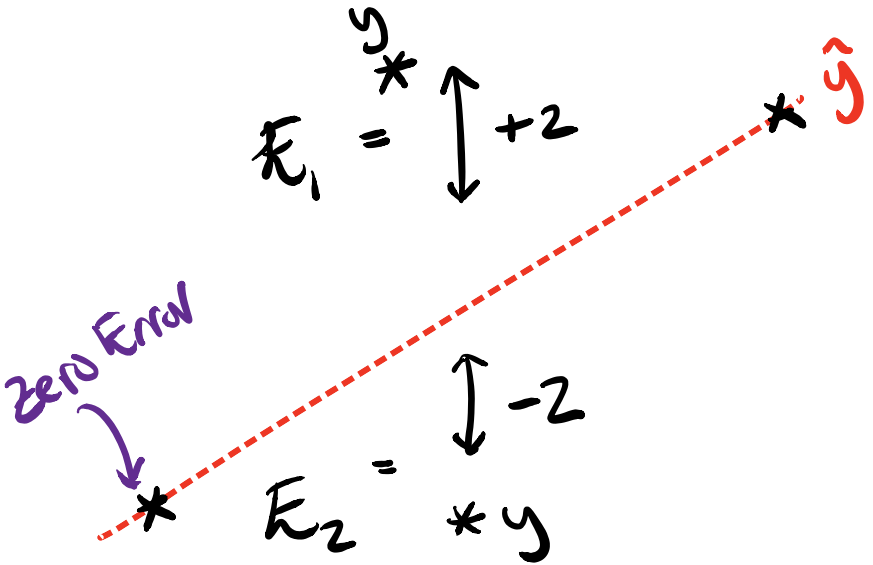
House Prices

Years $

x	y	$\hat{y}$	E
1	2.1		
2	3.3		
3	7.0		
4	9.5		

$Error = \hat{y} - y$

Why $l(\hat{y}-y) = (\hat{y}-y)^2$?



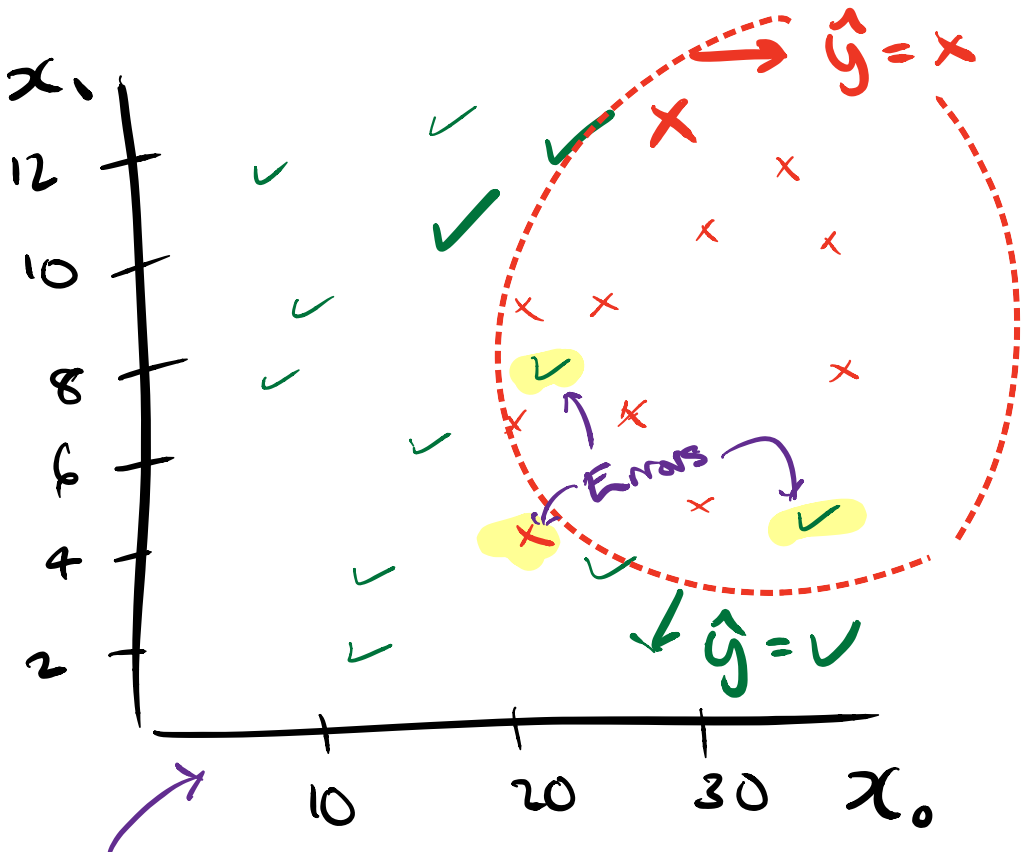
WRONG ANSWER!

$\sum E = E_1 + E_2$

$= 0$

$(-2)^2 + (2)^2 = 4 + 4 = 8$

Classification



$$f(x) = a(x_0 + x_1 - r)^2$$

Age (x_0)	Ticket (x_1)	Like? (y)	$\hat{y}$
18	10	✓	
30	12	x	
25	11	✓	

Axis shows $\mathbb{R}$ No., Eg. x

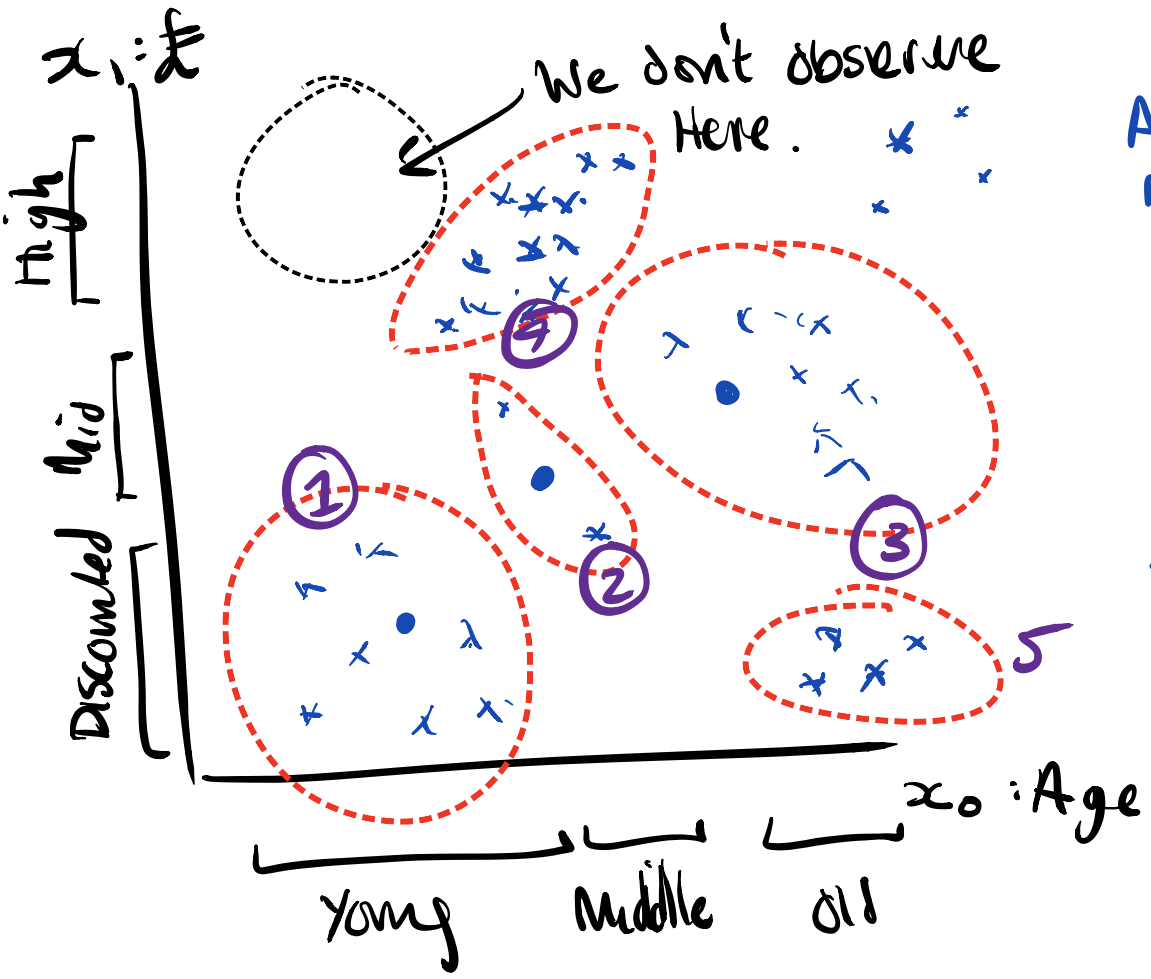
Unsupervised learning

- Compression
  - aka. "Dimensionality Reduction"
- Clustering
  - aka. "Grouping"

Clustering

① Machine Draws Group Boundaries
② → ② We Interpret

Add ③

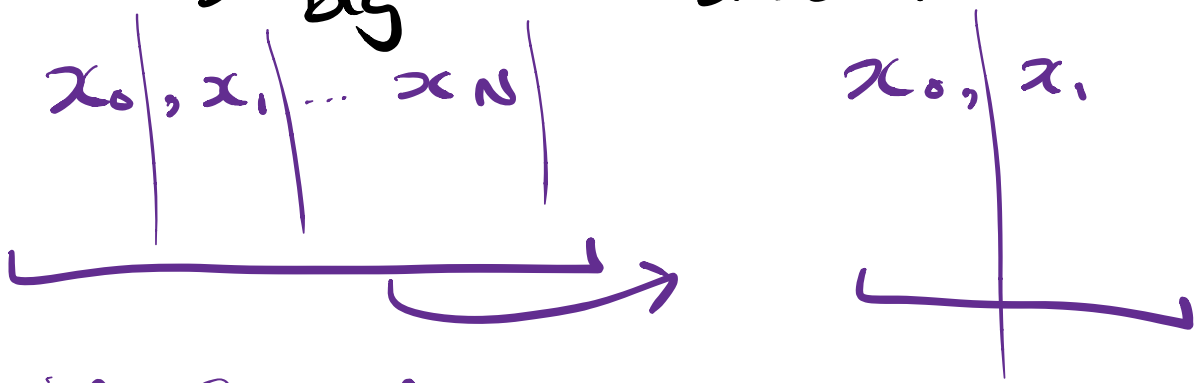


Age (x_0)	Ticket (x_1)	~~UJI~~
18	10	3
30	12	2
25	11	3

1: Children
2: Unemployed
3: Employed
4: High Kanes
5: Retired

Compression

$x_{BIG} \rightarrow x_{small}$.



Reducing columns

aka Dimensions

Dimension ≈ "Distinct Column"

Eg. Compression #1

$m = \text{mean}(x_0)$

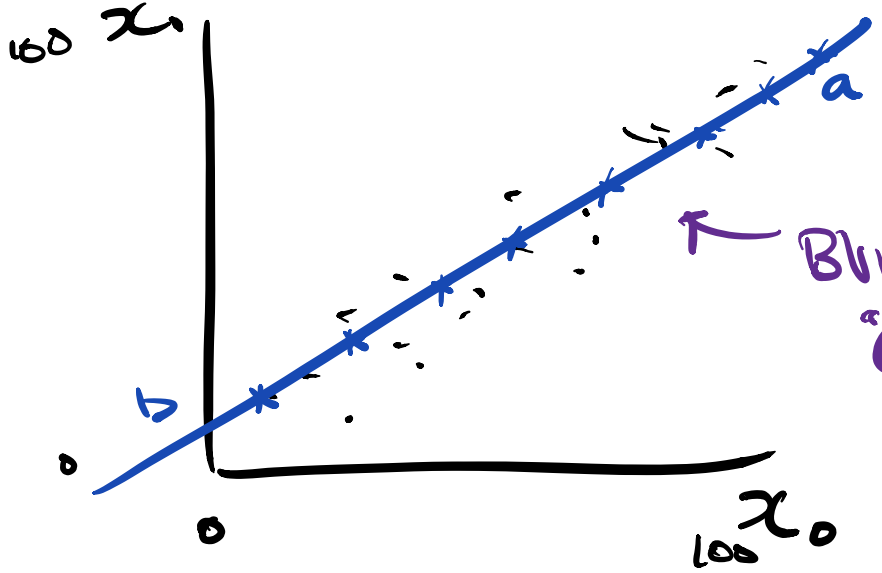
single value

column

Means Are Compression

Eg. Compression #2

2→ Linear Regression



] Real Information

x_0 | x_1

← Blue line "Captures" Most Information

0 to 100 →

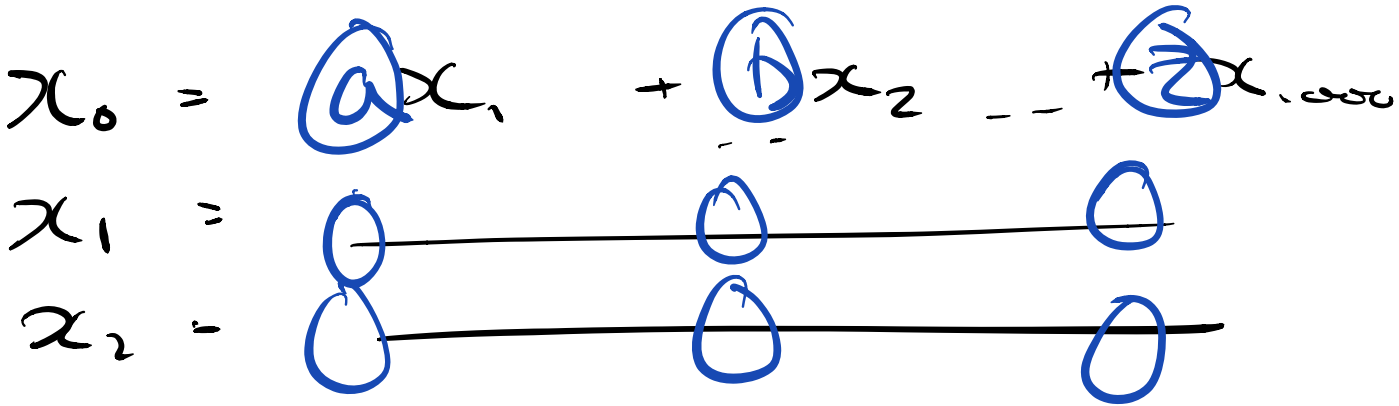
$x_1 = a x_0 + b$

↵ ← 0 to 100

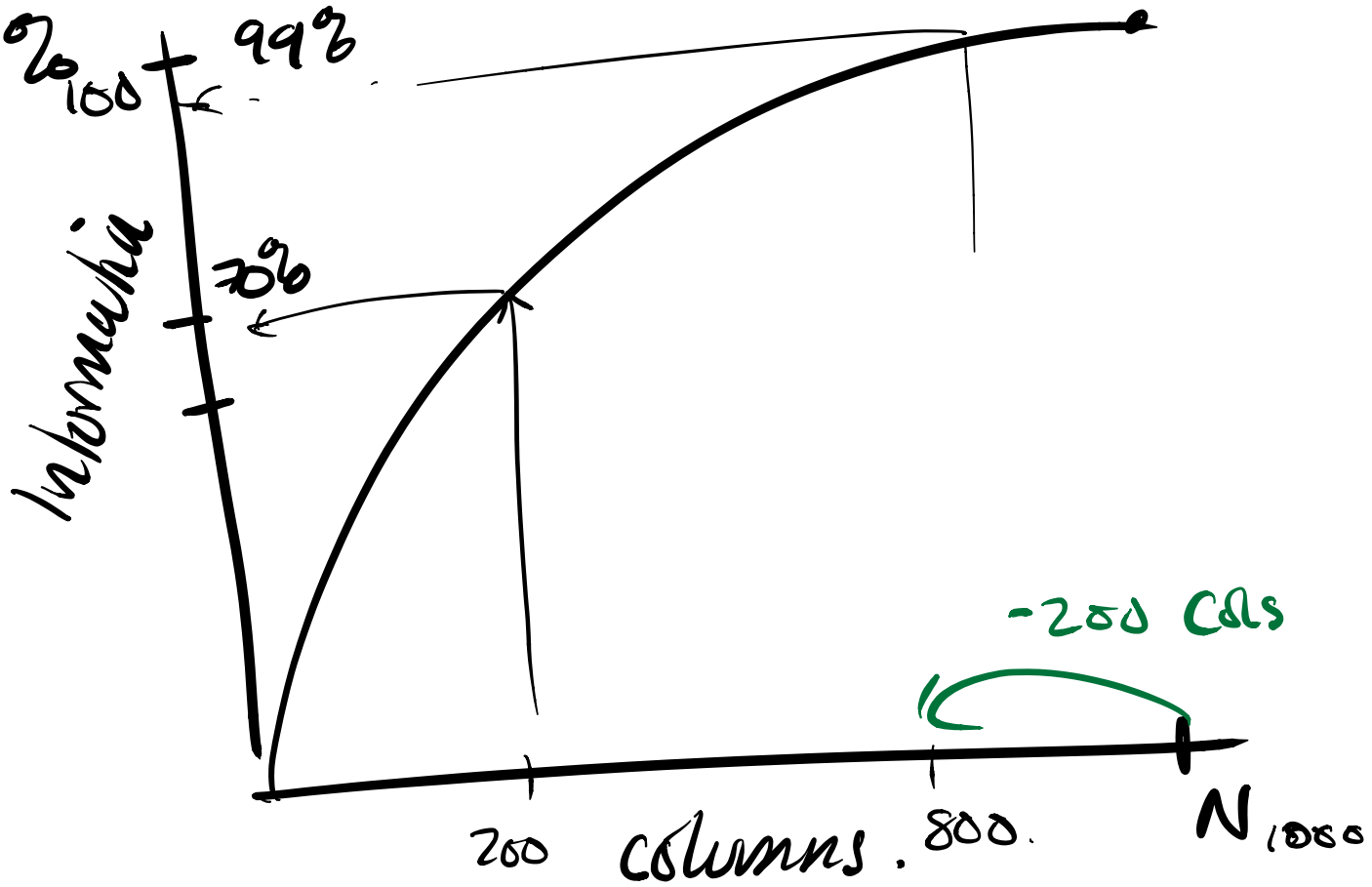
Aside: PCA

2⇒ Aprox. linear Reg. for All Cols

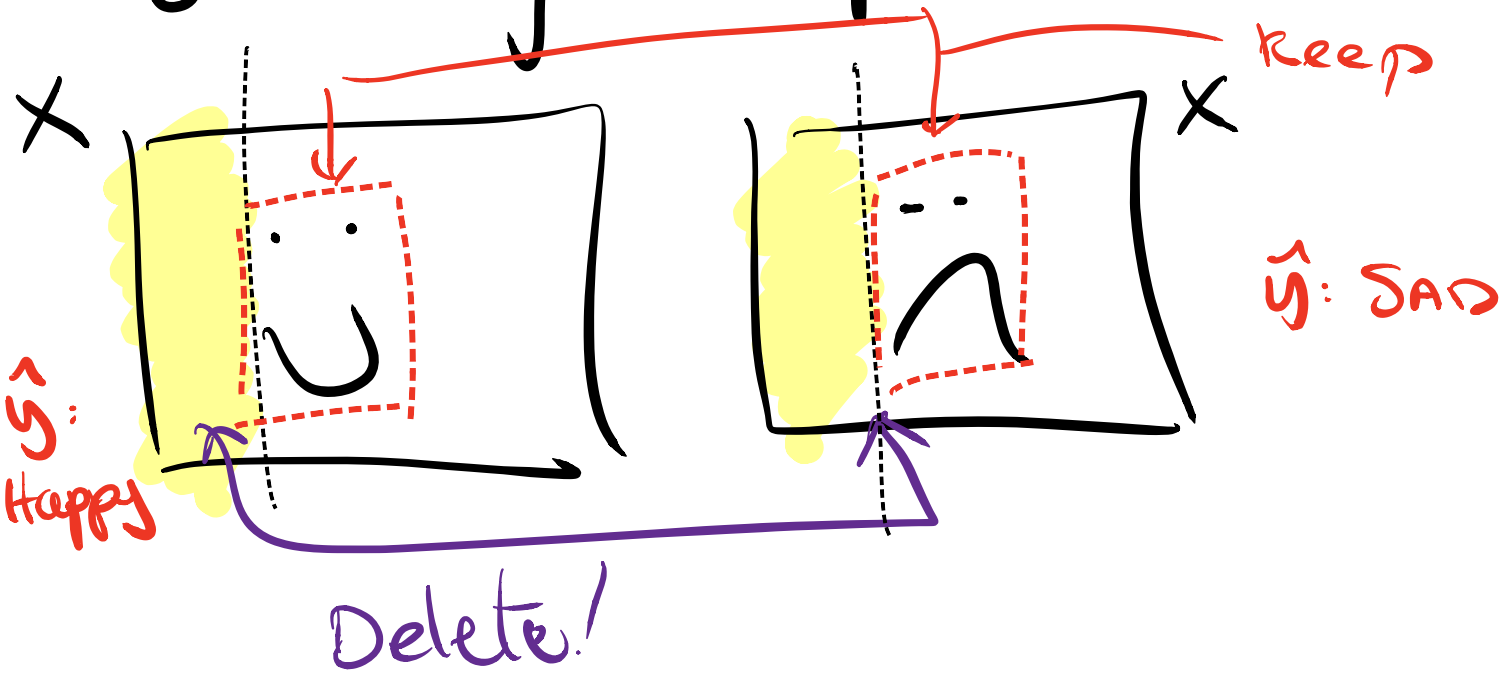
x_0 | x_1 | x_2 | ... | x_{1000}



Judgement: How to Compass



Eg. Image Compression



Goal: keep distinct areas, delete repeated areas

Aside: Images are Matrices