

Learning Algorithm:

$$S = \{ (x^{(i)}, y^{(i)}) \}_{i=1}^n$$

$$\forall_i \quad x^{(i)} \in \mathbb{R}^d$$

How do we get an encoding of input features?

Boolean Valued Data Yes/No

Did the patient experience pain?

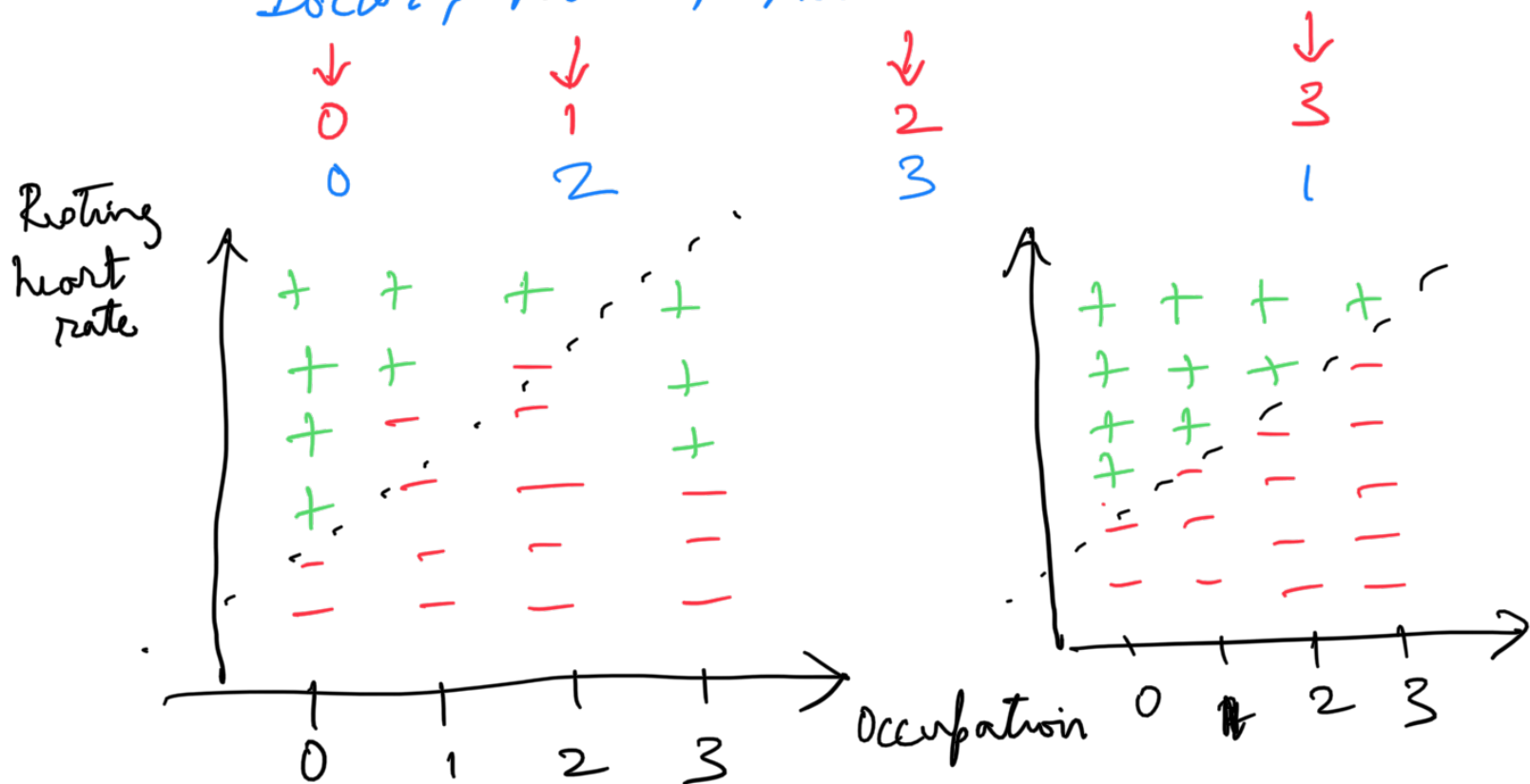
Is the patient diabetic?

Encoding is either $\{-1, +1\}$, $\{0, 1\}$
No \leftarrow \rightarrow Yes No \leftarrow \rightarrow Yes

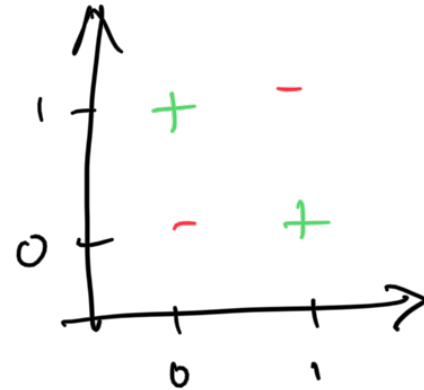
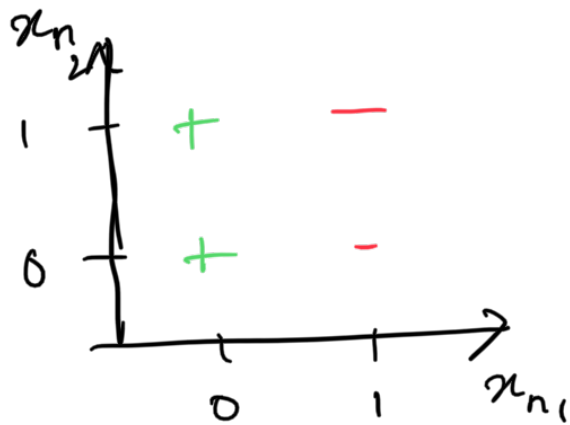
Categorical Data Finite Valued.

Occupation of the patient

Doctor, Nurse, Administrator, Pharmacist



		x_{n1}	x_{n2}
Doctor	0	0	0
Nurse	1	0	1
Administrator	2	1	0
Pharmacists	3	1	1



One Shot encoding

	Doctor	Nurse	Admin	Pharm
Doctor	1	0	0	0
Nurse	0	1	0	0

→ Medications: { diabetes, blood pressure, .. }

- ↳ No meds
- ↳ Diabetes
- ↳ Blood pressure
- ↳ Diabetes + Blood pressure

No Meds D BP D+BP

Factored Encoding: Boolean valued feature for each medication.

Ordinal Data Finite valued but ordered

Likert scales Strongly Agree Agree Neutral
 0 ← Disagree S. Disagree
 ↳ 4

Thermometer Encoding

	x_{n_1}	x_{n_2}	x_{n_3}	x_{n_4}	x_{n_5}
Strong Agree	1	0	0	0	0
Agree	1	1	0	0	0
Neutral	1	1	1	0	0
Disagree	1	1	1	1	0
S. Disagree	1	1	1	1	1

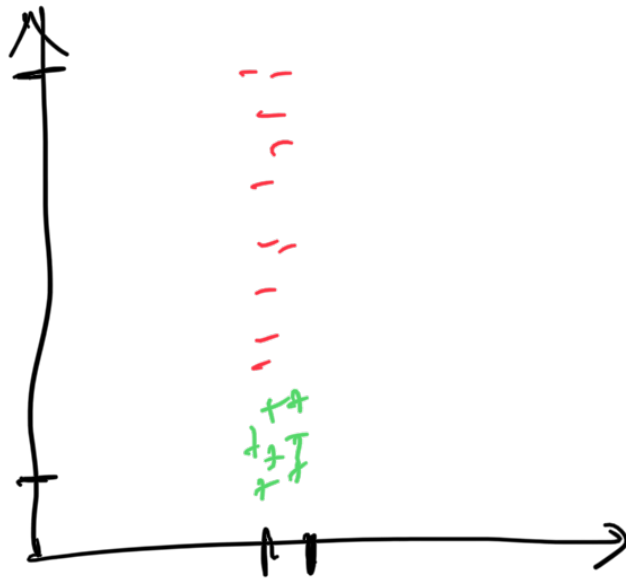
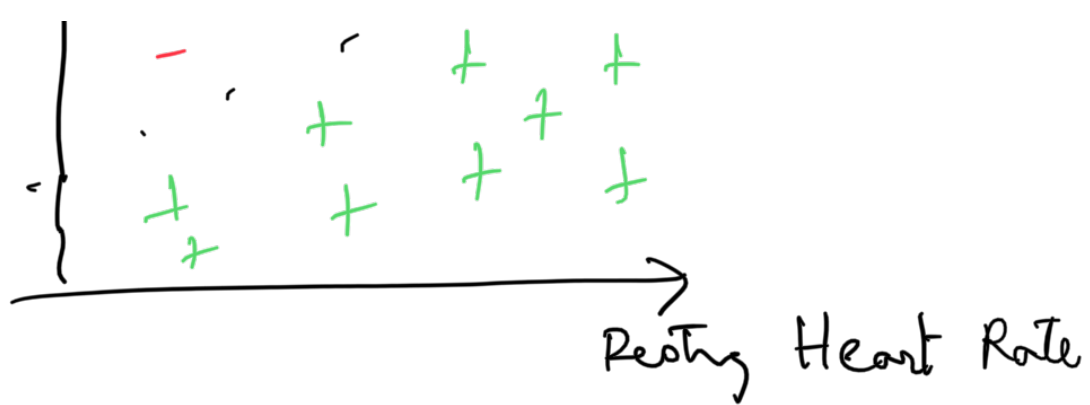
Imprecise Data

Age of patient $[20 - 29]$, $[30 - 39]$. . .
 ↓ ↓
 25 35

How an Internet mapping Glitch Turned
 a random Kansas Farm into a
 digital Hell (Kashmir Hill 2016)

Numerical Data





Data Normalization.

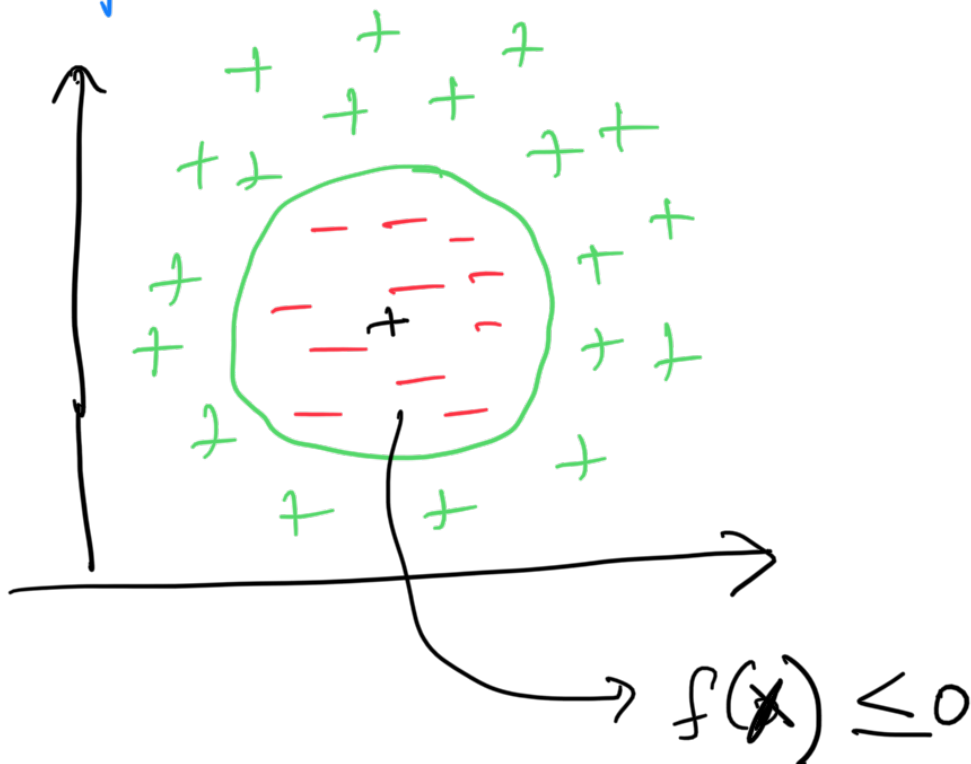
Feature x_i

$$\hat{x}_i = \frac{x_i - \text{Mean}(x_i)}{\text{Std}(x_i)}$$

Observed mean = 0

Std = 1

Auxiliary Features

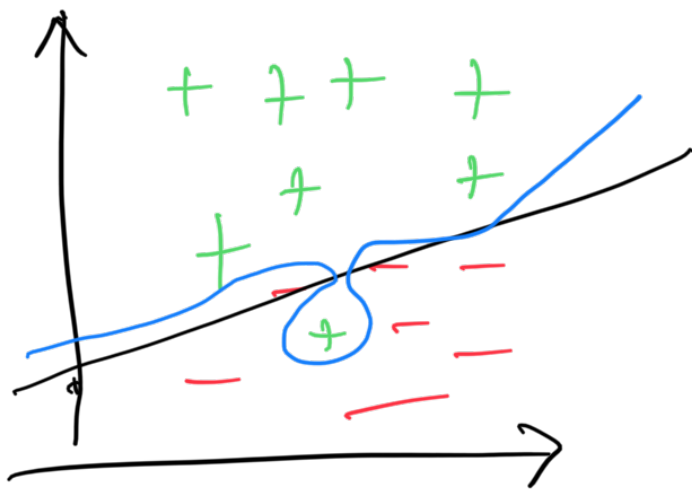


$$f(x) = (x_1 - c_1)^2 + (x_2 - c_2)^2 - x^2$$

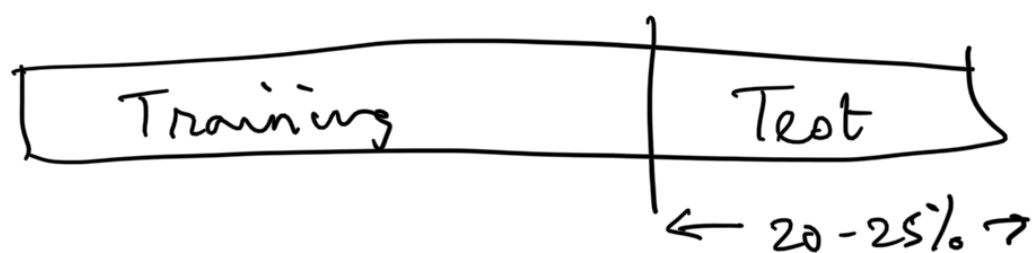
$$= a x_1^2 + b x_2^2 + c x_1 + d x_2 + e \leftarrow$$

New features $\in \mathbb{R}^4$
old features $\in \mathbb{R}^2$

	One feature	d features
0-th approximation	1	1
1st	$[1, x]$	$[1, x_1, \dots, x_d]$
2nd	$[1, x, x^2]$	$[1, x_1, \dots, x_d, \{x_i x_j\}_{i,j}]$
3rd	$[1, x, x^2, x^3]$	$[1, x_1, x_2, \dots, x_d, \{x_i x_j\}, \{x_i x_j x_k\}]$



Train - Test Data :



Randomly Shuffles the entire training set

