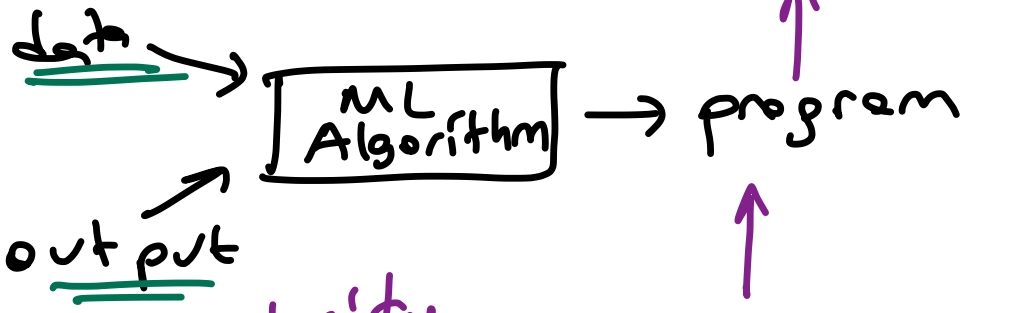
Algorithm vs ML Algorithm

- sorting
- searching
- ...

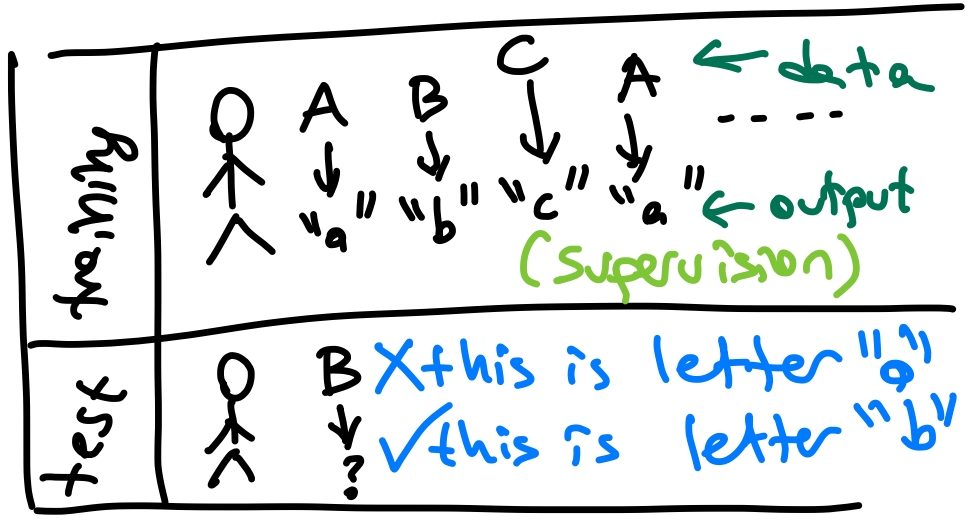
- given a list of #s
- task: to order these #s from smallest to largest

- fixed set of rules
- deterministic



- subjectivity (example data, past experience)

test data
- unseen data
- out-of-sample data



Machine Learning: programming computers to optimize learning a performance criterion using example data or past experience

- performance criterion → loss/error function → depends on the application
- example data → training set

scalar x × vector x → matrix X

inputs	predictions	error	predictions	truth
A	"a" ✓	4°C	10°C	14°C
B	"b" ✓	4°C	12°C	8°C
C	"o" ✗	4°C	8°C	12°C
E	"b" ✗	1°C	11°C	12°C
K	"k" ✓	2°C	12°C	10°C

absolute error $|\hat{y}_i - y_i|$

- $\hat{y}_i$ → estimated
- y_i → truth

Supervised Learning:

$\mathcal{X} = \{(x_i, y_i)\}_{i=1}^{N}$ → N → # of data points

$\mathcal{X}$ → training set

x_i → ith data point

y_i → ith label / ith target / output

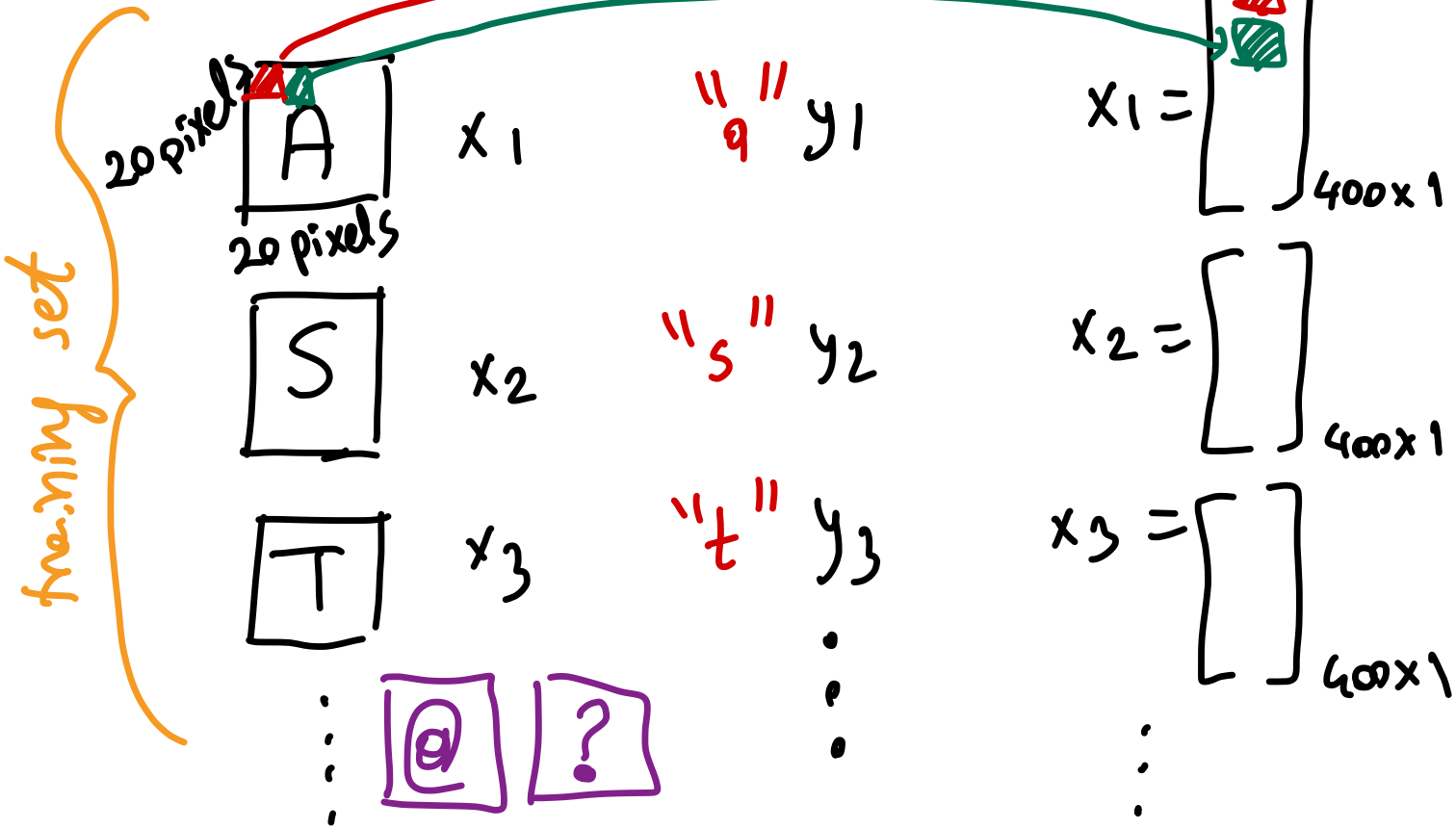
i → index for data points

$$\mathcal{X} = \{(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_N, y_N)\}$$

$$\mathcal{X} = \{([\,], \square)\}_{i=1}^{N}$$

Classification

training set



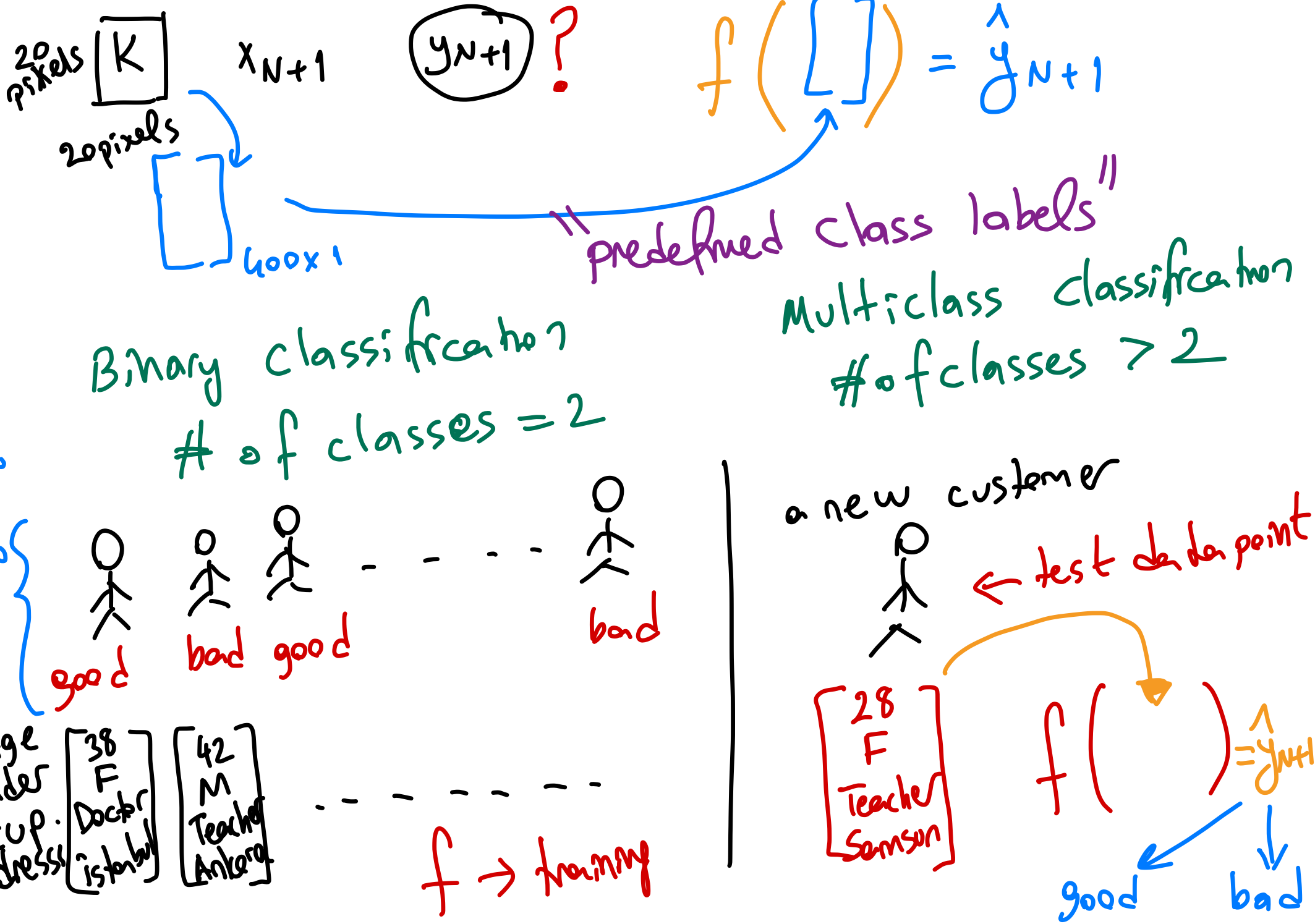
$y_1 =$ a

$y_2 =$ s

$y_3 =$ t

$x_i \in \mathbb{R}^{400} \quad \forall i$

$y_i \in \{a, b, \ldots, z\}$

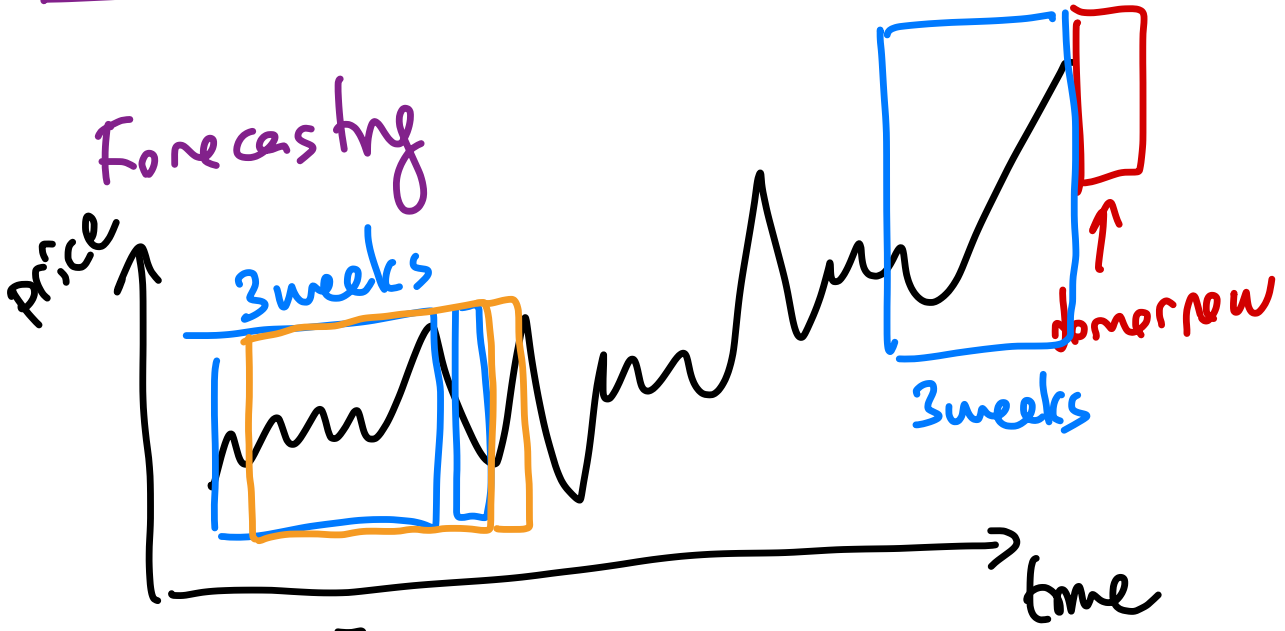

20 pixels
K
20 pixels
400x1
x_{N+1}
y_{N+1} ?
$f([\,]) = \hat{y}_{N+1}$
"predefined class labels"
Multiclass classification
# of classes > 2
Binary classification
# of classes = 2
training set
good
bad
good
bad
Age
Gender
occup.
Address
38
F
Doctor
Istanbul
42
M
Teacher
Ankara
f → training
a new customer
test data point
28
F
Teacher
Samsun
$f(\,) = \hat{y}_{N+1}$
good
bad

Regression

$$\mathcal{X} = \{(x_i, y_i)\}_{i=1}^{N} \qquad y_i \in \mathbb{R}$$

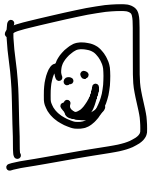
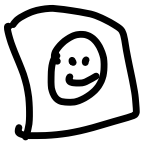
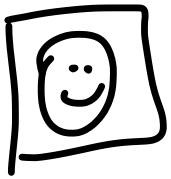
27,54 USD/TRY

Forecasting



$$X_1 = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{21} \end{bmatrix} \qquad y_1 = p_{22}$$

$$X_2 = \begin{bmatrix} p_2 \\ p_3 \\ \vdots \\ p_{22} \end{bmatrix} \qquad y_2 = p_{23}$$

Picture	Age
	24
	32
⋮	⋮
	18
	? 58

Unsupervised Learning

clustering

$$\mathcal{X} = \{x_i\}_{i=1}^{N}$$

NO CLASS LABELS!