Supervised & Unsupervised Machine Learning

History

“Field of study that gives computers the ability to learn without being explicitly programmed.”

Arthur Samuel (1959)

What do you think?

If the checkers program had been allowed to play only 10 games (instead of tens of thousands) against itself, a much smaller number of games, how would this have affected its performance?

- Would have made it better
- Would have made it worse

Supervised Learning

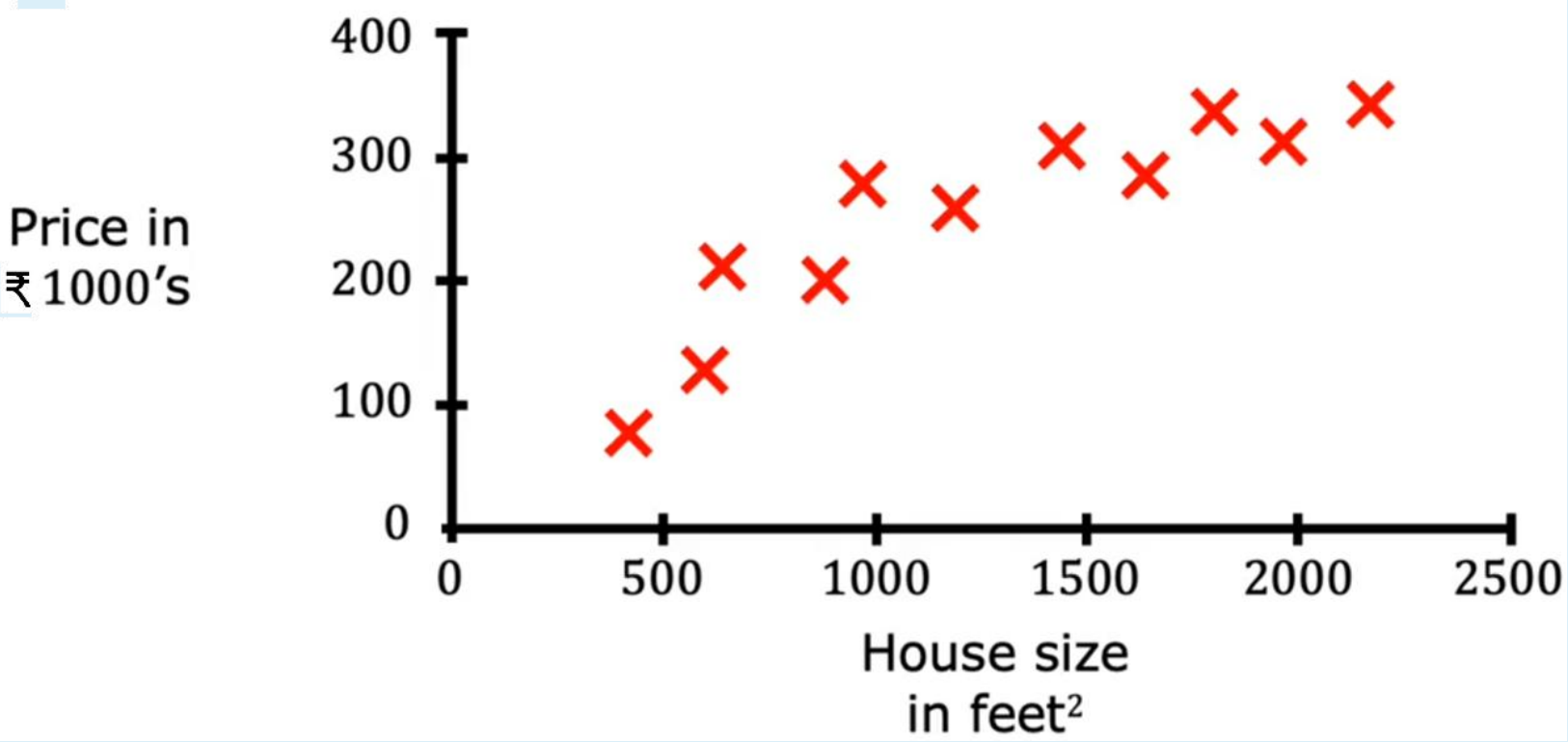
Learns from being given “Right answers”

Input X

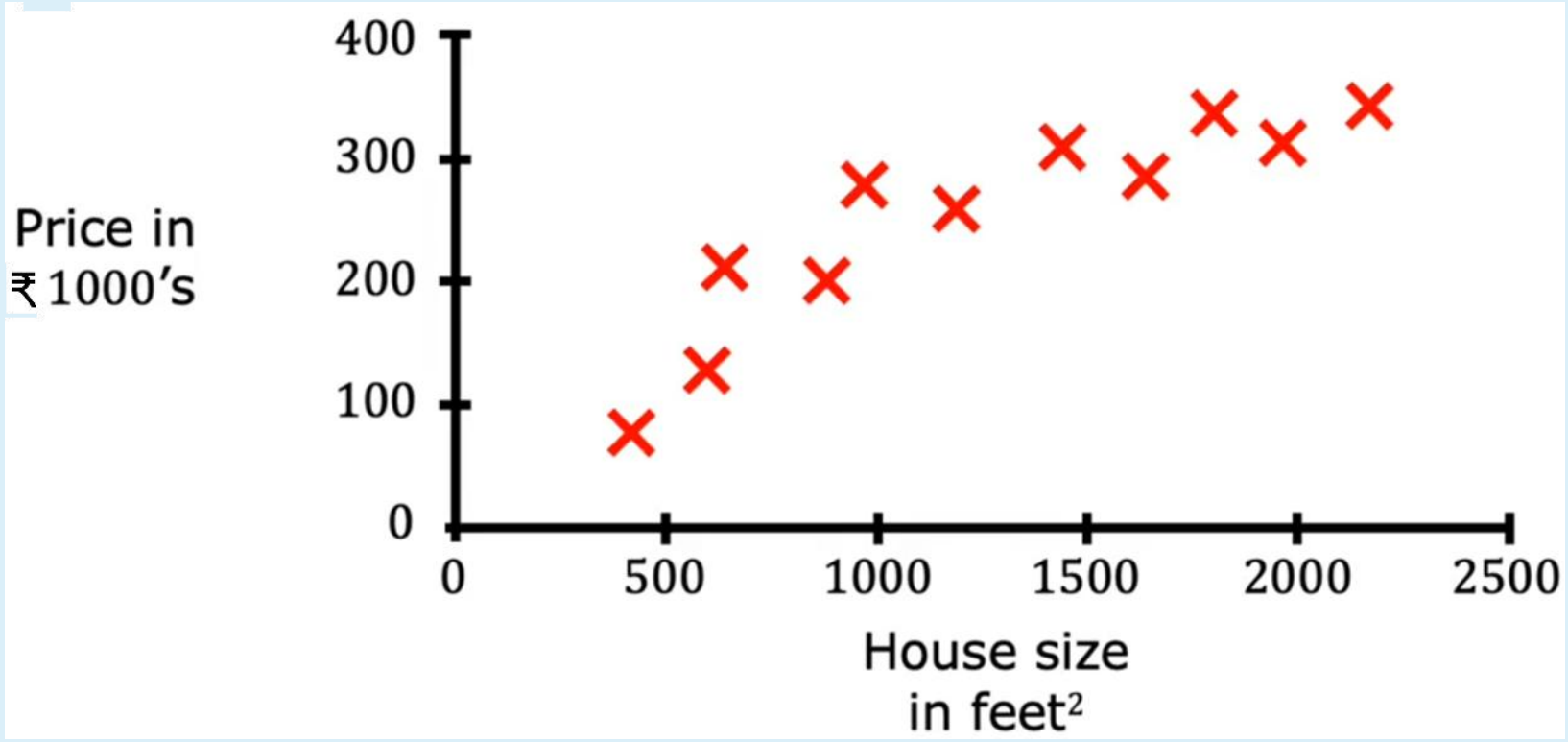
Output Label Y

Application	Input (X)	Output (Y)
Spam filter	email	Spam? (0/1)
Speech recognition	Audio	Text transcripts
Machine translation	English	Hindi
Online advertising	Ad, user info	Click? (0/1)
Self-driving car	Image, radar info	Position of the car
Visual inspection	Image of PCB	Defect? (0/1)

Housing price prediction



Housing price prediction **Regression**

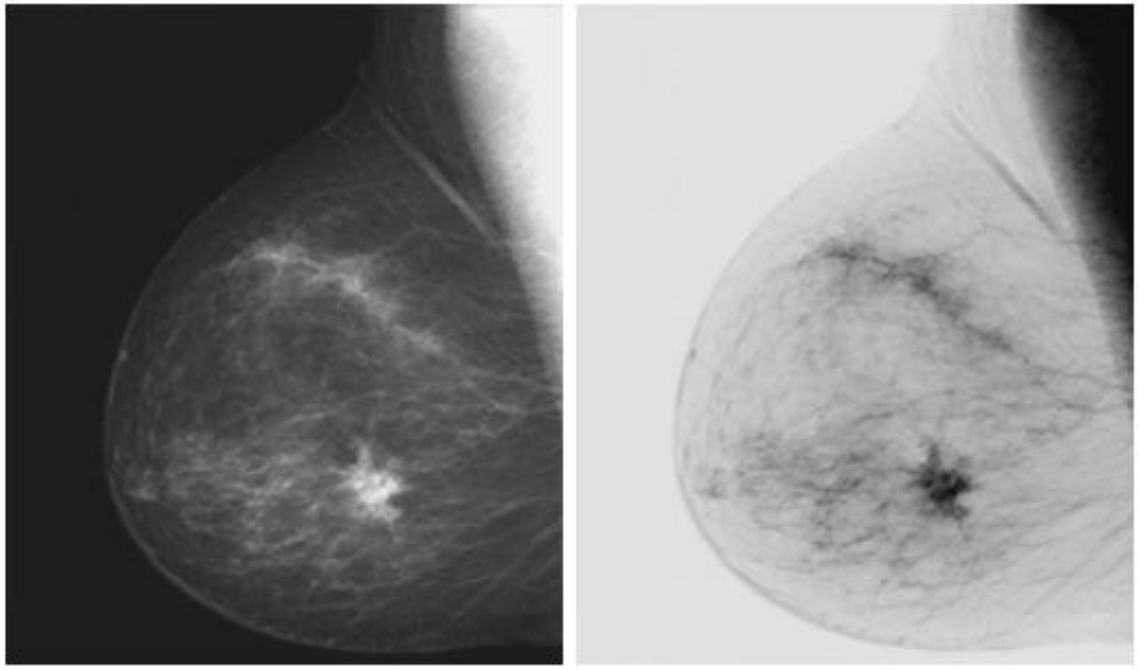


Regression can predict Infinitely many possible outputs

Image Negatives

Denote [0, L-1] intensity levels of the image.

Image negative is obtained by ***s= L-1-r***



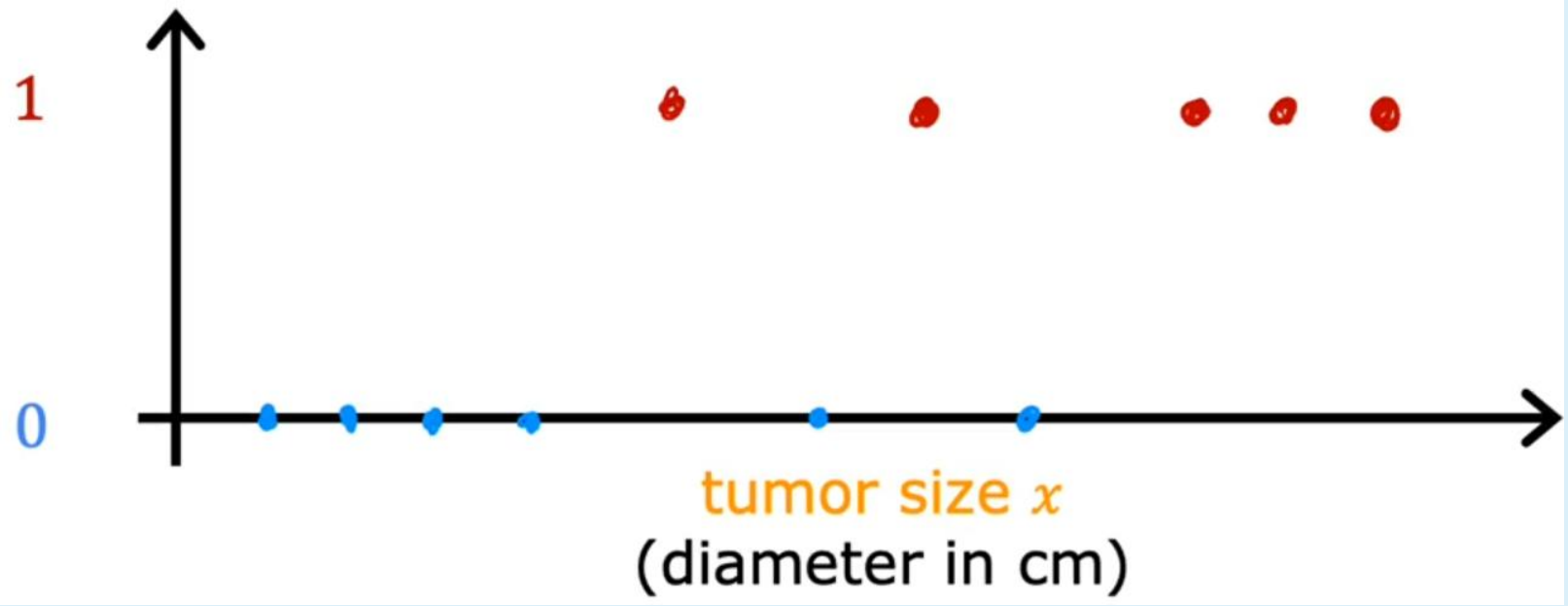
a b

FIGURE 3.4 (a) Original digital mammogram. (b) Negative image obtained using the negative transformation in Eq. (3.2-1). (Courtesy of G.E. Medical Systems.)

Breast cancer detection **Classification**

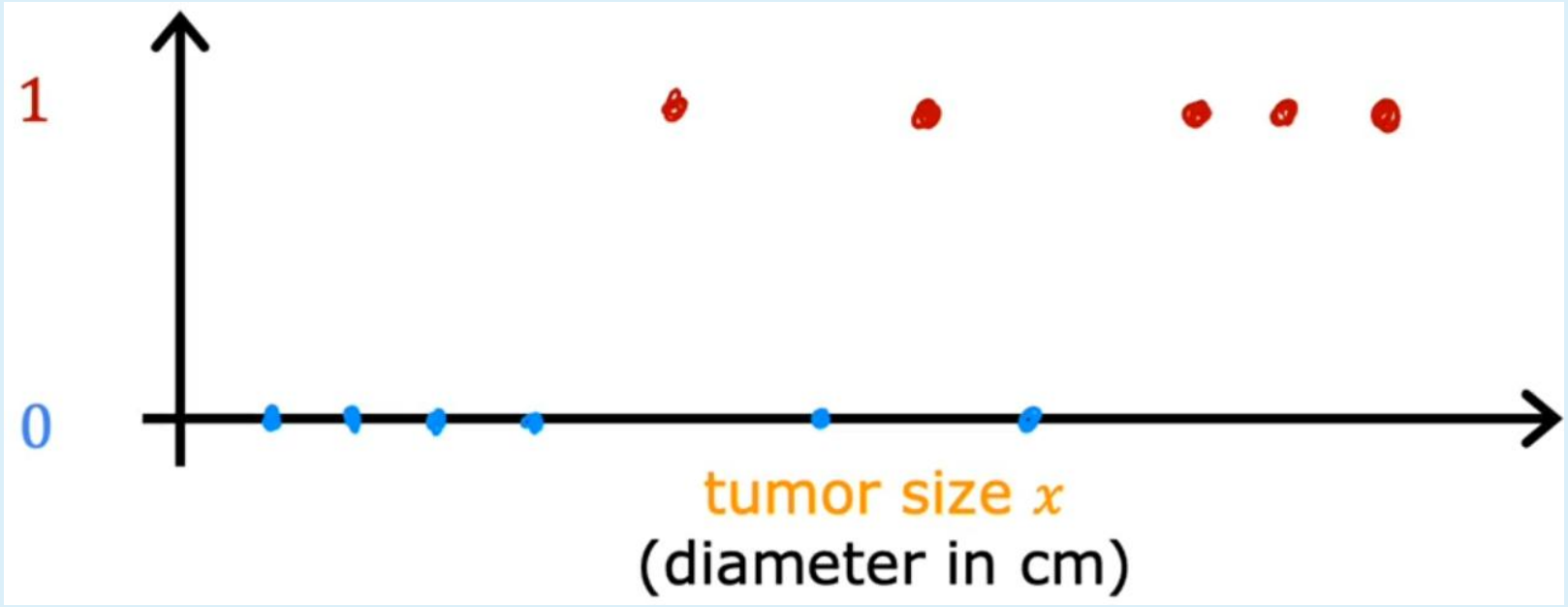
Size of the tumor

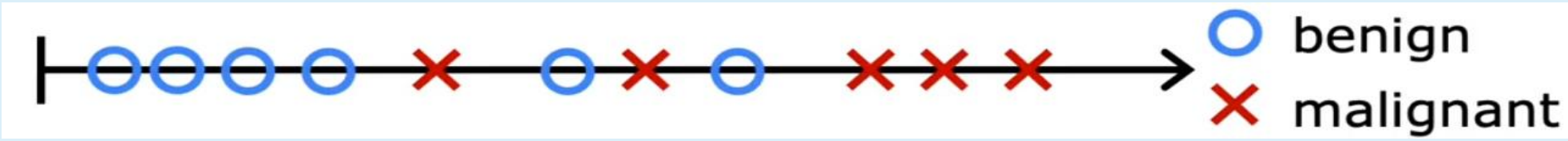
Category of the tumor (Malignant/benign)



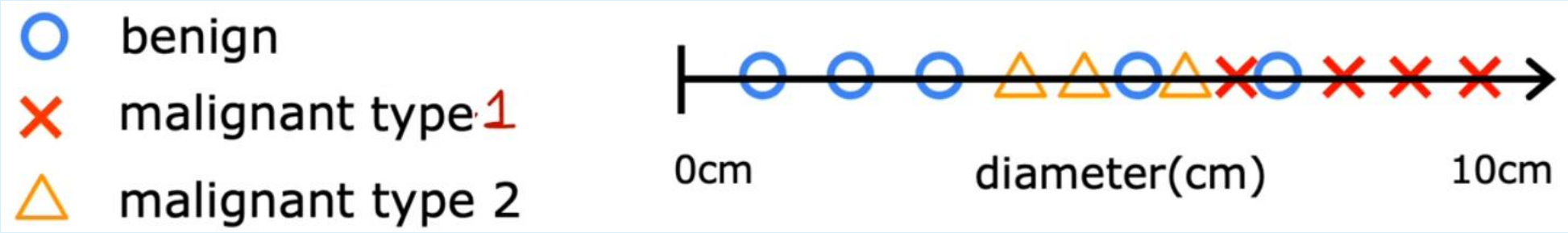
Size (cm)	Diagnosis	Output
1	B	0
4	M	1
2	B	0
6	M	1
7	M	1
…	…	…

Breast cancer detection Classification



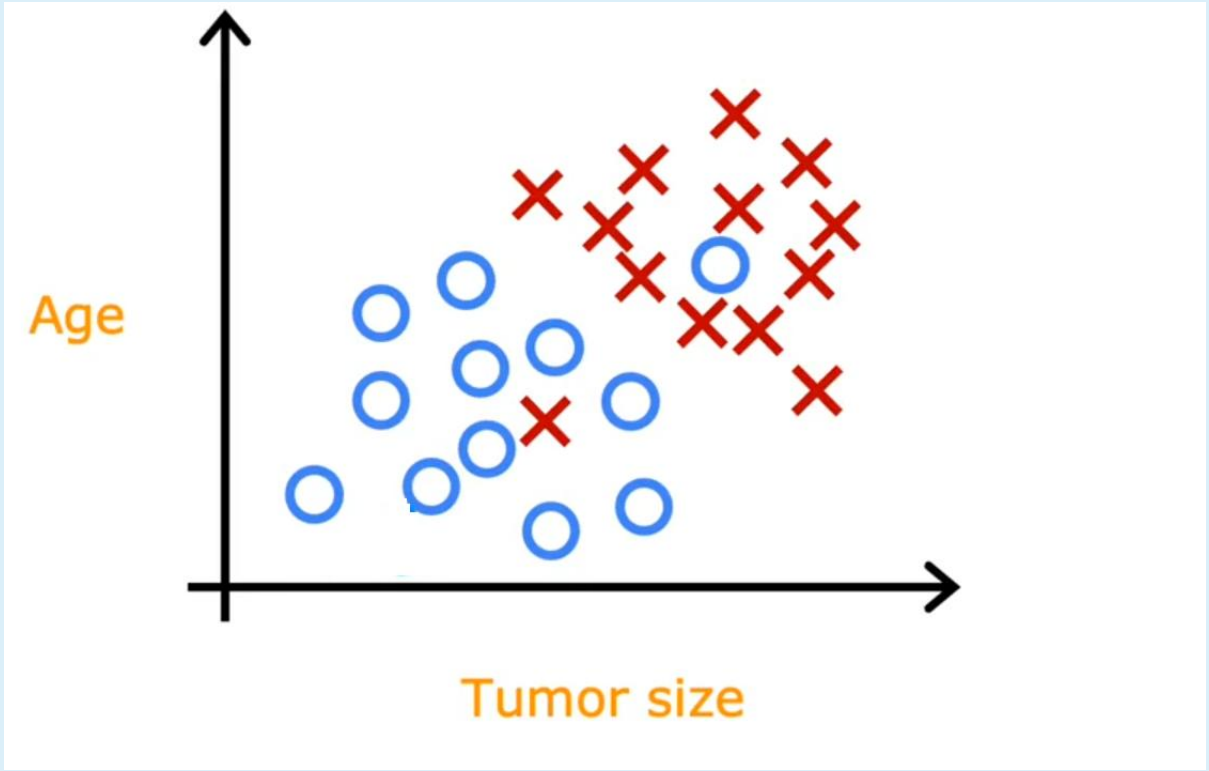


Breast cancer detection Classification



Classification predicts categories which are small number of possible output

Classification: Two or more inputs



Supervised Learning

Learns from being given the “right answers”

REGRESSION

Predict a number

Infinitely many possible outputs

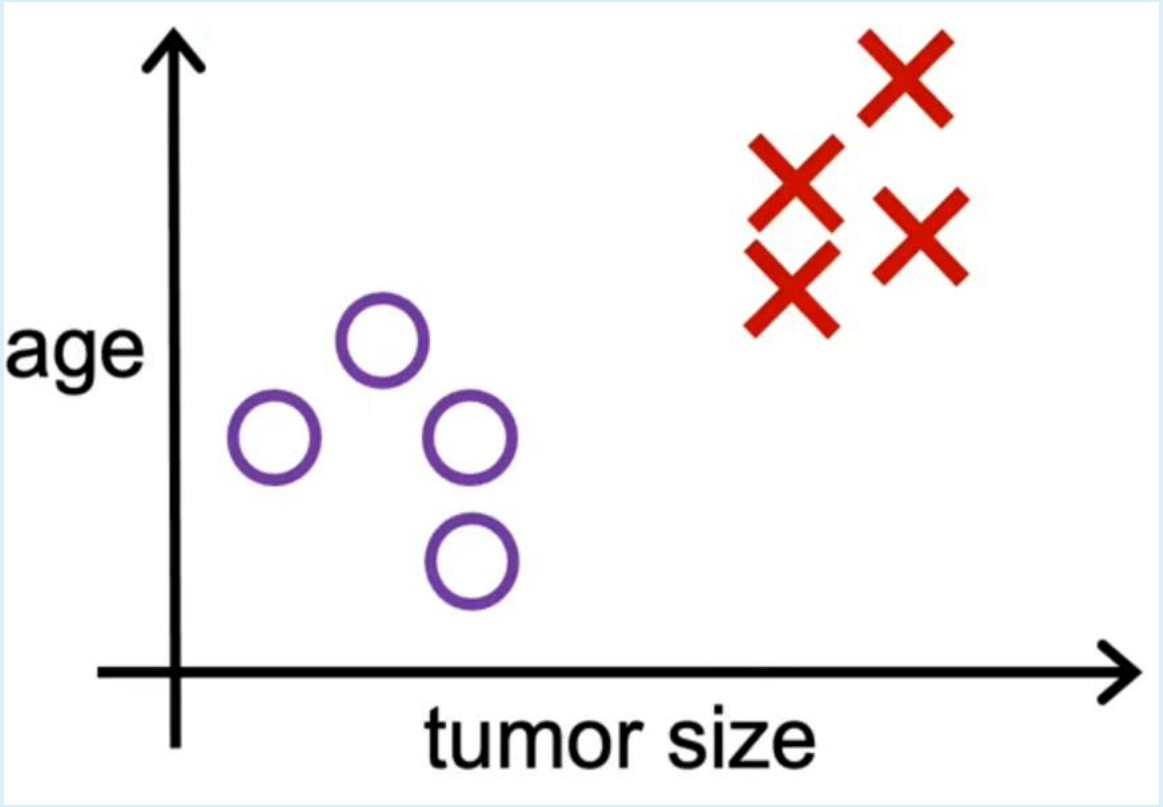
CLASSIFICATION

Predict categories

Small number of possible outputs

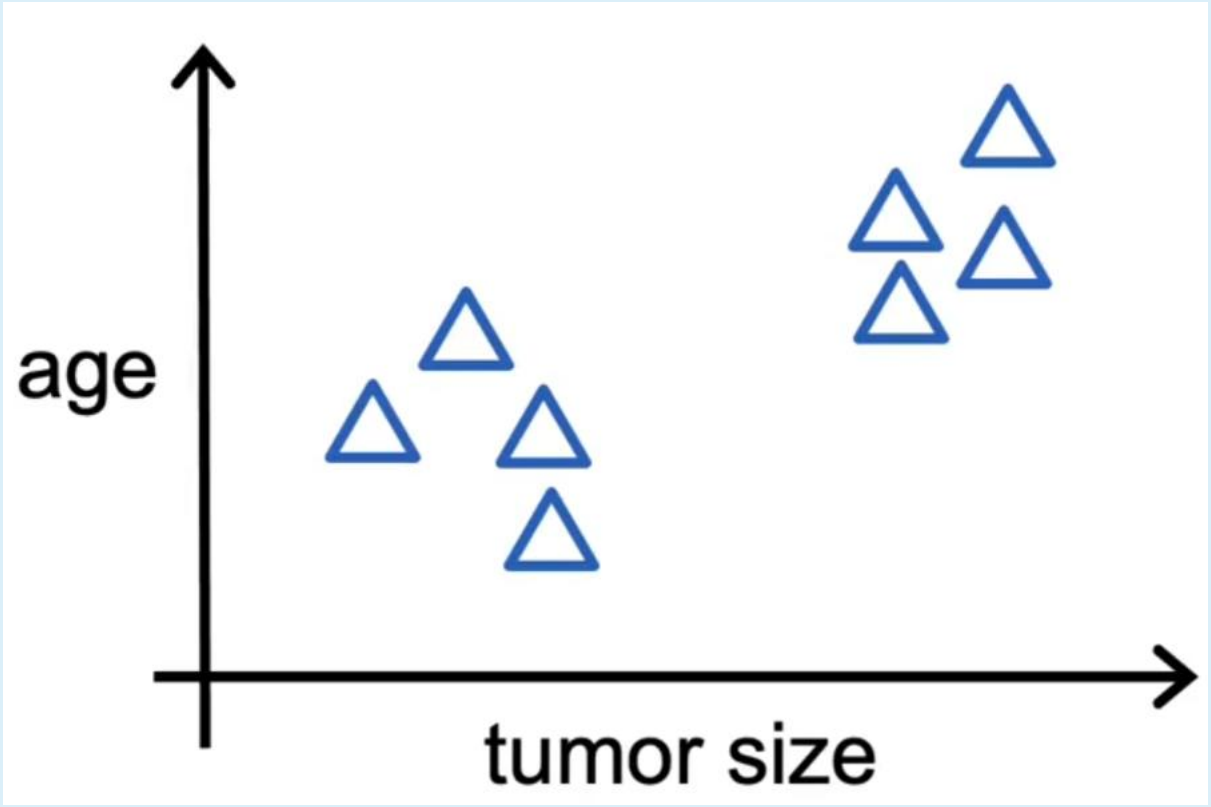
Supervised learning

Learns from the data labeled with the “right answers”



Unsupervised learning

Find something interesting in unlabeled data



Clustering: Google news

Hindustan Times

Apple AI features to not be free soon, company may launch monthly subscription

12 minutes ago

Business Today

Apple's big AI features might not stay free forever: Report

3 hours ago • Priya Singh

The Indian Express

As Apple aims to monetise AI, select Apple Intelligence features may stay behind paywall

2 hours ago

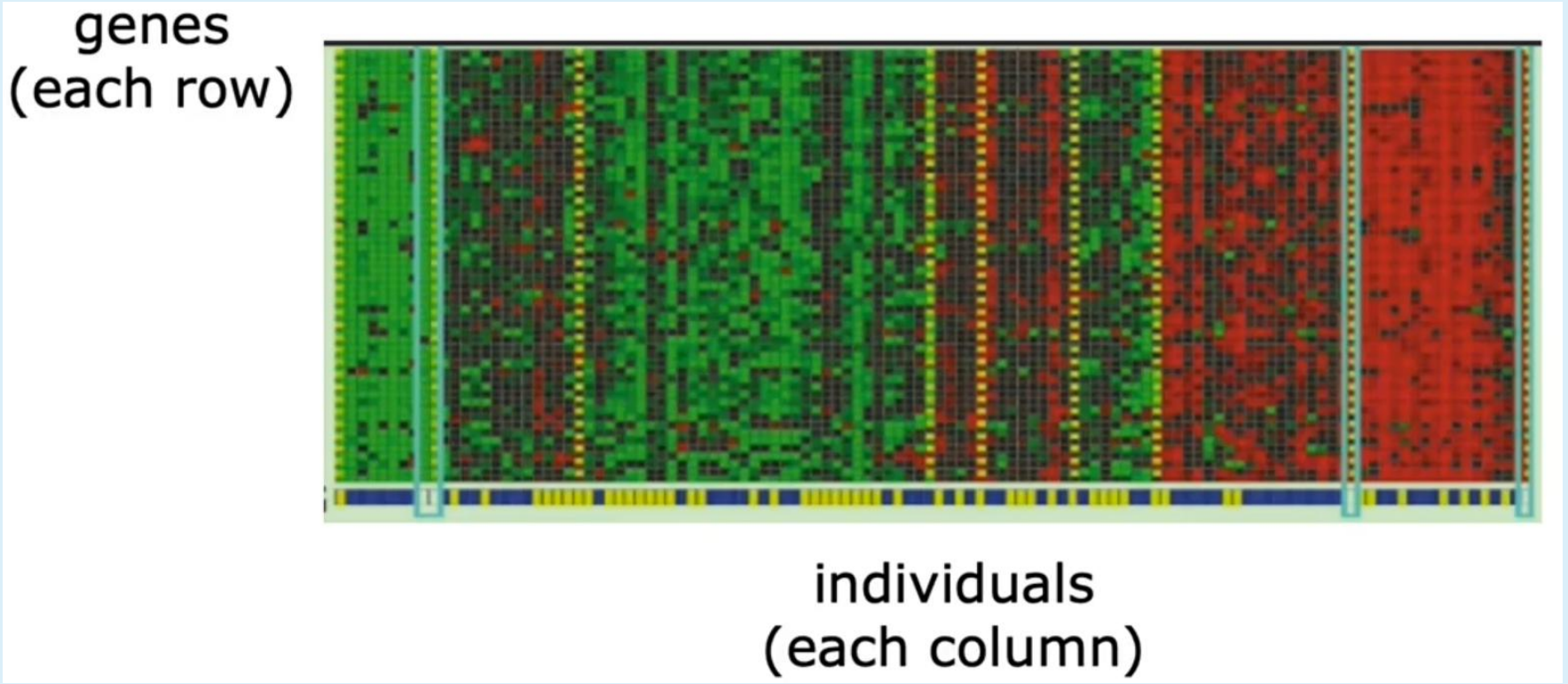
The Times of India

Apple eyes paid subscriptions for advanced AI features

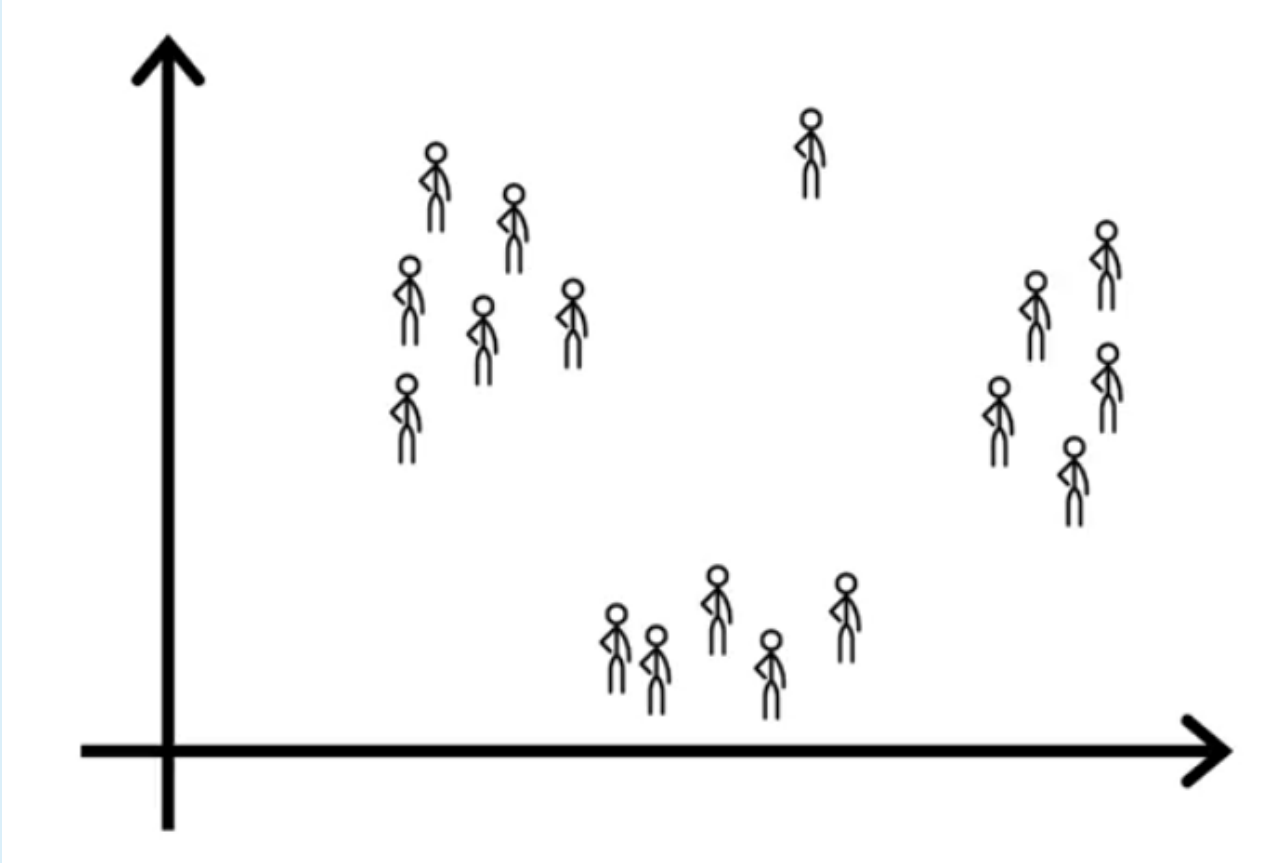
36 minutes ago

Full coverage

Clustering: DNA microarray



Clustering: Grouping customers



Unsupervised learning

Data only comes with input label X, but not output labels Y

Algorithm has to find structure in the data

Clustering: Group similar data point together

Anomaly detection: Find unusual data points

Dimensionality reduction: Compress data using few numbers

Of the following examples, which would you address using Unsupervised learning algorithm?

- ❑ Given email labeled as spam/not spam, learn a spam filter.
- ❑ Given a set of news articles found on the web, group them into sets of articles about the same story.
- ❑ Given a database of customer data, automatically discover market segments and group customers into different market segments.
- ❑ Given a dataset of patient diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not.