

# Introduction to Machine Learning

Eun-Sol Kim (김은솔)

Artificial Intelligence

# TAs

- ❑ Seokjun On(온석준), Lee Gyu Geun(이규근)
- ❑ For attendance-related inquiries, please contact the email below.
  - ❑ 예비군, 공결 등등
  - ❑ [mllab.ta.ai@gmail.com](mailto:mllab.ta.ai@gmail.com)

# Training and Testing

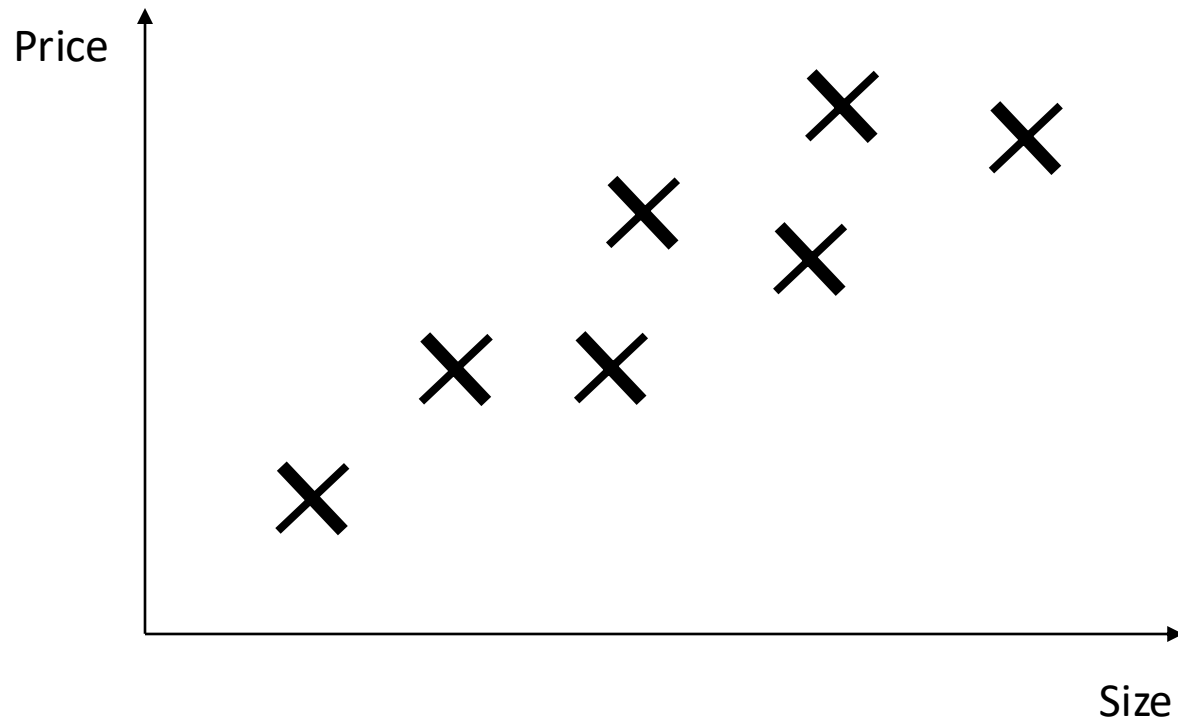
- ❑ The main purpose of the ML algorithm is to make suitable predictions for unseen inputs.
  - ❑ Training with given data
  - ❑ Testing with unseen and new data

# An Example

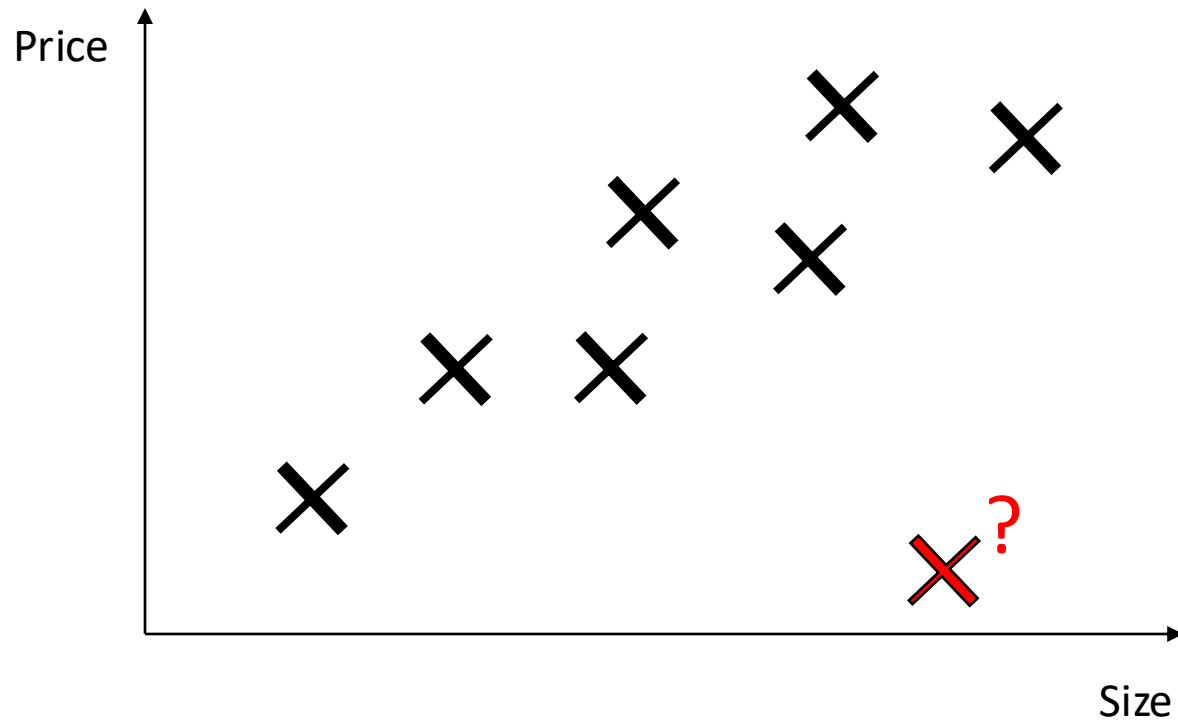
## ❑ Housing Price Prediction

No	Size of house	# of bedrooms	Age of house	Size of kitchen	Price
1	20	2	20	5	700
2	20	3	10	3	900
3	30	3	5	5	1400
4	35	3	10	8	1400
5	35	4	10	8	1450
6	40	4	5	4	1800
7	37	4	5	4	1700

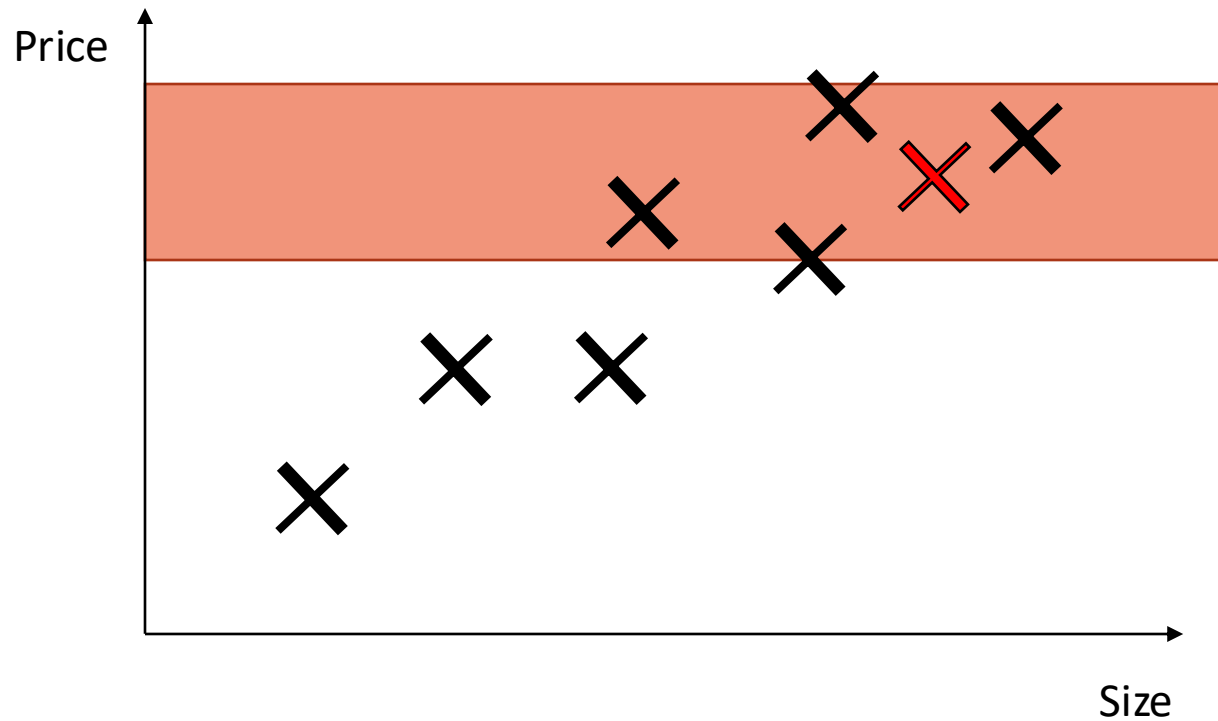
# Housing Price Prediction



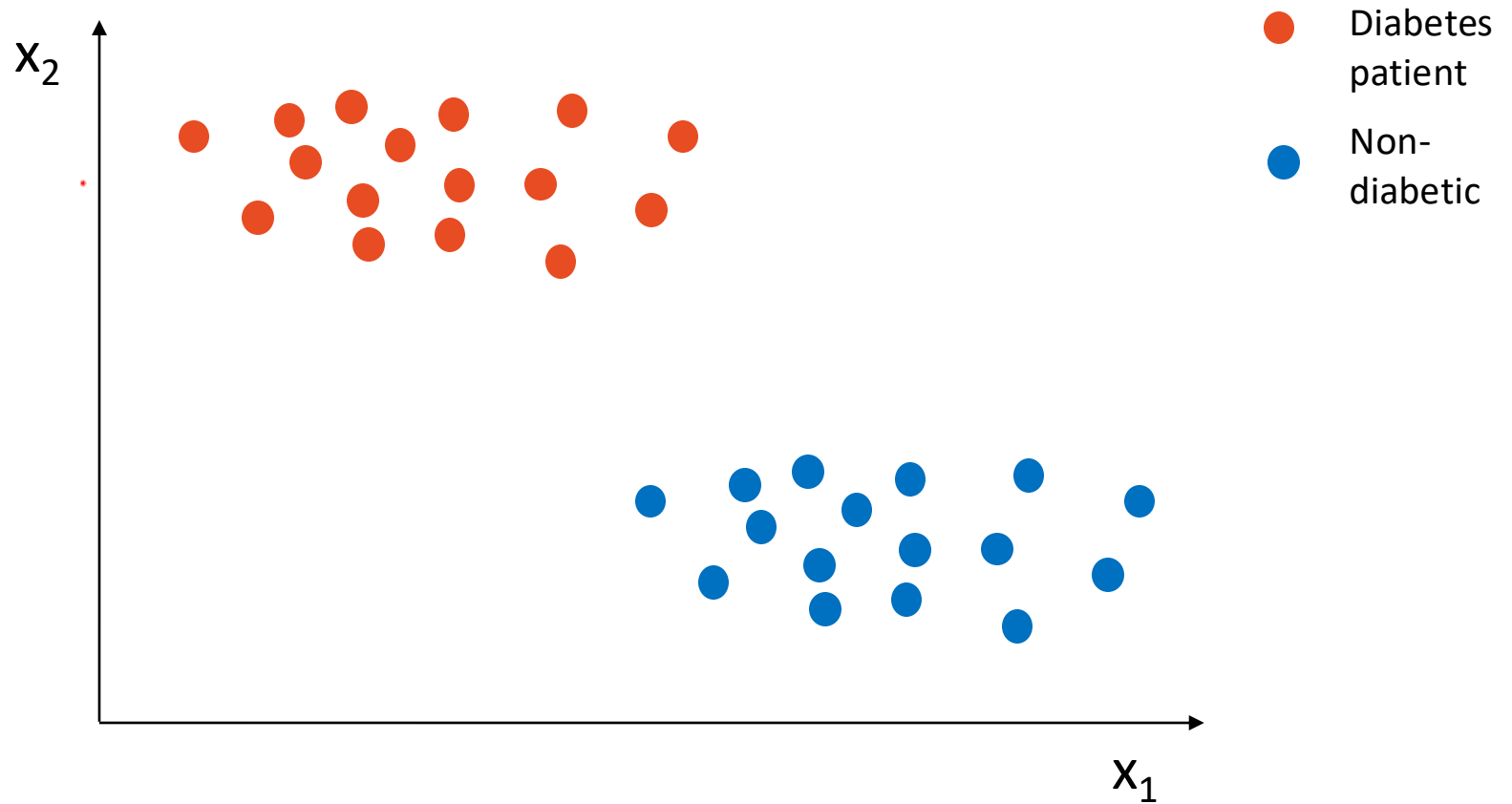
# Housing Price Prediction



# Housing Price Prediction

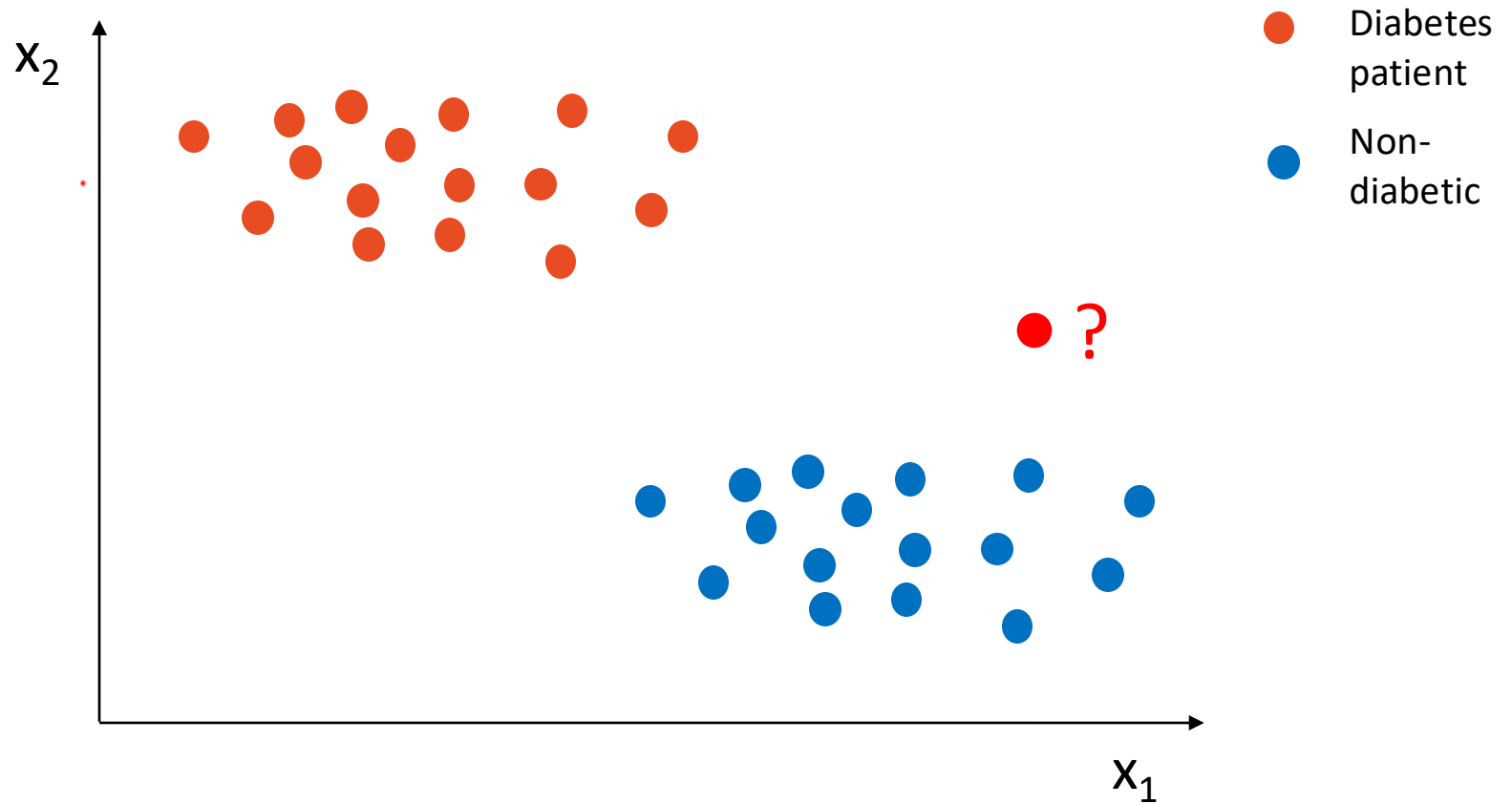


# Classification Example

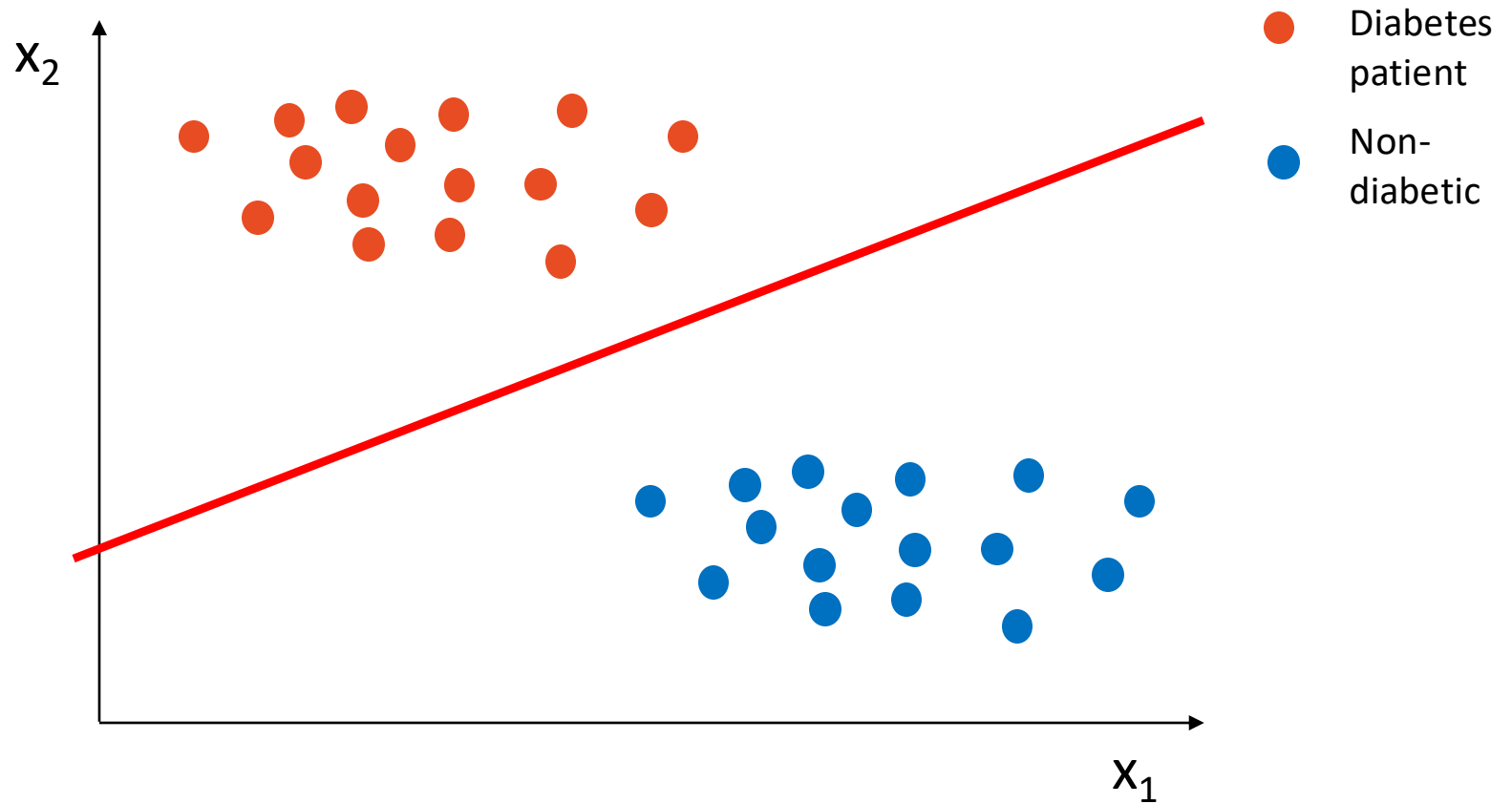




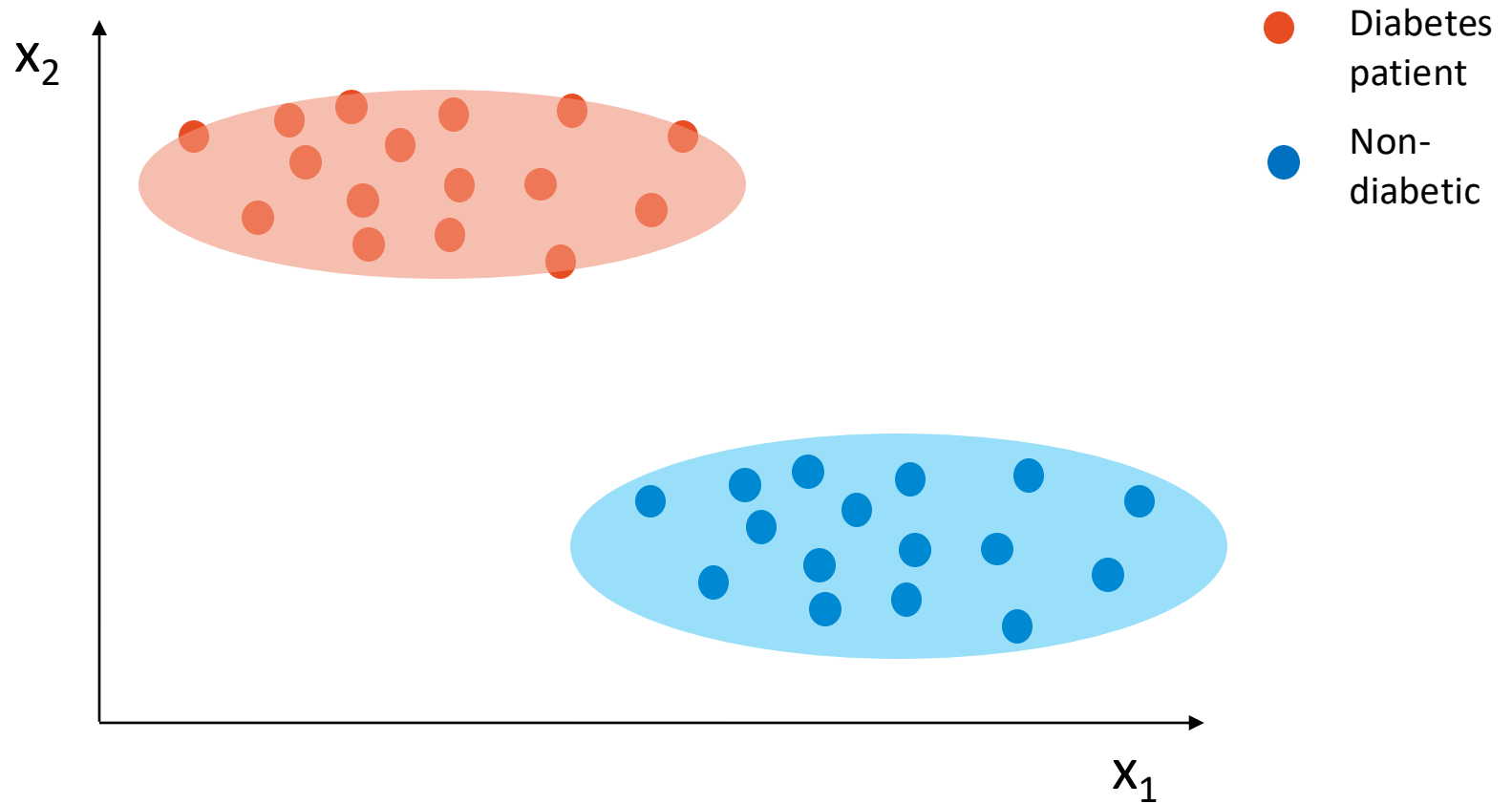
# Classification Example



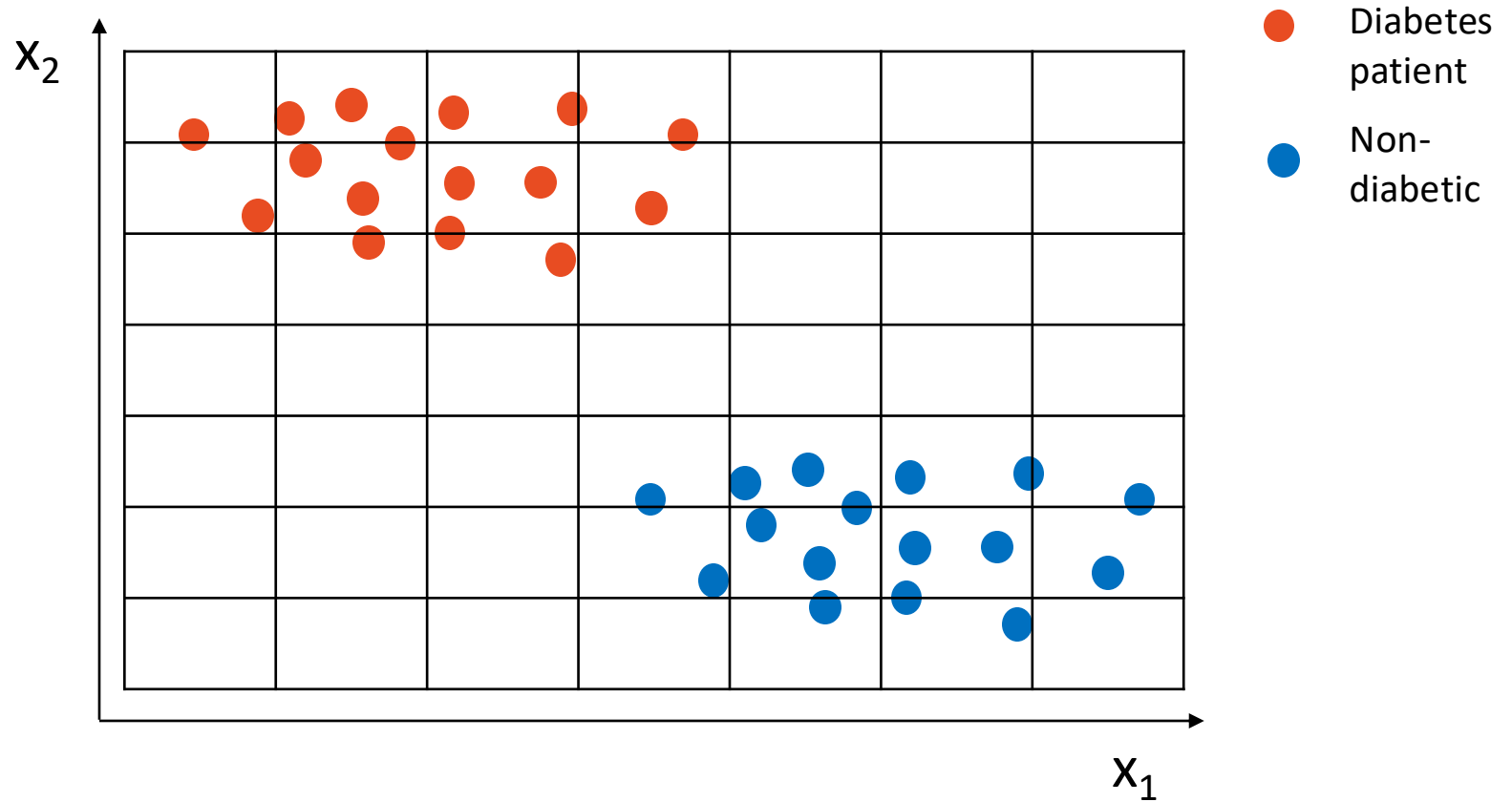
# Classification Example



# Classification Example

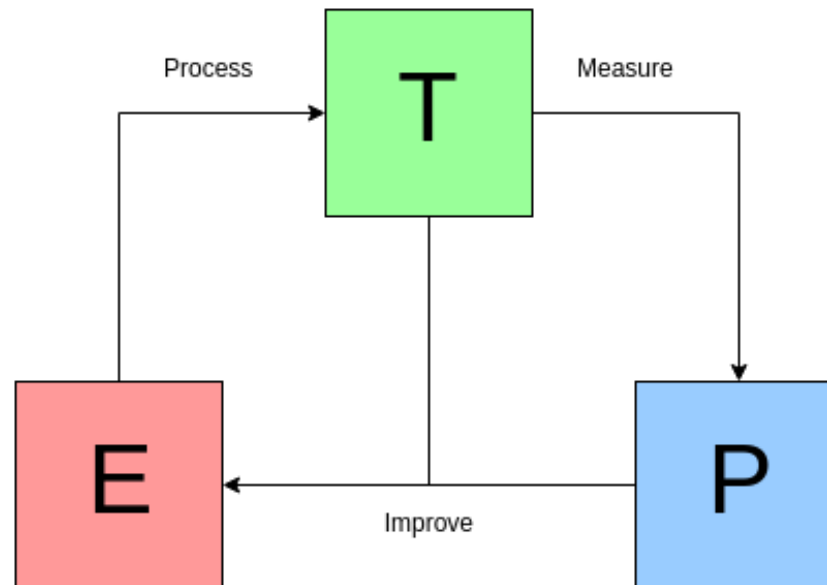


# Classification Example



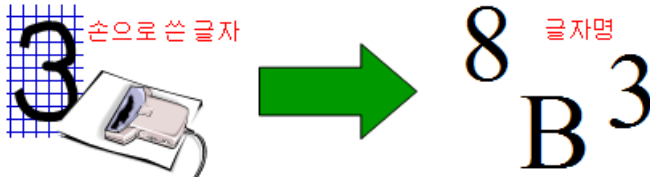
# Machine Learning

- ❑ Study of algorithms that
  - ❑ Improve their performance  $P$
  - ❑ At some task  $T$
  - ❑ With experience  $E$
- ❑ Learning = **improving  $P$  at  $T$  with  $E$**

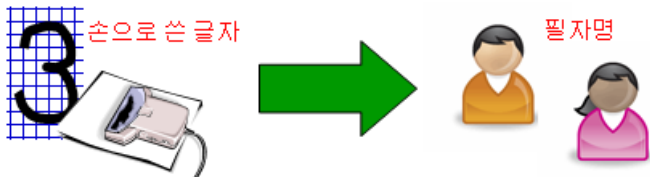


# T (Task) = Something Done (Output)

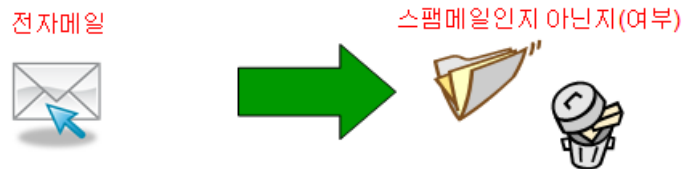
## ■ 손으로 쓴 글자 인식



## ■ 필자 인식



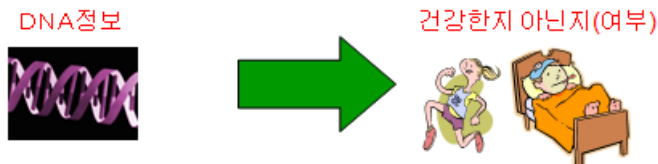
## ■ 스팸메일 필터



## ■ 컴퓨터 시각



## ■ 유전자 진단



☐ Classification

☐ Regression

☐ Diagnosing

☐ Diabetes

☐ Cancer

☐ Driving

☐ Car

☐ Forecasting

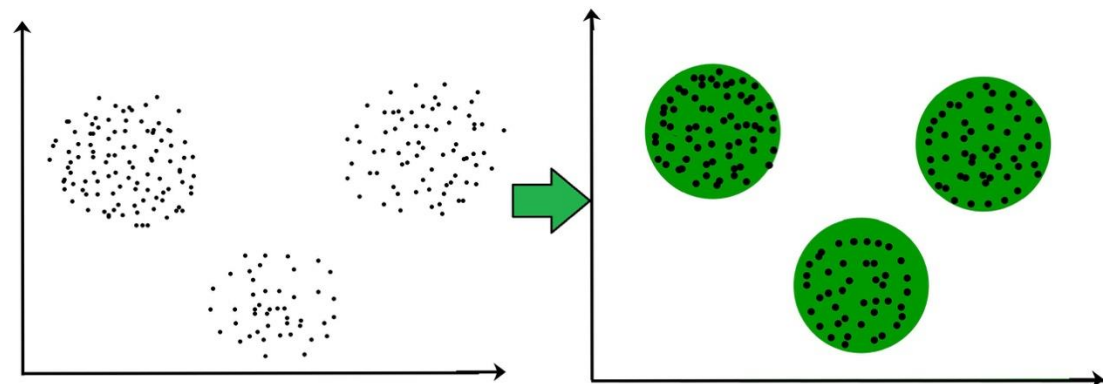
☐ Weather

☐ Stock price

# E (Experience) = (Training) Data

- ☐ Supervised learning
  - ☐ Learn w/ answers (class)
  - ☐ Classification
    - ☐ Yes/No, A/B/C/D, ...
  - ☐ Regression
    - ☐ Real number
- ☐ Unsupervised learning
  - ☐ Learn w/o answers
    - ☐ (no class)
  - ☐ Clustering
    - ☐ Feature selection
- ☐ Semi-supervised learning
  - ☐ Web image auto-tagging
- ☐ Reinforcement learning
  - ☐ AlphaGo, AlphaStar, Chess

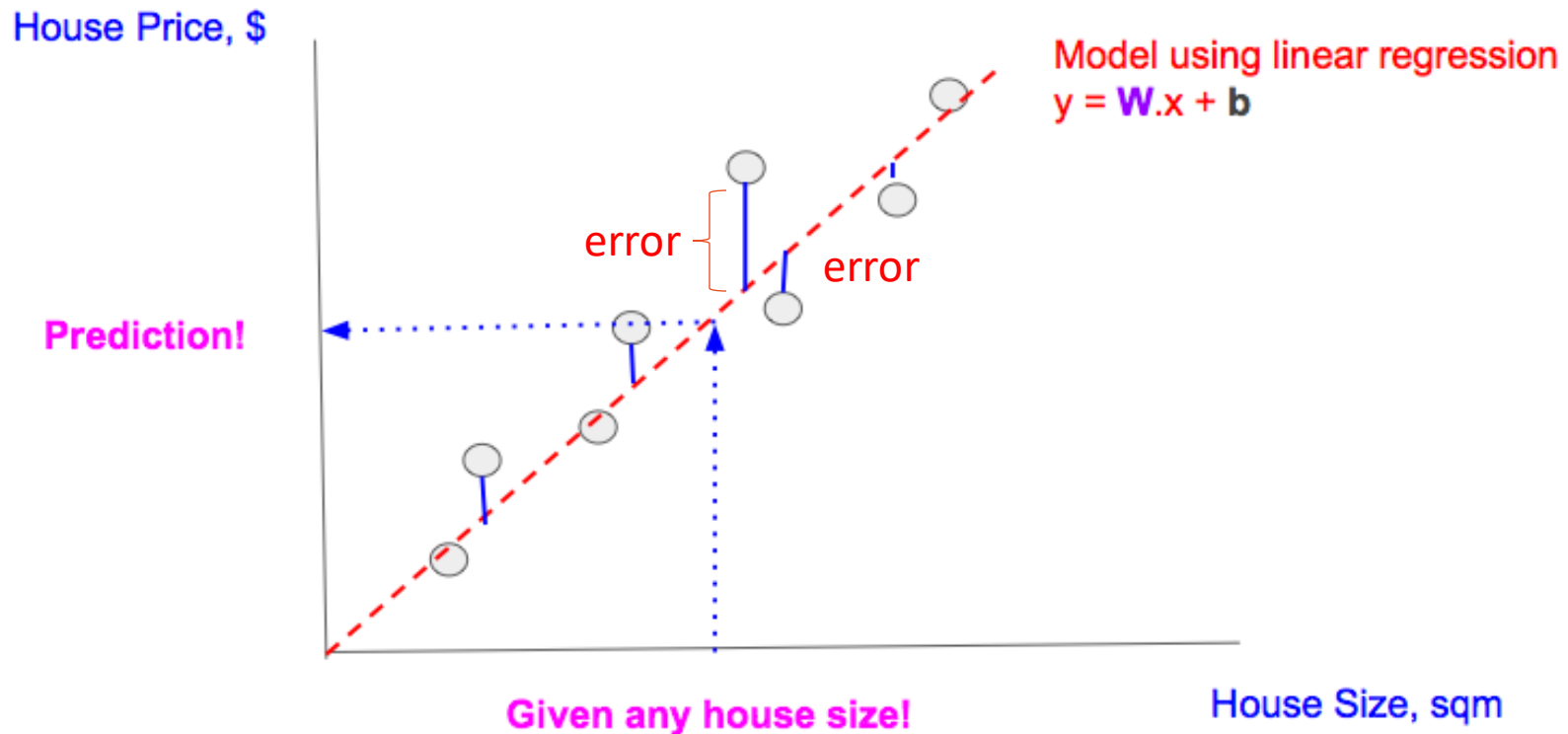
No.	sepal length Numeric	sepal width Numeric	petal length Numeric	petal width Numeric	class Nominal
48	4,6	3,2	1,4	0,2	Iris-setosa
49	5,3	3,7	1,5	0,2	Iris-setosa
50	5,0	3,3	1,4	0,2	Iris-setosa
51	7,0	3,2	4,7	1,4	Iris-versicolor
52	6,4	3,2	4,5	1,5	Iris-versicolor
53	6,9	3,1	4,9	1,5	Iris-versicolor
54	5,5	2,3	4,0	1,3	Iris-versicolor
55	6,5	2,8	4,6	1,5	Iris-versicolor



# P (Performance Measure) = Target (or Loss) Function

*“If you cannot measure it, you can not improve it.”*

Kelvin, Lord William Thomson (1824-1907)





# P (Performance Measure) = Target (or Loss) Function

*“If you cannot measure it, you can not improve it.”*

Kelvin, Lord William Thomson (1824-1907)

- Error rate
- Euclidean distance
- Log Probability
- Information theoretical measures
  - Mutual information
  - KL