Week 2. Introduction to Supervised Learning

Dr. Shuo Wang

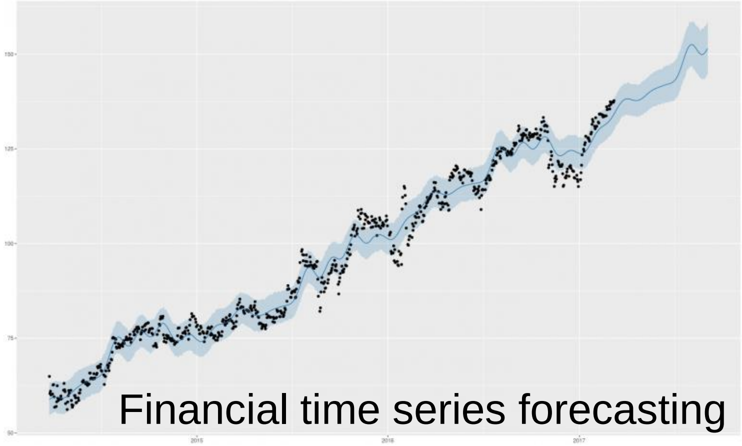
Overview

- Different forms of machine learning
- Supervised learning
- Regression and classification

Machine Learning Problems

Machine learning problems are those that require a model to be built automatically from data, e.g. to make classifications, estimations or predictions.

Speech recognition



Financial time series forecasting

Forms of Machine Learning

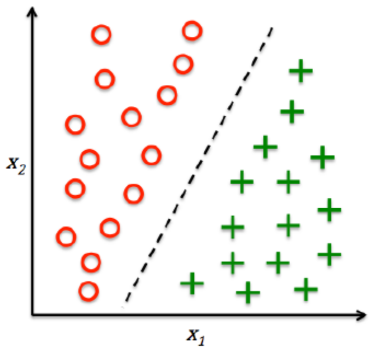
Three forms:

- Supervised learning
- Unsupervised learning
- Reinforcement learning

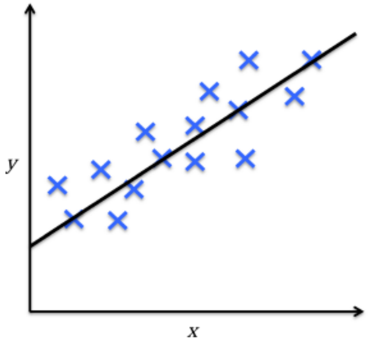
Supervised learning

- The most prevalent form
- Learning with a teacher
- Teacher: expected output, label, class, etc.
- Solve 2 types of problems:

classification



regression



Distinguishing Cats from Dogs

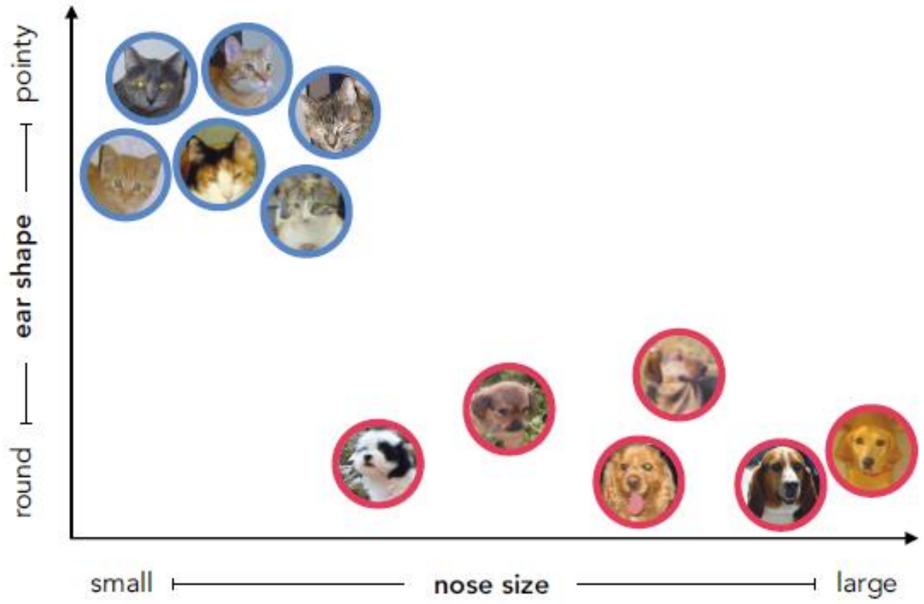


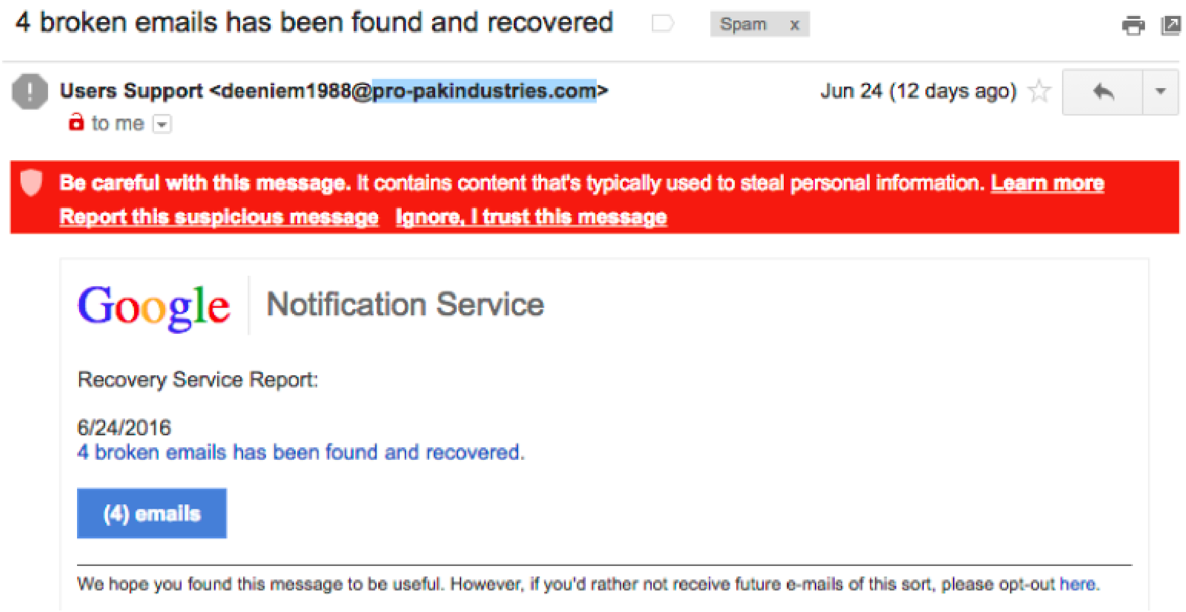
Fig 1.1/1.2 from Machine Learning Refined by Jeremy Watt

UNIVERSITY OF BIRMINGHAM

Example: Spam detection

- Input:
  Emails received

- Output:
  Spam, or not spam

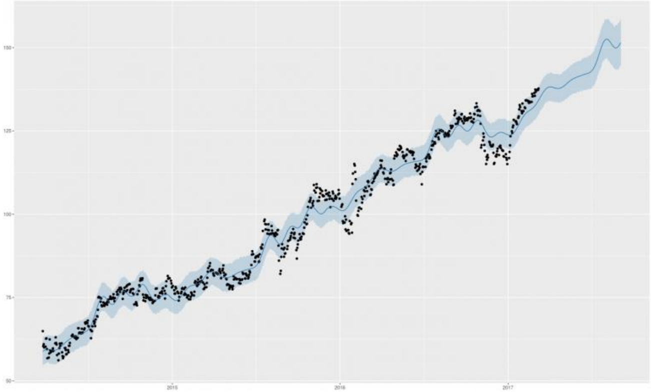


UNIVERSITY OF BIRMINGHAM

Example: Stock price prediction

- Input:
  Historical records of stock prices

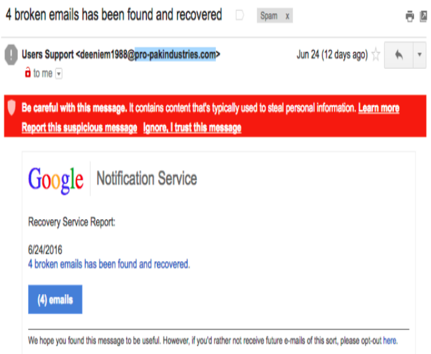
- Output:
  Next day's stock price

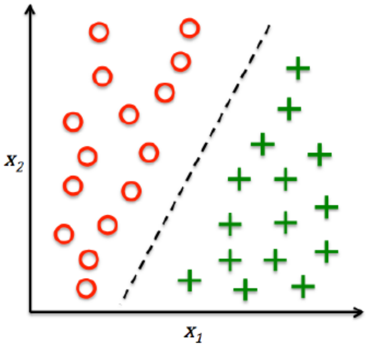


UNIVERSITY OF BIRMINGHAM

Types of supervised learning

Spam detection



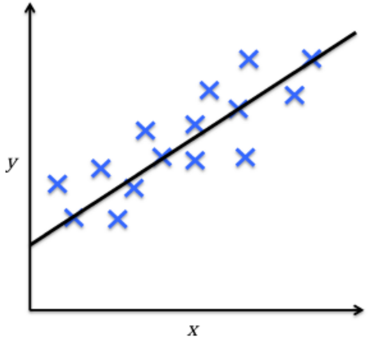


<u>Classification problem</u>

predict categorical class labels

e.g. the handwritten digit (multi-class)

Stock price prediction



<u>Regression problem</u>

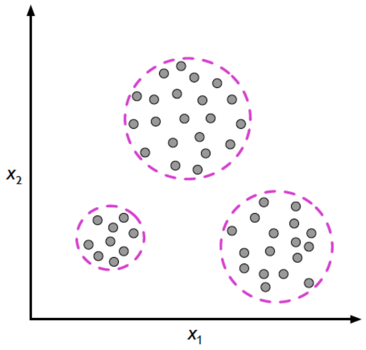
Prediction of a real value

e.g. students' grade scores

UNIVERSITY OF BIRMINGHAM

Unsupervised learning

- Learning without a teacher
- To find hidden structure/insights in data
- Clustering, e.g. product recommendation, sport strategy discovery

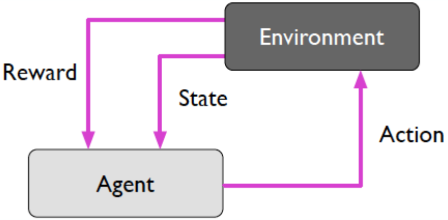


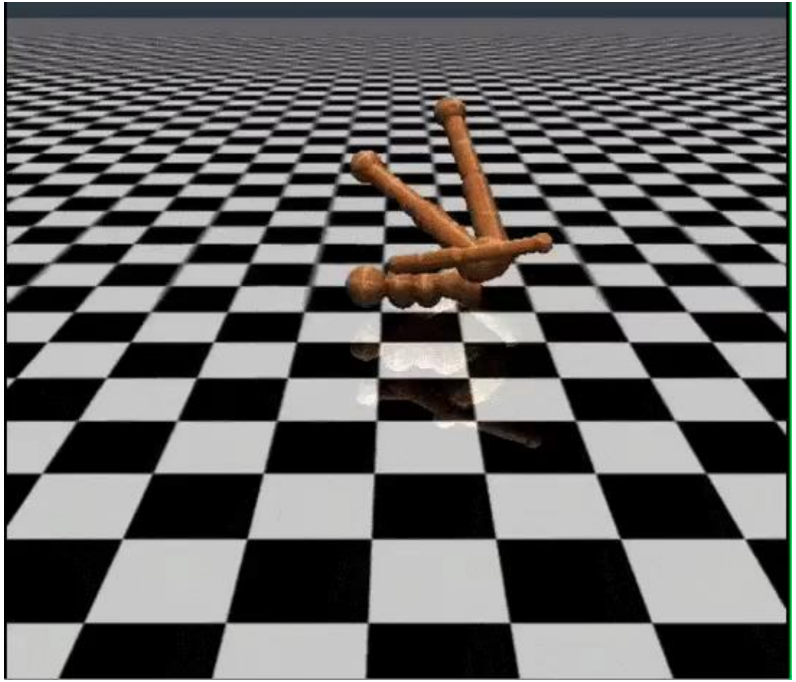
Clustering

RL

Reinforcement learning

- Learning with (delayed) feedback/reward
- Learn series of actions, e.g. chess, robots, …





UNIVERSITY OF
BIRMINGHAM

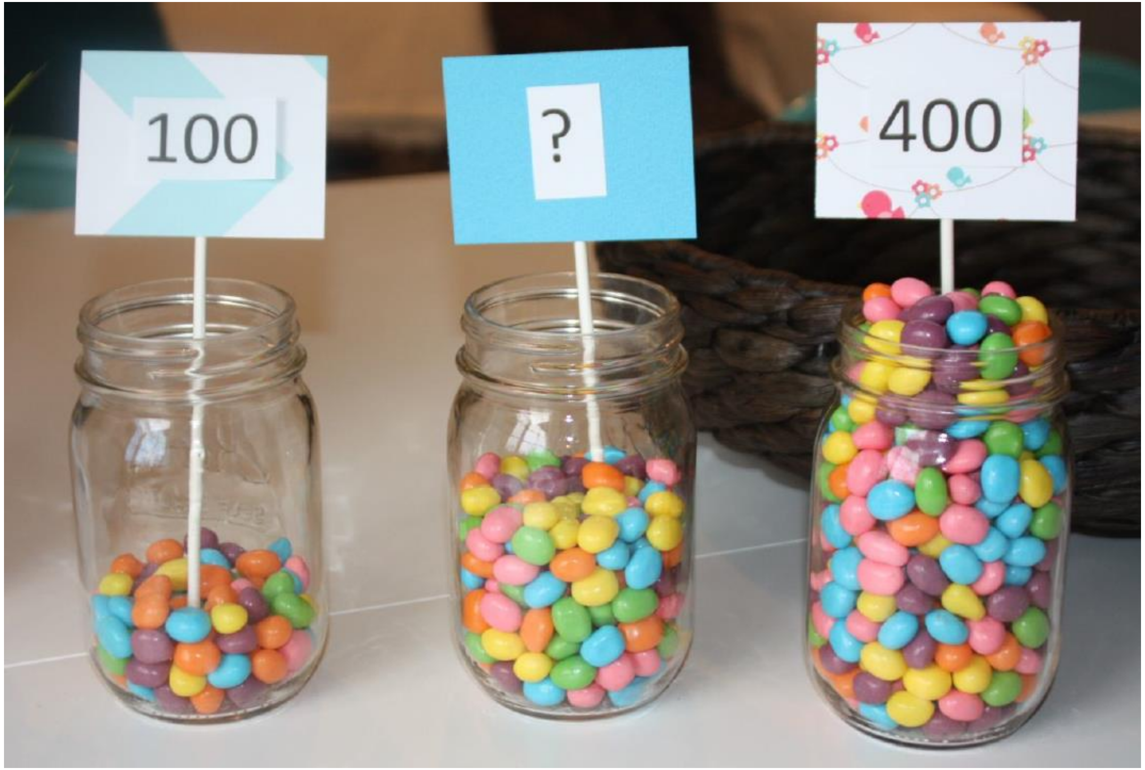
How does supervised learning work?

UNIVERSITY OF BIRMINGHAM

How does supervised learning work?

UNIVERSITY OF BIRMINGHAM

How does supervised learning work?



UNIVERSITY OF BIRMINGHAM

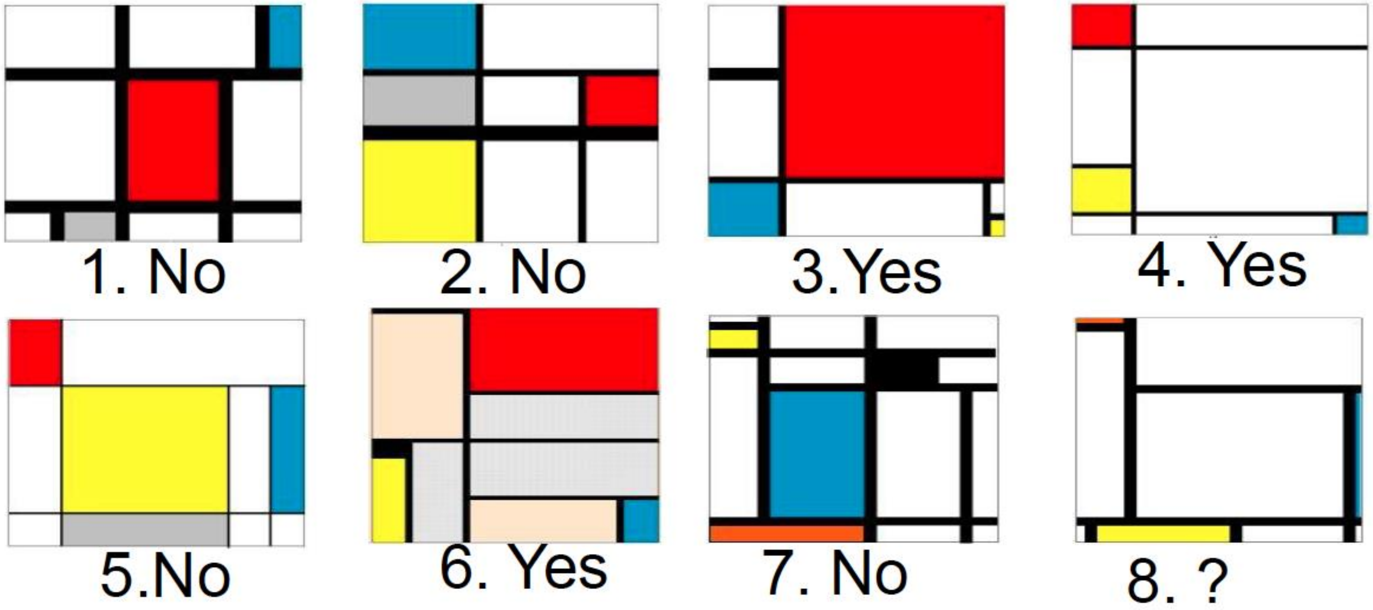
Formulate supervised learning

- Task:
  - Given some input x,
  - Predict an appropriate output y
- Goal: a function f such that $f(x) = y$

The learning process:

1) Have: examples of input-output pairs → training data
   $\left(x^{(1)}, y^{(1)}\right), \left(x^{(2)}, y^{(2)}\right), \ldots, \left(x^{(n)}, y^{(n)}\right)$
2) Supervised learning helps find a good f → training/modelling
3) Given a new input $x^{(n+1)}$, predict its output $y^{(n+1)}$ → prediction

Is painting 8 a genuine Mondrian?



UNIVERSITY OF
BIRMINGHAM

Annotated training data

Attributes

Labels

Examples

Number	Lines	Line types	Rectangles	Colours	Mondrian?
1	6	1	10	4	No
2	4	2	8	5	No
3	5	2	7	4	Yes
4	5	1	8	4	Yes
5	5	1	10	5	No
6	6	1	8	6	Yes
7	7	1	14	5	No

Painting 8

Number	Lines	Line types	Rectangles	Colours	Mondrian?
8	7	2	9	4	???

How quickly will your team complete a project?

(programming language)	(team expertise)	(estimated size)	...	(required effort)
Java	low	1000	...	10 p-month
C++	medium	2000	...	20 p-month
Java	high	2000	...	8 p-month
...	...	...	...	...

General notations we often use

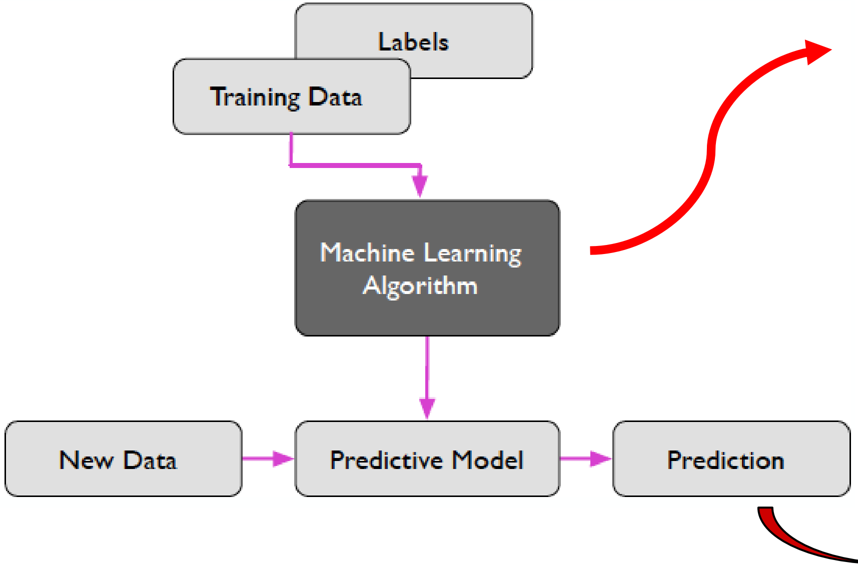
Lines	Line types	Rectangles	…	Mondrian?
$\boldsymbol{x}^{(1)}$				$y^{(1)}$
$\boldsymbol{x}^{(2)}$				$y^{(2)}$
$\boldsymbol{x}^{(3)}$				$y^{(3)}$
…				…

Vector notation:

$$\underbrace{\boldsymbol{x}^{(i)} = \left(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, \ldots, x_d^{(i)}\right)}_{\text{Attributes, d-dimensional}}$$

The input of the i-th example

Supervised learning workflow



Training phase:

$$(\boldsymbol{x}^{(1)}, y^{(1)}), (\boldsymbol{x}^{(2)}, y^{(2)}), \ldots, (\boldsymbol{x}^{(n)}, y^{(n)})$$

$$\Downarrow$$

$$f(\boldsymbol{x}) = y$$

Test/Use phase:

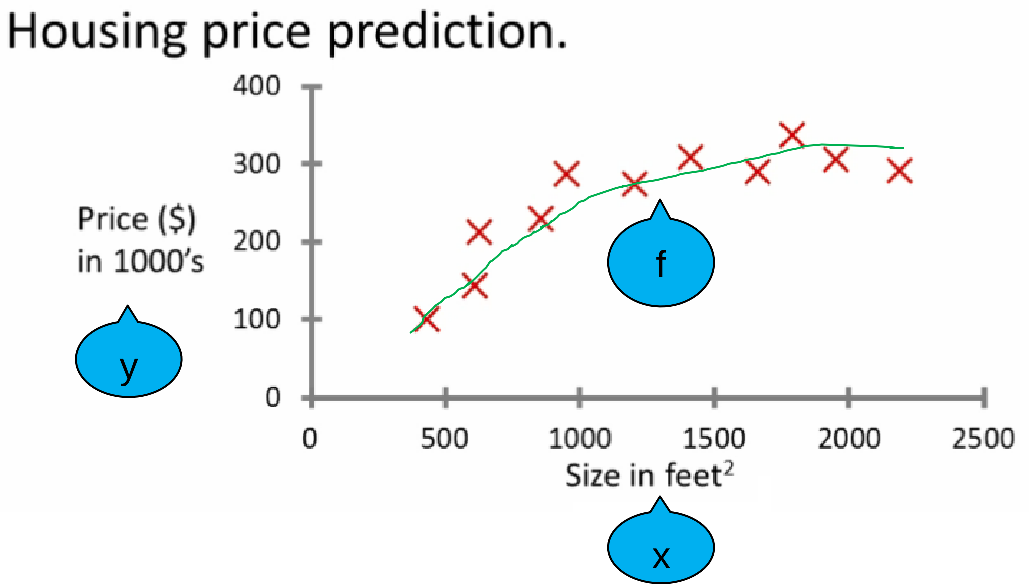
$$\boldsymbol{x} \rightarrow f(\boldsymbol{x}) \rightarrow y$$

UNIVERSITY OF BIRMINGHAM

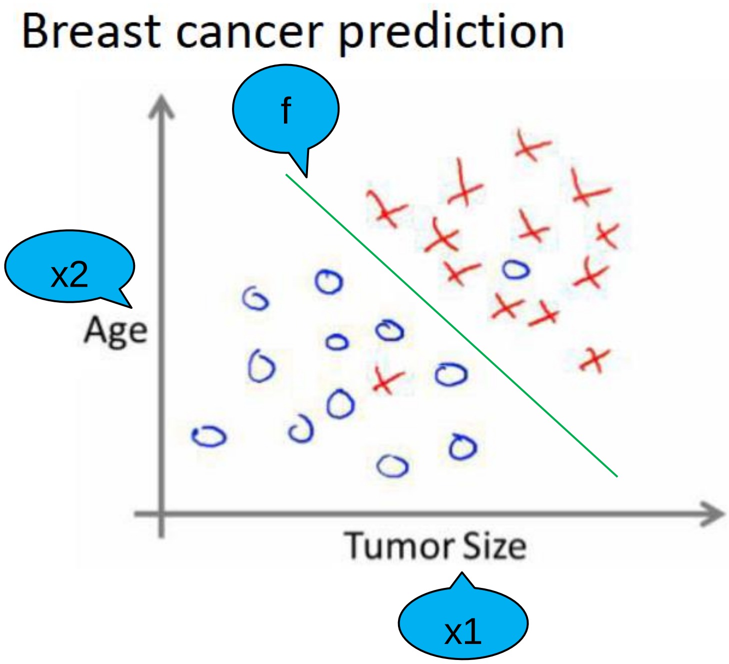
Source: Python machine learning, 2nd ed, Sebastian Raschka

Pictorially

Regression problem



Classification problem



UNIVERSITY OF BIRMINGHAM

Terminology in supervised learning

- Input = attribute(s) = feature(s) = independent variable(s)
- Output = target = response = dependent variable
- Function = hypothesis = predictor

Pause. Is this some magic?

So…

- There is an unknown function we are after.
- We are given the function values at n specific points only (training set)
- Is it really possible to find out the function values at other points?

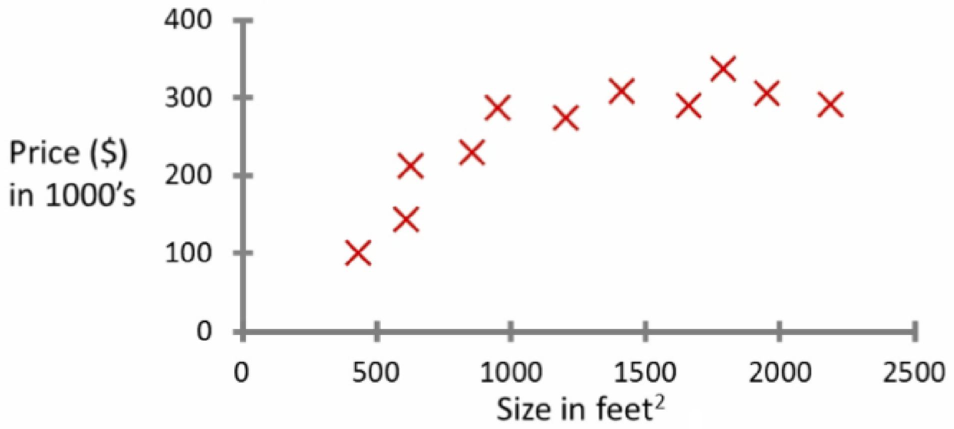
No!

- Unless we make the right **assumptions** about the unknown function.
- Each ML algorithm, implicitly or explicitly, makes assumptions.
- There is a zoo of ML algorithms, there is no best one.

How many predictors are there for each case?

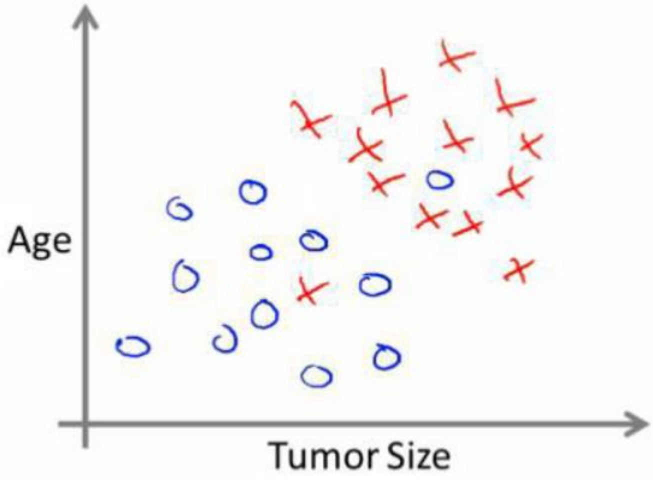
Regression problem

Housing price prediction.

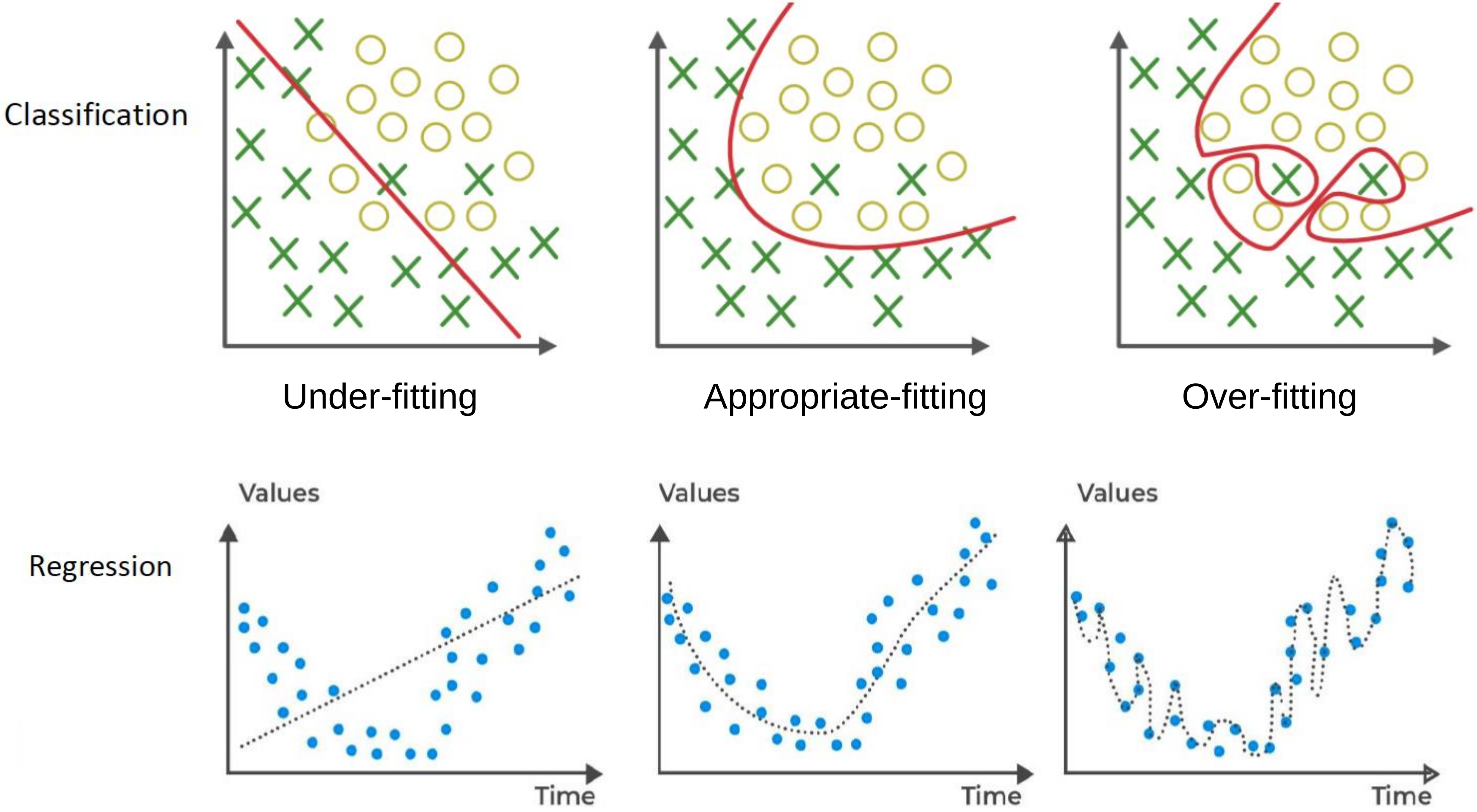


Classification problem

Breast cancer prediction



UNIVERSITY OF BIRMINGHAM



https://www.geeksforgeeks.org/regularization-in-machine-learning/

Overfitting and Underfitting

- Fitting the training data too well is BAD! Why?
- Remember the data you actually want to classify, or predict for, is not the same as the training data –so learning every irrelevant detail (noise) in a training data set will not help.
- Overfitting happens when the model is more complex than required.
- Underfitting happens when the model is simpler than required.

Applications of Supervised Learning

- Handwriting recognition
  - When you write an envelope, algorithms can automatically route envelopes through the post.
- Computer vision & graphics
  - When you go out during lockdown, object detection & vision tracking algorithms can automatically detect compliance with the rules.
- Bioinformatics
  - Algorithms can predict protein function from sequence.
- Human-computer interaction
  - Algorithms can recognize speech, gestures, intention.

Q/A

Teams Channel for Week2
Office Hours Tue/Thu 11am-12pm
See Canvas module homepage