



UNIVERSITY OF
BIRMINGHAM

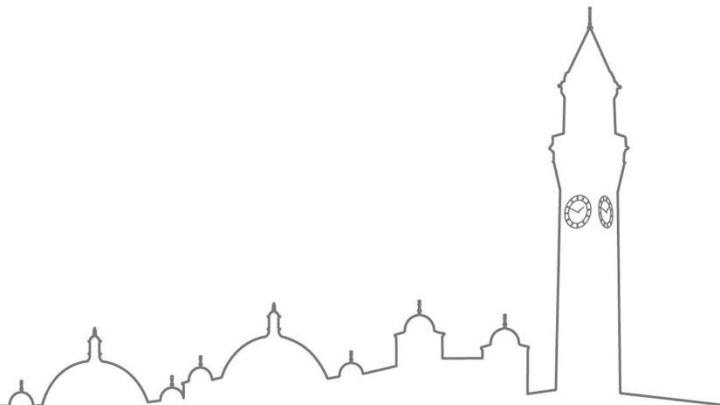
Week 2. Introduction to Supervised Learning

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.

Is this a cat?

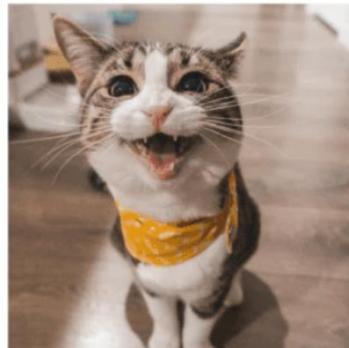
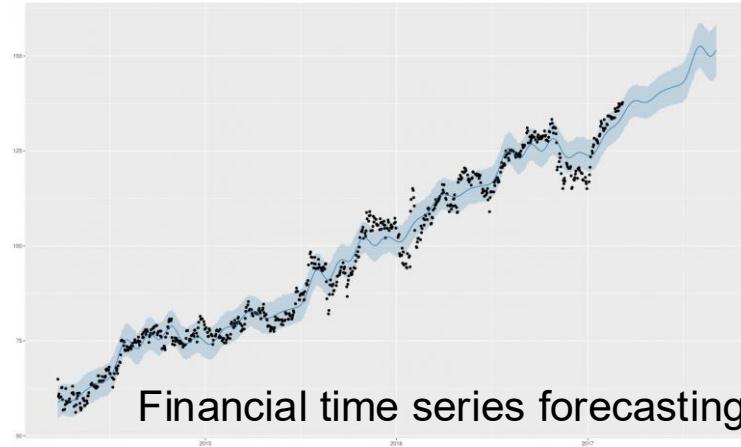


Image Classification



UNIVERSITY OF
BIRMINGHAM



Financial time series forecasting

Forms of Machine Learning

- Three forms (if we look at the *input* of the model), broadly speaking:
supervised learning, unsupervised learning, reinforcement learning
- Supervised learning
 - The most prevalent form
 - Learning with a teacher
 - Teacher: expected output, label, class, etc.



How does supervised learning work?

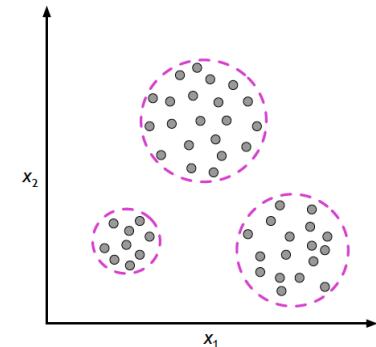


UNIVERSITY OF
BIRMINGHAM

Forms of Machine Learning

- Unsupervised learning

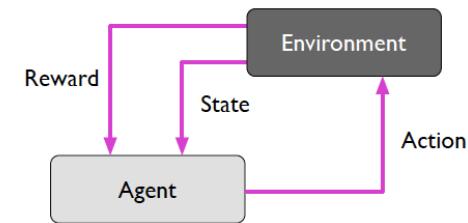
- Learning without a teacher
- To find hidden structure/insights in data
- Clustering, e.g. product recommendation, sport strategy discovery



Clustering

- Reinforcement learning

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



RL



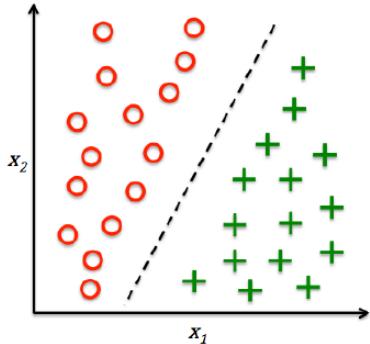
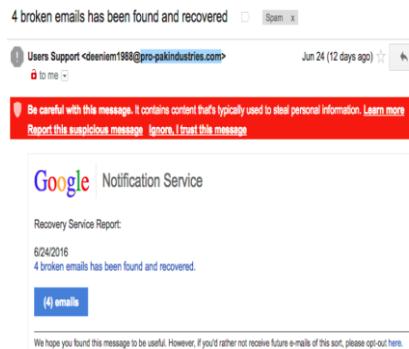
Forms of Machine Learning

- Three forms (if we look at the *input* of the model), broadly speaking: 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 (if we look at the *output* of the model): classification, regression problems

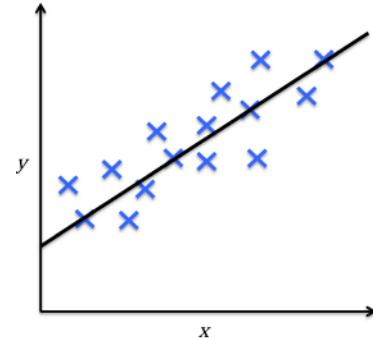


Types of supervised learning

Spam detection



Stock price prediction



Formulate supervised learning

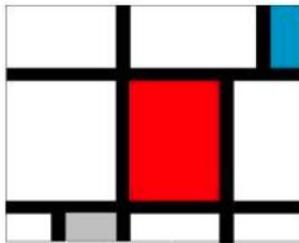
- Task:
 - Given some input x ,
 - Predict an appropriate output y
- Goal: a **function** f such that $f(x) = y$

The learning process:

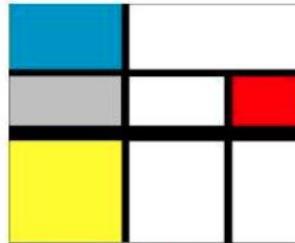
- 1) Have: examples of input-output pairs \rightarrow **training data**
 $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(n)}, y^{(n)})$
- 2) Supervised learning helps find a good $f \rightarrow$ **training/modelling**
- 3) Given a new input $x^{(n+1)}$, predict its output $y^{(n+1)} \rightarrow$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



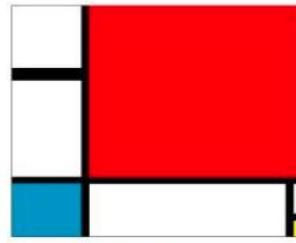
Is painting 8 a genuine Mondrian?



1. No



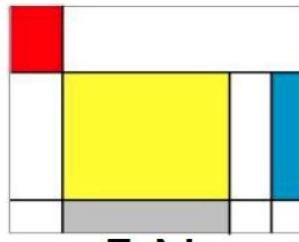
2. No



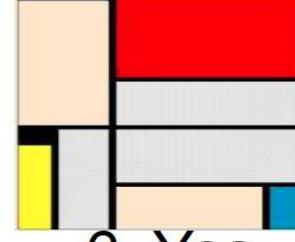
3. Yes



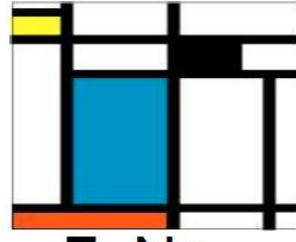
4. Yes



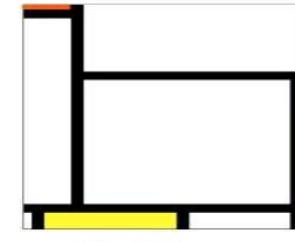
5. No



6. Yes



7. No



8. ?



Annotated
training data

Examples

Attributes

Labels

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	???

General notations we often use

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

Vector notation:

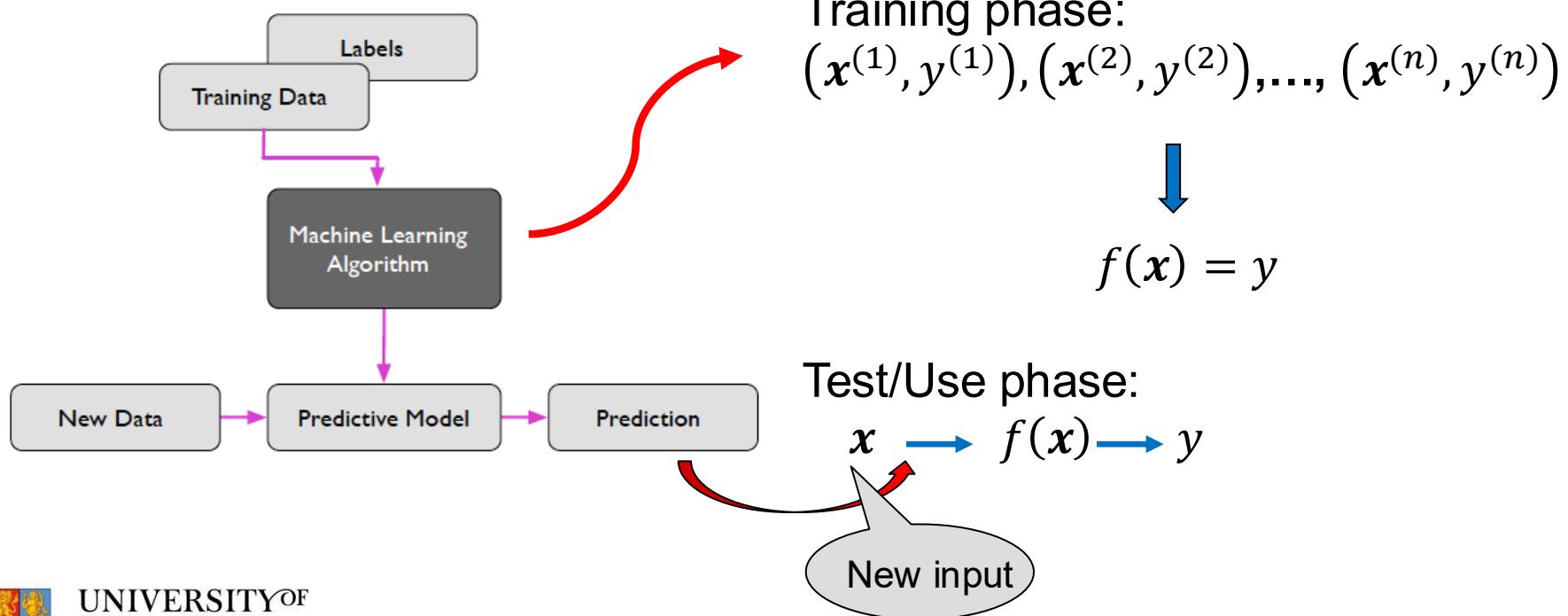
$$x^{(i)} = \left(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, \dots, x_d^{(i)} \right)$$

The input of the i-th example

Attributes, d-dimensional



Supervised learning workflow



Terminology in supervised learning

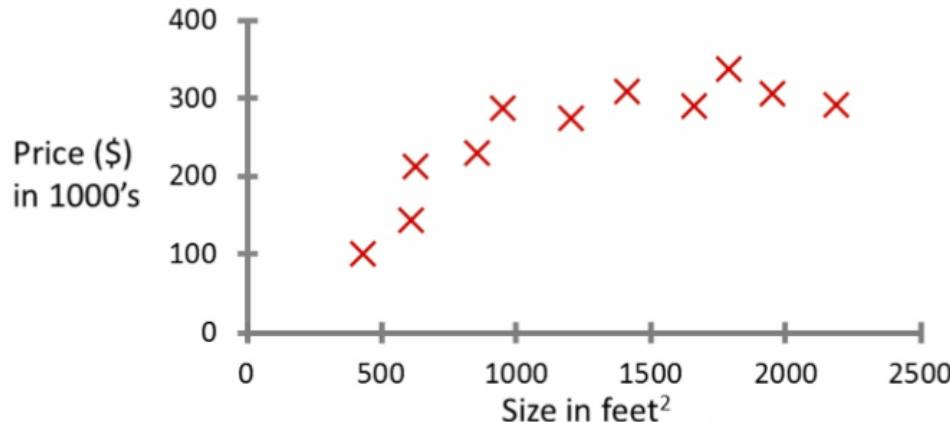
- Input = attribute(s) = feature(s) = independent variable(s)
- Output = target = response = dependent variable
- Function = hypothesis = predictor



Pictorially

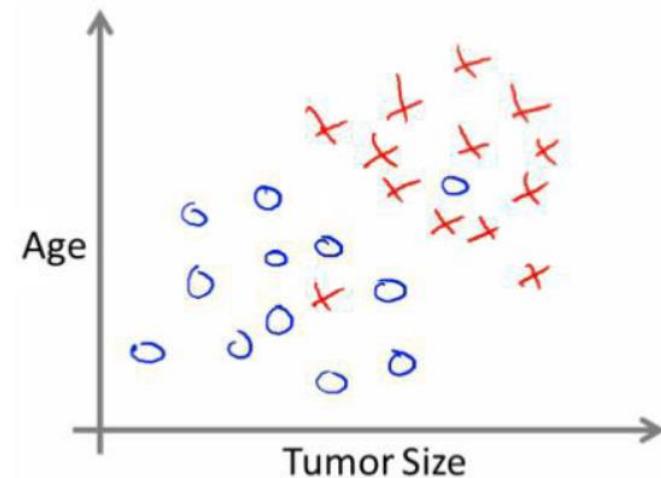
Regression problem

Housing price prediction.

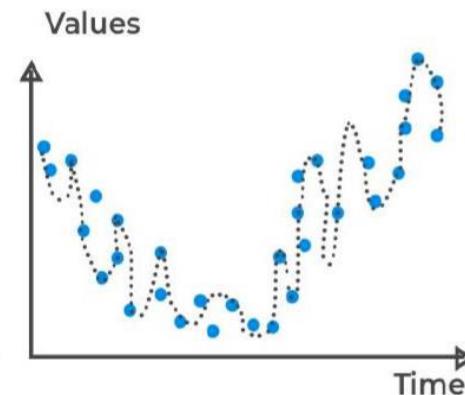
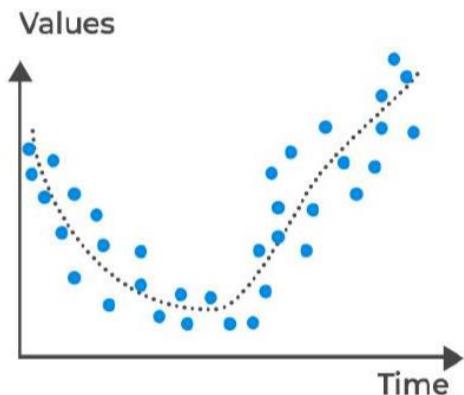
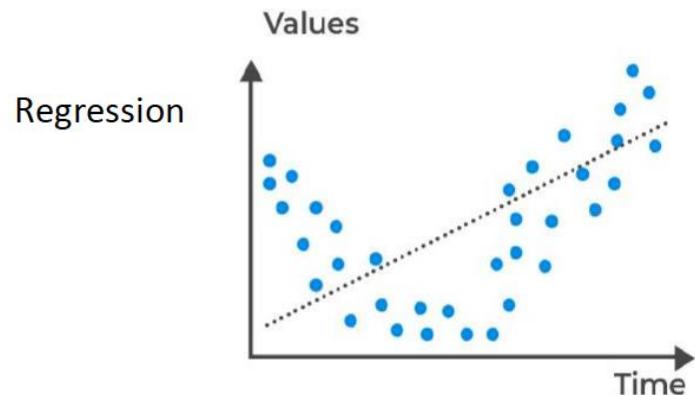
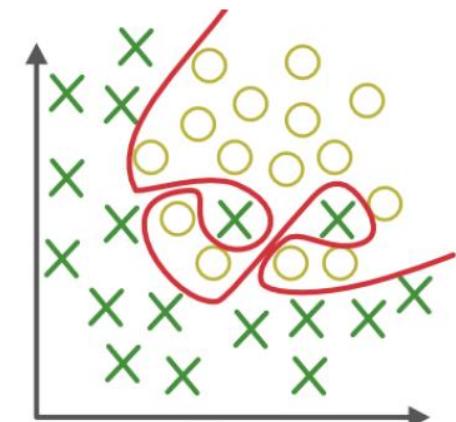
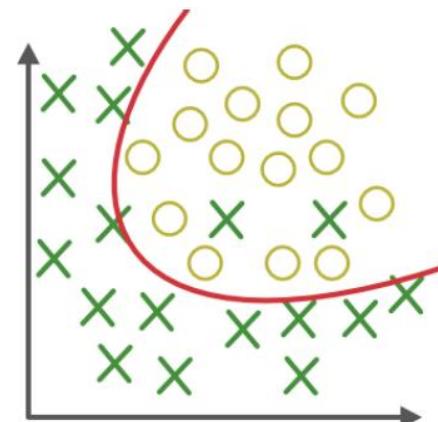
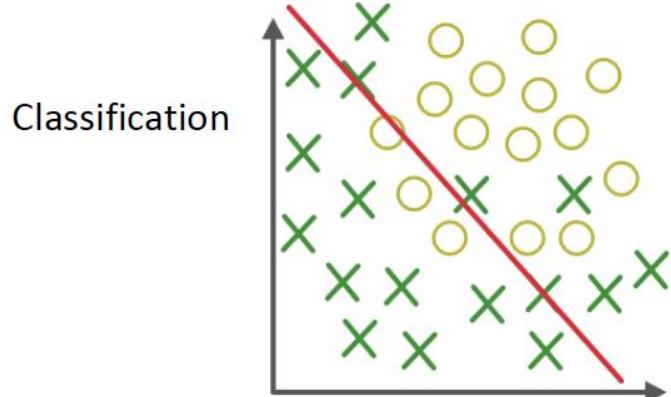


Classification problem

Breast cancer prediction



UNIVERSITY OF
BIRMINGHAM





UNIVERSITY OF
BIRMINGHAM

Q/A

Office Hour and Dropin Sessions
See Canvas module homepage

