

# LEARNING LAB, Part 2



## Machine Learning

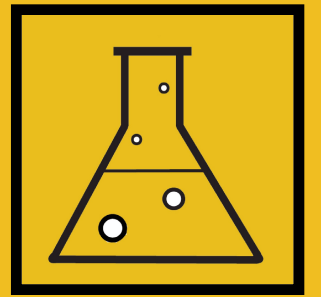
Machine Learning is a branch of Data Science which uses a whole suite of algorithms to understand data and make predictions from it.

**"Field of study that gives computers the ability to learn without being explicitly programmed".**



- ✓ Medical image analysis
- ✓ Driverless cars
- ✓ Voice and image recognition
- ✓ Artificial intelligence(?)

# Machine Learning



```
IF  
average (mostpixels)=Yellow  
THEN  
Lemon = True
```



## Why is it important?

Computers are great at following orders that have simple logical inputs or outputs. e.g.

Calculate the amount of tax I should pay...

There are certain problems that rule based programming **can't** deal with. e.g. What is

this a picture of?

**It would be enormously time consuming ,  
long winded and error prone to code  
rules to detect this!**

# Machine Learning



## Types of algorithms

Large array of algorithms available, depending on the problem you want to solve.

### Classification

#### Predict a class

Image recognition

Digit recognition

Spam email detection

### Regression

#### Predict a number

“How do house prices in Leeds vary with number of rooms, floors, neighbourhood”?

“How does the probability of diabetes correlate with diet, family history?”

# Machine Learning



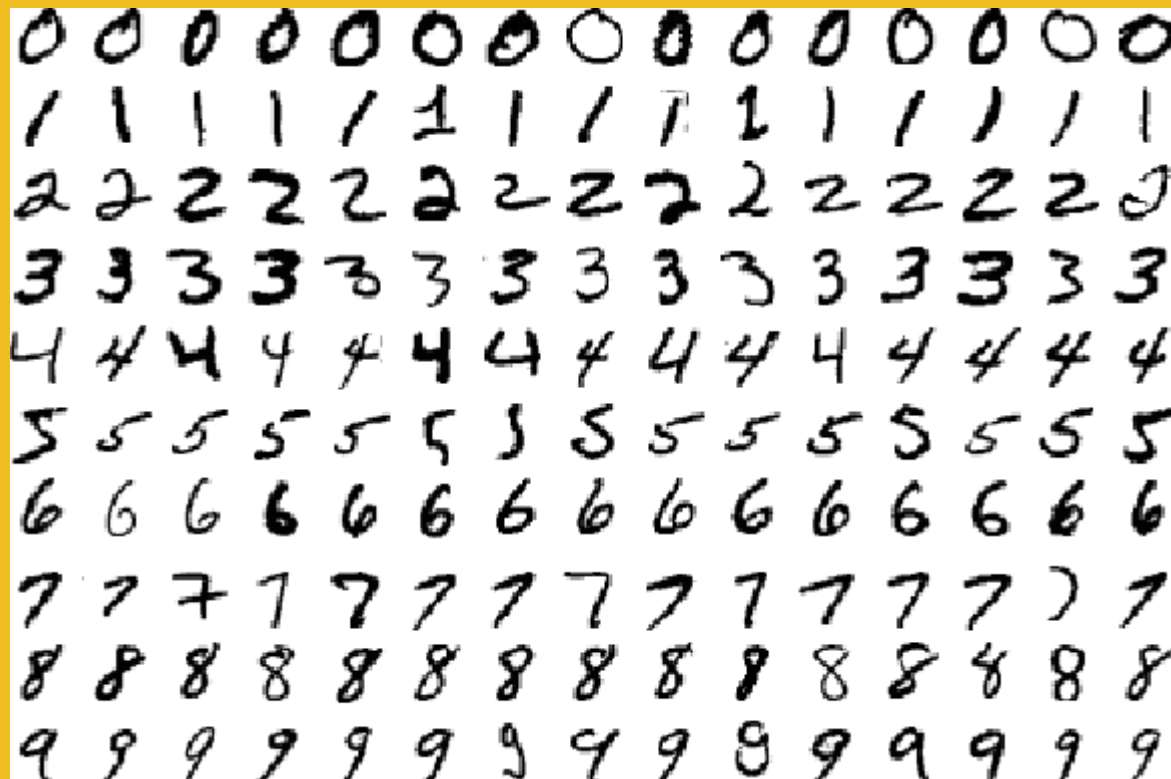
## Types of algorithms

Most of these are examples of supervised learning.

The dataset has the right answers within it, for the algorithm to learn from.

This is called the **training set**.

### Classification



### Regression

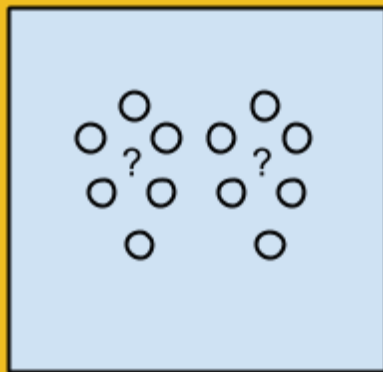
Rooms	Bathroom	Floorspace (m2)	Post Code	Price
3	2	600	LS1	£ 145,343.00
2	1	500	LS2	£ 130,321.00
1	1	400	LS1	£ 104,000.00
2	1	300	LS1	£ 134,000.00
2	1	300	LS2	£ 132,000.00

Source: Yan LeCun, MNIST

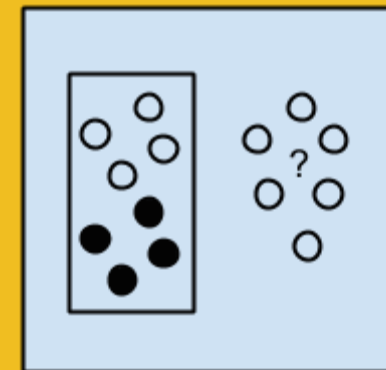
# Machine Learning



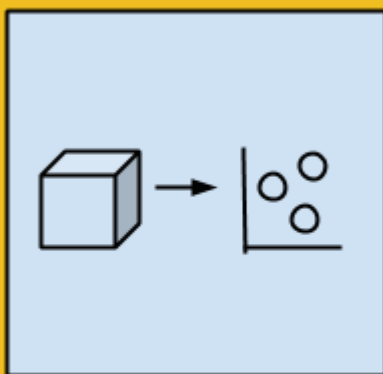
**There are lots of different algorithms – using them appropriately  
is key! ...**



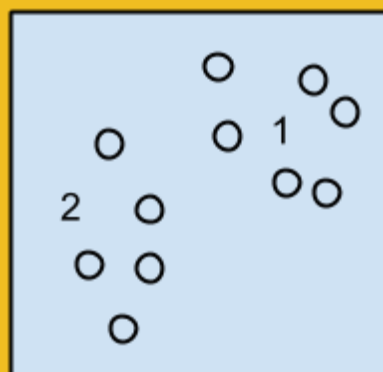
Unsupervised Learning  
Algorithms



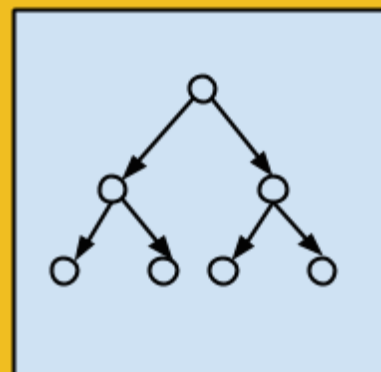
Supervised Learning  
Algorithms



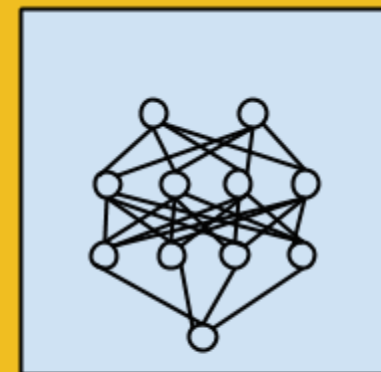
Dimensional Reduction  
Algorithms



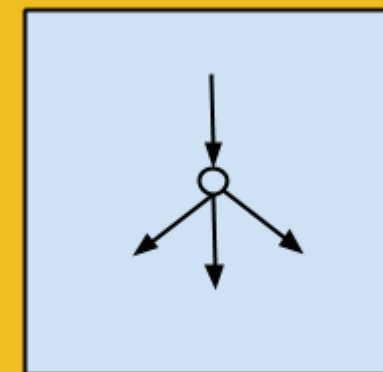
Clustering Algorithms



Decision Tree  
Algorithms



Deep Learning  
Algorithms



Artificial Neural Network  
Algorithms



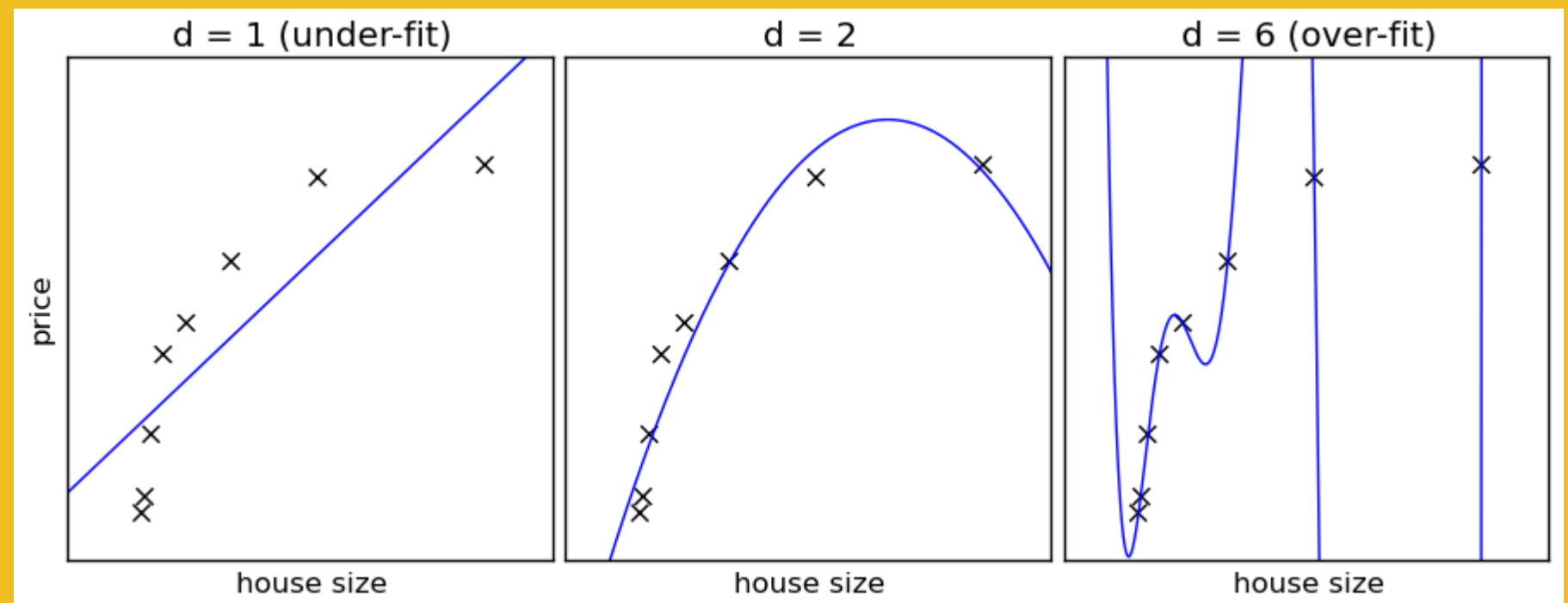
Bayesian Algorithms

# Machine Learning



**... As is training them appropriately!**

Linear regression example:

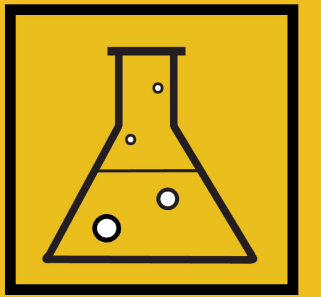


d=1 works poorly on training set and would likely work poorly on new, test data.

d=2 works well and is likely to generalize reasonably well, given data is within bounds.

d=6 Models 3 fits training data perfectly, but will fail on new data!

# Machine Learning



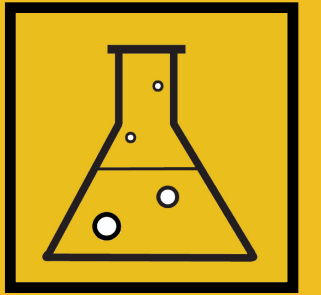
**DEMO**

## **Logistic Regression Classification**

Let's use Machine Learning to predict determine importance of factors in the development of diabetes.

This demo is in R, a powerful open source statistical programming tool that can seem a bit daunting. Feel free to follow along with the process or explore the code for yourself. The code is publicly available.

# Machine Learning



**Cool videos – shows what can be done!**

**[Drifting RC Car](#)**

**[OCR – early days](#)**

**[Real Car On Race Track](#)**



# Machine Learning



## Resources

[Wiki Complete Guide to Machine Learning](#)

[Andrew NG Machine Learning Course on Coursera \(enrolment ends today!\)](#)

[Microsoft ML Course on EDX](#)