The University Of Sheffield.

The University of Sheffield-EPSRC Winter School 2008

Mathematics for Data Modelling

Tony Dodd

t.j.dodd@shef.ac.uk

The University Of Sheffield.

Some Admin First

- Fire safety – continuously ringing bell
- Leave immediately by nearest fire exit and wait by St George's Church
- Coffee breaks/lunch
- Access to PCs/web and PC lab sessions
- Poster session
- Thursday dinner
- Accommodation ok?
- Any questions about the week?
- Questionnaires
- Payment/registration

The University Of Sheffield.

Aim

To provide researchers with an overview of data modelling.

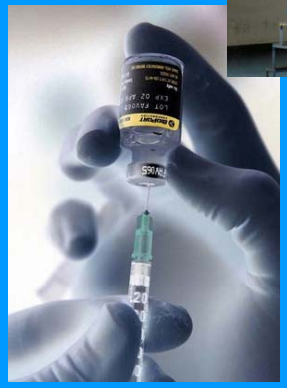
Why mathematics?

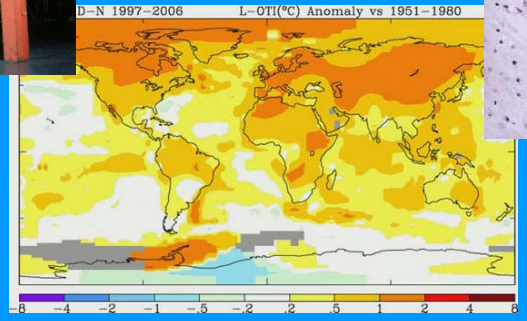
- Core to data modelling algorithms

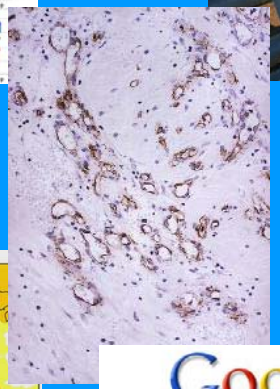
But, we will also cover how to model data well.

The University Of Sheffield.

Why data modelling?







The University Of Sheffield.

Why data modelling?

Increasingly important to success of many practical applications:

- Engineering
- Ecology
- Chemistry/chemical engineering
- Financial services
- Crime prevention
- Internet search
- Systems biology
- Medical diagnosis…

Engineers with experience in data modelling are in high demand!

The University Of Sheffield.

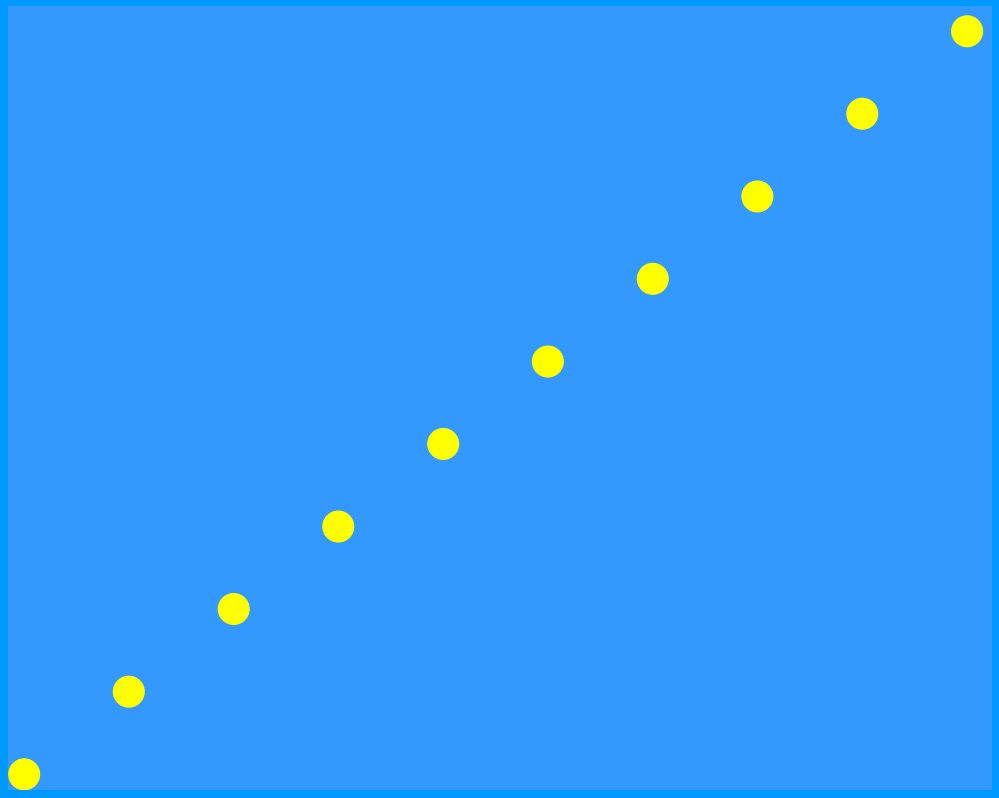
So what is data modelling?

Different things to different people.

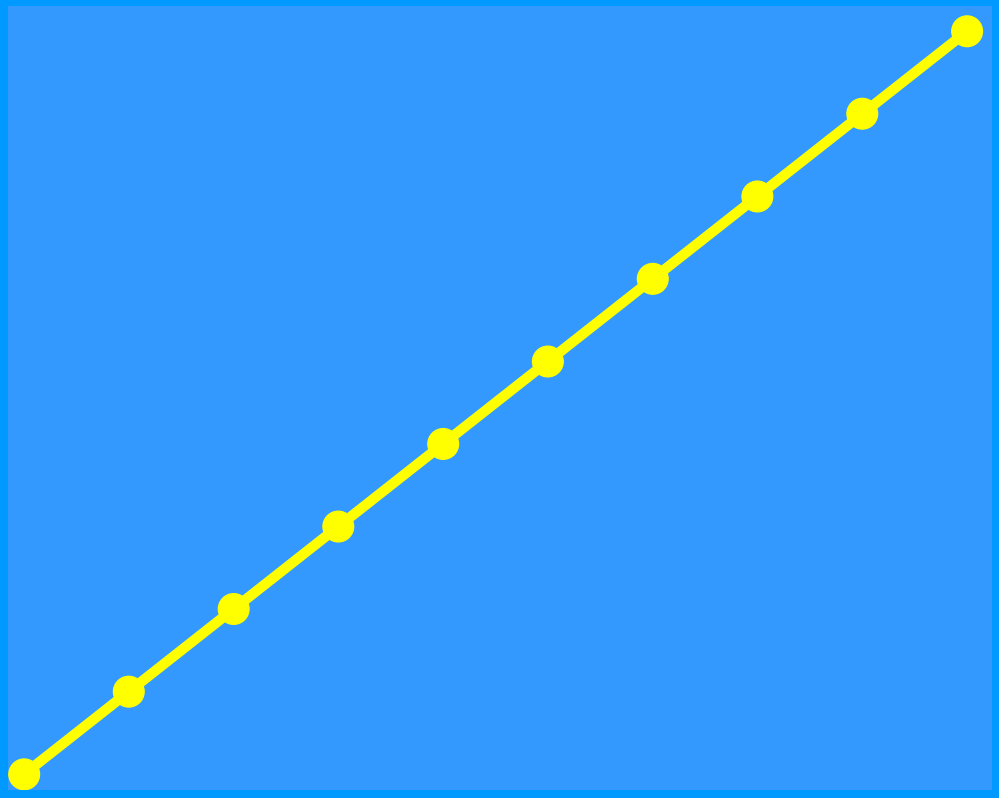
- Structuring and organising data.
- Physical models of data.
- Models to predict unseen data.

For this course consider some examples…

Example 1

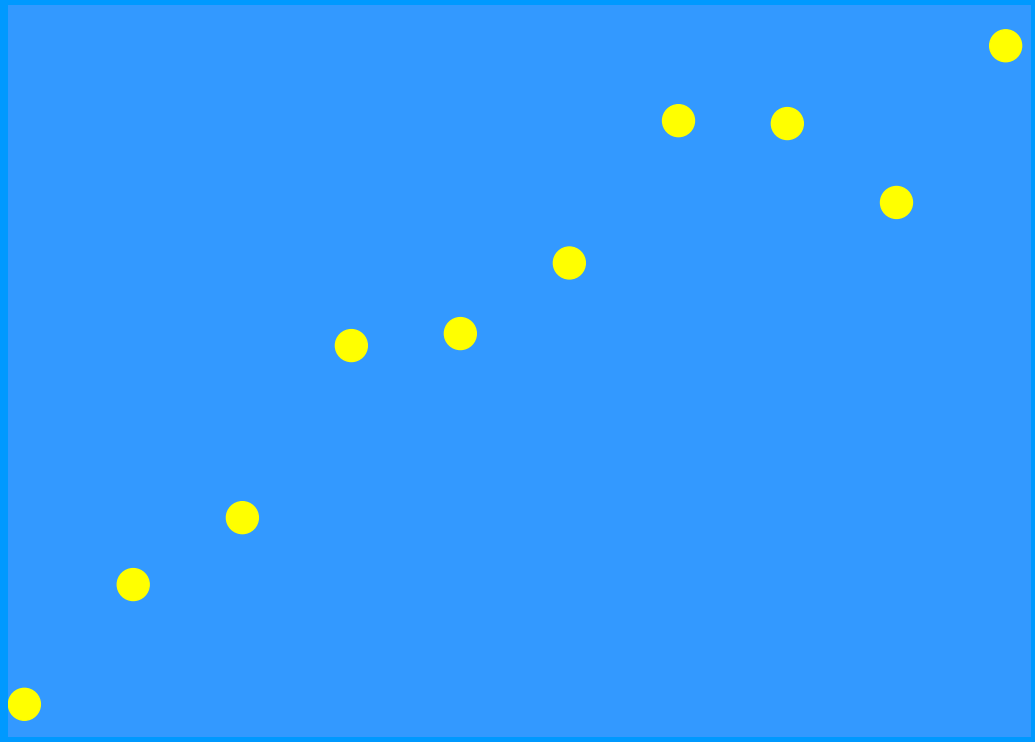


Example 1



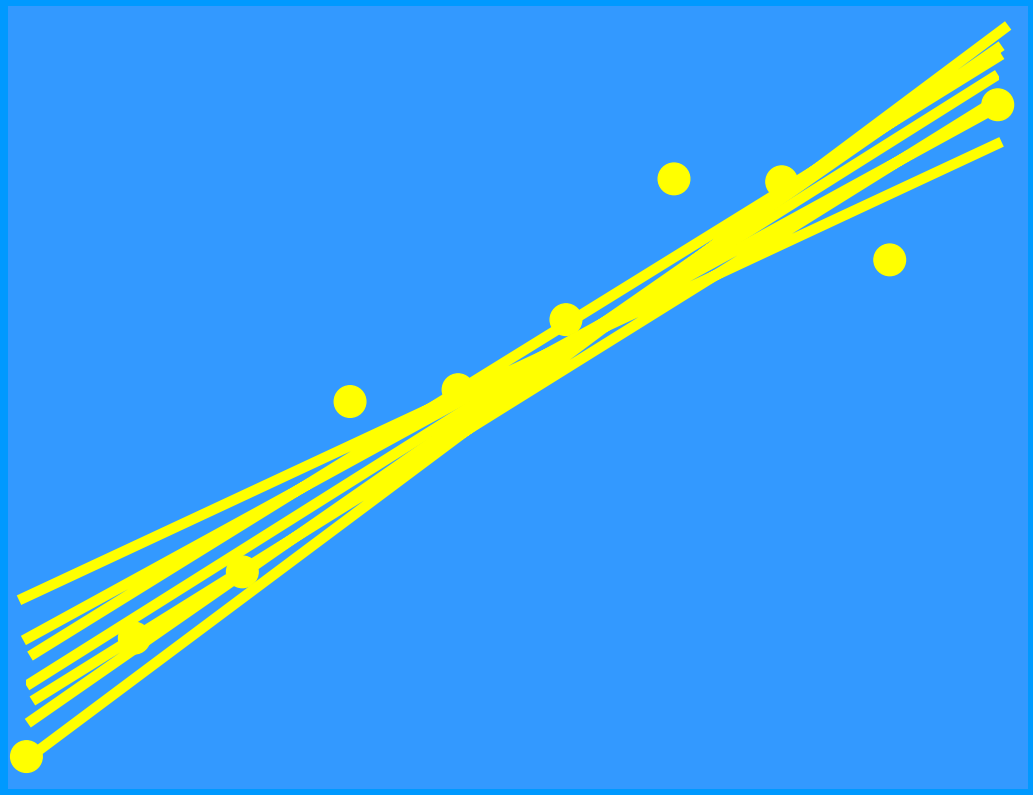
The University Of Sheffield.

Example 1



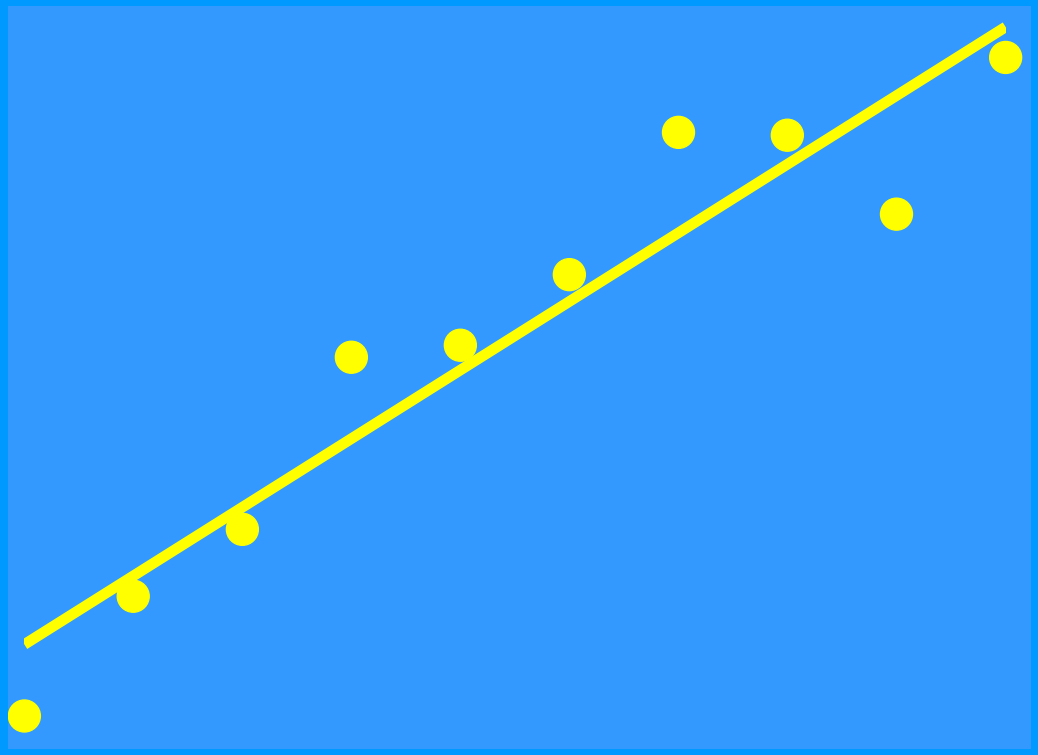
The University Of Sheffield.

Example 1



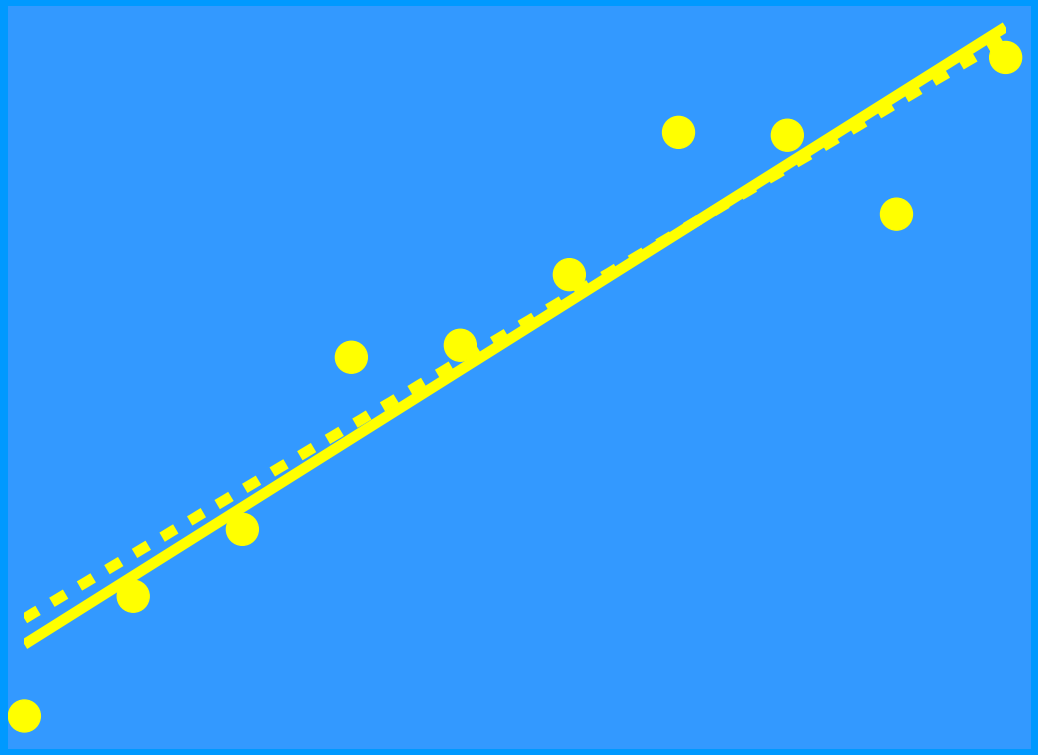
The University Of Sheffield.

Example 1



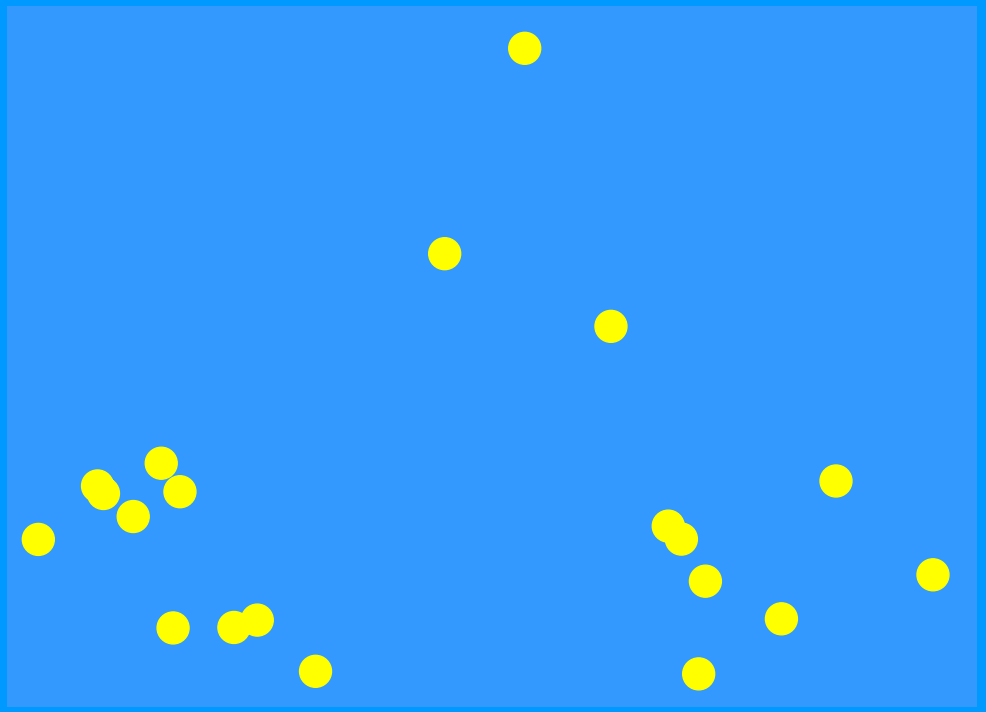
The University Of Sheffield.

Example 1



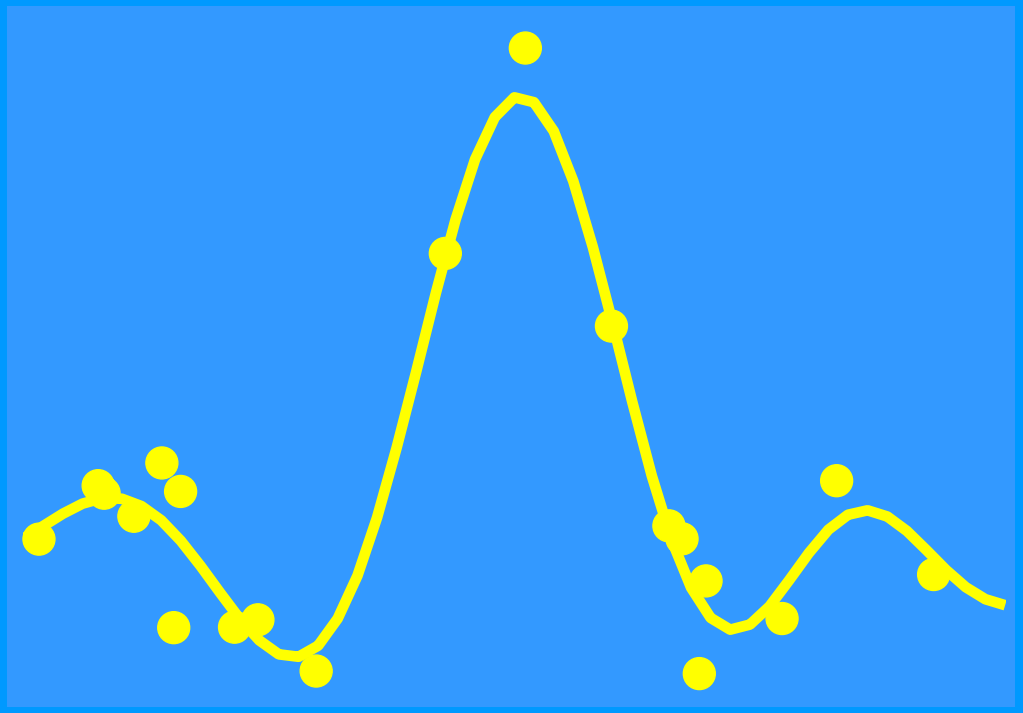
The University Of Sheffield.

Example 2



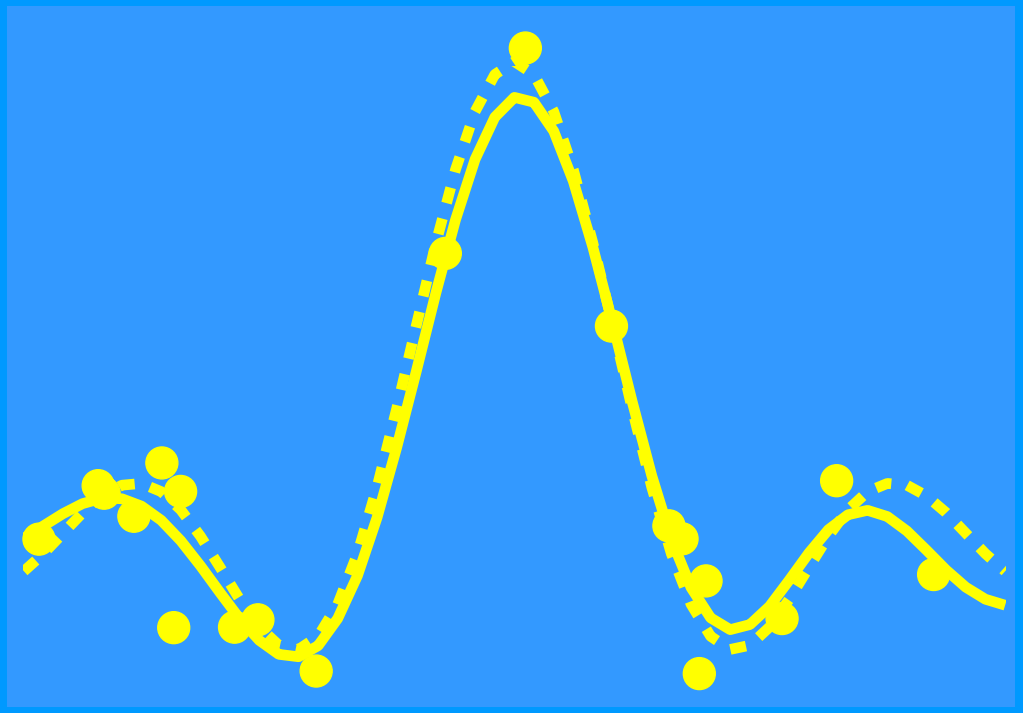
The University Of Sheffield.

Example 2



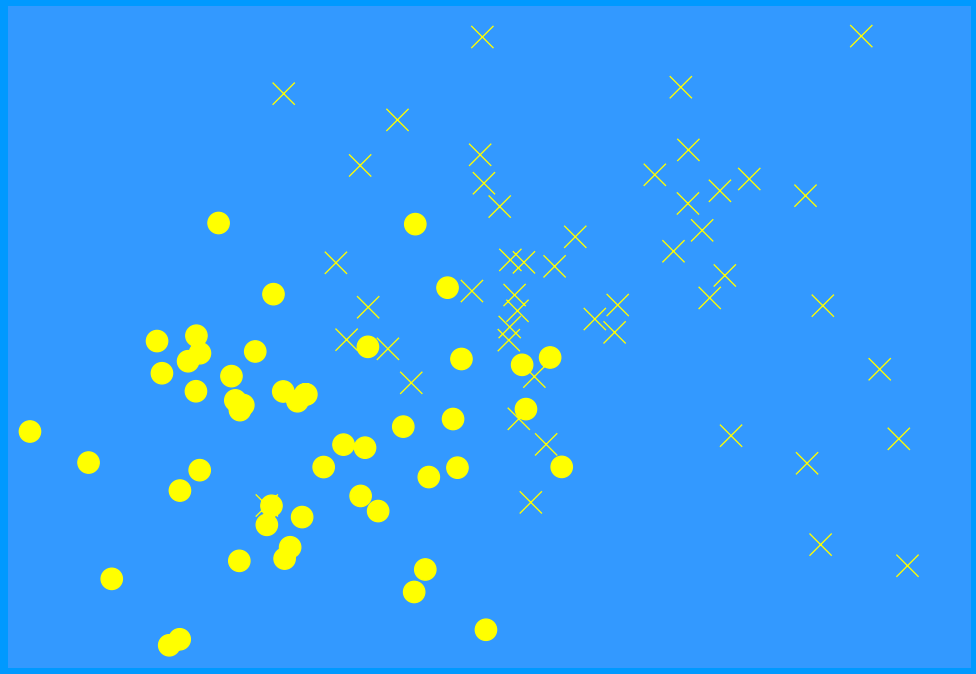
The University Of Sheffield.

Example 2



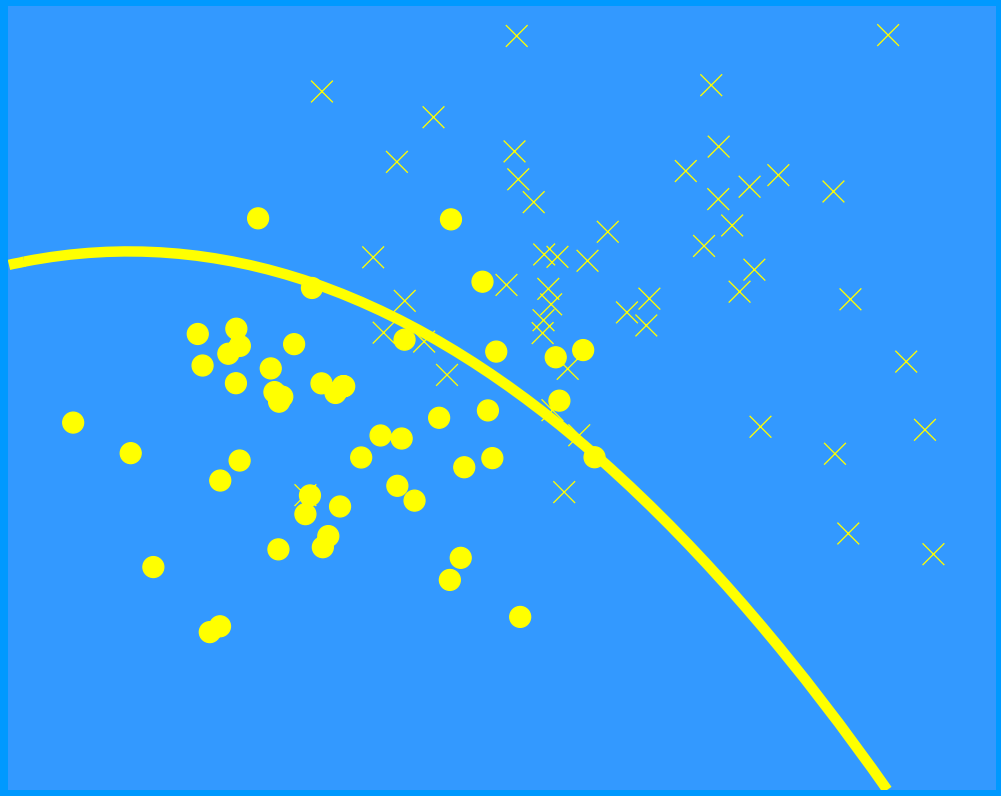
The University Of Sheffield.

Example 3



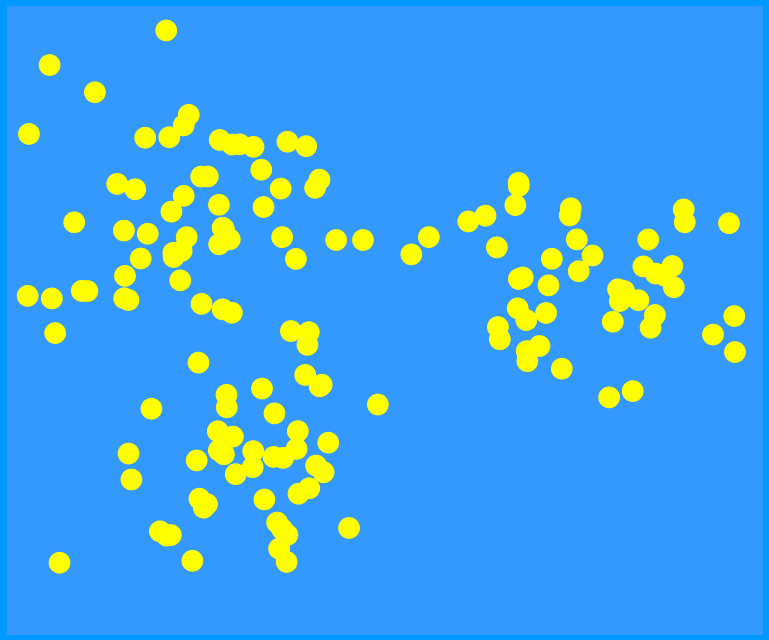
The University Of Sheffield.

Example 3



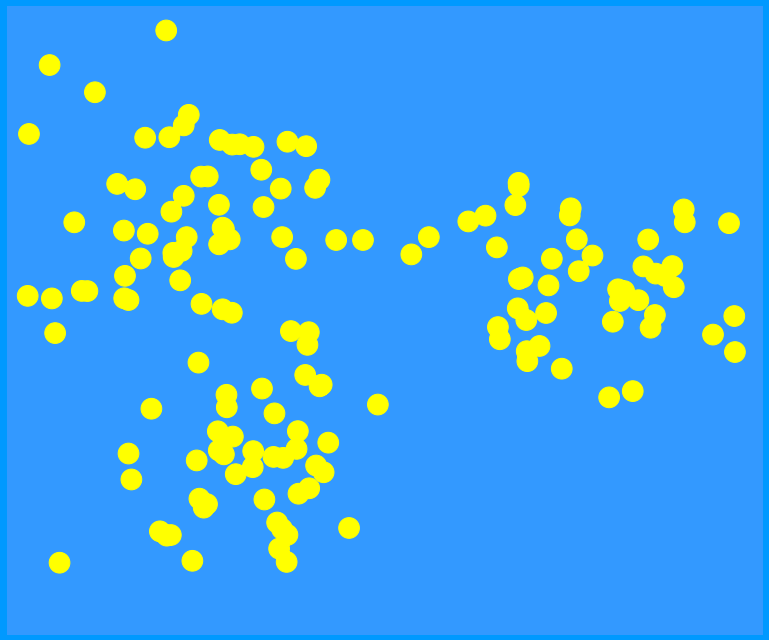
The University Of Sheffield.

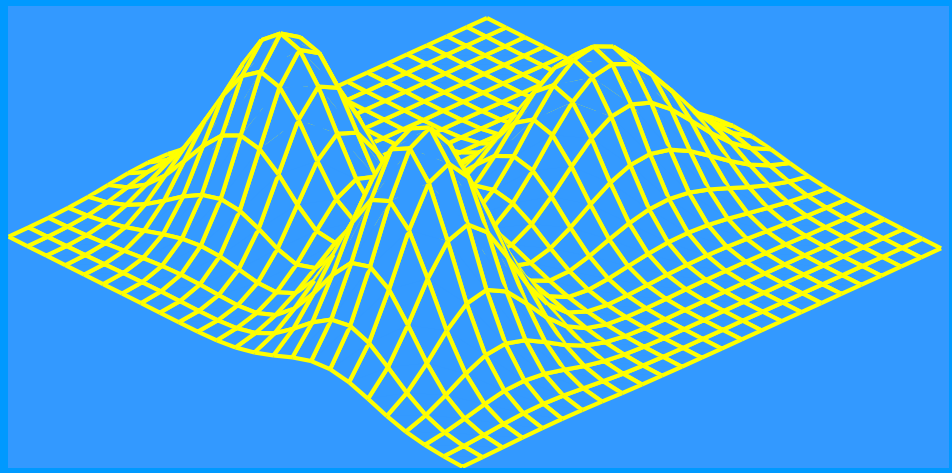
Example 4



The University Of Sheffield.

Example 4





The University Of Sheffield.

Data modelling problems

- Examples 1,2 – regression/curve fitting.
- Example 3 – classification/pattern recognition.
- Example 4 – density estimation.

This course - where do you put the line?

The University Of Sheffield.

Different types of learning

Supervised vs unsupervised

- Do you have target data?
- Learning with/without a teacher

Batch, incremental, sequential, online…

- Are all the data available initially?
- Are the data processed one at a time?

The University Of Sheffield.

The Winter School

- Day 1 – introduction, linear models, neural networks, how to model data well
- Day 2 – kernel methods, support vector machines
- Day 3 – unsupervised/semi-supervised
- Day 4 – Bayesian methods
- Day 5 – applications

The University Of Sheffield.

Notation

Inputs $\{x_i\}_{i=1}^N$, $x \in R^d$

Input variables $x_i = \left[x_{i,1}, x_{i,2}, \cdots, x_{i,m} \right]^T$

Outputs $\{y_i\}_{i=1}^N$ $y \in R$ regression

$y \in \{0,1\}$ classification

Targets $\left\{ z_i \right\}_{i=1}^N$ Possible values as per y

The University Of Sheffield.

Basic problem

Given

$$y = f(x)$$

$$z = y + e$$

where e is noise.

Estimate $\hat{f}$ from $\{x_i, z_i\}_{i=1}^{N}$.

Density estimation requires a more complicated notation – given as required.

The University Of Sheffield.

Finally…

- Ask questions.
- The course is for you.
- Use the breaks to network and discuss your work.
- Notes will be available at

http://www.datamodelling.group.shef.ac.uk/winterschool2008/lectures.php

- Videoed as well!