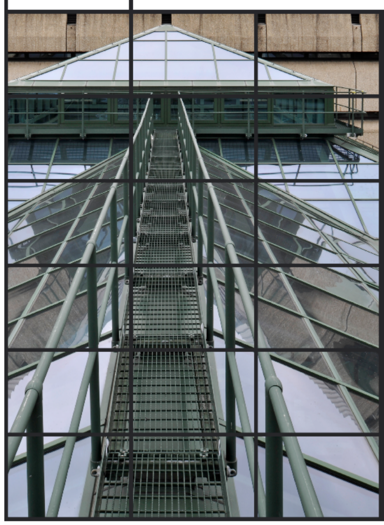
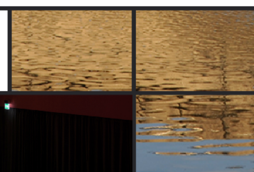
Introduction to AI





Mark Bell

30th March 2021

THE NATIONAL ARCHIVES

Lesson Objectives

Questions

What are Artificial Intelligence (AI) and Machine Learning (ML)?

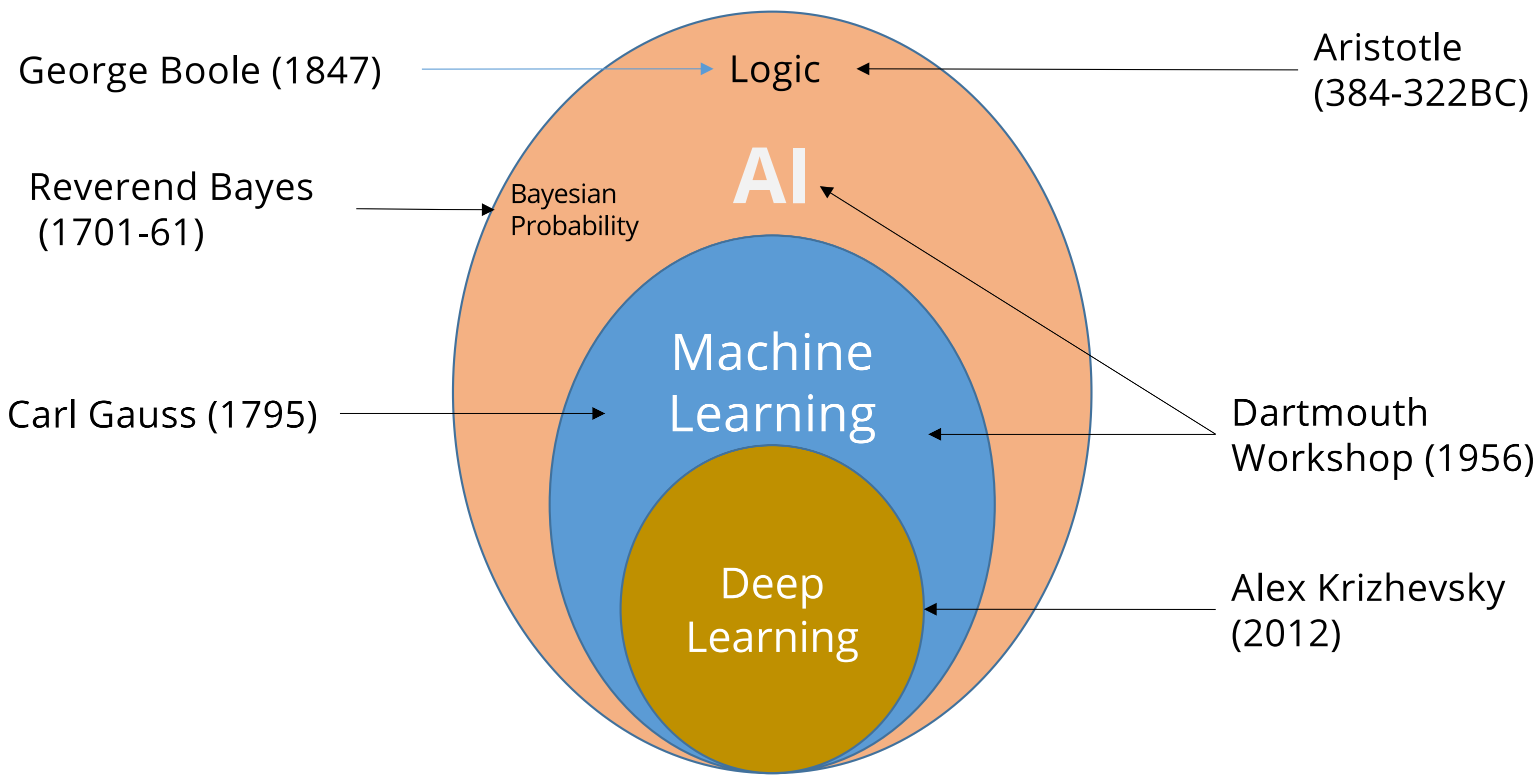
Objectives

Understand Machine Learning as a subfield of AI and name two others

Name four types of machine learning and describe the difference between supervised and unsupervised learning

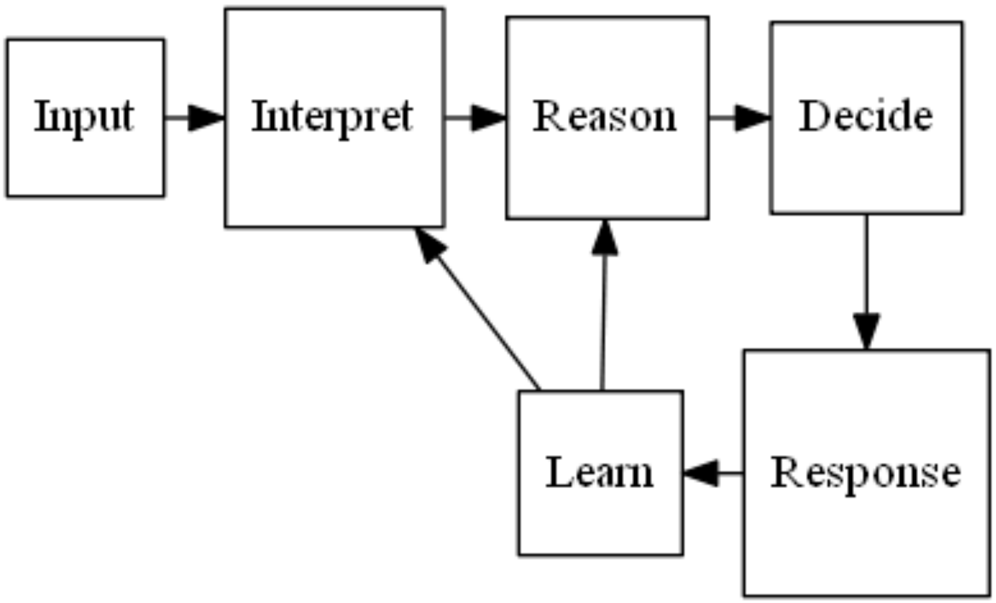
Describe the difference between a model of data, and a model trained on data

AI and Machine Learning



THE NATIONAL ARCHIVES

The components of an AI system



THE
NATIONAL
ARCHIVES

Why do we need machine learning?

With logic we can write rules to reason about data or make decisions.

- “**It is** raining therefore I will carry an umbrella”.

Logical rules are based on things being True or False but the world is not so clear cut. Probability lets us add doubt and uncertainty:

- “**It might** rain today, should I take an umbrella?”.

With logical rules and probability we can solve quite complex tasks, but there are limits:

- “Which paintings in our collection have umbrellas in them?”.

THE NATIONAL ARCHIVES

What is an umbrella?

Take a moment to think of how you would describe an umbrella using a set of rules:

What colour is it?
What shape is it?
How big is it?
What difference does it make if it is up or down?
Is a parasol an umbrella?

Describing something as intuitively simple as an umbrella is difficult because although we have a rough conceptual idea there isn't a fixed physical description. Even if you break it into component parts you still need to define them - what is a handle? what is a canopy? It becomes even more complicated if you try to specify the description in a way that a computer can interpret. Thankfully Machine Learning can come to our rescue.

What is Machine Learning?

Machine Learning is a set of technologies and methods for finding rules when they are too complex to define. There are four types of machine learning:

- Supervised: learning by example
- Unsupervised: puts data into groups without guidance
- Semi-supervised: a combination of supervised and unsupervised
- Reinforcement: learns by interacting with its environment

The fourth category has grabbed a lot of media attention as it is the basis for AlphaGo and driverless cars. This lesson will concentrate on Supervised and Unsupervised learning.

Activity

Which of the following do you think would use Machine Learning?

- a) Counting the number of people in a museum using information from entry and exit barriers?
- b) A search system that looks for images similar to a user submitted sketch.
- c) A system that recommends library books based on what other users have ordered.
- d) A queueing system that spreads people evenly between 5 ticket booths
- e) A program which extracts names from documents by finding all capitalised words and checking them against a list of known names
- f) A system which turns digitised handwritten documents into searchable text
- g) A robot which cleans the vases in a museum without bumping into them

THE NATIONAL ARCHIVES

Activity

Which of the following do you think would use Machine Learning?

- a) Counting the number of people in a museum using information from entry and exit barriers?
- b) **A search system that looks for images similar to a user submitted sketch.**
- c**) A system that recommends library books based on what other users have ordered.**
- d) A queueing system that spreads people evenly between 5 ticket booths
- e) A program which extracts names from documents by finding all capitalised words and checking them against a list of known names
- f) **A system which turns digitised handwritten documents into searchable text**
- g) **A robot which cleans the vases in a museum without bumping into them**

THE NATIONAL ARCHIVES

The two tasks of machine learning

Prediction:

- Generally numerical data
  - How much will temperature control in the archive cost if the summer is hot?
  - How many days will library borrowers keep books for?

Classification:

- Categorising or labelling data
  - Which paintings are of animals/architecture/people?
  - Which documents should be classified as sensitive?

THE NATIONAL ARCHIVES

Worked example – booking a holiday

Imagine you want to go on holiday next month and you want to predict the temperature for your destination of choice:

“Fantastic Land: latitude 1.23456; longitude -2.34567; Month: April; Temperature: ???”

Location	Latitude	Longitude	Month	Temperature
A	15.12345	32.873487	April	15C
B	1.11111	-1.99999	April	32C
C	6.55555	10.88888	April	-4C

THE NATIONAL ARCHIVES

Worked example – booking a holiday

Imagine you want to go on holiday next month and you want to decide if you would like your destination of choice based on previous holidays:

“Fantastic Land: latitude 1.23456; longitude -2.34567; Month: April; Temperature: 29”

Location	Latitude	Longitude	Month	Temperature	Rating
A	15.12345	32.873487	April	15C	Good
B	1.11111	-1.99999	April	32C	Bad
C	6.55555	10.88888	April	-4C	Good

Supervised learning

The previous examples were of supervised learning. In the supervised scenario a set of labelled examples are passed to a machine learning classifier which learns to identify relationships between features of the data and the labels, or a numerical output. The table shows some examples of features and labels for some supervised learning tasks:

Task	Application	Features	Outputs	Example of learning
Classify	Sentiment analysis	Word in sentence	+ve or –ve	Glad or happy indicative of +ve
Classify	Seasonal paintings	Image of paintings	Spring, Summer, Autumn, Winter	Red leaves predict Autumn
Predict	Child height	Age of child	Height of child	Height increases as age increases

THE NATIONAL ARCHIVES

Unsupervised learning

Unsupervised learning is not given any examples. Instead a target is suggested and the algorithm groups the data based on that target. The target is usually the number of groups wanted, and the algorithm will place data points into each group in order to maximum the similarity of group members. The following activity aims to give you an intuition for clustering, a commonly used form of unsupervised learning.

Person	Interests
A	politics, sport, nature
B	walking, cooking, quiz shows
C	baking, sewing, athletics
D	newspapers, biographies, history
E	football, rugby, cricket
F	fine dining, pub quizzes, bird watching

Supervised vs Unsupervised

Each paradigm has its advantages and disadvantages.

- Unsupervised learning:
  - identifying clusters of similar records in a set of data making it ideal for gaining a high level view of a new dataset.
  - Choosing the right number of clusters can be difficult and we have no control over the criteria for how clusters are formed.
- Supervised learning:
  - we define the categories in advance giving us control over the outputs.
  - The downside is the cost of labelling our data.
  - Consider the effort involved in the following tasks:

- Transcribing 100 pages of handwritten medieval documents
- Tagging each of 10000 images if they contain an umbrella
- Linking together daily visitor data with weather data currently held on someone else's website

THE NATIONAL ARCHIVES

Activity (not today)

Fill in the blanks with either “Supervised Learning”, “Unsupervised Learning”, “Prediction” or “Classification”:

- Estimating how much money a customer will spend in the museum shop is a _____ task
- A program to decide if a customer is a ‘big spender’ or a ‘browser’ would use a _____ algorithm
- Identifying four types of library visitor is an example of _____
- _____ requires labelled examples

THE NATIONAL ARCHIVES

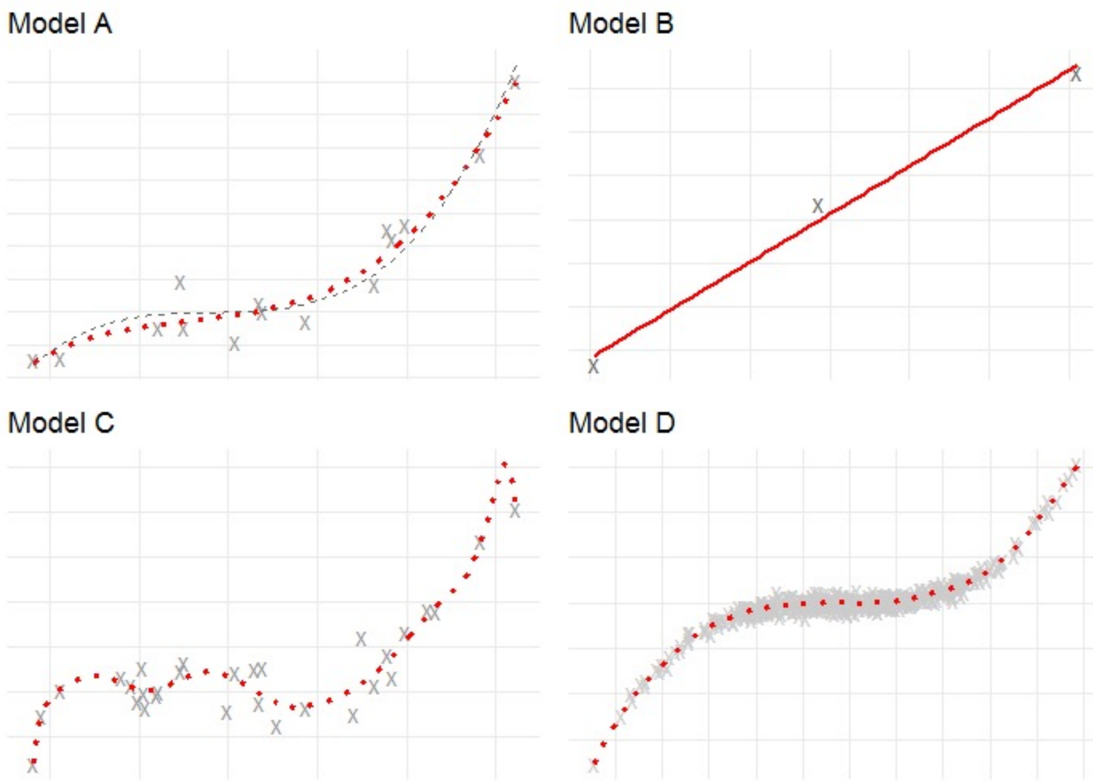
Models and algorithms

There is a lot of jargon and terminology in Machine Learning, and it can be confusing to a newcomer especially when some terms have more than one meaning. The term **model** is one such example. In this episode we will present two definitions of model, and also introduce the term **algorithm**.

Conceptual model:	the average
Algorithm:	add up numbers, divide by how many numbers
Trained model:	the average of a list of numbers
Use case:	predict height of next person to enter the room
Prediction:	trained model returns the average height

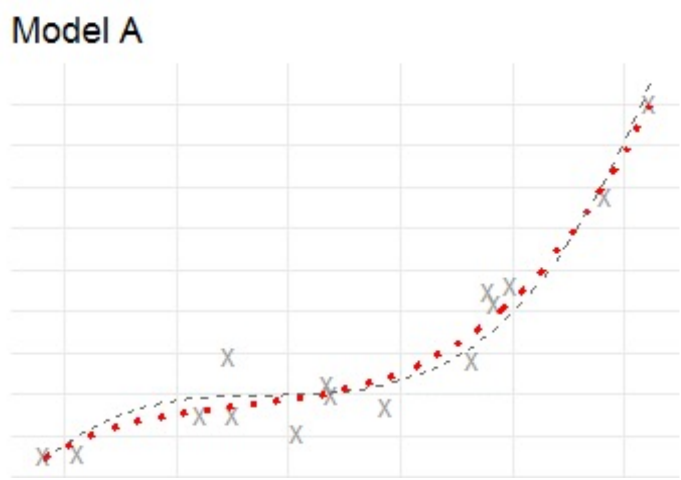
THE NATIONAL ARCHIVES

Training models with data



THE
NATIONAL
ARCHIVES

Model A

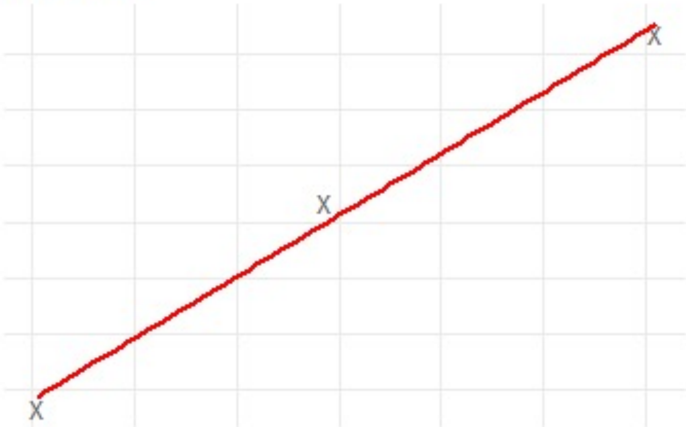


- 15 training data examples
- Moderate complexity model
  (allowed to curve twice)
- Fits very well to true function
  (not a coincidence)

THE NATIONAL ARCHIVES

Training models with data

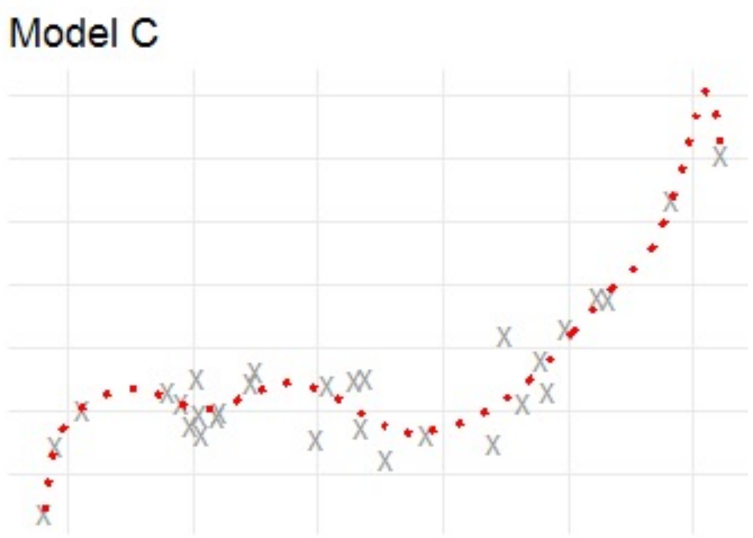
Model B



- 3 training data examples
- Lowest complexity model (no curves allowed)
- Does not fit well
- **Underfitting**
- Might be good enough

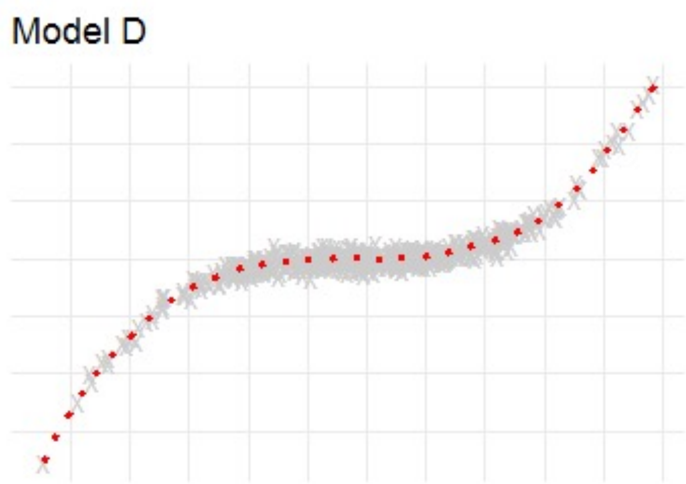
THE NATIONAL ARCHIVES

Training models with data



- 30 training data examples
- Highest complexity model
  (12 curves allowed)
- Fits too well
- **Overfitting**
- Will not generalise to new data

Training models with data



- 500 training data examples
- Highest complexity model (12 curves allowed)
- Fits very well
- Same result as model B but much more data needed

THE NATIONAL ARCHIVES

Model, algorithm, model

Conceptual model:	Line with two curves
Algorithm:	Least Squares (Gauss, 1795)
Trained model:	A formula to draw the line
Use case:	Predict Y from X
Prediction:	Plug X into formula return Y

THE NATIONAL ARCHIVES

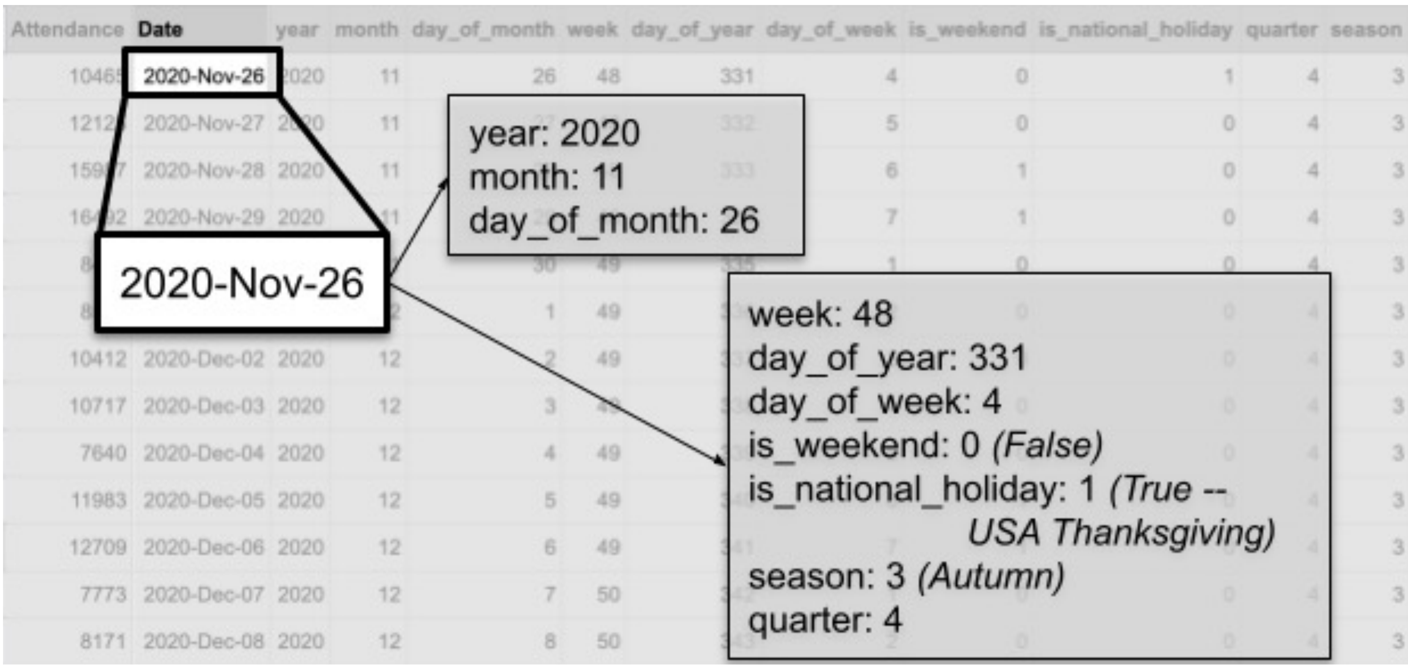
Feature Engineering

One of the trends seen in machine learning, especially in the field of **deep learning** is that the more training data available, the better the results, and the more complex problems can be solved. While this is true, it is not helpful if you don’t have much training data available and need to work with simpler conceptual models but still want to solve complex problems.

One solution to this is to use expert domain knowledge of the data to augment it with additional features. This process is known as **Feature engineering**. As a worked example, we will consider a column of dates.

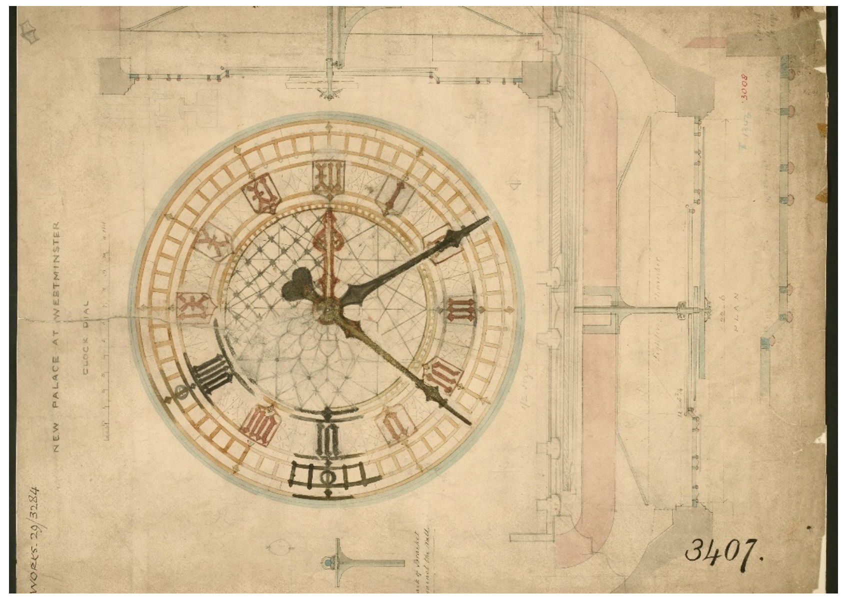
THE NATIONAL ARCHIVES

Feature Engineering example - Dates



THE
NATIONAL
ARCHIVES

Time's up!



Elevation for clock dial for Big Ben tower

THE NATIONAL ARCHIVES