# Key Stage 3 – All Sorted!

Notes for teachers

At a glance

Machine learning is a system where rather than a computer programmer deciding the best way to sort, organise, classify or use information, the computer program develops its own set of instructions (algorithm) based on information that users feed it. Scientists at the University of Oxford are working on ways to improve the speed and accuracy of these algorithms.
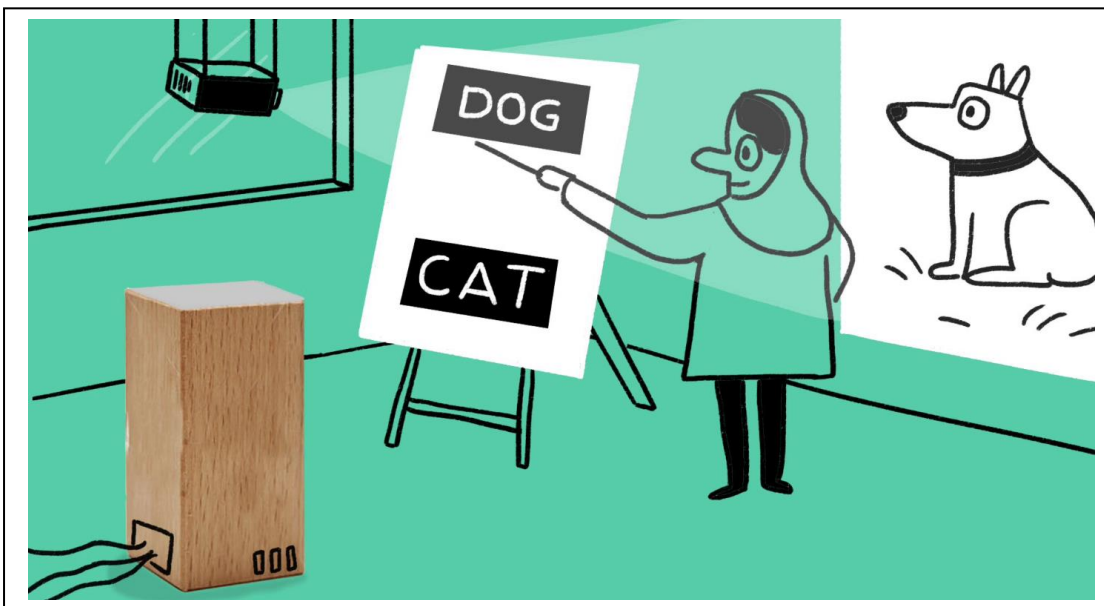
In this lesson students will develop their own rules/algorithms for sorting objects and also look at the effect that increasing the size of the data set has. It is also possible to use the second part of this lesson to look at creating and using classification keys.

Learning Outcomes

- Students can explain why it is important to use large data sets for machine learning
- Students can devise and use algorithms to sort objects
- Students realise outcome depends on data fed to it

Each student will need

- Paper and pen
- Printout of Thingies and Whatsits sheets
- Student Worksheet
- Liquorish Allsorts, Dolly Mixtures, Haribo or similarly varied sweets
- Access to the internet (optional)

http://www.oxfordsparks.ox.ac.uk/content/what-machine-learning

Lesson Activities

1. **Starter activity**
    - Watch the Oxford Sparks Machine Learning video
    - Based on the video, ask students in pairs to come up with three different things that machine learning is already used for or might be used for in the future.

        Things mentioned include:
        - Facial, text and speech recognition (including photo tagging/predictive text)
        - Spam filters
        - Online viewing or shopping recommendations
        - Credit card fraud recognition
        - Medical diagnosis
        - Social media

2. **Main activity: Thingies & Whatsits**
    - Two new types of creature have been discovered, Thingies and Whatsits.
    - Hand out or display Diagram 1 of Thingies and Whatsits.
    - Ask students in pairs or small groups to decide which group the next three creatures belong to and why.
    - Reveal the answers :
        - A- Whatsit
        - B- Thingy
        - C- Whatsit
    - Several groups are likely to have got the wrong answers. Emphasize that this is very likely at this stage and getting the right answer is probably just luck. In fact, in machine learning, getting the answer wrong helps a machine create better rules.
    - Show them diagram 2. How might they change their rule now?
    - Ask them to choose either a 'thingy' or 'whatsit' and draw a version that hasn't been seen in any of the pictures.
    - Ask a few students to show their creations.
    - Reveal the rule – **Whatsits have 4 straight sides and Thingies don't** (nothing else matters – colour, number of eyes etc). Do their new creations fit the rule?
    - [Optional] Students play a game with each other where they decide the rules for a new creature (a *whojamacallit*) and other students have to try to guess the rules either by asking questions or sketching a creature and finding out whether it is a *whojamacallit* or not.
    - Machine learning is very dependent on the data that we feed it. In order to produce accurate results we need to feed it a large amount of data but also data that is as varied as possible. For instance if we were creating a voice recognition program you would probably want to use a large number of words but also different accents or voice types. If we only fed it words spoken by men, for instance, it might struggle to understand women.

3. Main activity: Allsorts

- Give students the student hand out.
- Give the students 4-6 different liquorish allsorts, dolly mixtures, Haribo or similarly varied sweet (you could use photographs if you prefer).
- Ask them to create an algorithm to sort the sweets by asking a series of yes/no questions. e.g. Is it square? Is it yellow? Try to encourage students to focus on more general properties of the sweets rather than just asking 'is it a Cola bottle?' or similar.
- The simplest way of representing the algorithm is to use a flow chart, decision tree or similar graphical representation.
- Get students to swap their algorithm and sweets with other groups and check that they can use them to get the same answers.

4. Plenary

- 20Q is an app/website and stand alone toy that uses machine learning to guess the object that you are thinking of by asking a series of questions. The computer algorithm uses the answers that it gets from thousands of previous users to improve how quickly and accurately it can guess the object you are thinking of.
- Either get students to use the app/website themselves or run through a couple of objects as a class.
- The program reports back on how other users have answered the same questions. Ask students to decide whether they agree/disagree with some of the alternative answers given to questions.
- Although this algorithm is actually very good at playing 20 questions it does rely on the users answering correctly. If users answer incorrectly either by accident or on purpose it alters the accuracy of the result.
- Ask the students how this, and other machine learning systems, can improve the accuracy of its results. They should realise that having more users/data will usually improve the results as the wrong answers are 'diluted' by much larger numbers of correct answers.

Weblinks

- Oxford Sparks Machine Learning Page http://www.oxfordsparks.ox.ac.uk/content/what-machine-learning
- Twenty Questions game website. http://20q.net/