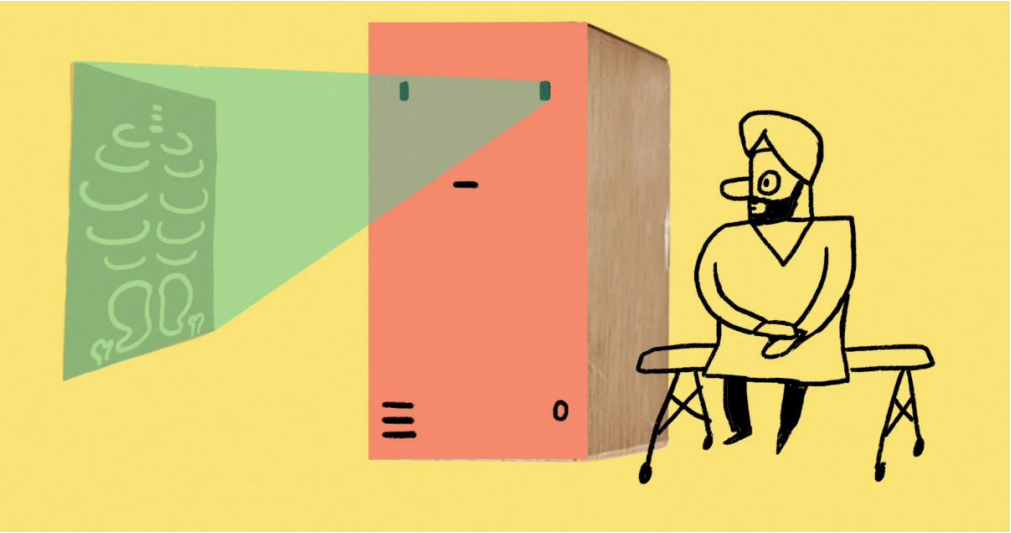
Key Stage 5 – Testing Testing

Notes for teachers

At a glance

Machine learning is a process where machines or rather, computer code running on machines, is created that allows the code to develop its own methods to categorise information based on data that we feed into it. Scientists at the University of Oxford are working on ways to improve the speed and accuracy of these systems. In the future, many decisions are likely to be made, not by people, but by computer algorithms. These algorithms might be used for a range of different 'tests' from recognising faces to diagnosing medical conditions. It might be possible in the future to run far more tests quickly and cheaply. However, is this always desirable?



False negatives/positives

In this lesson, students discover some of the uses of machine learning in addition to exploring the numbers behind false negatives and false positives and some of their surprising consequences.

Learning Outcomes

- Students can give examples of machine learning
- Students can describe what is meant by a false positive and a false negative and solve simple problems involving them
- Students recognize that false positives/negatives can lead to counterintuitive results

Each student will need

- Student Worksheet
- Access to internet or 20Q app (optional)

Possible 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: Testing Testing**
    - Explain that there are lots of different things that machine learning is already used for/might be used for in the future. We're going to take a look at three of those things:
        - Credit card fraud
        - SPAM filtering
        - Medical testing
    - Consider a SPAM filter on an email system. Ask students to describe ways that a SPAM filter might filter email incorrectly. Basically there are two ways it can make a mistake - it can send through SPAM email to your inbox or it can filter out email that you wanted to read. NB There are also two different correct things it can do i.e. filter out unwanted emails and send through wanted emails.
    - If it 'tests' an email and decides it is SPAM this is a positive test result – if it is wrong (i.e. it isn't really SPAM this would be a FALSE POSITIVE)
    - If it tests an email and decides it isn't SPAM this is a negative test result – if it is wrong (i.e. it really s SPAM even though it decided it wasn't this is a FALSE NEGATIVE)
    - Give out student handouts and get students to read through the credit card example and then fill in their own answers for the remaining questions. The first problem looks at SPAM filters, the second at medical testing. The answers to the questions are given below.

3. **Plenary**
    - Point out that tests that are performed by humans also have errors. One of the potential advantages of using computers is that they can get better at performing the tests i.e. they can **learn**. For instance, the more you use predictive text on your phone the better it gets at predicting what **you** want to type. Similarly you can change the settings in SPAM filters to teach the computer to always show you emails from particular people or to let it know when something you've received is SPAM. That way it gets better at recognising SPAM in the future and only showing you emails that you want to see. However, in both cases you need to correct it in order for it to improve. This is a way to reduce false positives and false negatives.
    - Ask students to discuss in pairs or small groups the possible advantages and disadvantages of

using computers to perform medical tests by, for instance, looking at a slide to detect cancer cells.

- Some things to think about might include:
    - Whether computers or humans are better at particular tasks
    - The cost of using people vs the cost of using computers
    - Whether you can find out why/how the decision has been made
    - Computers don't get bored and can run 24 hours a day
    - Computers using machine learning have to have good data fed to them in the first place
- [Optional] 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.
- Ask students to try a few rounds of the game and pay attention to any reports of how their answers differ from previous users. The accuracy of the guesses depends on the accuracy of the previous answers to the questions by other users.

Weblinks

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

Answers to student worksheet

SPAM

	SPAM filter blocks email (positive)	SPAM filter lets email through (negative)	Total
SPAM email	1800 90% of 2000- correctly processed	200 10% of 2000 - FALSE NEGATIVE	2000 20% of 10000
Normal Email	80 1% of 8000 – FALSE POSITIVE	7820 99% of 8000 – correctly processed	**8000** 80% of 10000

In this case, which of these describes as a false negative?

ANSWER - a) When a SPAM email is incorrectly identified a proper email.

Medical testing

Four million 20-24 year olds, incidence of disease is 0.01%, test is 99% accurate:

a) How many will have the disease? **40,000**
b) How many will not have the disease? **3,960,000**

If all of the people in the group are tested (test 99% accurate)

c) How many test results will be false negatives? (ie test for disease comes back negative when it should be positive) 1% of 40,000 = **400**

d) How many test results will be false positives? (ie the test says they have the disease when they don't) 1% of 3,960,000 = **39,600**

e) What percentage of all of the positive tests are false positives?
Number of positive tests for people who DO have the disease = 99%x40,000=**39,600**
Which is the same as the number of false positives**.** i.e. **50%** of the positive results are false.

It probably wouldn't be sensible to test all 20-24 year olds if half of the test results will be wrong (although students may consider it worth the risk/worry/expense if it might prevent the disease).

Over 80s - 25% have the disease, test 99% accurate

If we consider 1000 over 80 year olds

- 250 will have the disease
- 750 will not

Total number of positive tests = 99%x250 + 1%x750 =247.5+7.5=255
Total number of false positives = 7.5
% of false positives = 7.5/255 = 3% (to the nearest %)

This is a much lower proportion of false positive results than for the younger age group (3% vs 50%) so it would make much more sense to concentrate on testing older people.