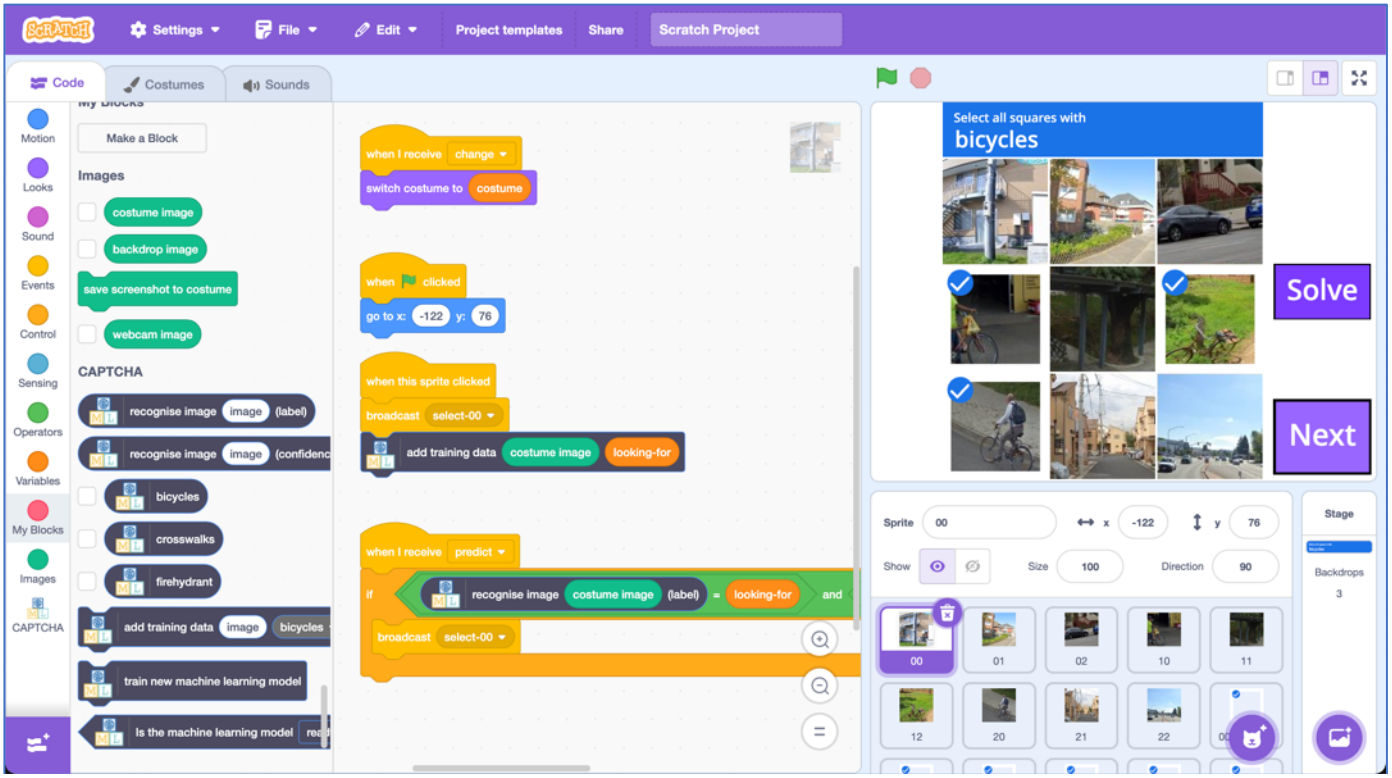
CAPTCHA

In this project you will make a project that can solve CAPTCHAs.

A “CAPTCHA” is a test used to recognise if a user is a human or a bot.

The aim of a CAPTCHA is to give the user something to do that is very easy for a human to do, but difficult for a computer to do.

Computers are getting better at solving CAPTCHAs all the time. In this project, you will try to create a computer program that can solve simple CAPTCHA challenges.



1. Go to https://machinelearningforkids.co.uk/ in a web browser

2. Click on "**Get started**"

3. Click on "**Try it now**"

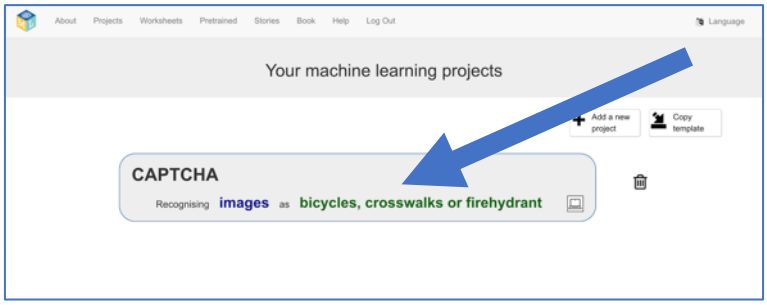
4. Click on "**Projects**" on the top menu bar

5. Click the **"Copy template**" button.

6. Find the "CAPTCHA" template and click on "**Import**"

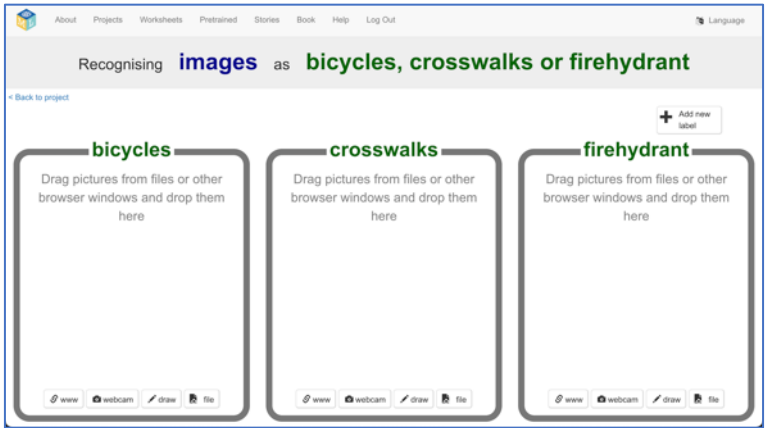
7. Click on the "**Store on your computer**" button

8. You should see "**CAPTCHA**" in the list of your projects. Click on it.



9. Start by reviewing the training buckets in the template. Click "**Train**"

You will be training the computer to recognise CAPTCHAs of bicycles, crosswalks, and fire hydrants.



10. Click on "**< Back to project**"

11. Click on "**Make**"

12. Click the "**Scratch 3**" button

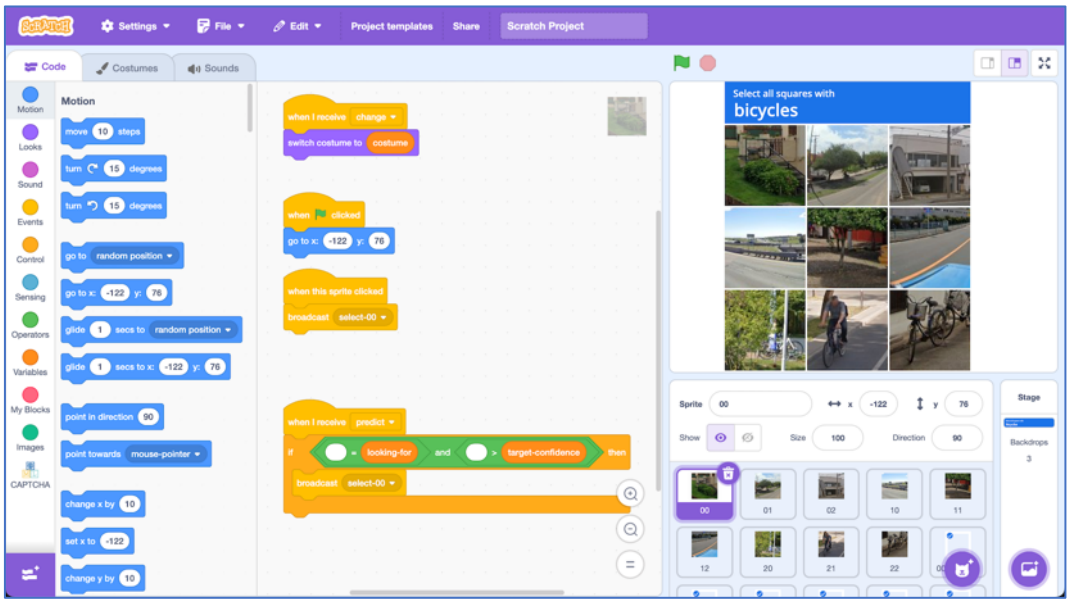
13. Click "**straight into Scratch**"

The page will warn you that you haven't trained a model yet, but that is okay as you will be using Scratch to collect your training data.

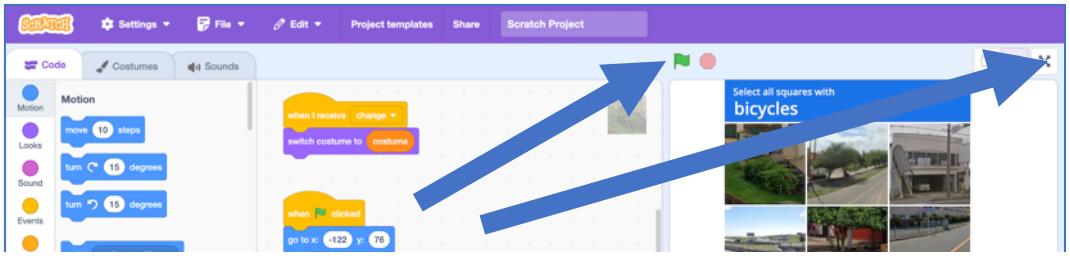
14. In the Scratch window, click on "**Project templates**"

15. Click on the "**CAPTCHA**" template

This is a very large template, as it contains many images.
Wait for the project to download.



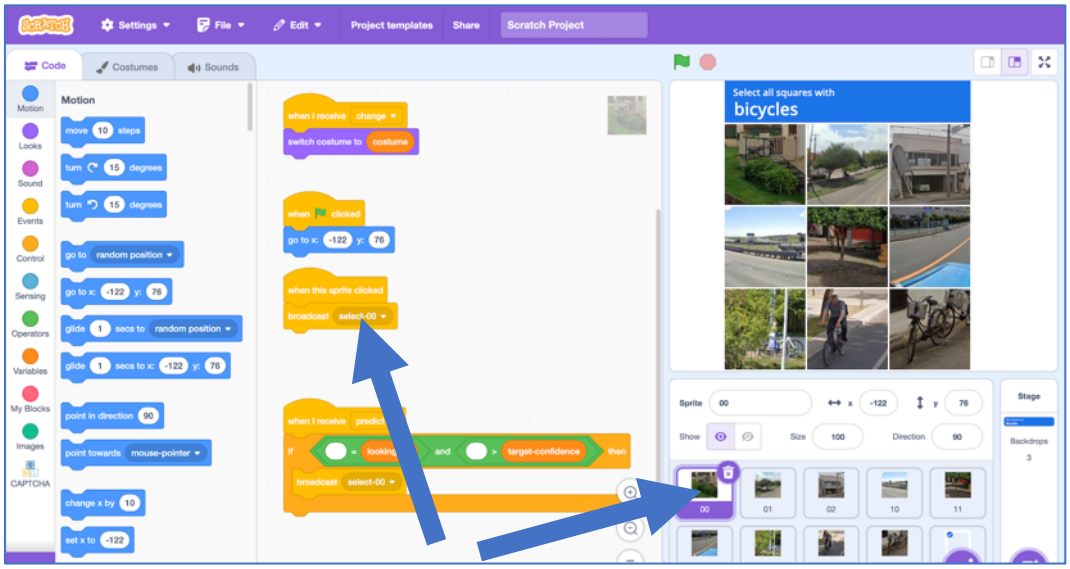
16. Click on the **full-screen** button, and then click on the **Green Flag**



17. Try out a couple of CAPTCHA's

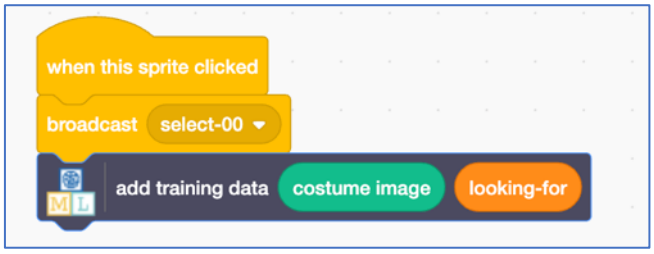
Click on the squares that you think match the prompt.
*Click on the **Next** button to try another CAPTCHA*
The next step is to start collecting images you can use to train the computer.

18. Click on the **00** sprite and find the "**when this sprite clicked**" code



19. Update the code so that it looks like this

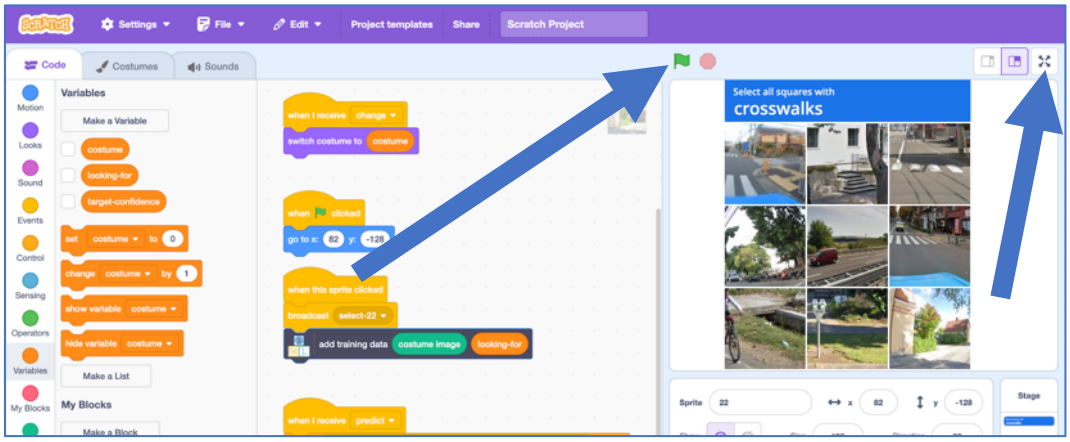
When you click on the sprite, this will add the image to the training data



20. Click on the **01** sprite. Make the same change to the code.

21. Repeat for the remaining grid sprites: **02**, **10**, **11**, **12**, **20**, **21**, **22**

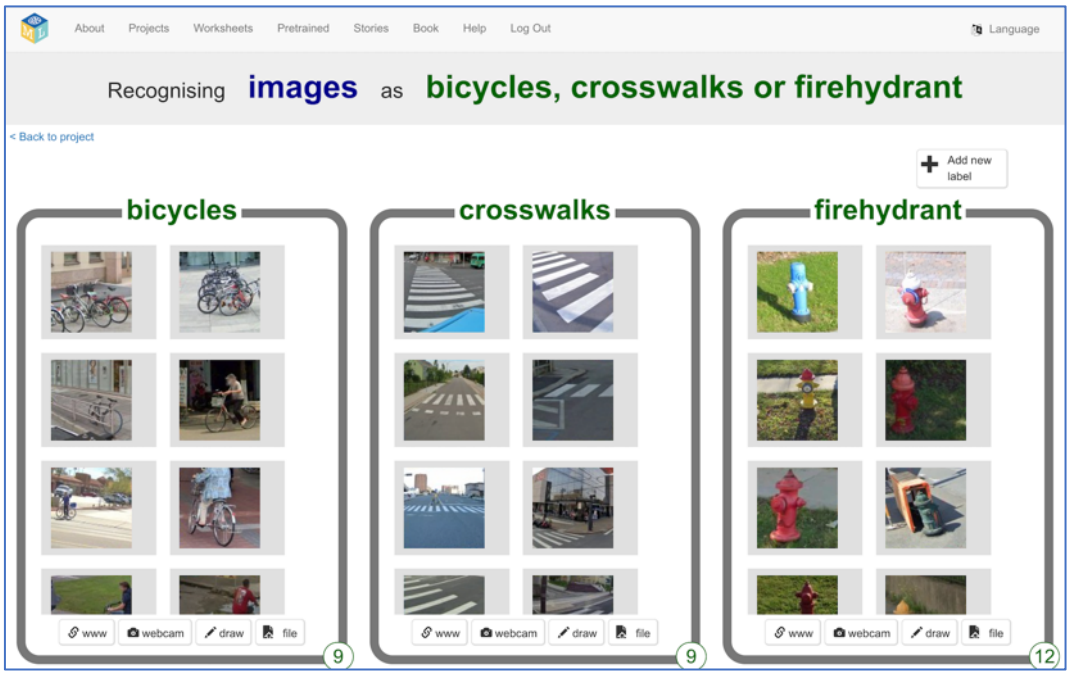
22. Click on the full-screen button and Green Flag again



23. Complete **ten** CAPTCHA challenges

24. In the training tool tab, review the training data you collected

*Click on "**< Back to project**" and then click on "**Train**" to get to this page*



25. Click on "**< Back to project**"

26. Click on "**Learn & Test**"

27. Click on the "**Train new machine learning model**" button

If the button is not there, you will need to complete another CAPTCHA first

What have you done so far?

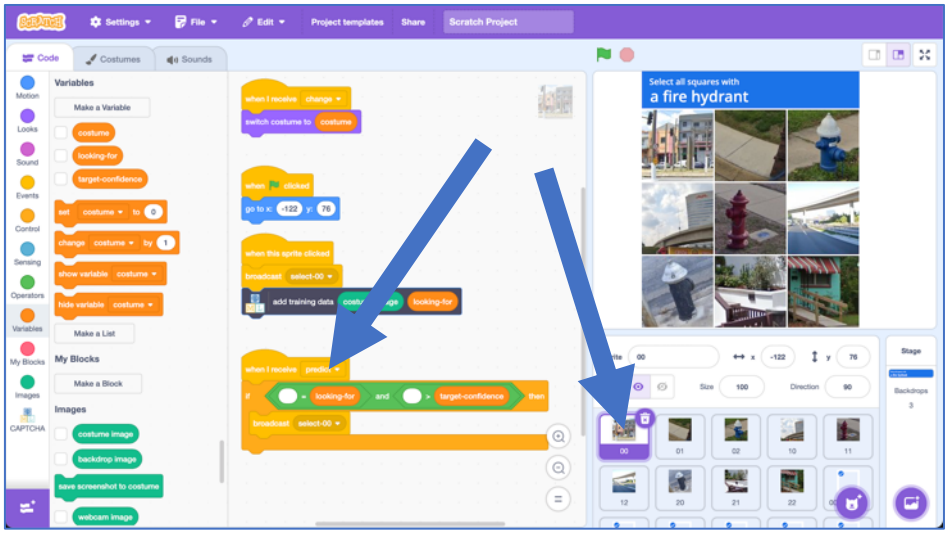
You've started to train a computer to recognise images from CAPTCHA challenges. You are doing it by collecting examples. These examples are being used to train a machine learning "model".

This is called "supervised learning" because of the way you are supervising the computer's training.

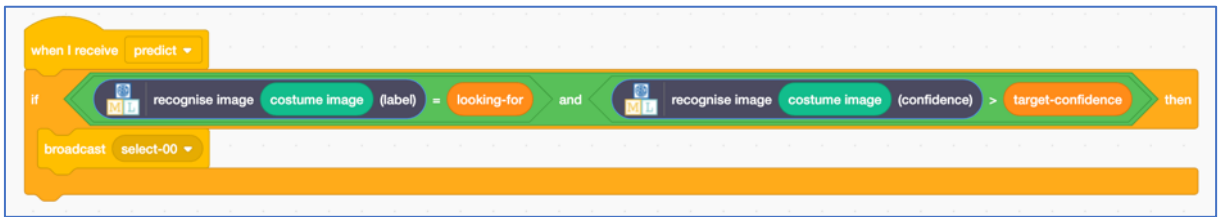
The computer will learn from patterns in the examples you've given it, such as colours, shapes and patterns. These will be used to be able to recognise new CAPTCHA images.

The next step is to modify your Scratch project to use the model to attempt to solve new CAPTCHAs.

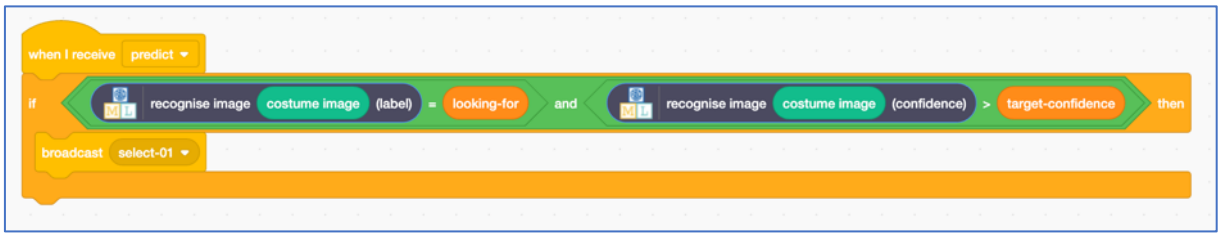
28. In the Scratch tab, click on the **00** sprite.
Find the "**when I receive predict**" code.



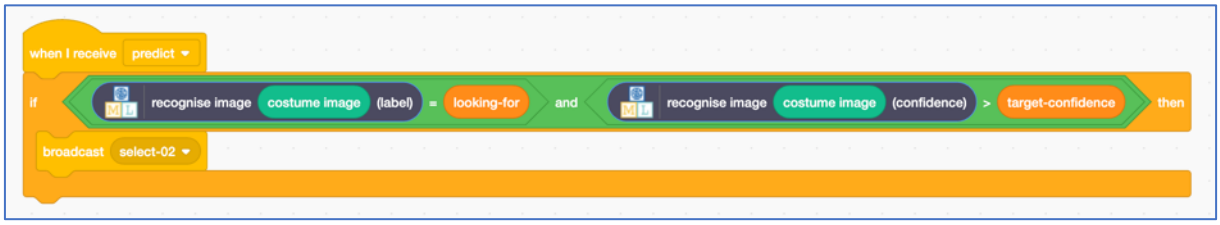
29. Update the code so that it looks like this:



30. Click on the **01** sprite. Make the same change to the code.

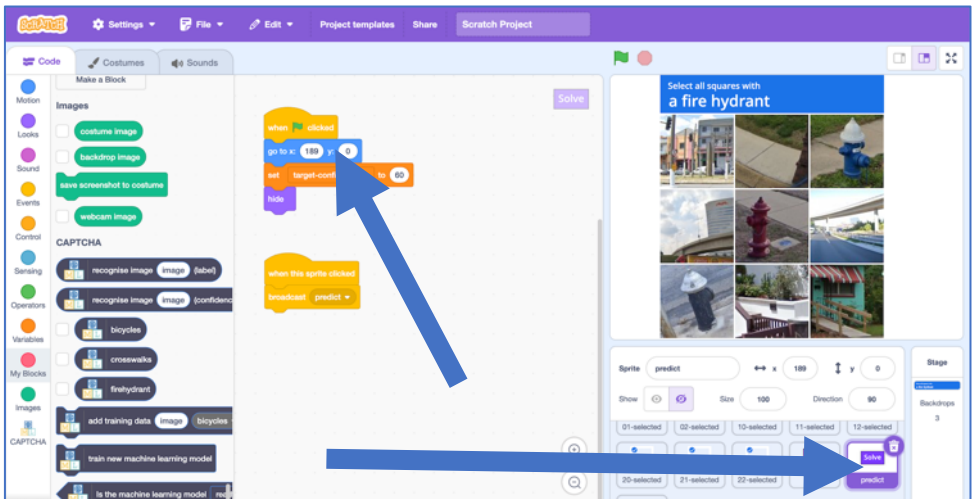


31. Click on the **02** sprite. Make the same change to the code.

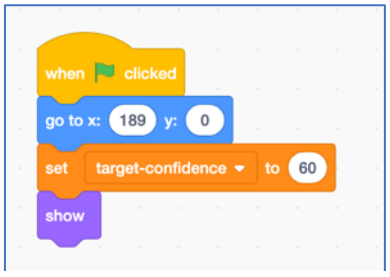


32. Repeat for the remaining grid sprites: **10**, **11**, **12**, **20**, **21**, **22**

33. Click on the "**predict**" sprite and find the Green Flag code

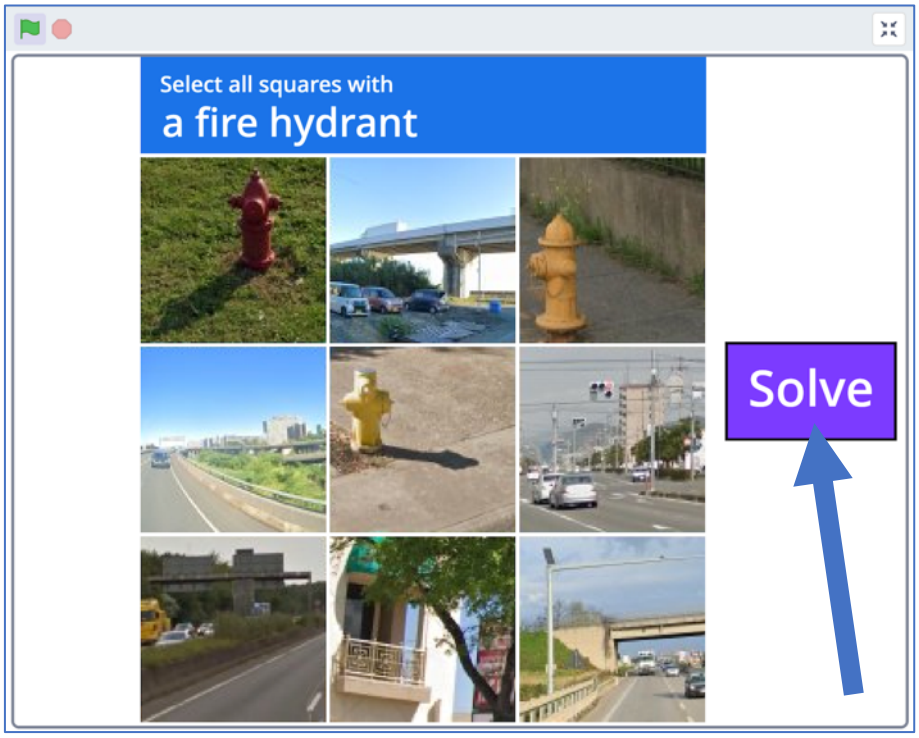


34. Replace the "**hide**" block with a "**show**" block



35. Click on the **full-screen** button and then the **Green Flag**

36. Click on the "**Solve**" button in the project and your machine learning model will select the grid squares that it predicts are a match



37. Click "**Next**" to get the next CAPTCHA challenge.
Click "**Solve**" again to test your model.

38. How good is your machine learning model at solving CAPTCHAs?
With only a small number of examples to learn from, it will likely be making several mistakes.
More training examples should help.

39. Try solving **ten** more CAPTCHA challenges yourself, without clicking on "**Solve**", to add to your training examples.

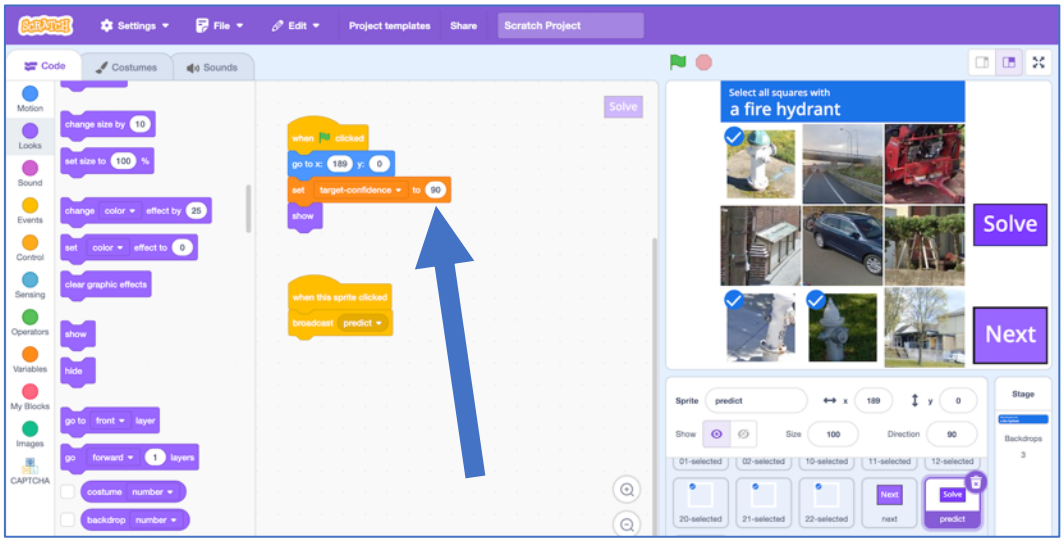
40. Review the extra training examples in the "**Train**" page.

41. Go to the "**Learn & Test**" page.

42. Click on the "**Train new machine learning model**" button again to train a new model with the additional training examples

43. In the Scratch window, try using the "**Solve**" button to solve more CAPTCHA challenges.
Have the additional training examples improved the model?

44. Click on the "**predict**" sprite.
Increase the "target-confidence" value to **90**.
This means your model will only select a grid square when it is at least 90% confident that it contains a match.



45. In the Scratch window, use the "**Solve**" button to solve more CAPTCHA challenges.
Does the increased confidence score mean it makes fewer mistakes?

What have you done?

You've explored ways to change how a machine learning project behaves.

Increasing the number of training examples is often a good way to improve the performance of a machine learning model.

Increasing the confidence threshold often reduces the number of incorrect selections the model makes, but may also prevent the model from making some correct selections.

Did you know?

"CAPTCHA" stands for:
Completely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part

What is a Turing Test?

This was an idea by English scientist Alan Turing in 1950.

Turing proposed a way of measuring if a computer is demonstrating intelligence, by testing if someone could tell it apart from a human performing the same task.

What is a CAPTCHA?

A CAPTCHA is a test given to a user to ask them to prove they are a human and not a computer. The idea is to select a task that is easy for a person to do, but difficult for a computer.

This is done to reduce problems caused by bots such as spam.

Can computers solve CAPTCHAs?

All images used in this project are taken from real-world CAPTCHA services. As you've seen, computers can be trained to perform some CAPTCHA tasks, however they do make mistakes.

CAPTCHA services can deal with this by varying the challenges, to include things that machine learning systems have not learned to do. If your machine learning model was given a CAPTCHA to identify buses, it would not be able to do this until you collected training examples of buses.