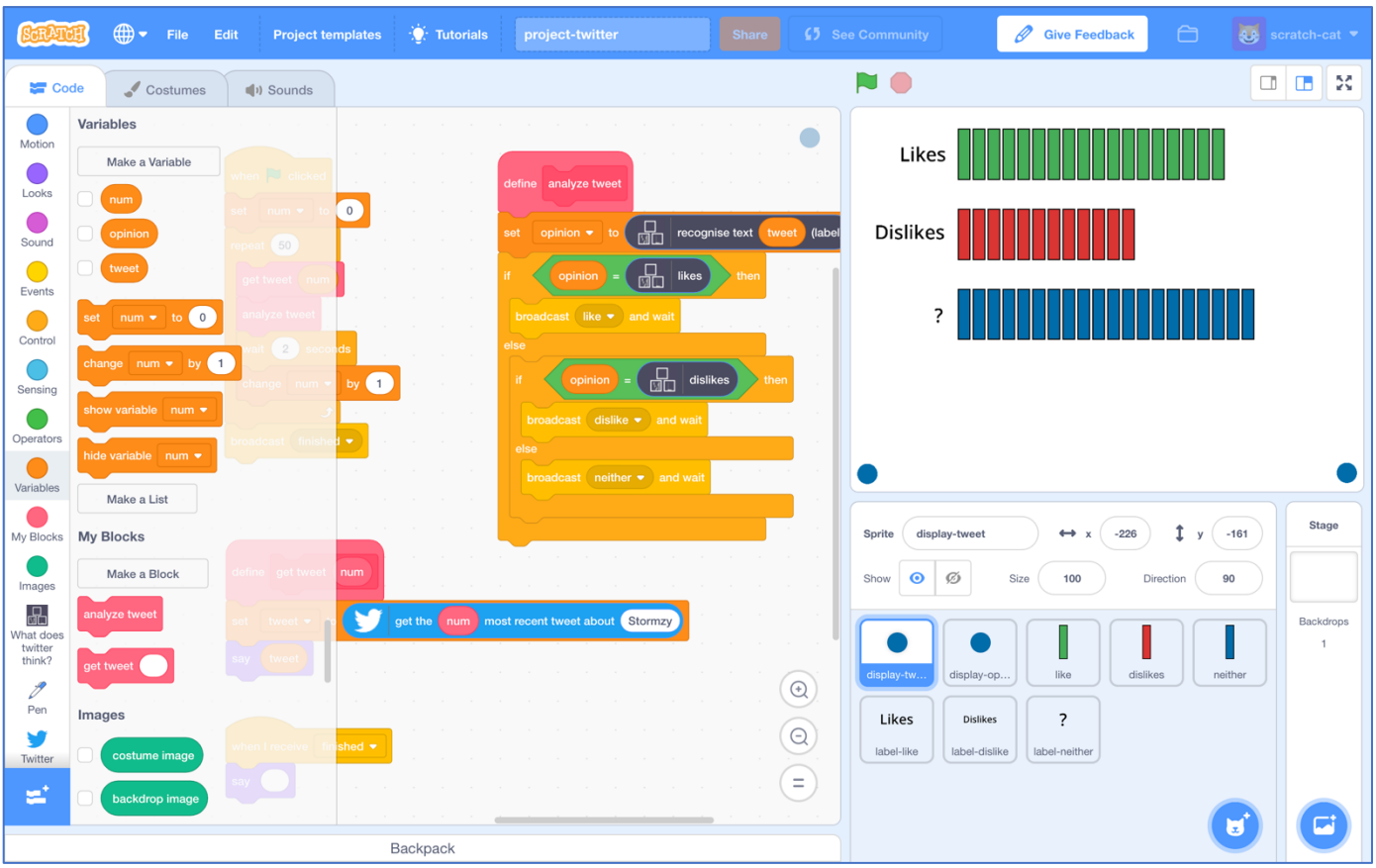
What does Twitter think?

In this project you will use machine learning to estimate what people think about a topical issue of your choice.

You'll train a machine learning model to recognise positive and negative comments about your topic, by collecting examples from social media.

You'll use your machine learning model in Scratch to analyze public discussion and represent this in a live graph.



1. Choose a topic that you'll use for the project

In this project, you'll be analysing what people on Twitter think about something.

Choose something topical that you think people will be talking about. It could be a new movie, a TV show, or something that is in the news.

***Check your idea with your teacher or group leader** before continuing.*

(For the rest of the screenshots in this worksheet, I'll be using Stormzy as I wrote this soon after Stormzy was announced as headlining the music festival Glastonbury).

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

3. Click on "**Get started**"

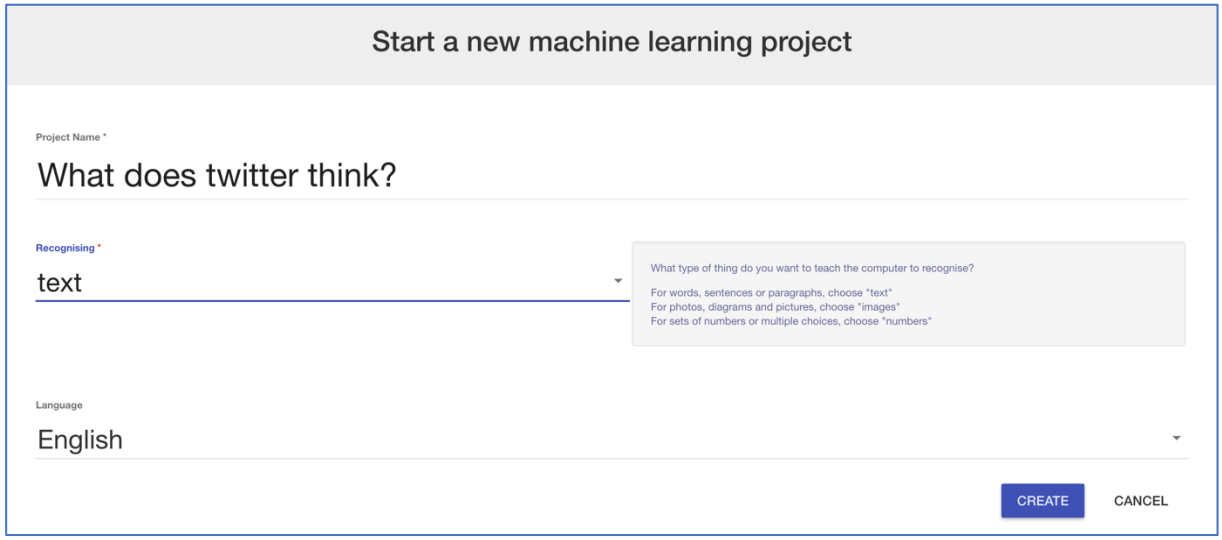
4. Click on "**Log In**" and type in your username and password

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

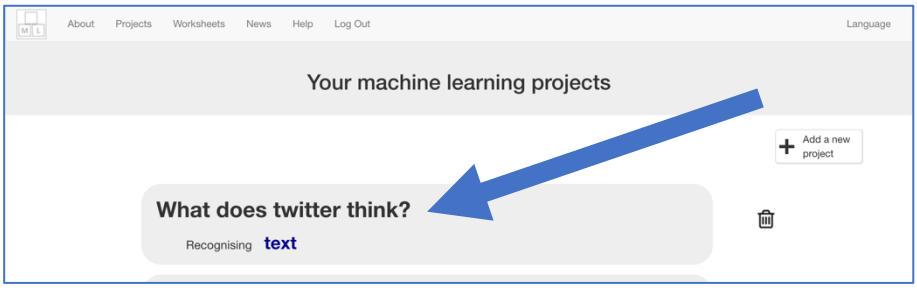
6. Click the **"+ Add a new project**" button.

7. Name your project "What does twitter think?" and set it to learn how to recognise "**text**".

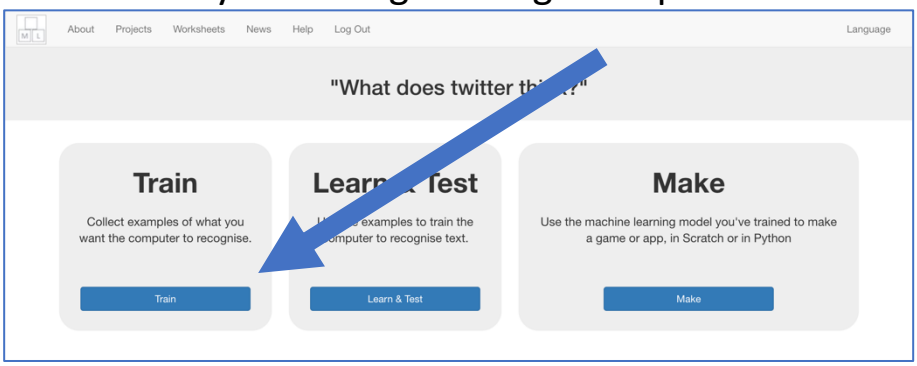
Click the "**Create**" button



8. You should now see "**What does twitter think?**" in the list of your projects. Click on it.

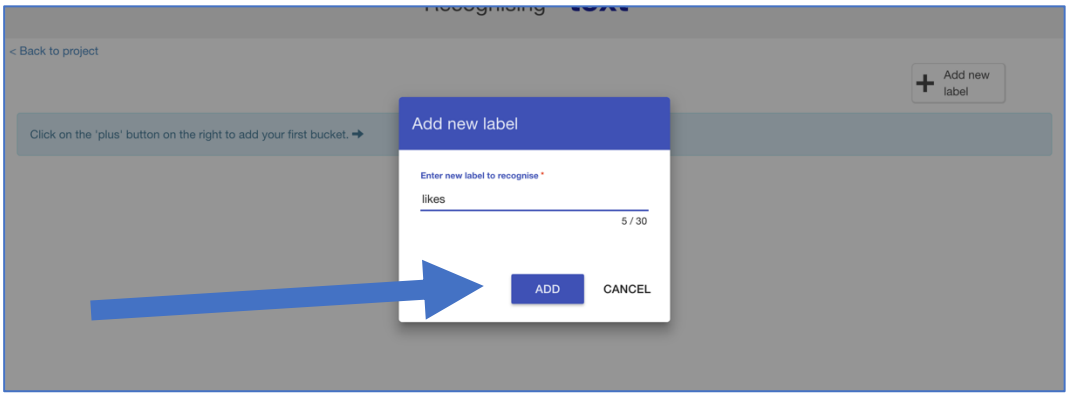


9. Start by collecting training examples. Click "**Train**"

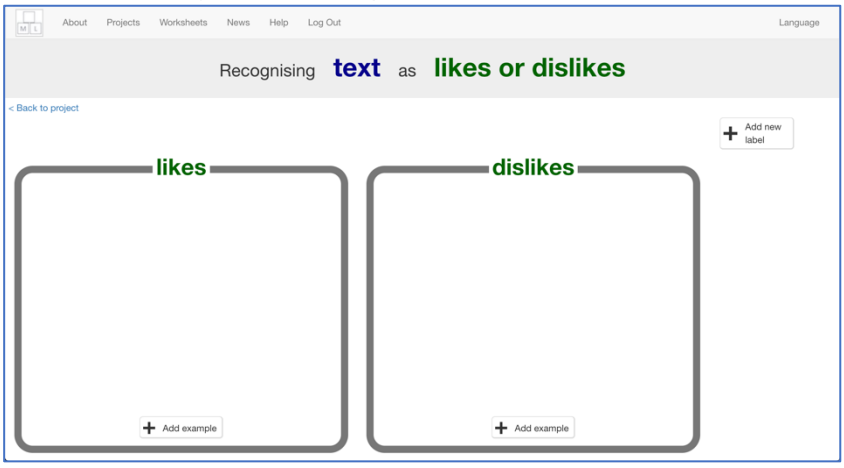


10. First, create a space to store examples of positive comments. Click "**+ Add new label**"

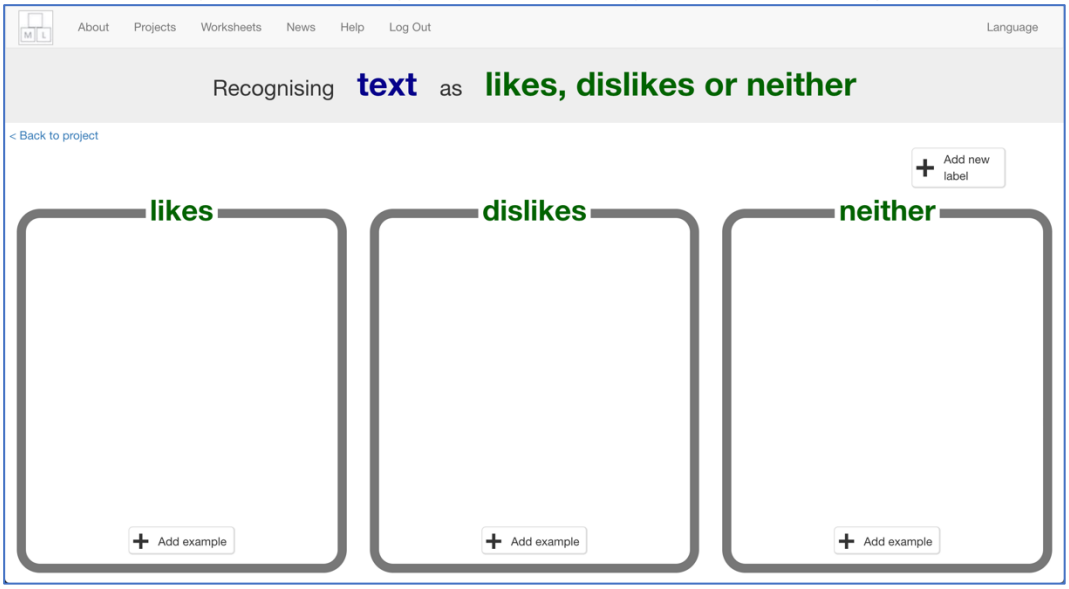
11. Call this bucket "**likes**" and click "**Add**"



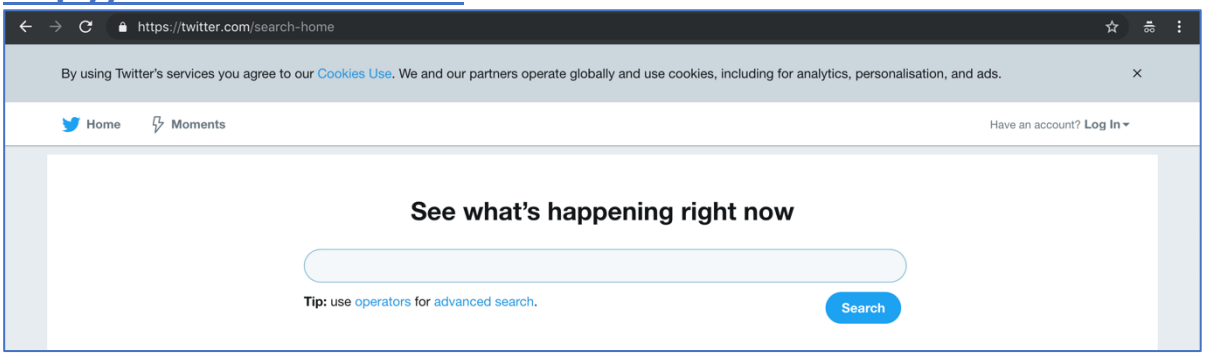
12. Click the "**+ Add a new label**" button again, and create a space to store examples of negative comments, called "**dislikes**"



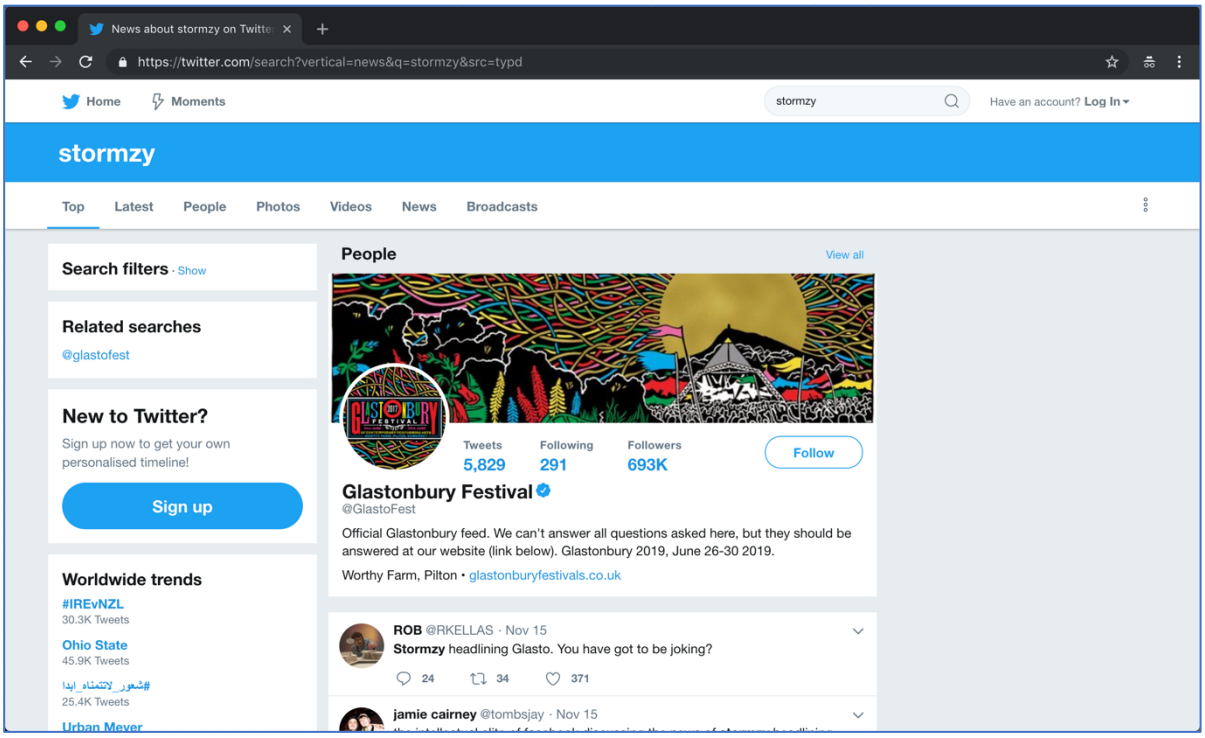
13. Click the "**+ Add a new label**" button again, and create a space to store examples of things that aren't positive or negative, called "**neither**"



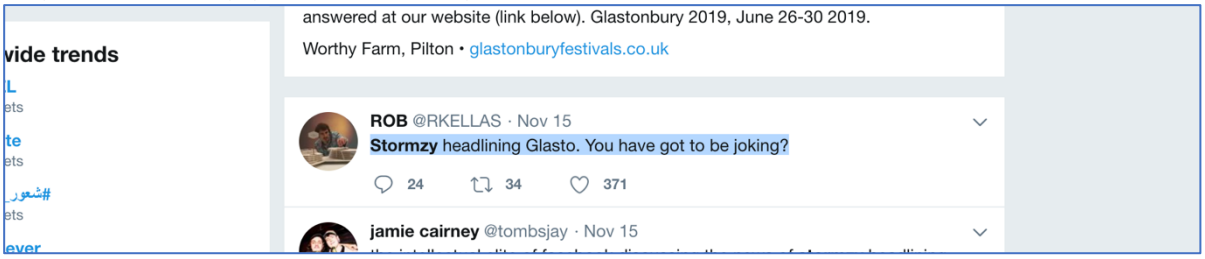
14. Open a new web browser window and go to http://search.twitter.com



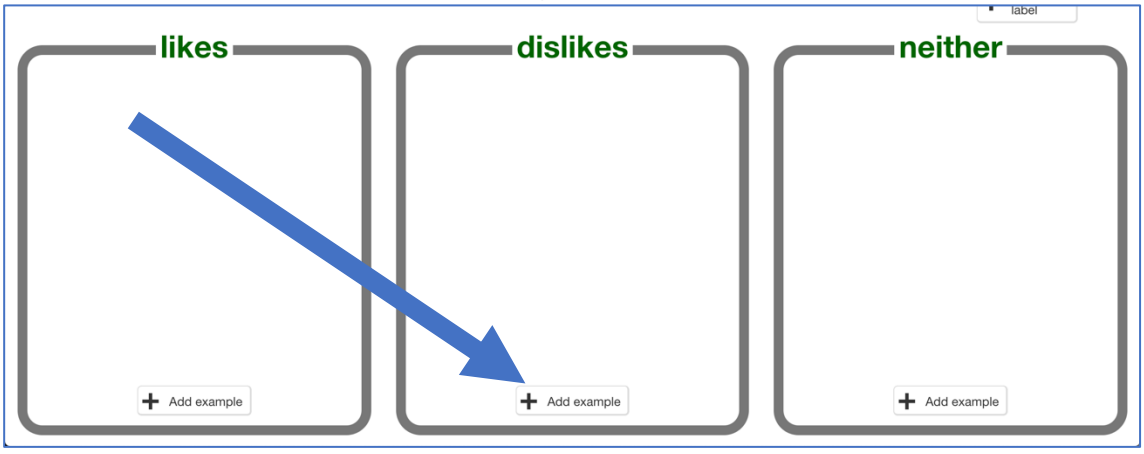
15. Search for the topic you're using for this project



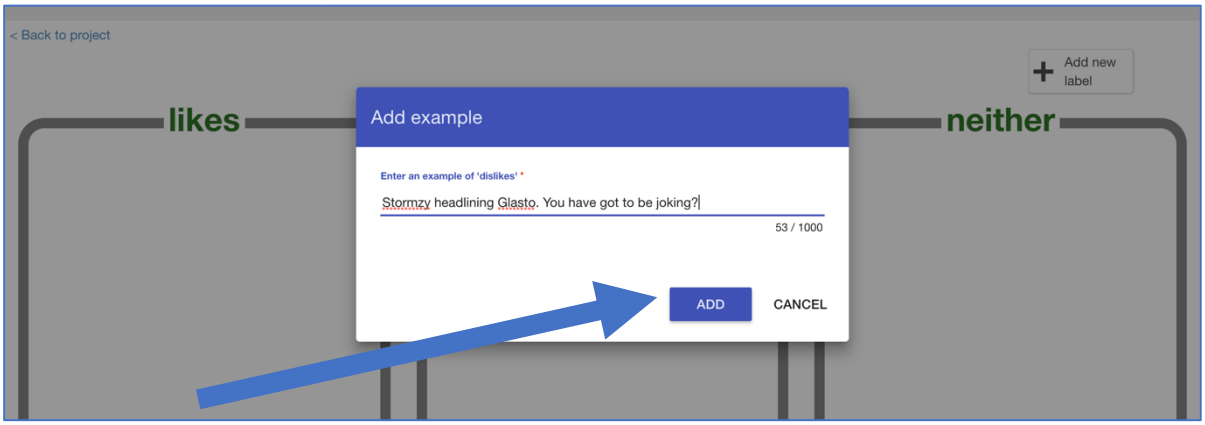
16. Find an example of someone saying something negative about your topic, and copy it to the clipboard



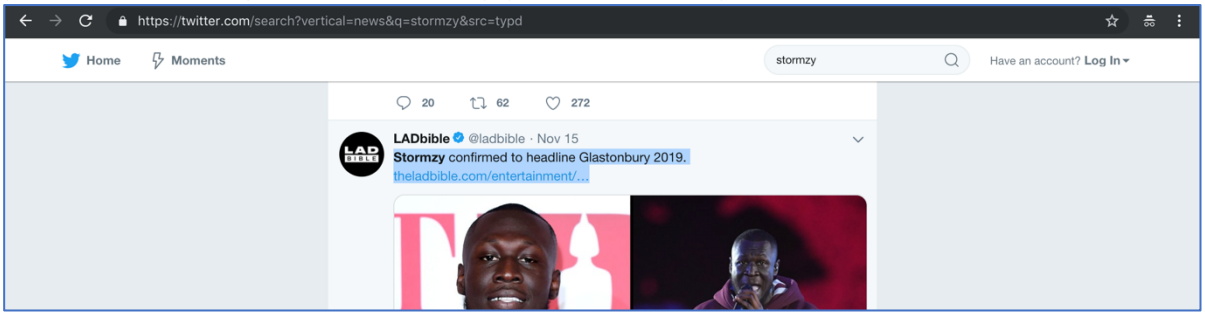
17. Click on the "**+ Add example**" button in the "**dislikes**" bucket



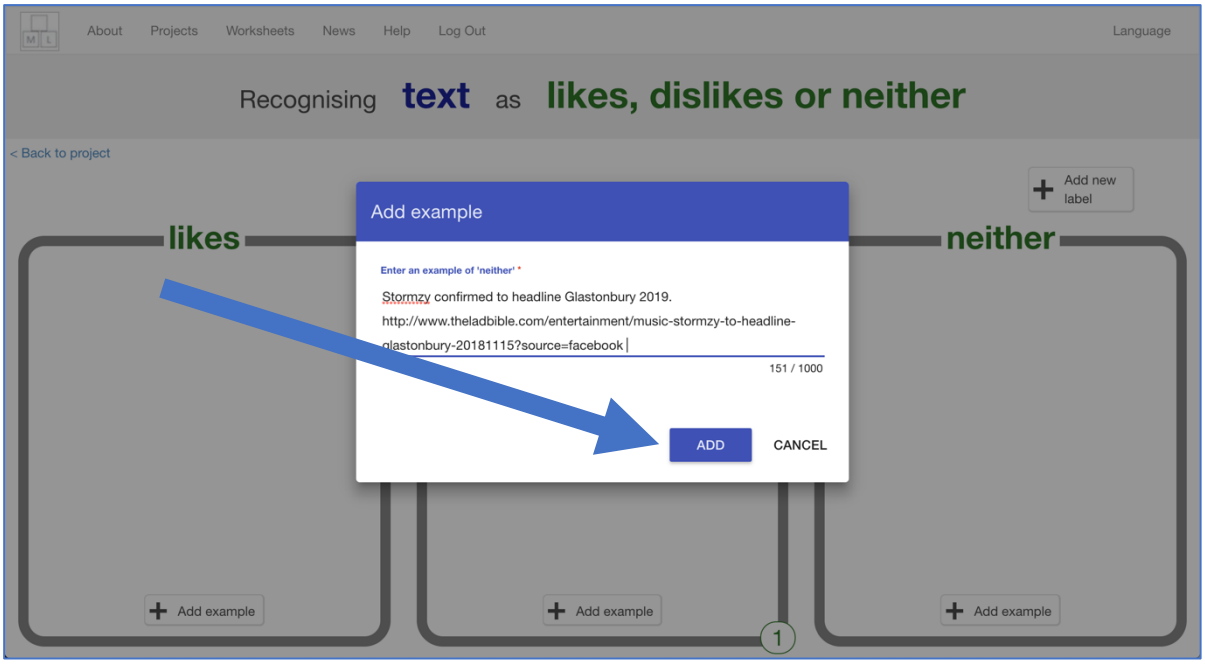
18. Paste the negative comment in the box, and click "**Add**"



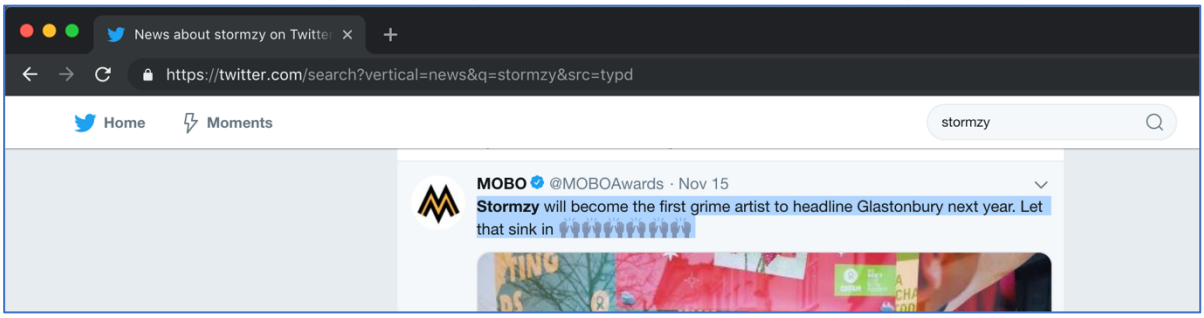
19. Find an example of someone mentioning your topic, that isn't really positive or negative, and copy it to the clipboard



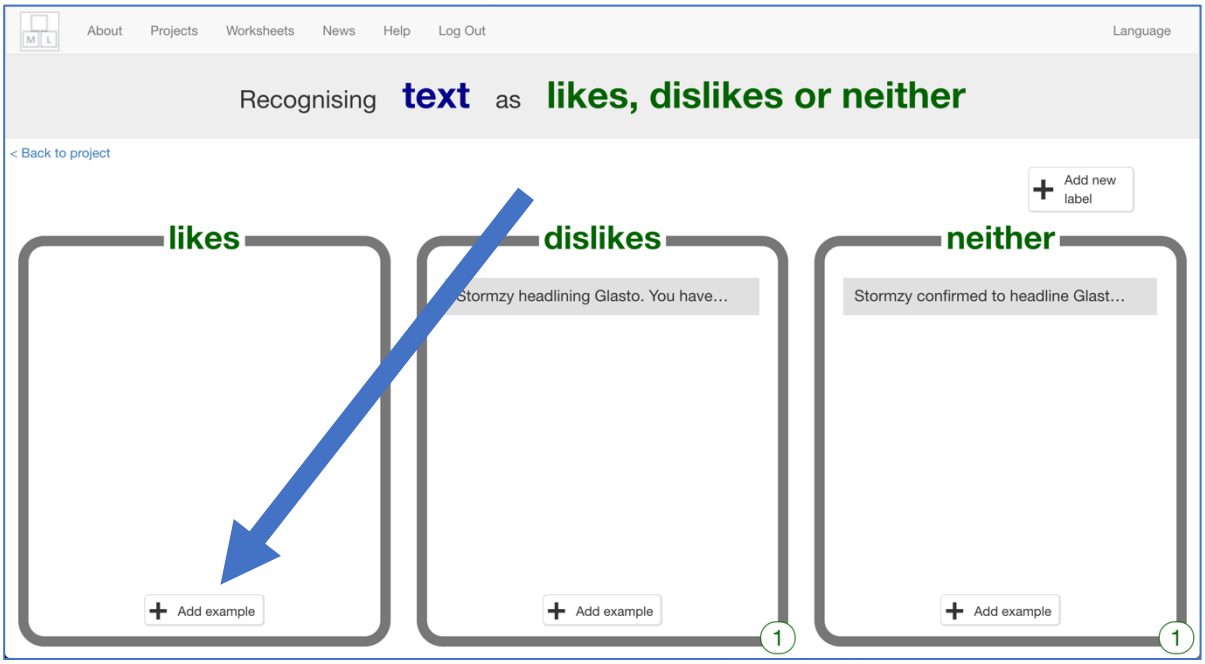
20. Click on the "**+ Add example**" button in the "**neither**" bucket
Paste the comment in the box, and click "**Add**"



21. Find an example of a positive comment about your topic and copy it



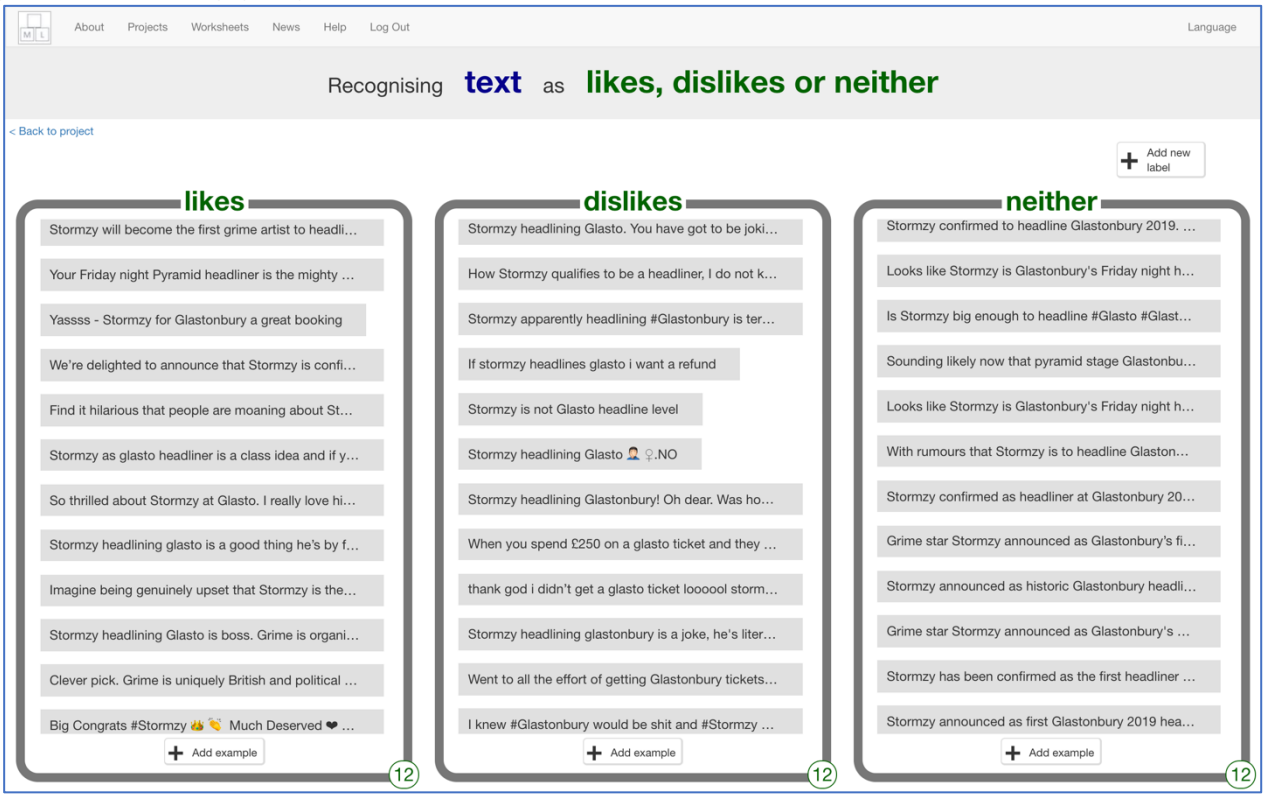
22. Click the "**+ Add example**" button in the "**likes**" bucket



23. Paste the positive comment in the box and click "**Add**"

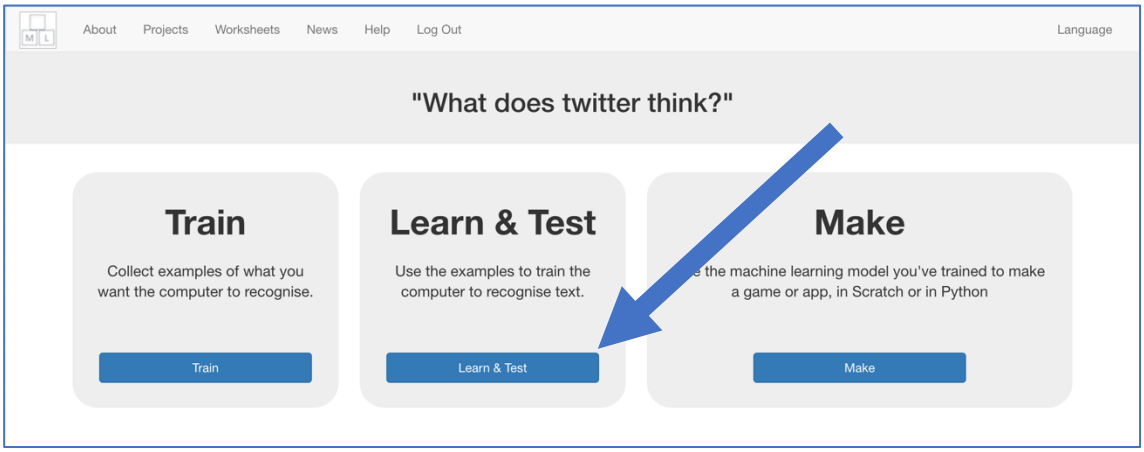
24. Repeat to fill all three buckets with examples

The more examples, the better your project will work, but the minimum for a working project is about 5 in each bucket.



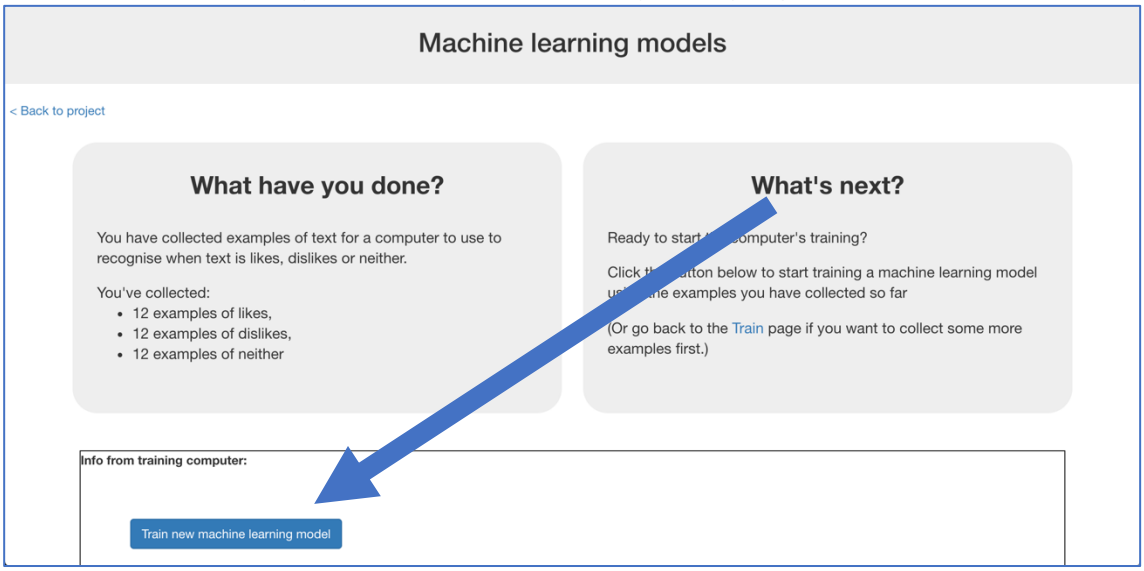
25. Click the "**< Back to project**" link

26. Next, use the examples you've collected to train a machine learning model. Click "**Learn & Test**"



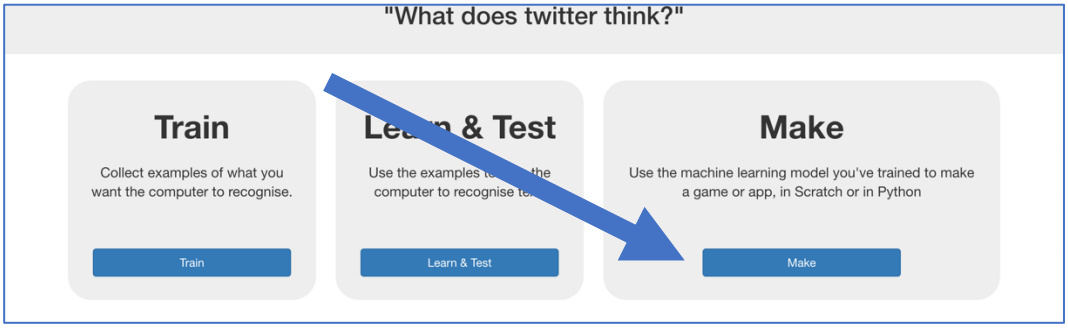
27. Click on the "**Train new machine learning model**" button

This will take a minute or two to train. While you're waiting, you could try the multi-choice quiz at the bottom of the page.

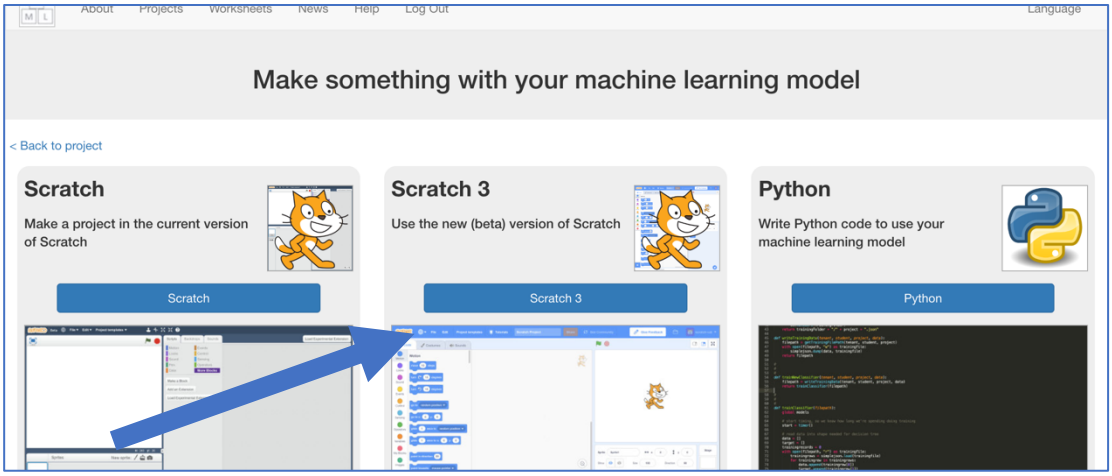


28. Click on the "**< Back to project**" link

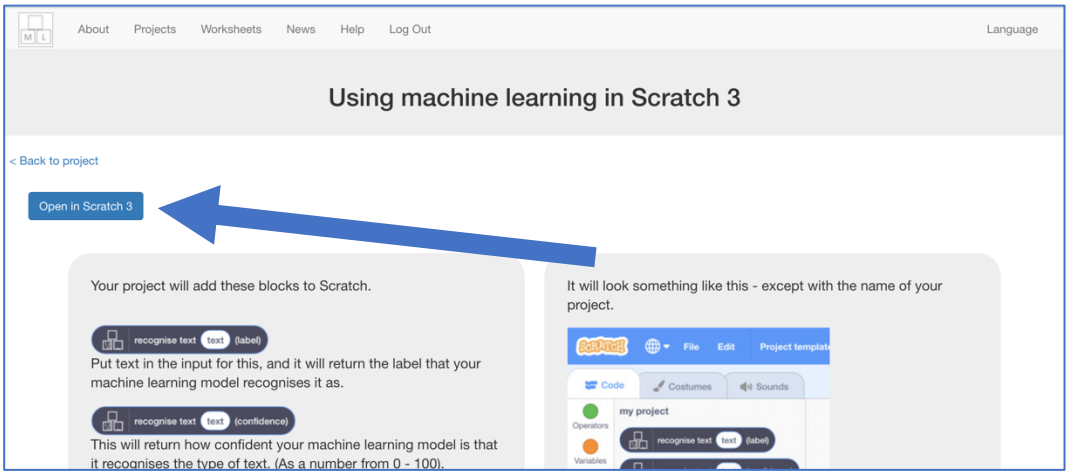
29. Next, we'll use Scratch to analyze tweets. Click "**Make**"



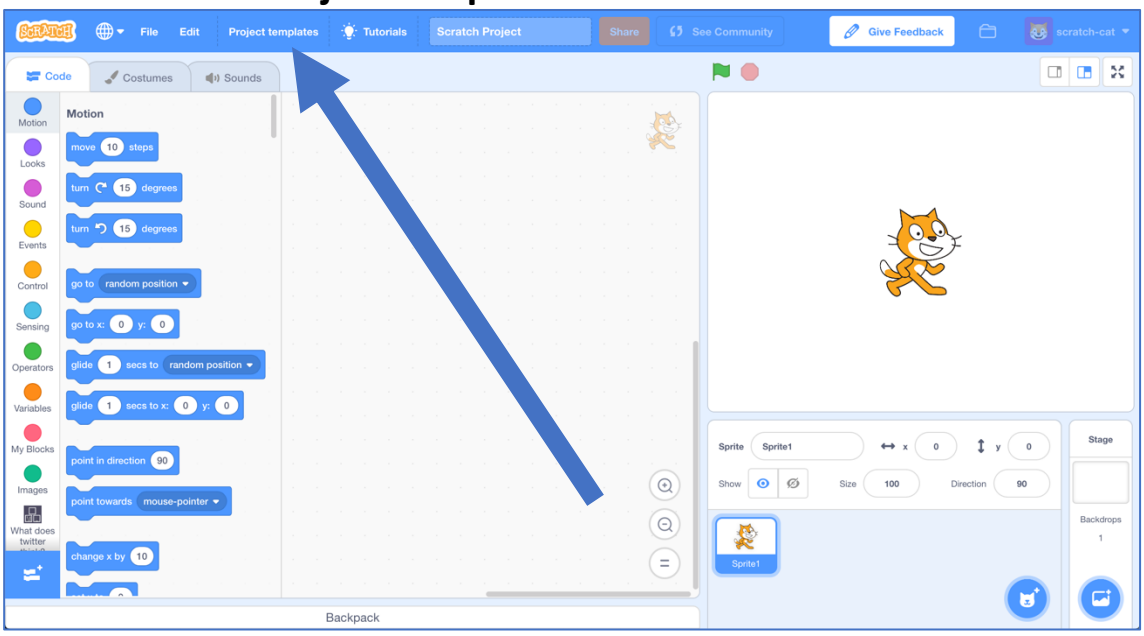
30. Click "**Scratch 3**"



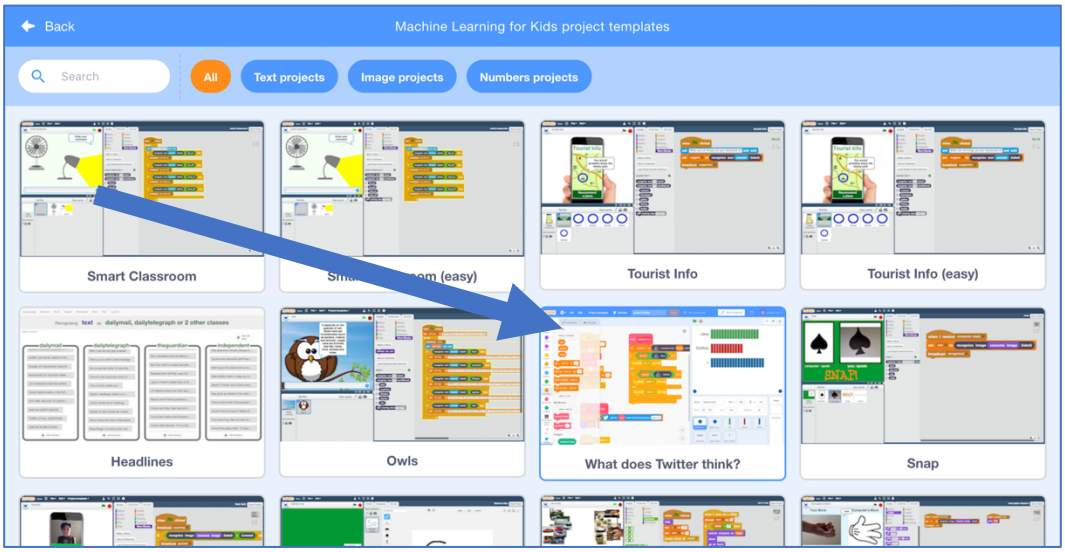
31. Click “**Open in Scratch 3**”



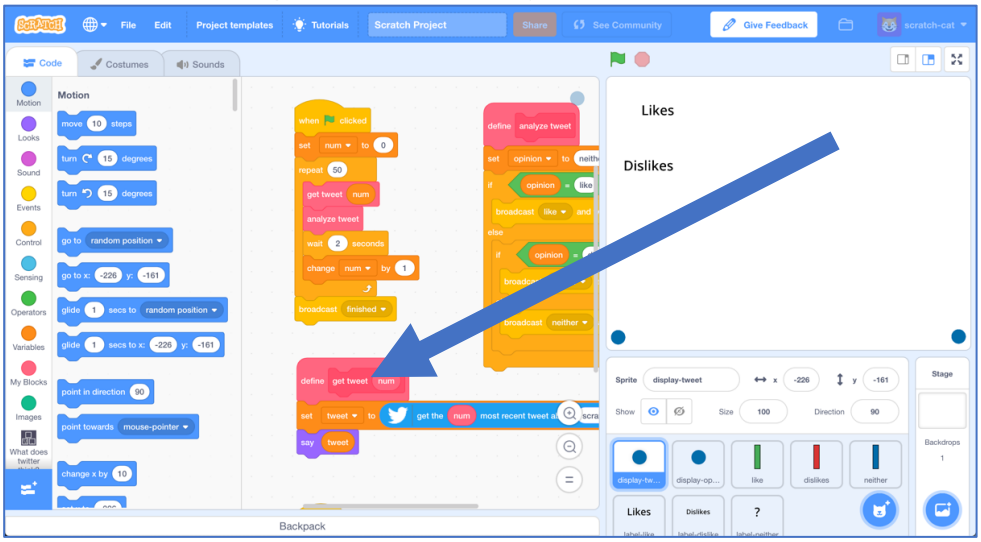
32. Click on “**Project templates**”



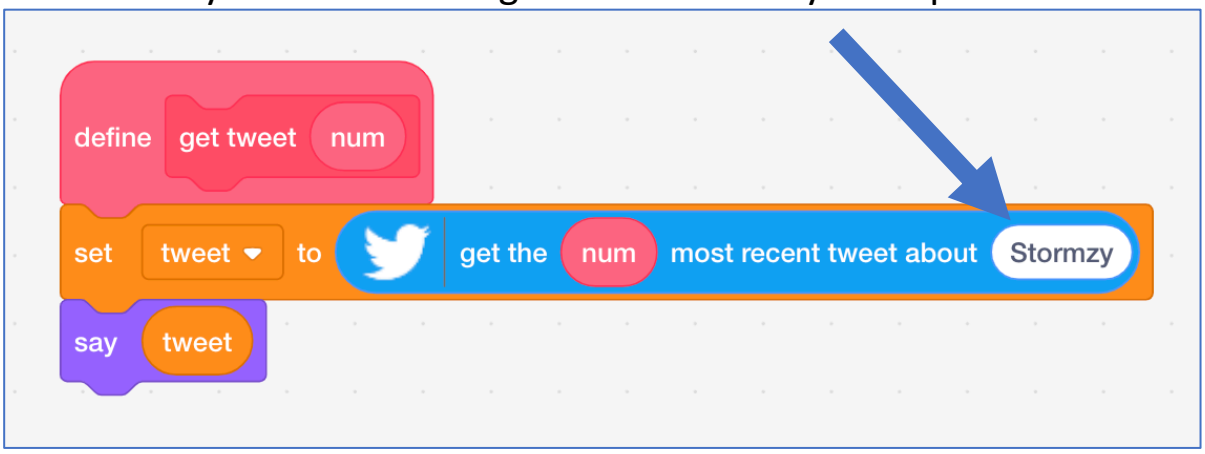
33. Find the “**What does Twitter think?**” project template and click on it



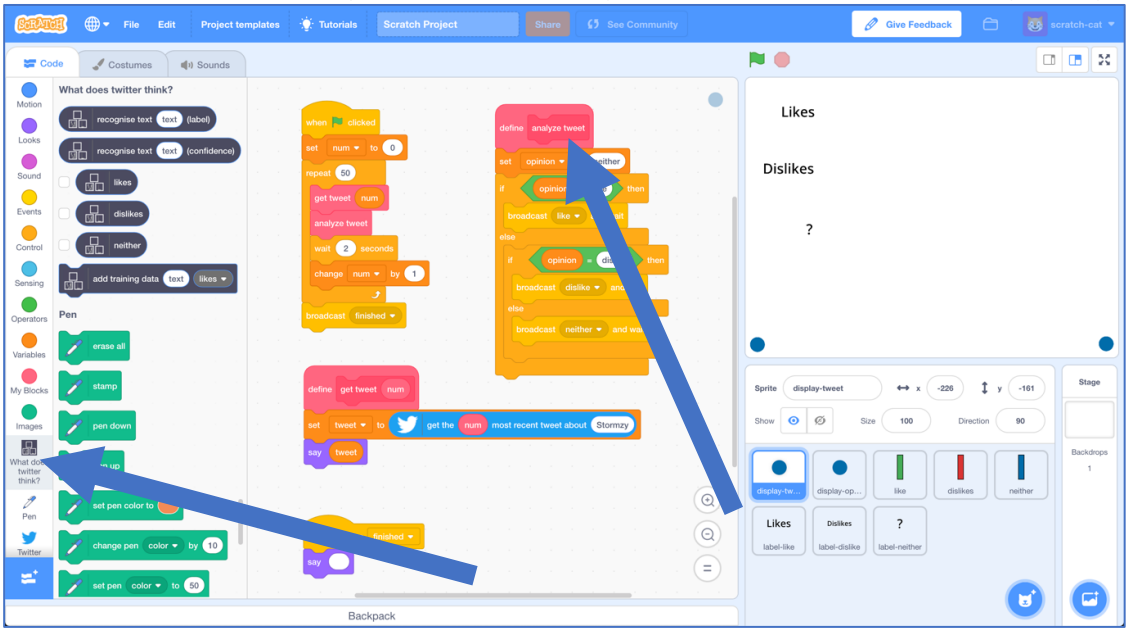
34. Find the “**get tweet**” script



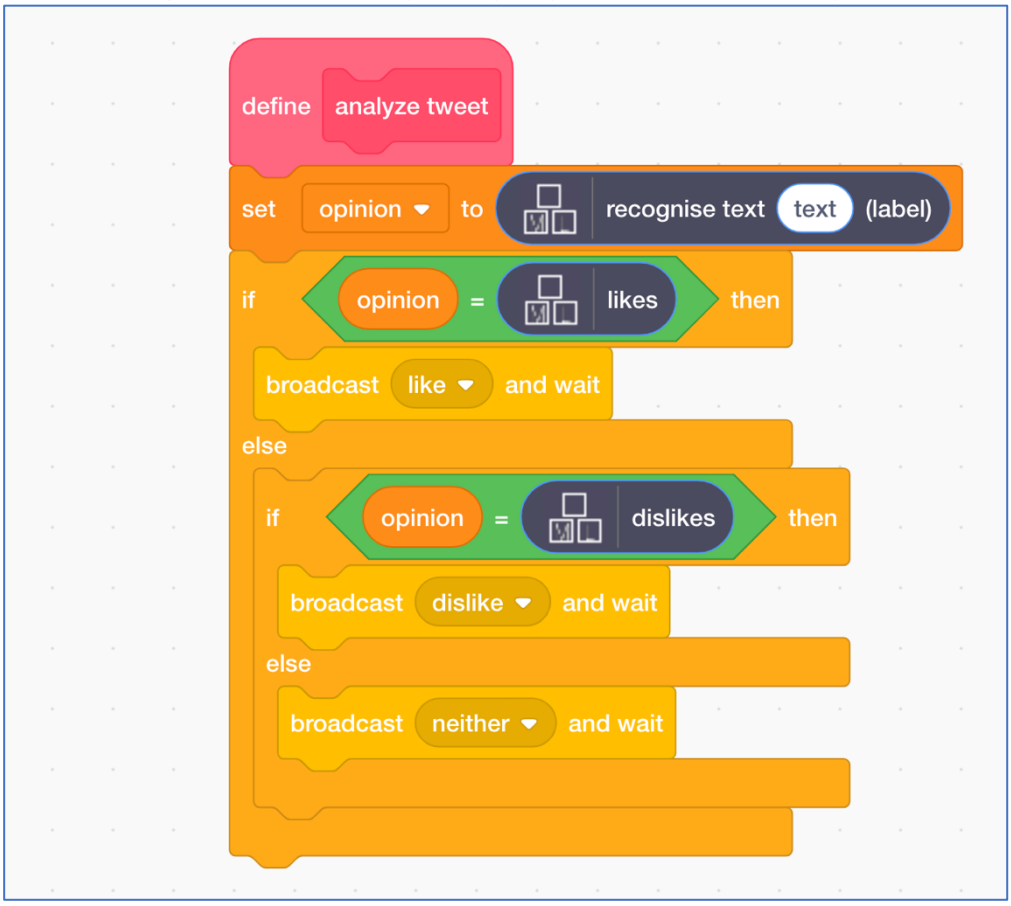
35. Modify it so that it will get tweets about your topic



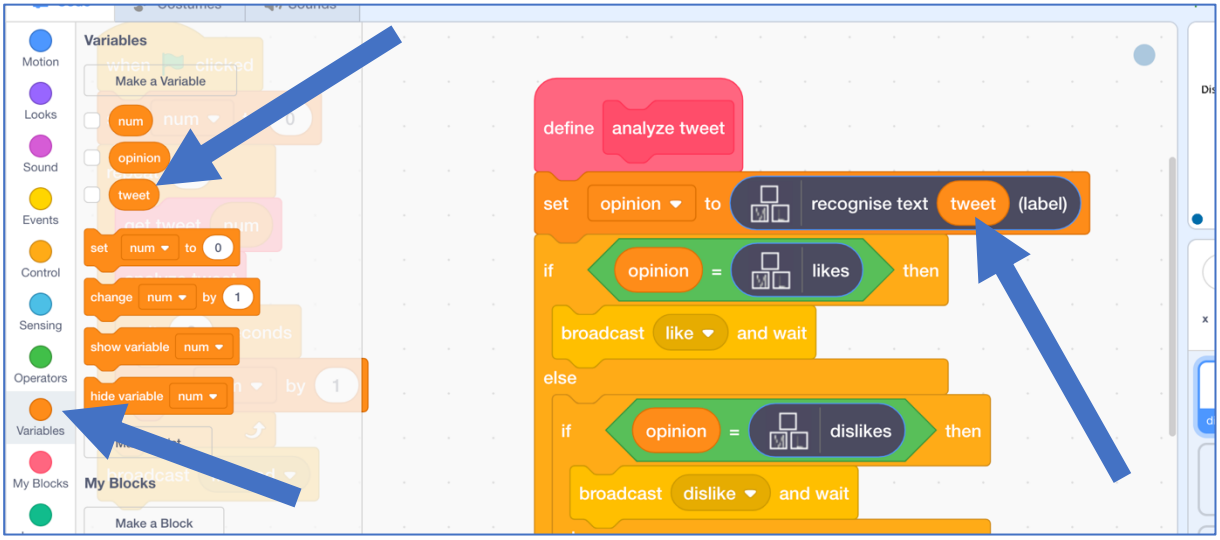
36. Find the “**analyze tweet**” script and the blocks from your project



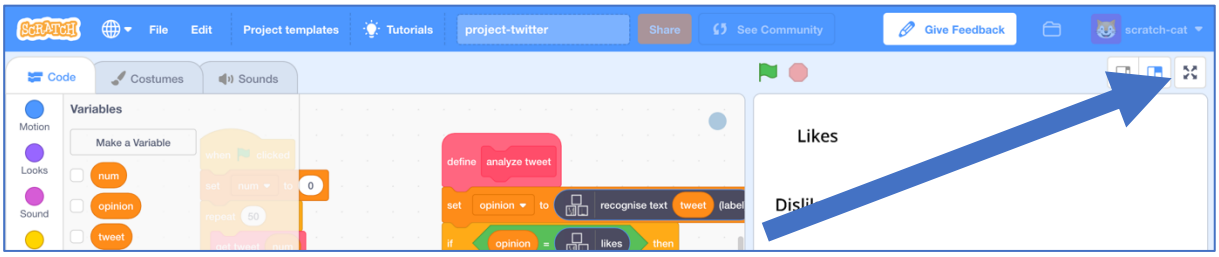
37. Drag blocks from your project into the script



38. Click on "**Variables**" in the left-hand side, and drag "**tweet**" into the "**recognise text**" block so that your machine learning model will analyze the next tweet

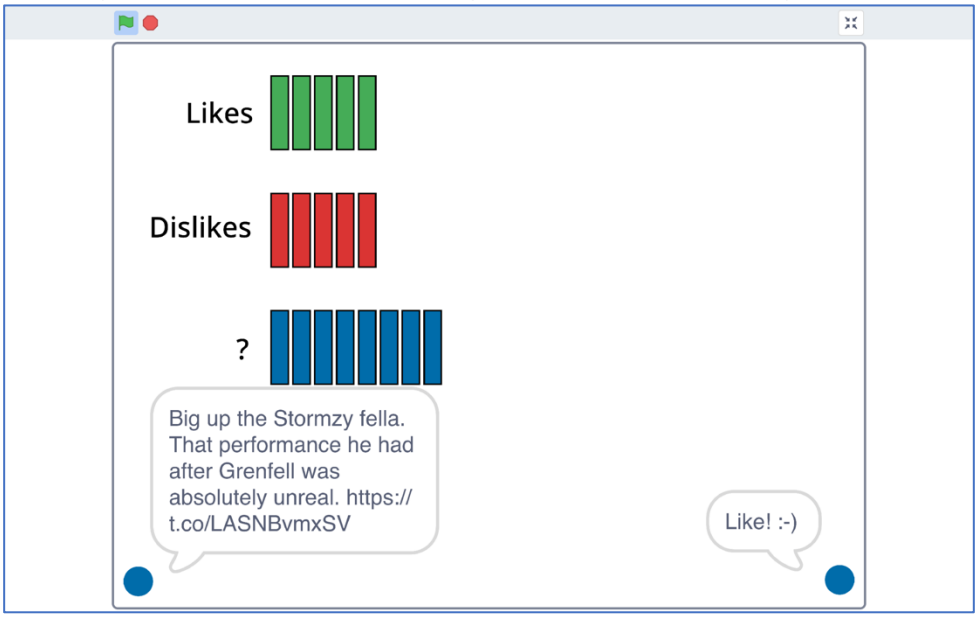


39. Time to test! Click the full-screen button.

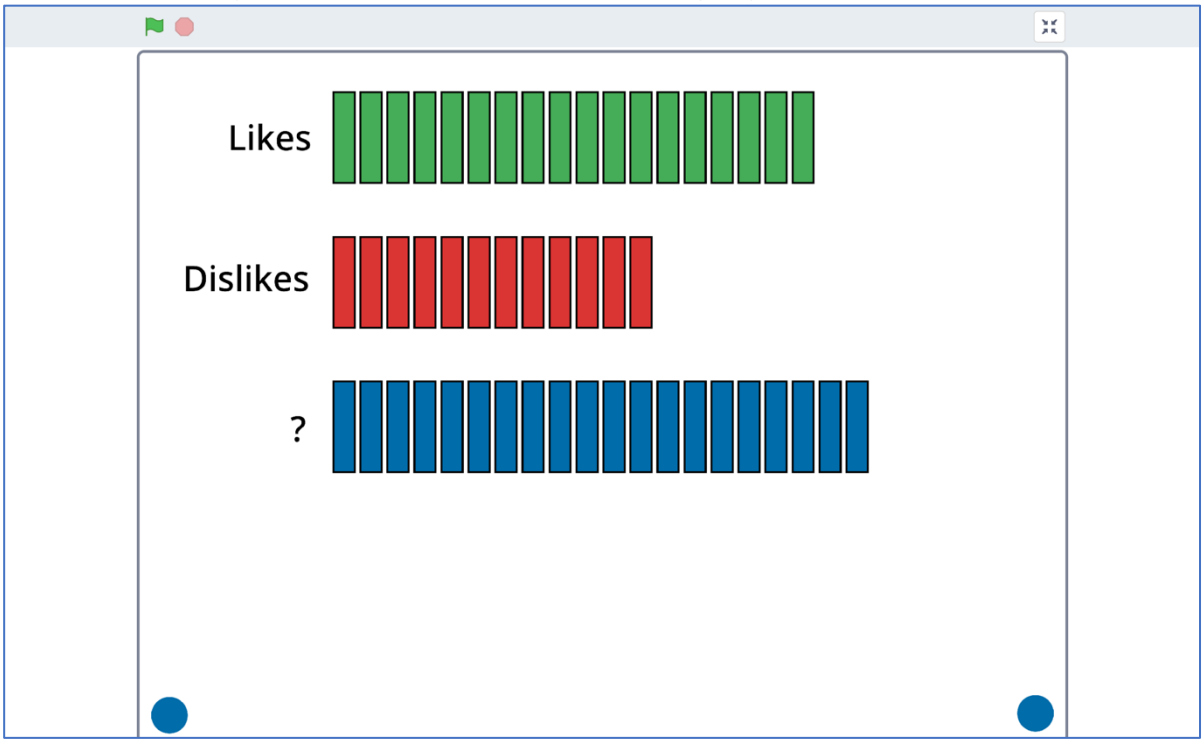


40. Click the green flag

The script will fetch 50 tweets about the topic, and draw a graph based on what your machine learning model thinks they are



41. The script will finish with the final graph



42. Save your project.

*Click **File** -> **Save to your computer***

What have you done?

You're using a type of natural language processing called sentiment analysis to measure the discussion about a topic on social media.

This is a very common usage of machine learning, to analyze what people think about everything from companies, retail products, and world issues.

With a small number of examples, your project will get a lot wrong, but the more examples you give it, the better it should get.

Even then, it will still make mistakes, but by making it easier to measure a very large number of messages quickly, this technique is still useful to give a quick estimate of the public mood.

Ideas and Extensions

Now that you've finished, why not give one of these ideas a try?

Or come up with one of your own?

Make your model more accurate

As the Scratch script plays, it displays what your machine learning model thought of each tweet. You'll probably disagree with some decisions your model makes.

Try to improve this by adding more examples in the "**Train**" page. Make sure you click the "**Train new machine learning model**" button again, to use those new examples. Then run your Scratch script again to see what difference it makes.

Write a Scratch script to train your model

Copying examples from another web browser is slow. Can you write a Scratch project to make this easier?

Use the "**get tweets**" block and the "**add training data**" block to make a project that will show you tweets, and if you press the "L" add them to the "likes" bucket and if you press the "D" add them to the "dislikes" bucket.

This will make it easier for you to collect training examples.

Use confidence scores

The confidence score block will tell you how sure your machine learning model is that it has correctly measured a tweet. You could use this so that the graph isn't updated unless the model is very confident.