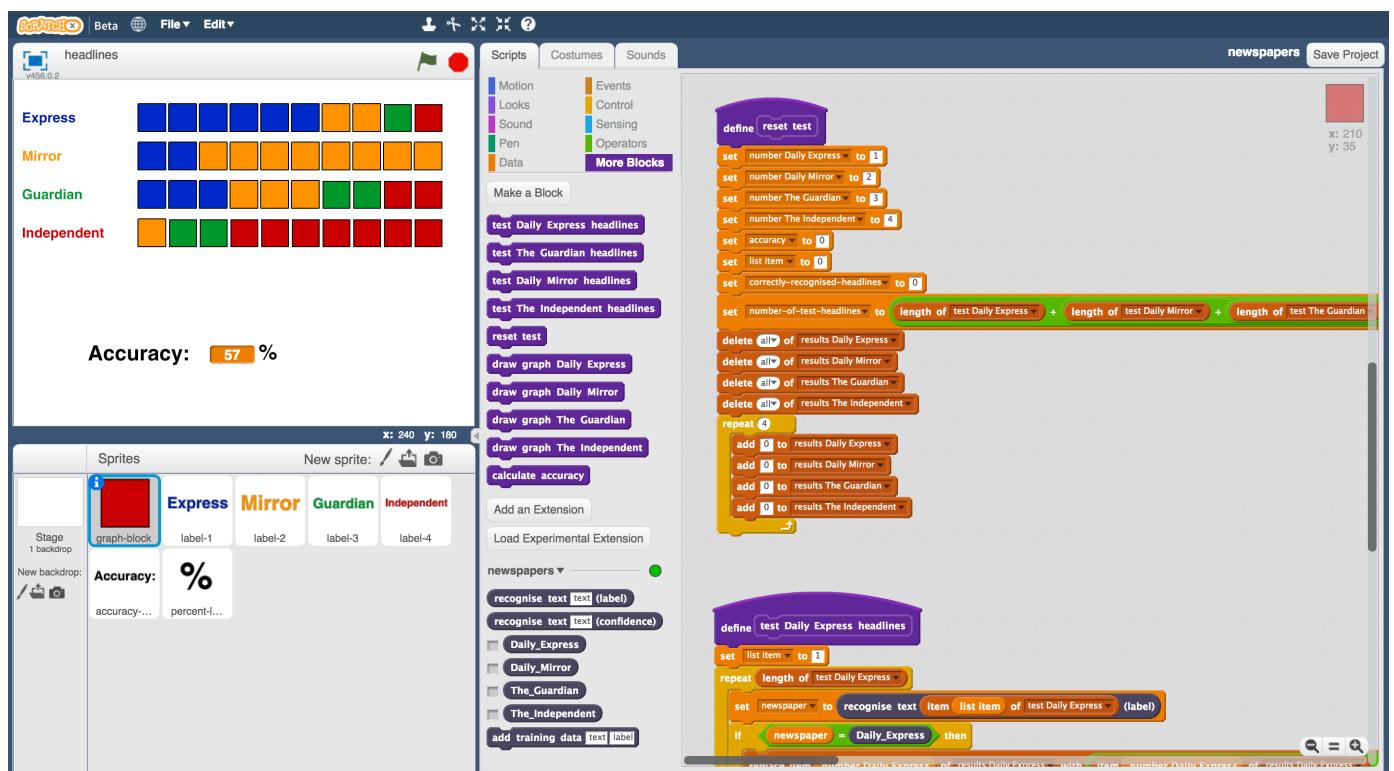


Headlines

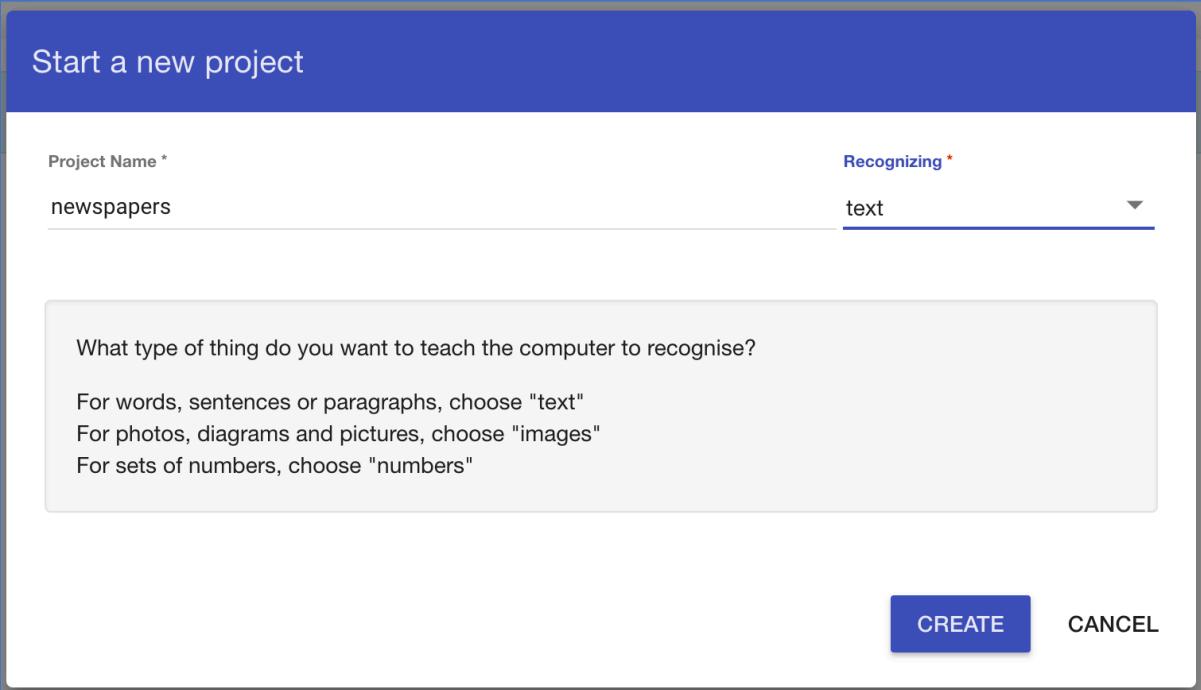
In this project, you will train a computer to recognise headlines from different newspapers.

You will collect examples of the front-page headlines from different newspapers. You'll use those to train a machine learning model that can learn from the use of language to be able to recognise which newspaper a headline is from.

You'll also learn about the importance of testing in machine learning projects, and create a simple testing tool in Scratch.



1. You'll need the **headlines.sbx** starter file for this project.
If you haven't got this, ask your teacher or group leader.
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
If you don't have a username, ask your teacher or group leader to create one for you.
If you can't remember your username or password, ask your teacher or group leader to reset it for you.
5. Click on “**Projects**” on the top menu bar
6. Click the “**+ Add a new project**” button.
7. Name your project “newspapers” and set it to learn how to recognise “**text**”



The screenshot shows a 'Start a new project' dialog box. At the top, there's a blue header bar with the text 'Start a new project'. Below this is a white form area. On the left, there's a 'Project Name *' field containing 'newspapers'. On the right, there's a 'Recognizing *' dropdown menu with 'text' selected. A large text input area below contains the question 'What type of thing do you want to teach the computer to recognise?' followed by three options: 'For words, sentences or paragraphs, choose "text"', 'For photos, diagrams and pictures, choose "images"', and 'For sets of numbers, choose "numbers"'. At the bottom right of the dialog are two buttons: a blue 'CREATE' button and a white 'CANCEL' button.

8. You should now see “newspapers” show up in the list of your projects. Click on it.

The screenshot shows a web interface titled "Your machine learning projects". At the top, there is a navigation bar with links: ml-for-kids, Welcome, About, Projects, Worksheets, News, Help, and Log Out. Below the navigation bar, there is a button labeled "+ Add a new project". A project card is displayed with the title "newspapers" and the subtitle "Recognising text". To the right of the project card is a trash can icon. The entire interface is enclosed in a blue border.

9. Click the “Train” button

The screenshot shows a project details page for "newspapers". At the top, there is a navigation bar with links: ml-for-kids, Welcome, About, Projects, Worksheets, News, Help, and Log Out. Below the navigation bar, the project name "newspapers" is displayed. There are three main buttons: "Train", "Learn & Test", and "Scratch". The "Train" button has the sub-instruction "Collect examples of what you want the computer to recognise." and a "Train" button. The "Learn & Test" button has the sub-instruction "Use the examples to train the computer to recognise text." and a "Learn & Test" button. The "Scratch" button has the sub-instruction "Use the machine learning model you've trained to make a game in Scratch." and a "Scratch" button. The entire interface is enclosed in a blue border.

10. Choose four national newspapers.

Pick different types. For example, two “tabloids” and “broadsheets”.

To save having to buy actual physical newspapers, you can use an online collection of newspaper front pages.

For the rest of this worksheet, I'll be using:

Daily Express, Daily Mirror, The Guardian, The Independent and getting the headlines from:

<https://www.thepaperboy.com/uk/front-pages.cfm>

11. Use the “+ Add new label” button to create a bucket for each newspaper you’re using.

The screenshot shows a web interface for training a machine learning model. At the top, a navigation bar includes links for ml-for-kids, Welcome, About, Projects, Worksheets, News, Help, and Log Out. Below the navigation, a title reads "Recognising **text** as **Daily_Express, Daily_Mirror or 2 other classes**". There are four rectangular training buckets labeled "Daily_Express", "Daily_Mirror", "The_Guardian", and "The_Independent". Each bucket has a "Add example" button at the bottom left. In the top right corner of the main area, there is a button labeled "+ Add new label".

12. In another web browser window, find newspaper front pages. *Arrange the windows so your training buckets are next to the book site.*

This screenshot displays two separate web browser windows. The left window is identical to the one in step 11, showing four training buckets for news classification. The right window is titled "Newspaper Headlines for Mon..." and displays three newspaper front pages: "DAILY EXPRESS", "DAILY MAIL", and "MIRROR". Each front page includes a headline, a small image, and a link to "Click Here to Switch to Front Pages View". Below each front page, there is a list of links: "Visit Daily Express Web Site", "Read the Digital Edition of Daily Express at PressDisplay", and "View Daily Express Twitter Feed".

13. Agree on some rules for the examples of headlines that you'll collect, so that the training is consistent.

For the rest of this worksheet, I'll be using:

- * Newspapers from weekdays only (Mon-Fri) not weekend editions
- * Only the biggest main headline from each newspaper

14. Click the “+ Add example” button in one of the buckets and copy the newspaper headline into the “Add new example” dialog.

The screenshot shows a user interface for training a machine learning model to recognize newspaper headlines. On the left, there's a list of four news sources: Daily_Express, Daily_Mirror, The_Guardian, and The_Independent. Each source has a "Daily_Express" or "Daily_Mirror" label above it and a "Add example" button below its respective list item. A central modal window titled "Add new example" is open, showing the headline "FURY AT PLOT TO GET NEW EU VOTE". Below the headline, there's a progress bar at 31 / 1000. At the bottom of the modal are "ADD" and "CANCEL" buttons. To the right of the modal, there's a summary section titled "Newspaper Headlines for 14 August 2017" with two examples shown:

- 8/14/2017: FRONT PAGE: FURY AT PLOT TO GET NEW EU VOTE**
FEARS were raised last night that opponents of Brexit will push for a second referendum. Labour veteran David Miliband provoked outrage by insisting there should be the terms of Britain's final exit deal with the...
 - Visit Daily Express Web Site
 - Read the Digital Edition of Daily Express at PressDisplay
 - View Daily Express Twitter Feed
- 8/14/2017: Front Page: WAR ON HEART DEATHS**
MILLIONS are to be offered checks at GP surgeries and pharmacies in an NHS drive to detect disease and early deaths. The strategy aims to raise dramatically the detection conditions including high blood pressure, high...
 - Visit Daily Mail Web Site
 - Read the Digital Edition of Daily Mail at PressDisplay
 - View Daily Mail Twitter Feed

At the bottom right, there's a link to "Visit Daily Mirror Web Site".

15. Repeat this for all four newspapers on this day

This screenshot shows the same machine learning interface after adding more examples for each newspaper. The four news sources are now listed with their respective headlines and "Add example" buttons:

- Daily_Express**: FURY AT PLOT TO GET NEW EU V...
- Daily_Mirror**: ANT: MY FIGHT TO SAVE MARRIA...
- The_Guardian**: Trump under fire from all sides as ...
- The_Independent**: May's plan to pack Brexit committ...

On the right, there are four corresponding news snippets with links to visit the websites:

- 8/14/2017: FRONT PAGE: Total Football podcast**
Politicians from all sides have rounded on Donald Trump under fire from all sides as a woman dies at far-right rally...
 - Visit The Daily Telegraph Web Site
 - Read the Digital Edition of The Daily Telegraph at PressDisplay
 - View The Daily Telegraph Twitter Feed
- 8/14/2017: FRONT PAGE: Trump under fire from all sides as woman dies at far-right rally**
Politicians from all sides have rounded on Donald Trump under fire from all sides as a woman dies at far-right rally...
 - Visit The Guardian Web Site
 - Read the Digital Edition of The Guardian at PressDisplay
 - View The Guardian Twitter Feed
- 8/14/2017: Front Page: Charlottesville's innocent victim**
Heather Heyer, 32, was killed when a car driven by a suspect drove at peaceful demonstrators...
 - Visit The Independent Web Site
 - Read the Digital Edition of The Independent at PressDisplay
 - View The Independent Twitter Feed
- 8/14/2017: Front Page: Scottish runners lead the way as Team GB Ian**
Scots Zoey Clark, left, and Eilidh Doyle, second left, celebrate with teammates Emily Diamond and Lavai Nielsen after winning...
 - Visit THE SCOTSMAN Web Site
 - Read the Digital Edition of THE SCOTSMAN at PressDisplay
 - View THE SCOTSMAN Twitter Feed

16. Repeat this for five days of newspaper headlines.

If you're using front pages from thepaperboy.com you can use the calendar at the bottom of the page to switch to a different day.

Remember the rules: Weekdays only. Biggest headline you see.

Tip: Use the 'Front Page Headline View' to get some headlines as text you can copy to save you some typing.

17. You should now have examples from one week of newspapers Keep the newspaper site window open – you'll need it again later.

The screenshot shows a web-based machine learning project interface. At the top, there's a navigation bar with links: ml-for-kids, Welcome, About, Projects, Worksheets, News, Help, and Log Out. Below the navigation, a title reads "Recognising **text** as **Daily_Express, Daily_Mirror or 2 other classes**". A link "[< Back to project](#)" is visible. On the right, there's a button "+ Add new label". The main area contains four boxes, each representing a news outlet:

- Daily_Express**: Contains headlines like "FURY AT PLOT TO GET NEW EU VO...", "WINE IS KEY TO A LONGER LIFE", "HOUSE PRICES UP BY £10,000", "2.37 MILLION EU MIGRANTS WOR...", and "SLAUGHTERED ON THE STREET".
- Daily_Mirror**: Contains headlines like "ANT: MY FIGHT TO SAVE MARRIAGE", "SO VULGAR (EVEN FOR YOU, SIR ...)", "FREEZE FARES", "BOMB HERO 'STOLE FROM VICTIMS'", and "MASSACRE".
- The_Guardian**: Contains headlines like "Trump under fire from all sides as wo...", "'Racism is evil', says Trump, 48 hour...", "EU chiefs condemn 'fantasy' of UK ...", "FA under pressure to explain payme...", and "Terror strikes Barcelona".
- The_Independent**: Contains headlines like "May's plan to pack Brexit committee...", "Government demands tariff-free trad...", "Care home crisis: places must doubl...", "Government crackdown on universit...", and "Terror in Barcelona: 12 dead in van a...".

Each box has a "+ Add example" button at the bottom.

18. Click the "< Back to project" link.

19. Click "Learn & Train".

20. Click "Train new machine learning model".

As long as you've collected enough examples, the computer should start to learn how to recognise headlines from the examples you've given to it.

21. Wait for the training to complete. This might take a few minutes.

22. Try the model out by putting a headline into the **Test** box

The screenshot shows a web-based machine learning application. At the top, it says "Machine learning models". Below that, there are two sections: "What have you done?" and "What's next?". The "What have you done?" section contains text about training a model to recognize newspaper names and a list of collected examples. The "What's next?" section provides instructions for testing the model with new text. At the bottom, there is a large input field with placeholder text "Try putting in some text to see how it is recognised based on your training.", a text input box containing "REVOLT OVER £36BN EU DIVORCE", and a "Test" button.

What have you done so far?

You've started to train a machine learning model to recognise patterns in the way that different newspapers use words, phrasing and punctuation in their front-page headlines.

But is it working well? How much has it learned so far? Does it need more examples to be able to recognise headlines correctly? What is it good at? What is it bad at? Which newspapers does it get confused between?

To answer those questions, you need to test the machine learning model. Next, you'll build a test in Scratch to do this.

23. Click the “< Back to project link”. Then click “Scratch”.

Using machine learning in Scratch

< Back to project

Your project will add these blocks to the **More Blocks** tab in Scripts.

recognise text [text] (label)
Put text in the input for this, and it will return the label that your machine learning model recognises it as.

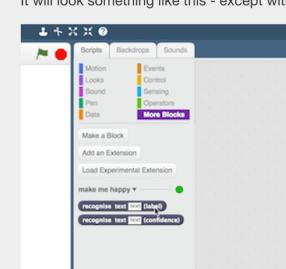
recognise text [text] (confidence)
This will return how confident your machine learning model is that it recognises the type of text. (As a number from 0 - 100).

Daily_Express Daily_Mirror The_Guardian The_Independent
These blocks represent the labels you've created in your project, so you can use their names in your scripts.

This means you can do something like this:



It will look something like this - except with the name of your project.



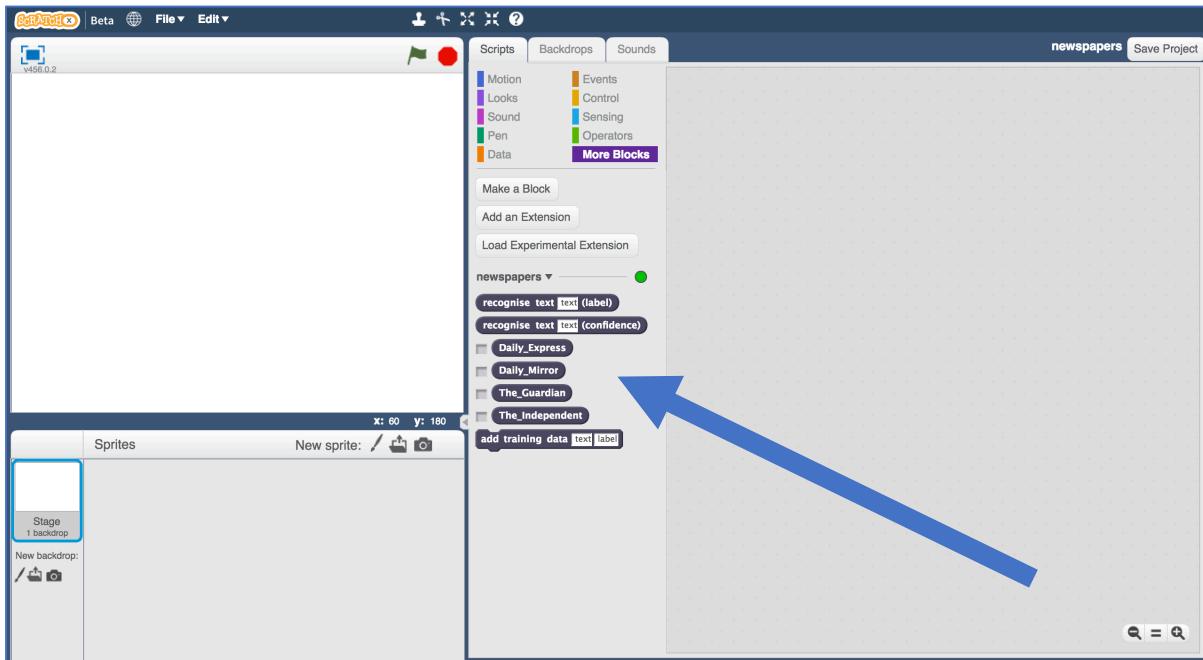
The coloured circle next to your project name tells you if your machine learning model is okay.

- **Green** means your model is trained and ready to go
- **Yellow** means your model hasn't finished training yet
- **Red** means something went wrong. Go back to the [Learn & Test](#) page to see what went wrong with training.

[Open in Scratch](#)

24. Click the “Open in Scratch” button to start making the test.

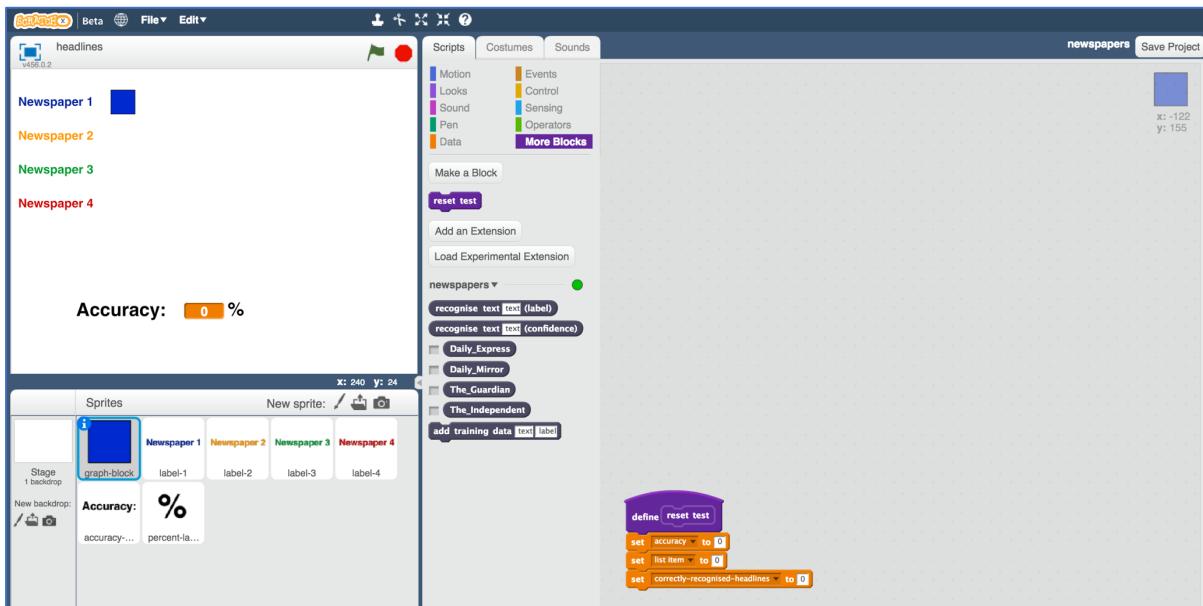
You should see new blocks in the “More blocks” section from your “newspapers” project.



25. Open the headlines.sbx starter project.

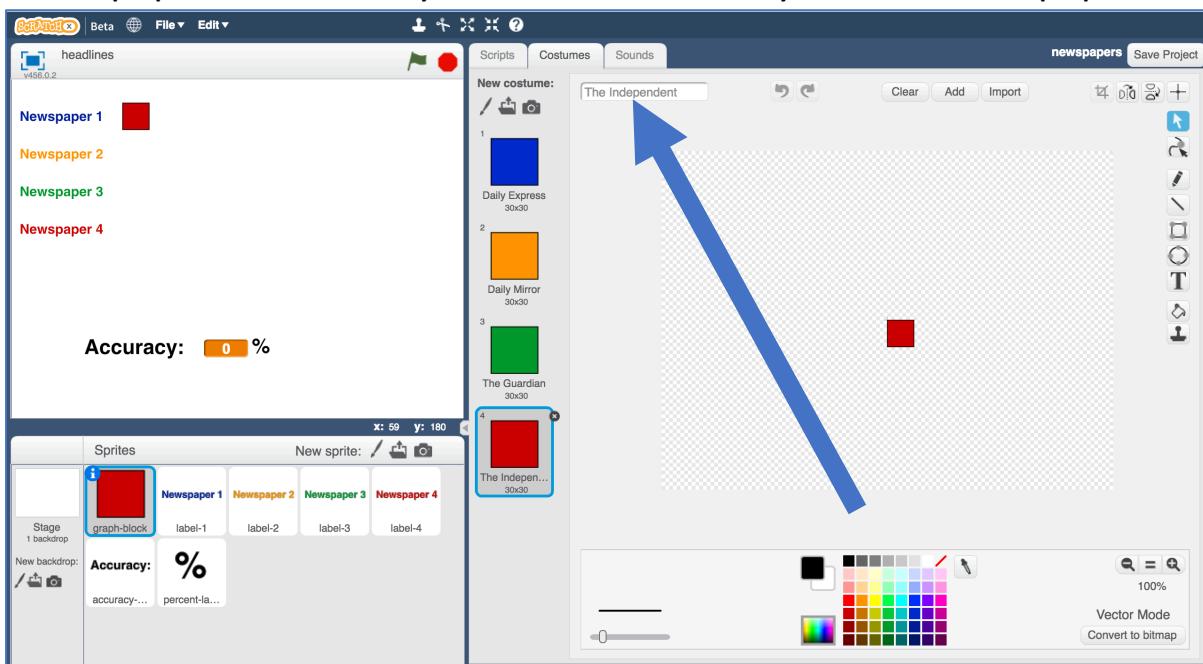
Use **File -> Load Project**

Click **OK** if it asks if you want to replace the current project

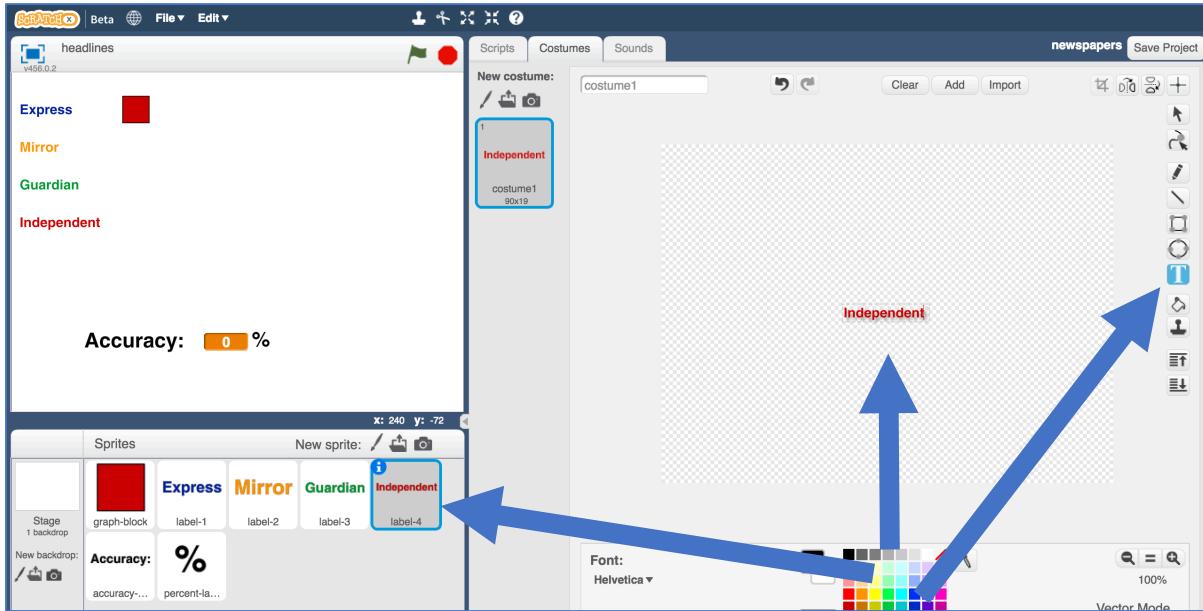


26. Click on the Costumes tab of the graph-block sprite.

Rename the costumes, so instead of being called “newspaper-1”, “newspaper-2”, etc. they have the names of your four newspapers.



27. Update the text in the label-1, label2, label-3, label4 sprites
Make sure that they match the order, and the colours, for the names you put in the graph-block sprite
Click on the costume using the T text tool to change the text.
You can use a shorter version of the names if it helps them fit better.

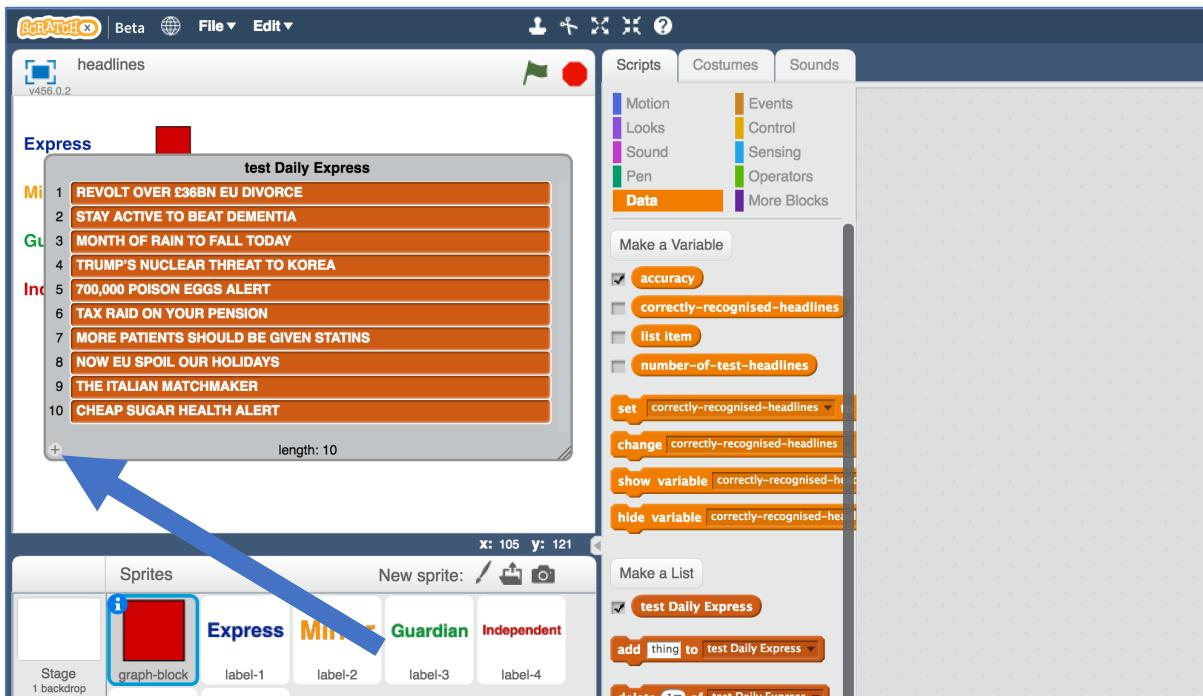


28. You will need **10** more front-page headlines from each of your four newspapers.
Note: It is very important that none of these are examples you used to train the model.
Ten more examples means another two week's worth of newspapers.

29. Click the Data tab, and then click the **Make a List** button
Create a list called “test Daily Express” for all sprites.
If you’re using different newspapers, update the name to match your first newspaper.

30. Enter your 10 test headlines from that newspaper.

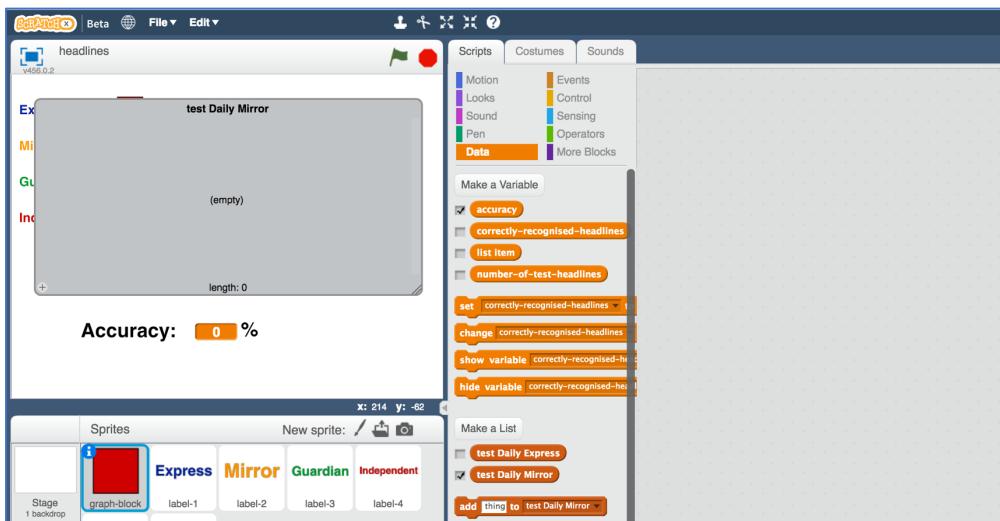
*Resize the list window to make it bigger: that'll make it easier to type into
Click the + button in the bottom left corner to add each new test headline*



31. Untick the “test Daily Express” list so that it is hidden

32. Click the Make a List button again, and create a second list for your second newspaper. Call it “test Daily Mirror”

If you're using different newspapers, update the name to match your second newspaper.

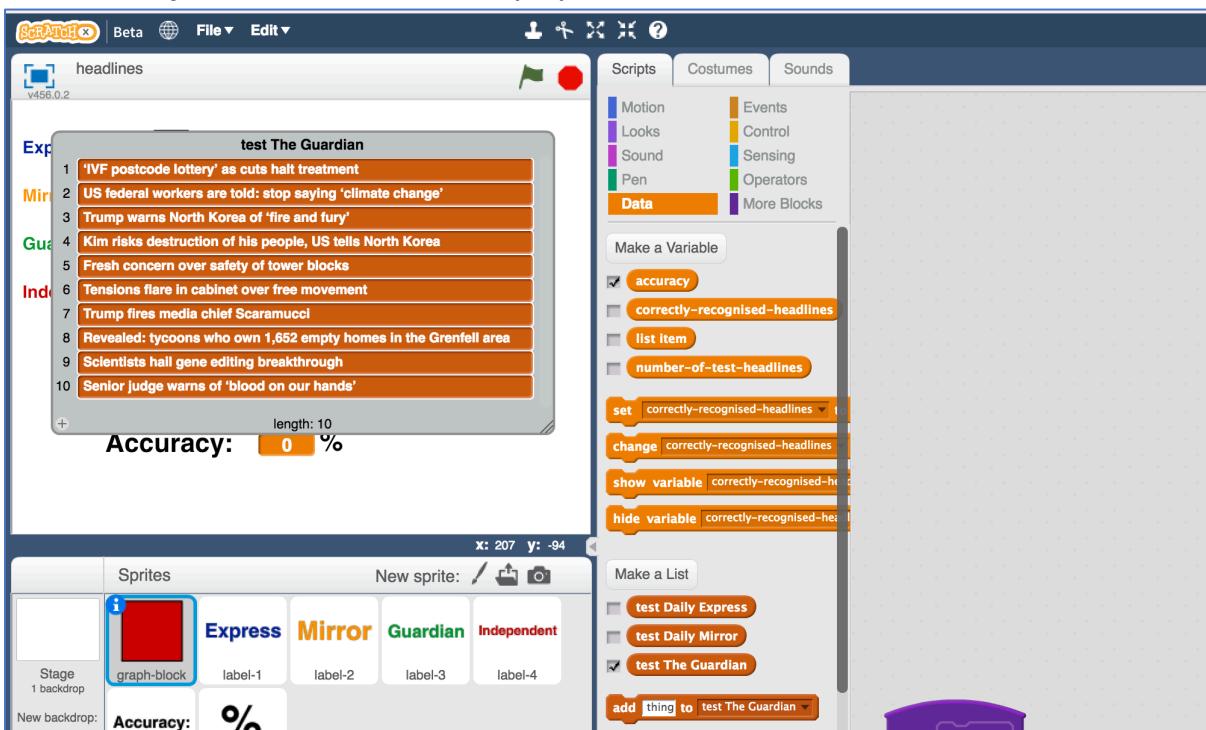


33. Type the 10 test headlines from your 2nd newspaper into this list.

34. Untick the list so that the second list is hidden like the first

35. Repeat for the third newspaper

Create a new list with a name that starts with "test". Add the ten test headlines from the third newspaper to the list.



36. Hide the third list by unticking it

37. Repeat for the fourth newspaper

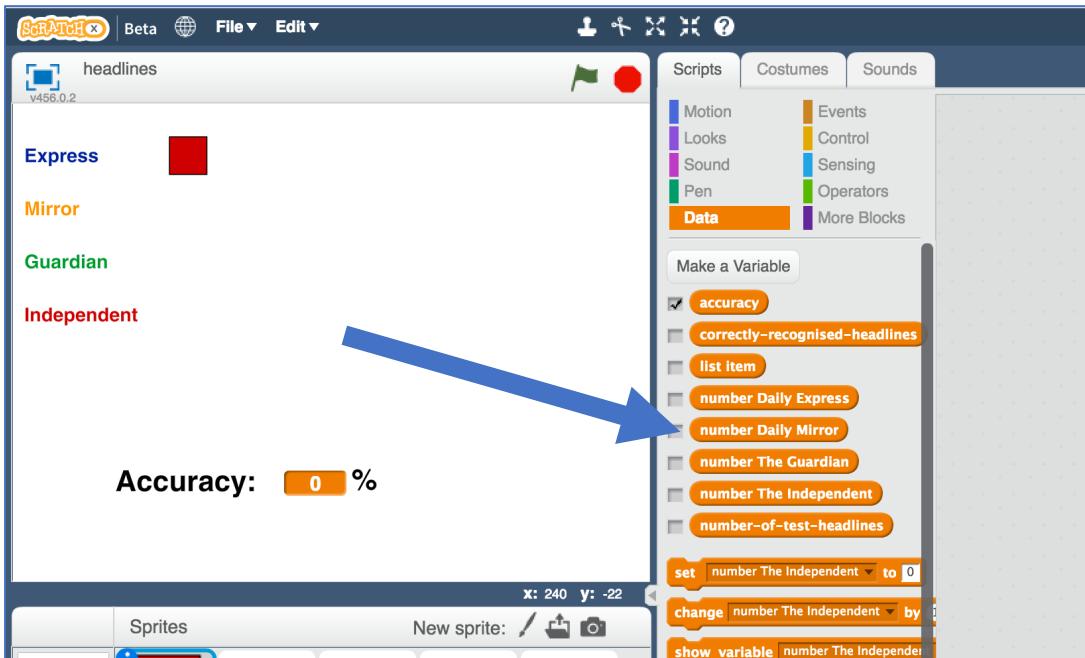
Create a new list with a name that starts with "test". Add the ten test headlines from the fourth newspaper. Untick the list to hide it.

38. Create four new variables

Create a variable by clicking the **Make a Variable** button on the Data tab.

Create them for all sprites. Untick them so that they are hidden.

Name them “number Daily Express”, “number Daily Mirror”, etc. (using the names of your four newspapers).

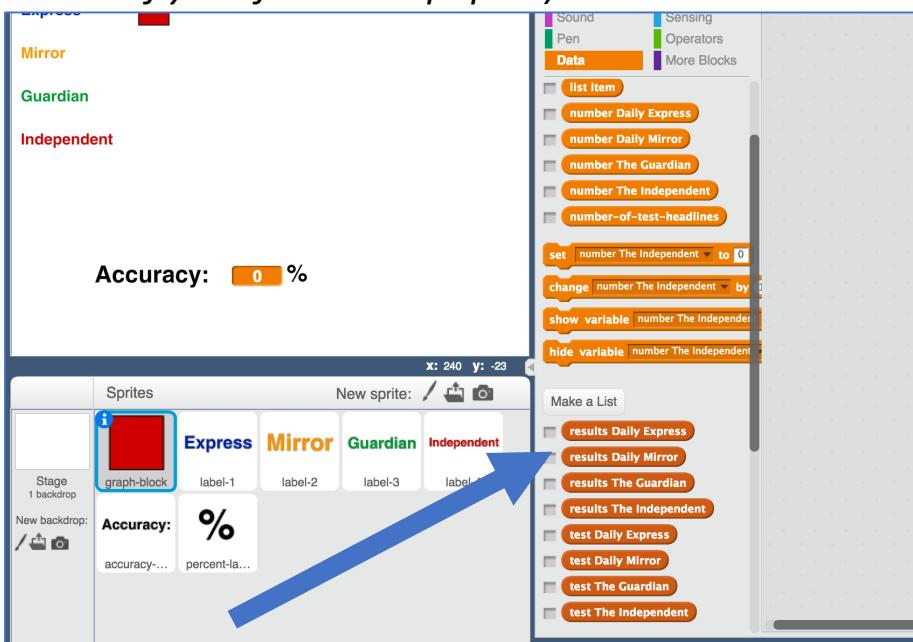


39. Create four new lists

Create a list by clicking the **Make a List** button on the Data tab.

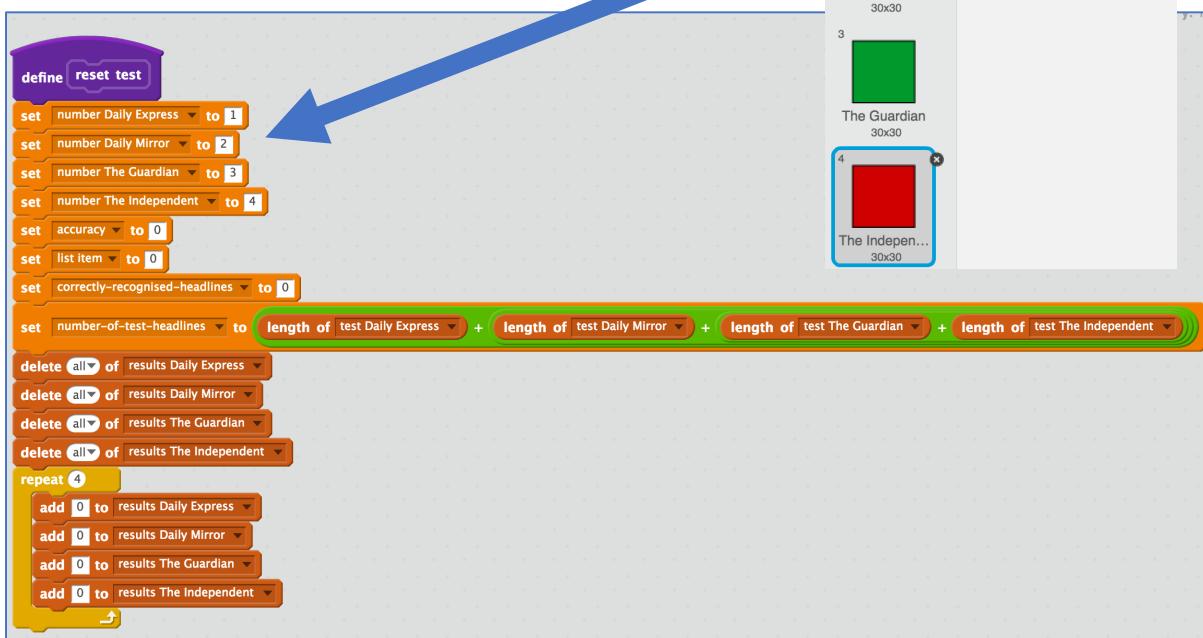
Create them for all sprites. Untick them so that they are hidden.

Name them “results Daily Express”, “results Daily Mirror”, etc. (using the names of your four newspapers).



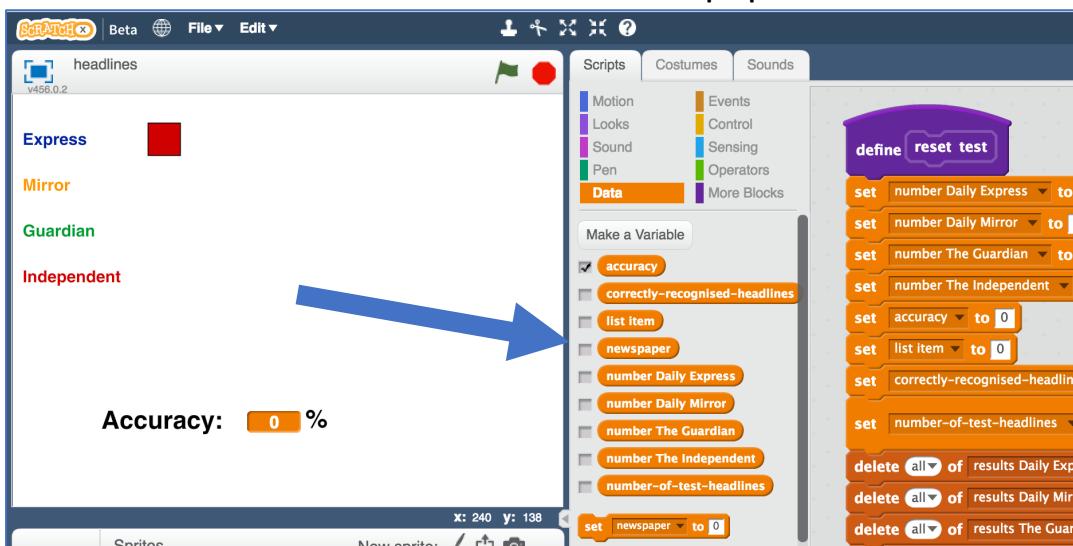
40. Update the script for the **reset test** block to reset the data

The numbers for the first four blocks (1,2,3,4) should match the costume numbers for the newspapers in graph-block. This also means they should be in the same order as the labels on the stage.

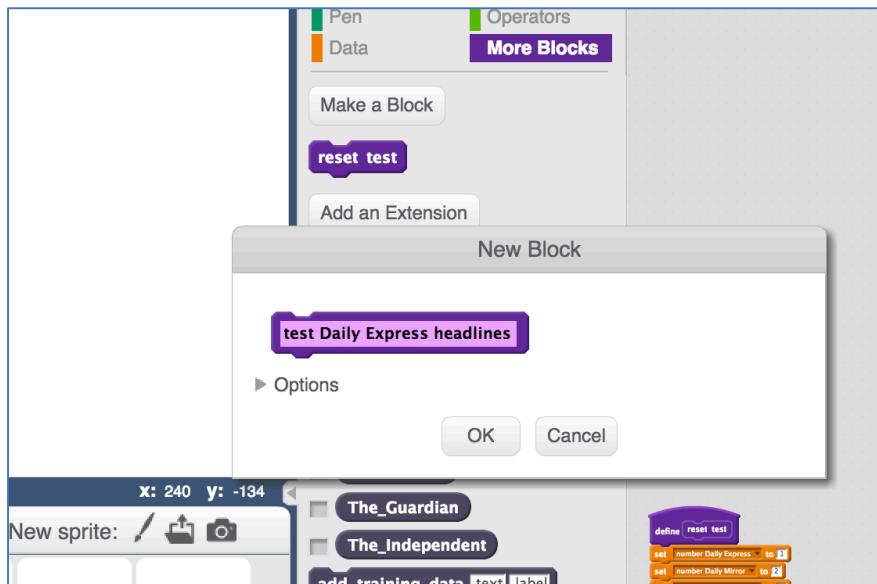


The final '4' in the repeat block is because there are four newspapers.

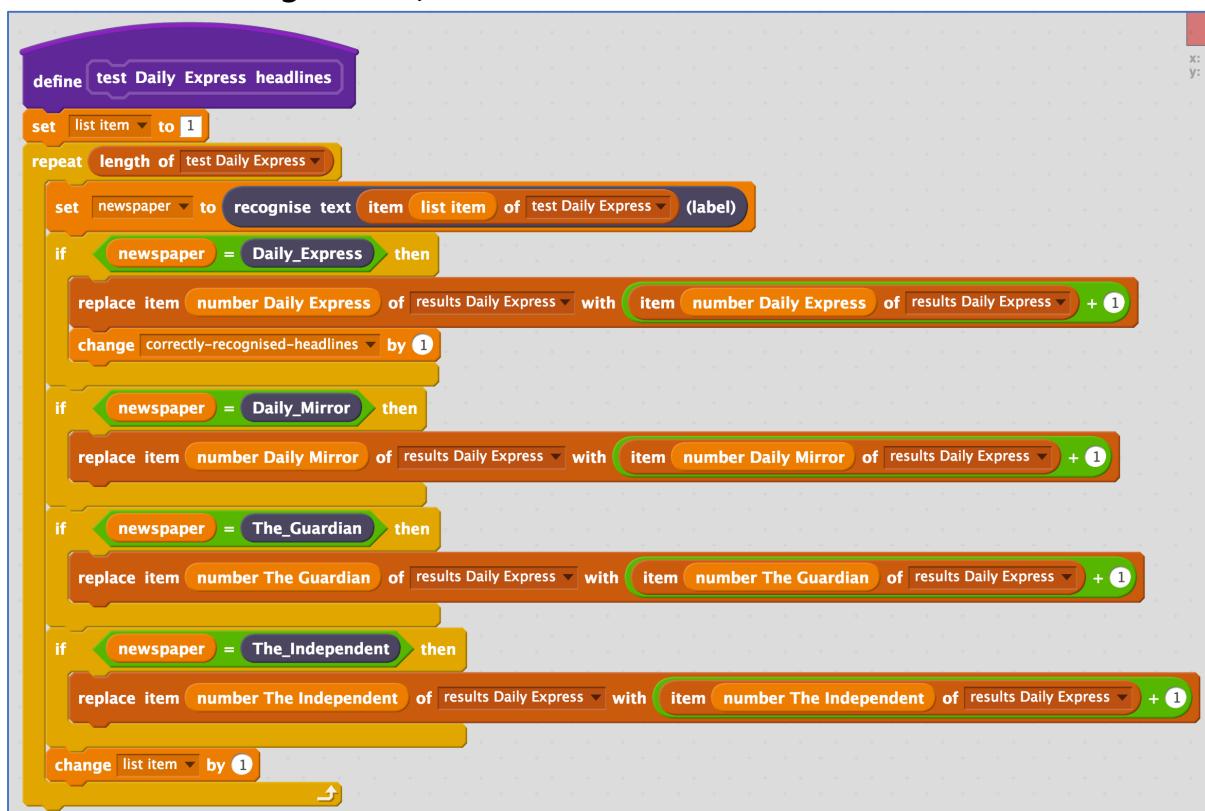
41. Create a new variable called “newspaper”. Untick it to hide it.



- 42.** Create a new block to test your first newspaper
*Click on the graph-block sprite. Click the **Scripts** tab.
 Click the purple **More Blocks** item. Click the **Make a Block** button.
 Name it “test Daily Express headlines”*



- 43.** Enter the following script into the new custom block
It will run your test headlines from the Daily Express through your machine learning model, and count the results it returns.



44. Create a new block to test the headlines from your second newspaper

Click on the graph-block sprite.

*Click on the **Scripts** tab. Click the purple **More Blocks** item.*

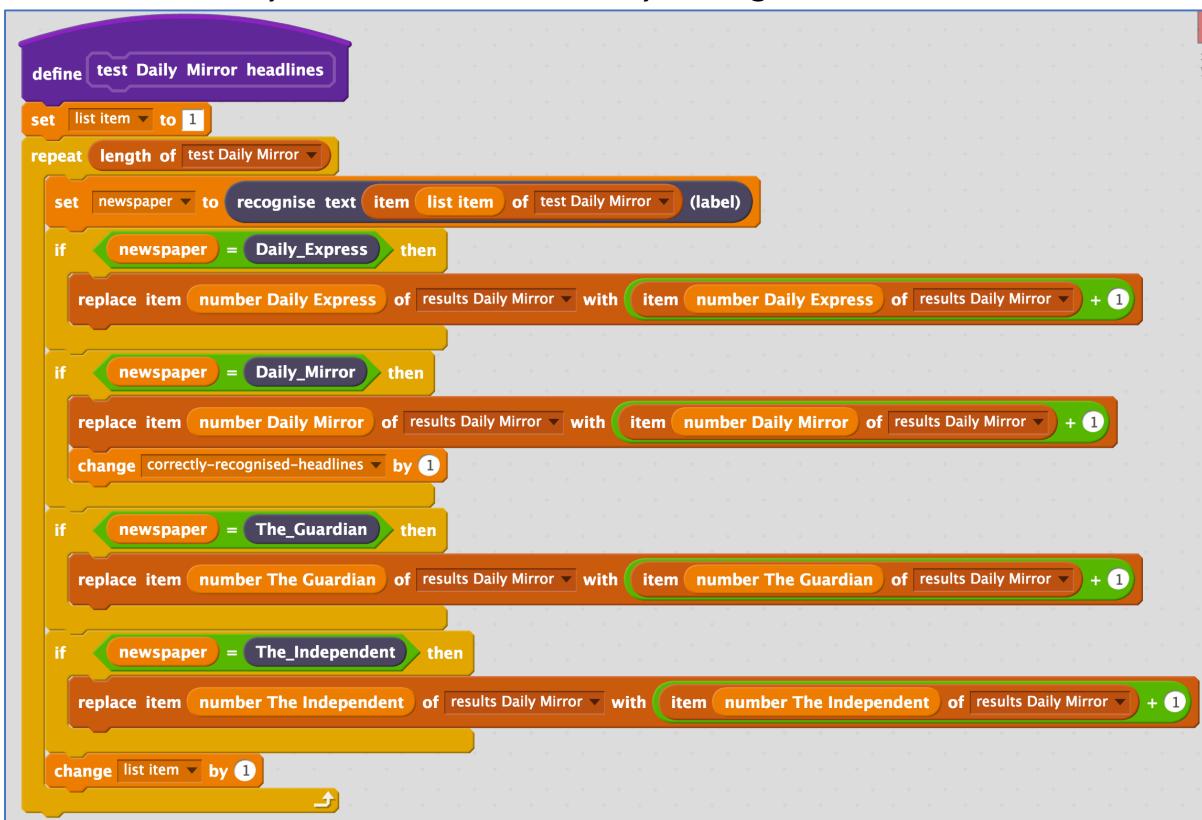
*Click the **Make a Block** button. Name it “test Daily Mirror headlines”*

45. Add the following script to the custom block.

Notice how similar it is to the last one.

You can save yourself some time by duplicating the contents of your last script (right-click on the contents and click on Duplicate).

If you do this, make sure you update it to use the Daily Mirror lists and also make sure you move the correctly-recognised-headlines block.



46. Create two more custom blocks to test the third and fourth newspapers.

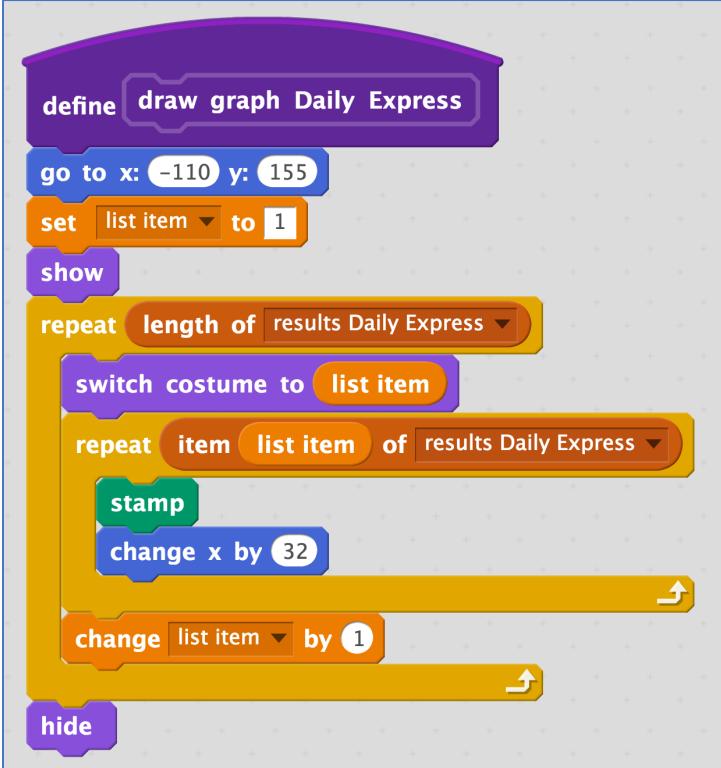


- 47.** Create a new block to draw a graph of the results from the test of the headlines from the first newspaper

Click on the graph-block sprite.

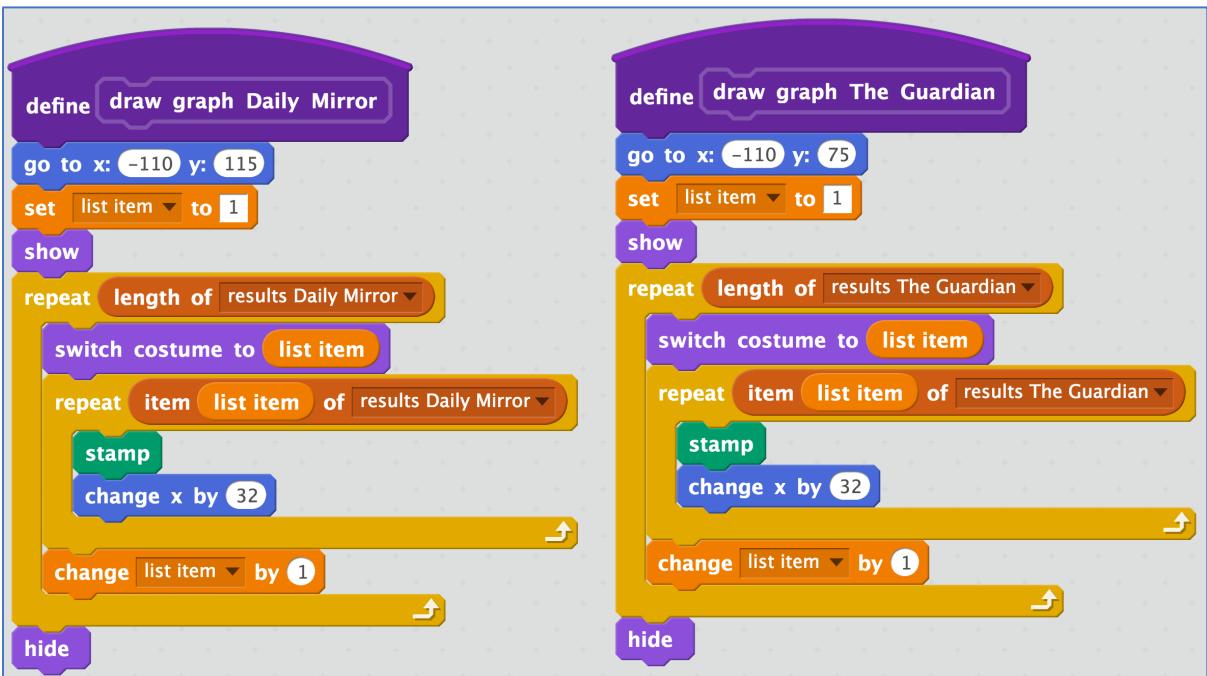
*Click on the **Scripts** tab. Click the purple **More Blocks** item.*

*Click the **Make a Block** button. Name it “draw graph Daily Express”*



```
define draw graph Daily Express
  go to x: -110 y: 155
  set [list item v] to [1]
  show
  repeat (length of [results Daily Express v])
    switch costume to [list item v]
    repeat (item [list item v] of [results Daily Express v])
      stamp
      change x by (32)
    end
    change [list item v] by (1)
  end
  hide
```

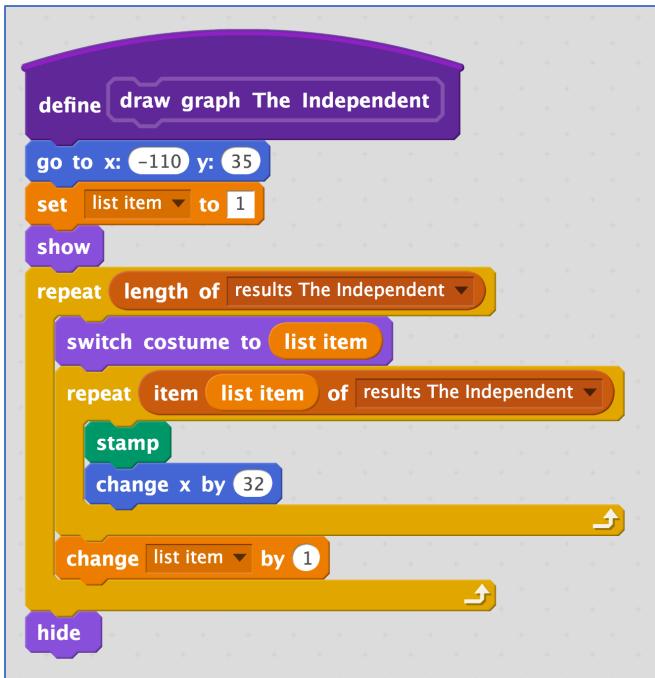
- 48.** Create blocks to draw graphs for the other three newspapers.



```
define draw graph Daily Mirror
  go to x: -110 y: 115
  set [list item v] to [1]
  show
  repeat (length of [results Daily Mirror v])
    switch costume to [list item v]
    repeat (item [list item v] of [results Daily Mirror v])
      stamp
      change x by (32)
    end
    change [list item v] by (1)
  end
  hide
```

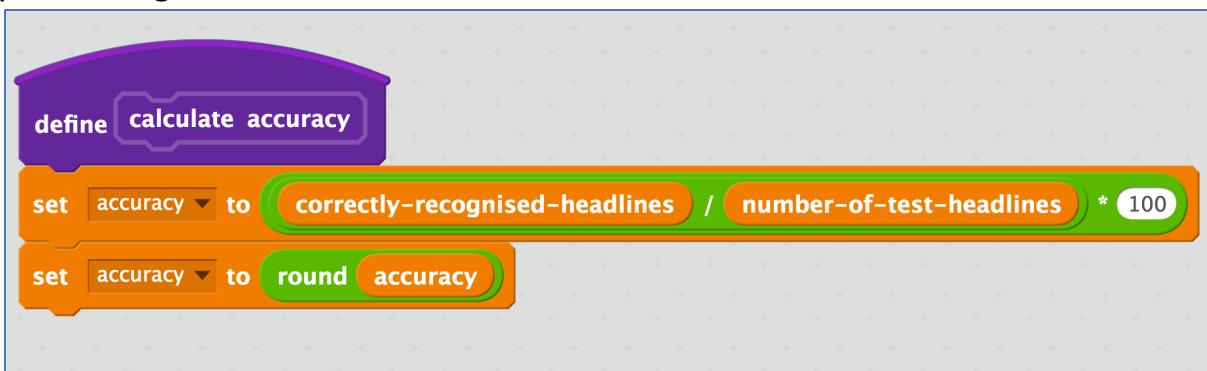


```
define draw graph The Guardian
  go to x: -110 y: 75
  set [list item v] to [1]
  show
  repeat (length of [results The Guardian v])
    switch costume to [list item v]
    repeat (item [list item v] of [results The Guardian v])
      stamp
      change x by (32)
    end
    change [list item v] by (1)
  end
  hide
```



- 49.** Create a new block to calculate how accurate your machine learning model is

Create this block in the graph-block sprite along with all the other scripts. This calculates how many headlines are correctly recognised as a percentage.

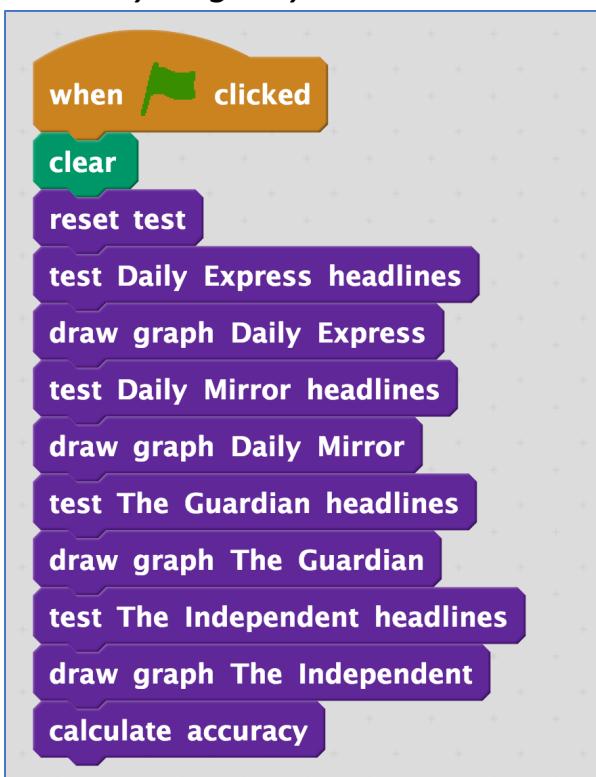


50. Time to put it all together!

Create this script in the graph-block sprite.

It will test your machine learning model using all of your test headlines from all the papers.

It will draw a graph to show you the results, and calculate the overall accuracy to give you a score.

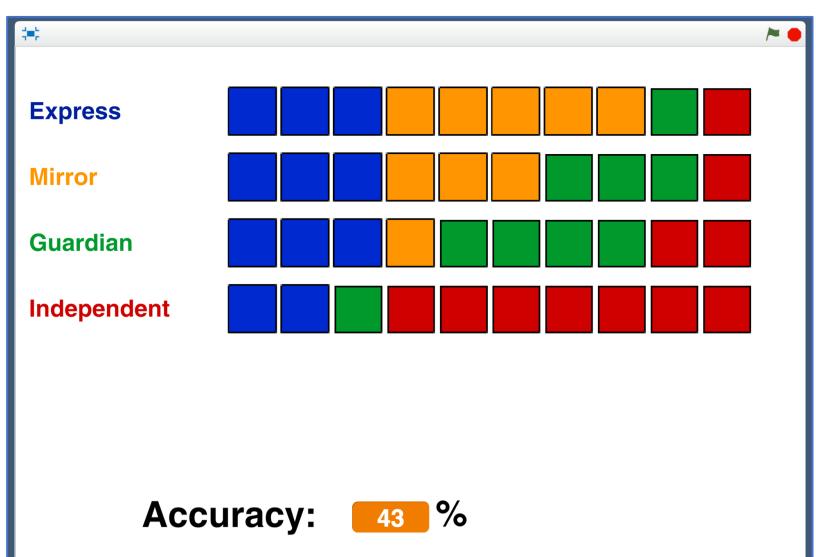


51. Run the test

Click the green flag, and watch it run.

Each row is made up of 10 blocks.

The 10 blocks represent the 10 test headlines from a newspaper.



Independent



I collected 10 headlines from front-pages of *The Independent*.

My machine learning model “thought” 2 of them were from the *Daily Express* (blue).

It “thought” that 1 of them was from *The Guardian* (green).

It correctly recognised the other 7 as being headlines from *The Independent*.

That was pretty good!

Express



I collected 10 headlines from front-pages of the *Daily Express*.

The machine learning model correctly recognised 3 of them.

It thought 5 were from the *Daily Mirror*, 1 was from *The Guardian*, and it thought 1 looked like a headline from *The Independent*.

Not so good. It looks like it's having trouble telling the difference between *Daily Express* and *Daily Mirror* headlines.

Accuracy: 43 %

Overall, my machine learning model correctly recognised 43% of the test headlines.

What have you done so far?

You've started to train a machine learning model to recognise newspaper front-page headlines. You've given it a very small number of examples to learn from – only a single week's worth of newspapers.

You've created a test for your model in Scratch. This test is like an exam for the model, testing how well it does on headlines it hasn't seen before.

It marks how well the machine learning model does. It draws graphs to help you understand what it is good at, and what it's not good at.

The next step is to try and improve the test result by giving our machine learning model some more training. With more examples of headlines, it will hopefully do even better.

52. Save your Scratch project

Click on File -> Save Project

53. Close the Scratch web browser window

It's important you do this in between giving your machine learning model additional training.

54. Click on the “< Back to project” link.

55. Click the Learn button

56. Repeat steps 14-16. Collect additional headlines from another couple of weeks of newspapers.

Make sure you don't use any headlines that are in your Scratch test.

If you show it an example that you'll use to test the model, that won't be a fair test.

You should now have fifteen example headlines from each newspaper.

The screenshot shows the 'ml-for-kids' platform interface. At the top, there is a navigation bar with links: Welcome, About, Projects, Worksheets, News, Help, and Log Out. Below the navigation bar, a title reads 'Recognising **text** as **Daily_Express, Daily_Mirror or 2 other classes**'. There are four main sections, each representing a news source:

- Daily_Express**: Headlines include: '2.37 MILLION EU MIGRANTS WORK IN BRITAIN', '91F BLAST OF HEAT ON WAY', 'BREAKTHROUGH ON ALZHEIMER'S', 'BRITAIN'S WIDE OPEN TO ILLEGAL MIGRANTS', 'DRINK COFFEE TO LIVE LONGER', 'EU CAN WHISTLE FOR OUR MONEY', 'EU IN CRISIS OVER BOAT MIGRANTS', 'EU TOLD: YOU CAN'T FRIGHTEN BRITAIN', 'FURY AT PLOT TO GET NEW EU VOTE', 'HOUSE PRICES UP BY £10,000', 'NEW OPERATIONS CRISIS HITS NHS', 'NEW WAY TO HALT DIABETES', 'SLAUGHTERED ON THE STREET', 'STATINS: NEW HEALTH ALERT', 'WINE IS KEY TO A LONGER LIFE'. A '+ Add example' button is at the bottom.
- Daily_Mirror**: Headlines include: "'FRAUDSTER WHO PUT 30 TONS OF HORSE ME...', 'ANT: MY FIGHT TO SAVE MARRIAGE', 'BLUNDER DOCS STERILISED ME AS I GAVE BIRTH', 'BOMB HERO 'STOLE FROM VICTIMS'', 'BRITISH NAVY... SWEDISH STEEL', 'CAM: IT'S SELFISH TO GIVE OUR HEROES PAY R...', 'DEMENTIA TIMEBOMB', 'FREEZE FARES', 'GIVE HEROES A DECENT RISE', 'MASSACRE', 'MPs TO VOTE ON MAX'S LAW', 'OUR CHARLIE IS STILL FIGHTING & SO ARE WE', 'SO VULGAR (EVEN FOR YOU, SIR GREEDY)', 'TORY MPS N-WORD SHAME', 'TRAGIC REFUGEES CRISIS'. A '+ Add example' button is at the bottom.
- The_Guardian**: Headlines include: 'Delay leaving single market - business chiefs', 'EU chiefs condemn 'fantasy' of UK plan to mirror t...', 'FA under pressure to explain payment after bullyin...', 'Labour threat to defeat May on Brexit bill', 'May appeals to Labour for policy ideas', 'May under fire as teacher pay rise held at 1%', 'PM's EU repeal bill dismissed as 'power grab'', 'Revealed: big tobacco's dirty battle for the African ...', 'Safety alert on 1,300 "failing" nursing homes', 'Scale of public sector salary cuts revealed', 'Student debt to hit £57,000 for the poorest', 'Terror strikes Barcelona', 'Trump under fire from all sides as woman dies at fa...', 'UK ditches its "cake and eat it" Brexit stance', "'Racism is evil", says Trump, 48 hours after fatal vio...'. A '+ Add example' button is at the bottom.
- The_Independent**: Headlines include: 'Brexit IT "horror" puts £34bn border tax at risk', 'Brexit minister caught in "tear down EU" rant', 'Care home crisis: places must double to meet soar...', 'Corbyn's summer tour', 'End of austerity? May softens on pay freeze', 'Government crackdown on universities handing ou...', 'Government demands tariff-free trade to avert bor...', 'Grenfell residents face deportation in a year', 'May abandons climate change priority at G20', 'May's plan to pack Brexit committee with Tories', 'Pay cap "dead in water" as firefighters land new deal', 'Teachers furious as pay is capped at 1 per cent', 'Terror in Barcelona: 12 dead in van attack', 'The devastation: Mosul finally liberated from Isis', "'Ripe for abuse": Brexit bill lets ministers scrap righ...'. A '+ Add new label' button is at the top right, and a '+ Add example' button is at the bottom.

57. Click the “< Back to project” link. Click the **Learn & Test** button.

58. Click the **Train new machine learning model** button

Wait for the training to complete. This might take a few minutes.

The screenshot shows a web interface for training a machine learning model. At the top, there is a navigation bar with links: ml-for-kids, Welcome, About, Projects, Worksheets, News, Help, and Log Out. Below the navigation bar, the title "Machine learning models" is displayed. On the left, a section titled "What have you done?" contains a message about collecting examples of text for four categories: Daily_Express, Daily_Mirror, The_Guardian, and The_Independent. It lists 15 examples for each category. On the right, a section titled "What's next?" asks if the user is ready to start training. It includes a button to start training and a note that the user can go back to the Train page to collect more examples if desired.

59. Click the “< Back to project” link. Click the **Scratch** button.

60. Click the **Open in Scratch** button.

61. Load your test Scratch project

Click on File -> Load Project

62. Test your new model by clicking the Green Flag.

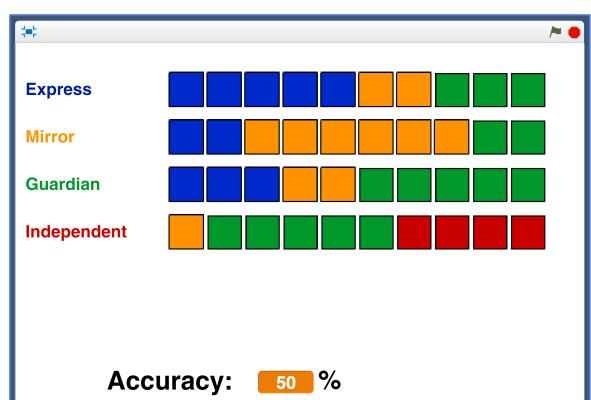
Compare the results with the results you got before.

My results after fifteen examples were a little better overall.

It was a little better at recognising

Express and Guardian headlines, and much better at recognising Daily Mirror headlines.

But it got a lot worse at recognising Independent headlines – getting wrong several that it was recognising correctly before!

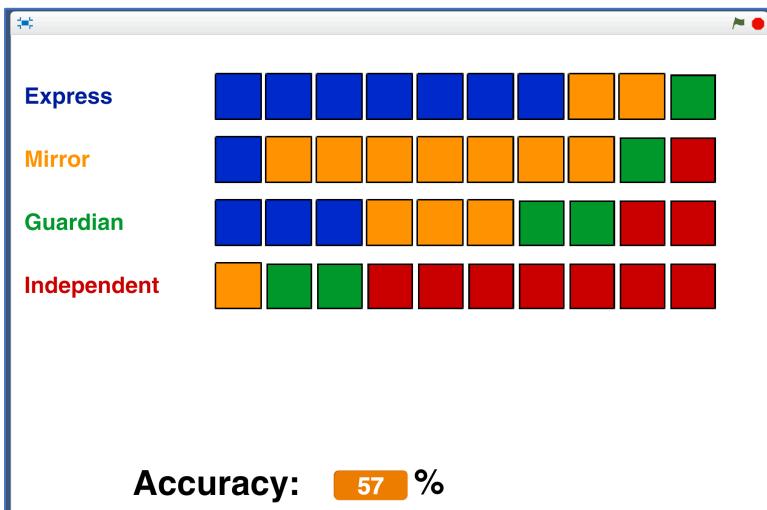


63. Repeat steps 53 – 62 again.

*Close Scratch. Go back to the Training tool and collect more examples.
Train a new machine learning model.*

Go back into Scratch and run your test to see how your new model is doing.

The screenshot shows the 'Machine learning models' page. At the top, there's a navigation bar with links: ml-for-kids, Welcome, About, Projects, Worksheets, News, Help, and Log Out. Below the navigation is a title 'Machine learning models'. Underneath the title, there's a link '< Back to project'. The page is divided into two main sections: 'What have you done?' and 'What's next?'. The 'What have you done?' section contains text about training a model to recognize news sources and a list of collected examples. The 'What's next?' section contains instructions for testing the model and links to Scratch and the Train page.



A little better overall.

It got worse at recognising Guardian headlines, though.

What have you done?

You've trained a machine learning model to recognise the use of language in the front page headlines of four national newspapers.

You've tested the model to assess the performance of the machine learning model.

You repeated this test at various stages of your model's training to see how the performance improved as you continued to train it.

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?

Look for other patterns in newspapers

Instead of comparing tabloid papers with broadsheet newspapers, try some other project ideas.

What about comparing local newspapers with national newspapers? Or weekend editions with week day papers?

Try different numbers of newspapers

What happens if you add additional newspapers? Is it still as good? What happens if you remove some, and only train it to tell headlines from two newspapers? Does it do better then?

Use a bigger test

Try running the test with 20 headlines from each newspaper instead of 10. What difference does that make?

Use the confidence block in Scratch

Think about how you could include the confidence that your machine learning model has in its predictions. Could your test treat predictions with high confidence differently from predictions that have a low confidence?