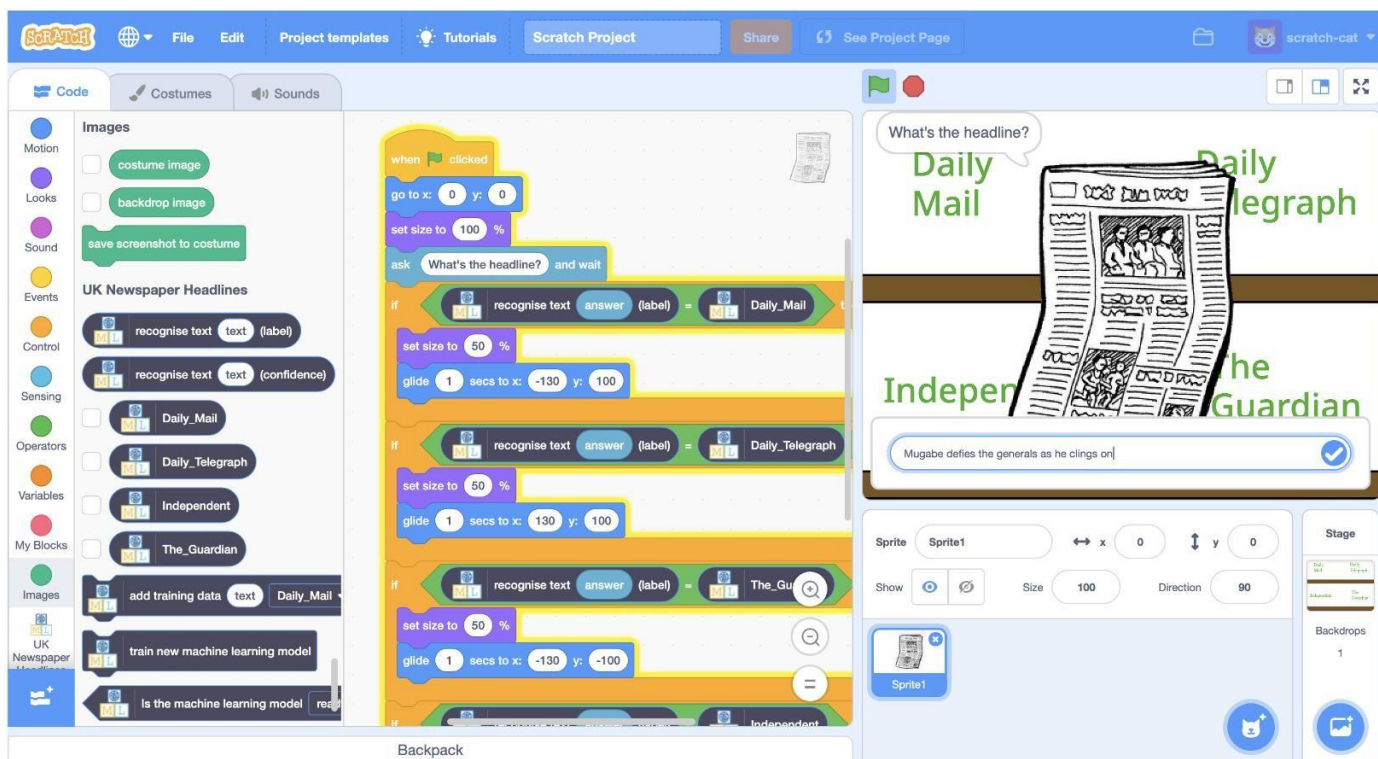




Estantes de periódicos

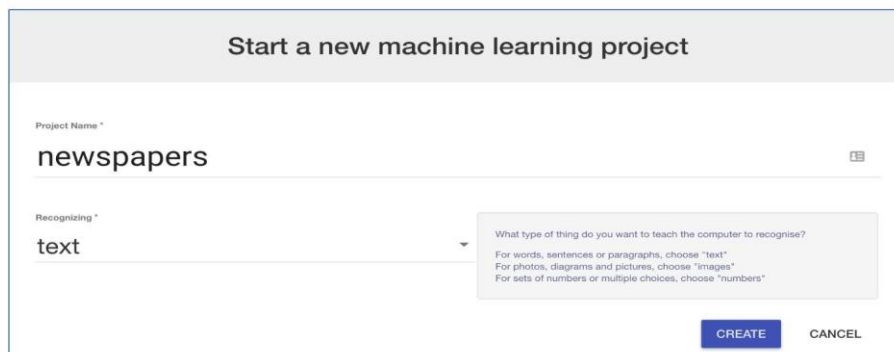
En este proyecto, entrenarás a un ordenador para que reconozca titulares de diferentes periódicos.

Recogerá ejemplos de los titulares de la primera página de diferentes periódicos. Los usarás para entrenar un modelo de aprendizaje automático que pueda aprender del uso del lenguaje para poder reconocer de qué periódico es un titular.



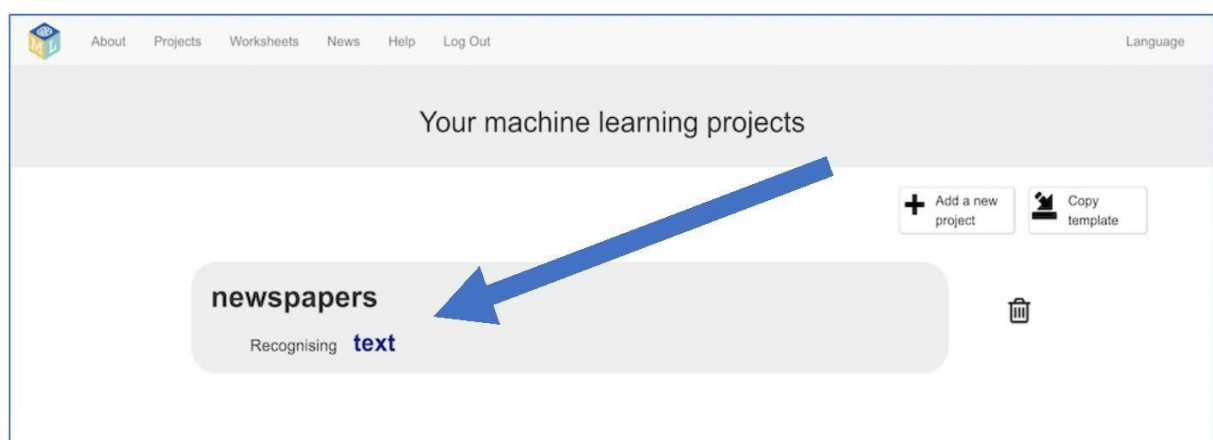
Esta hoja de trabajo de proyecto está bajo una licencia de Creative Commons Reconocimiento-Licencia de Compartir-Alike
<http://creativecommons.org/licenses/by-nc-sa/4.0/>

1. Ir a <https://machinelearningforkids.co.uk/> en un navegador web
2. Haz clic en "**Comenzar**".
3. Haz clic en "**Iniciar sesión**" y escribe tu usuario y contraseña. *Si no tienes usuario, pregunta a tu profesor o líder de grupo.*
4. Pulsa en "**Proyectos**" en la barra de menú superior.
5. Haz clic en el botón "**Añadir un nuevo proyecto**".
6. Pon nombre a tu proyecto "periódicos" y configúralo para reconocer "**texto**". Pulsa "**Crear**".

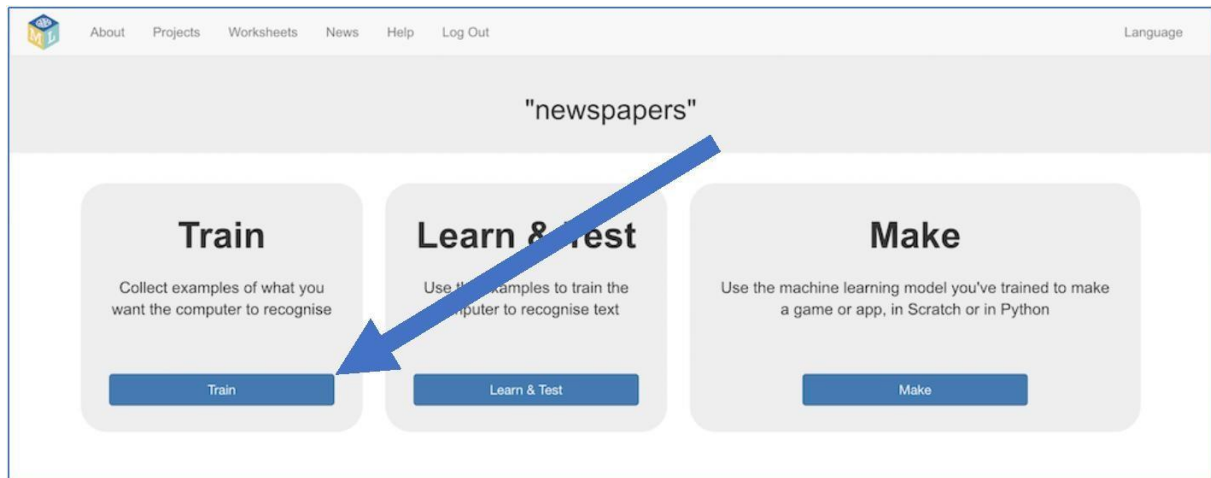


The screenshot shows a form titled "Start a new machine learning project". It has two main input fields: "Project Name *" with the value "newspapers" and "Recognizing *" with the value "text". A tooltip is visible next to the "Recognizing *" field, providing instructions: "What type of thing do you want to teach the computer to recognise? For words, sentences or paragraphs, choose 'text'. For photos, diagrams and pictures, choose 'images'. For sets of numbers or multiple choices, choose 'numbers'". At the bottom right, there are two buttons: "CREATE" and "CANCEL".

7. Ahora debería ver los "periódicos" en la lista de proyectos. Haz clic en él.



8. Haz clic en el botón "Entrenar".

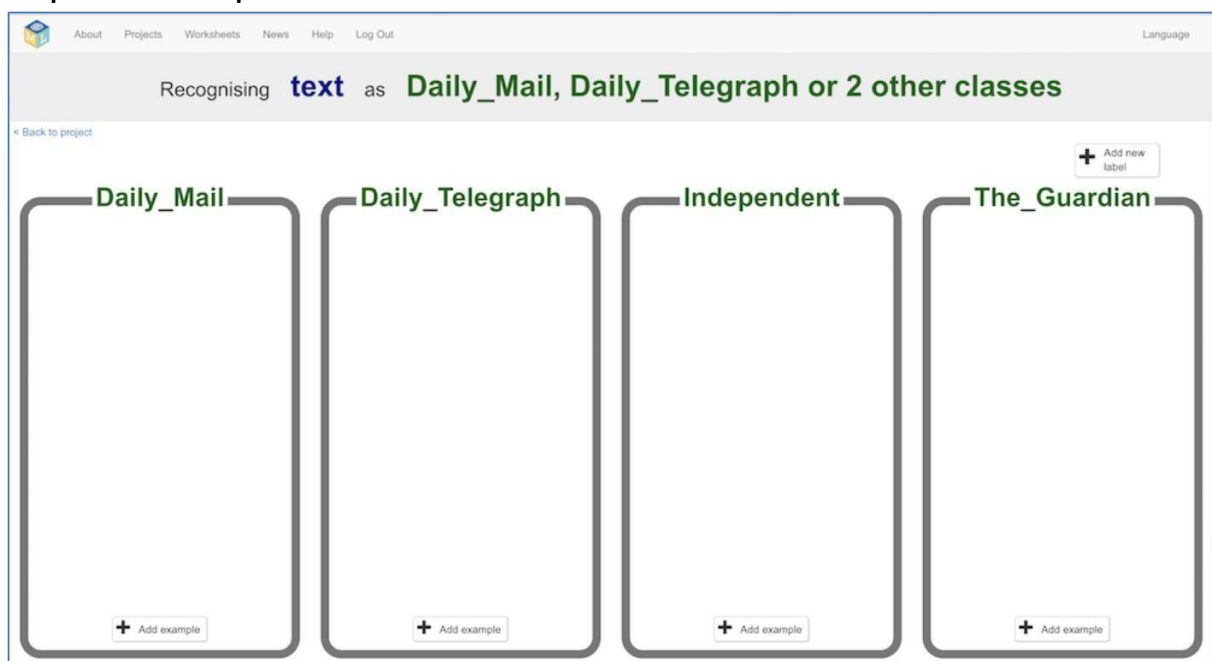


9. Elige cuatro periódicos nacionales.

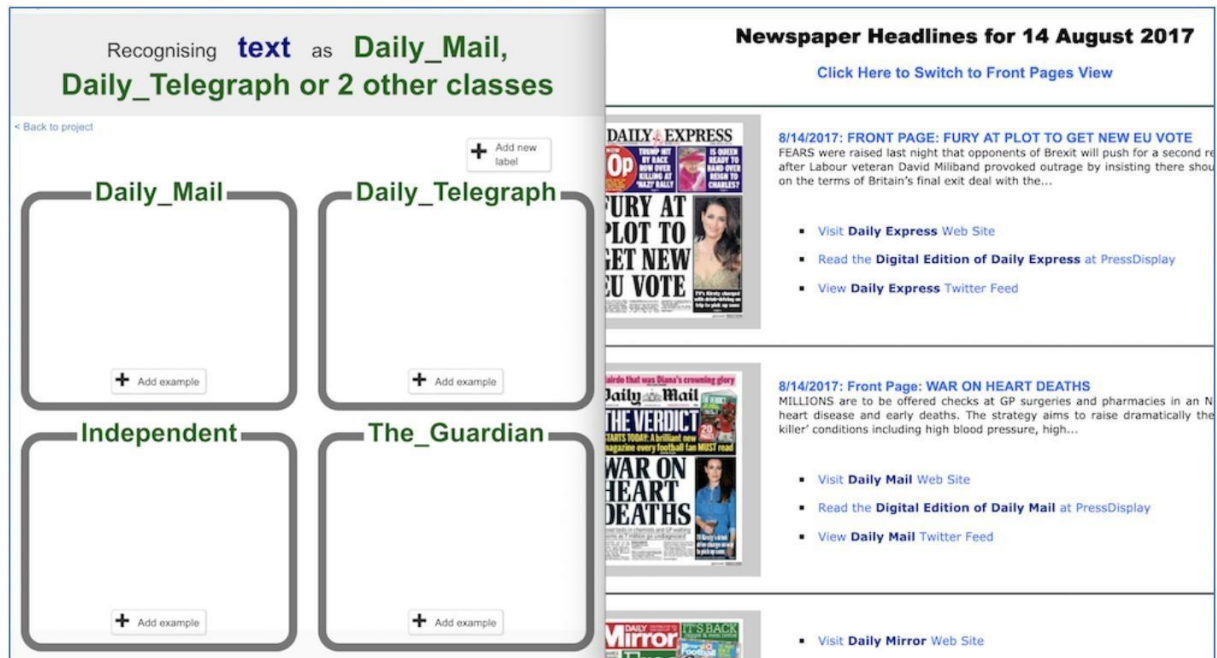
Para ahorrarte el tener que comprar periódicos físicos reales, puedes utilizar una colección online de portadas de periódicos.

Para el resto de esta hoja de trabajo, voy a utilizar: Daily Mail, Daily Telegraph, The Guardian, The Independent y he obtenido los titulares de:
<https://www.thepaperboy.com/uk/front-pages.cfm>

10. Utiliza el botón "+ Añadir nueva etiqueta " para crear un cubo para cada periódico que estás usando.



11. En otra ventana del navegador, busca portadas de periódicos. Organiza las ventanas para que los cubos de entrenamiento estén al lado del periódico.

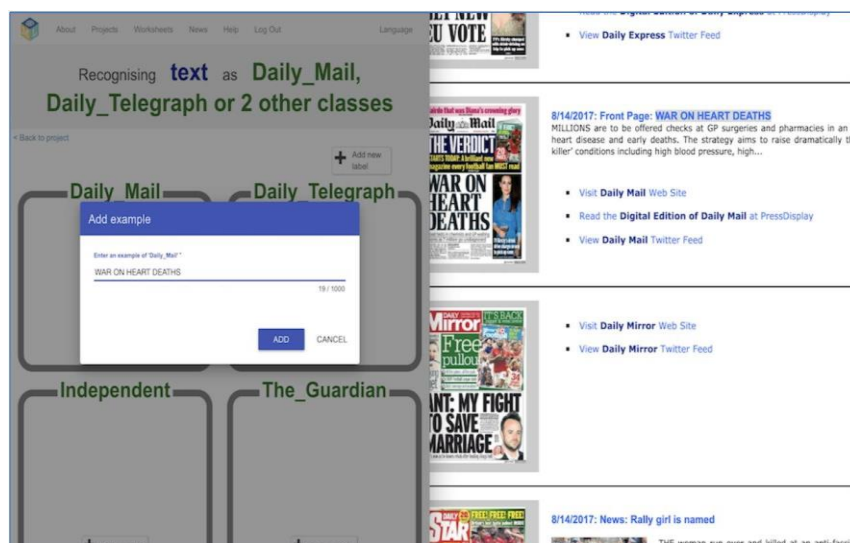


12. Acordar algunas reglas para recoger los ejemplos de los titulares, para que la formación sea coherente.

Para el resto de esta hoja de trabajo, estaré utilizando:

- * Periódicos de días de semana sólo (Lun-Vie) no ediciones de fin de semana
- * Sólo el titular principal más grande de cada periódico

13. Pulsa el botón "+ Añadir ejemplo" en uno de los cubos y copia el titular del periódico en el diálogo de "Añadir ejemplo".



14. Repite esto para los cuatro periódicos de este día.

Recognising **text** as **Daily_Mail**, **Daily_Telegraph** or 2 other classes

< Back to project

+ Add new label

Daily_Mail

WAR ON HEART DEATHS

+ Add example

Daily_Telegraph

Rees-Mogg: Cut stamp duty now

+ Add example

Independent

May's plan to pack Brexit commit...

+ Add example

The_Guardian

Trump under fire from all sides a...

+ Add example

8/14/2017: FRONT PAGE: Trump under fire from all sides as woma rally

Politicians from all sides have rounded on Dona explicitly to condemn white supremacy groups domestic terrorism after a woman was killed when antiracism protesters at the weekend, as the US...

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 Nazi 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 medal

15. Repite esto durante cinco días de titulares de periódicos. Si estás utilizando titulares de thepaperboy.com, puedes usar el calendario en la parte inferior de la página para pasar a otro día. Recuerda las reglas: sólo los días de lunes a viernes. El titular más grande que veas. Sugerencia: Utiliza la vista "Front Page Headline View" para obtener los encabezados como texto.

16. Ahora deberías tener ejemplos de una semana de prensa. Mantén la ventana del sitio del periódico abierta-lo necesitarás más tarde.

About Projects Worksheets News Help Log Out

Language

Recognising **text** as **Daily_Mail**, **Daily_Telegraph** or 2 other classes

< Back to project

+ Add new label

Daily_Mail

BREASTFEEDING LESSONS AT...

HARRY: NO ROYAL WANTS TH...

SHAME OF INSURANCE GIANTS

MASSACRE OF HOLIDAY FAMI...

WAR ON HEART DEATHS

+ Add example

Daily_Telegraph

BBC's gender pay gap revealed

Draw up your will in a text message

Set young free online 'to save the...

France's new hope puts cloud ov...

Rees-Mogg: Cut stamp duty now

+ Add example

Independent

May abandons climate change pri...

Government demands tariff-free t...

NHS faces £1bn Brexit bill for tre...

Male fertility crisis is 'wake-up call...

May's plan to pack Brexit commit...

+ Add example

The_Guardian

'Racism is evil', says Trump, 48 h...

May intensifies personal attack a...

Ban from 2040 on diesel and petr...

Backbenchers tell May: sack the ...

Trump under fire from all sides as...

+ Add example

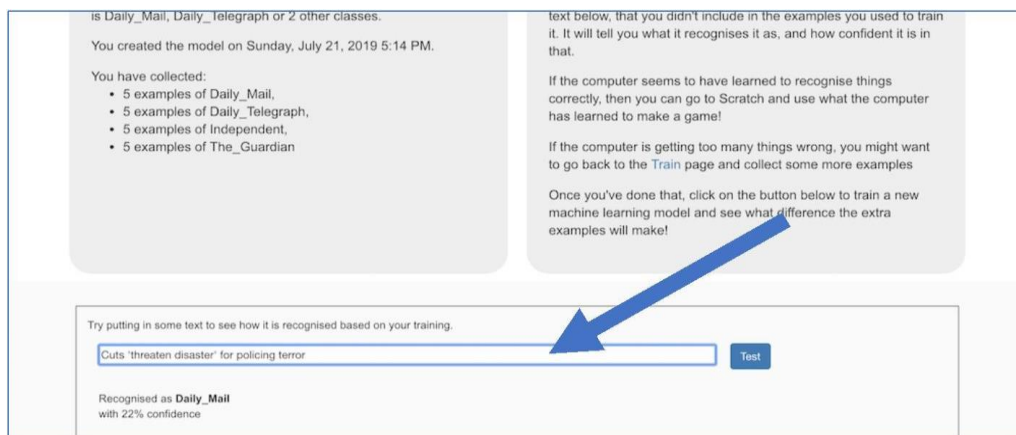
17. Pulsa el enlace "<Volver al proyecto". Luego haz clic en "Aprender & Probar".

18. Haz clic en "Entrenar un nuevo modelo".

Mientras hayas recogido suficientes ejemplos, el ordenador debería empezar a aprender cómo reconocer titulares a partir de los ejemplos que le has dado.

19. Espera a que se complete el entrenamiento. Esto puede tardar unos minutos.

20. Prueba el modelo poniendo un titular en la caja de pruebas.



The screenshot shows a web interface for training a machine learning model. It includes a list of collected examples (5 from Daily_Mail, 5 from Daily_Telegraph, 5 from Independent, and 5 from The_Guardian) and a 'Test' button. A blue arrow points to the 'Test' button. Below the button, the text 'Cuts "threaten disaster" for policing terror' is entered in the input field, and the output shows 'Recognised as Daily_Mail with 22% confidence'.

21. Pulsa el enlace "<Volver al proyecto".

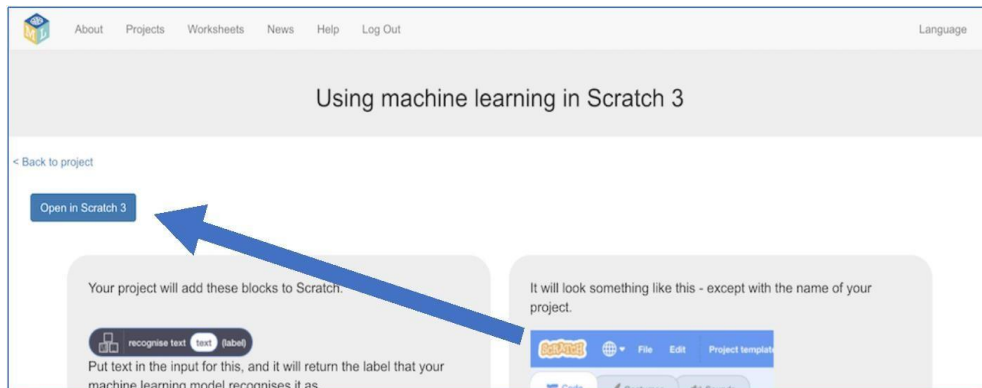
¿Qué has hecho hasta ahora?

Has empezado a entrenar un modelo de aprendizaje automático para reconocer los patrones en la forma en que los diferentes periódicos usan palabras, frases y puntuación en sus titulares de primera plana.

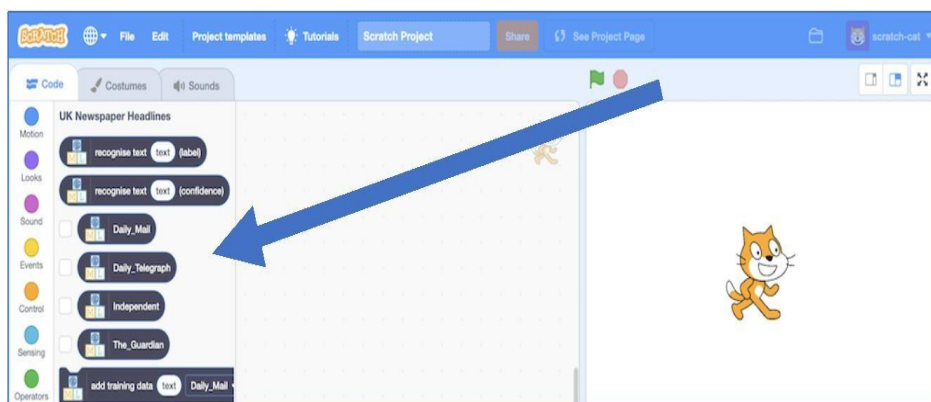
¿Pero está funcionando bien? ¿Cuánto ha aprendido hasta ahora? ¿Necesita más ejemplos para poder reconocer los titulares correctamente? ¿En qué es bueno? ¿En qué es malo? ¿Con qué periódicos se confunde?

22. Haz clic en el botón "Crea" y luego haz clic en el botón "Scratch 3".

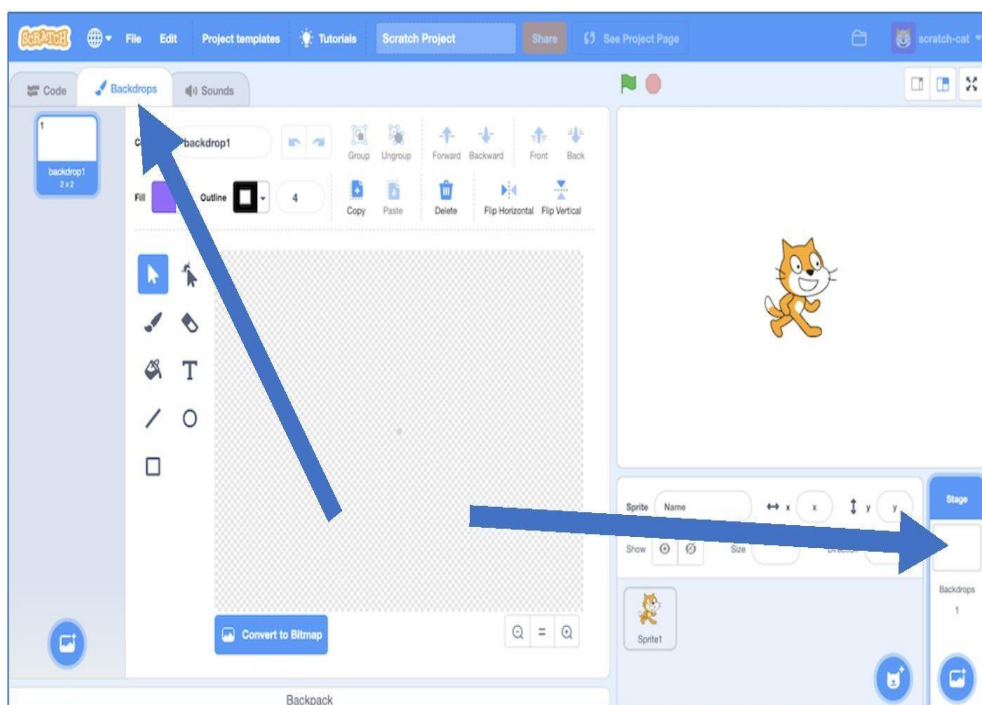
23. Haz clic en el botón "Open in Scratch 3".



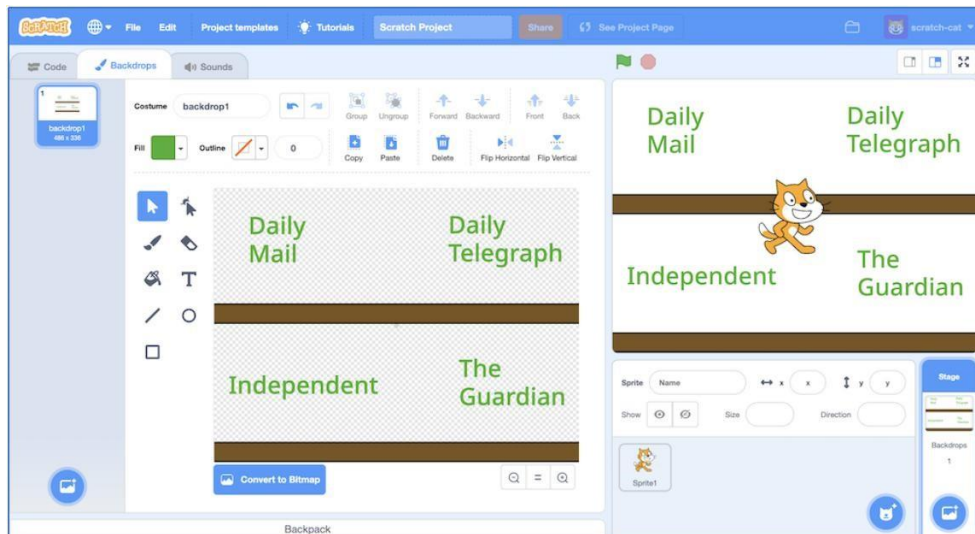
24. Deberías ver los bloques (en la sección inferior) del proyecto.



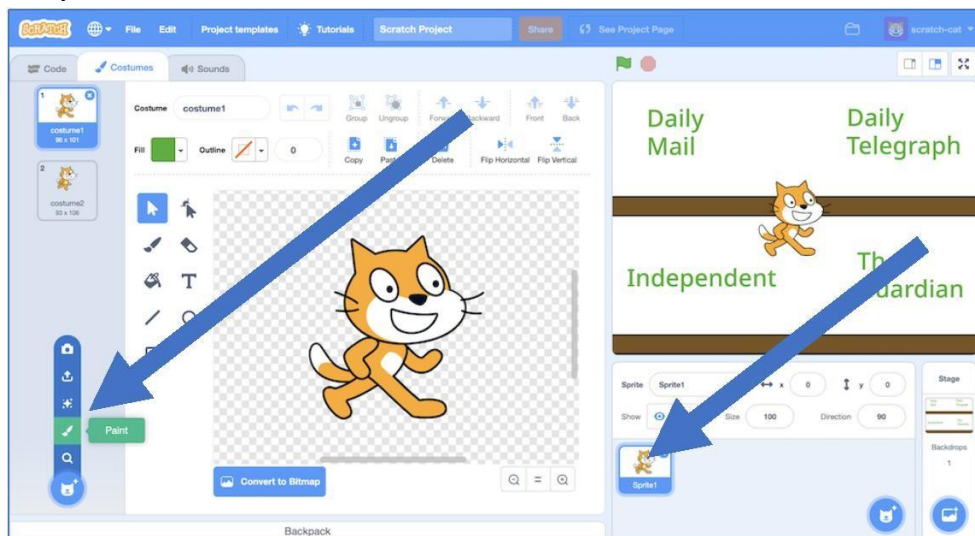
25. Haz clic en el Stage y, a continuación, haz clic en la ficha Backdrops.



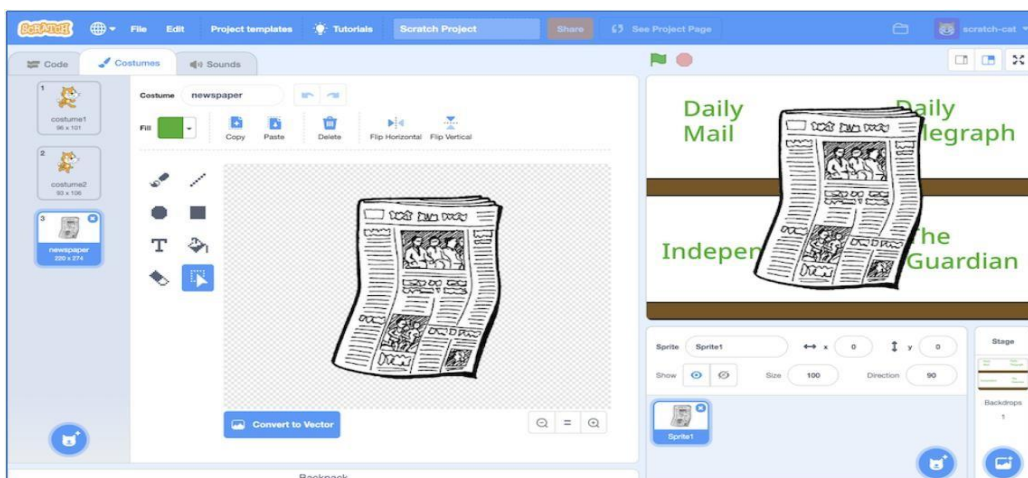
26. Dibuja algunos estantes para tus periódicos.



27. Haz clic en el sprite de gato y haz clic en Paint para dibujar un nuevo Sprite.

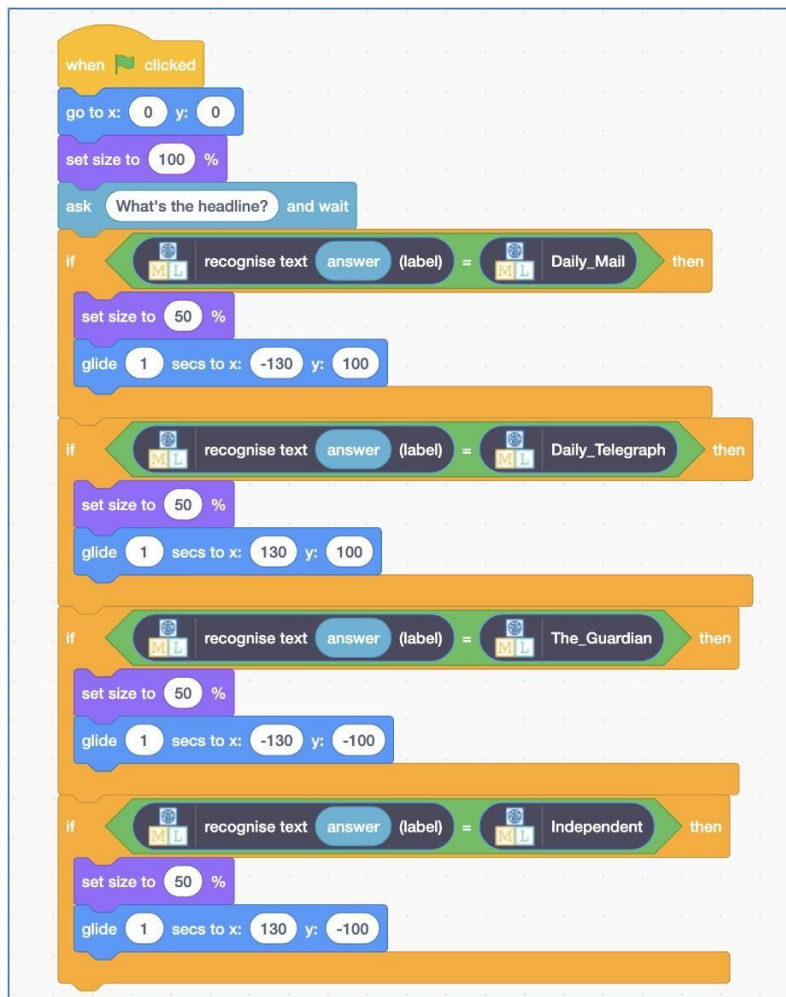


28. Dibuja un periódico sencillo.

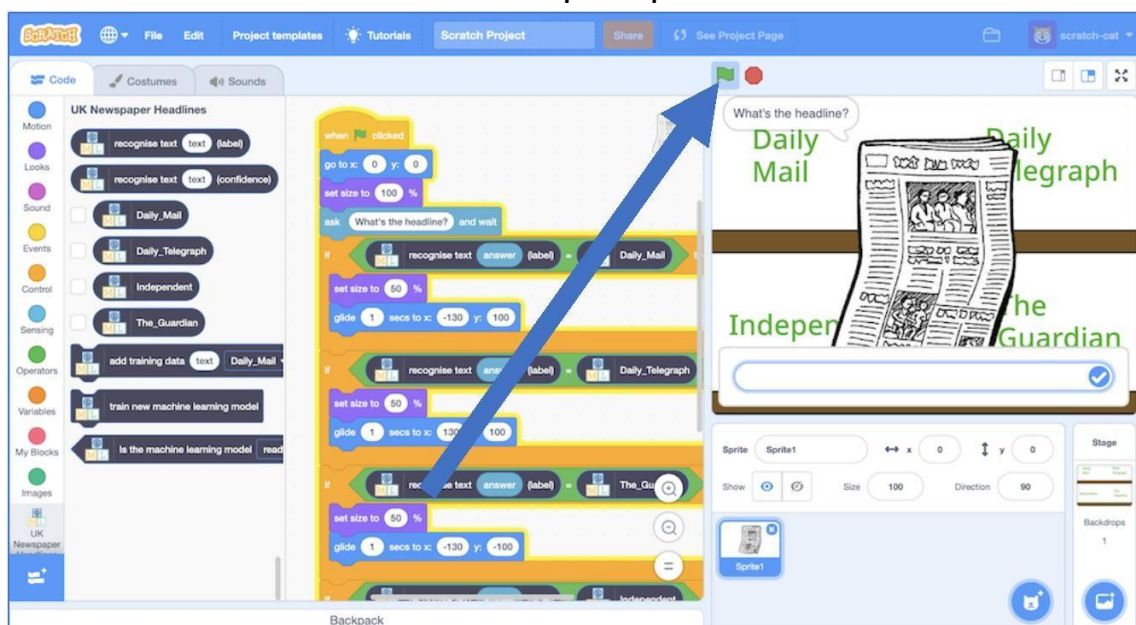


29. Crea el siguiente script.

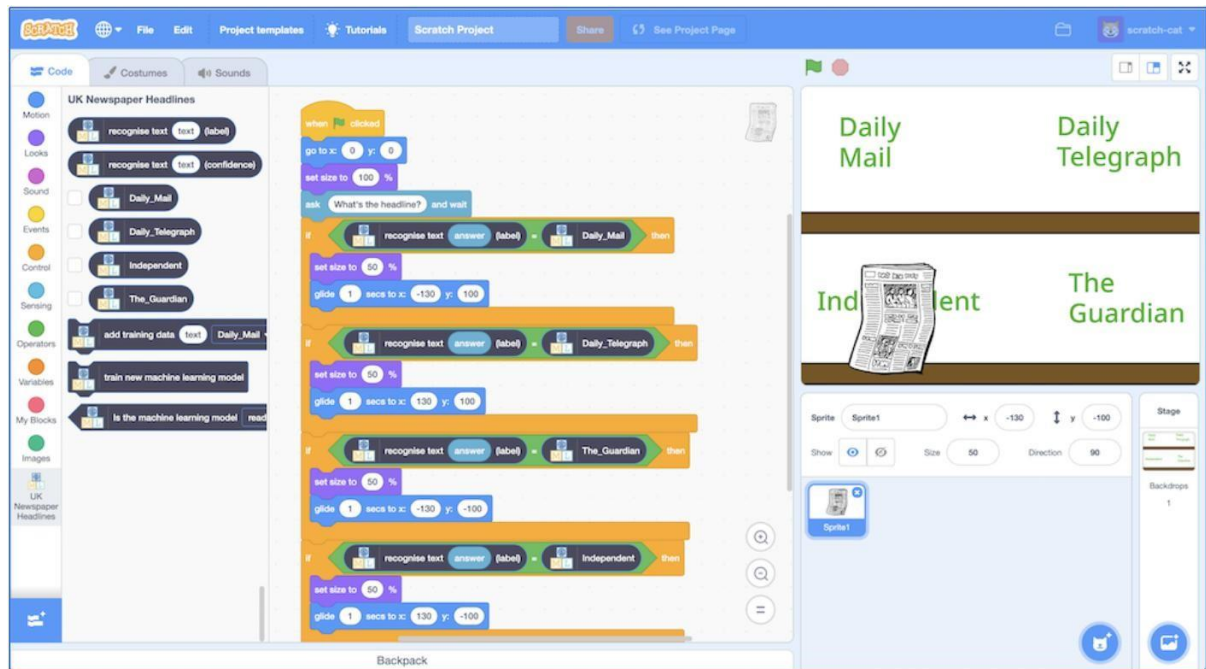
Cuando escribes un titular, este script predice de qué periódico se trata y lo pone en el estante correcto.



30. Haz clic en la Bandera Verde para probar.



- 31.** Intenta escribir diferentes titulares y ver cómo se comporta el ordenador.
Asegúrate de usar titulares que no usaste para entrenar el modelo.



- 32.** Guarda el proyecto Scratch
Pulsa Archivo-> Guardar en el sistema

¿Qué has hecho hasta ahora?

Has empezado a entrenar un modelo de aprendizaje automático para reconocer los titulares de primer plano. Le has dado un número muy pequeño de ejemplos para aprender de - sólo una semana de los periódicos.

Has creado un proyecto en Scratch para probar su modelo en titulares que no ha visto antes. Probablemente estarás viendo que se equivoca mucho.

El siguiente paso es tratar de mejorar los resultados dando a nuestro modelo de aprendizaje automático un poco más de entrenamiento. Con más ejemplos de titulares, esperemos que lo haga mejor.

33. Vuelve a la ventana de entrenamiento.

34. Pulsa el enlace "<Volver al proyecto".

35. Haz clic en el botón Train

36. Repite los pasos 13-15.

Recopila titulares adicionales de otro par de semanas de periódicos. Ahora debería tener quince titulares de ejemplo de cada periódico.

The screenshot shows a web application interface for training a text classification model. The top navigation bar includes links for 'About', 'Projects', 'Worksheets', 'News', 'Help', and 'Log Out'. The main header displays the current task: 'Recognising text as Daily_Mail, Daily_Telegraph or 2 other classes'. Below the header, there are four columns, each representing a different newspaper class: 'Daily_Mail', 'Daily_Telegraph', 'Independent', and 'The_Guardian'. Each column contains a list of example headlines. The 'Daily_Mail' column has 10 examples, while the others have 8. Each column has an 'Add example' button at the bottom. The interface also includes a '< Back to project' link on the left and an '+ Add new label' button on the right. The bottom of each column shows a count of examples: 71 for Daily_Mail and 69 for the others.

37. Pulsa el enlace "<Volver al proyecto". Puls el botón **Aprender y probar**.

38. Haz clic en el botón **Train new machine learning model**.
Espera a que se complete el entrenamiento. Esto puede tardar unos minutos.

39. Vuelve a la ventana Scratch.

Si la cerraste accidentalmente, puedes volver a ella haciendo esto:

- * Pulsa el enlace "< Volver al proyecto"*
- * Haz clic en el botón "Hacer"*
- * Haz clic en el botón "Scratch 3"*
- * Haz clic en el botón "Open in Scratch 3"*
- * Abrir el proyecto guardado antes, con File-> Load desde tu ordenador*

40. Repite los pasos 30-31 de nuevo.

Mira cómo está haciendo su nuevo modelo.

¿Tu modelo predice de qué periódico proviene un titular más a menudo que antes?

¿Qué has hecho?

Has entrenado un modelo de aprendizaje automático para reconocer el uso del lenguaje en los titulares de la primera página de cuatro periódicos nacionales.

Has probado el modelo para evaluar el rendimiento del modelo de aprendizaje automático.

Has comprobado esto en varias etapas de la formación de tu modelo para ver cómo ha mejorado el rendimiento a medida que continuabas entrenándolo.

Ideas y extensiones

Ahora que has terminado, ¿por qué no dar una prueba a una de estas ideas?

¿O se te ocurre uno a tí?

Buscar otros patrones en los periódicos

En lugar de comparar los periódicos sensacionalistas con los periódicos de gran formato, prueba otras ideas de proyecto.

¿Qué tal si comparamos periódicos locales con periódicos nacionales? ¿O ediciones del fin de semana con los diarios de la semana?

Prueba con diferentes números de periódicos

¿Qué pasa si añades más periódicos? ¿Sigue siendo tan bueno? ¿Qué pasa si se quitan algunos ejemplos, y sólo se entrenan para contar titulares de dos periódicos? ¿Lo hace mejor entonces?

Usar historias de periódicos en vez de titulares

¿Qué pasa con el texto de las historias de los periódicos en vez de los titulares? ¿Puedes entrenar un modelo de aprendizaje automático para reconocer el estilo de escritura en la historia?