

EdwAIR

iPort聲控燈實驗

學習目標

1. 讓學生能夠了解什麼是語音識別及其原理
2. 讓學生知道語音識別日常例子
3. 讓學生能夠製作自己的AI小燈
4. 讓學生知道語音識別的**限制及倫理問題**

1

什麼是語音識別？

甚麼是自動語音識別（ASR）？

- 使電腦自動識別人類的語音內容，將其翻譯成文本。



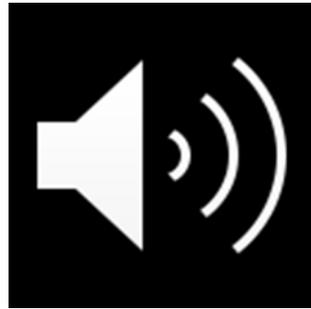
自動語音識別 (ASR)

- 涉及**模型訓練**和**測試**兩個過程
- 「輸入－處理－輸出」

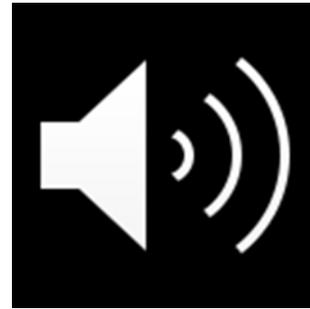


小活動

- 我們人類如何識別語音？ 聆聽以下錄音…



[連結1](#)



[連結2](#)

你分別聆聽到什麼？ 兩者有什麼分別？

人工智能模型如何識別語音？

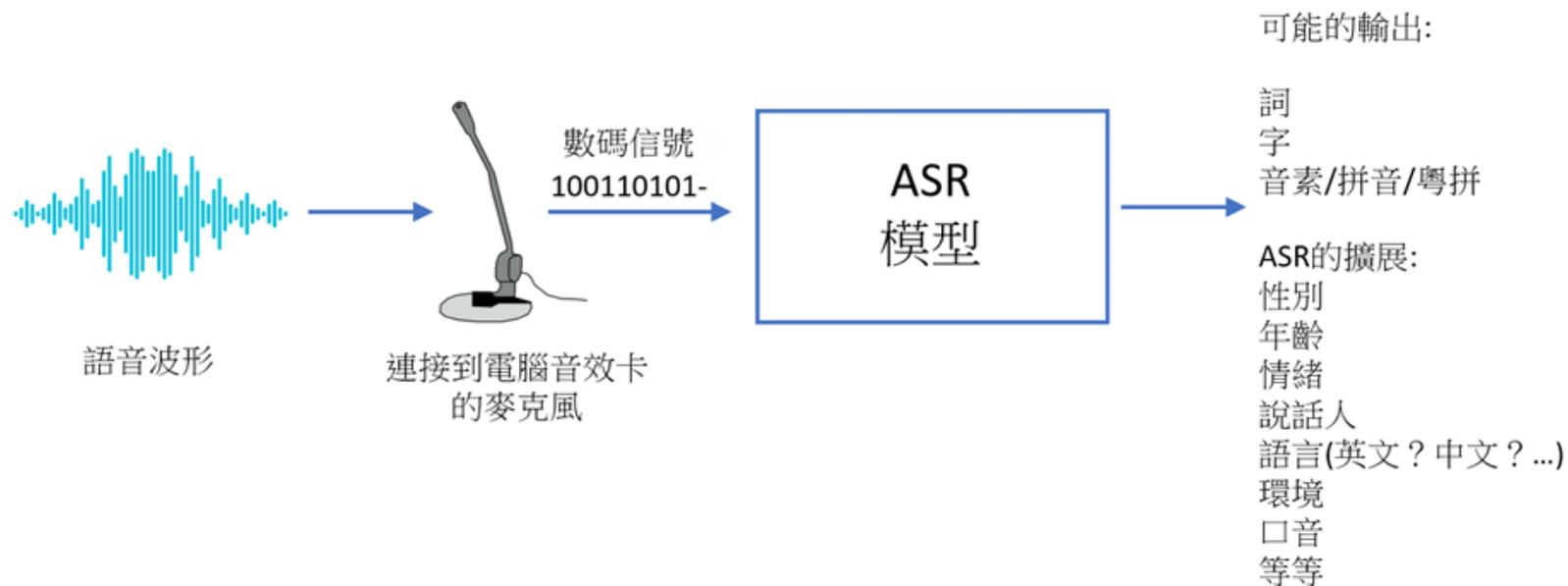
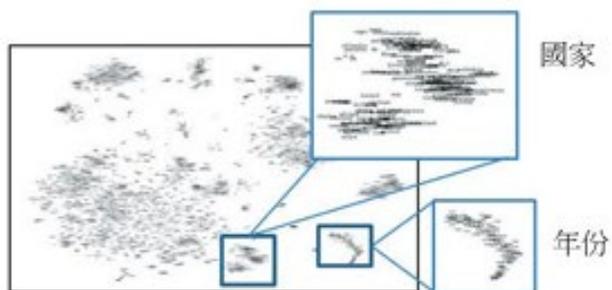


圖 4.1 自動語音識別管道

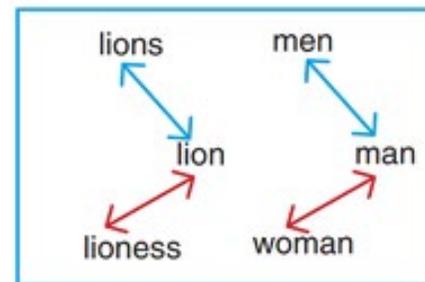
人工智能模型如何識別語音？

例子1 相似度 \Leftrightarrow 距離



語義的相似度於向量空間
中被編程為距離

例子2 詞語之間的關係 \Leftrightarrow 角度

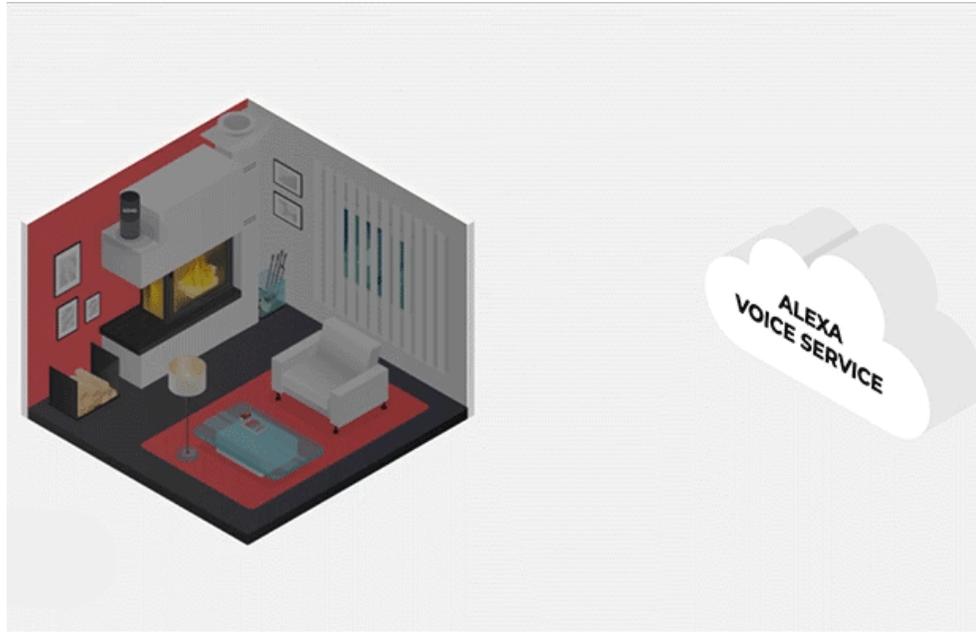


相同的關係角度也相同

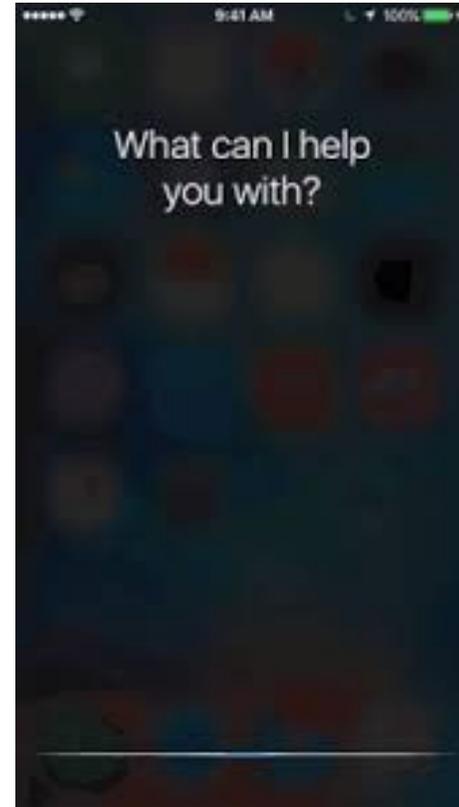
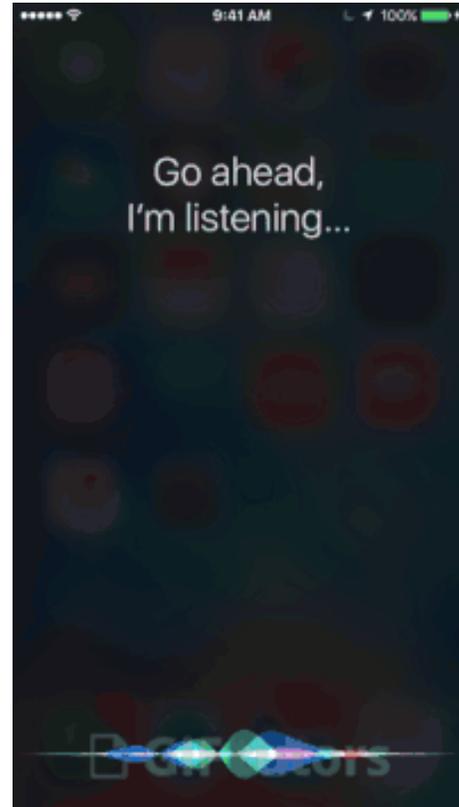
2

語音識別的應用

1. 智能家居語音控制設備

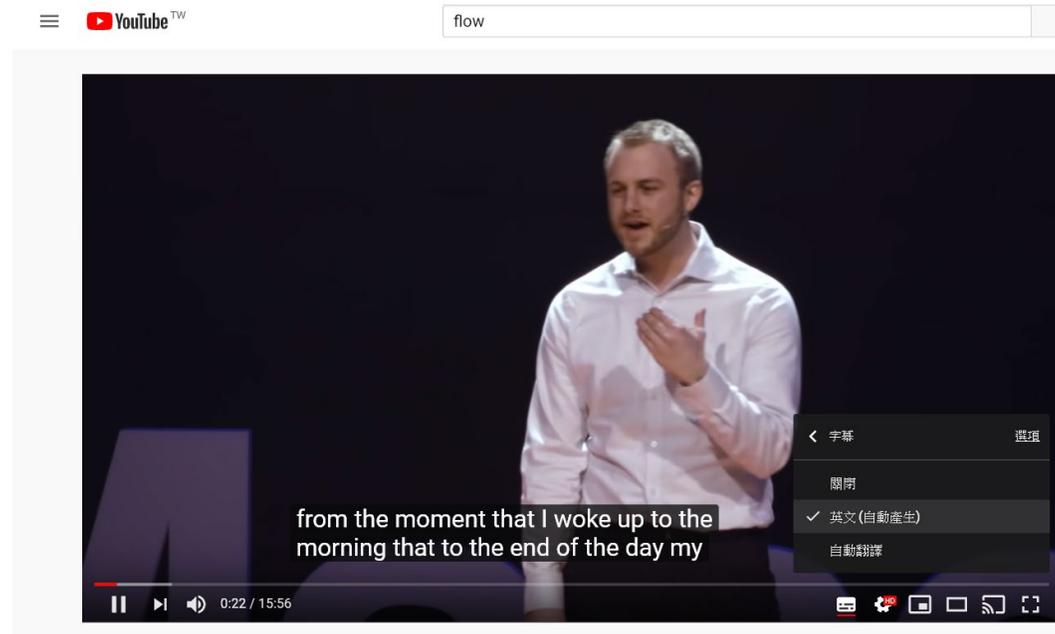


2. 手機應用



3. 語音轉錄

手機語音輸入法



在影片中自動生成字幕

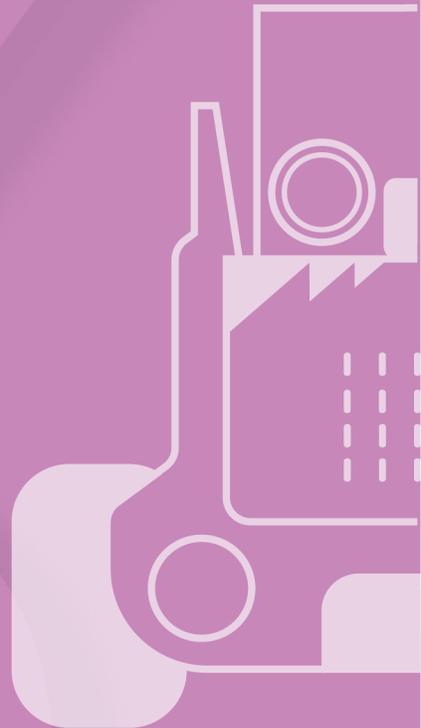
3

製作 AI 小燈

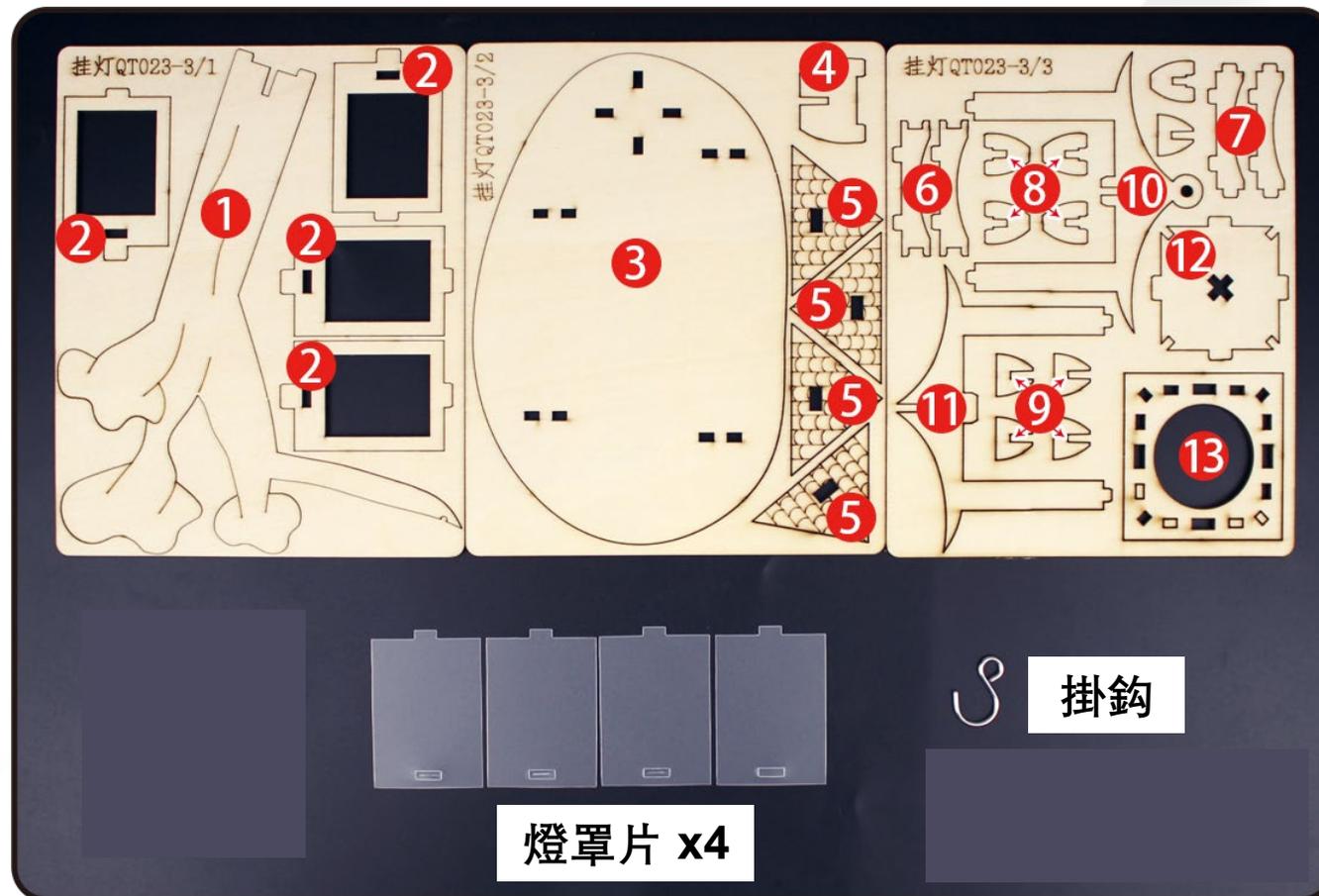
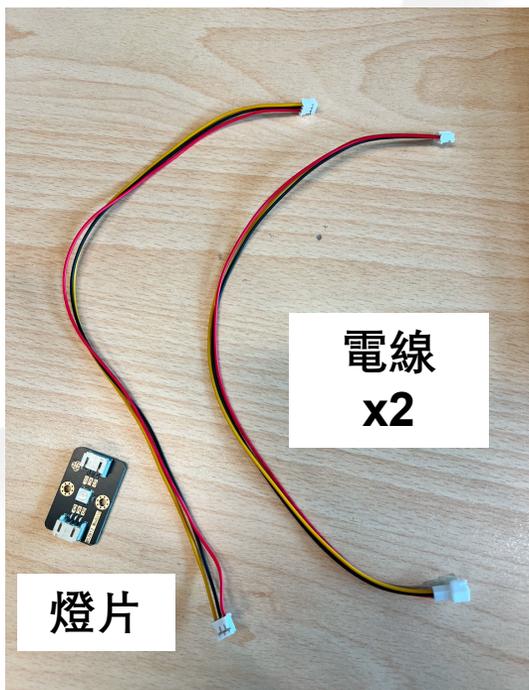
製作 AI 小燈



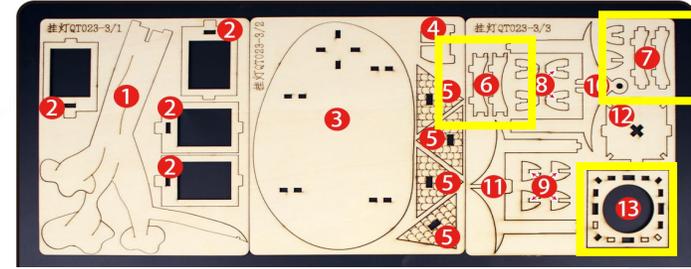
組裝小燈



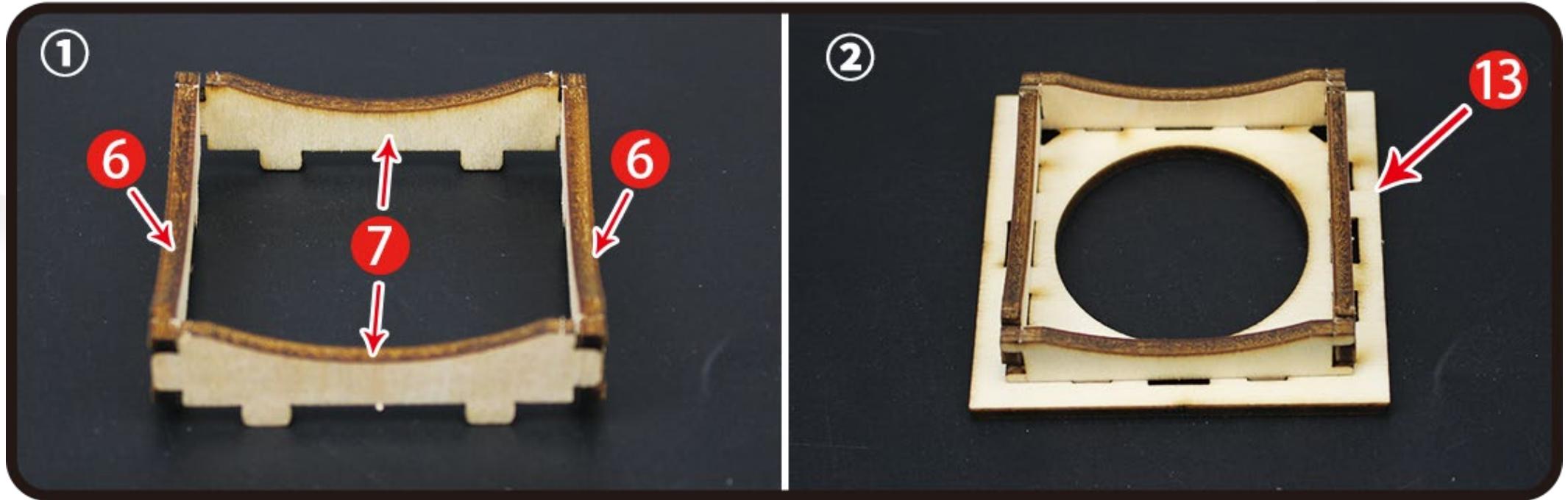
1. 檢查材料



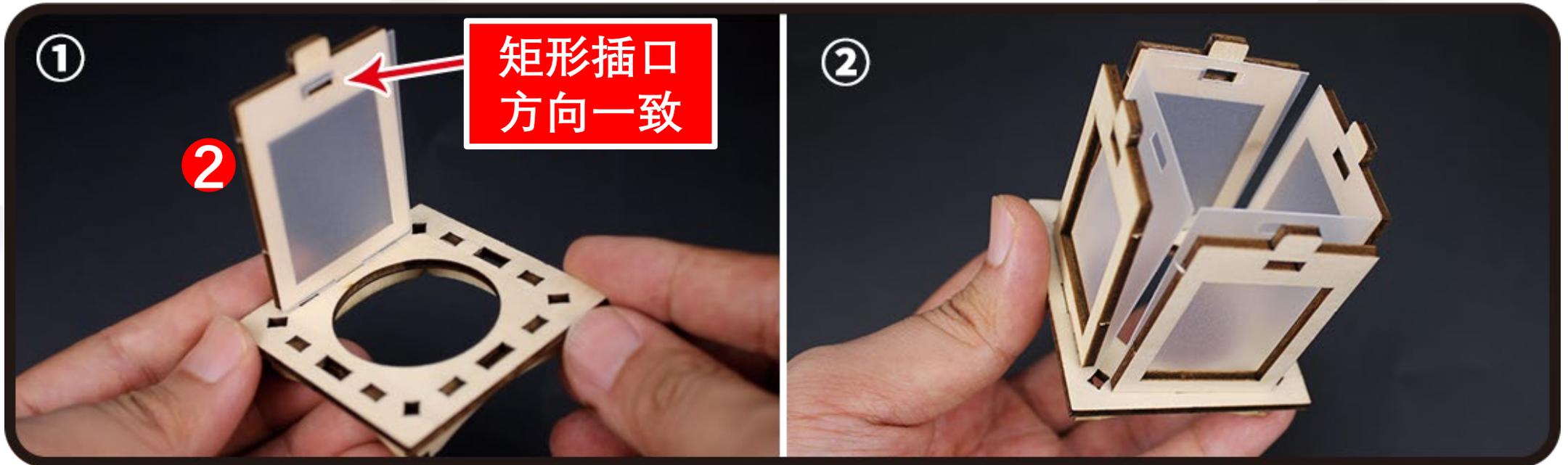
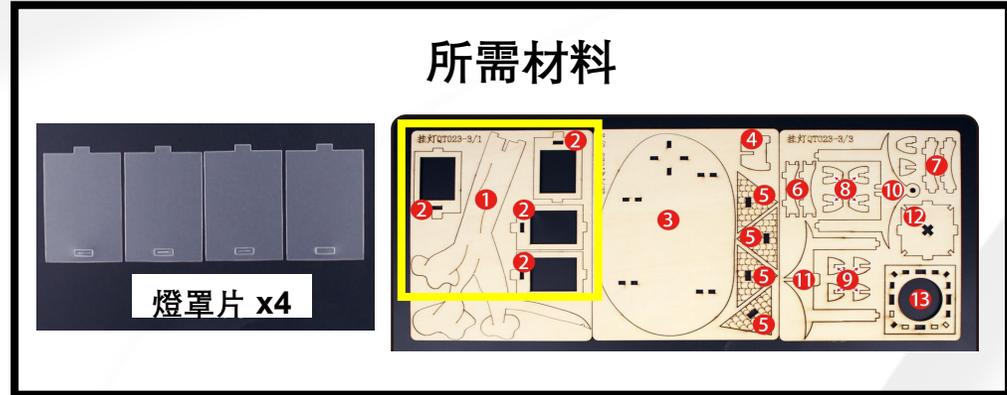
所需材料



2. 拼砌底座

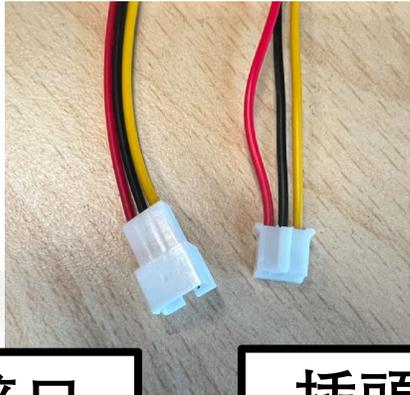


3. 安裝燈罩



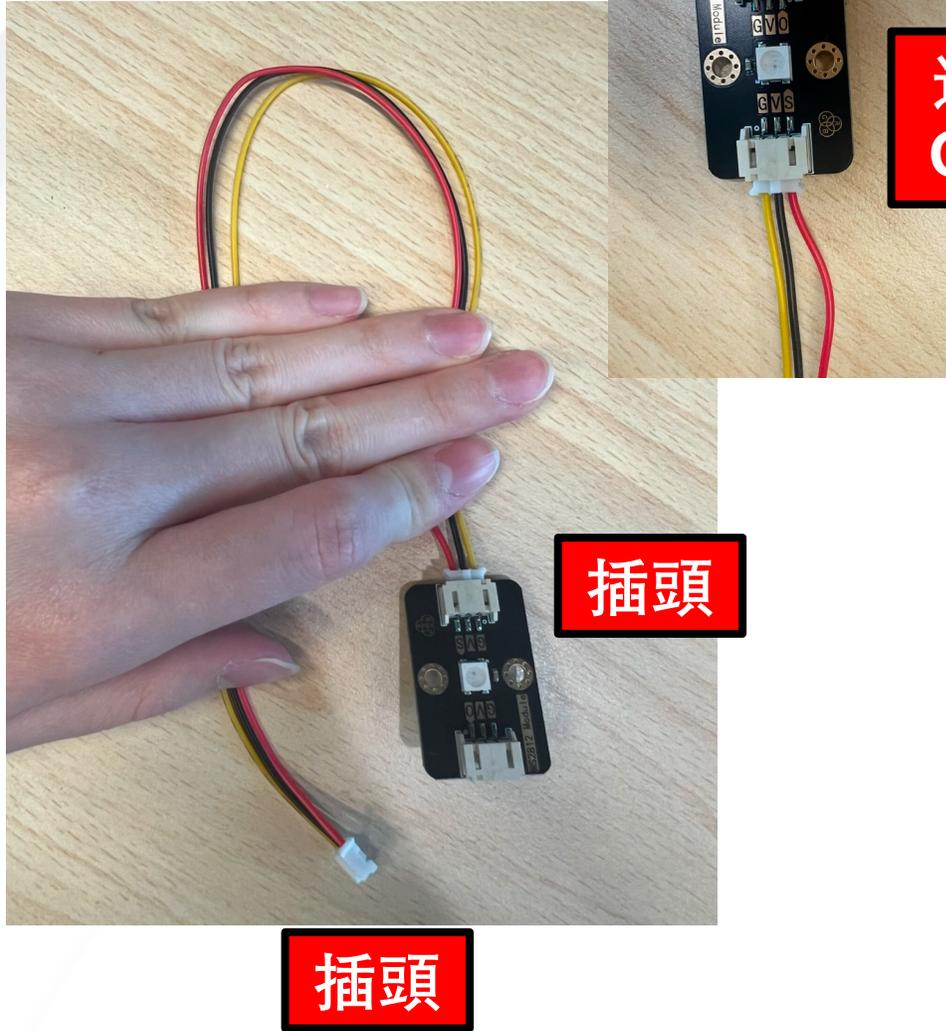
4. 連接燈片

電線有兩種尾端 ...



接口

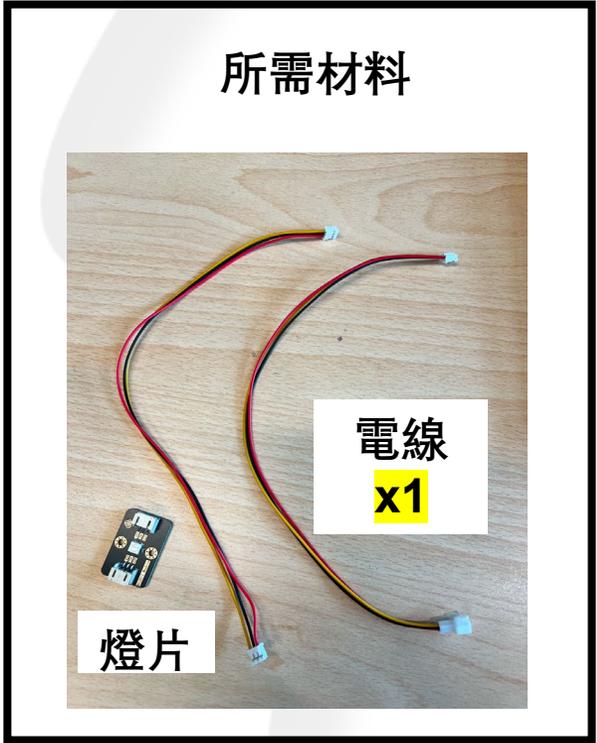
插頭



插頭

插頭

連接至 印着 GVS 的一端

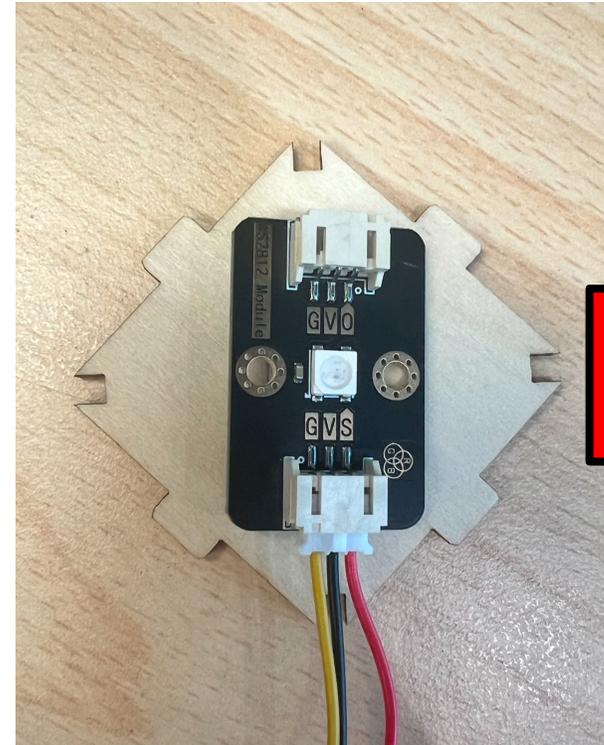
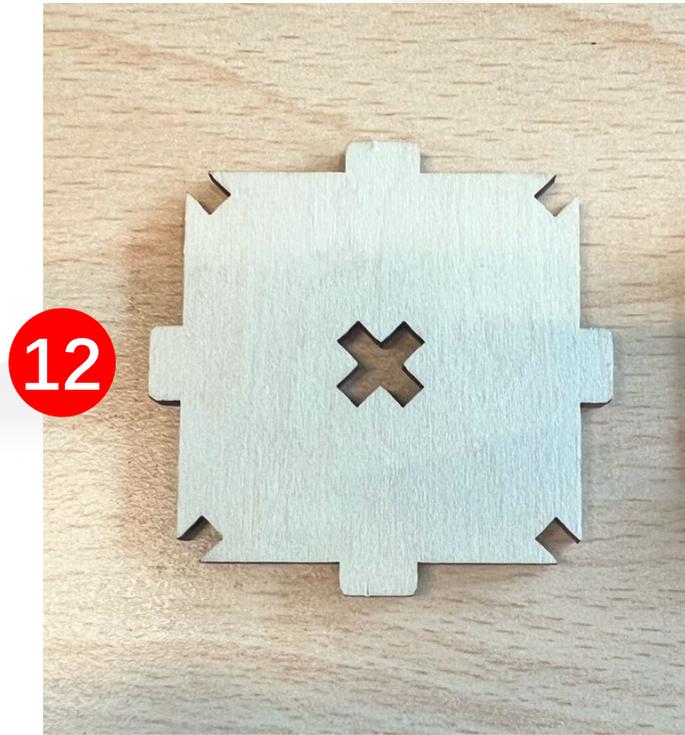
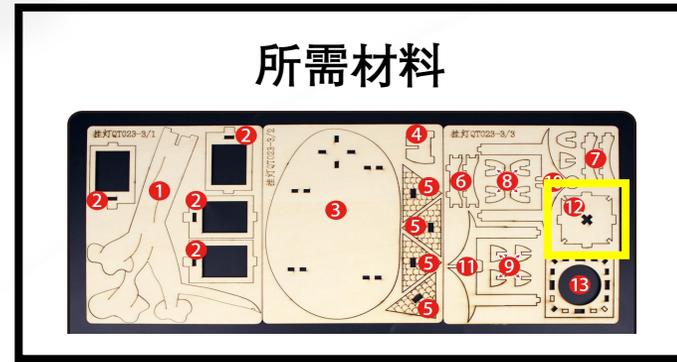


所需材料

電線 x1

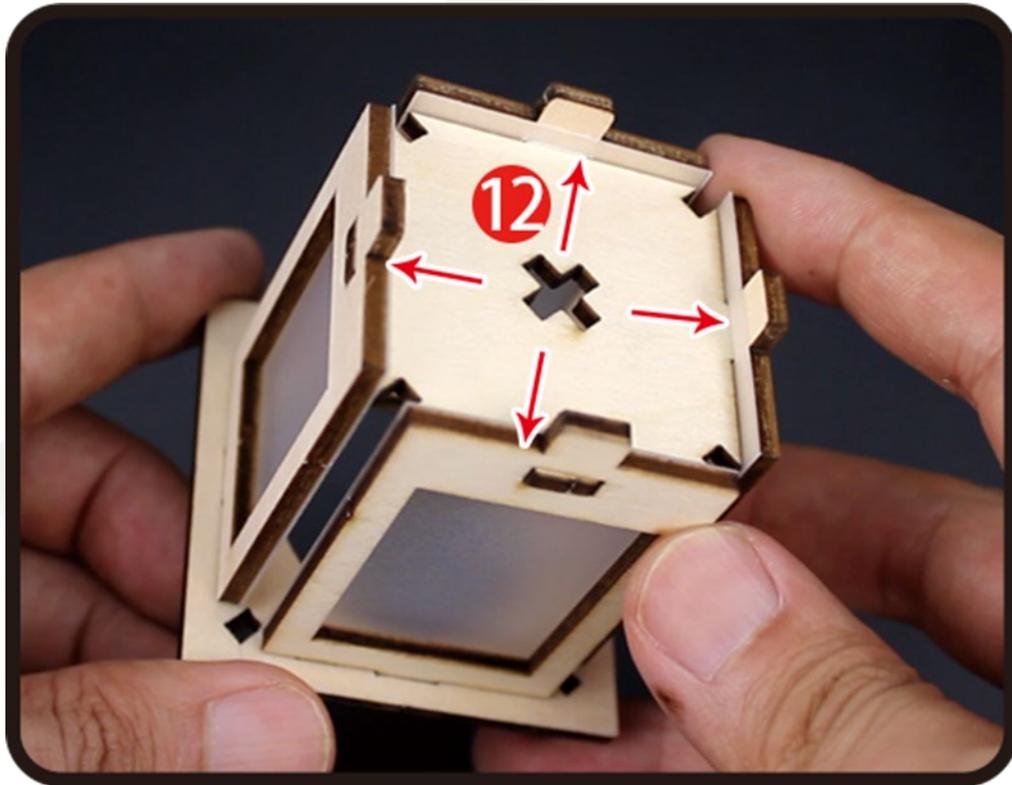
燈片

5. 拼貼燈片



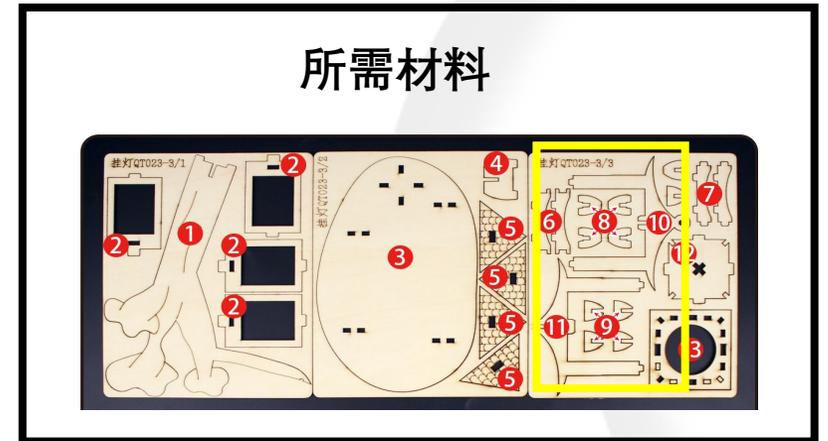
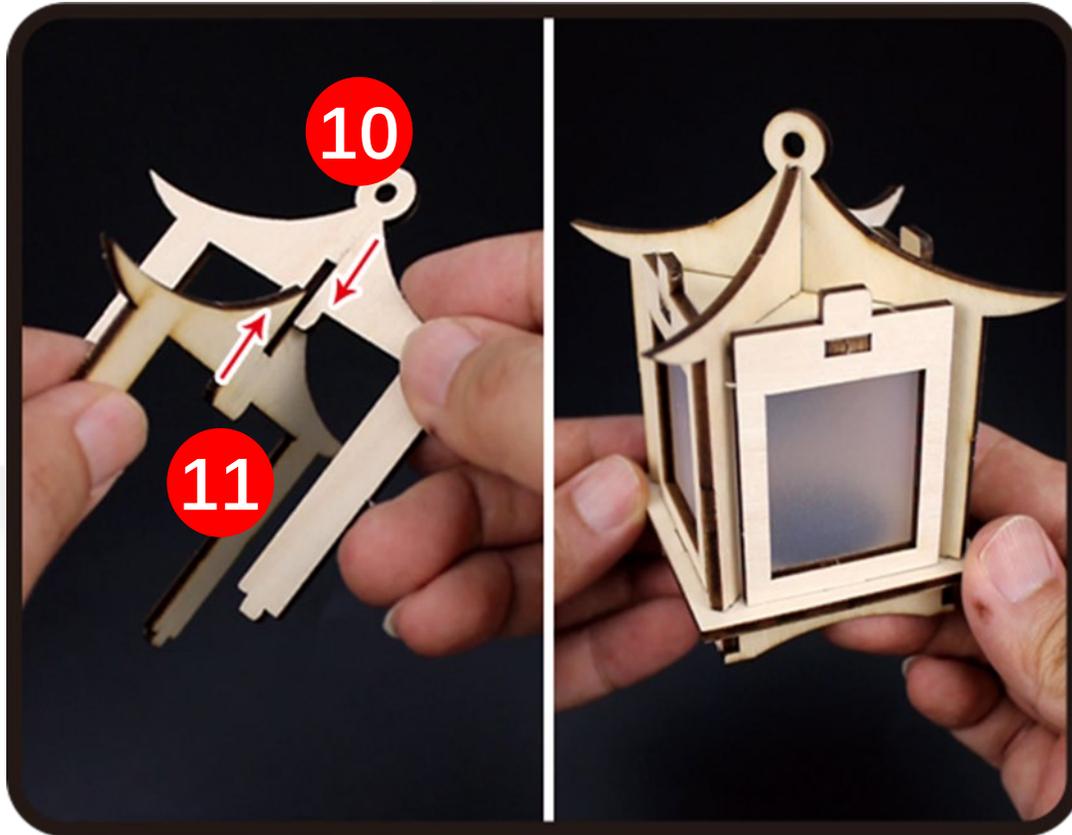
將木板旋轉
雙面膠粘貼

6. 安裝燈片木板

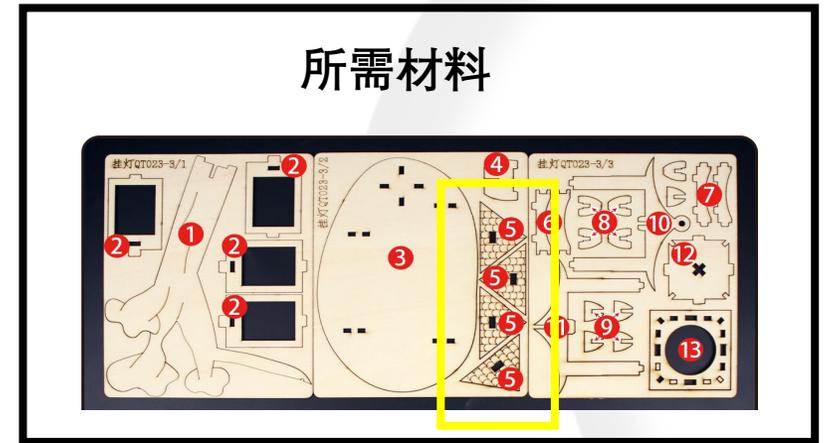


電線拉向
燈座方向

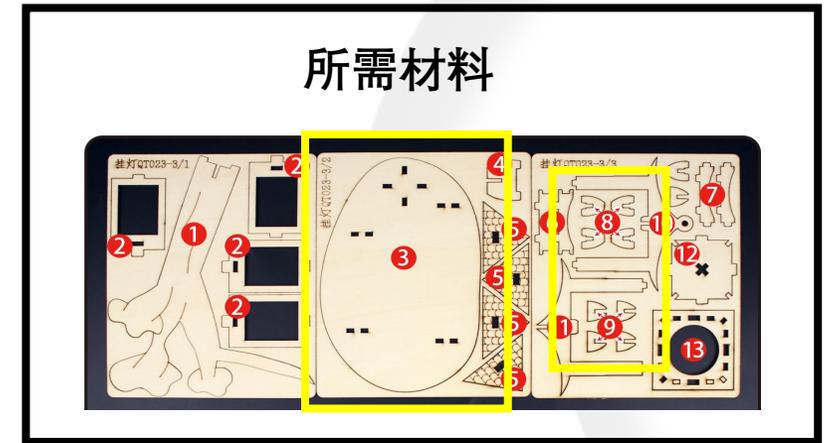
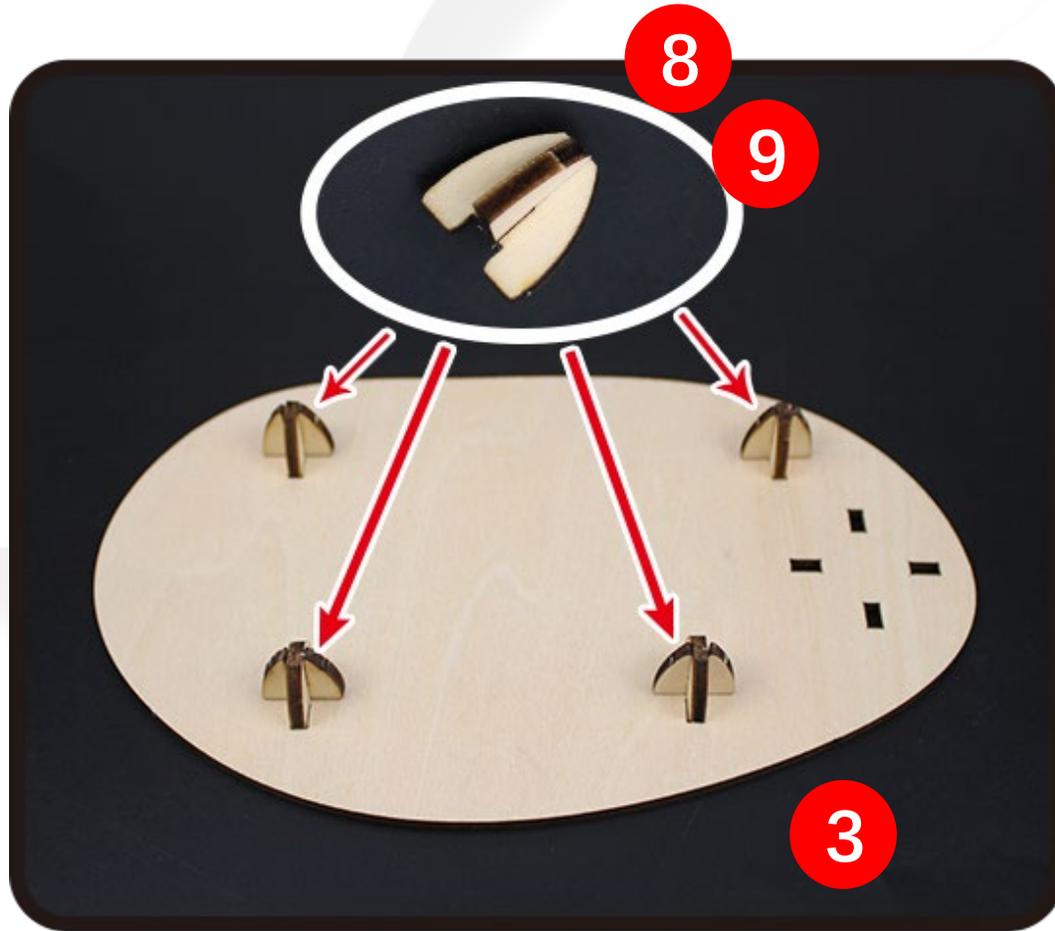
7. 拼砌頂樑



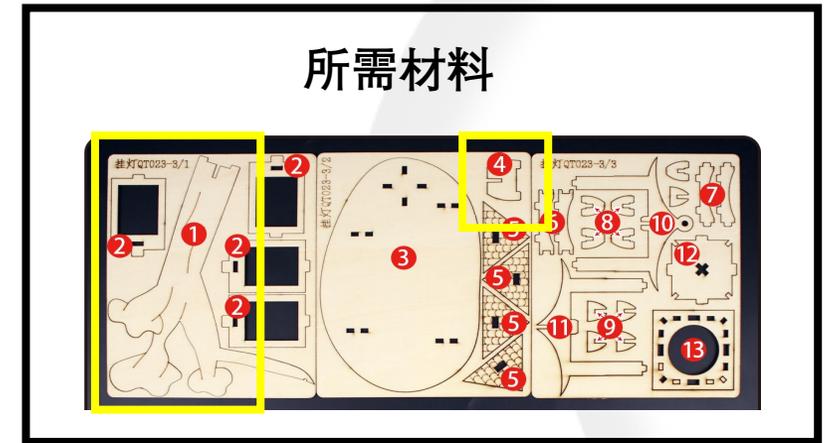
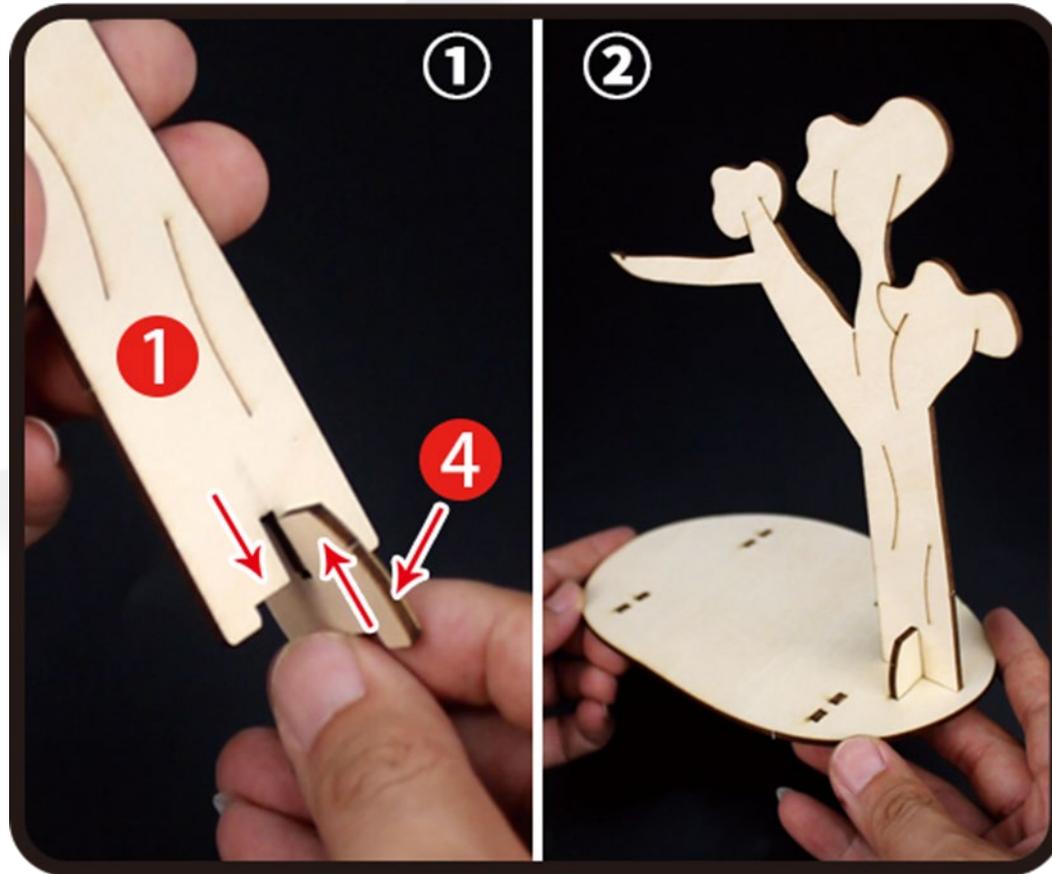
8. 安裝瓦片



9. 拼砌底台



10. 安裝樹景



11. 掛上小燈



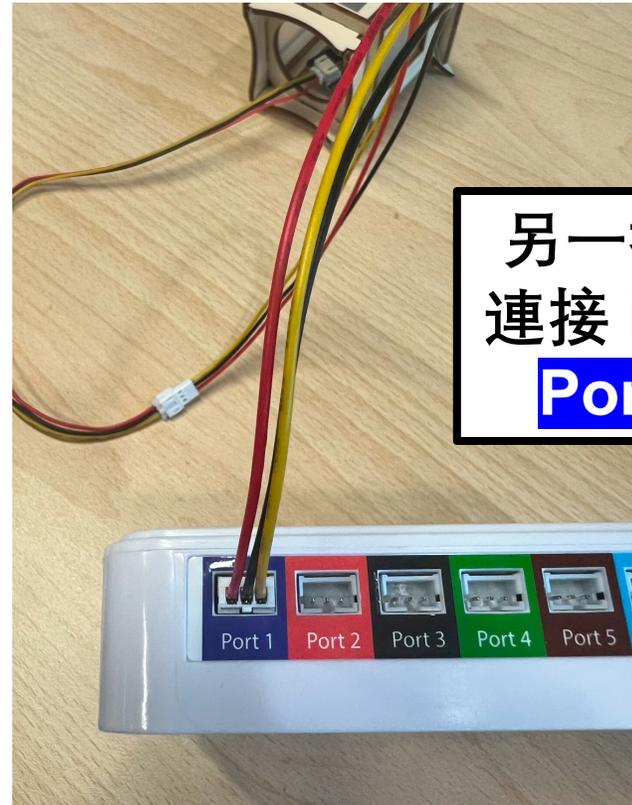
所需材料

掛鈎

12. 將小燈與 iPort 連接



連接第二條電線
(接口對插頭)



另一插頭
連接 iPort
Port 1

製作語言識別AI模型

數據輸入

注意：兩組數據的音頻有什麼分別

Background Noise

20 Samples

Mic Upload

Turn on

10 Samples

Mic Upload

Turn off

10 Samples

Mic Upload

AI模型訓練

Training

Train Model

Model

Open in iPort Customize UI

Microphon...

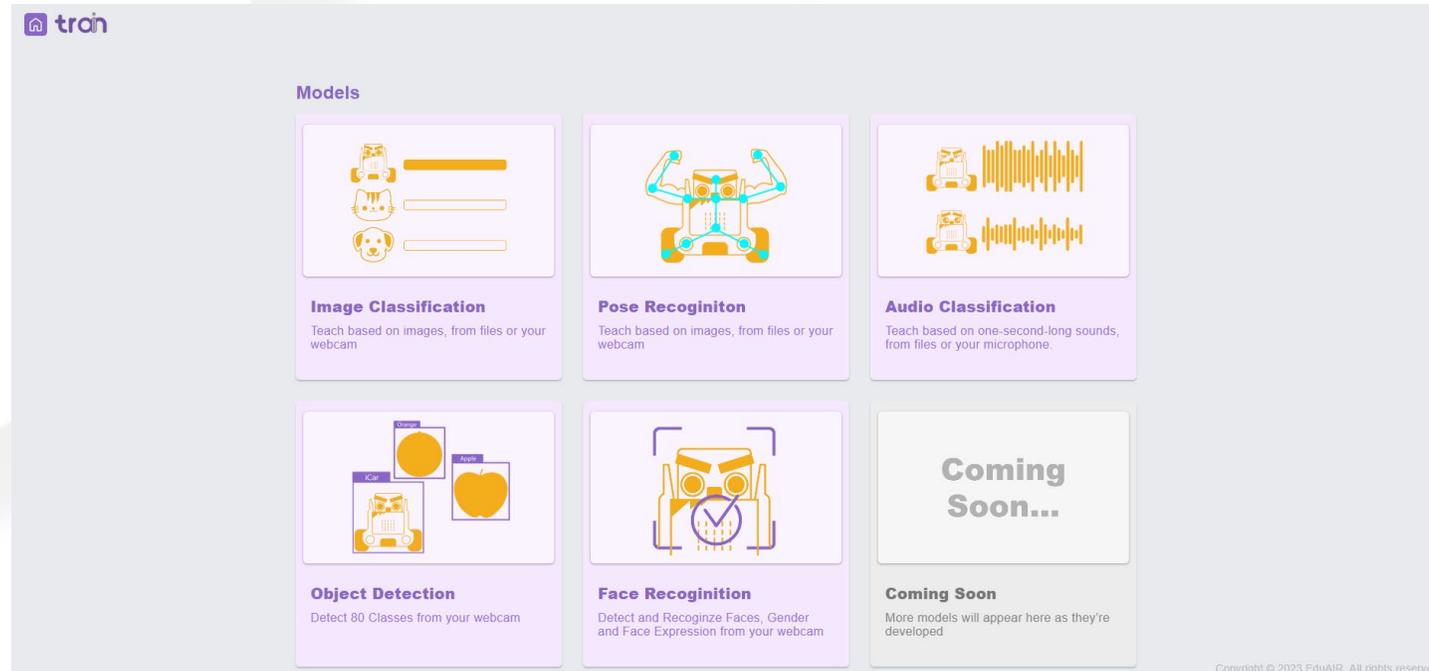
Classify every 0.5 second

Backgrou 23%

Turn off 9%

Turn on 68%

輸入以下網址



<https://itrain.eduairhk.com/>

點選 “Audio Classification”

train

Models

- Image Classification**
Teach based on images, from files or your webcam
- Pose Recogniton**
Teach based on images, from files or your webcam
- Audio Classification**
Teach based on one-second-long sounds, from files or your microphone.
- Object Detection**
Detect 80 Classes from your webcam
- Face Recognition**
Detect and Recognize Faces, Gender and Face Expression from your webcam
- Coming Soon...**
More models will appear here as they're developed

Copyright © 2023 EduAIR. All rights reserved.

點選 “Add Class”

The screenshot displays a user interface for managing machine learning classes. On the left, there are two class cards: "Background Noise" and "Class 2". Each card has a header with a three-dot menu icon and a section titled "Add Samples:" containing "Mic" and "Upload" buttons. Below these cards is a button labeled "Add Class", which is circled in black. To the right of the class cards is a "Training" panel with a gear icon and a "Train Model" button. Further right is a "Model" panel with a message: "You must train a model on the left before you can preview it here." Lines connect the class cards to the "Training" panel, and the "Training" panel to the "Model" panel.

新增一個 “Class”

The screenshot displays a user interface for managing machine learning classes. On the left, there are three class cards: 'Background Noise', 'Class 2', and 'Class 3'. Each card has a header with a title and a three-dot menu icon, followed by an 'Add Samples:' section with 'Mic' and 'Upload' buttons. The 'Class 3' card is circled in black. Below the class cards is an 'Add Class' button. On the right, there are two panels: 'Training' with a 'Train Model' button and a gear icon, and 'Model' with a text instruction: 'You must train a model on the left before you can preview it here.'

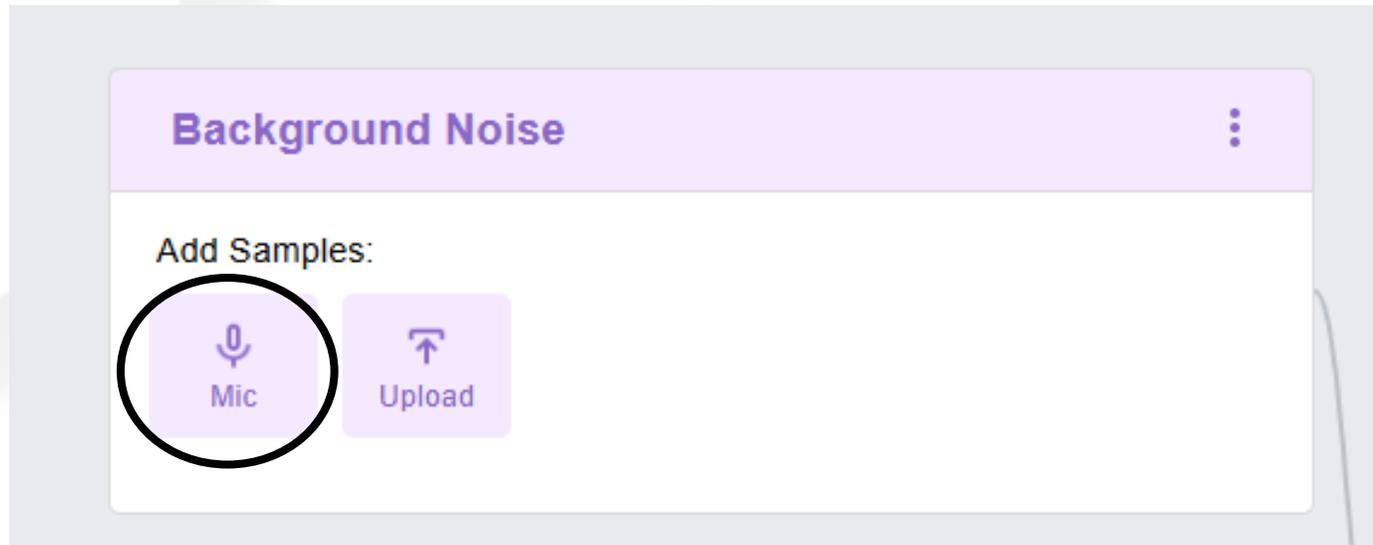
重新命名 “Class 2 & 3”



Class 2
→ **Turn on**

Class 3
→ **Turn off**

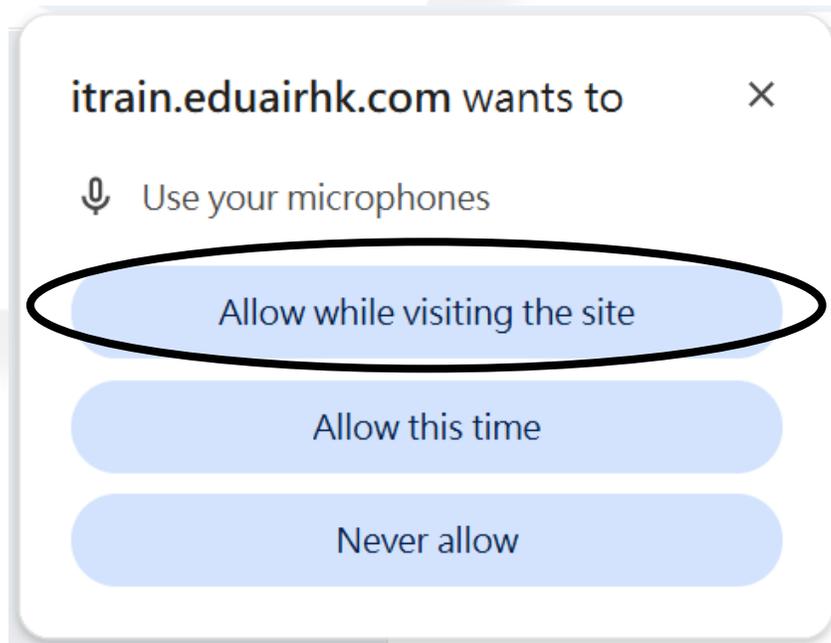
收集 “Background Noise” 數據



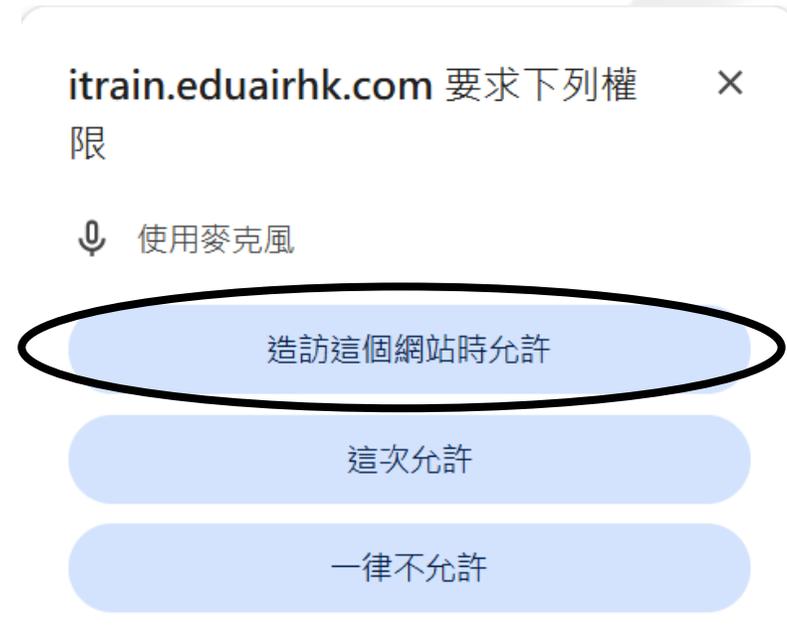
點選 “Mic”

允許電腦使用麥克風

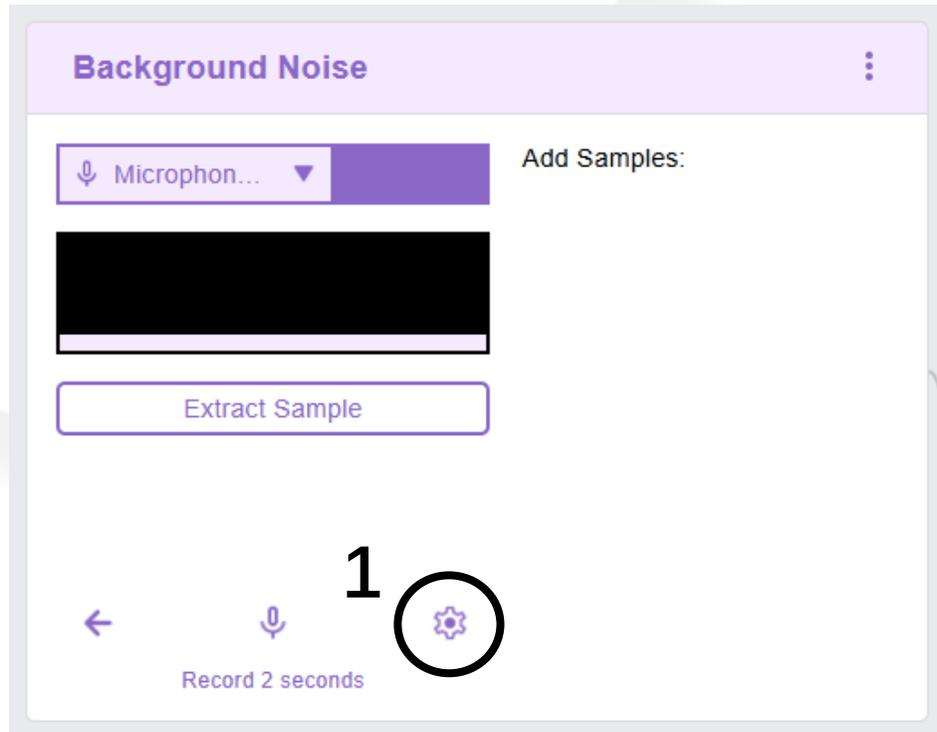
英文版本



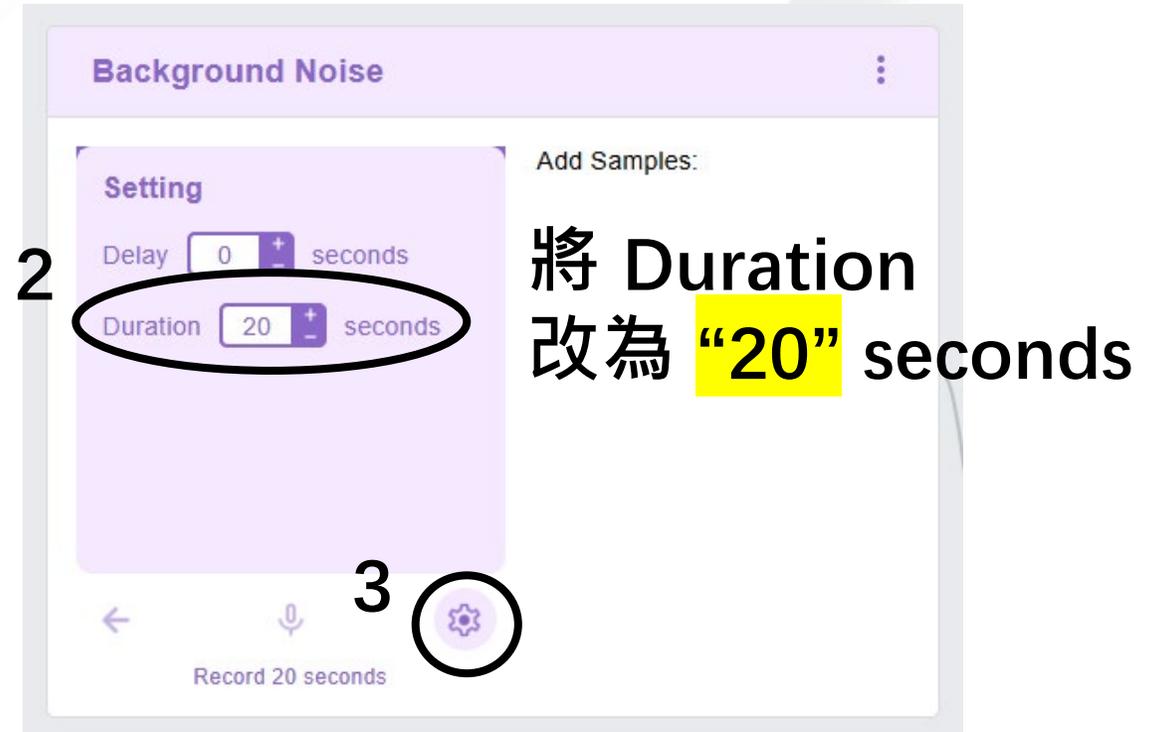
中文版本



收集 “Background Noise” 數據

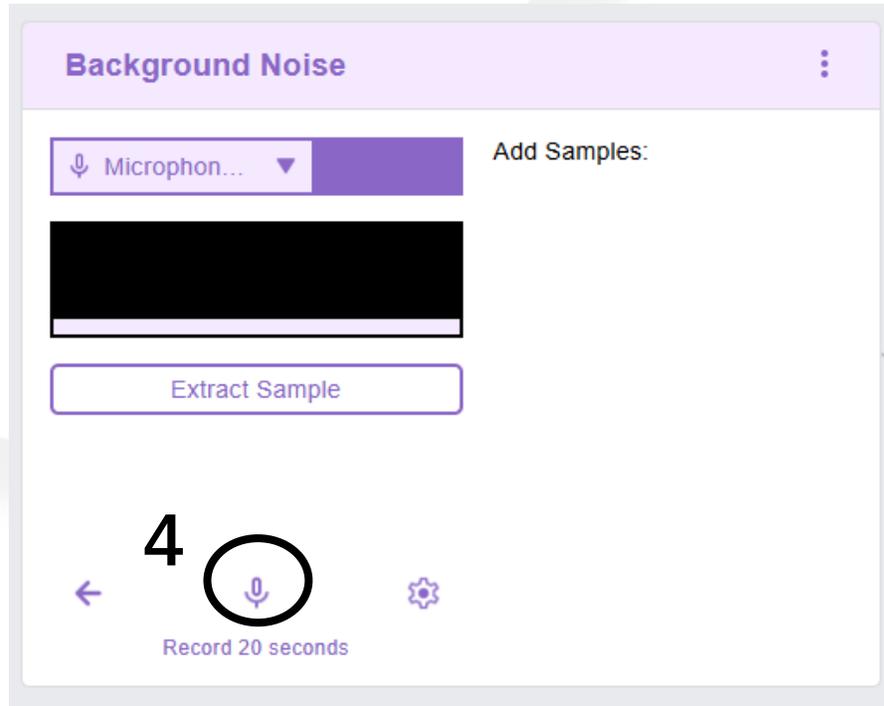


點選設定符號

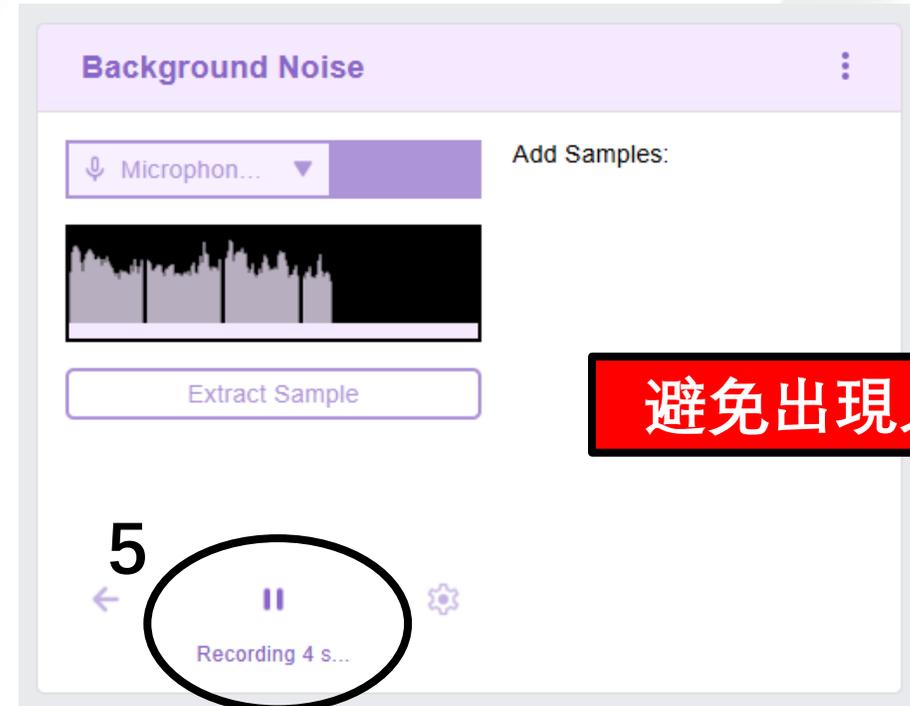


再次點選設定符號

收集 “Background Noise” 數據

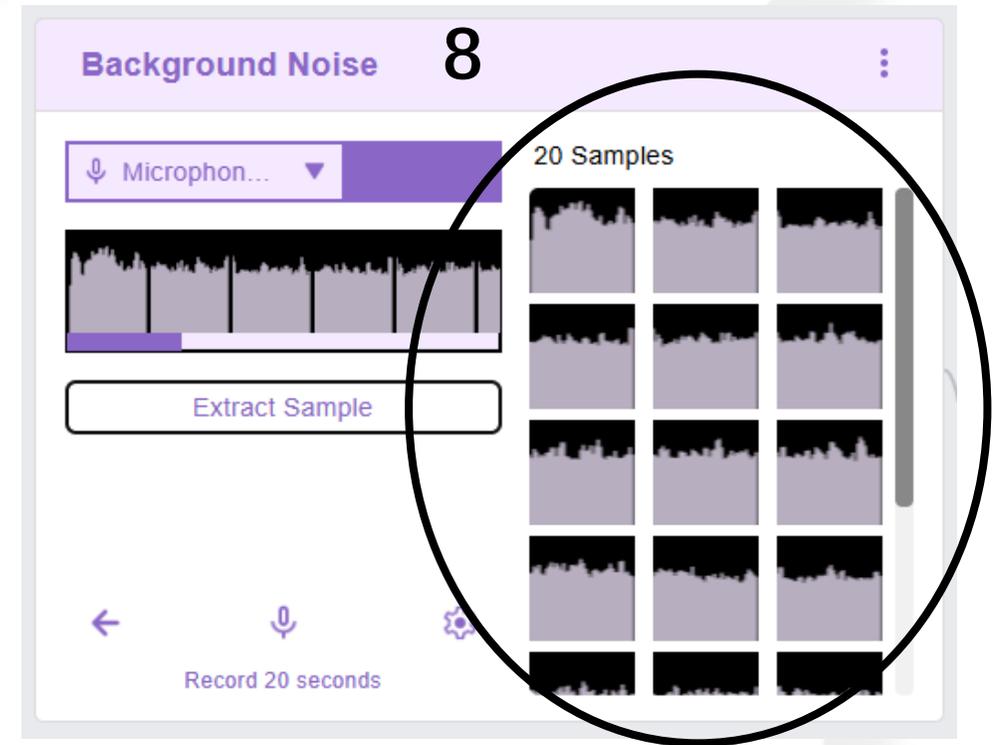


點選麥克風符號



等待聲音錄製 20秒, 不用按任何按鍵

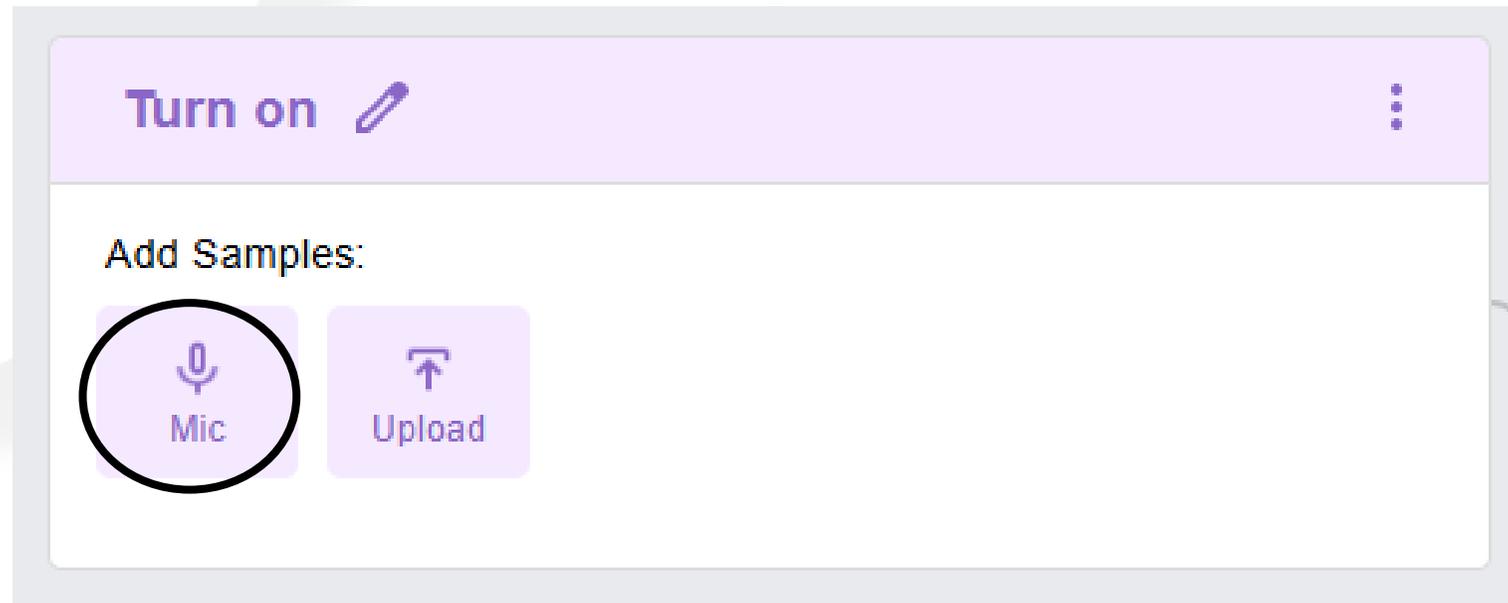
收集 “Background Noise” 數據



待聲音錄製完成後，檢查以上標示

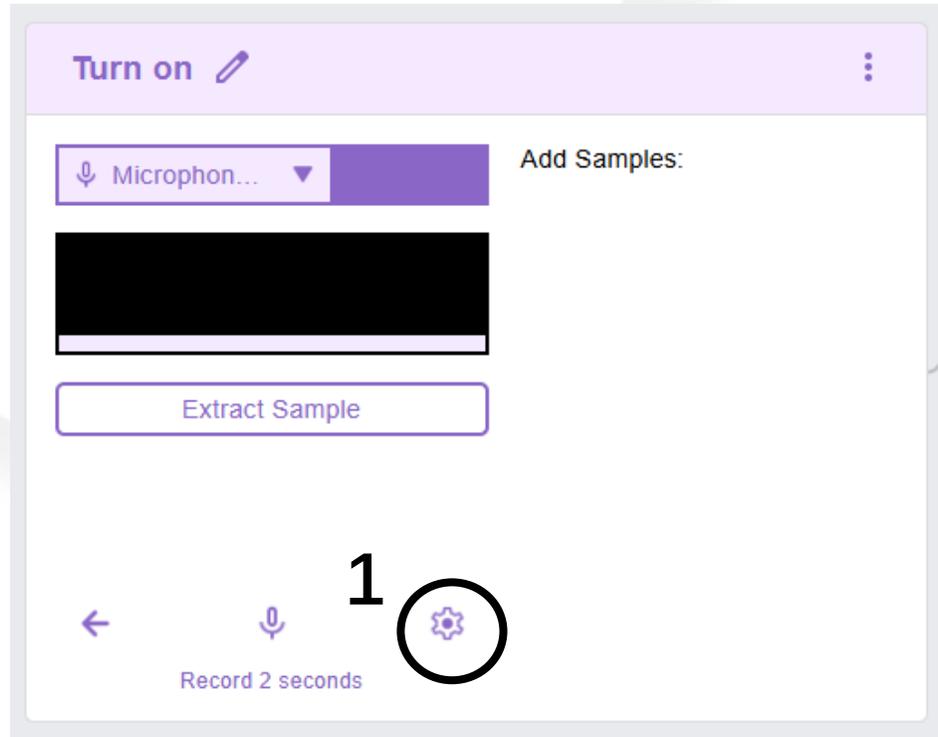
出現20組聲音數據

收集 “Turn on” 數據

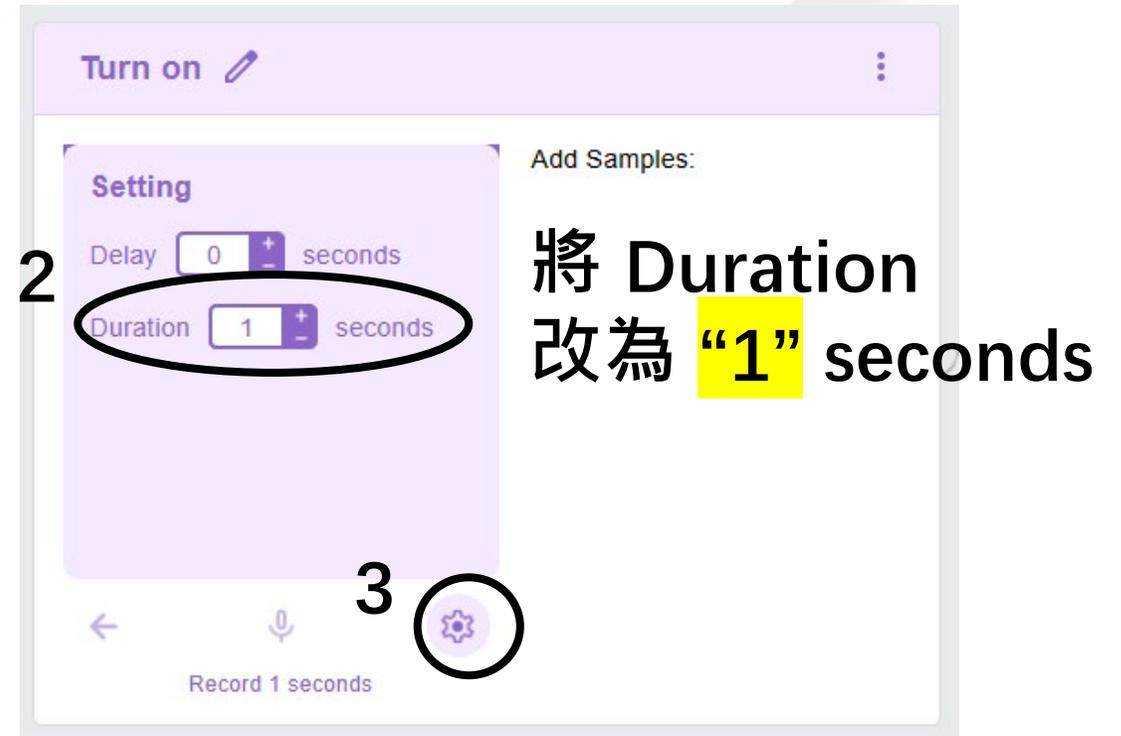


點選 “Mic”

收集 “Turn on” 數據

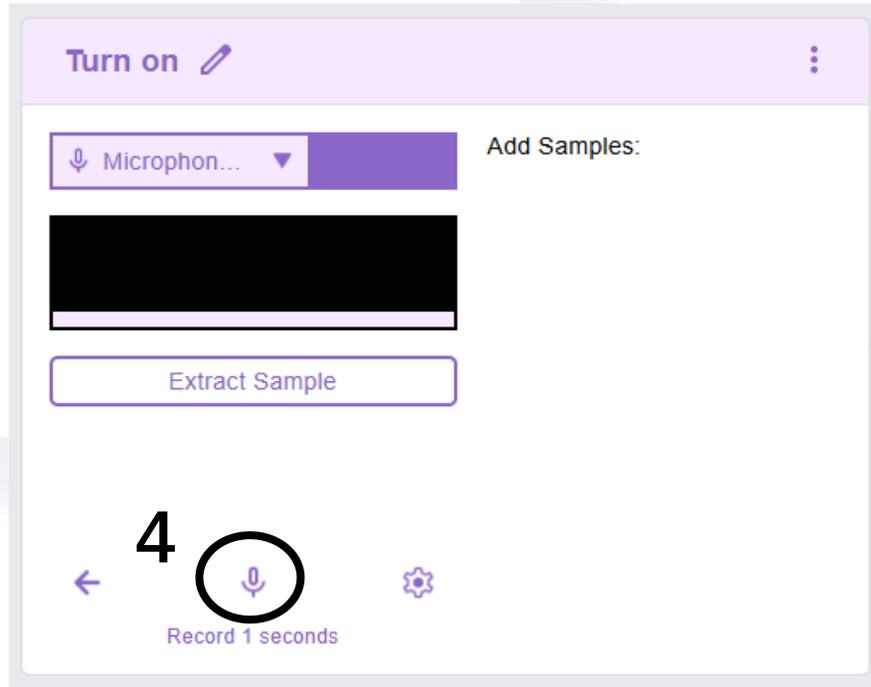


點選設定符號

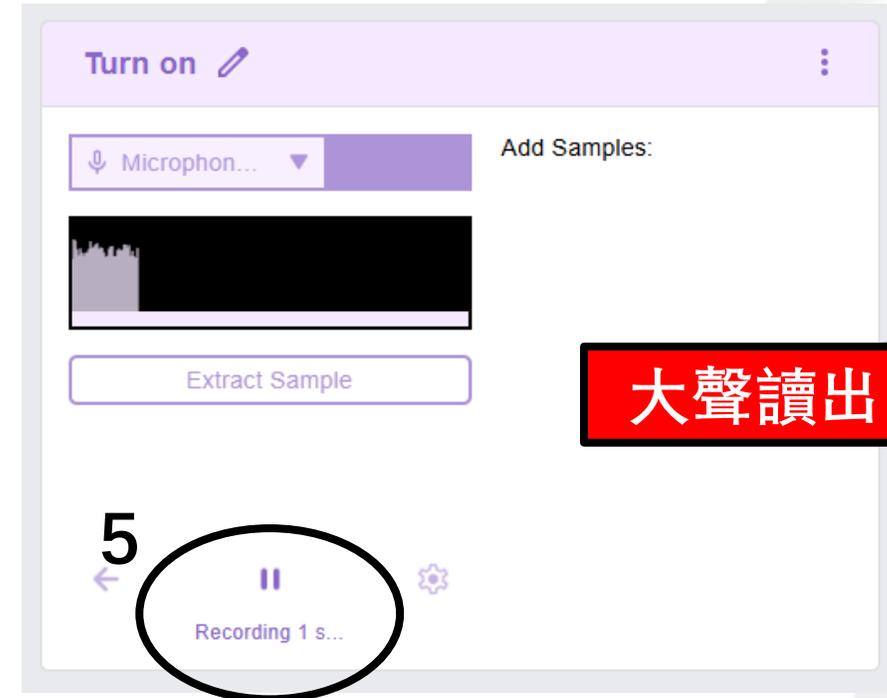


再次點選設定符號

收集 “Turn on” 數據



點選麥克風符號



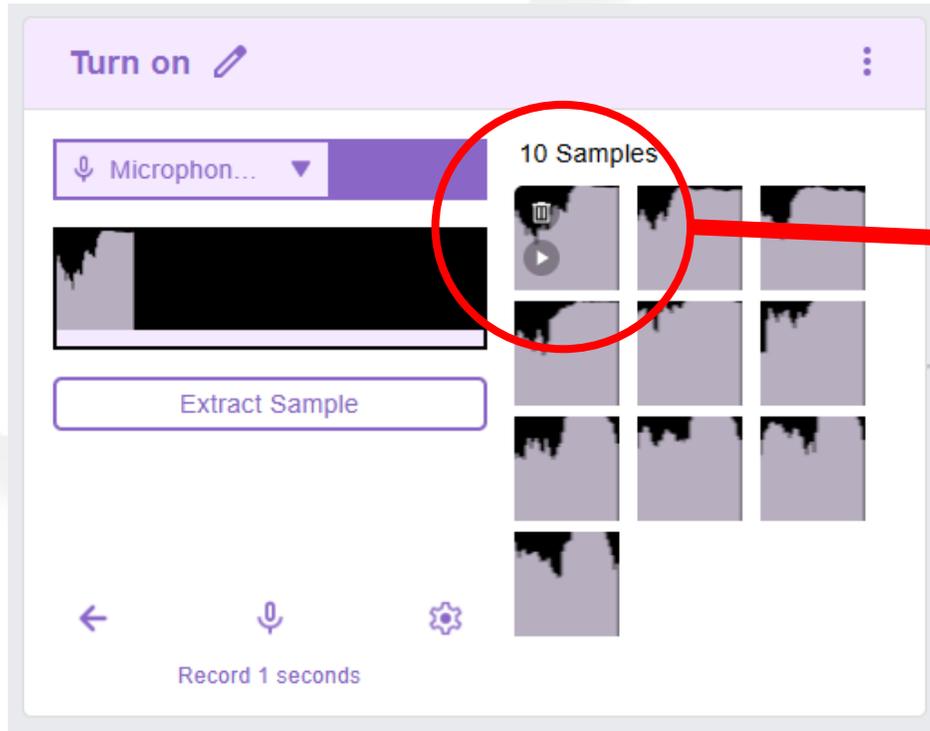
聲音將會錄製 1秒

收集 “Turn on” 數據



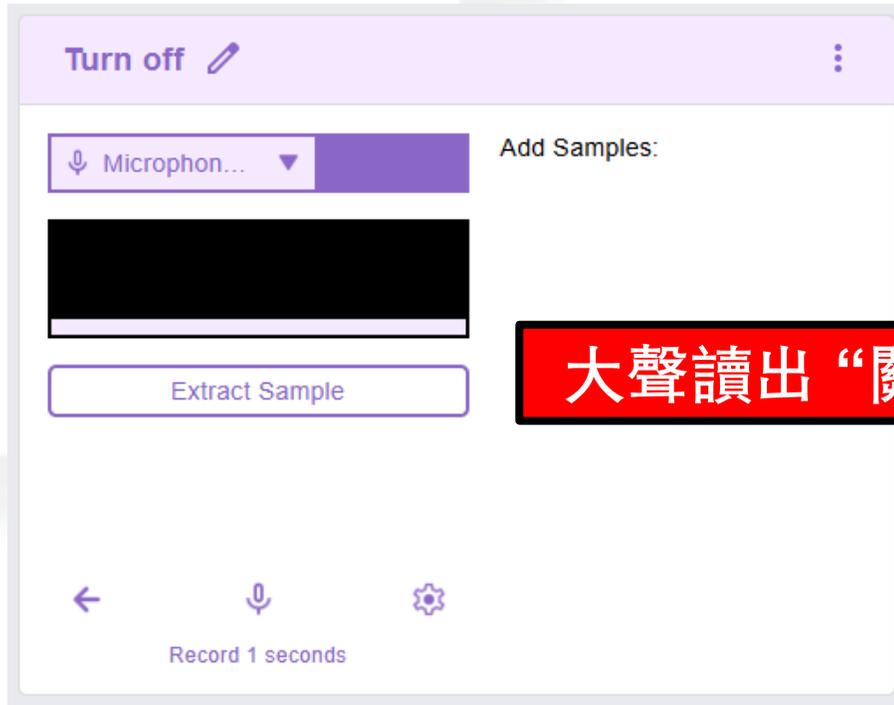
待聲音錄製完成後，檢查以上標示

收集 “Turn on” 數據

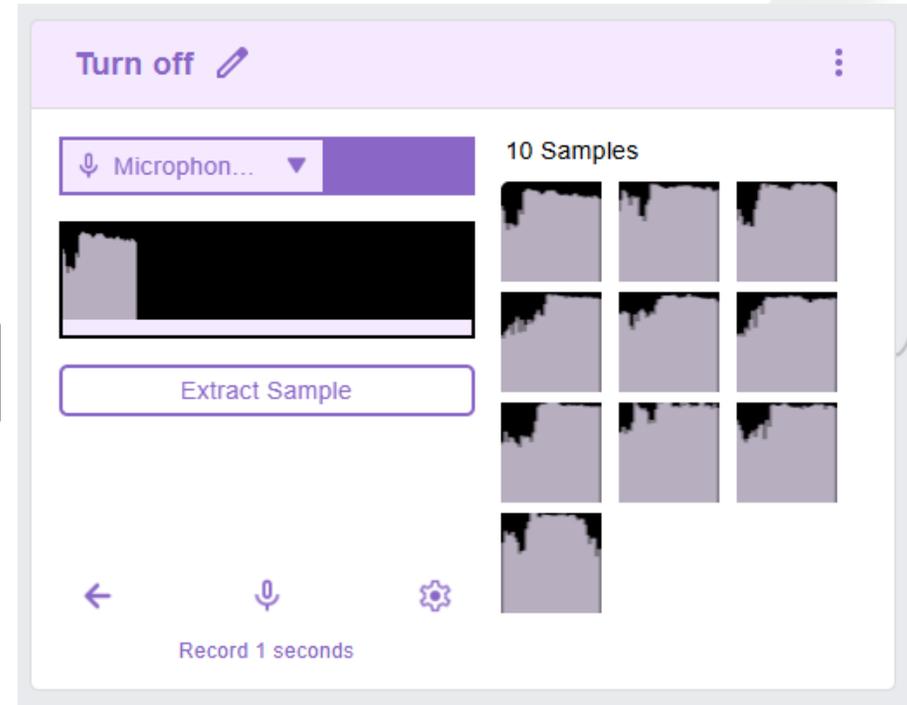


重複收集10組聲音數據

收集 “Turn off” 數據



大聲讀出“關”



與 “Turn on” 相同方法，錄製10組聲音數據

完成三種聲音數據收集

Background Noise
20組聲音數據

Turn on
10組聲音數據

Turn off
10組聲音數據

The screenshot shows a web interface for audio data collection and training. On the left, there are three data collection panels:

- Background Noise:** A purple header with a menu icon. Below it, it says "20 Samples". There are "Mic" and "Upload" buttons, followed by five waveform visualizations.
- Turn on:** A purple header with a pencil icon and a menu icon. Below it, it says "10 Samples". There are "Mic" and "Upload" buttons, followed by five waveform visualizations.
- Turn off:** A purple header with a pencil icon and a menu icon. Below it, it says "10 Samples". There are "Mic" and "Upload" buttons, followed by five waveform visualizations.

At the bottom left of the interface is an "Add Class" button. On the right side, there are two panels:

- Training:** A white panel with a gear icon and a "Train Model" button.
- Model:** A white panel with the text: "You must train a model on the left before you can preview it here."

Lines connect the three data collection panels to the "Train Model" button, indicating that data from these classes is used for training.

點選 “Train Model”

The screenshot displays a user interface for training a model. On the left, there are three sections for adding background noise:

- Background Noise**: 20 Samples. Includes a 'Mic' button, an 'Upload' button, and a waveform visualization.
- Turn on**: 10 Samples. Includes a 'Mic' button, an 'Upload' button, and a waveform visualization.
- Turn off**: 10 Samples. Includes a 'Mic' button, an 'Upload' button, and a waveform visualization.

At the bottom left, there is an 'Add Class' button. On the right side, there are two panels:

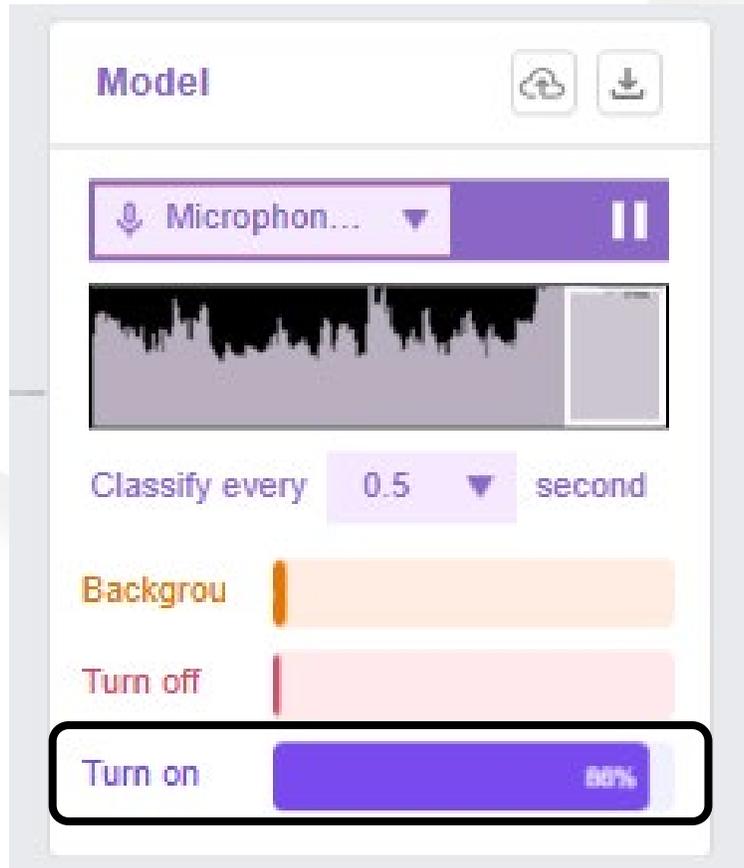
- Training**: Contains a 'Train Model' button, which is circled in black.
- Model**: Contains the text: "You must train a model on the left before you can preview it here."

完成語言識別 AI 模型

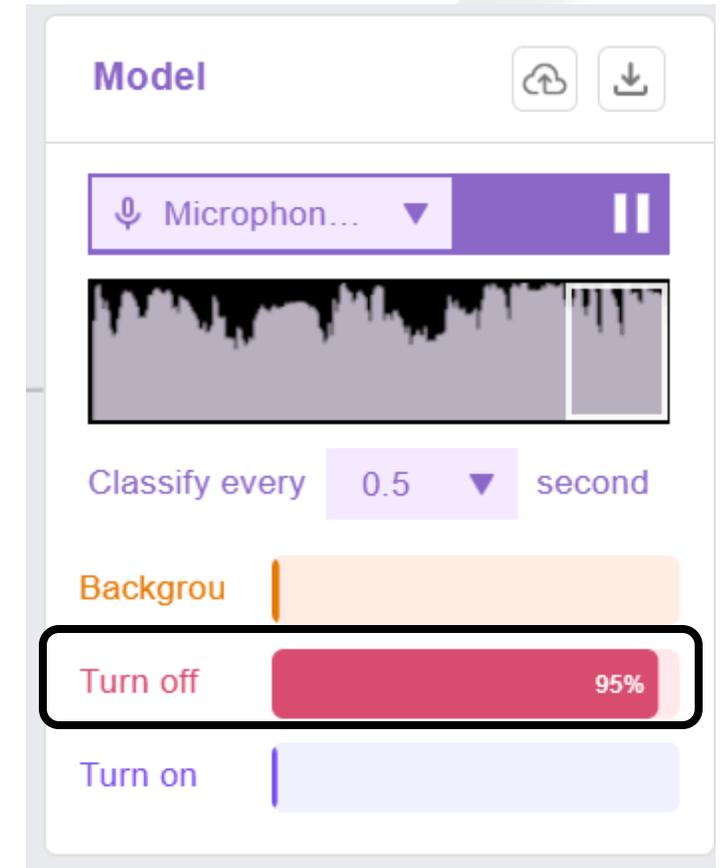
The screenshot displays a software interface for training an AI model. On the left, there are three sections for collecting audio samples: 'Background Noise' (20 Samples), 'Turn on' (10 Samples), and 'Turn off' (10 Samples). Each section includes a 'Mic' button, an 'Upload' button, and a waveform visualization. A 'Training' panel in the center contains a 'Train Model' button. On the right, a 'Model' preview window is shown, featuring a microphone control, a waveform, a 'Classify every 0.5 second' setting, and three progress bars for 'Backgrou' (23%), 'Turn off' (9%), and 'Turn on' (68%).

彈出 AI 模型預覽視窗

檢查 AI 模型

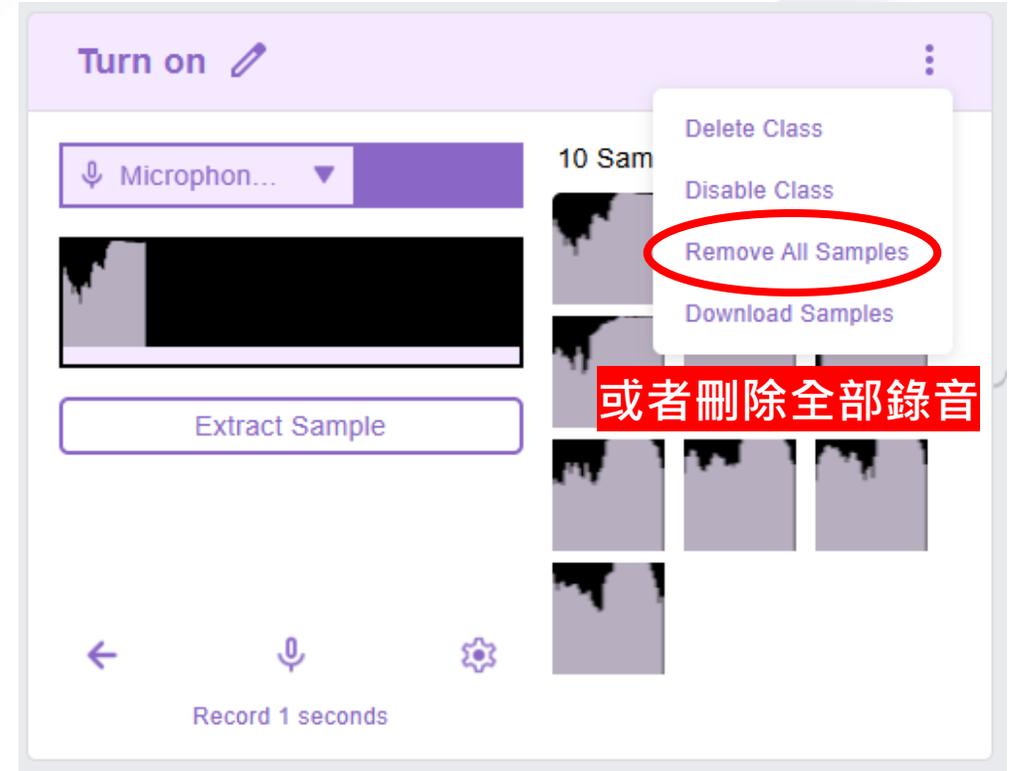


測試讀出“開”及“關”
檢查聲音的輸出結果
達到 80 % 以上



檢查 AI 模型

若然聲音的輸出結果無法達到 80% ...



再重新進行聲音錄製

檢查 AI 模型

Background Noise

20 Samples

Mic Upload

Turn on

10 Samples

Mic Upload

Turn off

10 Samples

Mic Upload

Add Class

Training

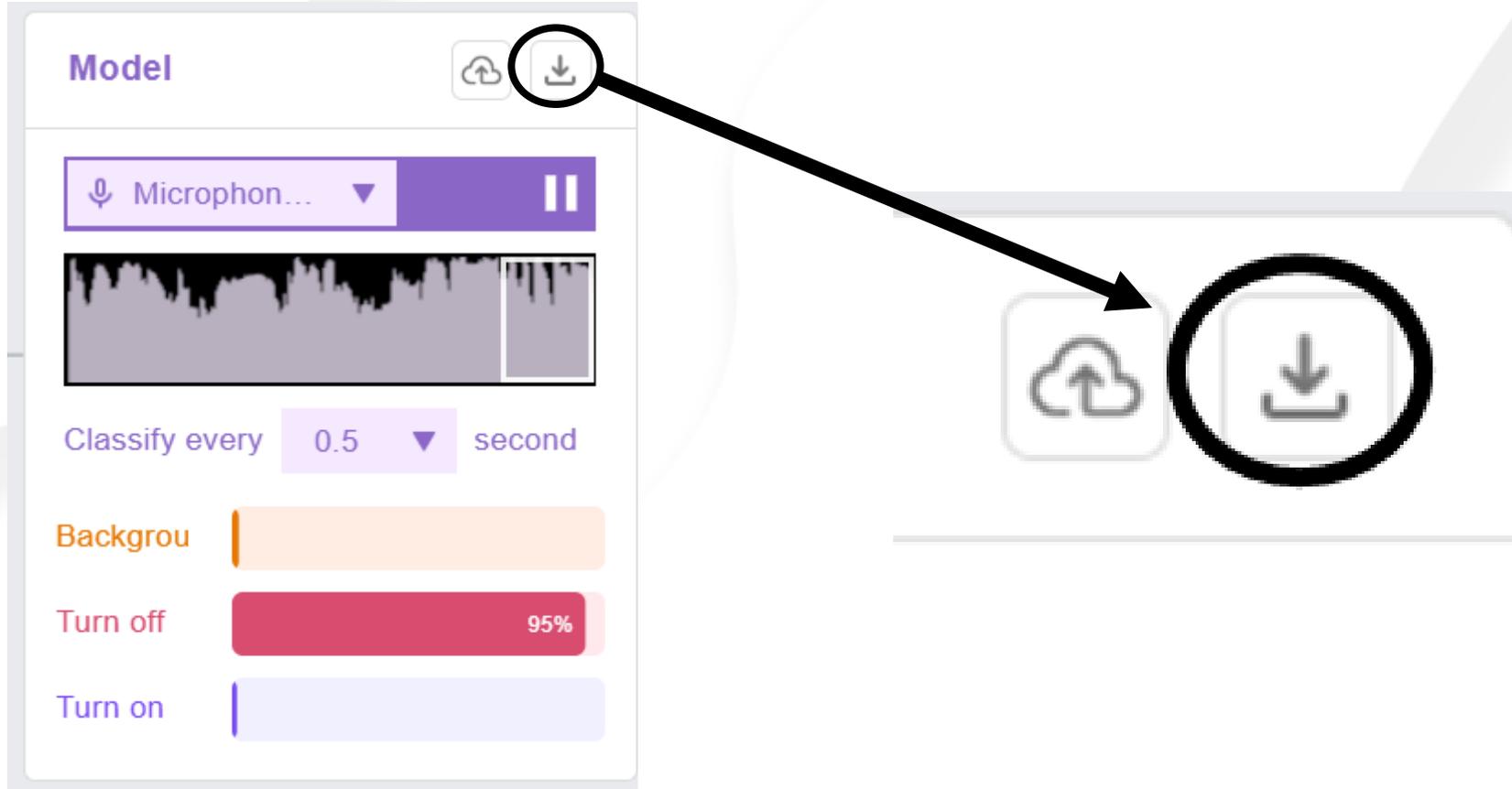
Train Model

Model

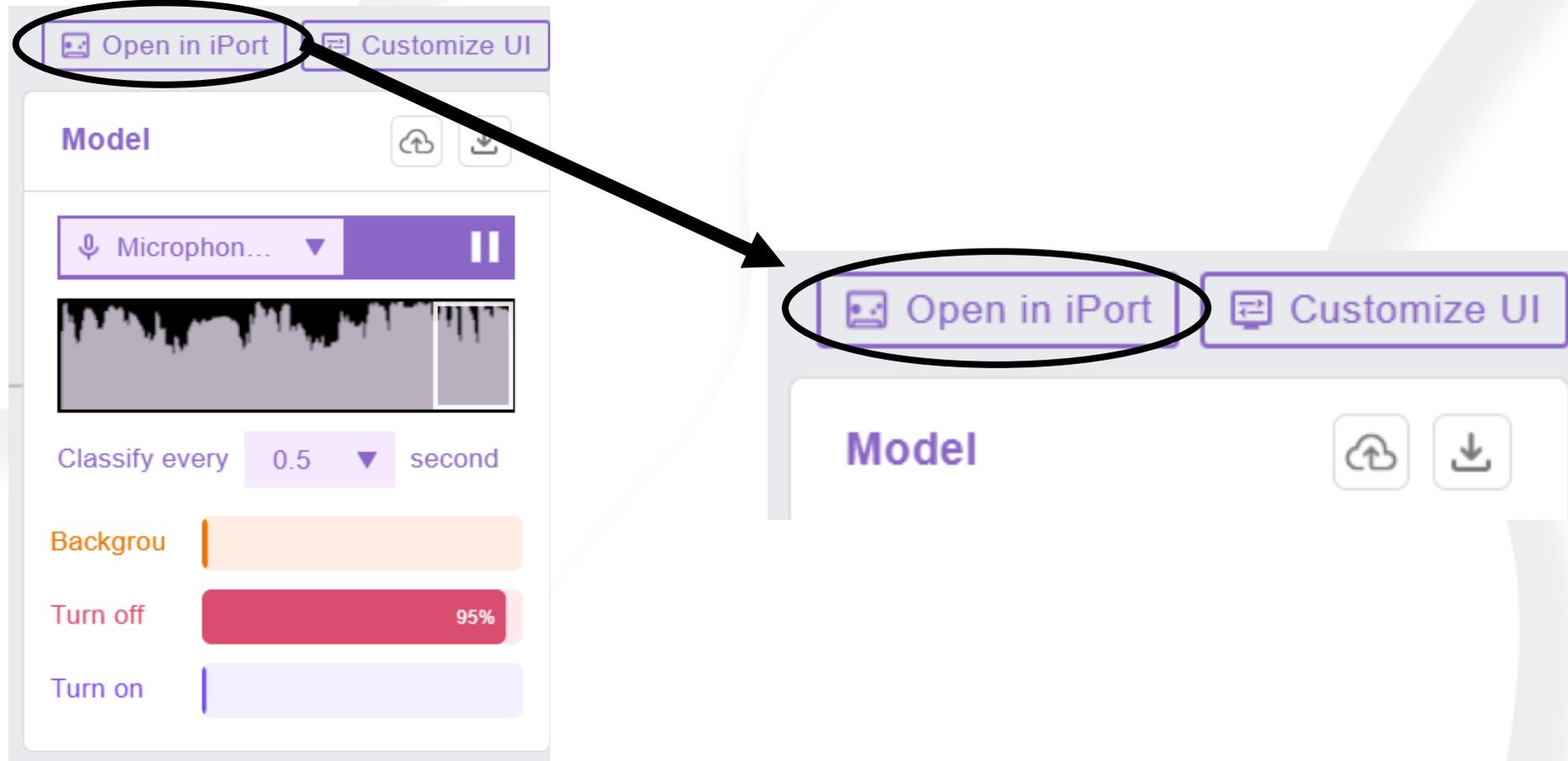
You must train a model on the left before you can preview it here.

整理數據後，
重新訓練模型

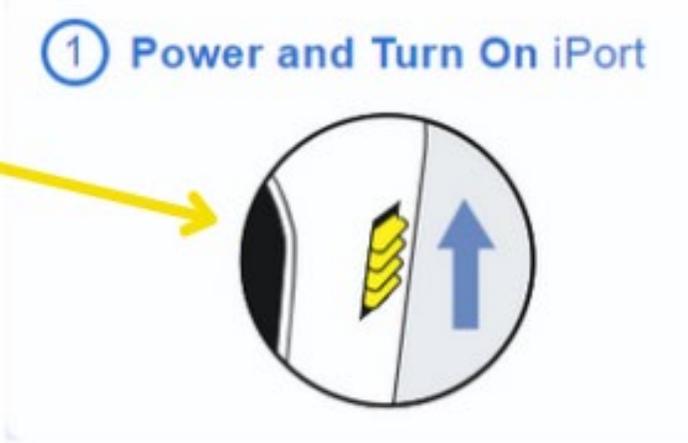
點選下載符號



點選 “Open in iPort”



開啟 iPort



把黃色開關向上推

連接 iPort USB



USB 連接器



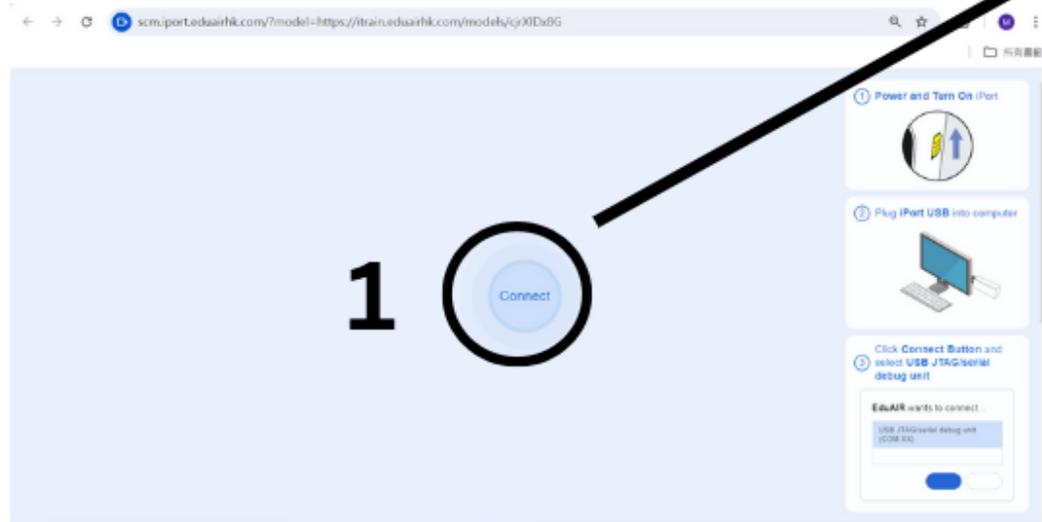
② Plug iPort USB into computer



把USB插進電腦

連接 iPort USB

點選“Connect”符號



scm.iport.eduairhk.com 要求與序列埠連線

USB JTAG/serial debug unit (COM10) - 已配對

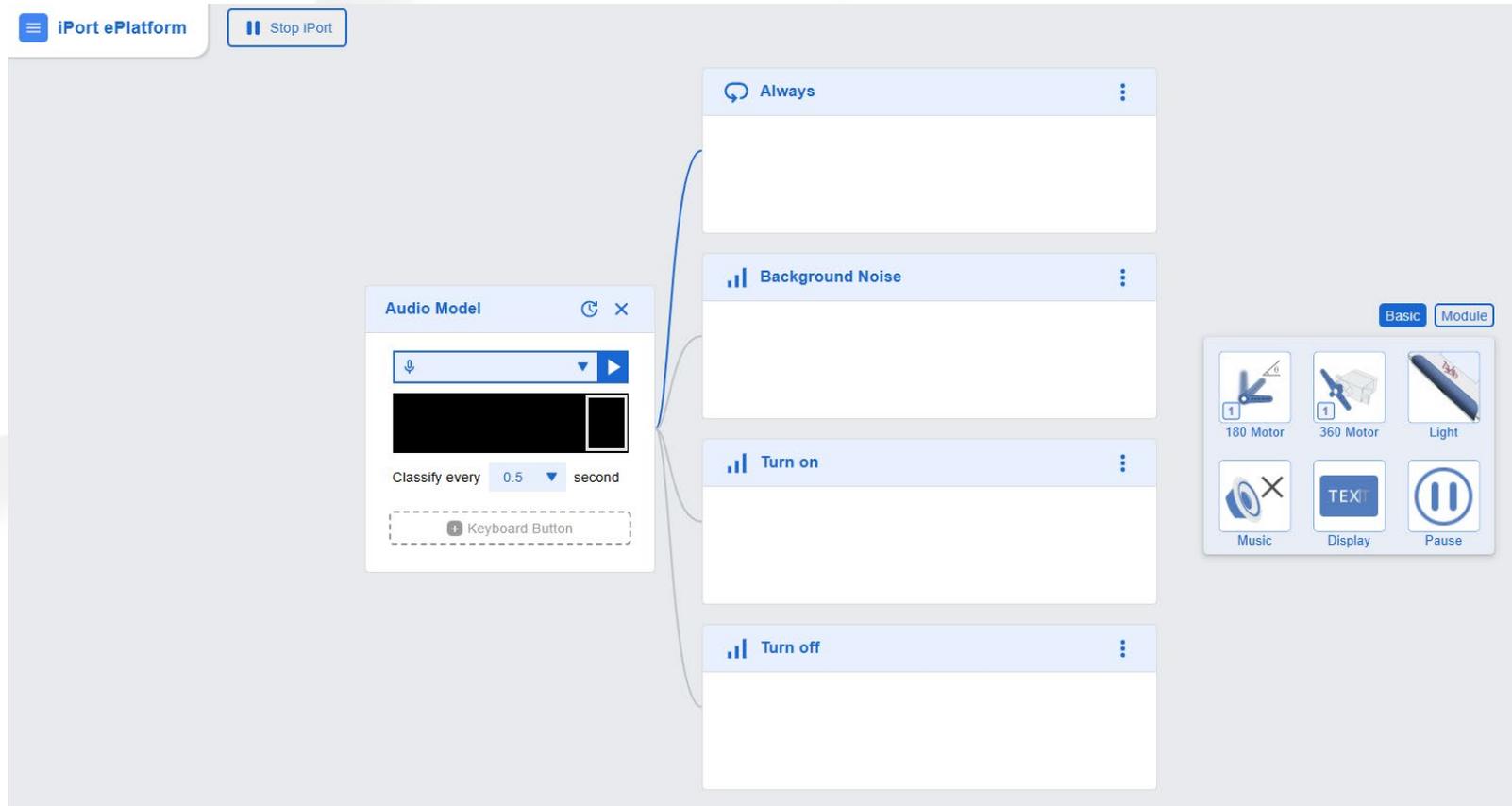
2

在電腦選取 **USB JTAG /serial debug unit**

點選“連線/Connect”



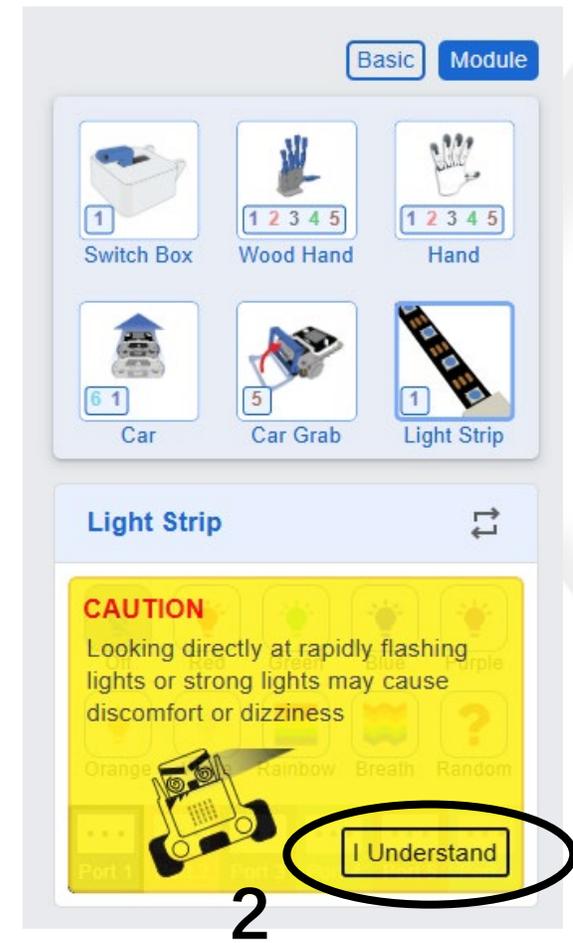
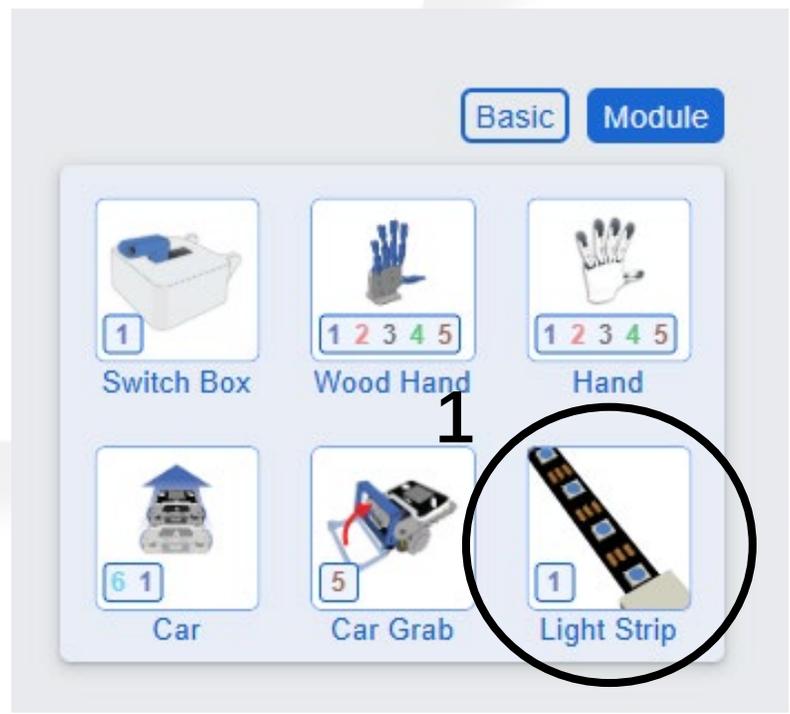
進入 iPort ePlatform



點選 “Module”

The screenshot displays the iPort ePlatform interface. At the top left, there is a menu icon and the text "iPort ePlatform". To its right is a "Stop iPort" button. The main workspace contains a vertical sequence of four blocks: "Always", "Background Noise", "Turn on", and "Turn off". A blue line connects the "Background Noise" block to an "Audio Model" dialog box on the left. The "Audio Model" dialog has a microphone icon, a play button, and a "Classify every 0.5 second" setting. Below this is a "Keyboard Button" option. To the right of the main workspace is a "Basic" module palette. A "Module" button in this palette is circled in black. A large black arrow points from this circled button to a larger callout box on the right. This callout box shows a grid of modules: "180 Motor", "360 Motor", "Light", "Music", "Display", and "Pause". The "Module" button in the "Basic" palette is also circled in black within this callout box.

點選 “Light Strip”



輸入燈光指令

Basic Module

1 Switch Box 2 Wood Hand Hand

6 1 Car 5 Car Grab 1 Light Strip

Light Strip

1 Off Red Green Blue Purple

Orange White Rainbow Breath Random

Port 1 Port 2 Port 3 Port 4 Port 5 Port 6

檢查已選取 Port 1

點選完成的
燈光指令

選取喜愛的燈光

輸入燈光指令

將燈光指令
拉放到
“Turn on”



輸入燈光指令



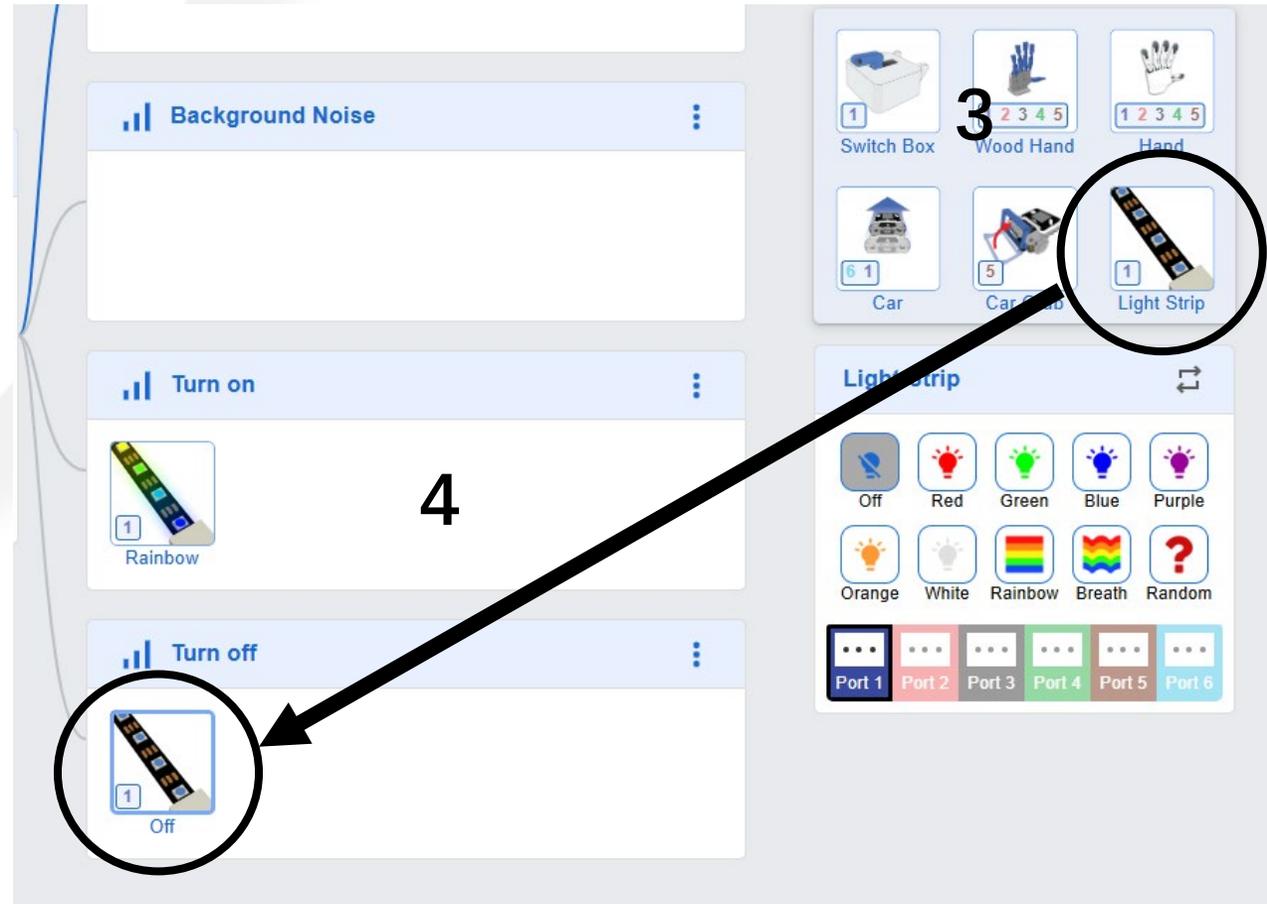
重新點選燈光
指令按鍵

選取 “Off”

檢查已選取 Port 1

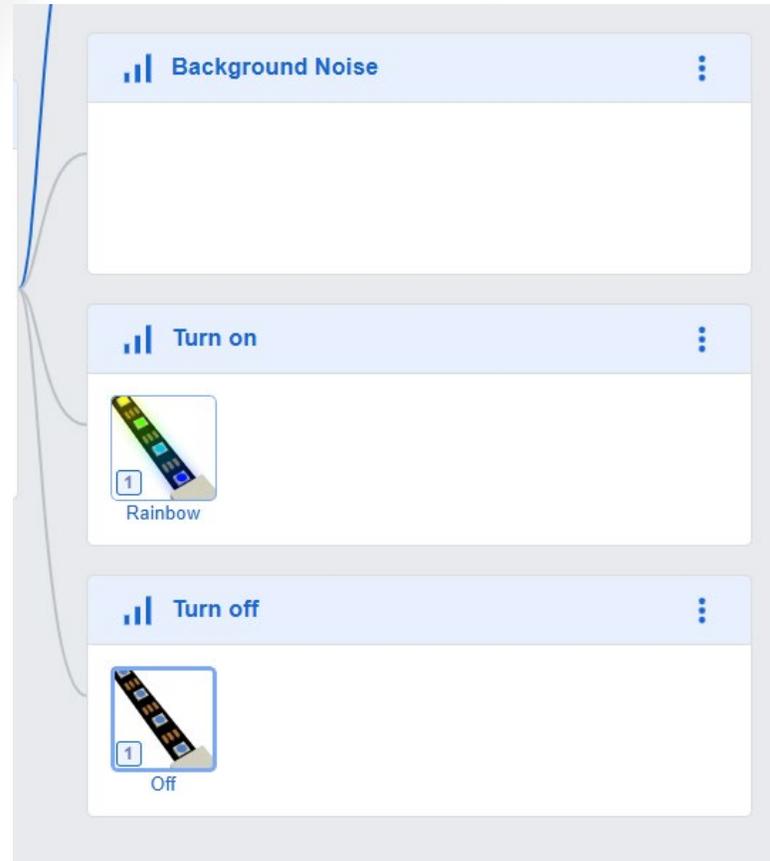
輸入燈光指令

將燈光指令
拉放到
“Turn off”



點選完成的
燈光指令

檢查燈光指令



The screenshot shows a control panel with three sections:

- Background Noise:** A section with a blue header containing a signal strength icon and the text "Background Noise". The area below is empty.
- Turn on:** A section with a blue header containing a signal strength icon and the text "Turn on". Below the header is an icon of a light bar with a small square containing the number "1" in the bottom-left corner. The text "Rainbow" is centered below the icon.
- Turn off:** A section with a blue header containing a signal strength icon and the text "Turn off". Below the header is an icon of a light bar with a small square containing the number "1" in the bottom-left corner. The text "Off" is centered below the icon.

4

語音識別的限制及潛在問題

語音識別的限制

語音受以下因素影響：

- 說話者的特徵：

說話風格、年齡、口音、詞彙、措辭、停頓

- 環境因素：背景噪音

- 需要精良的錄音設備

語音識別的限制

■ 試想像聲控開關有什麼潛在問題？



私隱問題？

你是否接受AI長期
監聽你的言行？

隱私問題

科技公司可能未經我們同意作出以下行為：

1. 使用你的敏感或機密信息
2. 無論你是否在跟虛擬助手說話，智能揚聲器和虛擬助手都可能在任何時候監聽你的說話
3. 錄製我們的語音成他們的訓練數據



來源：[Arstechnica](#)

Apple Siri Eavesdropping Puts Millions Of Users At Risk

 **Kate O'Flaherty** Senior Contributor @
Cybersecurity
I'm a cybersecurity journalist.



來源：[Forbes](#)

5

課堂總結

學習總結

1. 什麼是語音識別？

- 電腦自動將人類的語音內容轉換為相應的文字用於日常生活當中。

2. 語音識別有什麼應用？

- 控制家居設備
- 手機應用
- 語音轉錄

學習總結

3. 語音識別有什麼限制？

- 說話者的特徵
- 環境因素
- 需要精良的錄音設備

4. 語音識別有會帶來什麼倫理問題？

- 私穩問題