

Prediction of heart disease

Decision tree

Objective	1
Dataset	1
Create, train, learn, and test	2
Implementation	14

Objective

We want to create a model that predicts whether a person has a high risk of heart disease based on certain demographic and diagnostic metrics of the individual.

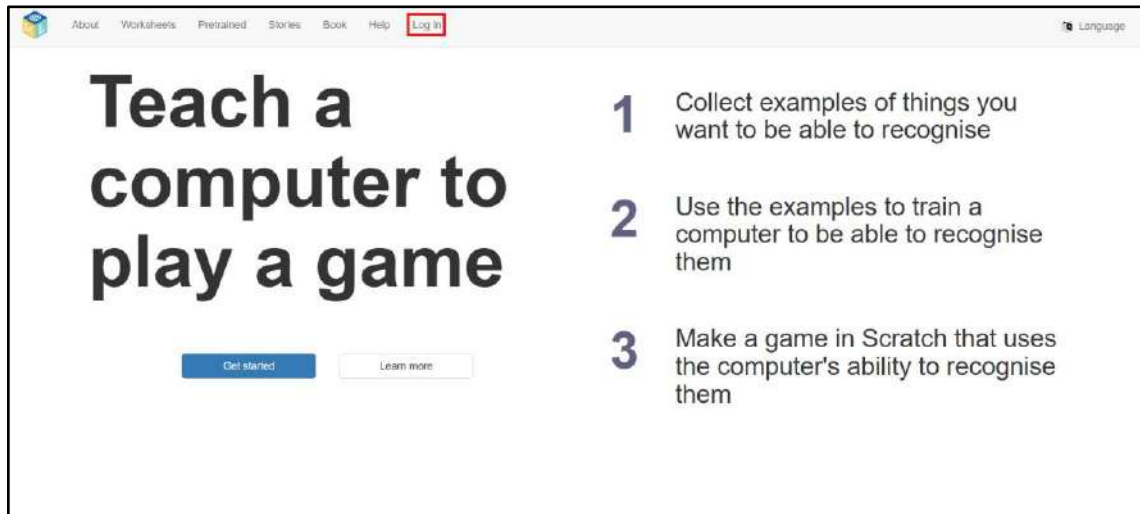
Dataset

- https://huggingface.co/datasets/Shambhavi625/heart_diseases/
- We will keep some of the rows and columns of the original dataset and change the column names to meet the needs of ML4Kids
- The table below shows the columns and their possible values as well as a description of them

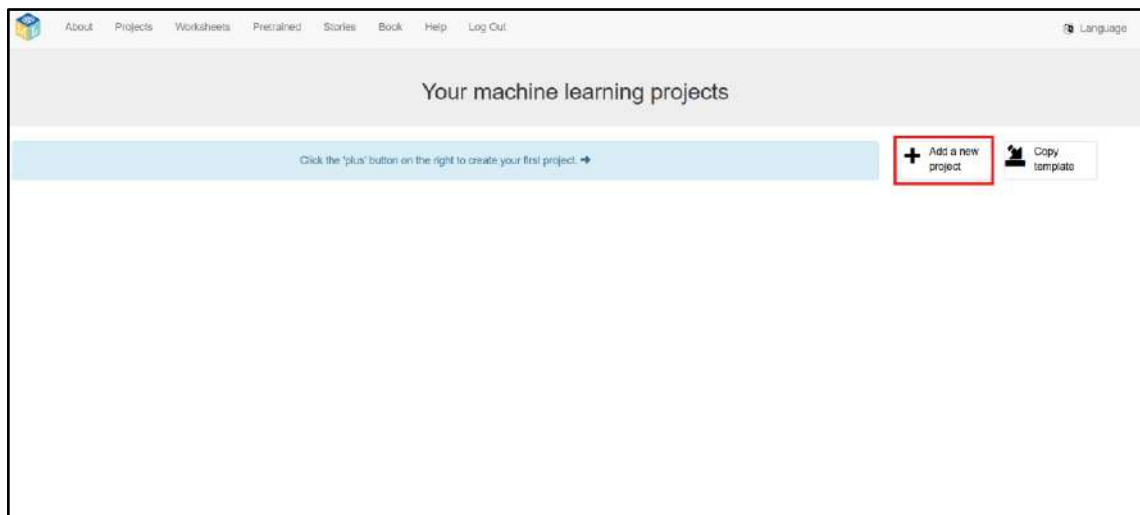
Column	Meaning	Possible values	Price Description
Age	Age of the patient	Numeric (30–80)	Higher age → higher risk of heart disease
Sex	Biological sex	M, F	M = male , F = female
ChestPainTyp	Type of chest pain	TA, ATA, NAP, ASY	TA = Classic heart attack-type pain ATA = Atypical angina - Similar but not school heart pain NAP = Non-angina pain - Chest pain that is NOT caused by the heart ASY = Asymptomatic - No chest pain at all
RestingBP	Blood pressure at rest	Numeric (~100–180)	~120 °C normal 130–139 °C increased ≥140 °C hypertension
Cholesterol	Cholesterol level	Numeric (~100–600)	<200 good 200–239 marginal ≥240 high (risk of artery blockage)
FastingBS	Fasting blood sugar	0, 1	0 = normal 1 = high (possible diabetes)
RestingECG	Electrocardiogram result at rest	Normal, ST, LVH	Normal = healthy ST = possible ischemia LVH = thick heart muscle
MaxHR	Maximum heart rate	Numeric (~60–200)	Higher = better heart performance Lower = possible heart problem
ExerciseAngi	Chest pain during exercise	Y, N	Y = (Yes) pain when exercising (bad sign) N = (No) no pain
Oldpeak	How much oxygen deficiency does the heart show during the effort	Numeric (0–5+)	0 = normal 1–2 = moderate problem >2 = severe cardiac stress

Create, train, learn, and test

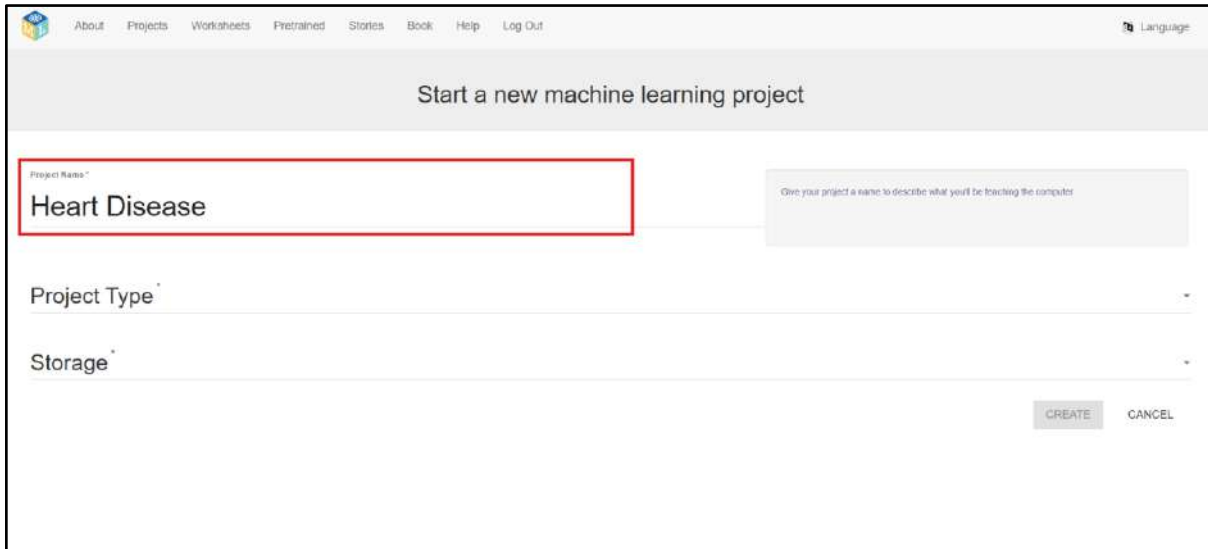
- Open ML4kids by visiting the link:
 - <https://machinelearningforkids.co.uk/>
- Log in – "Log in"



- Click "Add a new project"

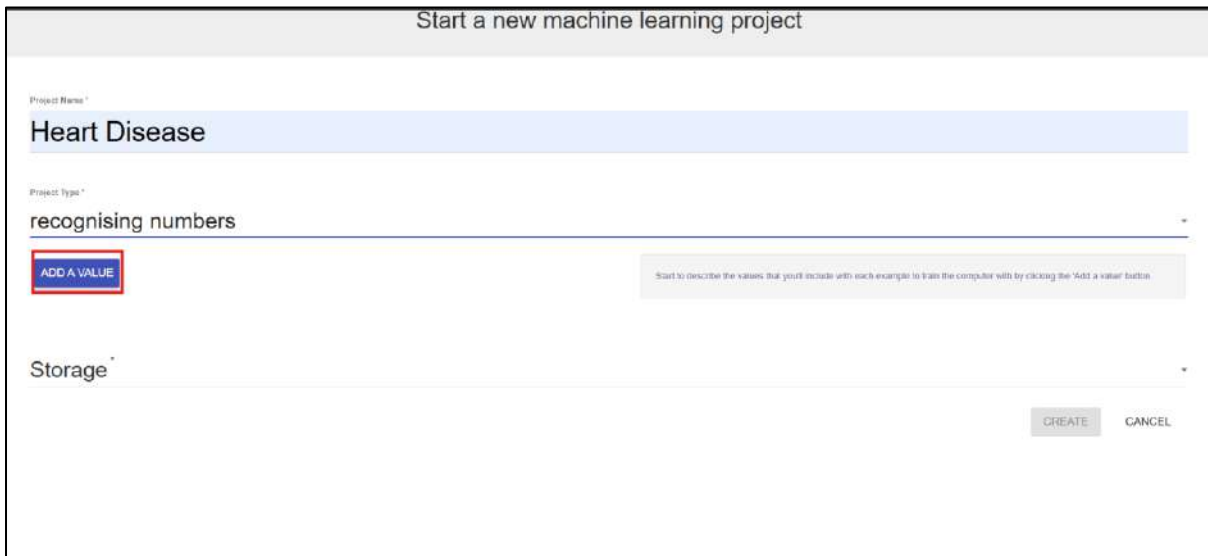


- Add "Project Name"



The screenshot shows the 'Start a new machine learning project' form. The 'Project Name' field is highlighted with a red rectangle and contains the text 'Heart Disease'. To the right of this field is a grey box with the text 'Give your project a name to describe what you'll be teaching the computer'. Below the 'Project Name' field are two dropdown menus: 'Project Type' and 'Storage'. At the bottom right of the form are two buttons: 'CREATE' and 'CANCEL'.

- As "Project Type" select:
 - "Recognising numbers"
- Click "ADD A VALUE"



The screenshot shows the 'Start a new machine learning project' form. The 'Project Name' field contains 'Heart Disease'. The 'Project Type' dropdown menu is set to 'recognising numbers'. Below the 'Project Type' field is a blue button labeled 'ADD A VALUE', which is highlighted with a red rectangle. To the right of this button is a grey box with the text 'Start to describe the values that you'll include with each example to train the computer with by clicking the "Add a value" button'. Below the 'Project Type' field is the 'Storage' dropdown menu. At the bottom right of the form are two buttons: 'CREATE' and 'CANCEL'.

- For each of the columns in the table above add the name of the value and the "**Type of value**"
- The possible types of value in ML4Kids are "**number**" and "**multiple-choice**", so if the price type is not numeric, you need to choose "**multiple-choice**" for everything else. If you select "**multiple-choice**", you must manually specify the possible values.
- First we add the value "**Age**" which is a numerical value
- We select "**Type of value**" -> **number**
- And then we click on "**ADD ANOTHER VALUE**"

Start a new machine learning project

Please choose where you would like to store your project

Project Name*

Heart Disease

Project Type*

recognising numbers

Value 1*

Age

Type of value*

number

If Age can be described as numbers, choose "number"

If it can be described as choosing from a few options, choose "multiple-choice"

ADD ANOTHER VALUE

Storage*

- Now we will add the value "**Sex**" which is not a numerical value, so its formula is "**multiple-choice**"
- A "Choices" field appears

Please choose where you would like to store your project

Project Name*

Heart Disease

Project Type*

recognising numbers

Value 1*

Age

Type of value*

number

If Age can be described as numbers, choose "number"

If it can be described as choosing from a few options, choose "multiple-choice"

Value 2*

Sex

Type of value*

multiple-choice

Choices

add a choice

If Sex can be described as numbers, choose "number"

If it can be described as choosing from a few options, choose "multiple-choice"

ADD ANOTHER VALUE

Storage*

CREATE CANCEL

- We will add "**F**" as options
- We click on "**enter**"
- Then write "**M**"
- And click "**enter**"

- After adding the first two values (**Age**, **Sex**), your screen should look like this:

Please choose where you would like to store your project

Project Name *

Heart Disease

Project Type *

recognising numbers

Value 1 *
Type of value *
Age
number

Value 2 *
Type of value *
Sex
multiple-choice

Choices
F M
add a choice

ADD ANOTHER VALUE

Storage *

- We do the same for the rest of the column names, after entering all the columns your screen should look like the following:

Please choose where you would like to store your project

Project Name *

Heart Disease

Project Type *

recognising numbers

Value 1 *
Type of value *
Age
number

Value 2 *
Type of value *
Sex
multiple-choice

Choices
F M
add a choice

Value 3 *
Type of value *
ChestPainTyp
multiple-choice

Choices
TA ATA
NAP ASY
add a choice

Value 4 *
Type of value *
RestingBP
number

Value 5 *
Type of value *
Cholesterol
number

Value 6 *
Type of value *
FastingBS
number

Value 7 *
Type of value *
RestingECG
multiple-choice

Choices
Normal ST
LVH
add a choice

Value 8 *
Type of value *
MaxHR
number

Value 9 *
Type of value *
ExerciseAngi
multiple-choice

Choices
Y N
add a choice

Value 10 *
Type of value *
Oldpeak
number

If Oldpeak can be described as numbers, choose "number".
If it can be described as choosing from a few options, choose "multiple-choice".

ADD ANOTHER VALUE

You are not allowed more than 10 values in a project.

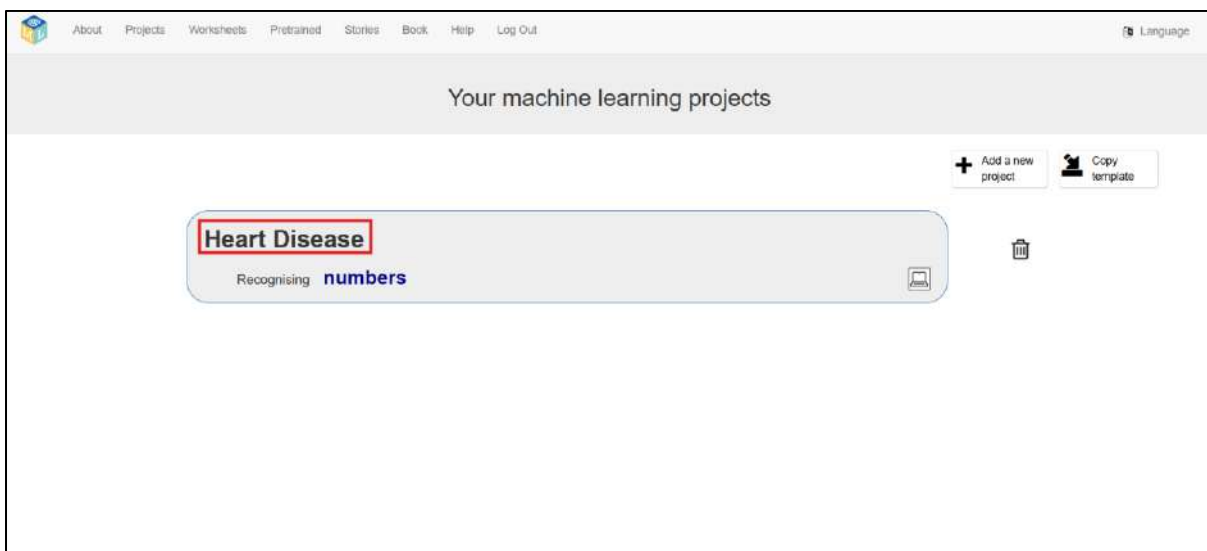
Storage *

CREATE CANCEL

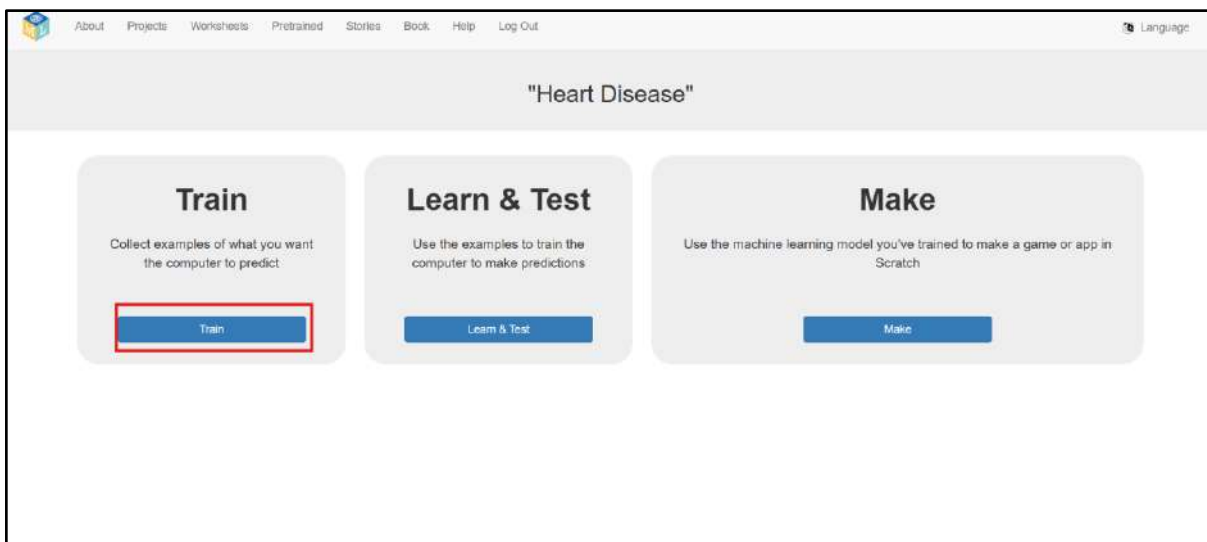
- In the "Storage" field, select:
 - "In your web browser"
- Click "CREATE"



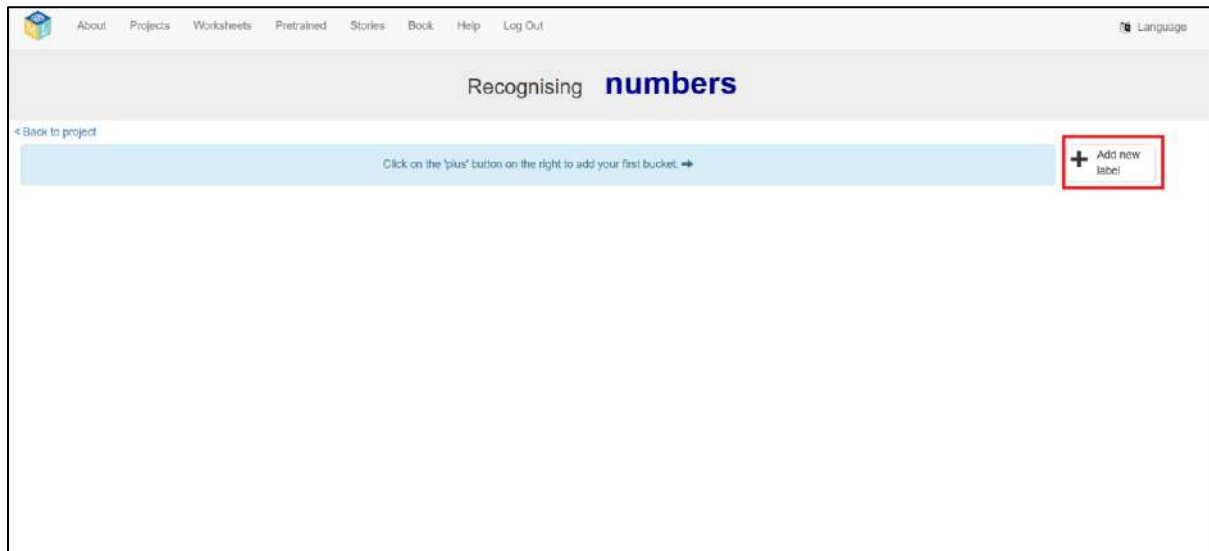
- Click on the project name



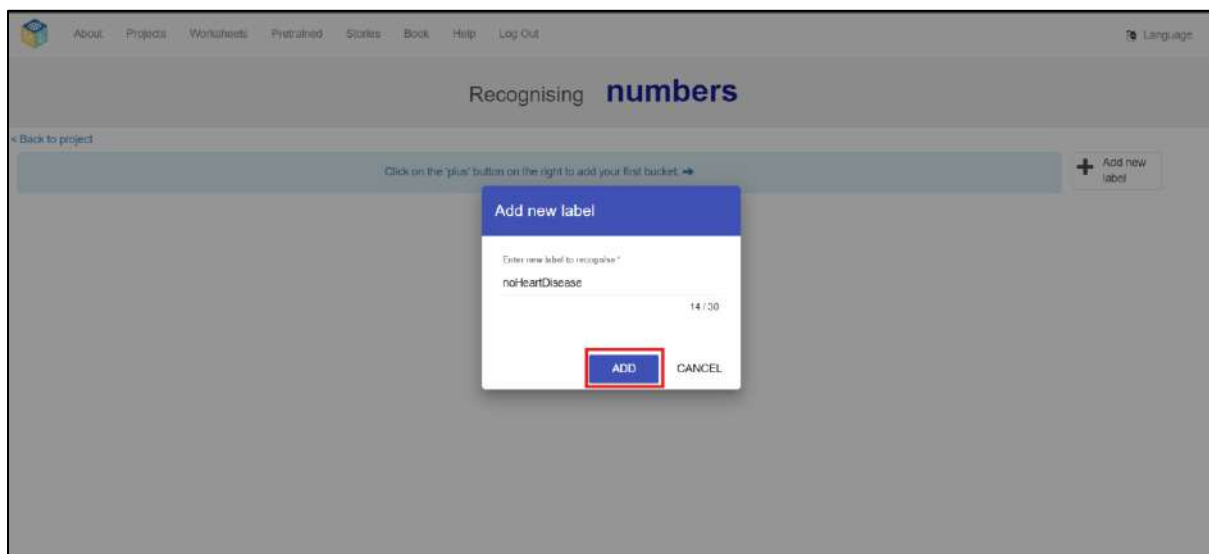
- Click "Train"



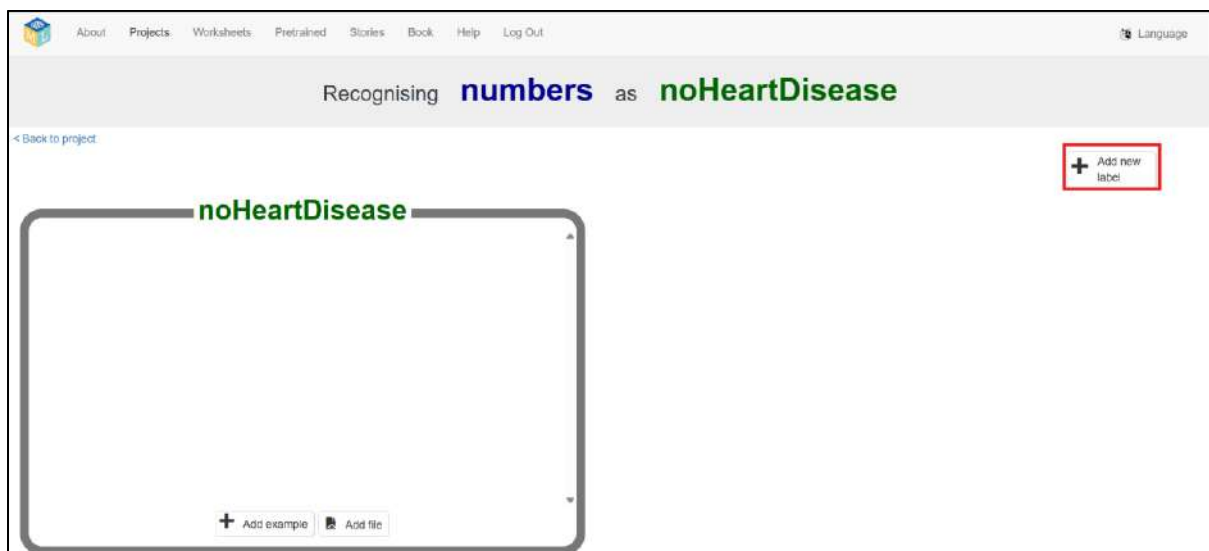
- Click on "Add new label"



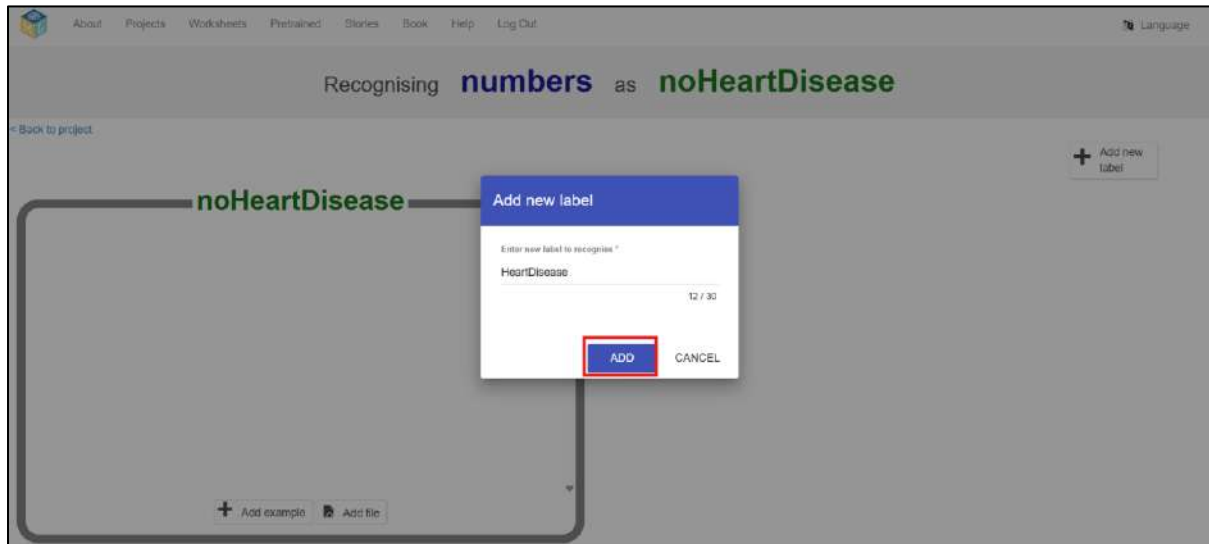
- Name it "noHeartDisease"
- Click on "ADD"



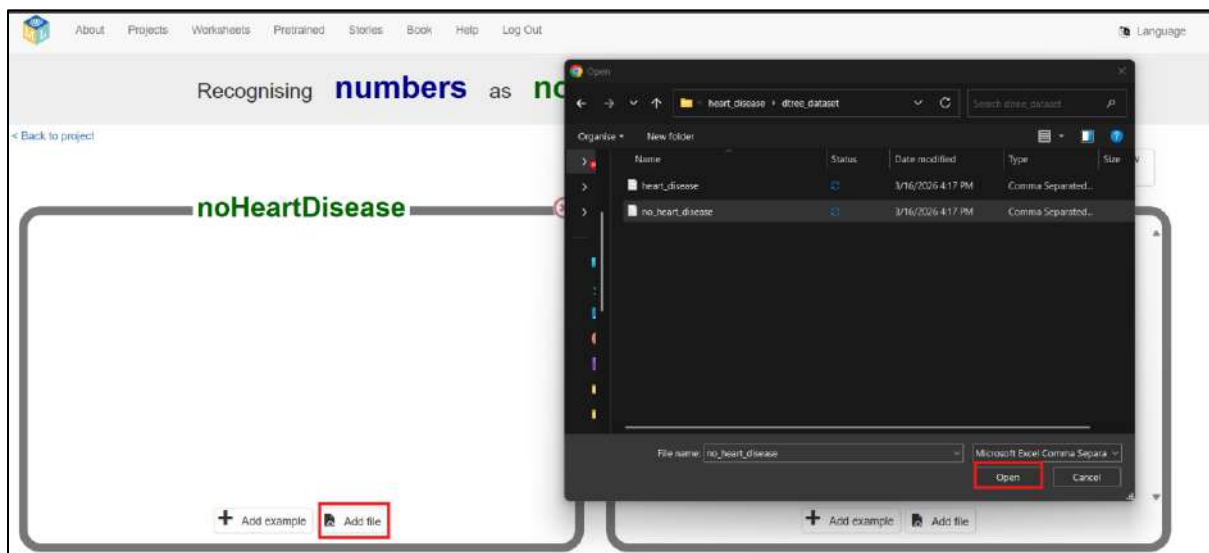
- Click "Add new label" again



- And name the second label "**HeartDisease**"
- And click on "**ADD**"



- Download the dataset files placed here: [Training Dataset](#)
- Inside the noHeartDisease tag click on "Add file"
- Select "no_heart_disease.csv"
- Click "Open"



- Do the same for the other tag, but now select the "**heart_disease.csv**" file

- After this step your screen should look like the following

Recognising **numbers** as **noHeartDisease** or **HeartDisease**

< Back to project

+ Add new label

noHeartDisease

Age: 40	Age: 37	Age: 54
Sex: M	Sex: M	Sex: M
ChestPainTyp: ATA	ChestPainTyp: ATA	ChestPainTyp: NAP
RestingBP: 140	RestingBP: 130	RestingBP: 150
Cholesterol: 289	Cholesterol: 283	Cholesterol: 195
FastingBS: 0	FastingBS: 0	FastingBS: 0
RestingECG: Normal	RestingECG: ST	RestingECG: Normal
MaxHR: 172	MaxHR: 98	MaxHR: 122
ExerciseAngi: N	ExerciseAngi: N	ExerciseAngi: N
Oldpeak: 0	Oldpeak: 0	Oldpeak: 0

Age: 39	Age: 45	Age: 54
Sex: M	Sex: F	Sex: M
ChestPainTyp: NAP	ChestPainTyp: ATA	ChestPainTyp: ATA
RestingBP: 130	RestingBP: 130	RestingBP: 140
Cholesterol: 289	Cholesterol: 283	Cholesterol: 195
FastingBS: 0	FastingBS: 0	FastingBS: 0
RestingECG: Normal	RestingECG: ST	RestingECG: Normal
MaxHR: 172	MaxHR: 98	MaxHR: 122
ExerciseAngi: N	ExerciseAngi: N	ExerciseAngi: N
Oldpeak: 0	Oldpeak: 0	Oldpeak: 0

+ Add example Add file Download

65

HeartDisease

Age: 49	Age: 48	Age: 37
Sex: F	Sex: F	Sex: M
ChestPainTyp: NAP	ChestPainTyp: ASY	ChestPainTyp: ASY
RestingBP: 160	RestingBP: 138	RestingBP: 140
Cholesterol: 180	Cholesterol: 214	Cholesterol: 207
FastingBS: 0	FastingBS: 0	FastingBS: 0
RestingECG: Normal	RestingECG: Normal	RestingECG: Normal
MaxHR: 156	MaxHR: 108	MaxHR: 130
ExerciseAngi: N	ExerciseAngi: Y	ExerciseAngi: Y
Oldpeak: 1	Oldpeak: 1.5	Oldpeak: 1.5

Age: 58	Age: 49	Age: 38
Sex: M	Sex: M	Sex: M
ChestPainTyp: ATA	ChestPainTyp: ASY	ChestPainTyp: ASY
RestingBP: 130	RestingBP: 130	RestingBP: 140
Cholesterol: 289	Cholesterol: 283	Cholesterol: 195
FastingBS: 0	FastingBS: 0	FastingBS: 0
RestingECG: Normal	RestingECG: ST	RestingECG: Normal
MaxHR: 172	MaxHR: 98	MaxHR: 122
ExerciseAngi: N	ExerciseAngi: N	ExerciseAngi: N
Oldpeak: 0	Oldpeak: 0	Oldpeak: 0

+ Add example Add file Download

35

- Click on "Back to project"

Recognising **numbers** as **noHeartDisease** or **HeartDisease**

< Back to project

+ Add new label

noHeartDisease

Age: 40	Age: 37	Age: 54
Sex: M	Sex: M	Sex: M
ChestPainTyp: ATA	ChestPainTyp: ATA	ChestPainTyp: NAP
RestingBP: 140	RestingBP: 130	RestingBP: 150
Cholesterol: 289	Cholesterol: 283	Cholesterol: 195
FastingBS: 0	FastingBS: 0	FastingBS: 0
RestingECG: Normal	RestingECG: ST	RestingECG: Normal
MaxHR: 172	MaxHR: 98	MaxHR: 122
ExerciseAngi: N	ExerciseAngi: N	ExerciseAngi: N
Oldpeak: 0	Oldpeak: 0	Oldpeak: 0

Age: 39	Age: 45	Age: 54
Sex: M	Sex: F	Sex: M
ChestPainTyp: NAP	ChestPainTyp: ATA	ChestPainTyp: ATA
RestingBP: 130	RestingBP: 130	RestingBP: 140
Cholesterol: 289	Cholesterol: 283	Cholesterol: 195
FastingBS: 0	FastingBS: 0	FastingBS: 0
RestingECG: Normal	RestingECG: ST	RestingECG: Normal
MaxHR: 172	MaxHR: 98	MaxHR: 122
ExerciseAngi: N	ExerciseAngi: N	ExerciseAngi: N
Oldpeak: 0	Oldpeak: 0	Oldpeak: 0

+ Add example Add file Download

65

HeartDisease

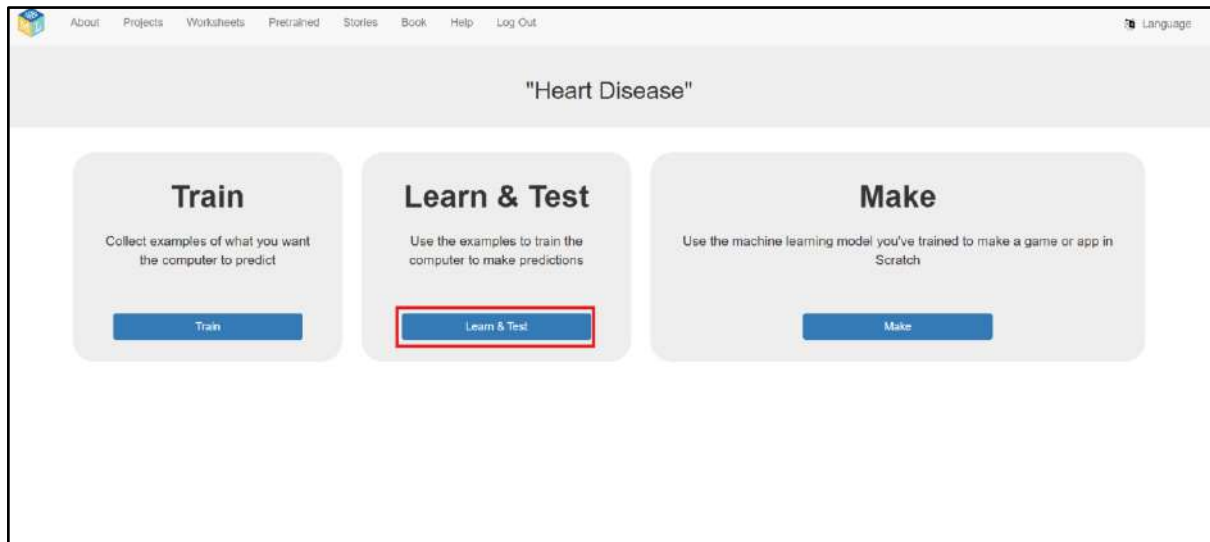
Age: 49	Age: 48	Age: 37
Sex: F	Sex: F	Sex: M
ChestPainTyp: NAP	ChestPainTyp: ASY	ChestPainTyp: ASY
RestingBP: 160	RestingBP: 138	RestingBP: 140
Cholesterol: 180	Cholesterol: 214	Cholesterol: 207
FastingBS: 0	FastingBS: 0	FastingBS: 0
RestingECG: Normal	RestingECG: Normal	RestingECG: Normal
MaxHR: 156	MaxHR: 108	MaxHR: 130
ExerciseAngi: N	ExerciseAngi: Y	ExerciseAngi: Y
Oldpeak: 1	Oldpeak: 1.5	Oldpeak: 1.5

Age: 58	Age: 49	Age: 38
Sex: M	Sex: M	Sex: M
ChestPainTyp: ATA	ChestPainTyp: ASY	ChestPainTyp: ASY
RestingBP: 130	RestingBP: 130	RestingBP: 140
Cholesterol: 289	Cholesterol: 283	Cholesterol: 195
FastingBS: 0	FastingBS: 0	FastingBS: 0
RestingECG: Normal	RestingECG: ST	RestingECG: Normal
MaxHR: 172	MaxHR: 98	MaxHR: 122
ExerciseAngi: N	ExerciseAngi: N	ExerciseAngi: N
Oldpeak: 0	Oldpeak: 0	Oldpeak: 0

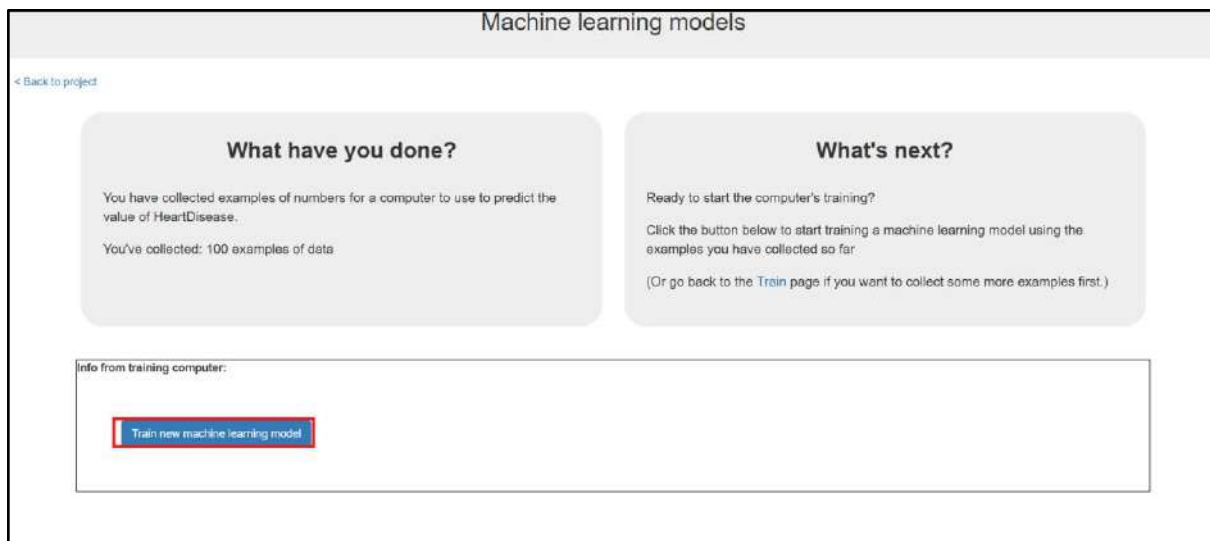
+ Add example Add file Download

35

- Click on **"Learn & Test"**



- Click on **"Train new machine learning model"**



- If your model was successfully trained, it will be available for you to try on

- If you click on "**Describe your model!**" You can see the decision tree and an explanation of it

You have collected:

- 65 examples of noHeartDisease,
- 35 examples of HeartDisease

If the computer seems to have learned to recognise things correctly, then you can go to Scratch and use what the computer has learned to make a game!

If the computer is getting too many things wrong, you might want to go back to the **Train** page and collect some more examples

Once you've done that, click on the button below to train a new machine learning model and see what difference the extra examples will make!

Try putting in some numbers to see how it is recognised based on your training.

Age
 Sex
 ChestPainTyp
 RestingBP
 Cholesterol
 FastingBS
 RestingECG
 MaxHR
 ExerciseAngi
 Oldpeak

[About](#)
[Projects](#)
[Worksheets](#)
[Pretrained](#)
[Stories](#)
[Book](#)
[Help](#)
[Log Out](#)

Language

Understand your machine learning model

[← Back to model](#)

The technique used to create your machine learning model is called a **Decision Tree Classifier**.

This is not the only way to train a machine learning model. We're using it because it's very quick and easy to train, and it is one of the easiest techniques to understand. This page shows you the decision tree that was created based on the training examples that you have collected.

When you test your model, the computer starts at the top of the tree, and follows a path until it reaches the bottom. The **class** at the bottom of the tree is the prediction that the machine learning model makes.

At each box in the tree, it reads the test described at the top of the box. If your test values pass the test described in the box, it follows the arrow to the left. If it doesn't pass the test, it follows the arrow to the right.

The **samples** shown in each box tells you how many examples in your training data matches that part of the decision tree.

The **value** shown in each box tells you how many examples in your training data passed the test shown at the top (following the left arrow) and how many examples didn't pass the test at the top (following the right arrow).

Try out your machine learning model to see how it uses the decision tree to make predictions

Age
 Sex
 ChestPainTyp
 RestingBP
 Cholesterol
 FastingBS
 RestingECG
 MaxHR
 ExerciseAngi
 Oldpeak

11

- You can also test your model. Download the file placed here [Test dataset](#). This file contains new data that has not been used to train the model.
- Add the values of the first row in the appropriate fields and click test, the value on the label will be the prediction of the model

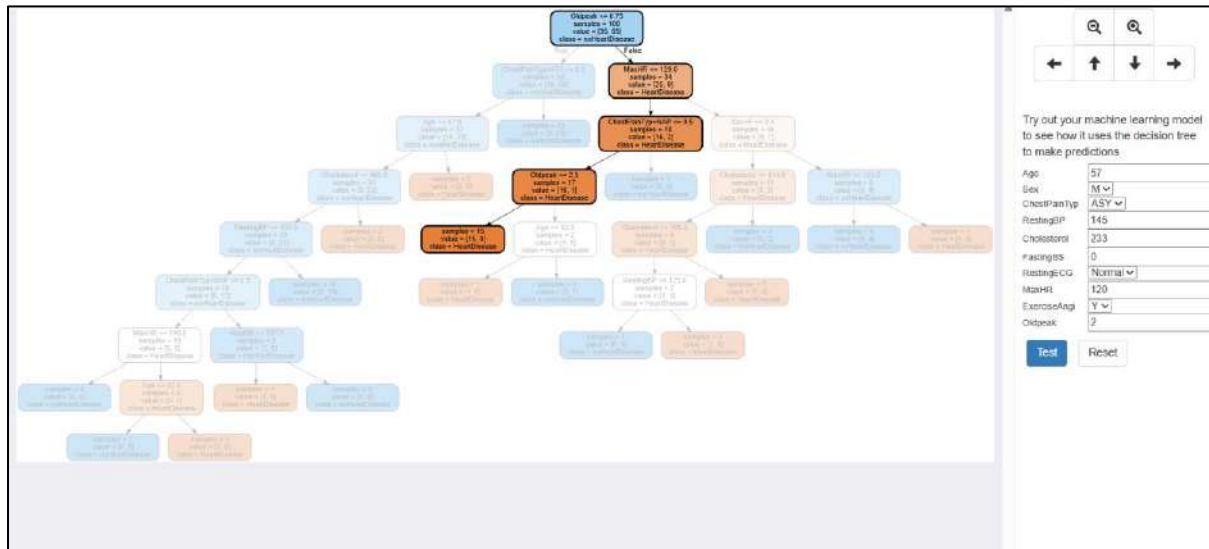
The screenshot displays a machine learning interface. On the left, a decision tree model is visualized, showing splits based on features like Age, ChestPainTyp, and RestingBP. In the center, a CSV file named 'test_data.csv' is open, showing the first six rows of data. On the right, a form allows users to input values for the features: Age (57), Sex (M), ChestPainTyp (ASY), RestingBP (145), Cholesterol (233), FastingBS (0), RestingECG (Normal), MaxHR (120), ExerciseAng (Y), and Outprik (2.0). The 'Test' button is highlighted with a red box.

Age	Sex	ChestPainTyp	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAng	Outprik
57	M	ASY	145	233	0	Normal	120	Y	2.0
42	F	ATA	120	210	0	Normal	160	N	0.0
50	M	NAP	140	250	0	ST	135	M	0.5
63	M	ASY	150	260	1	Normal	105	V	2.5
38	F	ATA	110	190	0	Normal	170	N	0.0

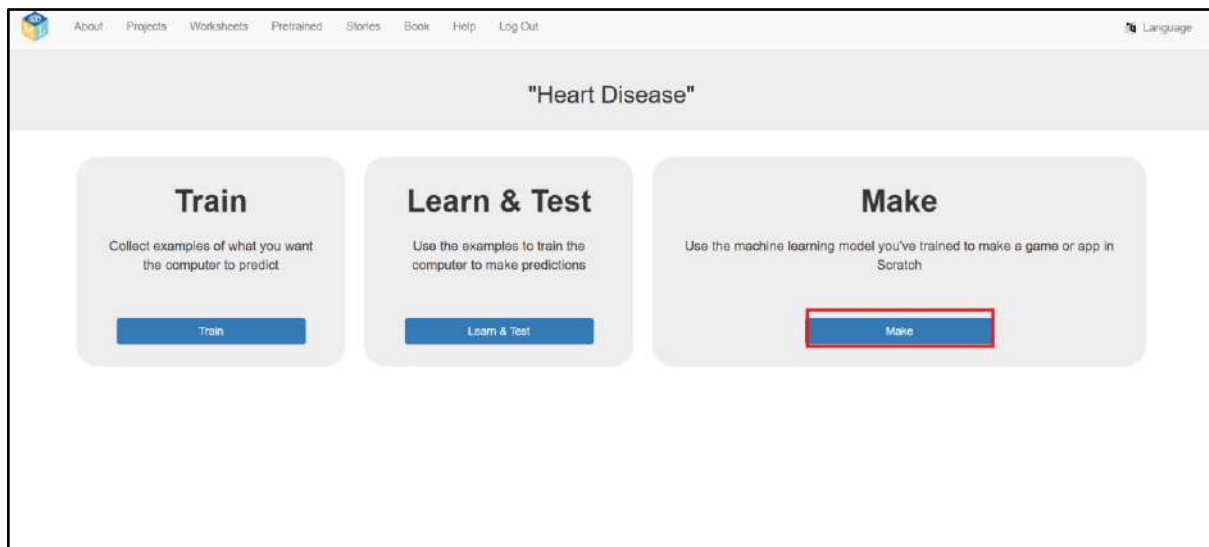
- Click on "Test"

This screenshot shows the same interface after the 'Test' button has been clicked. The decision tree model is still visible on the left. The prediction form on the right now shows the predicted value for 'Outprik' as 2.0, which is highlighted with a red box. The 'Test' button is also highlighted with a red box.

- The model correctly predicts **"heartDisease"**. As we can see in the **Test Dataset**, the label is also **"heartDisease"** for this data row

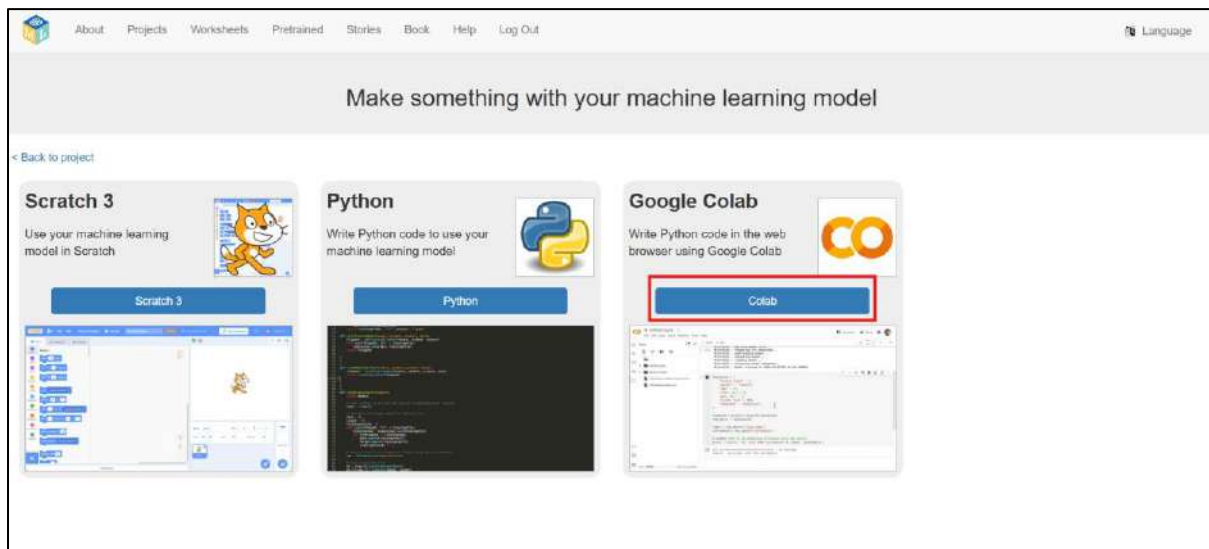


- Click on **"Back to model"**
- Click on **"Back to project"**
- Click **"Make"**



Implementation

- Click on "**Colab**"



- Follow steps 1-7 to enter the code at the end in **colab.google**
- So you can test your model through the Python programming language

