

Project Report

Fork vs Spoon Classifier — an AI image classifier for returned canteen cutlery

LRN International GCSE Artificial Intelligence (7923) — Paper 2, Component B

PRACTICE DRAFT - the accuracy figures below are placeholder values, not measured results

Candidate name	Muhammad Usama
Candidate number	2345
Centre name	_____
Project title	Fork vs Spoon Classifier
Tool used	Edge Impulse (WebAssembly export)
Date	2 September 2026

1. Explanation of the AI System and Objective

My system is an image classifier that looks at a photograph of a returned utensil and decides whether it is a fork, a spoon, or neither. I built it with Edge Impulse and deployed it as a web page: the user picks a photograph, the page resizes it to 96 x 96 pixels, and the model returns one of the three labels with a confidence score for each.

The setting is a school canteen cutlery return point, where staff currently sort every item by hand. AI is a good fit because cutlery varies in size, finish, reflection and angle, so fixed rules based on outline or brightness break as soon as an item is held differently or photographed against a different tray. Learning the shapes from examples handles that variation. The people who would use it are canteen staff, and the point of the 'neither' class is that a real return point receives knives, cups, wrappers and trays as well.

2. Step-by-Step Process

Step 1: Problem definition

The model had to take one still photograph and return one label from a fixed set, together with a confidence score between 0 and 1. There are three labels: fork, spoon and neither. The third exists because a real return point receives plenty of things that are neither. The output is a single winning label plus the confidence, which the web page shows as a bar for every class so that a close call is visible rather than hidden.

Step 2: Choosing the tool

I used **Edge Impulse**. I chose it for three reasons. First, it trains models in the browser with no coding, so all of my time went into the data and the testing rather than into setting up a Python environment. Second, it lets me upload my own dataset instead of only using its sample projects, which the syllabus expects. Third, and most importantly for this project, it exports the finished model as WebAssembly, so the classifier runs on my own web page rather than on somebody else's site.

Step 3: Collecting the data

I photographed the cutlery myself rather than downloading a dataset, so the images match the setting the model is meant for. **Source: my own photographs, taken by me on my own phone. Licence: none is needed - the images are my own work, they show cutlery and household objects only, and no person is identifiable in any of them.** There are three labels: **fork**, **spoon** and **neither**.

Step 4: Preparing the data

I removed shots that were out of focus or where the item was cut off at the edge of the frame, and made sure the backgrounds and lighting varied rather than being one desk in one light. I then split the images **80 / 20** into a training set and a separate test set. The split matters because a model scored on the images it was trained on would only be proving that it can memorise; the test set is the only honest measure.

Step 5: Training the model

Edge Impulse trained a **MobileNetV2** image model using transfer learning. Every image is resized to 96 x 96 pixels in colour, which is 9,216 pixels and 27,648 values going into the network. Transfer learning was used because the pre-trained network already recognises edges, metal texture and outline, and only the last layers had to learn my three classes. It is the only realistic option with a dataset I photographed myself.

Step 6: Testing the model

The trained model was scored on the held-back test images, which it had never seen during training. Test accuracy: **82%**. Because I took the photographs myself, the test set is the only guard against the model learning my desk rather than my cutlery.

Step 7: Improving the model

Testing showed a clear pattern of failure: version 1 was accurate on the plain backgrounds it was trained on but failed on unfamiliar ones - cutlery on wood, or held at an angle rather than laid flat. I photographed more items on varied backgrounds and at varied angles, and retrained as version 2. I then ran the same hard images through the new version and compared the confidence scores, rather than only looking at the headline accuracy.

Step 8: Deploying the model

In Edge Impulse I used **Deployment - WebAssembly - Build**, then unzipped the download. I copied `edge-impulse-standalone.js`, `edge-impulse-standalone.wasm` and `run-impulse.js` next to my own `index.html`. The page calls `Module.init()`, reads `get_properties()` to discover the label names, the input size and the number of features, then packs each pixel of the uploaded image into a single integer as (red << 16) | (green << 8) | blue before calling the classifier. Because the page reads everything from the model, nothing about the classes is hard-coded. The `.wasm` file has to be served over http with the content type `application/wasm`, so the site is hosted on a real web host (uzlabz.com) rather than opened as a local file.

3. Evaluation of Results

Test accuracy: **82%**

Test accuracy on the unseen images: **82%**. With three classes, random guessing scores about 33%, so that is the baseline to judge it against.

It handled well: a single utensil laid flat, filling most of the frame, on a background it had seen before. It found hard: unfamiliar backgrounds, strong reflections on polished metal, and items held at a steep angle so the head of the fork or the bowl of the spoon is foreshortened - which is exactly the part that carries the difference.

The confidence score tells me how strongly the image matched one class over the others, not whether the answer is right. The most useful behaviour is on the 'neither' class: when a knife is shown, a good result is a spread of low scores rather than a confident 'spoon'. A high confidence on a wrong label is the failure that would actually cause trouble at a return point, because nobody would think to check it.

4. Screenshots and Evidence

Each screenshot below is captioned with the input, the predicted label, the confidence and the time it was taken.

[paste screenshot here — a clear, easy example]

Caption: A correct result — input _____, predicted _____, confidence ____%, taken _____.

[paste screenshot here — a different class from the first]

Caption: A second correct result — input _____, predicted _____, confidence ____%, taken _____.

[paste screenshot here — one of the failures described in Step 7]

Caption: A hard or failed case, before improvement — input _____, predicted _____, confidence ____%, taken _____.

[paste screenshot here — the improvement, for direct comparison]

Caption: The same case after retraining — input _____, predicted _____, confidence ____%, taken _____.

5. Ethics and User Consideration

Privacy: inference runs locally in the browser, so no photograph of anyone's tray is uploaded or stored. In a canteen this matters more than it first appears - a camera at a return point can easily catch people as well as cutlery, and keeping the processing on the device means those frames are never collected anywhere.

Limitation: three labels only, so a knife falls into 'neither' rather than being identified. The model also judges one item per photograph, so a tray with a fork and a spoon on it gets one answer, not two.

Bias: it was trained on the cutlery I own, so a canteen with a different set - different length, finish, or plastic rather than metal - would score worse until its own photographs were added. The fix is to retrain on photographs from the actual return point before deploying it there, rather than assuming my cutlery stands for everyone's.

Note on Tools and Honesty

The dataset, the training, the testing and the judgements in this report are my own work. I used an AI assistant to help with the layout and CSS of the web page that displays the model, and to check the spelling and grammar of this document. The model itself was trained by me in Edge Impulse, and every number quoted here is one I measured.