



Echo Runner

A browser game steered entirely by an AI keyword-spotting model

Muhammad Usama · Candidate No: 2345

LRN International GCSE Artificial Intelligence (7923) — Paper 2



The Problem

- The system listens through the microphone and decides which of three commands was spoken: left, right or stop.
- Fixed rules are hard because every voice, accent and microphone produces a different signal for the same word.
- So the model learns the words from recorded examples, and a fourth 'backgroundnoise' label lets the game ignore silence and conversation.



Tool: Edge Impulse

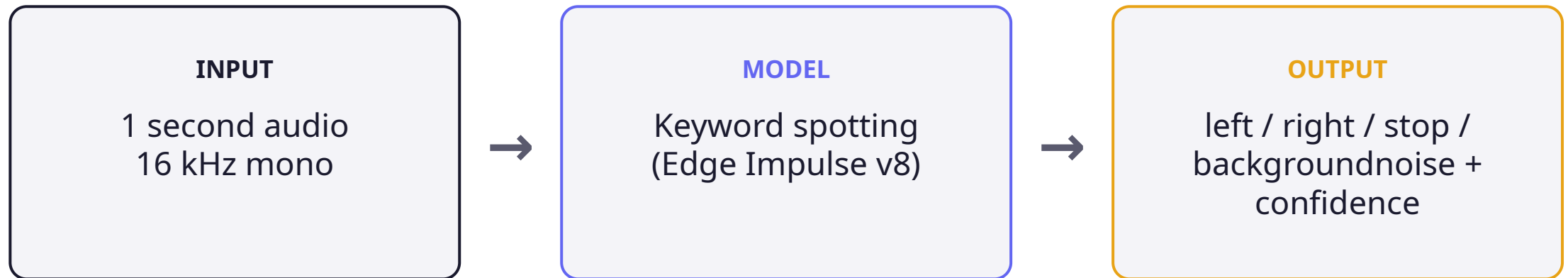
- A browser-based platform for building audio models with no coding.
- I moved to it from Teachable Machine because I needed control over the training data and a model that runs offline.
- It exports to WebAssembly, so the model loads from my own site and the audio never leaves the machine.



The Data

- Dataset: recordings I made myself, of my own voice - no third-party licence needed.
- Four labels: left, right, stop and backgroundnoise.
- Recorded at 16 kHz in one-second windows, then split 80/20 into training and test data.

How It Works



A sliding one-second window is sent to the model ten times a second, and the winning word sets the runner's movement state.



Training and Testing

- The model was trained on the recorded command samples.
- It was then tested on separate audio it had never heard.
- Test accuracy: 81%
- Measured on the v8 export: 3.4 ms per classification, 10 predictions per second, 8.07 MB model.

Improving the Model



BEFORE

- Digital silence was classified as 'left' at 99.6%
- A muted mic made the character run on its own

AFTER

- Added a silence gate and pre-filled the window with faint noise
- Retrained and exported version 8
- Start-to-first-guess fell from 1.58 s to 0.93 s

Test → find failures → add examples → retrain → compare.



Ethics and Users

- Privacy: the audio is processed in the browser and never recorded or uploaded - the buffer is overwritten continuously.
- Limitation: exact digital silence still returns 'left' at 99.6%, so the silence gate is a permanent safeguard, not a patch.
- Bias: measured, not assumed - in one session 'right' fired 39 times against 'left' 3 times, because my 'left' training data is weaker. More speakers is the real fix.



Reflection

- Hardest part: the bugs that only appeared with other people's voices and microphones, not mine.
- What I learned: measuring beat guessing every time - I assumed inference would be too slow for a phone, and it turned out to be 3.4 ms.
- Next improvement: record more speakers, especially for 'left', rather than lowering the threshold to compensate.