Magic Wand

Video link: https://youtube.com/shorts/62Gs7uI1v44

Part 1: Data Collection

Discussion: Why should you use training data collected by multiple students rather than using your own collected data only? Think about the **effectiveness** and **reliability** of your wand.

- Answer: Using training data from multiple students makes the wand more accurate and reliable. If the model is trained on only one person's data, it might only work well for that person.

Part 2: Edge Impulse Model Development

1. **Add a processing block. Read through the available options and pick one for your impulse. Justify your option.**

- Answer: I chose spectral analysis because gestures change over time, and looking at the frequency patterns helps me see the differences between them more clearly.

2. **Add a learning block. Read through the available options and pick one for your impulse. Justify your option.**

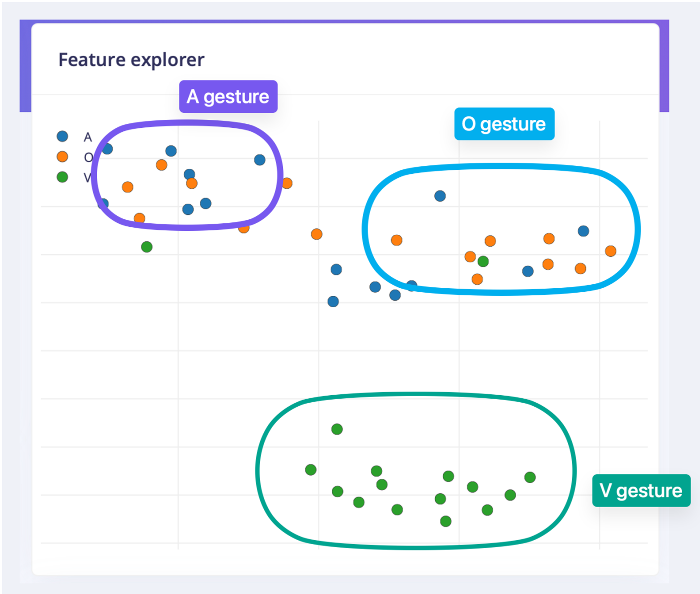
- Answer: I chose the Classification learning block because my goal is to recognize three distinct gestures: A, O, and V. Since I used spectral features as input, the classifier can detect patterns in frequency data and assign the correct gesture label more accurately.

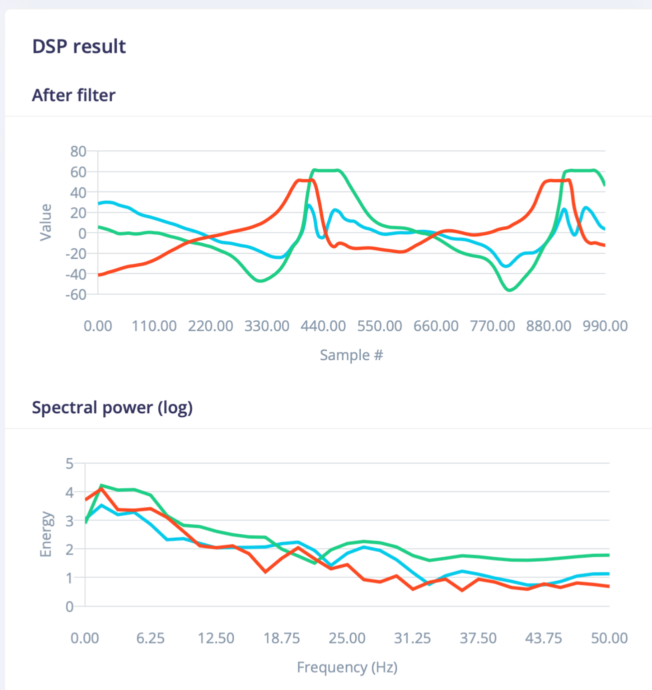
3. **Discussion: Discuss the effect of window size. Consider**
   1) **the number of samples generated**
   2) **the number of neurons in your input layer of neural network**
   3) **effectiveness when capturing slow-changing patterns**

- Answer: I chose a 1000 ms window because it generates fewer samples overall, but each one contains more information. This increases the number of input neurons, but helps capture slow-changing gestures more effectively. I think this tradeoff improves model performance for my use case.
-

4. **Choose your DSP block in the sidebar. Take a screenshot of your generated features, and sketch a rough decision boundary between classes. Explain why do you believe the generated features are good enough.**

- Answer: I think the generated features are pretty good because the V gesture forms a clear and separate cluster, making it easy to classify. A and O gestures have some overlap, but their clusters are still mostly distinct. With a proper model, they should still be separable with good accuracy.

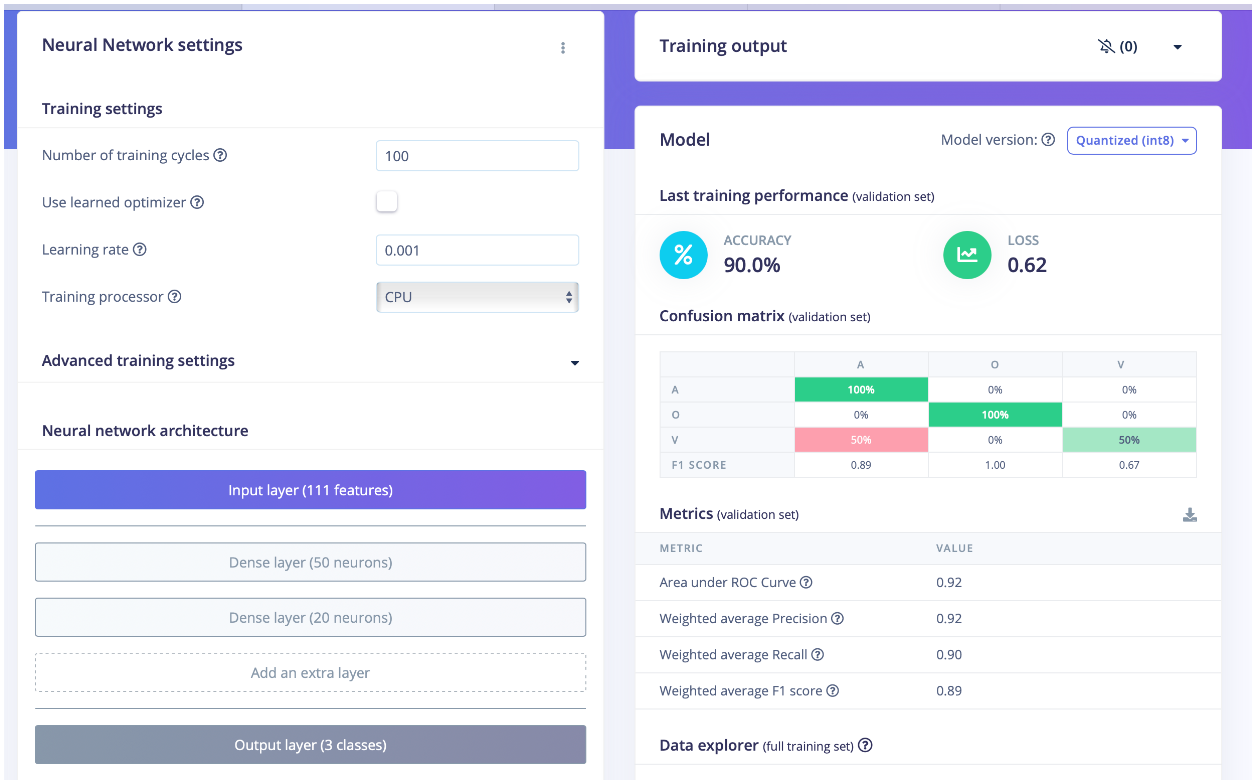




5. **Choose your ML block in the side bar. Tune the number of training epochs, learning rate, and neural network architecture until you are satisfied with the learning performance. Report the learning performance, your choices of hyper-parameters, and architecture.**

- Answer: At first, I used a smaller number of training cycles and fewer neurons in the dense layers, but the model's accuracy was low. So I kept adjusting the training cycles, learning rate, and network size to increase model capacity and improve performance. This final setup gave me a good balance between training time and accuracy, especially for classifying gestures A and O.

-



-

6. **Use "Live classification" and "Model testing" in sidebar to test your model performance. Please clearly document all metrics being used, e.g., accuracy, TP, FP, F1, etc.**

- Answer:
  - Overall Accuracy: 73.3%
  - True Positives (TP):
    - A: 9/10
    - O: 7/10
    - V: 6/10
  - False Positives (FP):
    - 2 instances where V was misclassified as A
  - F1 Scores (based on model testing):
    - A: 0.85
    - O: 0.78
    - V: 0.67
    - Weighted Average F1 Score: 0.77

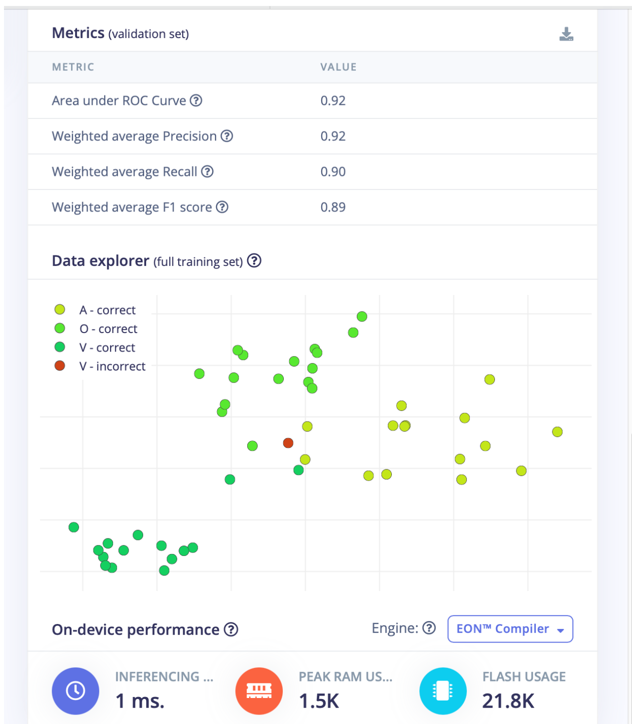
The model performs very well on gestures A and O, but still struggles with gesture V, which has lower consistency during classification. Overall, the real-time test results align closely with the training performance.

7. **Discussion:** Give at least two potential strategies to further enhance your model performance.

- Answer:

  1. Collect more training data for gesture O and A: More diverse samples can improve the model's ability to generalize. It helps reduce overfitting and improves recognition accuracy for these gestures.
  2. Add more dense layers or increase neurons: A deeper or wider network can better capture complex patterns. This may help the model distinguish between similar gestures more effectively.

8. Performance analysis and metrics



I evaluated the model using Edge Impulse's validation tools. It achieved strong results:

- AUC: 0.92
- Precision: 0.92
- Recall: 0.90
- F1 Score: 0.89

The Data Explorer shows most samples were classified correctly. One V gesture was misclassified, which matches my observation that V tends to be less consistent.
On-device performance is efficient:

- Inference time: 1 ms
- RAM usage: 1.5 KB
- Flash usage: 21.8 KB

9. **Challenges faced and solutions**

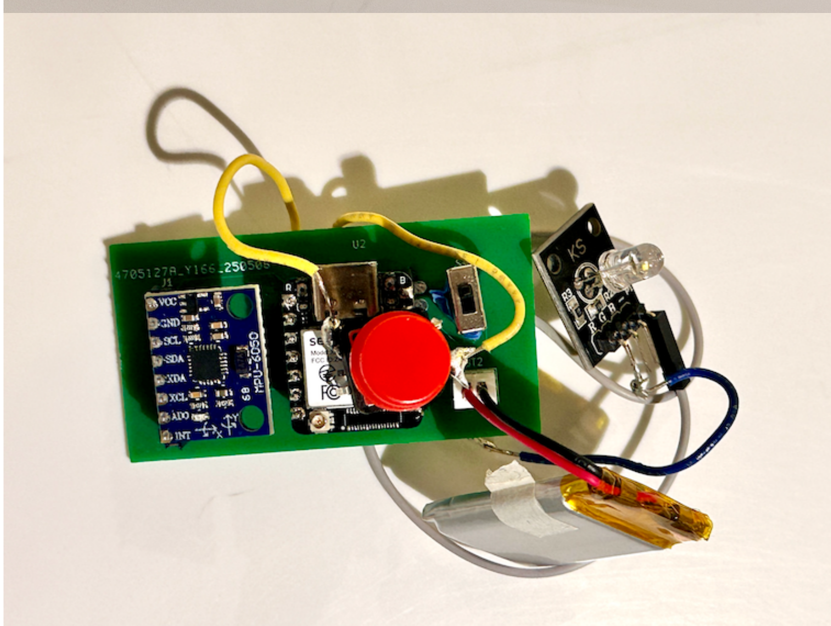
1)Inconsistent gesture V results： Gesture V was harder to classify accurately. I realized this was due to variations in how I performed it.

Solution: I collected more consistent and clearer V samples to improve model recognition.

2)Low accuracy in early training: My first few models had poor performance. The architecture was too simple.

Solution: I increased the number of training cycles and adjusted the neural network size to boost accuracy.

10. **Hardware pictures, connection, data collection and accuracy**



Hardware Setup: 3D printing

The image below shows the hardware setup of the wearable magic wand. It includes:

- MPU6050 motion sensor
- Push button for triggering gesture recording
- Switch and LED
- 3.7V battery
-ESP32S3

All components are connected and powered for real-time data collection and inference.

```
Detected: V (100.0%)
Detected: A (56.2%)
Detected: V (51.9%)
Detected: A (59.4%)
Detected: O (53.9%)
Detected: O (50.4%)
Detected: O (61.7%)
Detected: A (84.0%)
Detected: V (77.3%)
Detected: A (98.4%)
Detected: V (55.0%)
Detected: A (76.6%)
Detected: O (48.4%)
Detected: A (52.3%)
Detected: V (80.5%)
Detected: A (78.1%)
```