Assignment 5: Build and Deploy a Keyword Spotting Model using Edge Impulse

Q1) Does the model perform as accurately as expected on your smartphone? List a few methods to improve the model's accuracy.

The model's accuracy was not met as expected as the labelling was not as accurate while testing with the given data set. But the model perform some more better using the data uploaded. To improve the accuracy we can

- Collect more diverse training data with variations in accents, background noise, and pitch.
- Use data augmentation techniques to simulate real-world scenarios.
- Fine-tune the model's architecture to balance between complexity and inference time.
- Apply transfer learning or additional regularization techniques to improve generalization.

Q2) When building a model for resource-limited hardware, how do you balance fast inference times with acceptable model accuracy? What trade-offs did you encounter?

Balancing Fast Inference Times and Acceptable Accuracy:

Trade-offs:

  - Reducing the model size for faster inference may lead to lower accuracy.
  - Increasing accuracy often requires a more complex model, which may be slower and consume more power.

Approaches:

- Use model quantization (e.g., 8-bit integer quantization) to reduce computation without significantly affecting accuracy.
- Optimize the architecture to use fewer parameters, such as depthwise separable convolutions.
- Deploy models with adaptive settings to balance performance based on available resources.

Q3) Share your experience deploying the model to your smartphone and Arduino board. Mention any technical difficulties or interesting observations.

While deploying the model to smartphone I was very excited to see the results. The model worked well accurately identifying keywords. Edge Impulse significantly simplified the process of training and deploying models, especially for beginners. This exercise provided a practical understanding of how keyword-spotting systems can be implemented in everyday devices, such as voice-activated assistants.

Since I was an online participant i couldn’t deploy to Arduino Nano 33 BLE Sense.