**Machine Learning for IoT——HW2**



|  |  |  |
| --- | --- | --- |
| *Gaetano Salvatore Falco* | *Zafar Abdirasulov* | *Kuerxi Gulisidan* |
| *S280209* | *S301367* | *S304915* |

**Exercise 1 - Training & Deployment of a “Go/Stop” Classifier**

**1.1 Training & Optimization**

In order to find compliant hyper-parameters, we applied the following strategies:

We reduced the module size with an aggressive width scaling, with the parameter alpha set to 0.14

We then changed the parameters for the pre-processing of the audio, according to the table below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Downsampling\_rate** | **frame\_length\_in\_s** | **frame\_step\_in\_s** | **num\_mel\_**  **bins** | **lower\_**  **frequency** | **upper\_**  **frequency** |
| 16000 | 0.04 | 0.04 | 25 | 20 | 4000 |

Training hyper-parameters:

|  |  |  |  |
| --- | --- | --- | --- |
| **Batch\_size** | **Initial\_learning \_rate** | **end\_learning\_rate** | **epochs** |
| 32 | 0.015 | 1.e-9 | 60 |

For the model we used Convolution 2D layer with 128 filters, and we adopted the width scaling and weights pruning

And our model ends up with the following constraints met:

|  |  |  |
| --- | --- | --- |
| **Accuracy** | **Latency** | **Model size** |
| 97.5% | 4.9 ms | 24.7 kB |

**1.2 Deployment & Integration in Smart Batt**

In order to complete this laboratory, we designed a simple script as follows:

1. At the beginning of time, the system is not monitoring and it is receiveing the input stream as an audio, calling the callback function
2. The callback function checks if there is noise in the 1s received audio, if silence is detected it does nothing
3. When there is noise, we use the function “calculate\_next\_state\_FSM” to calculate if a “go” or “stop” is received, by invoking the tflite model and testing if the predicted laber for go or stop is higher than the given threshold of 95%
4. When “go” is detected, our global variable “state” goes to true and we can log the information we collect to redis
5. If there is silence or there is an unrecognised sound or when the predicted label is not higher than the threshold, the global variable “state” remains unchanged.