
Algorithm 2 OnMiniBatchUpdate_Callback.

Input: $MODEL$ as the full-precision input model.

Input: E_{size} as the target exponent bits size.

Input: M_{size} as the target mantissa bits size.

Input: D_{train} as the training data set.

Input: D_{val} as the validation data set.

Input: N_{ep} as the number of epochs.

Input: B_{size} as the mini-batch size.

Output: $MODEL$ as the quantized output model.

```
1: // Quantize
2:  $MODEL \leftarrow \text{QUANTIZETRAINING}(MODEL, E_{size}, M_{size})$ 
3: if  $1 < epoch$  then
4:     // Update model after first epoch
5:      $mse_v \leftarrow \text{Evaluate}(MODEL, D_{val})$ 
6:     if  $mse_v < mse_i$  then
7:          $\text{Update}(MODEL)$ 
8:          $mse_i \leftarrow mse_v$ 
9:     end if
10: end if
```
