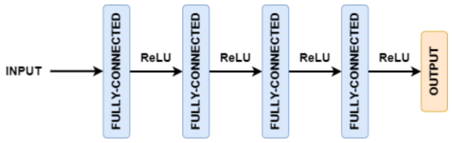
Assignment 1: Feedforward neural networks *(approximate training run-time on a CPU: 5 min)*

Implement a feedforward neural network that solves the speech classification task. This neural network must have four hidden layers with 128 neurons each. Apply a rectified linear unit (ReLU) activation function at the output of each hidden layer.



In addition, consider a minibatch size of 1,024 samples and, as the optimizer, employ stochastic gradient descent with a learning rate of 0.01, momentum of 0.9 and weight decay of 0.00001. Train your model for 50 epochs. Use “torch.manual_seed(0)” to make your results reproducible.

1) What is the size of your input layer and why? What output layer activation function and training loss function have you chosen and why?
2) Plot the training and validation losses, as well as the training and validation accuracies, as a function of the training iteration[1]. Do you think that overfitting has occurred? If your answer was yes, how may you solve it? (Just explain with words)
3) What is the test accuracy?
4) Calculate, by hand, the total number of parameters of the model, and indicate, step by step, how you reached your solution.

Hints:

- Do not forget to normalize your speech feature matrices, e.g., `X_train` ← `(X_train – np.mean(X_train)) / np.std(X_train)`
- Shuffle your training data before model training.
- Accuracy is defined as the ratio between the number of correct classifications and the total number of classifications.
- To check that you correctly calculated by hand the number of parameters of your model, you may want to use the following snip of code:

```
def get_no_params(model):
    nop = 0
    for param in list(model.parameters()):
        nn = 1
        for s in list(param.size()):
            nn = nn * s
        nop += nn
    return nop
```

[1] Note that your training data set is comprised of ⌈9,489 training samples / 1,024 samples per minibatch⌉ = 10 minibatches. Hence, your training algorithm runs for 50 epochs × 10 minibatches per epoch = 500 iterations.