

Rastiello

Assignment 08

Pham

Bartmer-Freund

Rubinic

Exercise 1

Since all activation functions are identities so that $f(x) = x$ holds for all activation functions, every hidden layer will only affect the input by multiplying with their weight. The output at the end will be proportional to the input.

Therefore, we can use a single layer instead of those hidden layers to represent the linear transformation resulting into a single-layer neural network.