

Network number	$C_{training}$	C_{valid}	C_{test}	Epochs before end
1	0.3365	0.3980	0.3925	120
2	0.2698	0.3069	0.3043	120

Figure 1: Classification error and number of epochs for each network

You can see in table 1 that the addition of complexity from network 2 compared to network 1 resulted in a significant increase of classification performance. This did not have any effect on overfitting, since both networks learned for the whole 120 epochs.