Figure 1: Classification error for each network as a graph of epochs

Network	best epoch	C validation set	C test set	C training set
1	16	0.6219	0.6304	9
2	7	0.6034	0.6048	9
3	17	0.5511	0.5559	9
4	19	0.5349	0.5356	9

Figure 2: table with the classification error

It is clear that the addition of a second layer improves the learning performance significant, while adding neurons to a layer did not show a big boost in performance for the case of a hidden layer size of 1 and 2. The classification error of the test set is slightly less than of the validation set, pointing to the possibility of overfitting. Training set shows a classification error of 9.0, pointing to a bug in the code. Here it would be expected, that the classification error is the lowest. The data shows, that added complexity does not necessarily result in improved performance. Furthermore it shows that it is easy to overfit a network and one should be vary when analyzing their results.