SVM and Neural Networks

Bjarki Sigurdsson

Abdulrahman Abdulrahim

Miguel de la Colina

Group 5

May 9, 2017

ABSTRACT

THE PURPOSE OF THE EXERCISE WAS TO GET FAMILIAR WITH NEURAL NETWORK AND SUPPORT VECTOR MACHINES AND TO GET A GRASP OF HOW THEY WORK AND HOW THEY FUNCTION WHEN USING REAL DATA-SETS IN THIS CASE THE CIPHERS.

1 NEURAL NETWORKS

For this exercise we use two-person data, reduced to the first 5 PCs and then split 50-50 to person-dependent training and test sets. We experiment with varying hidden layer structure as well as learning rates. We use the RSNNS package with `Std_Backpropagation` learning and a maximum of 1000 iterations throughout. For classification, we choose the output with the highest value in a winner-takes-all manner.

We train a neural network using the package's `mlp` function and plot it using a custom function from [1]. The resulting network is shown in figure 1. We also plot its error for each iteration as shown in figure 2. We see that this error converges with time as the network reaches a local minimum error.

---

[1] https://gist.githubusercontent.com/fawda123/7471137/raw/466c1474d0a505ff044412703516c34f1a4684a5/nnet_plot_update.r
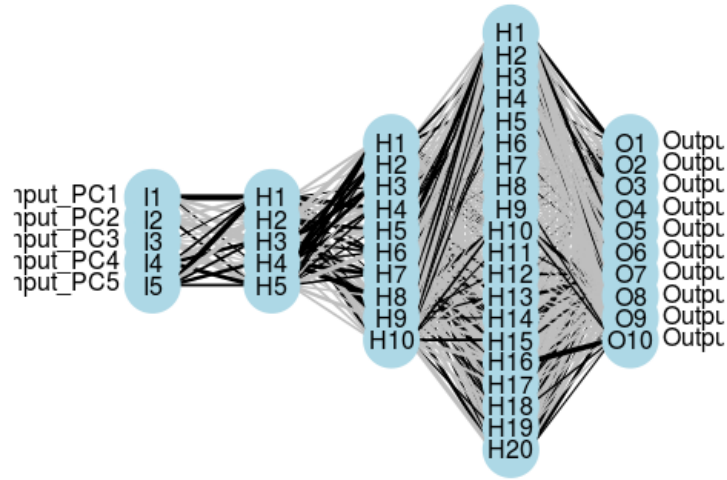
Figure 1: Neural network with hidden layer sizes c(5, 10, 20).

The results for varying the hidden layer sizes and learning rates are shown in table 1.

We see that there is a lot of freedom in the structure as various different structures yield similar results. We also see that there is a minimum number of neurons required for reasonable results. We also see that depth alone does not yield better results.

Learning rate affects the rate of gradient descent during training. A high learning rate will lead to fast convergence but may overlook local minima. A lower learning rate, by comparison, will require more iterations to converge. In our case, both learning rates result in the network converging to similar error rates.
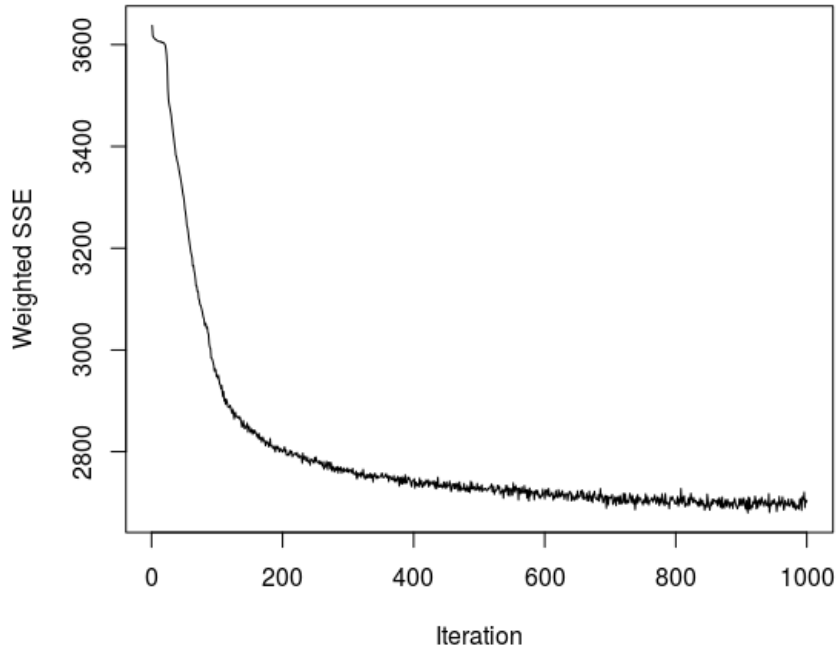
Figure 2: Error rate vs. iteration number.

Table 1: Varying structure and learning rate.

Hidden Structure	Learning Rate	Accuracy
c(1)	.2	.15325
c(5,10,20)	.2	.4255
c(10, 20)	.2	.49525
c(20, 40)	.2	.48975
c(10,10,10,10,10,10,10,10,10)	.2	.098
c(100)	.2	.4985
c(200)	.2	.4905
c(100)	.1	.49525
c(100)	.7	.4935

2 SVM