MÄLARDALENS HÖGSKOLA
ESKILSTUNA VÄSTERÅS

Assignment 1:

Samer Kakoore | ske19003
Rami Abukhader|Rar19002

1. Structure of the ANN:

The ANN consisting of three layers. The input layer with 19 input nodes, a hidden layer with 11 input nodes and the output layer with one node only. Nodes are connected to the next or previous layer nodes. Each layer beyond the output layer has a temporary node that never updates during the backpropagation. The actual classification result can be read from the node of the network output layer. If the predicted output larger or equal to 0.5 then the classification is 1 which is positive, otherwise the classification is 0 which means its negative. The initial value of the weights is between -1 and 1.

The network stops training on one of those three conditions:

1. If the error is less than 0.3
2. If the accuracy of the validation set becomes worse

2. Equations:

During the backpropagation, we calculate the gradients, the weight changes from the following formula($\Delta Wij = \eta.\delta j.\chi ij$) where η is the learning rate specifying the step size in the gradient search. And we can figure out if we should increase or decrease the weights to reduce the error function of the whole network by calculating the partial derivative $\delta_j = -\frac{\partial E_d}{\partial net_j} \cdot \chi_{ji} = \frac{\partial E_d}{\partial net_j} \cdot \frac{\partial net_j}{\partial \omega_{ji}} = \frac{\partial E_d}{\partial \omega_{ji}}$

of the error function from this formula, $E_d = \frac{1}{2} \sum_{k \in outputs} (T_k - O_k)^2$, while $-\frac{\partial E_d}{\partial net_j}$ referred to as the error term of unit j. Then we calculate the error term of the output layer from this equation $\delta_j = \left(t_j - o_j\right) o_j \left(1 - o_j\right)$ where $o_j \left(1 - o_j\right)$ is the derivative of the sigmoid function of the network and $\left(t_j - o_j\right)$ the error value of node j. (o_j) is the actual output of the network and (t_j) is the target output.

We count the error term at the hidden layer

$$\frac{\partial E_d}{\partial net_j} = o_j \left(1 - o_j\right) \cdot \sum_{k \in downstream(j)} \delta_k . \omega_{kj} .$$

3. Accuracy:

The overall accuracy of the network on the testing set is 76.88% that consists of 133 correct predictions out of 173.

4. Validation Figure:

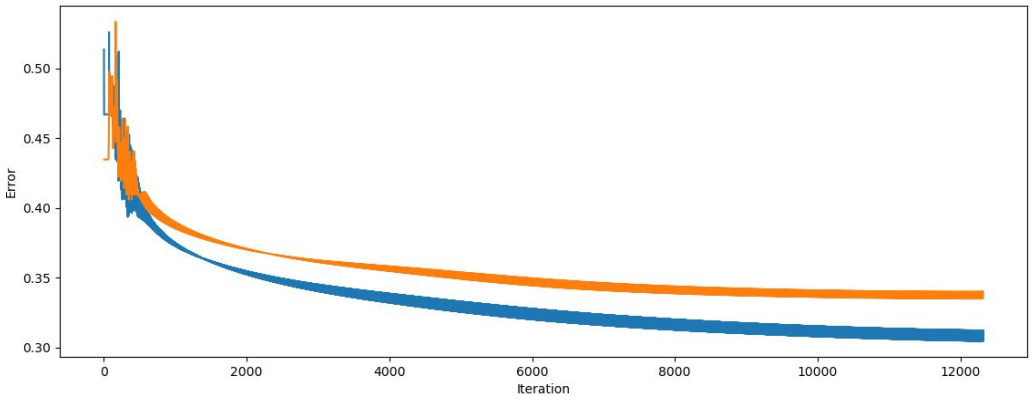

Figure 1: shows the training and validation iteration and how the error decreases.