

Cost Function or Loss Function

The squared error cost function (also known as the mean squared error) finds the sum of the squared difference between the estimated target and the actual target for a real-valued problem, while the cross-entropy cost function finds the difference between the predicted class from the probability estimates of the actual class label in a classification problem.

Regardless of the cost function used, when the error loss is small, we say that the cost is minimized. In Figure 29-3, the correct output of the example passed into the network is **2.3**. After the output values are evaluated from the feedforward training, the squared error cost function is used to assess the quality of the network’s output.

Remember that the MSE finds the average cost over all the data samples in the training dataset. In the example illustrated in Figure 29-3, we used just one data sample to demonstrate how the cost function works.

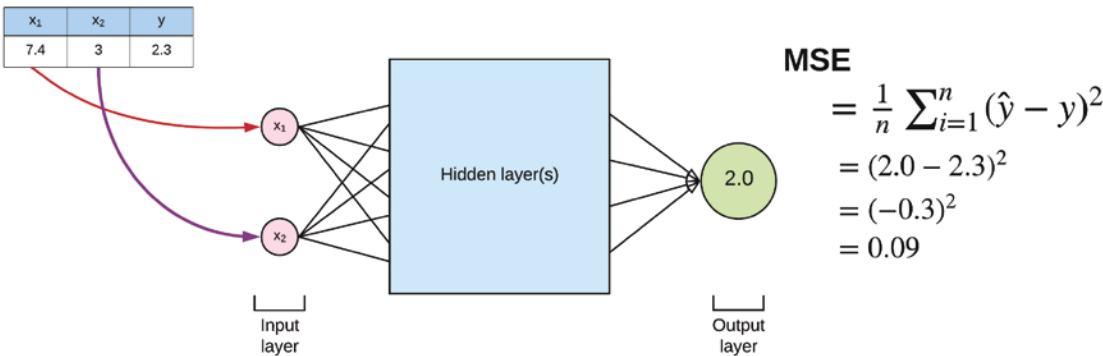


Figure 29-3. *MSE estimate of the neural network*

One-Hot Encoding

In a classification problem, one-hot encoding is the process of transforming the class labels of the target variable into a matrix of binary variables. The one-hot encoder assigns 1 when the output belongs to a particular class and 0 otherwise. An illustration of one-hot encoding is shown in Figure 29-4.