A Neural Network Approach to Estimating Material Properties

Thomas Martinetz and Thomas Poppe
Siemens AG
Corporate Research and Development
Otto-Hahn-Ring 6
81730 München, Germany

A neural network approach to the problem of estimating physical properties of a material based on the material's chemical composition is presented. The network, a multilayer perceptron, consists of sigmoidal hidden units and a linear output unit arranged in a feedforward architecture. As a component of a process optimization system which is applied in production processes with a priori unknown and eventually drifting characteristics, fast on-line adaptation of the network is performed. A first application has been the estimation of the "relative yield stress" of different steel qualities, which is necessary for optimizing the rolling process at a hot line rolling mill. On an independent test data set the neural network approach achieved a reduction of the average estimation error of about 15% compared to the current state-of-the-art method.

1. Introduction

Process optimization requires knowledge about the relevant properties of the processed material. Depending on the material transformation process to be controlled, physical properties of the material like its heat capacity, its viscosity, its heat conductivity, or its hardness (just to mention a few) determine the optimal choice for the control parameter values. In most cases, however, the respective material property cannot be measured directly but must be estimated based on the thermodynamic state of the material, i.e., its chemical composition, its temperature, the given pressure, and eventually geometric quantities. The quality of the estimation result determines to a great extent the cost effectiveness and the product quality of the production process.

To be able to estimate material properties based on the thermodynamic state variables, the respective physical relationship has to be known. A common approach is to try to describe this relationship through physical models. However, in most cases the underlying physics is too intricate and/or not understood sufficiently to allow the design of feasible physical models which yield satisfying estimation results. In addition, the development of physical models is time consuming, requires precise knowledge about the usually very complex physical processes, and each model is specific for each material and each material transformation process.

To increase cost effectiveness and product quality also of intricate material transformation processes, an approach is necessary which *learns* the underlying physical relationship instead of modeling it based on specific prior knowledge. In addition, it would be highly desirable to have an approach which is generic and can be applied to a variety of materials and transformation processes. In the following we demonstrate that neural networks as adaptive modeling schemes have the desired capabilities. We describe the application of a neural network to the problem of estimating the *relative yield stress* (plasticity) of steel plates based on the steel plates' chemical composition, temperature, and shape. Knowledge about the relative yield stress is necessary for optimizing rolling processes, in our case the rolling of steel at a hot line rolling mill.

2. The Neural Network Architecture

The neural network has to model the relation

$$\alpha = F(C, Si, Mn, P, S, Al, N, Cu, Cr, Ni, Sn, V, Mo, Ti, Nb, B, d, b, T_i, T_f)$$

between the relative yield stress α of the steel plate and the concentrations of the sixteen chemical additives $C, Si, ..., B$, the steel plate's thickness d and its width b. T_i and T_f denote the temperature of the steel plate before and after the rolling, respectively. These two temperatures serve as a measure for the actual rolling temperature T, which cannot be determined explicitly. The concentration of the sixteen chemical additives $C, Si, ..., B$ is obtained from a material analysis during the steel cooking.

Figure 1 shows the neural network architecture, a three-layer feedforward network consisting of ten sigmoidal hidden units and one linear output unit. Each hidden unit receives the same twenty-dimensional input vector $\mathbf{x} = (C, Si, ..., Nb, B, d, b, T_i, T_f)$. The weights of the hidden units i, $i = 1, ..., 10$, are denoted by $\mathbf{w}_i = (w_{i1}, ..., w_{i20})$, and the weights of the linear output unit are denoted by $\mathbf{w} = (w_1, ..., w_{10})$. The thresholds of the hidden units and the output unit are denoted by θ_i and θ, respectively. Hence, when the network receives the input $\mathbf{x}$ which carries the information about the steel plate to be rolled, the network generates the output

$$\mathcal{N}_{\mathcal{W}}(\mathbf{x}) = -\theta + \sum_{i=1}^{10} w_i \, \sigma\left(\sum_{j=1}^{20} w_{ij} x_j - \theta_i\right)$$

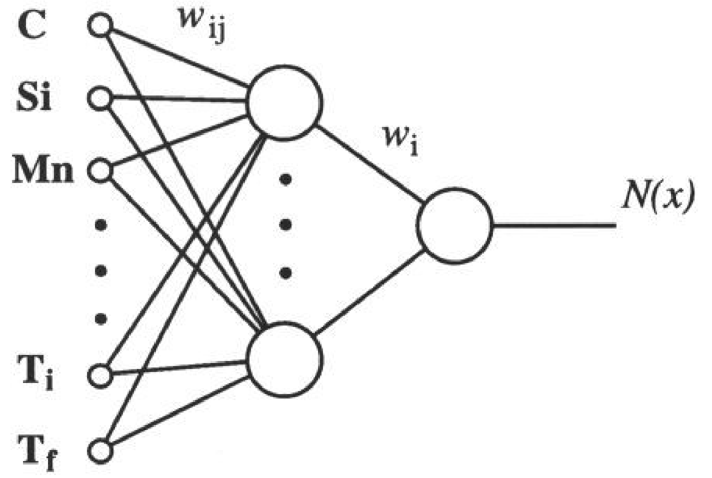

Figure 1: The architecture of the neural network. The network receives 20 inputs and consists of 10 sigmoidal hidden units plus one linear output unit.

as an estimation for the relative yield stress of the steel plate, with $\sigma(.) = 1/(\exp(-.) + 1)$ forming the sigmoidal output of the hidden units. The index $\mathcal{W}$ denotes the dependence of the network output $\mathcal{N}_{\mathcal{W}}(\mathbf{x})$ on the set $\mathcal{W} = (\mathbf{w}_i, \theta_i, \mathbf{w}, \theta)$ of all network weights and thresholds.

The estimation error of the network has to be minimized by adapting the network weights $\mathcal{W} = (\mathbf{w}_i, \theta_i, \mathbf{w}, \theta)$. This is achieved through pattern-by-pattern training, i.e., with each pattern μ through gradient descent on the square error

$$E(\mathcal{W}) = (\alpha^\mu - \mathcal{N}_{\mathcal{W}}(\mathbf{x}^\mu))^2 .$$

$\mathbf{x}^\mu$ comprises the chemical composition, thickness, width, and temperature of the μ-th steel plate, the actual relative yield stress of which was α^μ. With each new data pairs $(\mathbf{x}^\mu, \alpha^\mu)$ the network weights are adjusted through gradient descent on $E(\mathcal{W})$, which yields, by calculating

$$\Delta\mathcal{W} = -\eta \frac{\partial E(\mathcal{W})}{\partial \mathcal{W}}, \tag{1}$$

the backpropagation learning rules [1, 2].

3. The Performance

For testing the performance of the neural network approach and comparing it with the current state-of-the art method, 38442 measured data pairs $(\mathbf{x}^\mu, \alpha^\mu)$ from a rolling mill were made available by the steel manufacturer. The data pairs were ordered chronologically, corresponding to the order the steel plates were rolled. The first 10000 data pairs formed the training set which was used for a preadaptation of the network. The following 28442 data pairs were used for on-line testing and training.

$\langle E_{cur} \rangle$	$\langle E_{net} \rangle$	Δ
39.57%	33.70%	14.9%

Table 1: The relative RMS error of the neural network and the current method.

The on-line performance of the network was tested by sequentially presenting data pairs $(\mathbf{x}^\mu, \alpha^\mu)$ from the test set, in the same chronological order as the steel plates were rolled. With each steel plate the estimation error of the neural network and the current-state-of-the-art method, respectively, was recorded, and an adaptation step of the network weights was performed. Then the next data pair was presented, etc.. After 28442 data pairs the average estimation error of the neural network and the current-state-of-the-art method on these 28442 steel plates was calculated. The test was performed in the laboratory, however, the result is equivalent to the average estimation error the neural network would have achieved if it had really been applied at the rolling mill.

The achieved estimation performance is shown in Table 1. $\langle E_{net} \rangle$ denotes the root mean square (RMS) estimation error of the neural network on the data of the test set, relative to the standard deviation of the test data. $\langle E_{cur} \rangle$ denotes the relative RMS estimation error of the current state-of-the-art method on the test set, and Δ is the achieved improvement. The neural network approach achieves an improvement of 14.9% over the current state-of-the-art method.

4. Discussion

The results obtained with the straight-forward neural network approach are very promising. In the application described, the estimation of the relative yield stress of steel, the improvement of the estimation quality is so significant that the neural network approach will replace the current method and soon be a component of a commercially available process optimization system for rolling mills.

There are a couple of reasons for the favorable results with the neural network approach. The main reason is the on-line adaptation of the network. The network weights are permanently adjusted to the changing characteristics of the rolling mill and the drifts of the measuring devices for the chemical composition, thickness, width and temperature of the steel plate. Particularly the calibration of the measuring devices is not very reliable because of the very hazardous environment at a hot line rolling mill. The presented approach based on a neural network is able to compensate for these drifts due to its adaptability.

References

[1] Werbos P (1974) "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences." Ph.D. thesis, Harvard Univ. Committee on Applied Mathematics.

[2] Rumelhart DE, Hinton GE, Williams RJ (1986) "Learning Representations by Back-Propagating Errors." Nature, 323:533–536.