The use of a neural network for determining the temperature of a vapor film destruction in uncooled and saturated water-ethanol mixture in a cylindrical geometry

M M Vinogradov¹ and A R Zabirov ¹,²
¹National Research University “Moscow Power Engineering Institute”,
14 Krasnokazarmennaya Street, 111250, Moscow, Russia.
²Joint Institute for High Temperatures,
Bld.2 Izhorskaya, 13, 125412 Moscow, Russia.
E-mail: Demon_D@inbox.ru, Zabirov.Arslan@gmail.com

Abstract: The article presents the construction of an artificial neural network to determine the temperature of destruction of a vapor film in subcooled and saturated liquids of water-ethanol mixtures. To train the neural network, the results obtained on cylindrical samples of stainless steel, copper and nickel are used. In total, about 260 experimental points were used, which is sufficient to build a specific computational model. This article discusses a model of a neural network of the multilayer perceptron type. The trained neural network model shows a greater generalizing ability than the theoretical model for determining the temperature of vapor film destruction.

1. Introduction

Fairly long ago it was noticed that if the material is heated enough, and then quickly cooled, the strength of the product increases. The described process is called metal hardening. This process is especially important for carbon steels, because when the carbon steel is cooled so quickly, austenite (anisotropic modification of carbon steel) does not have time to form, but instead martensite (anisotropic modification) is formed. The main difference between austenite and martensite is that martensite has a higher strength than austenite. The quality of carbon steel obtained as a result of hardening directly depends on the intensity of heat transfer.

The highest heat transfer rates can be obtained in the nucleate boiling regime. Therefore, liquid cooling of metals is used for quenching. But when a high-temperature body is cooled in a liquid, a protective vapor layer is formed around it. Naturally, the formation of such a protective heat-insulating layer greatly reduces heat transfer. Typical heat transfer coefficient (HTC) values for saturated liquid film boiling are about 500 W/m²K. This HTC of the film boiling regime is two orders of magnitude lower than that of the nucleate boiling regime.

In 1986, Kenning and Hewitt found that film boiling of a subcooled liquid not only increases the HTC, but also changes the temperature at which the vapor film breaks down. Moreover, such a shift in the temperature of destruction of the vapor film can exceed the critical temperature of the liquid, when direct contact of the liquid with the wall is impossible.

Since 2012, a large number of different articles have dealt with the topic of neural networks. This direction appeared in 1963, but once came to an end due to a lack of computing power. Nowadays, computer technologies have become many times more powerful, so interest in this area has
reappeared. The direction of neural networks is an attempt to simulate the functioning of the brain of a person or any other intelligent creature. The work is based on the analogy of signal transmission by nerve cells through dendrites and axons. In this analogy, all biological aspects have been removed to the good of the general operating principle. The signals entering through the dendrites are amplified or weakened, entering the body of the neuron, added with all these incoming signals. If this total signal exceeds a certain value, then the body of the neuron is excited and sends a signal through the axon, which transmits this signal to other dendrites and other bodies of the neuron.

A neural network is a type of machine learning method. Its main advantage is the construction of a kind of mathematical model that can have a high accuracy in determining values. The obtained accuracy of determining the neural network model can exceed the accuracy of the theoretical determination. Because of the latter aspect, scientists in various fields are trying to use machine learning methods to get more accurate results than a physical model gives. An example of this is the submitted articles [1–4].

The review article [1] analyzes the use of a neural network in thermophysical and hydraulic problems. Examples of such problems are the determination of the critical heat flux, the heat transfer coefficient and the hydrodynamic in a multiphase flow. As a result of this analysis, it is shown that the resulting neural network models can be used to determine all of the abovementioned values, but they should be used with caution.

The authors of articles [2] and [3] use machine learning methods to study the growth of vapor bubbles. The resulting model allows determining the centers of bubbles and boundaries in the picture, which allows studying the dynamics of changes in bubbles. The resulting model is highly accurate and will greatly simplify the processing of experimental results.

In article [4], the authors use neural networks to increase the speed of solving the SPACE algorithm. To do this, the authors solve some problems using the SPACE algorithm, collecting a dataset for training and testing a neural network. The result of the trained model is about 10 times increase in the speed of the SPACE algorithm while achieving the same accuracy as when using this algorithm.

The authors of [5] used neural networks to determine the temperature of destruction of the vapor film on unsteady cooling of subcooled water temperature of the vapor film destruction at unsteady cooling of subcooled water. To train the neural network, the authors collected data from various articles, which indicate the experimental values of the destruction of the vapor film. In conclusion, the authors noted that the error of the resulting neural network model is about 6%.

The accuracy obtained in article [5] looks too good to be true. The reason for this is that the error in determining the temperature of destruction of the vapor film, like many other thermodynamic quantities associated with boiling, is about 30%. It may be assumed that the final result of the trained neural network model may give an error of about 6%. To test this result, we will build our own neural network model. To construct the model, the results obtained on cylindrical samples of stainless steel, copper and nickel were used. In total, there were about 260 experimental points, which was enough to build a certain model of a neural network.

When comparing the results of the trained neural network and the theoretical model, we use the equation, which is described in details in the following articles [6–9]:

\[ K = 4.5 \cdot 10^{-4} \cdot \frac{\Delta h_{v/ga}}{\nu \Delta T_{w/\rho c \lambda} L G} \left( \frac{\rho}{\nu} \right)^{1/4} \left( \frac{\nu}{g \rho} \right)^{1/2} \left( \frac{Pr^7 \nu}{G r^2} \right)^{1/6} \left( \frac{\Delta T_{w/sub}}{\Delta T_{w/0}} \right)^{0.5} \]  

In (1) the following parameters are used: \( \sigma \) is the surface tension; \( \nu \) is the kinematic viscosity; \( \Delta h_{v/ga} \) is the latent heat of evaporation; \( T_0 \) is the spinodal temperature; \( \Delta T_0 \) is the temperature difference \( (T_w - T_0) \); \( T_w \) is the wall temperature; \( \Delta T_{w/sub} \) is the liquid subcooling; \( \rho \) is the liquid density; \( \rho_p \) is the solid density; \( c_w \) is the heat capacity of a solid; \( \lambda_w \) is the thermal conductivity of a solid; \( g \) is the acceleration of gravity; \( D \) is the linear scale of the cooled body; \( Pr \) is the Prandtl number; and \( Gr \) is the Grashof number.
The error in determining the formula (1), as indicated earlier, is 30%. This high error is caused by poor repeatability of the results. For example, when a stainless-steel cylinder is cooled with pure ethanol at a liquid temperature of -10 °C, the vapor film destruction temperature changes from 202 to 245°C.

2. Neural network architecture

One of the simplest neural network architectures is the multilayer perceptron, which is shown in Fig. 1. The principle of operation of a neural network is as follows: input signals (x) are multiplied by the corresponding weights (w) and fed to the body of the neuron. In the body of the neuron itself, all received signals are added (z) and transferred to the activation function (f), and the signal of this activation function is fed to the next level of the neural network. Mathematically this is written as equation (2):

\[ y_i = f\left(\sum_{j=1}^{n} x_j w_{ij} + b_i\right) = f(z) \]  

(2)

Where \( x \) is the input to the neural network; \( w \) is the entry weight; \( b \) is the displacement neuron; \( z \) is the result of summation; and \( f \) is the activation function.

![Figure 1. Standard neural network view.](image)

The main task of any neural network is to adjust its weights, which requires training. Learning will take place with a teacher, and the difference of the squares is used as an indicator of the error (J).

\[ J = (T - Y)^2 \]  

(3)

Where \( Y \) is the value of the output of the neural network; and \( T \) is the expected result. The task of learning is to reduce the cost function. To do this, it is necessary to determine how to change the weights of the neural network, and this can be achieved by taking the derivative with respect to the weights. This process is called gradient descent assignment. Conventional gradient descent has many disadvantages, so we will use a modification of Adam's gradient descent [9]. There is also a theory that allows more successfully and quickly training a neural network model, taking certain initial weights, which is indicated in the next article [10]. Also, in the process of training the neural network, retraining occurs - this is when training tries to describe the data of the training sample too accurately. To prevent this process, regularization and normalization of the input parameters are used.

There are 2 types of regularization: “L1” and “L2”. The basic principle of regularization is to decrease the values of the weights. L1 regularization is used when it is necessary to remove unnecessary connections in the neural network, when many different types of input data are used, which are not known whether they are important for the output value. L2 regularization is used when all input metrics are associated with output metrics.

Normalization of the input values of the neural network occurs across all layers of the neural network. Normalization is performed for each input separately according to the following formula.

\[ X'_{i,k} = \frac{x_{i,k} - \bar{x}_i}{\sigma_i + \varepsilon} \]  

(4)

\[ \bar{x}_i = \frac{1}{m} \sum_{k=1}^{m} x_{i,k} \]  

(5)
\[ \sigma_i^2 = \frac{1}{m-1} \sum_{k=1}^{m} (X_{i,k} - \bar{X}_i)^2 \] (6)

Where \( k \) is the number of the example; \( i \) is the input number; and \( \varepsilon \) is the an infinitely small value \((10^{-8})\). The main advantage of normalization is the ability to quickly accelerate gradient descent and get rid of the dimension of the input quantity.

3. Neural network result

In this work, the following input parameters are used: liquid subcooling, Prandtl number, Grashof number, thermal activity of the material and complex \( \Delta h_{LG} \sigma \)/\( v \). Hidden layers use boolean activation functions, and the output layer uses linear function. Two hidden layers were used, with the number of neurons in each equal to 5. After training, the neural network acquired the following form, shown in Figure 2.

![Figure 2. Neural network architecture.](image)

The neural network training process is shown in Figure 3. This figure shows that the learning process takes about 22,500 iterations. In this case, the standard deviation of the test sample is about 2100, and on the training set – 1500.

![Figure 3. Learning process.](image)

Comparison of the results of the neural network and experimental values is shown in Figure 4. From Figure 4, it may be seen that the error of such a trained neural network is 30%. It can also be seen from this figure that some points fall outside the 30% error for the training and testing set.
Let us compare the obtained neural network model with the calculated formula (1) when cooling a stainless-steel cylinder in ethanol. The comparison of the results is shown in Figure 5. The trained neural network in this example shows a more accurate result than the theoretical model of equation (1).

**Figure 4.** Comparison of neural network results and experimental values. The dotted line shows an error of 30%.

**Figure 5.** Comparison of experimental results and theoretical equation (1).

**Conclusions**
The obtained results of the trained neural network show a more accurate result than the theoretical model of equation (1). Note that on the training and test set, the trained neural network model has an error of about 30%. The reason for this error is the poor repeatability of the results of experiments in the same mode on different samples.

Based on the results of constructing various architectures of neural networks, it has been determined that for a neural network of the multilayer perceptron type, it is sufficient to use 1 or 2 hidden layers. Among the many activation functions, the logic function and the activation function for hidden layers show the best result.
Acknowledgements
The work was carried out at the Department of Engineering Thermophysics of the National Research University "MPEI" with financial support from the Russian Science Foundation (Grant No. 20-79-10363).

References
[1] Shevchenko S A, Shevchenko R A et al. 2020 Artificial neural networks in thermophysical problems of nuclear power: Issues of justification of applicability (Moscow)
[2] Seredkin A, Plokhikh I et al. 2020 Science and Artificial Intelligence conference 5–8 doi: 10.1109/S.A.I.ence50533.2020.9303175
[3] Serdyukov V, Malakhov I and Surtaev F 2020 Journal of Visualization doi:10.1007/s12650-020-00660-z
[4] Hae M P, Jong H L and Kyung D K 2020 Annals of Nuclear Energy 141
[5] Ammar M B and Shikha A E 2020 International Journal of Heat and Mass Transfer 150
[6] Yagov V V, Zabirov A R, Kanin P K and Denisov M A 2017 Journal of Engineering Physics and Thermophysics 2(90) 266–75 doi:10.1007/s10891-017-1564-3
[7] Dedov A V, Zabirov A R, Sliva A P, Fedorovich S D and Yagov V V 2019 High Temperature 1(57) 72–82
[8] Zabirov A R, Kanin P K, Vinogradov M M and Sharafutdinov A M. 2018 Journal of Physics: Conference Series 1128
[9] Zabirov F, Kanin P, Vinogradov M, Sliva A, Dedov A, Fedorovich S and Yagov V 2018 Materials today: Proceedings 26171–8
[10] Diederik P K and Jimmy L B 2015 Adam: A Method for Stochastic Optimization in ICLR
[11] He K, Zhang X, Ren S and Sun J 2015 Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification