Application of the Artificial Neural Networks for the Prediction of Reactivity of Molecules in Radical Reactions

TUMANOV V.E., GAIFULLIN B.N.
Department of Computing and Informational Resources
Institute of Problems of Chemical Physics RAS
142432, Chernogolovka, Semenov ave, 1
RUSSIAN FEDERATION
tve@icp.ac.ru

*Abstract: In article an application of a feed-forward artificial neural network (ANN)***,** *trained on experimental sample, for an estimation of reactivity of organic molecules in radical reactions is considered. Results of training and a prediction of the network are discussed.*

Key Words: feed-forward artificial neural network, subject-oriented science intelligence system, rate coefficients, reactivity of organic molecules.

1 Introduction

Currently artificial neural network (ANN) is widely used in solving applied problems of automated processing of scientific data. The main fields of ANN application in chemical and biochemical studies are given in a review [1]. Most works in this area are devoted to the correlation between the structure of chemical compounds and the physicochemical properties or biological activity they showed. In the physical chemistry, the main directions of ANN application are the simulation of chemical processes and the simulation of the dynamic properties of the molecules and the systems.

On the one hand, the physical chemistry of radical reactions accumulated large amount of experimental data on the reactivity (specific reaction rate or activation energies) of molecules in radical reactions [2, 3]. On the other hand, the experiments to quantify the reactivity of molecules in radical reactions are an expensive and time-consuming task. Carrying out the quantum chemical calculations is time consuming, and the resulting data for these calculations are not sufficiently reliable. Therefore, the development of ANN based on existing experimental data to predict the reactivity of organic molecules in radical reactions is the vital task.

Knowledge of the reactivity of organic molecules in the radical reactions is necessary for the development of new organic materials, the design of new drugs, design of technological processes, planning and conducting a scientific experiment, the training of students and graduate students.

This paper discusses the use of feed-forward artificial neural network to predict the reactivity of organic molecules in the bimolecular radical reactions in the liquid phase.

2 Problem Formulation

Experimentally, the activation energy (E) or a classical potential barrier (E_e) determines the reactivity of organic molecules in a radical reaction:

$$E_e = E - 0.5(hLv_i - RT), \qquad (1)$$

v_i is a frequency of the stretching vibrations for the bond being broken, R is the gas constant, h is the Planck constant, L is the Avogadro number, and T is the reaction temperature (K).

Specific rate constant (k) of chemical reaction is calculated by the formula:

$$k = nA_0 \exp(-E/RT) \qquad (2)$$

where: A_0 is collision frequency per one reaction linkage, n is the number of equireactive bonds in a molecule.

When designing the information space for ANN predictions of the reactivity the functional relationship between the reactivity of the chemical reaction and the thermochemical properties (enthalpy of reaction - ΔH) is used.

N.N. Semenov was the first to pay attention to the functional relationship between the reactivity and reaction enthalpy (known as Polanyi – Semenov's ratio [4]):

$$E = B - \gamma \Delta H \qquad (3)$$

where B and γ – empirical coefficients.

The works [5, 6] proposed the empirical models of elementary radical reaction, which allowed constructing non-linear correlation dependences between the classical potential barrier of the radical reaction and its thermochemical properties:

- approximation of the above mentioned dependence in the work [5] by the parabola:

$$br_e = \alpha \sqrt{E_e - \Delta H_e} - \sqrt{E_e} \qquad (4)$$

- approximation of the above mentioned dependence in the work [6] in the form of the tacitly set curve:

$$br_e = D_{ei}^{1/2} \ln\left(\frac{D_{ei}^{1/2}}{D_{ei}^{1/2} - E_e^{1/2}}\right) + \alpha D_{ef}^{1/2} \ln\left(\frac{D_{ef}^{1/2}}{D_{ef}^{1/2} - (E_e - \Delta H_t)^{1/2}}\right)$$

(5)

Under the proposed empirical models assuming the harmonic stretching vibrations, the reaction of the radical abstraction R°· + R_1H → RH +$R°_1$ (where R°·and $R°_1$ - alkyl radicals, and RH and R_1H – hydrocarbon molecules) has the following parameters [5]:

1. Enthalpy $\Delta H_e = D_i - D_f + 0.5\ (hL\nu_i - hL\nu_f)$ including the energy difference of zero-point vibrations of broken and formed bonds (it represents a change in the potential energy of the system). Here ν_i – frequency of vibration of the molecule along the broken bond, ν_f – frequency of vibration of the molecule along the formed bond, D_i – - bond dissociation energy of the broken bond, $D_{ei} = D_i + 0.5hL\nu_i$, D_f – bond dissociation energy of the formed bond, $D_{ef} = D_f + 0{,}5hL\nu_f$.
2. The classical potential barrier of the activation E_e (1), which includes the zero-point energy of the broken bond.
3. The parameters $b = \pi(2\mu_i)^{1/2}\nu_i$ и $b_f = \pi(2\mu_f)^{1/2}\nu_f$, that describe the potential energy dependence of the atoms vibration amplitude along the breaking (*i*) and forming (*f*) valence linkage. $2b^2$ – the force constant of the linkage, μ_i – the reduced mass of the atoms due to breaking bond, μ_f – the reduced mass of the atoms due to forming bond.
4. The parameter r_e, which is the integrated stretching of breaking and forming bonds in the transition state.
5. The parameter α (α^2 which is equal to the ratio of force constants of breaking and forming bonds).
6. Pre-exponential factor A_0 per equireactive bond in the molecule.

According to statistically determined value of br_e, based on formula (4), it is possible to estimate the value of the classical potential barrier by the formula:

$$\sqrt{E_e} = \frac{br_e}{1-\alpha^2}\left(1 - \alpha\sqrt{\left(1 - \left(1-\alpha^2\right)\frac{\Delta H_e}{(br_e)^2}\right)}\right) \qquad (6)$$

Thus, we can assume that the dependence of the classical potential barrier E_e of the thermochemical characteristics of the reagents and the kinetic characteristics of the radical reactions can be represented as the functional relation:

$$E_e = F(\Delta H_e, T, nA_0, \alpha) \qquad (7)$$

Then the task of ANN works in predicting the values of the classical potential barrier E_e as a functional relation of the thermochemical and kinetic characteristics of the reagents with subsequent calculation of the activation energies and specific reaction rate by the formulae (1) and (2) reduce to the approximation of unknown functional relation (7).

3 Artificial Neural Networks for the Prediction of Reactivity of Molecules

To approximate the dependence (7) we used feed-forward artificial neural network [7] with a typical architecture shown in Fig. 1. We used the ANN having four inputs, 3 inner layers, each of 7 neurons and one output.

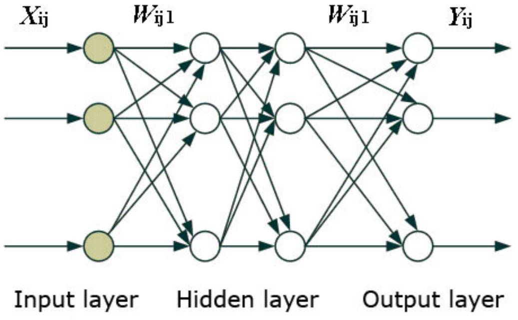

Fig.1. Typical architecture of feed-forward artificial neural network.

ANN work is set by the formulae:

$$\begin{aligned} NET_{jl} &= \sum_i w_{ijl} x_{ijl}, \\ OUT_{il} &= \Phi(NET_{jl} - \theta_{jl}), \\ x_{ij(l+1)} &= OUT_{il}, \\ \delta &= 0.5 \sum_j \sum_k \left(y_j^k - d_j^k \right)^2. \end{aligned} \qquad (8)$$

where the index i will always denote the input number, j — number of neurons in the layer, l — number of the layer; x_{ijl} — i- th input of j-th neuron in the layer l; w_{ijl} — weighting factor of the i-th input neuron number j in layer l; NET_{jl} — signal NET j-th neuron in layer l; OUT_{jl} — the output signal of the neuron; θ_{jl} — the threshold of neuron j in the layer l; x_{jl} — the input column vector of the layer l.

ANN input vector is set as the vector $x_0=\{T, D_{ei}, D_{ef}, nA_0, \alpha\}$, output data is equal to E_e.

The method of back propagation of the error [7] was used as training procedure. Activation function is a sigmoid function and is set by the following formula:

$$f(x) = \frac{1}{1+e^{-\beta x}} \qquad (9)$$

The parameter $\beta > 0$ was chosen experimentally.

For ANN training 3000 iterations were required on training set of 295 samples. Training set was constructed from the elemental radical reactions R$^{\cdot}$ + RH in the liquid phase, where R$^{\cdot}$ - a radical and RH – a hydrocarbon molecule.

Table 1 shows the comparison of the predictions of the values of the classical potential barrier of the reaction using ANN (E_{ANN}), the experimental values of activation energy (E) and the values calculated by the formula (6) of the classical potential barrier (E_e).

Table 1. Training results of ANN

Reaction	***E***	$\boldsymbol{E_e}$	$\boldsymbol{E_{ANN}}$
$CH_3 + CH_2ClBr$	27.17	36.57	30.97
$C_6H_5 + (CH_3)_4C$	23.82	43.17	19.00
$CCl_3 + CH_3(CH_2)_5CH_3$	46.88	67.35	52.35
CH_3 + *cyclo*-$[(CH_2)_6]$	44.17	46.80	42.13
C_6H_5 + *cyclo* -$[(CH_2)_5]$	27.99	31.04	27.48
CH_3 + *cyclo*-$[CH(CH_3)(CH_2)_4]$	30.23	39.67	33.18
$CCl_3 + C_6H_5CH_3$	44.30	52.62	40.88
$C_6H_5 + C_6H_5CH_3$	20.67	24.45	16.01

The error of the values prediction of the classical potential barrier of the radical reaction using ANN in the control sample (of 20 samples) was 3.34 ± 2.0 kJ / mol, which is within the experimental error (± 4 kJ / mol). The error of values prediction of the classical potential barrier for the radical reaction (6) on the same control sample was 9.5 ± 7.0 kJ / mol. ANN predicts better than the calculation by formula (6). This is due to the size of the statistical error br_e, which defines the class of radical reactions. Thus, the ANN better approximates the functional dependence (7) by calculating the weight matrix relations.

4 Conclusion

It was the first time when feed-forward ANN was used for the approximation on the experimental data of the functional dependence of the classical potential barrier of the chemical reaction from the thermochemical characteristics of the reagents and reaction kinetic parameters.

The results of the prediction of the reactivity of liquid-phase reactions of the hydrocarbons with hydrocarbon radicals are within the limits of the experimental error.

References:

[1] X1. I.I. Baskin, V.A. Palyunin, N.S. Zefirov Application of artificial neural networks in chemical and biochemical studies, *Bulletin of Moscow University. Chemistry Series.* 1999. T. 40. No. 5. C. 323-326. (in Russian)

[2] X2. E.T. Denisov, V.E. Tumanov, T.G. Denisova, T.I. Drozdov, T.S. Pokidova *The implementation of the Bank of kinetic constants of radical liquid-phase reactions on IBM PC / AT.* The Institute of Chemical Physics Chernogolovka. Preprint. 1992. 58 p. (in Russian)

[3] X3. Mallard W.G., Westley F., Herron J.T., Hampson R.F. *NIST Chemical Kinetics Database - Ver. 6.0.* NIST Standard Reference Data. 1994. Gaithersburg, MD.

[4] X4. N.N. Semenov, *About Some Problems of Chemical Kinetics and Reactivity*. Publisher Academy of Sciences of the USSR. Moscow. 1958. - 686 p.

[5] X5. E.T. Denisov New empirical models of free radical abstraction reactions. *Uspekhi Khimii*. 1997. V. 66. N. 10. p. 953-971.

[6] X6. E.T. Denisov, V.E. Tumanov Transition-State Model as the Result of 2 Morse Terms Crossing Applied to Atomic-Hydrogen Reactions. *Zhurnal Fizicheskoi Khimii.* 1994. V. 68. No. 4. p. 719-725.

[7] X7. I. Gladyshev. *Analysis and processing of the data: special guide*. St. Petersburg: Piter 2001. 752 p. (in Russian)