DE GRUYTER
OPEN

DOI: 10.1515/eces-2017-0008 ECOL CHEM ENG S. 2017;24(1):107-118

Krystyna MACEK-KAMIŃSKA[1] and Sławomir STEMPLEWSKI[2*]

APPLICATION OF NEURAL NETWORKS IN DIAGNOSTICS OF CHEMICAL COMPOUNDS BASED ON THEIR INFRARED SPECTRA

ZASTOSOWANIE SIECI NEURONOWYCH DO DIAGNOSTYKI ZWIĄZKÓW CHEMICZNYCH NA PODSTAWIE ICH WIDM W PODCZERWIENI

Abstract: The paper presents possibilities of using the so-called „finger-print" identification method and artificial neural network (ANN) for diagnosis of chemical compounds. The construction of a tool specifically developed for this purpose and the ANN, as well as the required conditions for its proper functioning were described. The identification of chemical compounds was tested in two different ways for proving correctness of the assumptions. First of all, initial studies were carried out with the objective to verify the proper functioning of the developed procedure for IR spectrum interpretation. The second research stage was to find out how the properties of artificial neural networks will satisfy identification or differentiation in case of spectra with very similar structures or for mixtures consisting of several chemical compounds. Interpretation of infrared spectra of mono-constituent substances was successfully performed for both - the training and test data. Interpretation process of infrared spectra of bi-component substances, based on the example of structurally related compounds obstructing identification process, should also be described as positive. The model was able to interpret spectra of mixtures, which were previously registered into the database. Unfortunately, the program is not always able to determine which chemical substances reflect their presence in the infrared spectrum of ternary mixtures. During the research tests, it was also noted that the more complex the structure of a substance being present in the mixture was, the more difficult the interpretation of the spectra to be carry out properly by the program was. On the other hand, positive results were obtained for mixtures of compounds with not so complex structure. It must be emphasized that the results so far are promising and more attention should be paid to them in further studies.

Keywords: artificial neural networks, identification, spectroscopy, spectra, infrared

Introduction

Infrared spectroscopy becomes more and more important tool for analysis of chemical compounds. The infrared radiation is a part of the electromagnetic spectrum comprised between the range of visible radiation and microwave radiation. There are many infrared spectrum identification methods but their implementation still carries out a high risk of error [1-5].

[1] Institute of Electromechanical Systems and Industrial Electronics, Opole University of Technology, ul. Prószkowska 76, 45-758 Opole, Poland, phone +48 77 449 80 32
[2] Faculty of Mathematics, Physics and Computer Science, University of Opole, ul. Oleska 48, 45-052 Opole, Poland, phone +48 77 452 72 16
[*] Corresponding author: sstemplewski@math.uni.opole.pl

Currently used diagnostic software compares the entire spectrum of chemical compounds or ranges selected by a user with spectra available in the library being developed by the user or purchased from the manufacturer of the software. This is the so-called „finger-print" interpretation method [6]. The cons and pros of this identification method is that there is no need and at the same time no learning opportunities to recognize spectra of chemical compounds. In case the library searching algorithm of the program cannot cope with the proper recognition of a compound's spectrum and incorrectly interprets it, than the user is not able to correct proper functioning of the program. In such a case, regardless the spectrum preparation method, the program always gives the same error. Conversely, if the spectrum of the compound being identified is not found in the database, than the result shows the spectrum with its closest similarity, which does not mean that the compound found in the library has similar structure to the compound contained in the sample. Due to the above reasons, it is reasonable to apply the advantages of artificial intelligence and infrared spectroscopy (*Fourier Transform Infrared Spectroscopy* - FTIR) by verification of the spectra interpretation method and combining it with artificial neural networks (ANN). This will enable identification of the substance and allows to resolve problems that occurred by using the previous diagnostic software.

Such configuration gives effective diagnostic tool supporting by making proper decision when identifying chemical structure. Similarities being used by chemical compound identification methods were applied by using the ANN model to develop the spectrum interpretation process of selected chemical substances characterized by large nonlinearities.

Infrared spectroscopy

The region of most interest for chemical analysis is the mid-infrared region (also called intermediate infrared) with the wavenumber range of 4000 to 400 cm^{-1}; however recently there raising interest in near-infrared (14290-4000 cm^{-1}) and far-infrared (400-100 cm^{-1}) was seen [7-9]. The entire spectrum of electromagnetic radiation is divided into several bands and the boundaries between them are designated by limits of corresponding experimental methods of producing and detecting radiation.

The absorption spectra - in terms of absorbance units (A) or transmittance (T) may be illustrated as ordinates, and depending on the wavelength λ or its reciprocal of the wavelength $1/\lambda$, i.e. wavenumbers - as abscissa. Transmittance is defined as the ratio of the light energy transmitted through the sample, while the absorbance is the logarithm to the base 10, of the reciprocal of the transmittance, T [10]:

$$A = log\frac{1}{T} \tag{1}$$

Band intensities are usually reported in semi-quantitative terms: s - strong, m - medium, and w - week.

Figure 1 illustrates an example of absorption spectrum of acetone [CH_3-C(O)-CH_3] registered as transmittance. Figure 2 shows the same absorption spectrum of acetone but presented as absorbance.

There are common two basic methods for diagnosis of compounds based on their infrared spectra [11]. The first one concerns determination of chemical functional groups (set of molecules) considering the position (wavenumber) of peaks of absorption bands characteristic for these functional groups occurring in spectra. The second method called

"fingerprint" involves comparing spectra of the test substance with catalogue reference substances until the maximum compatibility of these both spectra will be reached. To facilitate interpretation of infrared spectra, it was decided to verify the second diagnostic method. Thus, it was assumed that if two spectra have absorption bands in the same position and the individual wave height distributions are constant then it can be stated that both spectra belong to the same chemical substance.

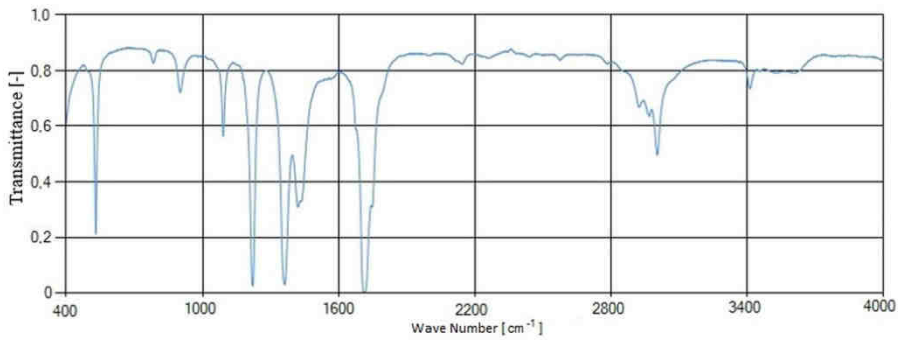

Fig. 1. Absorption spectrum of acetone presented as transmittance

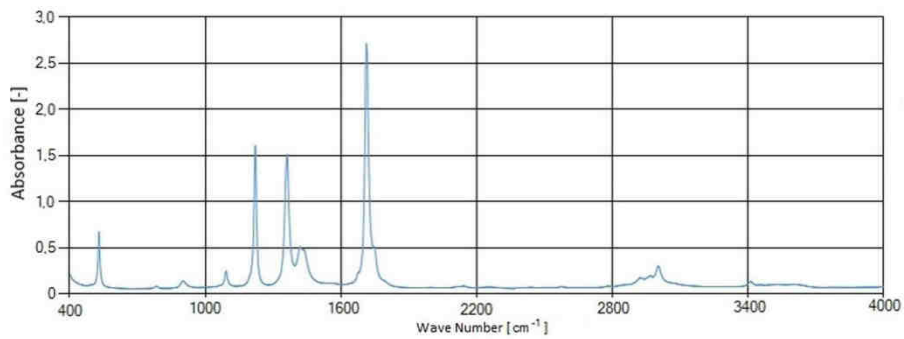

Fig. 2. Absorption spectrum of acetone presented as absorbance

Neural network model of IR spectrum interpretation process

The complexity of the process being analysed did not allow for its sufficiently accurate description using mathematical models of physical nature that would consider physical and chemical phenomena occurring there; and therefore, it was decided to apply alternative method for modelling by using artificial neural networks. The literature describes attempts of solving the problem of identification of the process [12-15]; however the focus was made on identifying segments of structures, rather than on complete individual substances. The authors also proposed a pioneering approach of identification of chemical compounds employing interdependences occurring in infrared spectra. The research studies allowed determining the input parameters of the analysed process and the neural network responses.

By identification of complex technological processes we deal with the problem of dimensionality and a fair degree of ambiguity associated with the existence of unidentified

interferences and lack of knowledge on accurate mathematical description, which necessitates the use of probabilistic methods of identification. These concern mainly linear static and dynamic objects, while the nonlinear problems, due to their enormous diversity and complexity, are solved by different methods and techniques of approximation. Despite the development of many efficient algorithms for linear and nonlinear equations with a known structure, there is a lack of general and efficient methodology for identification of nonlinear equations with unknown structure. Therefore, artificial neural networks with their attractive characteristics provide an opportunity for development of general methodology of identification of dynamic objects. Advantage of implementing the networks for the purpose of modelling and identification of objects relates to their ability of approximation of any nonlinearities and customization of the adopted structure on the basis of experimental data or other training image data. The result is that such networks are very commonly used for modelling sophisticated phenomena, the nature of which we are not able to examine, while there is an adequate knowledge on their input and output values.

The authors have proposed in the model the method for evaluating the impact of input parameters on output parameters relaying on assumption that the structure of neural network model is not known. This solution is commonly known in the literature, and it involves performing a series of experimental tests on the developed model without analysing its structure [16]. In order to gain more information about the meaning of particular components of the input vector, elimination of subsequent components as well as verification of the extent to which they affect values of output signals are performed simultaneously. There is also known an approach being used when the analysis seems to be time-consuming and the interpretation is subjective and ambiguous. In this method, the evaluation of the impact of particular parameters of the process on subsequent output vector components of the model is performed on the basis of dependent of polynomial nature.

The following describes components of input and output vectors of the model of neural object being analysed in this paper. Input vector components are position and absorption values of particular peaks (maximum points) in the spectrum.

First component of the vector is the maximum value occurring across the whole spectrum and it takes the value of 1 (under the adopted standardization). Subsequent 19 inputs are the values of the other absorption peaks. It is assumed that in a typical infrared spectrum of chemical compound about 20 characteristic points can be found. Broader designation is pointless, because it tends explicitly to volume increase of the neural network without providing any benefits. Despite the values of maximum points also their positions are being entered on the input of the artificial neural network. This is analogically performed as in the case of maximum points - the input 21 takes the value of the wavelength with the highest characteristic peak. Further 19 points are wavelengths, where the maximum points are also located on the curve (the so-called peaks):

w_1 - value of the highest characteristic point occurring in the spectrum,

w_2-w_{20} - value of subsequent characteristics occurring in the spectrum,

w_{21} - wavelength with the highest characteristic point in the spectrum,

w_{22}-w_{40} - wavelength on which subsequent characteristic points in the spectrum are located,

where w - the neural network input.

Output parameters of the artificial neural network model depend on the size of the database developed from the infrared spectra obtained during the experimental studies. Each output corresponds to the substance being registered in the database. It is also possible to indicate chemical compounds on outputs of the neural network in bit code units.

However, this solution seems to have basic limitation - inability to indicate the presence of a compound in the mixture composed of several substances. Therefore, application of the „single substance - single output" method seems to be the most appropriate one. The model used in research is called a multi-layer perceptron (MLP) and consists of four layers. Figure 3 illustrates an example of a multi-layered artificial neural network. The input layer neurons correspond to model inputs and are connected with hidden layers neurons, which are in turn connected with output layer neurons, representing the network response. The hidden layers contains 100 neurons each. The most popular algorithm used to optimize weights in the MLP neural network is the error back-propagation learning algorithm. The learning process took 83582 epoch.

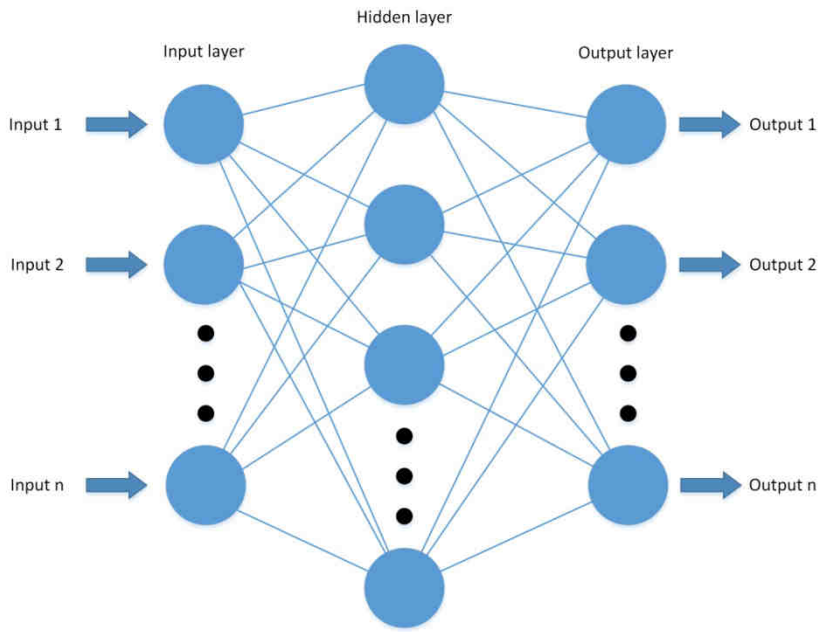

Fig. 3. Example of a multi-layered artificial neural network

Program for infrared spectra interpretation

The software for interpretation of infrared spectra was written in C#, in which the artificial neural network was programmed as well [17, 18]. The software for interpretation of infrared spectra was designed to confirm proper functioning of the modified method of identification. However, before the user will be able to use the elaborated dependence, first of all, the spectrum of chemical compound must be analysed. For the research study were used spectra of absorption in transmission system, however prior to the analysis of spectrum, it should be shifted to a spectrum in absorption system (Fig. 4). When the spectrum is already entered as absorption, it is possible to determine the so-called baseline that is the line connecting minima of absorption in the entire spectrum. It is determined according to principles adopted in analytical chemistry, under which the minimum value of absorption spectrum is the lowest value being reached by the baseline. On the right from

the minimum line it maintains the constant value of absorption, while on the left it tends to be combined with the last point in the spectrum, with the proviso that the line in particular ranges of the spectrum cannot achieve the absorption greater than the spectrum alone. Baseline is of major importance in initial analysis of the spectrum, as it is the value based on which the maximum points in each bandwidth (absorption maxima) are calculated.

The next stage of the initial analysis of the spectrum is determination of maximum points. First of all, the algorithm searches the inflection points, and then it selects from the designated intervals the maximum points with values not smaller than 0.05 of absorption units, starting from the baseline, since only then the maximum point value will have a real impact on the result of identification of chemical compound. Next, the software program selects up to twenty points with the highest value. Obviously, there are spectra having greater or smaller number of maximum points, but it is assumed that twenty points means a sufficient value of points to determine the origin of the infrared spectrum.

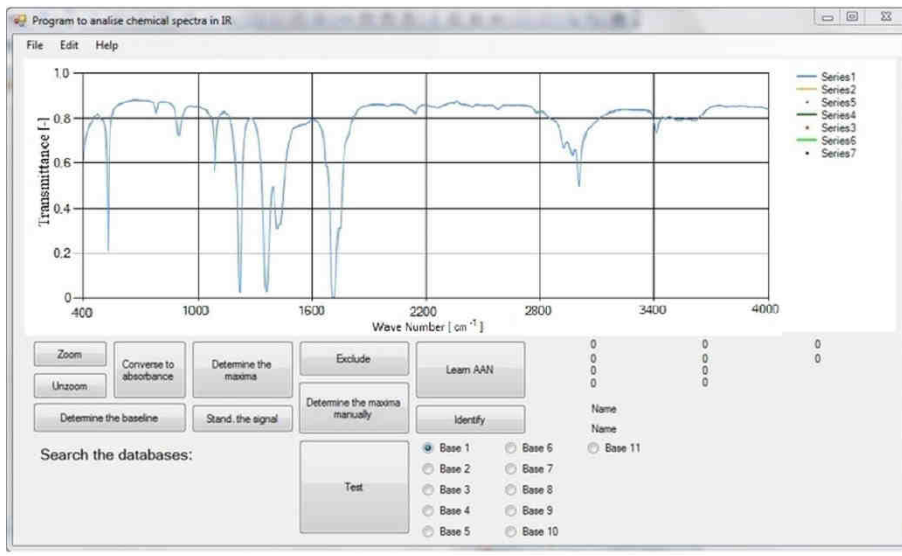

Fig. 4. The main interface with the spectrum recorded in transmission system

When the user decides that points designated by the software are correct, then he/she can go to the final stage regarding preparation of points to the process of identification of a chemical compound. Maximum points are normalized in such manner that the highest value is equated to 1, while the other values are changed according to the principle of mathematical proportion.

In order to enable the validation of identification process, before running the process should be run the function that automatically determines the chemical functional groups that can belong to the tested substance. With this option, results will be validated, that is the question will be answered: Whether the obtained result can indicate the selected chemical compound or maybe it belongs to the group (class) of chemical compound. Another advantage of this function is to indicate databases which should be considered when searching spectra.

The process of identification provides maximum points being designated at the inputs of artificial neural network. At outputs of the network we get the result, which is subsequently verified by the program. Verification means checking the class of chemical compounds (based on the identified functional groups) to which the substance indicated by ANN belongs and comparing it with other acceptable functional groups designated by the program at the stage of analysing the infrared spectrum.

Training the ANN is done by backward propagation of errors [19, 20]. The equation of activation is used in the software program as hyperbolic tangent:

$$f(x) = \frac{1}{1+e^{-\beta x}} - 1 \tag{2}$$

this is a bipolar sigmoid function, which means that at the output of the ANN, the obtained values will be in the range from –1 to 1.

Length of the training process of ANN depends on several factors, such as the number of spectra of substances entered into the database or differences between them, while the closest similarity of the substances are, the more difficult for the program is to search for the characteristic dependences (differences) in such manner that it will be able to differentiate them.

It should be noted that the error is computed using the derivative of the activation function described by the formula:

$$f'(x) = 1 - a^2 \tag{3}$$

wherein *a* is the function of activation (2).

Using the standard algorithm for training backpropagation, one of the key properties of artificial neural networks that is generalization of the achieved knowledge, failed to be fully accomplished. Although, the efficacy of identification of chemical substances based on the localization and values of maximum points in particular bands were high, than at the time of entering on the input slightly modified training data, the network was losing the ability to indicate the most dependent chemical compound. This action forced modification of the algorithm for training in such a way that it will train the network to identify compounds similar to those entered in the database. Several modifications were checked, but the best results were obtained by modifying the training algorithm in such a manner that at input of the network, the training data and the training data modified by a predetermined value were entered alternately. To prevent from training „by heart", the network's input data was entered randomly not iteratively. By applying this simple modification, improved results of ANN responses to a given signal were obtained than those obtained by standard method of training the network.

Experimental study

The identification of chemical compounds was tested in two different ways for proving correctness of the assumptions. First of all, initial studies were carried out with the objective to verify the proper functioning of the developed procedure for IR spectrum interpretation. For this purpose, mono-compounds, the so-called monosubstances were given under analysis within the database containing 50 spectra. The second research stage was to find out how the properties of artificial neural networks will satisfy identification or differentiation in case of spectra with very similar structures or for mixtures consisting of several chemical compounds. For the need of performing this research a database with 50 different classes of chemical compounds was developed.

Table 1 presents output values of the artificial neural network corresponding to particular spectra for training data. The Output 1 indicated Acetone as the result of classification, the Second Output indicated Atmer, and the Third Output indicated Benzene, the fourth Output to Ether, fifth Output to Methanol. As shown in the results, the presented artificial neural network correctly performed the process of spectrum identification, as evidenced by values on particular network's outputs.

Table 1
Values of selected artificial neural network outputs for particular chemical compounds (training data)

	Output 1	**Output 2**	**Output 3**	**Output 4**	**Output 5**
Acetone	**0.9990**	0.0003	0.0129	–0.0101	–0.0153
Atmer	0.0005	**0.9993**	–0.0163	–0.0071	0.0121
Benzene	0.0124	–0.0025	**0.9985**	0.0063	0.0005
Ether	–0.0004	–0.0022	–0.0005	**0.9997**	0.0017
Methanol	–0.0034	–0.0019	0.0027	–0.0084	**0.9994**

Table 2
Values of selected artificial neural network outputs for particular chemical compounds (test data)

	Output 1	**Output 2**	**Output 3**	**Output 4**	**Output 5**
Acetone	**0.9075**	0.0593	–0.0368	–0.0051	–0.0046
Atmer	–0.0614	**0.9846**	–0.0594	0.0899	–0.0269
Benzene	0.0504	0.0600	**0.9206**	–0.0558	0.1527
Ether	–0.0154	0.0305	–0.0199	**0.9541**	–0.0044
Methanol	0.0042	–0.0400	0.0774	0.0563	**0.9962**

Table 2 presents average output values of the artificial neural network corresponding to particular spectra for test data. Each data has 10 different spectra of the same chemical compound. Concluding, the network has correctly classified the spectra, both for training as well as for test data, proving the correctness of the posed theses.

In order to confirm the efficacy of the IR spectroscopy-based method for identification of chemical compounds by means of ANN, the research was carried out on mixture of two compounds - creatinine and amphetamine. The main reason for choosing these two substances was the similar structure of these two compounds, which causes overlap of the spectra absorption bands, thus impeding their identification (Fig. 5).

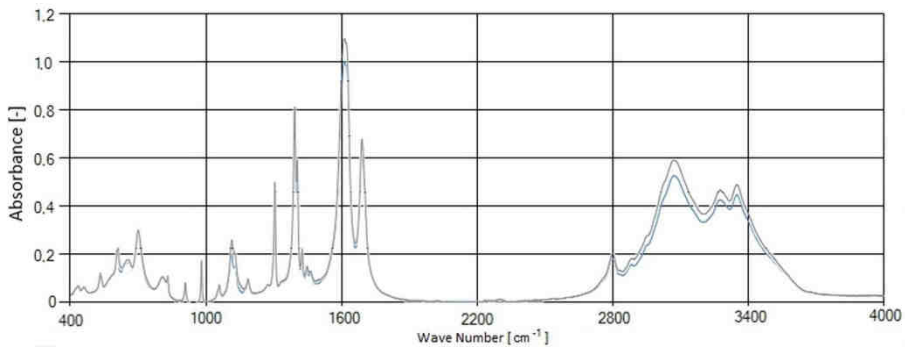

Fig. 5. Absorption spectra of a mixture of amphetamines (5%) and creatinine (95%) (grey), and of creatinine (100% - blue)

Due to the fact that the two spectra were produced by using the FT-IR method (Fourier Transform Infrared Spectroscopy), it is possible to submit them electronically in order to obtain spectra of mixtures with different compositions. In this manner, fourteen spectra of mixtures containing amphetamine spectra in the range from 5 to 95% in increments of 5% (5, 10, 15, 20, 25, 85, 90, and 95 percent) and of 10% (30, 40, 50, 60, 70, and 80 percent) were formed.

Both spectra of pure monosubstances have characteristic bands at similar wavenumbers, and the additional difficulty is that the bands at the time of mixing the spectra of these two substances overlap, blurring distinctions, which makes further identification of the mixture's components very difficult.

The developed database of 16 spectra was used to train the ANN to identify binary and single compound-mixtures. The ANN managed to recognize minimal distinctions among spectra, as used in their subsequent identification. Among 16 spectra of samples being tested, the ANN after training process was able to differentiate all of the tested spectra. It this case, it managed to perform much better than the used so far identification programs, which encountered problems when the mixture contained much less amphetamine than creatinine.

Table 3 shows output values of the artificial neural network for training data, while Table 4 presents output values for test data. As it can be seen, the network was able to search for characteristics that allow carrying out proper interpretation of the spectrum.

Table 3

Values of selected artificial neural network outputs for particular chemical compounds and their mixtures (training data)

	Output 1	**Output 2**	**Output 3**	**Output 4**	**Output 5**
Amphetamine	**0.9573**	–0.0048	–0.0038	–0.0042	–0.0134
Creatinine	–0.0147	**0.9805**	–0.0051	–0.0054	–0.0144
5% AMP	–0.0053	–0.0035	**0,9752**	–0.0045	–0.0058
60% AMP	0.0087	0.0042	0.0045	**0.9918**	0.0088
95% AMP	0.0117	–0.0071	–0.0072	–0.0081	**0.9590**

Table 4

Values of selected artificial neural network outputs for particular chemical compounds and their mixtures (test data)

	Output 1	**Output 2**	**Output 3**	**Output 4**	**Output 5**
Amphetamine	**0.9766**	–0.0637	–0.8221	–0.0252	–0.5446
Creatinine	–0.0923	**0.9782**	0.0552	0.1126	0.0080
5% AMP	0.0079	0.1569	**0.9802**	–0.1305	0.0945
60% AMP	0.0615	–0.0500	–0.1760	**0.9931**	–0.0886
95% AMP	–0.2788	–0.0906	0.0813	–0.0142	**0.9870**

The trained neural network for identification of mixtures consisting of chemical compounds is able to search for characteristics, which enable it to classify chemical compounds or a mixture thereof. If the neural network has not been trained with mixture patterns, it also indicates changes in output, flagging up possibility of occurrence of the substance even in its small amounts.

For the purpose of performing the research study, database was developed that consists of 10 mono-compounds. The objective of the study was to examine the efficacy of the process of identification of mixture spectra for three chemical mixtures.

Table 5

Artificial neural network output values for particular samples given on the network's input

	Spectrum A	Spectrum B	Spectrum C	Spectrum D
Sample 1 (A10 B10 C80)	–0.6522	–0.0180	0.9998	0.0725
Sample 2 (A10 B10 D80)	0.3243	–0.7361	0.7096	0.9997
Sample 3 (A10 B80 C10)	0.9706	0.9888	0.9845	0.6021
Sample 4 (A10 B80 D10)	0.9882	0.9999	0.5730	–0.7231
Sample 5 (A25 B25 C50)	–0.9854	0.7600	0.9998	–0.2973
Sample 6 (A25 B25 D50)	0.1596	0.3459	0.9999	0.9893
Sample 7 (A25 B50 C25)	0.9991	–0.9999	0.9998	–0.9877
Sample 8 (A33 B33 C33)	0.6431	0.9998	0.9999	–0.9111
Sample 9 (A33 B33 D33)	0.7199	–0.9998	0.9999	0.8895
Sample 10 (A50 B25 C25)	0.9999	0.9932	0.9995	–0.9719
Sample 11 (A50 B25 D25)	0.9985	–0.9999	0.9999	0.9434
Sample 12 (A80 B10 C10)	0.9999	–0.9925	0.9306	0.3704
Sample 13 (A80 B10 D10)	0.9999	–0.9896	0.8049	0.1113

The results presented in the table above indicate that the artificial neural network is not always able to identify accurately the mixture consisting of three substances. The accuracy of identification - in this case - depends on the structure of chemical compounds; the more complex the substance is chemically the greater difficulty for the neural network is to identify occurrence of other compounds.

To summarize the experimental studies performer, it must be emphasized that the infrared spectrum interpretation of chemical compounds consisting of more than one component is a very sophisticated analysis, and not so rare, even for a qualified expert in the field it may be difficult or impossible to perform it.

Responses from the neural network model for a given problem do not indicate quantitative presence of the substance occurring in the spectrum of the mixture, but only indicate the presence of chemical compound in the infrared spectrum.

Interpretation of infrared spectra of mono-constituent substances was successfully performed for both - the training and test data, which is presented in Tables 1 and 2 in this section. The model was able to classify spectra by assigning them to their corresponding chemical compounds. Responses from the neural network model were correct for the training data as well as for the test data. These results confirm that it was manager to obtain the most valuable property of artificial neural networks, which is generalization of the acquired knowledge.

Interpretation process of infrared spectra of bi-component substances, based on the example of structurally related compounds obstructing identification process, should also

be described as positive, which present Tables 3 and 4. The model was able to interpret spectra of mixtures, which were previously registered into the database. The experimental task was even more difficult, while there were almost no optical differences between the spectrum of a pure substance and of a substance having little quantity of the other compound. Responses from the neural network model, as in the analysis of mono-constituent substances, were correct for training and test data.

Analysis of ternary compounds was the final phase of the research study, and definitely it was the most sophisticated stage. Unfortunately, the program is not always able to determine which chemical substances reflect their presence in the infrared spectrum of ternary mixtures. During the research tests, it was also noted that the more complex the structure of a substance being present in the mixture was, the more difficult the interpretation of the spectra to be carry out properly by the program was. On the other hand, positive results were obtained for mixtures of compounds with not so complex structure. It must be emphasized that the results so far are promising and more attention should be paid to them in further studies.

Conclusions

Results presented in this research study provide evidence on the efficiency of the process of interpretation of infrared spectra based on the method using artificial neural network. The results achieved during the research study clearly confirm that application of artificial neural networks and diagnostics based on the so-called „fingerprint" method enable the development of neural network model, which allows for efficient identification of chemical compounds by using their infrared absorption spectra.

The main objective of the research was to improve and develop the process of identification of infrared spectra. This objective was achieved by verifying the identification method and combining it with artificial intelligence as artificial neural network, which resulted in development of neural network model allowing diagnosis of chemical compounds.

Selection of network configuration and training methods was introduced based on evaluation of average and standard deviations of the error rate, which were calculated for responses of the network to selected values of input signals. In order to restrict the ability to select the non-optimal configuration of the network for which very low values of errors were obtained due to single favourable selection of initial weights, the best configurations, according to the authors, were tested for repeatability of the obtained results.

The developed software can be a useful tool for quick diagnosis of chemical compounds, such as for forensic laboratories, by applying the advantages offered by the ANN to generalize the acquired knowledge, which means indicating chemical compounds similar to those found in the database of the program. The program indicates the direction for further research by classifying the substance being tested to a certain group of chemical compounds.

References

[1] Anilkumar GK. A subjective job scheduler based on a backpropagation neural network. Human-centric Computing Inform Sci. 2013;3:3-17. DOI: 10.1186/2192-1962-3-17.

[2] Babak MA, Sharafat AR, Setarehdan SK. An adaptive backpropagation neural network for arrhythmia classification using R-R interval signal. Neural Network World. 2012;6:535-548. DOI: 10.14311/NNW.2012.22.033.

[3] Balara D, Timko J, Žilková J. Application of neural network model for parameters identification of non-linear dynamic system. Neural Network World. 2013;2:103-116. DOI: 10.14311/NNW.2013.23.008.
[4] Klawun C, Wilkins CL. Neural network assisted rapid screening of large infrared spectral databases. Anal Chem. 1995;67(2):374-378. DOI: 10.1021/ac00098a023.
[5] Jalali-Heravi M. Neural networks in analytical chemistry. Methods Molecular Biol. 2008;458:78-118. DOI: 10.1007/978-1-60327-101-1_6.
[6] Polfer NC, Paizs B, Snoek LC, Compagnon I, Suhai S, Meijer G, et al. Infrared fingerprint spectroscopy and theoretical studies of potassium ion tagged amino acids and peptides in the gas chase. J Amer Chem Soc. 2005;127(23):8571-8579. DOI: 10.1021/ja050858u.
[7] McCarty GW, Reevesab JB, Reevesab VB, Follettc RF, Kimbled JM. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Sci Soc Amer J. 2002;66:640-646. DOI: 10.2136/sssaj2002.6400.
[8] Colthup NB, Wiberley SE, Daly LH. Introduction to Infrared and Raman Spectroscopy. New York and London: Academic Press; 1990. http://www.sciencedirect.com/science/book/9780121825546.
[9] Silverstein RM, Webster FX, Kiemle DJ. Spectrometric Identification of Organic Compounds. New York: John Wiley Sons; 2014.
[10] Naumann D. A novel procedure for strain classification of fungal mycelium by cluster and artificial neural network analysis of Fourier transform infrared (FTIR) spectra. Analyst. 2009;134(6):1215-1223. DOI: 10.1039/b821286d.
[11] Coates J. Interpretation of Infrared Spectra, A Practical Approach. New York: John Wiley Sons; 2000. DOI: 10.1002/9780470027318.a5606.
[12] Xia M, Huang R, Sun Y, Semenza GL, Aldred SF, Witt KL, et al. Identification of chemical compounds that induce HIF-1alpha activity. Toxicol Sci. 2009;112(1):153-63. DOI: 10.1093/toxsci/kfp123.
[13] Srinivasan GV, Ranjith C, Vijayan KK. Identification of chemical compounds from the leaves of Leea indica. Acta Pharm. 2008;58(2):207-14. DOI: 10.2478/v10007-008-0002-7.
[14] Aguilera N, Becerra J, Villaseñor-Parada C, Lorenzo P, González L, Hernándeza V. Effects and identification of chemical compounds released from the invasive Acacia dealbata Link. Chem Ecol. 2015;31:479-493. DOI: 10.1080/02757540.2015.1050004.
[15] Tanaka M, Kuriyama S, Itoh G, Kohyama A, Iwabuchi Y, Shibata H, et al. Identification of anti-cancer chemical compounds using Xenopus embryos. Cancer Sci. 2016;107:803-811. DOI: 10.1111/cas.12940.
[16] Janczak A. Identification of Nonlinear Systems Using Neural Networks and Polynomial Models: A Block-Oriented Approach. Berlin: Springer Science Business Media; 2005.
[17] Tadeusiewicz R, Chaki R, Chaki N. Exploring Neural Networks with C#. Boca Raton: CRC Press; 2015.
[18] Heaton J. Programming Neural Networks with Encog 3 in C#. St. Louis: Heaton Research; 2011.
[19] Gudise VG, Venayagamoorthy GK. Comparison of particle swarm optimization and backpropagation as training algorithms for neural networks. Swarm Intelligence Symposium, 2003. DOI: 10.1109/SIS.2003.1202255.
[20] Mohammadi N, Mirabedini SJ. Comparison of particle swarm optimization and backpropagation algorithms for training feedforward neural network. J Mathemat Computer Sci. 2014;12:113-123. http://www.isr-publications.com/jmcs/articles-711-comparison-of-particle-swarm-optimization-and-backpropagation-algorithms-for-training-feedforward-neural-network.