Application of a perceptron to solve the problem of analyzing the fluorescence spectrum of a DBMBF2 sensor in a mixture of aromatic hydrocarbons

Ilya Katanov
Samara National Research University
Samara, Russia
semargl42@gmail.com

Alexander Kupriyanov
Samara National Research University
Samara, Russia
akupr@ssau.ru

Yuriy Kononevich
Institute of Organoelement compounds. A. N. Nesmeyanov of the Russian Academy of Sciences
Moscow, Russia

Dmitry Ionov
Photochemistry center of the RAS, NRC "Crystallography and Photonics" of the RAS
Moscow, Russia

***Abstract*—The article deals with the problem of choosing the neural network architecture in the problem of analyzing signal data obtained by shooting spectra from fluorescent sensors, which are based on the formation of exciplexes between the boron Dibenzoyl methanate fluorophore (DBMBF2) and aromatic compounds. Attention is paid to the problem of selecting the structural features and parameters of the network in the process of training and testing on available data.**

***Keywords*—DBMBF2, aromatic compounds, neural network, fluorescent sensors**

I. Introduction

This work is based on the usage of data obtained from the sensor described in [1]. This sensor can selectively detect benzene and its derivatives in multicomponent mixtures of aromatic hydrocarbon compounds. It's work is based on the properties of the dibenzoyl methanate boron fluorophore (DBMBF2). During the operation of this sensor, changes in the DBMBF2 fluorescence spectra that appear due to the formation of complexes (exciplexes) between the fluorophore and aromatic compounds in the excited state are measured. As the output, the sensor provides spectrum shape data in the form of 2048 spectrum values, each of which represents the signal intensity at a specific frequency. In [1], processing and analysis of spectral data is performed based on the model [2] of changing of the fluorescence spectra shape of DBMBF2, which is adsorbed on silica gel. The multidimensional least squares method is used to determine the parameters of this model [5]. The resulting parameters are then applied to solve the inverse problem of calculating the concentration of hydrocarbons from a known form of the spectrum. However, an attempt to analyze data obtained from two or more chemosensitive elements that react differently to changes in gas concentration showed insufficient effectiveness of this method in the described task.

The problem of calculating the concentration of hydrocarbons by using a known form of the spectrum can be solved not only by using the least squares method. Other method to solve similar tasks were presented in work [4]. However, we intend to propose a different kind of solution for this problem. This data analysis task is characterized by a lack of information about the data structure, dependencies between data, and the distribution of analyzed indicators. Under these conditions, the best solution is to use neural networks to create a neural network model that can determine the concentration of hydrocarbons by the shape of the spectrum. This is due to the ability of neural networks to learn and model nonlinear processes while working with data that does not have clear relationships and dependencies.

However, the usage of neural networks is associated with a number of difficulties. In particular, we need to choose the network architecture that is appropriate for the task, as well as determine the values of hyperparameters that would allow us to solve the problem in the best way using the available data. With this approach, we will have to go through all the architecture and hyperparameters options, and choose a specific set that will give the best results of solving the problem among the presented options.

II. Choice of neural network architecture

In the described case, we solve the problem of predicting the gas concentration based on the available values of the spectrum shape taken from two sensors. At the same time, the spectrum data taken from specific sensor does not depend on data of other sensor in any way and does not form any clear sequence. Each element of the source data is a one-dimensional array of fluorescence intensities at different wavelengths, taken for a specific concentration of various hydrocarbons in the air.

Prediction problems are very often solved by using direct propagation networks. The basis of networks of this type is a multi-layer perceptron, which is widely used for data processing in modeling, identifying various situations, and predicting any events or values [9-14]. Combined with the format of the available data, a multi-layer perceptron is a very promising solution. For this reason, it was decided to choose an architecture based on the use of a multi-layer perceptron for further research.

III. Neural network hyperparameters selection

Neural network hyperparameters are parameters that are used during network training, but do not change in the process. These include parameters such as:

- Learning rate.
- Neuron activation functions.
- Optimization algorithm.
- The batch size (Batch Size).
- Number of neurons in hidden layers

The difficulty of choosing hyperparameters is that the choice must satisfy two conditions – solve the problem at the

Copyright © 2020 for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0)

CEUR-WS.org/Vol-2667/paper70.pdf

lowest values of prediction errors and provide sufficient generalizing ability of the network to avoid overfitting [6].

There are various approaches for selecting the values of neural network hyperparameters [7,8]. The most popular method is called Grid Search, which could be described as searching for combinations of all the proposed values of hyperparameters in the network and selection of the best combination based on a certain metric, such as the deviation of predicted values from true values, in ppm in our case. It was decided to use this approach in this work.

Each of the sensor elements used in the experiment provides data in form of an array of 2048 fluorescence intensity values at different wavelengths. A detailed study of the data allowed us to determine that out of 2048 channels provided by each sensor, values of only 882 responded to changes in concentration. Data from two sensors was used simultaneously during network training. Thus, since following above statements, the data used for network training is a one-dimensional array of 1764 elements representing fluorescence intensities captured by detectors from two sensor elements consisting of different chemosensory materials, characterized by surface modification of the carrier matrix or the use of various DBMBF2 derivatives[4], the number of input neurons was assumed to be equal to the number of available elements – 1764 to be exact. As options of the number of neurons in the hidden layers, the products of the number of input neurons by various powers of the number 2 were accepted. Thus, options like 220, 441, 882, 1764, 3528 neurons per hidden layer were accepted. 1,2,3 and 4 layers were the acceptable number of hidden layers.

The options of optimization algorithms was represented by such algorithms as Adam, Adagrad, Adadelta, SGD (Stochastic Gradient Descent).

The options of activation functions for hidden layers consisted of ReLU, LeakyReLU, Tanh (hyperbolic tangent), sigmoid, and linear function.

The learning speed was represented by values from 0.001 to 0.0001 with step of 0.0001.

Thus, all combinations of the parameters of neural networks were trained and validated using cross-validation on a samples of fluorescence intensity and concentration values obtained from sensors during the experiment described in [1], and among them one with the lowest average error of cross-validation, having values k=10, was selected.

As a result, the following parameters of the neural network prevailed:

- Number of hidden layers: 3;
- Number of neurons in hidden layers: 882, 882 and 220 in series;
- Activation function of hidden layers of neurons: ReLU;
- Learning rate: 0.0004;
- Optimization algorithm: Adagrad.

IV. Evaluation of the Resulting Network Model

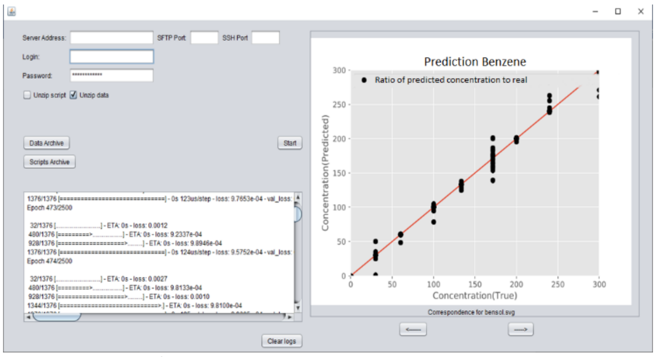
To evaluate the obtained hyperparameters, a software tool, which was used to train a neural network with the proposed parameters and available data, as well as to obtain output data of training results and network predictions, was implemented.

The software tool that was mentioned above consists of two components:

- A server-side program that directly trains the network using the capabilities of the supercomputer of the Samara national research University, by utilizing CUDA cores through Python libraries, TensorFlow and Keras to be exact;

- A client application written in the Java using libraries such as Swing, for providing GUI, JSch to connect to server and control training process and Apache Commons to obtain data after training finishes. This application loads the Python program, and transfers data, used to train the network, to the server. It also was controlling the start of training and receiving output results obtained as a result of training the network and predicting gas concentration values on testing data-sets.

Screenshot of the operation of this application is shown in figure 1.



Server Address:
SFTP Port:
SSH Port:
Login:
Password: ************
☐ Unzip script ☑ Unzip data

```
1376/1376 [==============================] - 0s 123us/step - loss: 9.7653e-04 - val_loss:
Epoch 473/2500

  32/1376 [..............................] - ETA: 0s - loss: 0.0012
 480/1376 [=========>....................] - ETA: 0s - loss: 9.2337e-04
 928/1376 [===================>..........] - ETA: 0s - loss: 9.8946e-04
1376/1376 [==============================] - 0s 124us/step - loss: 9.5752e-04 - val_loss:
Epoch 474/2500

  32/1376 [..............................] - ETA: 0s - loss: 0.0027
 480/1376 [=========>....................] - ETA: 0s - loss: 9.8133e-04
 928/1376 [===================>..........] - ETA: 0s - loss: 0.0010
1344/1376 [============================>.] - ETA: 0s - loss: 9.8100e-04
```

Fig. 1. Example of application work.

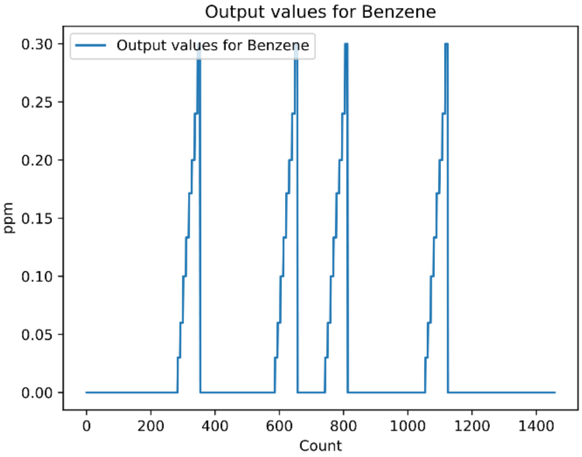


Fig. 2. Benzene concentrations per timeslot values used for k-fold cross-validation.

This application was used to test the previously obtained set of hyperparameters by training network on 1,376 examples of sensor measurements, with concentrations presented in fugures 2, 3 and 4.

Cross-validation by the K-Fold method, where k=10, was used to ensure validity of our results, which gave us the prediction results for benzene concentrations shown in figure 5.

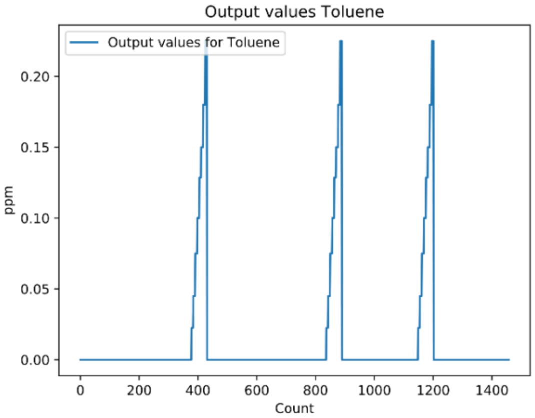


Fig. 3. Toluene concentrations per timeslot values used for k-fold cross-validation.

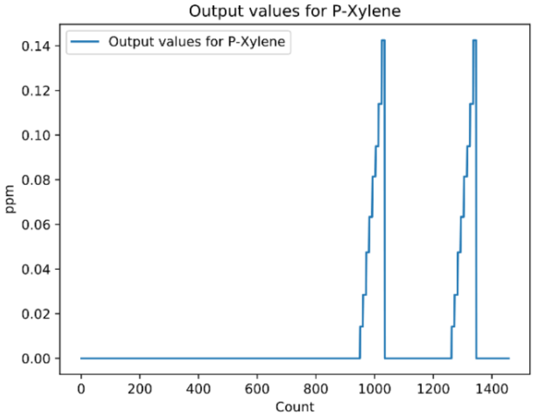


Fig. 4. P-Xylene concentrations per timeslot values used for k-fold cross-validation.

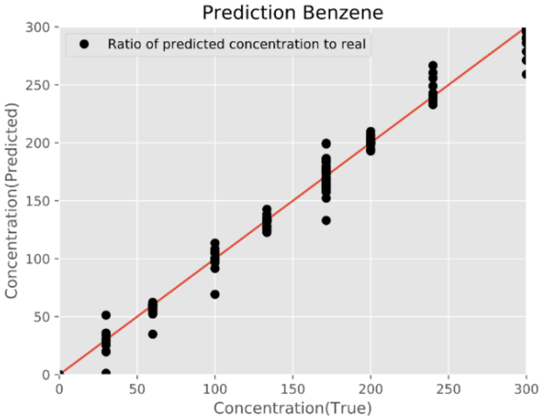


Fig. 5. The ratio of predicted benzene concentrations to real ones, the concentration is measured in ppm.

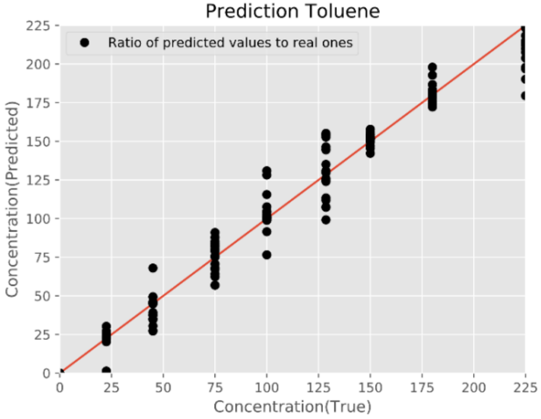


Fig. 6. The ratio of predicted toluene concentrations to real ones, the concentration is measured in ppm.

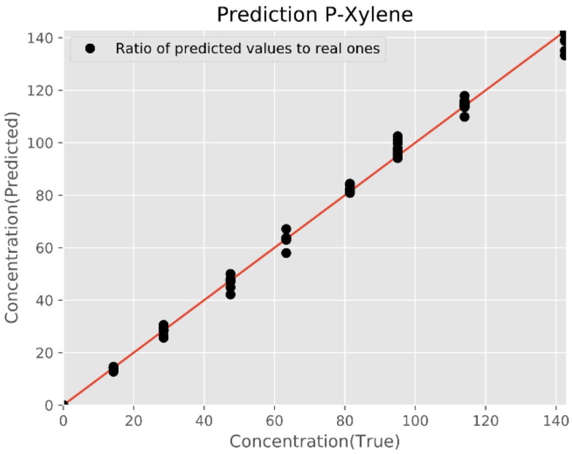


Fig. 7. The ratio of predicted p-xylene concentrations to real ones, the concentration is measured in ppm.

The same set of hyperparameters was applied to training neural networks used for prediction of concentrations of toluene and p-xylene. The results of k-fold cross-validation for those neural networks is shown in figures 6 and 7.

As you can see from the picture, the predicted concentration values are very close to the real ones. The average error rate during validation was 8 ppm, when measuring benzene with evenly distributed concentrations equal to 30, 60, 100, 133, 171, 200, 240 and 300 ppm. This confirms that the multilayer perceptron with the architecture presented in this paper can be used in the task of analyzing spectral data using two or more sensors, which will allow us to obtain a sufficiently high accuracy of predicting gas concentrations.

V. Improvements

Despite the fact that the obtained result of selecting hyperparameters in training already shows results close to real ones, it is possible to achieve even greater accuracy by selecting hyperparameters using evolutionary algorithms, such as the genetic algorithm, which has been growing in popularity in recent years when performing this task.

Acknowledgment

The research was supported by the Ministry of Science and Higher Education of the Russian Federation (Grant # 0777-2020-0017) and partially funded by RFBR, project number # 19-29-01135.

References

[1] D. Ionov, G. Yurasik, Y. Kononevich, V. Sazhnikov, A. Muzafarov and M. Alfimov, “Simple Fluorescent Sensor for Simultaneous Selective Quantification of Benzene, Toluene and Xylene in a Multicomponent Mixture,” Procedia Eng., vol. 168, pp. 341-345, 2016.

[2] D. Ionov, G. Yurasik, A. Antonov, V. Sazhnikov and M. Alfimov, “Model of formation of exciplexes of dibenzoylmethanate boron difluoride with aromatic hydrocarbons on the surface of silica,” High energy chemistry, vol. 49, no. 3, pp. 210-215, 2015.

[3] D. Ionov, V. Sazhnikov, G. Yurasik, A. Safonov, Y. Kononevich and M. Alfimov, “Exciplexes Of Fluorine And Methyl Derivatives Of Boron Dibenzoyl Methanate With Benzene And Toluene On The Surface Of Trimethylsilylated Aerosil,” High energy chemistry, vol. 52, no. 6, pp. 473-479, 2018.

[4] N. Vasilyev and A. Morozov, “Identification of substances according to strongly distorted measurement errors of the spectra,” Computer Optics, vol. 38, no. 4, pp. 856-865, 2014.

[5] M. Maeder and Y.-M. Neuhold, “Practical Data Analysis in Chemistry,” Data Handling in Science and Technology, vol. 26, pp. 1-326, 2007.

[6] E. Watanabe and H. Shimizu, "Relationships between internal representation and generalization ability in multi layered neural network for binary pattern classification problem," Proceedings of International Conference on Neural Networks (IJCNN), Nagoya, Japan, vol. 2, pp. 1736-1739, 1993. DOI: 10.1109/IJCNN.1993. 716989.

[7] V. Tsaregorodcev, “Determining the optimal size of the reverse propagation neural network by comparing the average values of modules weights of synapses,” Proceedings of the 14th international conference on Neurocybernetics, Rostov-on-don, pp. 60-64, 2005.

[8] A. Larko, “Optimizing the size of the reverse propagation neural network” [Online]. URL: http://www.sciteclibrary.ru/rus/catalog/pages/8621.html.

[9] V. Gridin, “Reverse propagation neural network size optimization,” Proceedings of the Scientific and practical seminar New information technologies in automated systems, Moscow, pp. 270-273, 2016,

[10] J. El Haddad, M. Villot-Kadri, A. Ismael, G. Gal-lou, K. Michel, “Artificial neural network for on-site quantitative analysis of soils using laser induced breakdown spectroscopy,” Spectrochimica Acta Part B: Atomic Spectroscopy, vol. 79-80, pp. 51-57, 2013.

[11] B. Kuang, Y. Tekin and A.M. Mouazen, “Comparison between artificial neural network and partial least squares for on-live visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content,” Soil and Tillage Research, vol. 146, pp. 243-252, 2015.

[12] A.M. Ahmed, O. Duran, Y. Zweiri and M. Smith, “Quantitative analysis of petroleum hydrocarbon contaminated soils using spectroscopy, spectral unmixing and deep neural networks,” Proceedings Image and Signal Processing for Remote Sensing, Berlin, Germany, vol. 10789, 2018.

[13] Y. Zhao, “A spectral analysis based heteroscedastic model for the estimation of value at risk,” The Journal of Risk Finance, vol. 19, no. 3, pp. 295-314, 2018.

[14] E.C. Ferreira, D.M.B.P. Milori, E.J. Fer-reira, R.M. da Silva, L. Martin-Neto, “Artificial neural network for Cu quantitative determination in soil using a portable Laser Induced Breakdown Spectroscopy system,” Spectrochimica Acta Part B: Atomic Spectroscopy, vol. 63, no. 10, pp. 1216-1220, 2008.