32nd European Modeling & Simulation Symposium
17th International Multidisciplinary Modeling & Simulation Multiconference

ISSN 2724-0029 ISBN 978-88-85741-44-7 © 2020 The Authors.
DOI: 10.46354/i3m.2020.emss.048

Machine Learning for Optical Sensing with Grating Nanostructures

E.D. Chubchev[1,*], I.A.Nechepurenko[1,2], A.V. Dorofeenko[1,2,3], K. A. Tomyshev[2], O.V.Butov[2] , D.P. Kulikova[1], E.M. Sgibnev[1], G.M. Yankovsky[1], A.V. Baryshev[1], A.S. Baburin[4], I.A. Rodionov[4]

[1]Dukhov Research Institute of Automatics (VNIIA), 22 Sushevskaya, Moscow 127055, Russia
[2]Kotelnikov Institute of Radioengineering and Electronics of Russian Academy of Sciences, 11-7 Mokhovaya, Moscow 125009, Russia
[3]Institute for Theoretical and Applied Electromagnetics of Russian Academy of Sciences, 13 Izhorskaya, Moscow 125412, Russia
[4]FMN Laboratory, Bauman Moscow State Technical University, 2/18 Rubtsovskaya emb., Moscow 105082, Russia

*Corresponding author. Email address: eugene.chubchev@yandex.ru

Abstract

Over the past decade, machine learning has found a large number of applications in physics. Machine learning algorithms can extract the most informative features of the data, reduce the dimensionality and increase the signal-to-noise ratio. This article discusses the use of machine learning algorithms to increase the accuracy of the optical sensors based on optical fiber Bragg grating sensor and the hydrogen sensor based on Wood anomaly in a diffraction grating. We show that application of machine learning algorithms to experimental data processing allows reaching high accuracy and reduce level of noise in optical sensors.

Keywords: optical sensors; machine learning; fiber Bragg gratings; gas sensors.

1. Introduction

Sensing is an important task for many scientific and engineering problems (Ignatov & Merzlikin, 2020; Nechepurenko et al., 2018; Tomyshev, Manuilovich, Tazhetdinova, Dolzhenko, & Butov, 2020). The important problem in sensing is sensor data processing. It allows both raising the sensitivity to a physically achievable limit and solving the problem of poor selectivity (Askim, Mahmoudi, & Suslick, 2013; Hotel, Poli, Mer-Calfati, Scorsone, & Saada, 2018; Johnson et al., 1997). Various methods are used in data processing, such as nonlinear regression, which is typically based on a theoretical analytical model. However, in some cases the parameters of interest can be determined imprecisely with the theoretical model (Kornienko et al., 2020) or cannot be derived from it because of its complexity.

In such cases, machine learning (ML) can be applied. ML is a branch of computer science, which goal consists in automatically building a model of data without being explicitly programmed. Machine learning algorithms can simplify the structure of the data, automatically find patterns in it, and solve

problems with higher accuracy.

Examples of ML algorithms are k-nearest neighbor method (Altman, 1992), decision tree-based algorithms such as gradient boosting (Friedman, 2001), principal component analysis (PCA) (Gorban, Kégl, Wunsch, & Zinovyev, 2008) and artificial neural networks (Goodfellow, Bengio, & Courville, 2016).

In sensing, ML algorithms are used to solve the problem of poor selectivity (Hotel et al., 2018; Kornienko et al., 2020) and to compensate sensor drift (Liu et al., 2019; Ma, Luo, Qin, Wang, & Niu, 2018; Wang et al., 2017; Zhao et al., 2019). However, ML algorithms are quite rarely used to increase the precision of the sensors with continuous sensor readings.

In this paper, we apply ML approach for processing experimental data obtained with two different optical sensors. We use dimensionality reduction method such as principal component analysis and polynomial regression to predict the parameter of interest. We show that ML algorithms allow attainment of higher precision and reduce the level of noise in sensor readings.

2. Application of Machine learning methods to optical sensors

2.1. TFBG sensor

We apply ML algorithms to data obtained in experiments with a tilted fiber Bragg grating (TFBG) sensor (Fig. 1) consisting of an optical fiber covered with a golden thin film and tilted Bragg grating inside the fiber core. The TFBG sensor working principle is the following. The fundamental mode excites some cladding modes and the surface plasmon (Chubchev, Nechepurenko, Dorofeenko, Vinogradov, & Lisyansky, 2020) when scattered by the tilted Bragg grating. The typical scattering spectrum is shown in Fig. 2a. The spectral position of surface plasmon resonance is highly sensible to the refractive index of surrounding medium. Our goal is to find the way to predict the change in refractive index by the spectrum.

Because of the model complexity, there is no precise analytical model for the dependence of the plasmon resonance frequency on the surrounding medium. For this reason, we apply ML algorithms to find the external refractive index from experimentally measured scattering spectrum.

We conducted experiment where the TFBG sensor was surrounded by distilled water. Then, the refractive index was changed by mixing the water with small amounts of isopropanol. Isopropanol was added 27 times. That is, 28 values of refractive index were obtained. The change in the water refractive index was calculated analytically. As the isopropanol was added, the associated spectra were measured. The number of spectra associated with each refractive index value varied from 5 to 11. The total number of spectra was equal to 202. Each spectrum consisted of 16001 values of signal amplitude at various wavelengths.

Before applying ML algorithms, we conducted Fourier transform of the spectra to remove noise and the slowly varying part of the spectra. In Fig. 2b, an example of the filtered spectra is shown. This procedure is described in details in … .

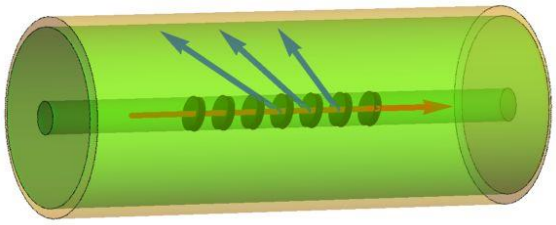

Figure 1. The scheme of the TFBG sensor.

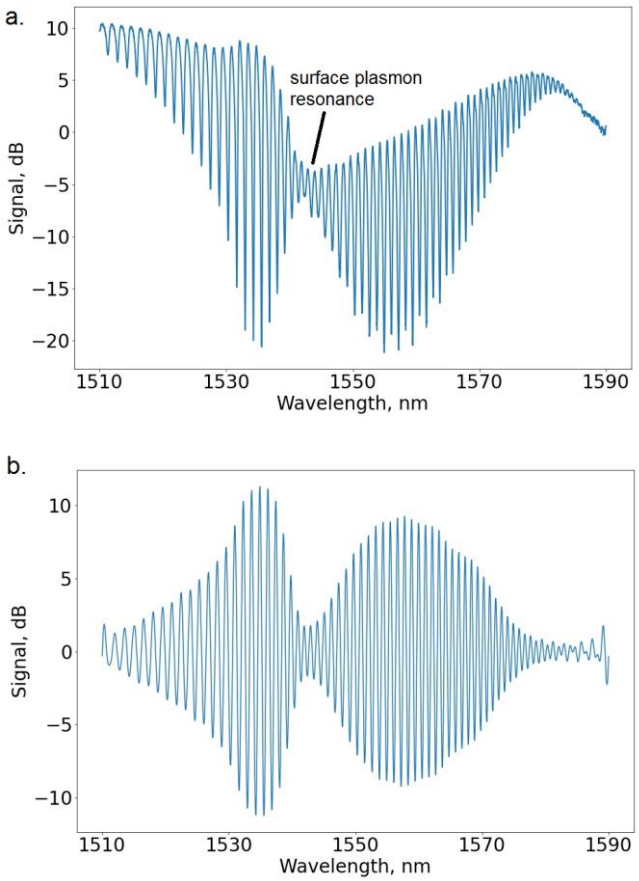

Figure 2. The example of the scattering spectrum.

To find parameters of the ML algorithm, i.e. to train, and check its accuracy, a set of the measured values of refractive index was randomly split into 2 sets in ratio of 4:1. The largest set named a training set was used to find the algorithm parameters the ML algorithm. The small set named test set was used to estimate the difference between the refractive index predicted by the ML algorithm and the refractive index calculated analytically.

Since the number of spectra is two orders of magnitude smaller than the number of the values in each spectrum in training set, the dimensionality reduction was applied. Otherwise, the number of parameters would be very high. Because of large number of parameters, ML algorithm can learn noise and demonstrate high error for test set despite the low error on the training set. To reduce the dimensionality of the spectrum, we used the principal component analysis (PCA). PCA allows approximating each spectrum with a linear combination of vectors named principal components:

$$S_k \approx \langle S \rangle + \sum_{i=1}^{q} a_{ki} P_i, \quad (1)$$

where a_{ki} are the decomposition components, P_i is the principal components, q is the number of principal components, $\langle S \rangle = \sum_{k=1}^{N} S_k / N$ is the averaged spectrum, and N is the number of spectra in the training set. The error of the approximation with q principal components, P_i, is minimal if P_i are the spectra correlation matrix eigenvectors corresponding to its q largest eigenvalues (Shalev-Shwartz & Ben-David, 2014). The elements of the correlation matrix are defined by the equation

$$C_{ij} = \frac{1}{N-1} \sum_{k=1}^{N} \left(S_k(\lambda_i) - \langle S \rangle(\lambda_i) \right) \left(S_k(\lambda_j) - \langle S \rangle(\lambda_j) \right). \quad (2)$$

To determine q, we calculate the eigenvalues of the correlation matrix (Fig. 3).

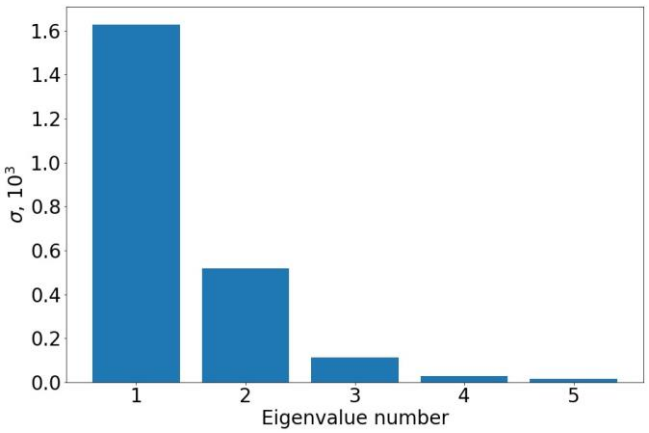

Figure 3. The histogram of the eigenvalues.

In Fig.4, the principal components of the training set are shown. Since the correlation matrix is symmetrical, the principal components are orthogonal. Therefore, the coefficients a_{ki} can be calculated via the equation $a_{ki} = \left(P_i, S_k - \langle S \rangle \right)$ at $\|P_i\| = 1$. The coefficients a_{ki} are used as new variables instead of the spectral intensities, so that the number of variables equals 3 instead of 16001.

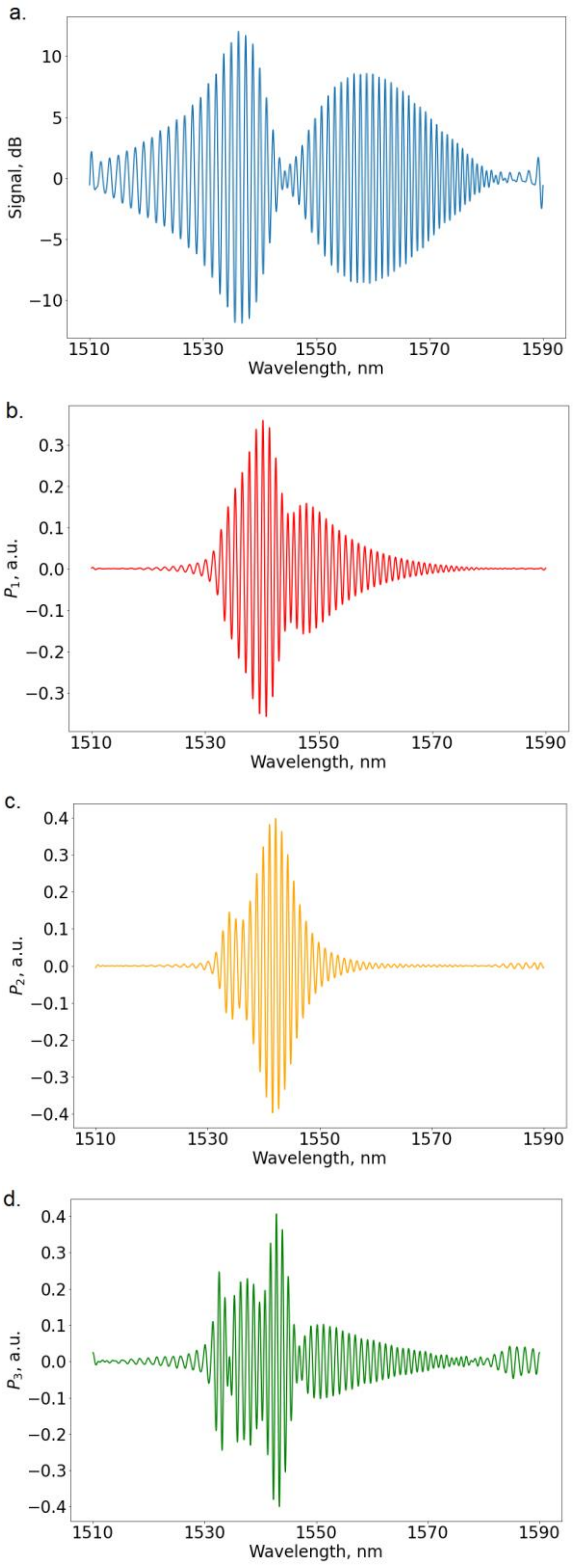

Figure 4. The averaged spectrum (a) and the first (b), second (c), and third (d) principal components corresponding to the 3 maximum eigenvalues.

To find the value of the refractive index from the principal component coefficients, we use polynomial regression. The degree of polynomial was equal to 7. The number of coefficients is equal to 120. The

coefficients of the polynomial are derived from the minimization of squared error with L_1 -regularization term:

$$E(\mathbf{c},\lambda)=\sum_{k=1}^{N}\left(n_k - f\left(\mathbf{a}_k,\mathbf{c}\right)\right)^2+\lambda\sum_{i=1}^{p}|c_i|\,. \qquad (3)$$

The regularization term $\lambda\sum_{i=1}^{p}|c_i|$ in Eq. (3) shrinks the majority of coefficients towards zero and does not let the model to learn noise.

2.2. Hydrogen dielectric Bragg grating sensor

Another sensor we consider is a dielectric Bragg grating made of tungsten trioxide (WO_3), which permittivity is sensitive to change in hydrogen concentration (Fig. 5). The Bragg grating structure creates a Wood anomaly in the transmission spectrum. The change of hydrogen concentration leads to the change of the frequency, width and depth of the Wood anomaly resonance. By measuring the transmittance spectra, the hydrogen concentration can be retrieved.

During the experiment with the sensor, the hydrogen concentration assumed values of 100 ppm and 1000 ppm. The number of measured spectra was 1538. A typical spectrum is shown in Fig. 6. The Wood anomaly resonant dip is located at the wavelength of 673 nm (Fig. 6)

A serious issue with the hydrogen sensor is the noise. The approximation of the resonant dip with asymmetric Fano resonance profile did not allow obtaining noise-free dependencies of the linewidth and the resonant wavelength on the hydrogen concentration. To reduce the level of noise, we apply PCA to the spectra around the Wood anomaly resonance.

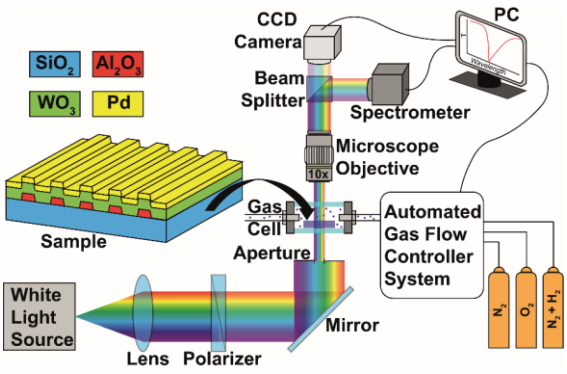

Figure 5. Optical scheme for testing gas sensing elements.

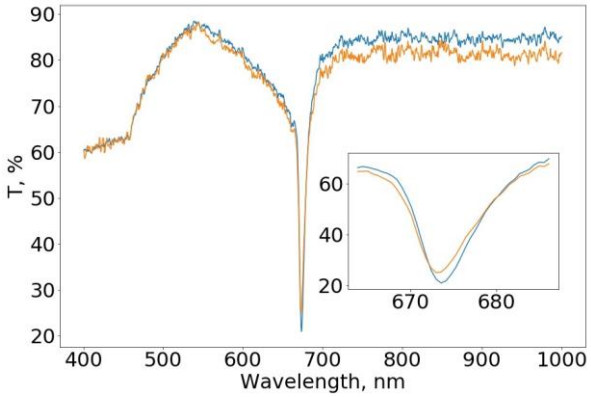

Figure 6. Examples of the transmittance spectrum. The resonance dip is shown in the inset.

Because of the high level of noise in non-resonant parts of spectra, we consider only resonant wavelength range. To process spectra, we apply PCA to resonant dip. The histogram of the correlation matrix eigenvalues is shown in Fig. 7.

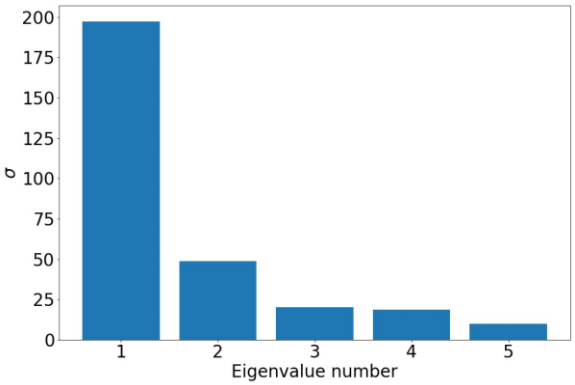

Figure 7. The eigenvalue histogram.

Since the only one eigenvalue is significantly larger than the others, it is enough to take only one principal component P_1 into account and the decomposition coefficient $a_1=\left(P_1,T-\langle T\rangle\right)$, where $\langle T\rangle=\sum_{i=1}^{N}T_i/N$ is the average transmission spectrum, where $N=1538$ is the number of spectra. $\langle T\rangle$ and P_1 are shown in Fig. 8.

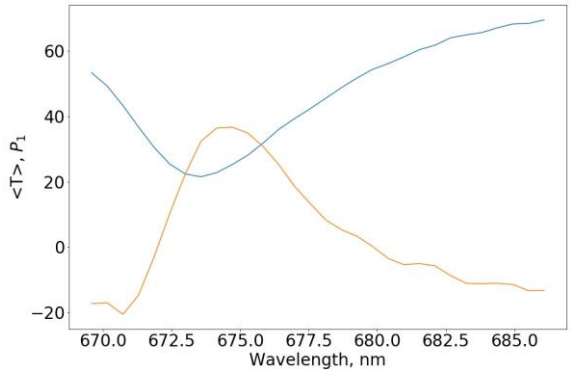

Figure 8. The averaged spectrum (blue)and the first principal component (orange).

3. Results and Discussion

In the case of TFBG sensor, we apply algorithm trained on the training set to the test set. Due to the regularization term in Eq. (3), only 7 out of 120 coefficients are not equal to zero. We have found that the refractive index change mean absolute error for the test set is about $2*10^{-5}$ RIU (Fig. 9) and is the same order of magnitude as the error on the training set.

In the case of the hydrogen sensor, PCA applied to the transmittance spectra leads to reduction of noise amplitude. In Fig. 10, we compare the value of decomposition coefficient for the first principal component *a* and the values of the transmittance at single Wood anomaly wavelength.

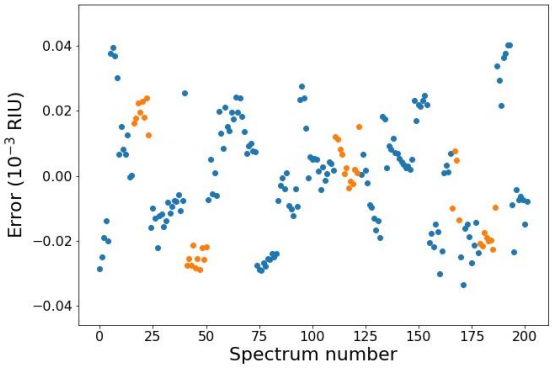

Figure 9. The difference between the analytically calculated values of refractive index and the refractive index predicted by polynomial regression. Blue and orange dots correspond to training and test set, respectively.

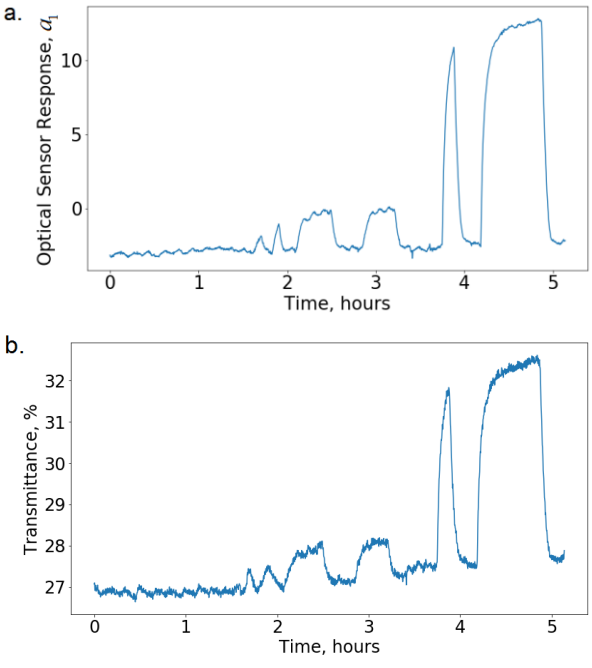

Figure 10. (a) The decomposition coefficient for the first principal component and (b) the transmittance at the Wood anomaly frequency vs time.

4. Conclusions

Optical sensors provide a complex response as compared to classical (resistive, etc.) ones. Namely, they give a whole spectrum instead of a single number. This makes machine learning (ML) methods very perspective in this field. We have shown that ML algorithms offer a great advantage in optical sensing. Using our experimental data from two different optical sensors, we demonstrate both enhancement of sensitivity and increase in the signal-to-noise ratio.

Funding

This work was supported by RFBR through the research project N 20-02-00504.

References

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician, 46*(3), 175-185.

Askim, J. R., Mahmoudi, M., & Suslick, K. S. (2013). Optical sensor arrays for chemical sensing: the optoelectronic nose. *Chemical Society Reviews, 42*(22), 8649-8682.

Chubchev, E., Nechepurenko, I., Dorofeenko, A., Vinogradov, A., & Lisyansky, A. (2020). Nanostructured optical waveguide with a highly confined mode. *JOSA B, 37*(8).

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*: MIT press.

Gorban, A. N., Kégl, B., Wunsch, D. C., & Zinovyev, A. Y. (2008). *Principal manifolds for data visualization and dimension reduction* (Vol. 58): Springer.

Hotel, O., Poli, J.-P., Mer-Calfati, C., Scorsone, E., & Saada, S. (2018). A review of algorithms for SAW sensors e-nose based volatile compound identification. *Sensors and Actuators B: Chemical, 255*, 2472-2482.

Ignatov, A. I., & Merzlikin, A. M. (2020). Two optical sensing elements for H2O and NO2 gas sensing based on the single plasmonic–photonic crystal slab. *Advanced Optical Technologies, 1*(ahead-of-print).

Johnson, S. R., Sutter, J. M., Engelhardt, H. L., Jurs, P. C., White, J., Kauer, J. S., . . . Walt, D. R. (1997). Identification of multiple analytes using an optical sensor array and pattern recognition neural networks. *Analytical Chemistry, 69*(22), 4641-4648.

Kornienko, V. V., Nechepurenko, I. A., Tananaev, P. N., Chubchev, E. D., Baburin, A. S., Echeistov, V. V., . . . Rodionov, I. A. (2020). Machine Learning for Optical Gas Sensing: A Leaky-Mode Humidity Sensor as Example. *IEEE Sensors Journal, 20*(13),

6954-6963.

Liu, T., Li, D., Chen, J., Chen, Y., Yang, T., & Cao, J. (2019). Active Learning on Dynamic Clustering for Drift Compensation in an Electronic Nose System. *Sensors, 19*(16), 3601.

Ma, Z., Luo, G., Qin, K., Wang, N., & Niu, W. (2018). Online sensor drift compensation for E-nose systems using domain adaptation and extreme learning machine. *Sensors, 18*(3), 742.

Nechepurenko, I., Andrianov, E., Zyablovsky, A., Dorofeenko, A., Pukhov, A., & Lozovik, Y. E. (2018). Absorption sensor based on graphene plasmon quantum amplifier. *Physical Review B, 98*(7), 075411.

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*: Cambridge university press.

Tomyshev, K., Manuilovich, E., Tazhetdinova, D., Dolzhenko, E., & Butov, O. (2020). High-precision data analysis for TFBG-assisted Refractometer. *Sensors and Actuators A: Physical*, 112016.

Wang, Y., Yang, A., Chen, X., Wang, P., Wang, Y., & Yang, H. (2017). A deep learning approach for blind drift calibration of sensor networks. *IEEE Sensors Journal, 17*(13), 4158-4171.

Zhao, X., Li, P., Xiao, K., Meng, X., Han, L., & Yu, C. (2019). Sensor Drift Compensation Based on the Improved LSTM and SVM Multi-Class Ensemble Learning Models. *Sensors, 19*(18), 3844.