УДК 544.6

IDENTIFICATION OF COMPONENTS OF COMPLEX MIXTURES USING MACHINE LEARNING METHODS

Belyaev V.E. (ITMO University), **Aliev T.A. (**ITMO University)
Domareva N.P. (ITMO University) **Petrova M.S.** (ITMO University)**.**
Research supervisor – Prof. Dr. Skorb E.V.
(ITMO University)

Introduction. Currently, the issue of quality control of manufactured products is acute in the food industry. Due to the constant development of this industry and the introduction of new techniques into production, it is necessary to monitor the products obtained for compliance with state standards. For example, whether the content of heavy metal ions in goods is not exceeded, or the content of fermented milk bacteria in dairy products. The one of these aspects that requires attention is the control of the antibiotics content in milk. We propose a new system that will allow determining the concentration of antibiotics in milk on the farm, which will reduce the number of defective products, which will save money for producers and consumers' health.

Main part. At the initial stage, various antibiotics were selected, such as cefazolin, ceftiofur, penicillin, streptomycin, tetracycline. Our research was aimed not only at qualitative analysis of the mixture, i.e., determining the presence of an antibiotic in milk, but also at quantitative analysis, which allows us to talk about the concentration of this antibiotic in the sample under study. To perform this task, at the data collection stage, we prepared samples of various concentrations, namely, the maximum permissible concentration of these antibiotics in milk, as well as values that differ by 10 and 100 times in greater and lesser sides. 20 experiments were conducted for each concentration, the total amount of data for one antibiotic was 100 output values.

The values obtained from a database that we use to create a model using the Python programming language to determine antibiotics in milk. Since we are faced with the task of classification, it was decided to use the gradient boosting method, namely the framework – Cat Boost. Various libraries, such as NumPy and Pandas, were used for initial processing and sorting of the source data. To test the results of machine learning, the source data was divided into two groups: 80% of the source data was used for training and the remaining 20% were used directly for tests.

Conclusions. Analyzing the result of machine learning, we can say that the created program analyzes the presence of an antibiotic in milk with an accuracy of at least 88%. This suggests that the basic scheme of the experiment works, and this technique can be used to analyze other components of complex mixtures, for example, the analysis of hormones in the same milk.

Literature:

1. Tuerk J. et al. Analysis of antibiotics in urine and wipe samples from environmental and biological monitoring—comparison of HPLC with UV-, single MS-and tandem MS-detection //Journal of Chromatography B. – 2006. – Т. 831. – №. 1-2. – С. 72-80.
2. Mistry A. et al. How machine learning will revolutionize electrochemical sciences //ACS energy letters. – 2021. – Т. 6. – №. 4. – С. 1422-1431.