**Сравнение подходов к улучшению репрезентативности спектроскопических данных с использованием вариационных автоэнкодеров**

***Мущина А.С.***

*Студент, 2 курс магистратуры,*

*физический факультет, Московский государственный университет*

*имени М.В. Ломоносова, Москва, Россия*

*E–mail:* [*anastasemusa@gmail.com*](mailto:anastasemusa@gmail.com)

В настоящей работе рассматривается обратная задача спектроскопии многокомпонентных водных растворов, направленная на определение концентраций различных ионов по спектральным данным (спектры комбинационного рассеяния, инфракрасного или оптического поглощения) [3]. Зависимость формы спектра от состава многокомпонентного раствора и концентраций ионов имеет сложный характер, и требуется анализ множества спектральных каналов одновременно. Такой анализ может быть выполнен с использованием методов машинного обучения, например, нейронных сетей. Однако эффективность этого подхода во многом определяется объемом и репрезентативностью доступных экспериментальных данных.

Мы исследуем возможность расширения обучающей выборки для регрессионной нейросети путем генерации дополнительных искусственных спектров с помощью вариационного автоэнкодера (ВАЭ) [2], что может способствовать повышению репрезентативности данных и, как следствие, снижению ошибки решения обратной задачи. Для такой генерации могут быть использованы различные подходы.

Применение обусловленного вариационного автоэнкодера (оВАЭ), обученного на экспериментальных данных, требует выбора стратегии подбора соответствующих наборов концентраций ионов для генерации спектров [1]. Подходы с использованием ВАЭ предполагают определение целевых концентраций ионов для искусственных спектров, что может быть выполнено с помощью регрессионной нейросети. Сгенерированные при помощи ВАЭ и оВАЭ спектры могут быть использованы различными способами совместно с экспериментальными данными при обучении регрессионных нейронных сетей, решающих обратную задачу. В данной работе проводится сравнительный анализ этих методов, а также обсуждаются их достоинства и ограничения.

**Литература**

1. A. Efitorov, T. Dolenko, K. Laptinskiy, S. Burikov, S. Dolenko. Use of Conditional Generative Variational Autoencoder Networks to Improve Representativity of Data in Optical Spectroscopy. Proceedings of Science, 410, art. 013 (2021). <https://doi.org/10.22323/1.410.0013>
2. D. P. Kingma, M. Welling, An Introduction to Variational Autoencoders, Foundations and Trends in Machine Learning: Vol. 12 (2019): No. 4, pp 307-392. <http://dx.doi.org/10.1561/2200000056>
3. Maurya, V. K., Singh, R. P., and Prasad, L. B., Comparative evaluation of trace heavy metal ions in water sample using complexes of dithioligands by flame atomic absorption spectrometry, Oriental Journal of Chemistry 34(1), 100-109 (2018). <http://dx.doi.org/10.13005/ojc/340111>