**Модели машинного обучения для прогнозирования физико-химических свойств глубоких эвтектических растворителей**

***Одегова В.С., Лавриненко А.К.***

*Студент 1 курса магистратуры*

*Национальный исследовательский университет ИТМО, Санкт-Петербург, Россия*

*E-mail:* *odegova.vs@gmail.com*

Глубокие эвтектические растворители (ГЭР), представляют собой двух- или трехкомпонентные смеси, которые имеют пониженную температуру плавления по сравнению с исходными веществами [1]. Чаще всего ГЭР состоят из природных компонентов, поэтому они биодеградируемы, а их эффективность экстракции сопоставима с известными органическими растворителями. Поэтому ГЭР рассматриваются как перспективные вещества для применения в различных областях химии, фармацевтики, металлургии, нефтяной и пищевой промышленности [1]. Из-за большого количества индивидуальных компонентов и их сочетаний выбор смеси с оптимальными свойствами для конкретной задачи эмпирическим путем очень долог и часто не приводит к оптимальному результату [2]. В своей работе мы предлагаем использование методов машинного обучения (МО) для прогнозирования физико-химических свойств ГЭР.

В литературе существуют примеры предсказания некоторых физико-химических свойства ГЭР, однако они применимы только для узкого круга бинарных ГЭР с определенной структурой [2]. В данной работе мы предсказывали плотность, так как она является ключевым технологическим параметром, а также влияет на эффективность экстракции [2]. Была собрана уникальная база данных по плотностям ГЭР, содержащая 4000 двух- и трехкомпонентных систем различных типов, классифицируемых по структуре исходных веществ [1] (рис. 1А). Из-за несбалансированности базы данных относительно некоторых типов ГЭР (рис. 1А) при разработке модели были использованы веса, для повышения важности малочисленных типов.

|  |  |
| --- | --- |
|  |  |
| **Рис. 1А.** Диаграмма распределения ГЭР с различной структурой в собранной базе данных | **Рис. 1Б.** *Метрика R2(доля дисперсии) для классических моделей МО*  |

Для построения моделей МО были использованы параметры (дескрипторы), которые описывают геометрические свойства молекул, количества структурных групп и связей, а также мольная доля и температура, при которой была измерена плотность. Нами была разработана модель на основе классических алгоритмов машинного обучения, которая способна предсказывать плотность бинарных и третичных ГЭР различного состава с точностью на кросс-валидации, учитывающей уникальность смеси, R2 = 0.85 (рис 1Б).

**Литература**

[1] Hansen B.B. et al. Deep Eutectic Solvents: A Review of Fundamentals and Applications // Chem. Rev. 2021. Vol. 121, № 3. P. 1232–1285.

[2] Abdollahzadeh, M., et al. Estimating the density of deep eutectic solvents applying supervised machine learning techniques. Sci Rep 2022 12, 4954