

УДК 665.6: 51-7

**МОДЕЛИРОВАНИЕ ПРОЦЕССА ТВЕРДОКИСЛОТНОГО
АЛКИЛИРОВАНИЯ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ
МАШИННОГО ОБУЧЕНИЯ**

**MODELING OF THE PROCESS OF SOLID ACID ALKILATION
USING MACHINE LEARNING ALGORITHMS**

**Е.А. Симановский, К.Ю. Устюжанин, Н.А. Руднев,
Р.Р. Шириязданов, С.Р. Хафизова, Т.В. Смольникова**

**Уфимский государственный нефтяной технический университет,
г. Уфа, Российская Федерация**

**Evgeny A. Simanovsky, Konstantin Y. Ustyuzhanin,
Nikolay A. Rudnev, Rishat R. Shiryazdanov,
Svetlana R. Khafizova, Tatyana V. Smolnikova**

Ufa State Petroleum Technological University, Ufa, Russian Federation

e-mail: ngb2008@mail.ru

Аннотация. В настоящее время в большинстве отраслей идет модификация и оптимизация производственных процессов и бизнес-процессов, они становятся более эффективными, экологичными и безопасными. Вместе с улучшением технологии в промышленности широко используются методы, позволяющие повысить эффективность производства с помощью информационных технологий. Эти тенденции называют цифровой трансформацией или цифровизацией производства. Цифровизация производства напрямую связана с построением математических моделей процессов или аппаратов, объединением их в общую автоматизированную саморегулирующуюся систему, которая позволит управлять производством.

В работе рассмотрено построение моделей алкилирования изобутана бутан-бутиленовой фракцией на цеолитсодержащих катализаторах с применением нейронных сетей в виде многослойного персептрона и сетей LSTM. Построенные модели в зависимости от параметров процесса (температуры, катализатора, процентного содержания бутиленов в сырье и т.д.) предсказывают состав продукта.

Для поиска модели, которая наиболее точно опишет процесс, был проведен ряд экспериментов, отражающий влияние таких факторов, как нормирование, количество эпох обучения, количество слоев и количество нейронов в слое.

Было проведено сравнение точности расчета, времени и удобства использования построенных моделей. Надо отметить, что точность и скорость построения моделей существенно зависят от архитектуры использованной нейронной сети. И поскольку пока нет строгих правил и зависимостей архитектуры сети от вида решаемых задач, то точность и качество зависят от опыта исследователя и поставленной задач.

Построенные в ходе работы модели алкилирования изобутана олефинами описывают процесс в статике с высокой точностью и не требуют создания сложных физико-химических моделей.

Было выявлено, что при повышении эпох обучения, точность предсказания возрастает, но существует риск переобучения модели. Также точность зависит от правильного подбора и расположения слоев и количества нейронов в них. Лучшая точность предсказания состава продукта была достигнута при использовании нейронных сетей, погрешность не превышала 1,71 %, что позволяет сделать вывод, о том, что их использование предпочтительно для построения подобных анализаторов качества.

Abstract. Currently, most industries are modifying and optimizing production processes and business processes, they are becoming more efficient, environmentally friendly and safe. Along with the improvement of technology

in industry, methods are widely used to increase the efficiency of production using information technology. These trends are called digital transformation or digitalization of production. Digitalization of production is directly related to the construction of mathematical models of processes or apparatuses, their combination into a common automated self-regulating system, which will allow to control production.

The paper discusses the construction of models of alkylation of isobutane with butane-butylene fraction on zeolite-containing catalysts using neural networks in shape of multilayer perceptrone and LSTM. Constructed models, depending on the process parameters (temperature, catalyst, the content of butylenes in raw material, etc.) predict the composition of the product.

A series of experiments were conducted to find the model that most accurately describes the process. Described experiments represents the influence of parameters like normalizing, number of training epochs, number of layers and number of neurons per layer.

A comparison of the calculation accuracy, time and usability of the constructed models was carried out. It should be noted that the accuracy and speed of building models significantly depends on the architecture of the neural network used. And since there are no strict rules and dependencies of the network architecture on the type of tasks to be solved, accuracy and quality depend on the experience of the researcher and is a creative task.

The models of isobutane alkylation with olefins constructed during the work describe the process in statics with high accuracy and do not require the creation of complex physico-chemical models.

It was found that with increasing epochs of learning, the accuracy of prediction improves, but there is a risk of retraining the model. Also, accuracy depends on the correct selection and location of layers and the number of neurons in them. The best accuracy in predicting the composition of the product was achieved using neural networks, the error did not exceed 1.71 %, which

allows us to conclude that their use is preferable for constructing such quality analyzers.

Ключевые слова: алкилирование; машинное обучение; нейронные сети; цеолиты; математическая модель; бутан-бутиленовая фракция

Key words: alkylation; machine learning; neural networks; zeolites; mathematical model; butane-butylene fraction

Введение

В настоящее время во всем мире идет модификация и оптимизация производственных процессов, они становятся более эффективными, экологичными и безопасными. Для решения задач оптимизации и управления производством все чаще применяются математические модели, построенные на основе современных алгоритмов машинного обучения. Наряду с развитием технологии в промышленности широко используются методы, позволяющие повысить эффективность производства с помощью информационных технологий. Такие тенденции называют цифровой трансформацией или цифровизацией производства. Цифровизация химического производства напрямую связана с построением математических моделей процессов или аппаратов, объединением их в общую автоматизированную саморегулирующуюся систему, которая позволит управлять производством [1–4].

Одним из перспективных способов получения высокооктанового компонента бензинов является процесс алкилирования изобутана олефинами с возможностью перевода на гетерогенные катализаторы. Этот процесс еще не реализован в промышленном масштабе. Проведенные ранее серии опытов [5] позволили накопить обширную базу экспериментальных данных. Переход от экспериментальных и опытно-промышленных установок к крупнотоннажному производству проходит

более эффективно благодаря созданию теоретической базы и математической модели процесса.

Существуют различные подходы к созданию математических моделей или цифровых двойников: физико-химический, черный ящик (методы математической статистики и машинного обучения), гибридный.

Инструментом для моделирования процесса алкилирования изобутана бутан-бутиленовой фракцией выбраны нейронные сети, это современные и перспективные способы моделирования, которые доказали свою целесообразность уже во многих отраслях промышленности.

В исследовании было проведено сравнение нескольких алгоритмов по скорости построения модели, точности и сложности создания.

Исследовательская часть

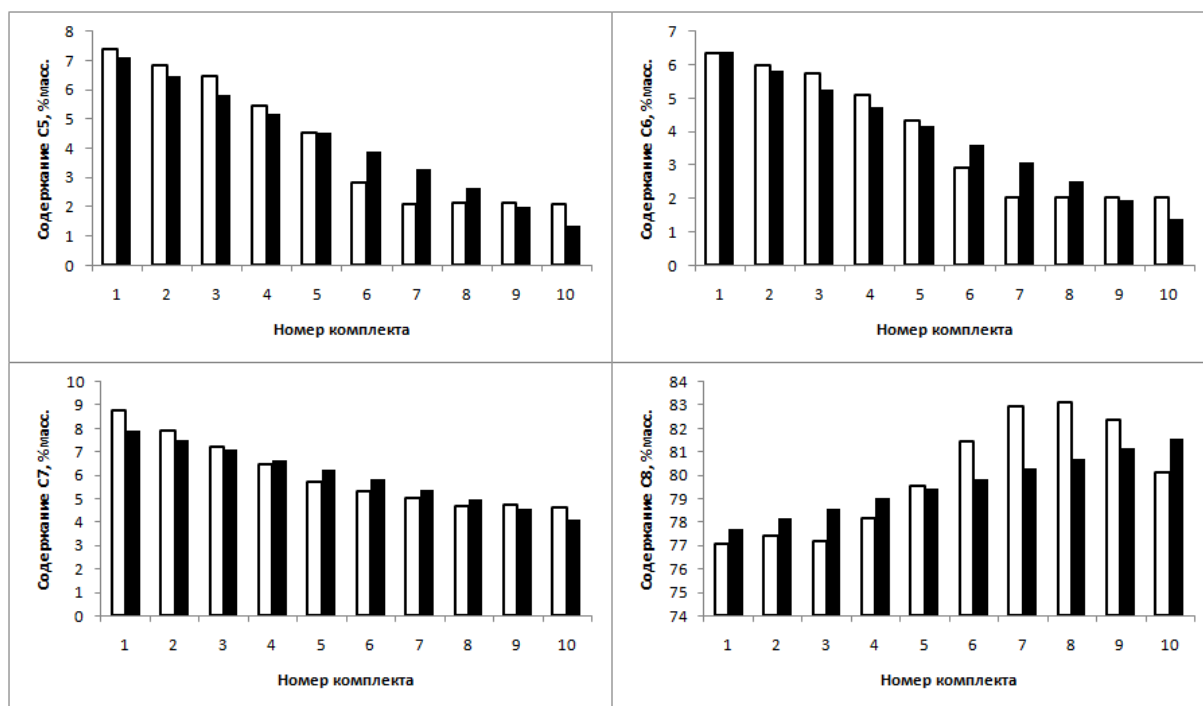
В качестве входных параметров были взяты - время работы катализатора, массовая концентрация бутиленов в сырье, температура процесса. В качестве выходных - массовая концентрация компонентов C₅, C₆, C₇, C₈ в продукте [6].

Исходными данными для построения модели являлись технологические измерения и результаты лабораторных анализов, полученные на пилотной установке. Для построения нейросетевых моделей дополнительное количество данных было получено интерполяцией экспериментальных данных. Все исходные данные были проверены на модели, описывающей физико-химическую модель алкилирования изобутана бутан-бутиленовой фракцией, разработанной на кафедре «Газохимия и моделирование химико-технологических процессов» УГНТУ. В таблице 1 представлена часть истинных значений параметров выходной выборки.

Таблица 1. Истинные выходные значения

№ комплекта	Концентрация, % масс.			
	C ₅	C ₆	C ₇	C ₈
1	7,3813	6,3627	8,7819	77,0513
2	6,8211	5,9686	7,9221	77,4126
3	6,4545	5,7499	7,2089	77,1529
4	5,4384	5,0812	6,4632	78,1781
5	4,5149	4,3217	5,7201	79,5571
6	2,8029	2,9197	5,3302	81,4328
7	2,1035	2,0199	5,0569	82,9629
8	2,1069	2,0231	4,7036	83,1232
9	2,1062	2,0215	4,7211	82,4055
10	2,0973	2,0088	4,6519	80,1173

Изначально была построена линейная регрессионная модель, отражающая зависимость концентраций C₅, C₆, C₇, C₈ от количества бутиленов в сырье, температуры и времени проведения процесса. На рисунке 1 представлены результаты этой модели.



«□» – истинные значения; «■» – предсказанные моделью

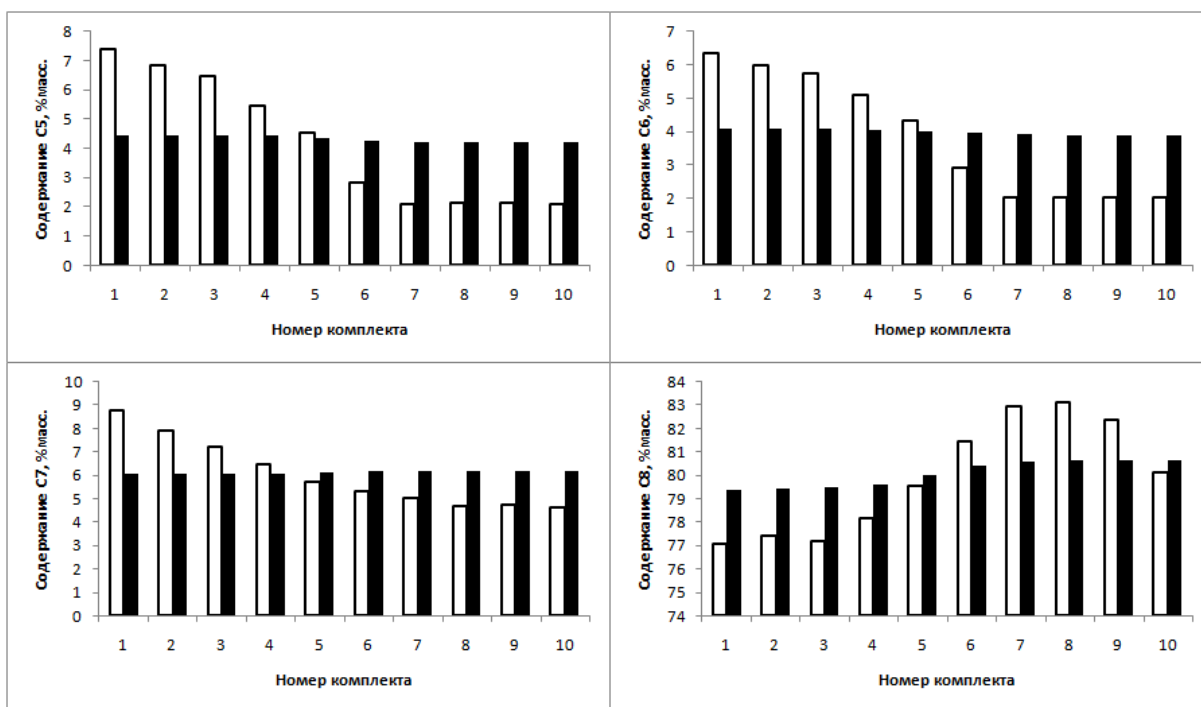
Рисунок 1. Сравнение истинных и предсказанных регрессионной моделью выходных значений концентрации изопарафинов C₅, C₆, C₇, C₈

Далее была разработана простейшая нейронная сеть прямого распространения с тремя скрытыми слоями; архитектура сети № 1 представлена в таблице 2. Обучающие данные подавались без нормирования.

Результат работы нейронной сети с архитектурой из таблицы 2 и количеством эпох обучения, равным 500, представлены на рисунке 2. За одну эпоху обучения через нейронную сеть полностью проходит обучающая выборка.

Таблица 2. Архитектура нейронной сети № 1

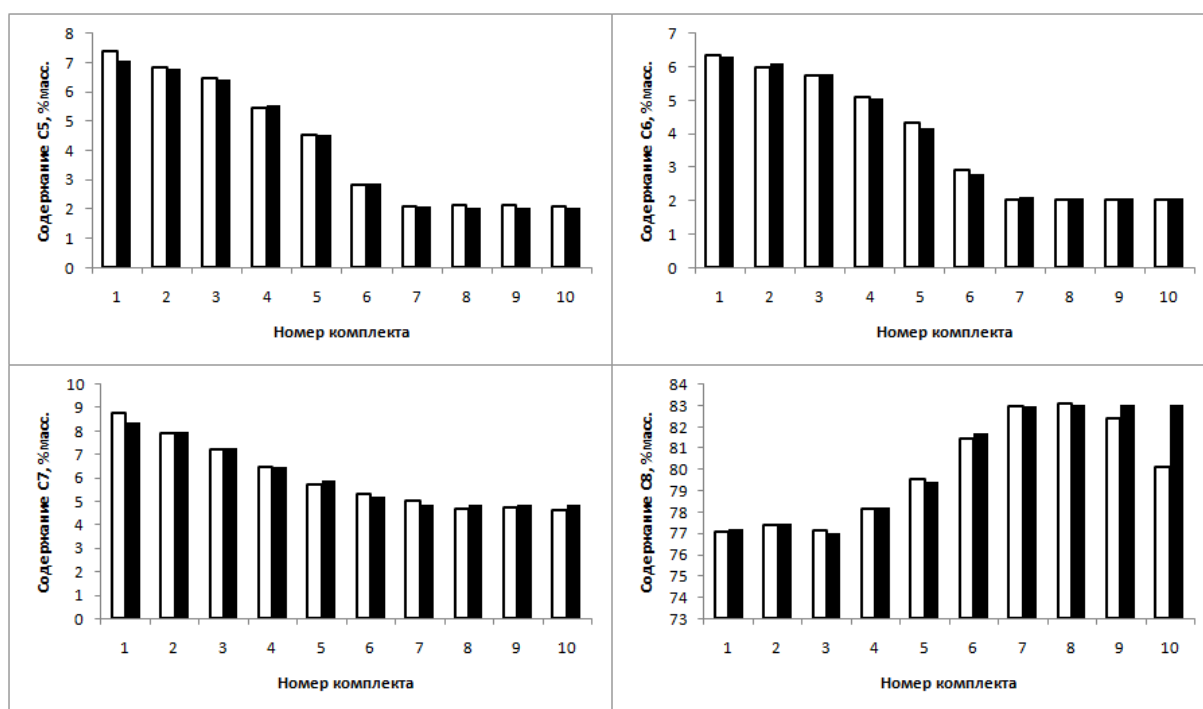
Номер слоя	Количество нейронов	Функция активации
1	8	—
2	40	сигмоидальная
3	20	сигмоидальная
4	10	линейная
5	4	—



«□» – истинные значения; «■» – предсказанные моделью

Рисунок 2. Сравнение истинных и предсказанных нейронной сетью № 1.1 выходных значений концентрации изопарафинов C₅, C₆, C₇, C₈

После этого число эпох при обучении было увеличено до 2000 (рисунок 3).



«□» – истинные значения; «■» – предсказанные моделью

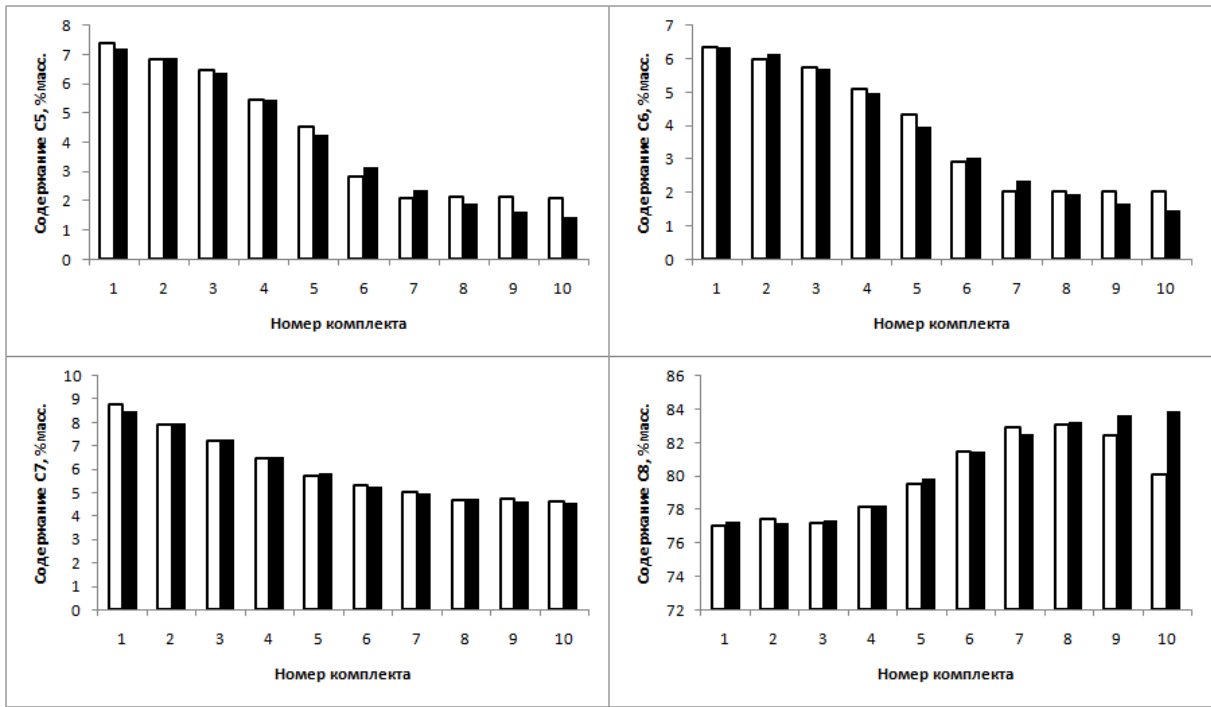
Рисунок 3. Сравнение истинных и предсказанных нейронной сетью №1.2 выходных значений концентрации изопарафинов C₅, C₆, C₇, C₈

Из-за повышения количества эпох сильно увеличивается время обучения, а также это может приводить к переобучению [5].

Нейронную сеть с архитектурой, описанной в таблице 2, и с количеством эпох обучения, равным 500, обучим еще раз, но уже с нормированием параметров. Результаты представлены на рисунке 4.

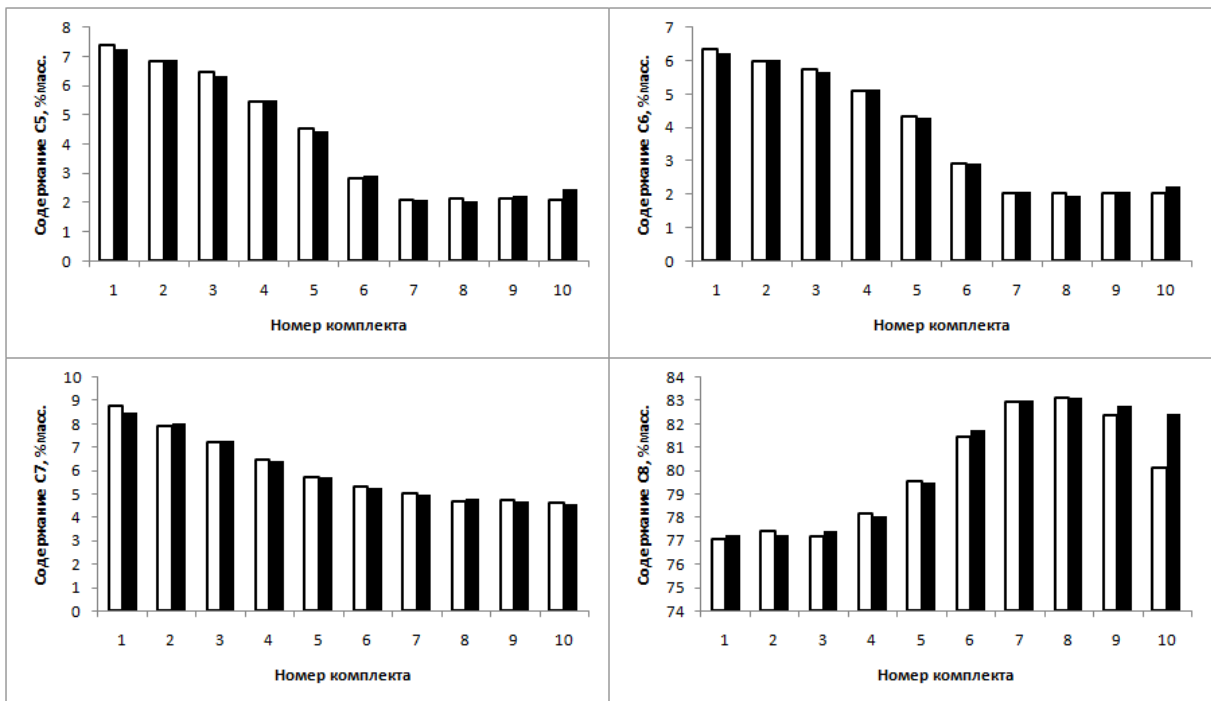
Результат значительно улучшился по сравнению с обучением без нормирования данных. Повысим количество эпох до 2000 и построим модель еще раз. Результаты представлены на рисунке 5.

При увеличении количества эпох нейронная сеть может переобучиться. Для проверки построим таблицу 3, где представлена зависимость точности от количества эпох обучения.



«□» – истинные значения; «■» – предсказанные моделью

Рисунок 4. Сравнение истинных и предсказанных нейронной сетью №1.3 выходных значений концентрации изопарафинов C₅, C₆, C₇, C₈



«□» – истинные значения; «■» – предсказанные моделью

Рисунок 5. Сравнение истинных и предсказанных нейронной сетью №1.4 выходных значений концентрации изопарафинов C₅, C₆, C₇, C₈

Таблица 3. Зависимость точности от количества эпох

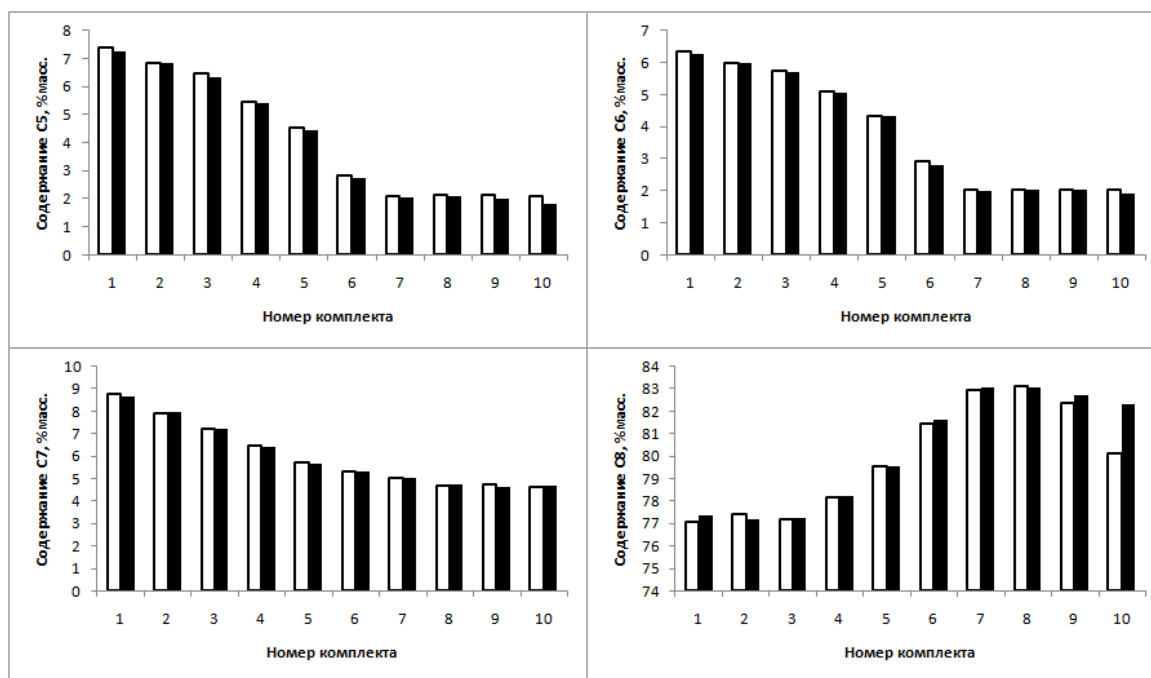
Количество эпох	500	2000	3000	5000	10000	20000
Средняя квадратичная ошибка	0,5094	0,1560	0,1973	0,1935	0,1673	0,0741

Архитектура нейронной сети была усложнена, добавлены новые слои. Новая архитектура сети № 2 представлена в таблице 4.

Таблица 4. Архитектура нейронной сети № 2

Номер слоя	Количество нейронов	Функция активации
1	8	—
2	40	сигмоидальная
3	80	сигмоидальная
4	40	линейная
5	20	сигмоидальная
6	10	линейная
7	4	—

На рисунке 6 представлена модель с архитектурой, приведенной в таблице 4, полученная при обучении с количеством эпох, равным 5000.



«□» – истинные значения; «■» – предсказанные моделью

Рисунок 6. Сравнение истинных и предсказанных нейронной сетью №1.5 выходных значений концентрации изопарафинов C₅, C₆, C₇, C₈

Уже давно существуют и часто используются рекуррентные нейронные сети. При обучении модели они учитывают не только значение поступивших параметров, но и свое прошлое значение, таким образом, можно сказать, что данные слои подходят для построения динамических моделей. На рисунке 7 представлена модель рекуррентного нейрона.

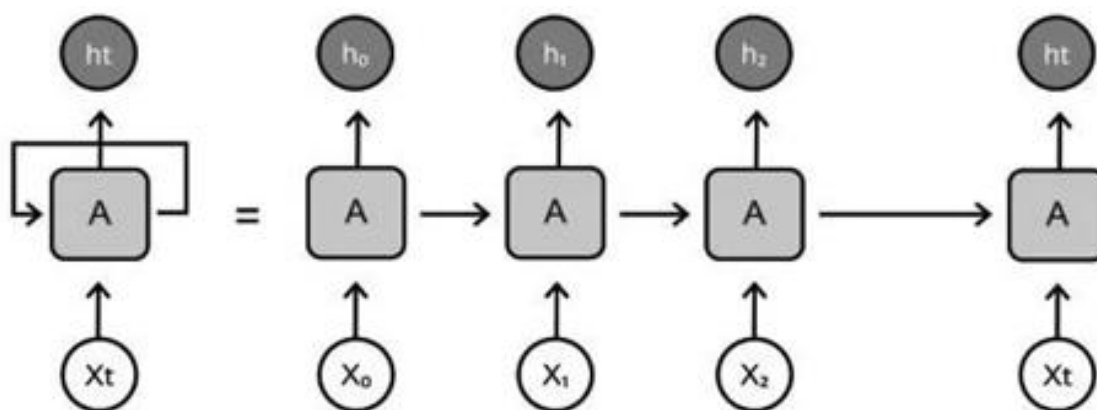


Рисунок 7. Модель рекуррентного нейрона

Наиболее часто при построении рекуррентных сетей используются RNN и LSTM-слои [7–9], обладающие долгосрочной памятью. Построим нейронную сеть с использованием LSTM-слоев.

Построенная LSTM-сеть работает не совсем корректно, это может быть связано либо с малым количеством данных, либо с тем, что модель статическая, то есть экспериментальные данные не отражают точное поведение процесса во времени и динамику процесса. Можно сделать вывод, что такая сеть не подходит для решаемой нами задачи.

Сравнение всех моделей приведено в таблице 5.

Процесс алкилирования изобутана бутан-бутиленовой фракцией на твердом катализаторе наиболее эффективно протекает при температуре 70 °С; массовой концентрации бутиленов в сырье 0,01; времени работы катализатора 30 ч. Сравнительная характеристика предсказания концентраций компонентов разными моделями представлена в таблице 6.

Таблица 5. Сравнительная таблица моделей

Архитектура	Нормирование	Количество эпох обучения	Средняя квадратичная ошибка	Время, с
Регрессионная	–	–	0,7811	0,0010
Архитектура № 1	нет	500	3,0412	2,6438
Архитектура № 1	нет	2000	0,2412	9,1333
Архитектура № 1	есть	500	0,5094	3,1488
Архитектура № 1	есть	2000	0,1560	10,3226
Архитектура № 1	есть	20000	0,0777	99,0260
Архитектура № 2	есть	500	0,2505	3,9834
Архитектура № 2	есть	2000	0,1357	12,0123
Архитектура № 2	есть	20000	0,0617	126,8060
LSTM-сеть	есть	500	3,6861	13,4654

Как видно из таблицы 6, наиболее точно предсказанные значения были получены при использовании нейронной сети с архитектурой, описанной в таблице 4, а все полученные значения концентраций близки к истинным значениям, ошибка составила не более 2 %.

Таблица 6. Сравнение точности моделей

Компонент	C ₅	C ₆	C ₇	C ₈
Истинные значения	2,0997	2,0283	4,8423	82,2619
Регрессионная модель	4,2560	3,9118	6,0599	79,6498
% ошибки	102,70	92,86	25,15	3,18
Архитектура № 1	2,0643	1,9606	4,8103	81,9154
% ошибки	1,69	3,34	0,66	0,42
Архитектура № 2	2,0934	2,0450	4,7593	82,441
% ошибки	0,30	0,82	1,71	0,22

Модель с архитектурой, представленной в таблице 2, также предсказала требуемые концентрации достаточно точно, но процент ошибки был немного выше. Самое худшее предсказание было у регрессионной модели.

В таблице 7 представлены основные параметры сравнения моделей.

Таблица 7. Сравнение параметров моделей

Параметр	Регрессионная модель	Нейронная сеть с архитектурой № 1	Нейронная сеть с архитектурой № 2	LSTM-сеть
Время Построения	Быстрое построение	Время построения зависит от множества факторов: количества эпох обучения, слоев нейронов, типа слоев, нейронов в каждом слое, количества комплектов обучающей выборки, мощности ЭВМ		
Удобство	Проста в построении, не требует применения ЭВМ с высокой производительностью	Возможность сохранения модели и настройка точности, в зависимости от требований, требуются мощные ЭВМ		
		Модель не сложна в построении	Модель не сложна в построении, но требуется учитывать наиболее оптимальное расположение слоев	
Точность	Низкая точность	Высокая точность, ошибка предсказания менее 5 %	Высокая точность, ошибка предсказания менее 2 %	Высокая точность

Выводы

Искусственный интеллект часто используется при моделировании процессов нефтепереработки. Модели на основе искусственного интеллекта чаще всего используют искусственные нейронные сети, позволяют максимально точно описывать реальные процессы и аппараты и прогнозировать требуемые параметры. При использовании искусственных нейронных сетей не требуется выбор оптимальной модели, а также они способны самообучаться и дообучаться с появлением новых или более точных данных, что позволяет со временем оптимизировать их и повышать точность.

Построенные в ходе работы модели алкилирования изобутана олефинами описывают процесс в статике с высокой точностью и не требуют создания сложных физико-химических моделей.

Было выявлено, что при повышении эпох обучения точность предсказания возрастает, но существует риск переобучения модели. Также точность зависит от правильного подбора и расположения слоев и количества нейронов в них. Более высокая точность предсказания состава продукта была достигнута при использовании нейронных сетей, погрешность не превышала 1,71 %, что позволяет сделать вывод, о том, что их использование предпочтительно для построения подобных анализаторов качества.

Список используемых источников

1. Chen X., Zhang H., Song Y., Xiao R. Prediction of Product Distribution and Bio-Oil Heating Value of Biomass Fast Pyrolysis // Chemical Engineering and Processing – Process Intensification. 2018. Vol. 130. P. 36-42. DOI: 10.1016/j.cep.2018.05.018.

2. Ge Z., Song Z., Ding S.X., Huang B. Data Mining and Analytics in the Process Industry: The Role of Machine Learning // IEEE ACCESS. 2017. Vol. 5. P. 20590-20616. DOI: 10.1109/ACCESS.2017.2756872.

3. Ye Y., Hu T., Zhang C., Luo W. Design and Development of a CNC Machining Process Knowledge Base Using Cloud Technology // The International Journal of Advanced Manufacturing Technology. 2016. Vol. 94. Issue 9-12. P. 3413-3425. DOI: 10.1007/s00170-016-9338-1.

4. Lalanda P., Morand D., Chollet S. Autonomic Mediation Middleware for Smart Manufacturing // IEEE Internet Computing. 2017. No. 21 (1). P. 32-39. DOI: 10.1109/MIC.2017.18.

5. Шириязданов Р.Р. Регенерация цеолитсодержащего катализатора алкилирования изобутана бутан-бутеновой фракцией сверхкритическим диоксидом углерода // Сверхкритические флюиды: теория и практика. 2011. Т. 6. № 1. С. 19-24.

6. Руднев Н.А., Трапезникова Е.Ф., Хафизова С.Р., Смольникова Т.В., Хамзин Ю.А., Шириязданов Р.Р. Разработка кинетической модели процесса алкилирования изобутана олефинами на цеолитсодержащих катализаторах // *Химия и технология топлив и масел*. 2018. № 4. С. 16-19.

7. Gers F.A., Schmidhuber J., Cummins F. Learning to Forget: Continual Prediction with LSTM // *Neural Computation*. 2000. Vol. 12. No. 10. P. 2451-2471. DOI: 10.1162/089976600300015015.

8. Zhao R., Yan R., Wang J., Mao K. Learning to Monitor Machine Health with Convolution Bi-Directional LSTM Networks // *Sensors*. 2017. Vol. 17. No. 2. P. 273-290. DOI: 10.3390/s17020273.

9. Malhotra P., Vig L., Shroff G., Agarwal P. Long Short Term Memory Networks for Anomaly Detection in Time Series // *Computational Intelligence, and Machine Learning – 2015: Proceeding of 23rd European Symposium on Artificial Neural Networks*. Bruges, Belgium. 2015. P. 89-94.

References

1. Chen X., Zhang H., Song Y., Xiao R. Prediction of Product Distribution and Bio-Oil Heating Value of Biomass Fast Pyrolysis. *Chemical Engineering and Processing – Process Intensification*, 2018, Vol. 130, pp. 36-42. DOI: 10.1016/j.cep.2018.05.018.

2. Ge Z., Song Z., Ding S.X., Huang B. Data Mining and Analytics in the Process Industry: The Role of Machine Learning. *IEEE ACCESS*, 2017, Vol. 5, pp. 20590-20616. DOI: 10.1109/ACCESS.2017.2756872.

3. Ye Y., Hu T., Zhang C., Luo W. Design and Development of a CNC Machining Process Knowledge Base Using Cloud Technology. *The International Journal of Advanced Manufacturing Technology*, 2016, Vol. 94, Issue 9-12, pp. 3413-3425. DOI: 10.1007/s00170-016-9338-1.

4. Lalanda P., Morand D., Chollet S. Autonomic Mediation Middleware for Smart Manufacturing. *IEEE Internet Computing*, 2017, No. 21 (1), pp. 32-39. DOI: 10.1109/MIC.2017.18.

5. Shiriyazdanov R.R. Regeneratsiya tseolitsoderzhashchego katalizatora alkilirovaniya izobutana butan-butenovoi fraktsiei sverkhkriticheskim dioksidom ugleroda [Regeneration by Supercritical Fluid of Zeolite-Containing Catalyst for Isobutane Alkylation with Butane-Butylene Fraction]. *Sverkhkriticheskie flyuidy: teoriya i praktika – Supercritical Fluids Theory and Practice*, 2011, Vol. 6, No. 1, pp. 19-24. [in Russian].

6. Rudnev N.A., Trapeznikova E.F., Khafizova S.R., Smol'nikova T.V., Khamzin Yu.A., Shiriyazdanov R.R. Razrabotka kineticheskoi modeli protsessa alkilirovaniya izobutana olefinami na tseolitsoderzhashchikh katalizatorakh [Creation of the Kinetic Model of Alkylation of Isobutane with Olefins on Zeolite-containing Catalysts]. *Khimiya i tekhnologiya topliv i masel – Chemistry and Technology of Fuels and Oils*, 2018, No. 4, pp. 16-19. [in Russian].

7. Gers F.A., Schmidhuber J., Cummins F. Learning to Forget: Continual Prediction with LSTM. *Neural Computation*, 2000, Vol. 12, No. 10, pp. 2451-2471. DOI: 10.1162/089976600300015015.

8. Zhao R., Yan R., Wang J., Mao K. Learning to Monitor Machine Health with Convolution Bi-Directional LSTM Networks. *Sensors*, 2017, Vol. 17, No. 2, pp. 273-290. DOI: 10.3390/s17020273.

9. Malhotra P., Vig L., Shroff G., Agarwal P. Long Short Term Memory Networks for Anomaly Detection in Time Series. *Proceeding of 23rd European Symposium on Artificial Neural Networks «Computational Intelligence, and Machine Learning – 2015»*. Bruges, Belgium, 2015, pp. 89-94.

Сведения об авторах

About the authors

Симановский Евгений Анатольевич, магистрант кафедры «Газохимия и моделирование химико-технологических процессов», УГНТУ, г. Уфа, Российская Федерация

Evgeny A. Simanovsky, Undergraduate Student of Gas Chemistry and Modeling of Chemical Engineering Processes Department, USPTU, Ufa, Russian Federation

e-mail: evg.sim.an@yandex.ru

Устюжанин Константин Юрьевич, ассистент кафедры «Газохимия и моделирование химико-технологических процессов», УГНТУ, г. Уфа, Российская Федерация

Konstantin Y. Ustyuzhanin, Assistant of Gas Chemistry and Modeling of Chemical Engineering Processes Department, USPTU, Ufa, Russian Federation

e-mail: ustyuzhanin.ky@gmail.com

Руднев Николай Анатольевич, канд. техн. наук, доцент кафедры «Газохимия и моделирование химико-технологических процессов», УГНТУ, г. Уфа, Российская Федерация

Nikolay A. Rudnev, Candidate of Engineering Sciences, Assistant Professor of Gas Chemistry and Modeling of Chemical Engineering Processes Department, USPTU, Ufa, Russian Federation

e-mail: ngb2008@mail.ru

Шириязданов Ришат Рифкатович, д-р техн. наук, профессор кафедры «Газохимия и моделирование химико-технологических процессов», УГНТУ, г. Уфа, Российская Федерация

Rishat R. Shiriyaздanov, Doctor of Engineering Sciences, Professor of Gas Chemistry and Modeling of Chemical Engineering Processes Department, USPTU, Ufa, Russian Federation

e-mail: petroleum9@bk.ru

Хафизова Светлана Ринатовна, канд. хим. наук, доцент кафедры «Газохимия и моделирование химико-технологических процессов», УГНТУ, г. Уфа, Российская Федерация

Svetlana R. Khafizova, Candidate of Chemical Sciences, Assistant Professor of Gas Chemistry and Modeling of Chemical-Technological Processes Department, USPTU, Ufa, Russian Federation

e-mail: hafizovs@yandex.ru

Смольникова Татьяна Васильевна, канд. хим. наук, доцент кафедры «Газохимия и моделирование химико-технологических процессов», УГНТУ, г. Уфа, Российская Федерация

Tatyana V. Smolnikova, Candidate of Chemical Sciences, Assistant Professor of Gas Chemistry and Modeling of Chemical-Technological Processes Department, USPTU, Ufa, Russian Federation

e-mail: SmolnikovaT@yandex.ru