УДК 658.5.012.7

И.С. Можаровский, С.А. Шевлягина

ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ АНСАМБЛЯ МОДЕЛЕЙ ПРИ ПОСТРОЕНИИ ВИРТУАЛЬНЫХ АНАЛИЗАТОРОВ ДЛЯ ОЦЕНКИ ПОКАЗАТЕЛЯ КАЧЕСТВА СЛАБОФОРМАЛИЗОВАННОГО ПРОЦЕССА[1]

Можаровский Игорь Сергеевич, *кандидат технических наук, окончил Владивостокский государственный университет экономики и сервиса, доцент инженерной школы Владивостокского государственного университета, младший научный сотрудник Института автоматики и процессов управления Дальневосточного отделения РАН. Имеет научные работы в области обработки данных, моделирования технологических объектов и построения виртуальных анализаторов. [e-mail: Mozharovskiy.Igor@vvsu.ru].*

Шевлягина Светлана Александровна, *кандидат технических наук, окончила Дальневосточный федеральный университет, старший научный сотрудник ИАПУ ДВО РАН, доцент ДВФУ. Имеет научные работы в области моделирования технологических процессов, сбора и обработки данных, построения виртуальных анализаторов и систем управления. [e-mail: samotylova@dvo.ru].*

Аннотация

В нефтеперерабатывающих и нефтехимических производствах одной из ключевых задач является выбор подходящей стратегии управления технологическими процессами с минимальными материальными и энергетическими затратами, при этом без снижения качества производимой продукции, которое важно быстро и оперативно оценить. Для мониторинга и контроля выпускаемой продукции в режиме реального времени широкое распространение получили виртуальные анализаторы (ВА), которые используются в качестве альтернативы поточным анализаторам и лабораторным измерениям. Однако в реальных условиях часто сложно построить ВА для оценки показателей качества для слабоформализованных процессов. Предлагается использовать ансамбли моделей (АМ) как способ повышения точности ВА для слабоформализованных процессов. АМ применяют набор моделей для получения более точных прогнозов. Для создания разнообразных моделей дополнительно используется внедрение шумовой составляющей в технологические переменные, которые являются входами при построении ВА. Приводится сравнительный анализ нескольких комбинаций АМ. Эффективность использования АМ при построении ВА для оценки качества выпускаемой продукции продемонстрирована на слабоформализованном процессе стабилизации и вторичной переработки бензина.

Ключевые слова: виртуальный анализатор, слабоформализованный процесс, ансамбли моделей, качество продукции, стабилизация и вторичная переработка бензина.

doi: 10.35752/1991-2927_2024_3_77_111

PROSPECTS OF USING MODEL ENSEMBLES IN SOFT SENSOR DESIGN FOR ESTIMATING THE QUALITY OF A WEAKLY FORMALIZED PROCESS

Igor Sergeevich Mozharovskii, *Candidate of Sciences in Engineering; graduated from the Vladivostok State University of Economics and Service, Associate Professor at the Engineering School of Vladivostok State University; a junior staff scientist at the Institute of Automation and Control Processes of the Far East Branch of the Russian Academy of Sciences; an author of scientific works in the field of data processing, modeling of technological objects and soft sensors design. e-mail: Mozharovskiy.Igor@vvsu.ru.*

Svetlana Aleksandrovna Shevliagina, *Candidate of Sciences in Engineering; graduated from Far Eastern Federal University; a senior staff scientist at the Institute of Automation and Control Processes of the Far East Branch of the Russian Academy of Sciences, Associate Professor at the Far Eastern Federal University; an author of scientific works in the field of modeling of technological processes, data acquisition and processing, soft sensors design and control systems. e-mail: samotylova@dvo.ru.*

1 Исследование выполнено в рамках государственного задания ИАПУ ДВО РАН (тема FWFW-2022-0002).

Abstract
One of the most important tasks in the petroleum refining and petrochemical industries is to select an appropriate control strategy for industrial processes. Such strategies are required to minimize material and energy costs without compromising the quality of manufacturing products, which must be evaluated quickly and promptly. Soft sensors (SS) are widely used as an alternative to in-line analyzers and laboratory measurements for real-time product monitoring and control. However, in real-world applications, it is often difficult to design SSs to estimate product quality for weakly formalized processes. The article proposes the use of model ensembles (ME) to improve the accuracy of SSs for weakly formalized processes. MEs use a set of models to produce more accurate predictions. In addition, the introduction of a noise component into the process variables used as inputs in the SS design is used to create a variety of models. A comparative analysis of several combinations of ME is presented. The efficiency of using ME in the SS design to assess the quality of output products is demonstrated on a weakly formalized process of stabilization and naphtha distillation.

Keywords: soft sensor, weakly formalized process, model ensemble, product quality, stabilization and naphtha distillation.

Введение

Использование виртуальных анализаторов (ВА) на производстве позволяет операторам в режиме реального времени оценивать требуемые показатели качества выходных продуктов, оперативно реагировать на технические нарушения и обеспечивать безопасный переход на новый режим функционирования установки. Дополнительным преимуществом внедрения ВА на производство является своевременное выявление и устранение бракованной продукции, что приводит к сокращению материальных и энергетических затрат при переработке некондиционной продукции. Причем, ВА являются ценным инструментом во многих отраслях промышленности, таких как целлюлозно-бумажные комбинаты, системы очистки сточных вод, нефтеперерабатывающая и нефтехимическая промышленность, цементные печи и др.

Согласно [1], перспективы использования ВА в промышленности подкреплены необходимостью выбора подходящей производственной политики с целью повышения производства конечной продукции и её качества с минимальными энергетическими и материальными затратами. При этом ВА позволяют оценить требуемые переменные (показатели качества продукции), которые не могут быть измерены автоматически или измеряются с большими временными затратами, неточно или спорадически, например, с помощью лабораторных анализов, в режиме реального времени [2, 3].

Существует большое количество трудностей при построении и обслуживании ВА, связанных с нелинейностью технологических процессов, вызванной изменениями условий функционирования и эксплуатации установки; изменением состава, подаваемого на установку, сырья; малого фрагмента данных; слабой формализованностью объектов исследования [1]. Под последним понимается система, которую невозможно описать конечным числом математических выражений, описывающих её состояние [4]. Несмотря на постоянное развитие техники и различных подходов, возникает задача построения ВА для оценки показателей качества сложного слабоформализованного процесса.

Нечеткие алгоритмы находят широкое применение в различных отраслях, в том числе и при построении ВА [5, 6]. Основным недостатком их применения является сложность их настройки, а именно составления базы нечётких правил. Для этого требуются эмпирические знания об объекте исследования, четкое понимание функциональных зависимостей между технологическими переменными и требуемыми показателями качества продукции и репрезентативный набор данных, который не всегда доступен. Большой интерес при построении ВА вызывают ансамбли моделей (АМ) в связи с простотой их реализации [7–9]. При этом АМ объединяют несколько моделей для получения надежных прогнозов.

Подход АМ заключается в получении нескольких моделей, обучающихся на различных случайных выборках данных из одного обучающего множества, а конечное значение ВА рассчитывается как взвешенное среднее значений выхода всех моделей [10]. Такая совокупность моделей обычно более точна, чем из отдельных моделей, составляющих АМ. Двумя популярными методами АМ являются Bagging [11, 12] и Boosting [13, 14]. Для повышения эффективности АМ можно использовать различные начальные условия для обучения моделей, манипулировать обучающим набором данных, использовать разные архитектуры и алгоритмы обучения.

В этой работе сравниваются различные стратегии построения ВА для оценки показателей качества для слабоформализованного процесса. Основное внимание уделяется построению АМ для улучшения ВА. Дополнительно используется внедрение шумовой составляющей во входные технологические переменные, способствующие разнообразить модели для достижения наилучшего результата. В статье тестируются различные стратегии АМ для создания надежных ВА. Демонстрация функционирования ВА показана на примере слабоформализованного массообменного технологического процесса стабилизации и перегонки бензина.

Описание объекта и постановка задачи

В качестве объекта исследования выбран слабоформализованный массообменный технологический процесс стабилизации и вторичной перегонки бензина (рис. 1). Отбираемая на атмосферной установке фракция НК-140 °С поступает на колонну К-1. С верха колонны К-1 пары поступают в конденсатор, где частично

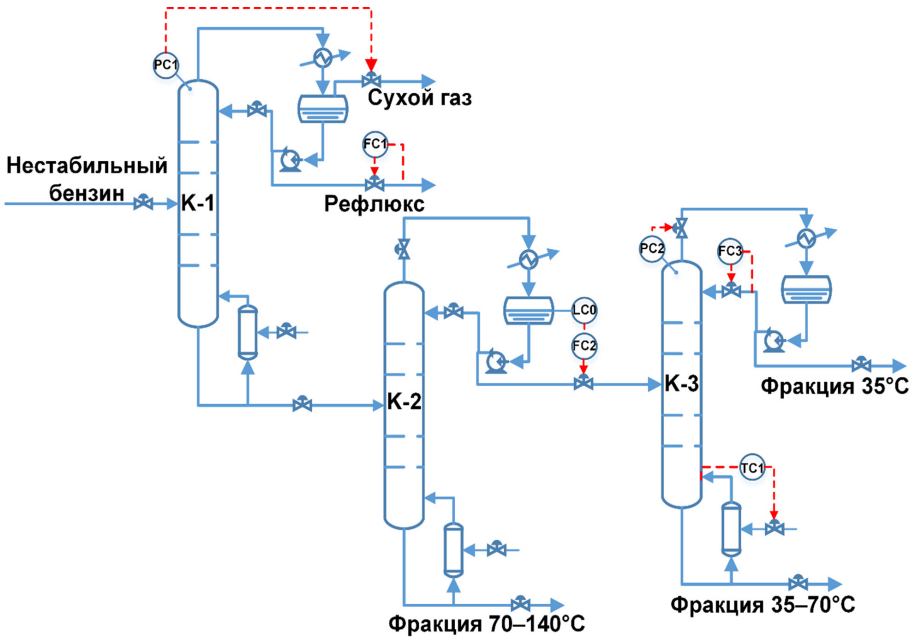

Рис. 1. Схема процесса стабилизации и перегонки бензина

конденсируются и собираются в емкости. Углеводородный газ (сухой газ) выводится в линию топливного газа. Часть рефлюкса возвращается в колонну в качестве орошения, а балансовый избыток выводится с установки (позиция FC1).

Стабильная фракция НК-140 °С из К-1 поступает в колонну К-2, где происходит её разделение на фракции НК-70 °С и 70–140 °С. С верха колонны К-2 пары через конденсатор поступают в емкость, откуда часть фракции подается на орошение в колонну К-2, а балансовый избыток через клапан (позиция FC2) откачивается на загрузку колонны К-3. В колонне К-3 происходит разделение фракций НК-35 °С и 35–70 °С. Фракция НК-35 °С с верха колонны К-3 через конденсатор поступает в емкость. Из емкости часть фракции НК-35 °С с помощью насоса через клапан регулятора расхода (позиция FC3) подается на орошение К-3, а балансовый избыток откачивается в промежуточный парк газофракционирующей установки. С низа колонны К-3 фракция 35–70 °С откачивается с установки.

Отбор входных переменных, использованных для построения ВА, выполнен на основе экспертной оценки и представлен в таблице 1.

На производстве контроль качества фракции НК 35–70 °С (определение концентрации суммы углеводородов C_1–C_4 во фракции НК 35–70 °С) осуществляется с помощью лабораторных исследований 1 раз в 8 часов. Несвоевременный контроль качества производимой продукции не позволяет оперативно обнаруживать брак на ранних стадиях производства. В связи с этим для исследуемого объекта необходима оперативная оценка концентрации суммы углеводородов C_1–C_4 во фракции НК 35–70 °С (низ колонны К-3).

Ансамбли моделей

Рассмотрим два из наиболее распространенных алгоритмов формирования АМ: бэггинг (Bagging) и бустинг (Boosting), которые генерируют и объединяют множество моделей, используя один алгоритм обучения.

Таблица 1

Описание технологических переменных

Переменные	Описание	Ед. измерения	Позиция
x_1	Давление верха колонны К-1	кгс/см^2	PC-1
x_2	Расход рефлюкса колонны К-1	т/ч	FC-1
x_3	Давление верха колонны К-3	кгс/см^2	PC-2
x_4	Температура куба колонны К-3	°С	TC-1
x_5	Расход орошения колонны К-3	м^3/ч	FC-3
x_6	Расход сырья, поступающий в колонну К-3	м^3/ч	FC-2
y	Сумма углеводородов C_1–C_4 во фракции НК 35–70 °С	мас. %	-

Bagging является самым простым в реализации методом. Обозначим исходную выборку как $Z = \left(\left[x_1 \quad y_1 \right], ..., \left[x_i \quad y_i \right], ..., \left[x_N \quad y_N \right] \right)$, где N – количество наблюдений в выборке. Из исходного набора данных Z случайным образом создаются B-бутстреп выборок Z_b^*, $b = 1, 2, ..., B$. На основе сформированных бутстреп-выборок обучаются и разрабатываются модели $\hat{f}_1^* (x), ..., \hat{f}_b^* (x), ..., \hat{f}_B^* (x)$. Результаты прогнозов полученных моделей усредняются.

Таким образом, прогноз ансамбля можно записать, как среднее B аппроксимирующих функций на основе набора B-бутстреп выборок:

$$\hat{f}_{bagging} (x) = \frac{1}{B} \sum_{b=1}^{B} \hat{f}_b^* (x).$$

Такой метод хорошо подходит к построению моделей, чувствительных (с точки зрения параметров модели и эффективности прогнозирования) к небольшим изменениям в обучающем наборе данных.

Что касается алгоритма Boosting, то он сложнее, чем Bagging. Основное отличие Boosting от ранее описанного алгоритма Bagging заключается в последовательном обучении нескольких базовых моделей с использованием весов w. Механизм определения индивидуальных весов будет зависеть от полученного прогноза. В случае получения прогноза, приближенного к реальному, индивидуальные веса будут уменьшаться, в случае определения некорректного значения прогноза – увеличиваться. При таком подходе используемая для обучения новой модели функция ошибок зависит от производительности предыдущих моделей. После того, как все базовые модели построены, их прогнозы объединяются в окончательный прогноз:

$$\hat{f}_{boosting} (x, w) = \sum_{b=1}^{B} w_b \hat{f}_b^* (x).$$

Алгоритм Boosting можно представить в следующем виде:

1. Всем N строкам обучающей выборки присвоить одинаковые веса $1/N$.

2. В цикле от b = 1 до B:

2.1. Обучить и получить модель $\hat{f}_b^* (x)$ на данных с соответствующими весами.

2.2. Вычислить взвешенную ошибку прогноза $\hat{f}_b^* (x)$.

2.3. Пересчитать веса на каждой строке в соответствии с ошибкой прогноза: уменьшить вес для корректно полученных прогнозов, увеличить – для некорректно полученных прогнозов.

3. Вычислить результирующий прогноз ансамбля как взвешенный прогноз базовых моделей.

В [15] предлагается улучшить производительность прогнозирования АМ за счет добавления шумовой составляющей к входным данным для формирования моделей на основе различных обучающих выборок. Вопрос правильного выбора шумовой составляющей в литературе до конца не освещен. Большие шумы искажают основную особенность данных, в то время как небольшие могут не оказывать достаточного воздействия.

К исходным данным технологических переменных добавляются шумы, представляющие собой случайные векторы, имеющие нормальное распределение $N\left(0, \sigma^2\right)$. Тогда матрица входных переменных будет определяться следующим образом:

$$X' = X + U, \quad U = \begin{bmatrix} u_{1,1} & \cdots & u_{1,6} \\ \vdots & \vdots & \cdots \\ u_{N,1} & \cdots & u_{N,6} \end{bmatrix}.$$

Следует отметить, что матрица шумовой составляющей U добавляется только в матрицу входов X и не добавляется в выходной вектор y.

Обсуждение результатов

В качестве обучающих и тестовых выборок использовались данные с установки (технологические параметры) и лаборатории y. Данные технологических параметров соотнесены со временем отбора пробы проведения лабораторного анализа и усреднены за последний час. Исходный набор промышленных данных предварительно обработан для исключения выбросов. После обработки набор данных включал 500 наблюдений и 6 входных переменных (табл. 1). Первые 80 % данных вошли в обучающий набор, оставшиеся 20 % использованы в качестве тестового набора для верификации полученных ВА. Таким образом, размеры обучающей и тестовой выборок составили 400x7 и 100x7 соответственно (с учетом выходной переменной).

При построении и выборе ВА для оперативной оценки суммы углеводородов C_1–C_4 во фракции НК 35–70 °С слабоформализованного процесса стабилизации и вторичной перегонки бензина рассмотрены следующие алгоритмы: метод наименьших квадратов (МНК), робастная регрессия (РР), метод опорных векторов (МОВ), нейронная сеть прямого распространения (НС), бэггинг деревьев решений и градиентный бустинг с МНК в качестве функции потерь (Least-squares boosting – LSBoost).

Эффективность ВА, полученных различными методами с подбором оптимальных параметров, определялась с использованием коэффициента детерминации R^2, абсолютной средней ошибки (САО) и среднеквадратической ошибки (СКО):

$$R^2 = 1 - \frac{\sum_{n=1}^{N} \left(y_n - \hat{y}_n \right)^2}{\sum_{n=1}^{N} \left(y_n - \overline{y} \right)^2},$$

$$\text{САО} = \frac{1}{N} \sum_{n=1}^{N} \left| y_n - \hat{y}_n \right|,$$

$$\text{СКО} = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y}_n)^2,$$

где y_n – измеряемое значение выходной переменной, мас. %; $\hat{y}_n$ – ее значение, полученное на основе ВА, мас. %; $\overline{y}$ – среднее значение наблюдаемой выходной переменной, мас. %.

Была определена структура НС, имеющая два скрытых слоя с 10 нейронами в каждом слое, с сигмоидной функцией активации. Для обучения НС использовался метод Левенберга-Марквардта. Для обучения АМ использовался весь набор данных обучающей выборки и 23 цикла обучения, скорость обучения равна 1. Для повышения точности прогнозов АМ использовалась стратегия добавления шумовой составляющей в исходную матрицу входов с дисперсией σ^2 = 0,03 как оптимально подобранной.

Результаты вычислений представлены в таблице 2.

Из представленных в таблице 2 результатов критериев точности наилучшими подходами к построению ВА для оценки требуемого показателя качества производимой продукции являются: НС и АМ. Бэггинг деревьев решений и градиентный бустинг показали прирост в точности при использовании стратегии добавления шумовой составляющей в исходную матрицу входов. Как показано в таблице 2, использование АМ (алгоритм Bagging) улучшает точность ВА для оценки требуемого показателя качества продукта слабоформализованного объекта в сравнении с ВА, построенным на основе РР, на ≈ 47 %, 9 % и 26 %. Добавление шумовой составляющей в обучающий набор входных данных позволил дополнительно повысить точность ВА для алгоритма Bagging до и после шумовой составляющей на ≈ 7 %, 4 % и 8 % по R^2, САО и СКО, соответственно.

На рисунках 2 и 3 представлены результаты функционирования ВА, построенного на основе РР, и ВА, построенного на АМ (алгоритм Bagging) с добавлением шумовой составляющей.

Из рисунков 2 и 3 видно, что концентрация суммы углеводородов C_1–C_4 во фракции 35–70 ºС имеет сильную изменчивость и может достигать 8 мас.%. Одна из причин – это изменение состава подаваемого сырья на установку стабилизации и вторичной перегонки бензина. Использование классических подходов множественной линейной регрессии, таких как РР, не позволяет построить ВА, который обеспечит надежную оценку требуемого показателя качества. Использование АМ позволяет периодически осуществлять оценку y на уровне 4–5 мас.%, что дает возможность рассматривать такой подход и в дальнейшем для таких типов процесса.

Заключение

При стабилизации и вторичной перегонке бензина важно быстро оценить концентрацию суммы углеводородов C_1–C_4 во фракции 35–70 ºС. Процесс является

Таблица 2

Критерии точности ВА на тестовой выборке

Критерий точности	МНК	РР	МОВ	НС	АМ		АМ с добавлением шума	
					Bagging	LSBoosting	Bagging	LSBoosting
R^2	0,461	0,357	0,490	0,518	0,526	0,515	0,562	0,554
САО	0,721	0,725	0,695	0,666	0,661	0,673	0,637	0,644
СКО	0,883	1,054	0,836	0,789	0,778	0,794	0,719	0,731

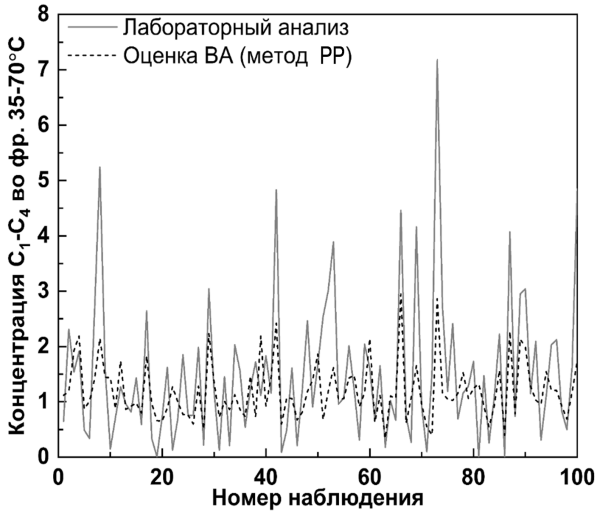

Рис. 2. Оценка концентрации суммы углеводородов C_1–C_4 во фракции 35–70 ºС, полученная РР

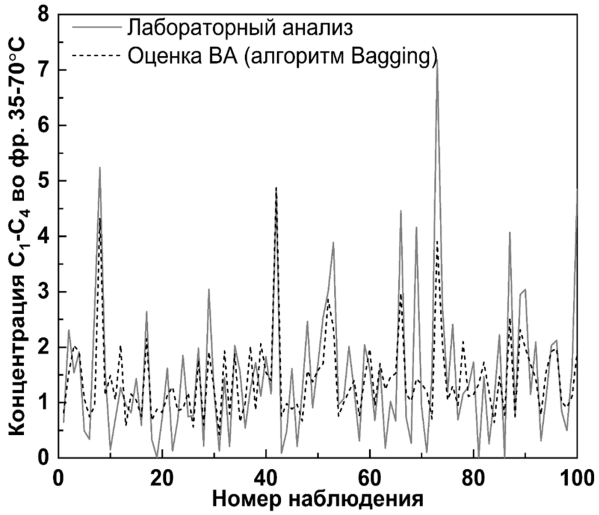

Рис. 3. Оценка концентрации суммы углеводородов C_1–C_4 во фракции 35–70 ºС, полученная на АМ

слабоформализованным, и используемые в промышленности методы множественной линейной регрессии не подходят для оценки качества выпускаемой продукции. В связи с этим в статье сравнивается ряд стратегий по улучшению прогнозных возможностей ВА на основе АМ.

Результаты показывают, что АМ с внесением шумовой составляющей на входные данные позволяют повысить точность ВА по сравнению с РР и другими АМ. Лучшие характеристики достигаются при использовании алгоритма Bagging. Точность полученных ВА недостаточна, это свидетельствует о необходимости проведения дополнительных исследований. Полученный результат позволяет предполагать, что рассматриваемые методы моделирования можно применять для построения ВА для слабоформализованных процессов.

СПИСОК ЛИТЕРАТУРЫ

1. Soft Sensors for Monitoring and Control of Industrial Processes / L. Fortuna, S. Graziani, A. Rizzo, M.G. Xibilia. Switzerland : Springer, 2007. 284 p.

2. Ma L., Wang M., Peng K. A Two-Phase Soft Sensor Modeling Framework for Quality Prediction in Industrial Processes with Missing Data // Journal of Process Control. 2023. Vol. 129. 103061.

3. Integrating Transfer Learning Within Data-Driven Soft Sensor Design to Accelerate Product Quality Control / S. Kay, H. Kay, M. Mowbray, A. Lane, C. Mendoza, P. Martin, D. Zhang // Digital Chemical Engineering. 2024. Vol. 10. 100142.

4. Walter E., Pronzato L. On the Identifiability and Distinguishability of Nonlinear Parametric Models // Mathematics and Computers in Simulation. 1996. Vol. 42. pp. 125–134.

5. Mozharovskii I., Shevlyagina S. A Hybrid Approach to Soft Sensor Development for Distillation-in-Series Plant under Input Data Low Variability // Measurement Science and Technology. 2024. Vol. 35, iss. 7. 076211.

6. Genetic Fuzzy System for Data-Driven Soft Sensors Design / J. Mendes, F. Souza, R. Araújo, N. Goncalves // Applied Soft Computing. 2012. Vol. 12, iss. 10. pp. 3237–3245.

7. Polikar R. Ensemble based systems in decision making // IEEE Circuits and Systems Magazine. 2006. Vol. 6, iss. 3. pp. 21–45.

8. Алексеева В.А. Использование методов машинного обучения в задачах бинарной классификации // Автоматизация процессов управления. 2015. № 3 (41). С. 58–63.

9. Юрков Д.А., Сокольчик П.Ю., Сташков С.И. Применение ансамблевых методов для прогнозирования качества продукции // Химия. Экология. Урбанистика : матер. всерос. науч.- практ. конф. (с междунар. участием), г. Пермь, 17–19 апреля 2024 г. : в 4 т. Пермь, 2024. Т. 3. С. 390–394.

10. Shao W., Tian X. Adaptive Soft Sensor for Quality Prediction of Chemical Processes Based on Selective Ensemble of Local Partial Least Squares Models // Chemical Engineering Research and Design. 2015. Vol. 95. pp. 113–132.

11. Breiman L. Bagging Predictors // Machine Learning. 1996. Vol. 24, iss. 2. pp. 123–140.

12. Bakır R., Orak C., Yüksel A. Optimizing Hydrogen Evolution Prediction: A Unified Approach Using Random Forests, lightGBM, and Bagging Regressor Ensemble Model // International Journal of Hydrogen Energy. 2024. Vol. 67. pp. 101–110.

13. Freund Y., Schapire R. A Short Introduction to Boosting // Japanese Society for Artificial Intelligence. 1999. Vol. 14, iss. 5. pp. 771–780.

14. Boosting Diversity in Regression Ensembles / M. Bourel, J. Cugliari, Y. Goude, J.M. Poggi // Statistical Analysis and Data Mining: The ASA Data Science Journal. 2024. Vol. 17, iss. 1. e11654.

15. Zhang G.P. A Neural Network Ensemble Method with Jittered Training Data for Time Series Forecasting // Information Sciences. 2007. Vol. 177, iss. 23. pp. 5329–5346.

REFERENCES

1. Fortuna L., Graziani S., Rizzo A., M.G. Xibilia. *Soft Sensors for Monitoring and Control of Industrial Processes.* Switzerland, Springer Publ., 2007. 284 p.

2. Ma L., Wang M., Peng K. A Two-Phase Soft Sensor Modeling Framework for Quality Prediction in Industrial Processes with Missing Data. *Journal of Process Control*, 2023, vol. 129, 103061.

3. Kay S., Kay H., Mowbray M., Lane A., Mendoza C., Martin P., Zhang D. Integrating Transfer Learning Within Data-Driven Soft Sensor Design to Accelerate Product Quality Control. Digital Chemical Engineering, 2024, vol. 10, 100142.

4. Walter E., Pronzato L. On the Identifiability and Distinguishability of Nonlinear Parametric Models. *Mathematics and Computers in Simulation*, 1996, vol. 42, pp. 125–134.

5. Mozharovskii I., Shevlyagina S. A Hybrid Approach to Soft Sensor Development for Distillation-in-Series Plant under Input Data Low Variability. *Measurement Science and Technology*, 2024, vol. 35, iss. 7, 076211.

6. Mendes J., Souza F., Araujo R., Goncalves N. Genetic Fuzzy System for Data-Driven Soft Sensors Design. *Applied Soft Computing*, 2012, vol. 12, iss. 10, pp. 3237–3245.

7. Polikar R. Ensemble Based Systems in Decision Making. *IEEE Circuits and Systems Magazine*, 2006, vol. 6, iss. 3. pp. 21–45.

8. Alekseeva V.A. Ispolzovanie metodov mashinnogo obucheniia v zadachakh binarnoi klassifikatsii [Use of the Machine Learning Methods in Binary Classification Problems]. *Avtomatizatsiia protsessov upravleniia* [Automation of Control Processes], 2015, no. 3 (41), pp. 58–63.

9. Yurkov D.A., Sokolchik P.Iu., Stashkov S.I. Primenenie ansamblevykh metodov dlia prognozirovaniia kachestva produktsii [Application of Ensemble Methods for the Product Quality Forecasting]. *Khimiia. Ekologiia. Urbanistika. Mater.*

vseros. nauch.-prakt. konf. (s mezhdunar. uchastiem), g. Perm, 17–19 aprelia 2024 g. v 4 t. [Proceedings of All-Russian Sci.-Pract. Conf. on Chemistry, Ecology, Urbanistics (with International Participation), city of Perm, 17-19 April, 2024, in 4 Volumes]. Perm, 2024, vol. 3, pp. 390–394.

10. Shao W., Tian X. Adaptive Soft Sensor for Quality Prediction of Chemical Processes Based on Selective Ensemble of Local Partial Least Squares Models. *Chemical Engineering Research and Design*, 2015, vol. 95, pp. 113–132.

11. Breiman L. Bagging Predictors. *Machine Learning*, 1996, vol. 24, iss. 2, pp. 123–140.

12. Bakir R., Orak C., Yüksel A. Optimizing Hydrogen Evolution Prediction: A Unified Approach Using Random Forests, Light GBM, and Bagging Regressor Ensemble Model. *International Journal of Hydrogen Energy*, 2024, vol. 67, pp. 101–110.

13. Freund Y., Schapire R. A Short Introduction to Boosting. *Japanese Society for Artificial Intelligence*, 1999, vol. 14, iss. 5, pp. 771–780.

14. Bourel M., Cugliari J., Goude Y., Poggi J.M. Boosting Diversity in Regression Ensembles. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 2024, vol. 17, iss. 1, e11654.

15. Zhang G.P. A Neural Network Ensemble Method with Jittered Training Data for Time Series Forecasting. *Information Sciences*, 2007, vol. 177, iss. 23, pp. 5329–5346.