



Научная статья

УДК 621.311

<https://doi.org/10.21285/1814-3520-2022-3-487-498>

Повышение точности прогноза электропотребления промышленного предприятия методами машинного обучения с помощью отбора значимых признаков из временного ряда

Никита Николаевич Сергеев^{1✉}, Павел Викторович Матренин²^{1,2}Новосибирский государственный технический университет, г. Новосибирск, Россия¹veegresatikin3102@gmail.com, <https://orcid.org/0000-0003-1534-9072>²matrenin.2012@corp.nstu.ru, <https://orcid.org/0000-0001-5704-0976>

Резюме. Цель – повышение точности прогнозирования электропотребления предприятия за счет анализа и предварительной обработки входных данных, а также оценка влияния выбора признаков на результаты различных моделей прогнозирования. В качестве объекта, для которого составляется прогноз, было выбрано деревообрабатывающее предприятие Нижнего Новгорода. В работе сравниваются два типа методов машинного обучения: нейросетевые и ансамблевые модели. Рассмотрен один из способов выделения наиболее значимых параметров (признаков) временного ряда для улучшения результата следующих ансамблевых моделей на базе деревьев решений: адаптивный бустинг (AdaBoost), градиентный бустинг (Gradient Boosting) и случайный лес (Random Forest). Определение наиболее важных признаков исходного временного ряда осуществлено с помощью расчета коэффициентов корреляции между электропотреблением в прогнозируемый час и в предшествующие часы. Установлено, что в рассматриваемом объекте для построения прогноза наиболее важными признаками являются значения потребляемой энергии в те часы, которые отстают от прогнозируемого часа на кратное число суток. В качестве дополнительного признака используется график планово-профилактических работ деревообрабатывающих станков. Результаты показали, что при выборе правильной методики отбора признаков ансамбли деревьев решений способны превосходить искусственные нейронные сети по результатам: наименьшая средняя ошибка нейросетевой модели на тестовой выборке составляет 7,0%, в то время как для одной ансамблевой модели (градиентного бустинга) получено значение ошибки в 5,5%. Показано, что использование графика ремонтных работ также повышает точность прогноза: для рассмотренных ансамблевых моделей снижение ошибки составило от 20 до 30%.

Ключевые слова: анализ временных рядов, прогнозирование электропотребления, промышленное предприятие, машинное обучение, искусственные нейронные сети, ансамбли деревьев решений

Финансирование: Исследование выполнено за счет гранта Российского научного фонда (проект № 22-79-00181).

Для цитирования: Сергеев Н. Н., Матренин П. В. Повышение точности прогноза электропотребления промышленного предприятия методами машинного обучения с помощью отбора значимых признаков из временного ряда // iPolytech Journal. 2022. Т. 26. № 3. С. 487–498. <https://doi.org/10.21285/1814-3520-2022-3-487-498>.

ENERGY INDUSTRY

Original article

Increasing the accuracy of forecasting the electricity consumption of an industrial enterprise by machine learning methods using the selection of significant features from a time series

Nikita N. Sergeev^{1✉}, Pavel V. Matrenin²^{1,2}Novosibirsk State Technical University, Novosibirsk, Russia¹veegresatikin3102@gmail.com, <https://orcid.org/0000-0003-1534-9072>²matrenin.2012@corp.nstu.ru, <https://orcid.org/0000-0001-5704-0976>

Abstract. This study aims to improve the accuracy of forecasting the electricity consumption of an enterprise based on an analysis and preliminary processing of input data, as well as at evaluating the effect caused by feature selection on the results of various forecast models. A woodworking enterprise located in Nizhny Novgorod was selected as a forecast object. Two types of machine learning methods, including neural network and ensemble models, were compared. An approach to selecting the most significant parameters (features) from a time series was considered in order to improve the results of the following ensemble models based on decision trees: adaptive busting (AdaBoost), Gradient Boosting and Random Forest. The most significant features of the initial time series were determined using the calculation of correlation coefficients between the values of electricity consumption in forecasted and previous hours. For the considered forecast object, the most significant features were established to be the consumed energy in hours lagging behind the forecasted hour by the multiple number of days. The schedule of repair works for woodworking machines was used as an additional feature. According to the obtained results, decision tree ensembles can surpass artificial neural networks provided that significant features are selected correctly. Thus, the smallest average error of a neural network model on a test sample comprised 7.0%, while an error of 5.5% was obtained for a Gradient Boosting ensemble model. The use of a repair schedule was demonstrated to additionally increase the forecast accuracy: for the considered ensemble models, the error reduced from 20 to 30%.

Keywords: time series analysis, electric power consumption forecasting, industrial enterprise, machine learning, artificial neural networks, decision tree ensembles

Funding: The research was funded by the grant of the Russian Science Foundation (project No. 22-79-00181).

For citation: Sergeev N. N., Matrenin P. V. Increasing the accuracy of forecasting the electricity consumption of an industrial enterprise by machine learning methods using the selection of significant features from a time series. *iPolytech Journal*. 2022;26(3):487-498. (In Russ.). <https://doi.org/10.21285/1814-3520-2022-3-487-498>.

ВВЕДЕНИЕ

Прогнозирование электрической нагрузки потребителей представляет собой важную задачу, решение которой необходимо прежде всего системному оператору для эффективного управления электроэнергетической системой.

Конкретные цели прогнозирования во многом определяются временным горизонтом прогноза [1]:

- краткосрочные прогнозы составляются на несколько часов или несколько дней вперед и в основном используются для регулирования энергетического баланса, а также при формировании ценовых заявок поставщиков и покупателей электрической энергии;

- среднесрочные прогнозы строятся на срок от одного месяца до одного года вперед и применяются для решений по техническому обслуживанию оборудования и сетей, определения цены на электроэнергию, заключения сделок по передаче электроэнергии и т. д.;

- долгосрочные прогнозы строятся на несколько лет вперед и обычно используются для оценки будущего спроса на электроэнергию, анализа необходимости сооружения новых линий электропередачи и др.

Участники оптового рынка электроэнергии и мощности особенно заинтересованы в высоком качестве краткосрочных прогнозов, по-

скольку вместе с оплатой электроэнергии по ценам оптового рынка необходимо оплачивать штрафы за отклонения от ценовых заявок на покупку электроэнергии [2]. Это говорит о необходимости поиска новых методов прогнозирования графиков нагрузки и совершенствовании существующих.

Прогнозирование графиков нагрузки включает в себя задачу анализа временных рядов. Современные методы прогнозирования можно разделить на классические статистические методы и методы машинного обучения. Первая группа включает в себя различные регрессионные, авторегрессионные и гибридные модели, например авторегрессии со скользящим средним ARMA/ARIMA [1, 3–6]. Эти методы сталкиваются с трудностями на практике в случае сильностохастического, вероятностного характера изменений нагрузки энергосистемы и отсутствия строгой периодичности.

К методам машинного обучения относят методы опорных векторов [5–8], решающие деревья [9, 10], ансамблевые методы [11, 12] и искусственные нейронные сети (ИНС) [6, 12–14]. Эффективность таких методов существенно зависит от объема, вида и информативности данных, на которых обучаются модели. Необходимость обучения моделей является главной особенностью и в то же время главным недостатком методов машинного

обучения. Для получения качественных результатов требуется большая выборка данных для обучения, и при этом данные должны быть достоверными и репрезентативными.

Отдельно следует выделить ИНС как наиболее сложные и разнообразные на данный момент алгоритмы машинного обучения. По своей сути нейронные сети являются нелинейными статистическими моделями. Существует множество различных архитектур нейронных сетей, каждая из которых имеет свои особенности и предназначена для решения определенного набора задач. Широкое применение для прогнозирования графиков электропотребления нашли нейронные сети прямого распространения (Feedforward Neural Network) [13], сверточные нейронные сети (Convolutional Neural Network, CNN) [14] и рекуррентные нейронные сети (Recurrent Neural Network, RNN) [6, 12].

ИНС имеют ряд преимуществ, благодаря которым они могут быть успешно применены в различных задачах и ситуациях:

– могут обучаться и моделировать сложные нелинейные зависимости;

– умеют выполнять обобщение (после завершения обучения и определения зависимостей в исходном наборе данных они могут применять изученные зависимости и делать предсказания на новом наборе данных);

– не требуют строгих ограничений к входным величинам (например, к их распределению), как правило, единственное требование к данным – принадлежность значений отрезку $[0; 1]$ или $[-1; 1]$, что можно получить достаточно просто с помощью нормализации числовых признаков и кодирования категориальных;

– способны выполнять несколько задач параллельно без ущерба производительности системы.

Отбор признаков представляет собой процесс получения подвыборки признаков, удовлетворяющих определенным критериям, из исходной выборки [15–17]; к таким критериям относятся коэффициент корреляции, евклидово расстояние, мера информации. Это позволяет исключить не имеющие значения данные, то есть уменьшить размер-

ность массива данных для обучения. При этом появляется возможность упростить модель, а также повысить ее точность и скорость обучения.

В большинстве исследований в области прогнозирования электропотребления с помощью машинного обучения используется стандартный набор признаков для обучения моделей: предыдущее электропотребление, порядковый номер дня в неделе, час суток, признак выходного (праздничного) дня, температура, количество осадков, облачность и т. д. Новизна данной работы заключается, в частности, в применении графика ремонтных работ для построения прогнозов.

МЕТОДЫ ИССЛЕДОВАНИЯ

Постановка задачи выглядит следующим образом: дана выборка входных данных (X, T, H, R) ; необходимо создать модель, которая на основе этих данных строит прогноз электропотребления на 24 часа вперед:

$$\hat{y} = f(X, T, H, R); \quad (1)$$

$$X = \{x_1, x_2, \dots, x_{24m}\}; \quad (2)$$

$$R = \{0, 1\}; \quad (3)$$

$$H = \{0, 1\}; \quad (4)$$

$$T = \{t_1, t_2, \dots, t_5\}, \quad (5)$$

где X – множество параметров, соответствующих ретроспективному почасовому потреблению мощности; x_1, x_2, \dots, x_{24m} – данные почасового электропотребления за m предыдущих суток, кВт; H – бинарная переменная, принимающая значение 1, если рассматриваемый час приходится на выходной или праздничный день, и 0 в обратном случае; R – бинарная переменная, принимающая значение 1, если рассматриваемый час попадает в период ремонтных работ, и 0 в обратном случае; T – множество параметров, соответствующих информации о време-

ни начала прогноза; t_1, t_2, \dots, t_5 – час суток, день недели, число, месяц и год часа, соответствующий началу прогноза; \hat{y} – прогноз электропотребления на сутки вперед.

В качестве метрик качества модели были выбраны средний модуль ошибки (MAE от англ. Mean Absolute Error) и средний модуль процентной ошибки (MAPE от англ. Mean Absolute Percentage Error):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|; \quad (6)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\%, \quad (7)$$

где n – количество часов; \hat{y}_i – прогноз электропотребления в i -й час; y_i – фактическое значение электропотребления в i -й час.

Для решения поставленной задачи были рассмотрены два вида моделей машинного обучения: ансамбли деревьев решений и глубокая полносвязная нейронная сеть.

Основной подход к решению задач обучения по прецедентам заключается в поиске такого набора параметров алгоритма из заданного набора параметров A , который даст минимум средней ошибки алгоритма на всей выборке X^m :

$$a = \arg \min_{a \in A} \frac{1}{m} \sum_{i=1}^m l_i(a(x_i), y^*(x_i)), \quad (8)$$

где i – номер образца; $l_i(a(x_i), y^*(x_i))$ – функция потерь, показывающая, насколько отклоняется ответ алгоритма $\hat{y}_i = a(x_i)$ от правильного ответа $y_i = y^*(x_i)$ на произвольном объекте $x \in X$. При этом вид функции потерь зависит от типа решаемой задачи. Поставленная задача представляет собой задачу регрессии, и в ходе исследования были рассмотрены следующие функции потерь:

– среднеквадратичная функция потерь (Mean Squared Error Loss, MSE), где

$$l_i(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2, \quad (9)$$

– средняя по модулю функция потерь (Mean Absolute Error Loss, MAE), где

$$l_i(y_i, \hat{y}_i) = |y_i - \hat{y}_i|, \quad (10)$$

– функция потерь Хьюбера (Huber Loss) [18], где

$$l_i(y_i, \hat{y}_i) = \begin{cases} \frac{1}{2}(y_i - \hat{y}_i)^2 & \text{для } |y_i - \hat{y}_i| \leq \delta, \\ \delta \left(|y_i - \hat{y}_i| - \frac{1}{2}\delta \right) & \text{иначе.} \end{cases} \quad (11)$$

В ходе работы было установлено, что в данной задаче наилучшие результаты при построении ансамблей деревьев решений получаются при использовании среднеквадратичной функции потерь (9), а при создании нейронной сети – средней по модулю функции потерь (10).

В работе были использованы следующие подходы к построению ансамблевых моделей: адаптивный бустинг над решающими деревьями (AdaBoost, AB), градиентный бустинг над решающими деревьями (Gradient Boosting, GB) и случайный лес (Random Forest, RF) [19].

Выбор ансамблевых методов обусловлен наличием бинарных признаков H и R , учитывающих выходные дни и проведение ремонтных работ в определенные моменты времени, соответственно. Перечисленные ансамблевые методы в качестве базовых алгоритмов используют решающие деревья, в которых для решения поставленной задачи составляется набор простых логических правил. Это позволяет им эффективно учитывать признаки, не являющиеся количественными.

Принципиально случайный лес отличается от остальных перечисленных методов типом используемого ансамбля. В случайном лесе для обучения применяется бэггинг, в котором обучающая выборка разделяется на несколько случайных подвыборок, при этом один и тот же элемент может попадать в раз-

ные подвыборки. Каждый отдельный регрессор в этом случае обучается независимо от остальных, и ответы всех решающих деревьев обобщаются и усредняются. Итоговая композиция строится как среднее всех N базовых алгоритмов $b_1(x), \dots, b_n(x)$, каждый из которых обучен на подвыборке X_1, \dots, X_N , соответственно:

$$a_N(x) = \frac{1}{N} \sum_{n=1}^N b_n(x). \quad (12)$$

В бустинге же решающие деревья строятся последовательно, каждый следующий регрессор зависит от результатов предыдущего и стремится исправить его ошибки. Окончательный ответ модели в этом случае определяется как средневзвешенная сумма базовых алгоритмов, для каждого из которого определяется коэффициент (вес) γ :

$$a_N(x) = \sum_{n=1}^N \gamma_n b_n(x). \quad (13)$$

Принципиальное отличие адаптивного бустинга от градиентного бустинга заключается в способе учета ошибок ансамбля. В адаптивном бустинге снижение ошибки достигается путем увеличения веса тех объектов обучающей выборки, на которых была допущена ошибка. В градиентном бустинге снижение ошибки ансамбля достигается с помощью градиентного спуска – каждый шаг делается вдоль направления наискорейшего убывания ошибки, и каждый регрессор строится так, чтобы получить максимальное снижение ошибки модели.

Стоит отметить, что использование всех доступных данных для обучения модели не всегда целесообразно, а в некоторых случаях это даже может негативно сказаться на качестве модели. Поэтому отбор наиболее информативных признаков при работе с такими моделями особенно важен.

В общем случае структуру любой нейронной сети можно представить в виде набора связанных между собой входного, скрытых и

выходного слоев.

Каждый нейрон скрытого слоя представляет собой вычислительную единицу, последовательно применяющую линейную функцию на входные переменные и нелинейную функцию активации на результат; i -й нейрон можно представить в виде уравнений как

$$\begin{aligned} z_i &= W^T x + b, \\ a_i &= g(z_i), \end{aligned} \quad (14)$$

где W – матрица весов (weights); x – вектор входных значений; b – свободный коэффициент (сдвиг, bias); g – нелинейная функция активации.

Существует множество видов функций активации: сигмоидная (Sigmoid), гиперболический тангенс, линейный выпрямитель (Rectified Linear Unit, ReLU), линейный выпрямитель «с утечкой» (Leaky ReLU) [20]. В данной модели применяются две такие функции: сигмоидная и ReLU.

Сигмоидная функция определяется как

$$g(z) = \frac{1}{1 + e^{-z}}. \quad (15)$$

Сигмоидная функция часто используется как функция активации последнего нейрона сети, поскольку она принимает значения из отрезка $(0; 1)$ и непрерывно дифференцируема на всей области определения, что очень важно для процесса обучения нейронной сети.

Уравнение ReLU имеет следующий вид:

$$g(z) = \max(0, z). \quad (16)$$

ReLU является наиболее распространенной на практике функцией активации нейронов скрытых слоев по нескольким причинам. Во-первых, эта функция проста с вычислительной точки зрения, что позволяет ускорить обучение нейронной сети. Во-вторых, ReLU в целом ведет себя как линейная функция активации; как правило, оптимизировать модель проще, когда ее вид близок к линейному. В-третьих, в отличие от сигмоид-

ной функции, не имеет недостатка в виде исчезающих градиентов: ее производная принимает постоянное значение на всей области определения.

В работе используется нейронная сеть прямого распространения (или глубокая нейронная сеть) с шестью скрытыми слоями. Подробно архитектура используемой сети приведена на рис. 1.

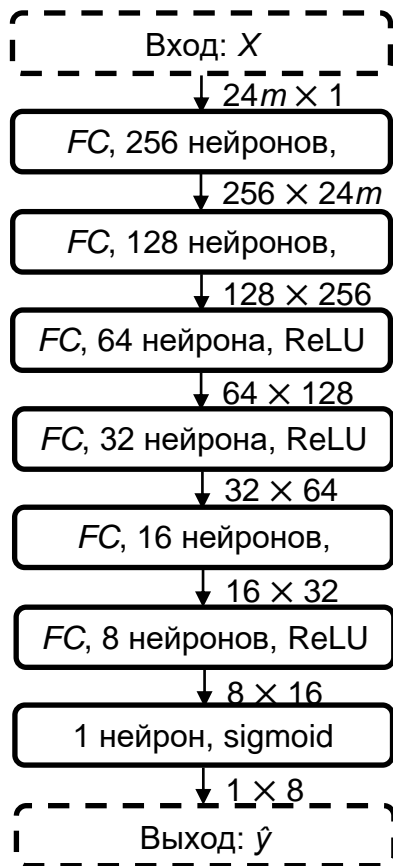


Рис. 1. Архитектура используемой нейронной сети (m – число суток, FC – полносвязный слой (fully connected), \hat{y} – предсказание модели)

Fig. 1. Architecture of the used neural network (m – number of days, FC – fully connected layer, \hat{y} – model prediction)

Минимум (или значение, близкое к минимуму) средней ошибки алгоритма на всей выборке $L(y, \hat{y})$ достигается путем оптимизации нейронной сети, то есть нахождением такой матрицы весов W , при которой значение $L(y, \hat{y})$ минимально. Для оптимизации нейронных сетей используется множество методов, но наиболее распространенным является метод обратного распространения

ошибки (Back-propagation) с двумя способами реализации: стохастический градиентный спуск (Stochastic Gradient Descent, SGD) и пакетный градиентный спуск (Batch Gradient Descent) [21].

SGD представляет собой самый тривиальный алгоритм оптимизации. Обновление параметров модели происходит после вычисления градиента функции потерь на каждом объекте и движения в противоположную сторону, т. е. в направлении наискорейшего убывания функции потерь:

$$W = W - \alpha dW, \quad (17)$$

где α – скорость обучения (learning rate) – коэффициент, задающий величину шага в направлении антиградиента – один из наиболее важных гиперпараметров модели, от которого зависит скорость сходимости алгоритма оптимизации; dW – матрица градиентов $dW = \partial L / \partial W$.

SGD обладает существенным недостатком: множество параметров нейронной сети в большинстве случаев расположено в многомерном пространстве, из-за чего функция потерь может быть намного сильнее чувствительна к изменениям в одном направлении и менее чувствительна к изменениям в другом. В результате появляются шумы, и процесс обучения замедляется.

Пакетный градиентный спуск вычисляет градиент на всех образцах вместе, что делает этот метод несколько стабильнее по сравнению с SGD, но в случае большого объема выборки данных каждая итерация может занимать слишком много времени, что, безусловно, приводит к снижению скорости обучения.

В качестве компромисса часто применяют мини-пакетный градиентный спуск (Mini-batch Gradient Descent) [22]. В этом случае обновление весов происходит после обработки заранее заданного количества образцов, что позволяет компенсировать недостатки двух других способов. В разработанной нейросетевой модели применяется данный способ для 32 объектов.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

В рассматриваемой задаче исходной информацией является почасовое потребление мощности промышленного деревообрабатывающего предприятия за 2 года (в период с 01.01.2019 по 31.12.2020), а также график планово-профилактических работ (ППР) деревообрабатывающих станков на этот период.

Для построения модели необходимо выделить наиболее значимые признаки из имеющейся выборки. С этой целью был рассчитан коэффициент корреляции между электропотреблением в прогнозируемый час и некоторыми признаками. На рис. 2 приведено изменение коэффициента корреляции Спирмена между электропотреблением в рассматриваемый час и в предыдущие часы за две недели.

Можно наблюдать возрастание коэффи-

циента корреляции между электропотреблением в текущий час и в часы, отстающие от текущего на кратное число суток (отмеченные точки на рис. 2); наибольшая корреляция наблюдается между текущим часом и часами, отстающими от текущего на 168 и 336 ч, то есть в семидневные интервалы. Таким образом, в технологическом процессе исследуемого предприятия отчетливо выделяются недельные циклы. Из этого можно сделать вывод, что для построения прогноза наиболее важными признаками являются значения потребляемой энергии в те часы, которые отстают от рассматриваемого на кратное число суток.

В табл. 1 приведены эмпирически подобранные гиперпараметры используемых в работе моделей. Исходная выборка была поделена на обучающую и тестовую в отношении 90 на 10.

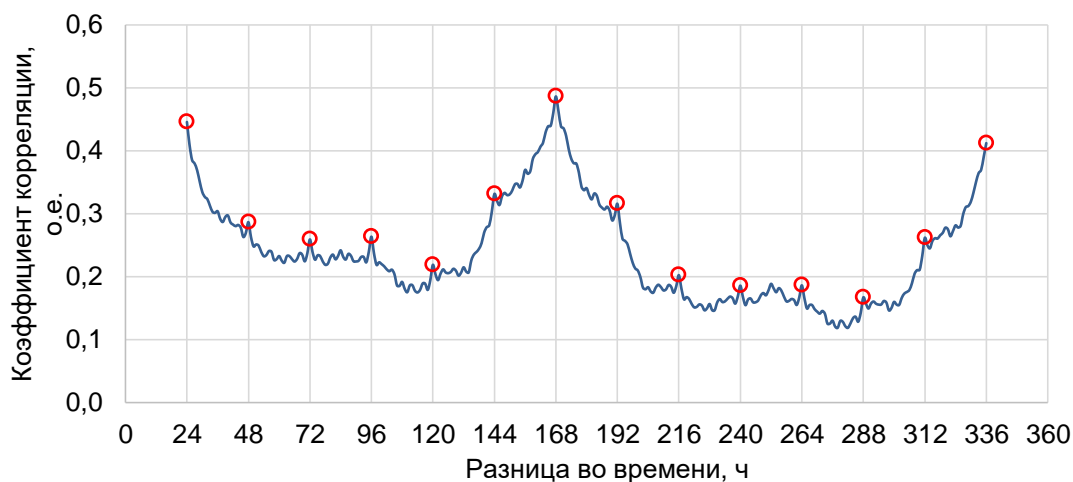


Рис. 2. Коэффициенты корреляции между электропотреблением в рассматриваемый час и в предыдущие часы
Fig. 2. Coefficients of correlation between the power consumption in the considered hour and the preceding hours

Таблица 1. Гиперпараметры моделей
Table 1. Model hyperparameters

Модель	Гиперпараметры
AdaBoost	максимальная глубина дерева – 8 размер ансамбля – 50
Gradient Boosting	максимальная глубина дерева – 5 размер ансамбля – 50
Random Forest	максимальная глубина дерева – 7 размер ансамбля – 30
Deep NN	количество скрытых слоев – 6 алгоритм оптимизации – adam коэффициент скорости обучения – 0,001 количество эпох – 100 число пакетов – 32

Для оценки влияния выбора признаков на точность прогноза ансамблевых моделей были рассмотрены 4 варианта исходных данных для обучения модели:

- 1) электропотребление за каждый час с учетом ремонтных работ;
- 2) электропотребление за каждый час без учета ремонтных работ;
- 3) электропотребление только за отобранные часы с учетом ремонтных работ;
- 4) электропотребление только за отобранные часы без учета ремонтных работ.

Для нейросетевой модели рассматривались только варианты 1 и 2, поскольку при обучении ИНС важно иметь как можно большую выборку данных [23].

Оценки точности моделей, обученных по каждому из перечисленных вариантов выбора признаков, приведены в табл. 2 и 3.

По полученным результатам можно видеть, что в ансамблевых моделях использование всего рассматриваемого временного ряда не дает существенного прироста точности по сравнению с использованием данных только по наиболее значимым часам, а в некоторых случаях может приводить и к сниже-

нию точности.

Учет графика ППР также позволяет заметно повысить точность прогноза. Так, для наилучшей по качеству модели Gradient Boosting снижение ошибки составило 1,4 процентных пункта (с 6,9 до 5,5) или 20%. На рис. 3 показаны истинный график электропотребления и графики прогноза электропотребления, полученные с помощью градиентного бустинга на тестовой выборке как с использованием данных по графику ППР, так и без них.

На рис. 3 приведен наглядный пример того, как учет ППР позволяет данной модели достигать высокой точности на участках низкого электропотребления, связанных с проведением крупных ремонтных работ. В то же время без этих данных модель на указанных промежутках сильнее отклоняется от истинного графика.

В нейросетевой модели использование графика ППР не привело к существенному снижению ошибки по сравнению с ансамблевыми моделями. Здесь снижение ошибки составило всего 0,5 процентных пункта, или 6,7%.

Таблица 2. Оценка точности моделей MAPE, %
Table 2. MAPE model accuracy metric, %

Модель	Используемые признаки			
	Все часы, с ППР	Все часы, без ППР	Отобранные часы, с ППР	Отобранные часы, без ППР
AdaBoost	6,2	8,1	6,3	7,6
Gradient Boosting	5,6	7,1	5,5	6,9
Random Forest	5,8	7,6	5,8	7,5
Deep NN	7,0	7,5	–	–

Таблица 3. Оценка точности моделей MAE, кВт
Table 3. MAE model accuracy metric, kW

Модель	Используемые признаки			
	Все часы, с ППР	Все часы, без ППР	Отобранные часы, с ППР	Отобранные часы, без ППР
AdaBoost	206,5	249,2	212,4	233,6
Gradient Boosting	180,5	209,7	179,1	205,6
Random Forest	190,9	223,1	189,5	219,9
Deep NN	196,5	207,6	–	–



Рис. 3. Сопоставление истинного графика электропотребления и прогнозов
Fig. 3. Comparison of the true curve of power consumption and forecasts

ЗАКЛЮЧЕНИЕ

Был выполнен сравнительный анализ ансамблевых и нейросетевых моделей машинного обучения в задаче прогнозирования электропотребления предприятий при различных подходах к отбору признаков. С помощью определения коэффициентов корреляции были определены наиболее значимые часы рассматриваемого временного ряда. Для сравнения были построены модели на нескольких вариантах исходных данных для обучения.

Было установлено, что использование только отобранных часов вместо всего временного ряда для построения прогноза с помощью ансамблевых моделей не приводит к снижению точности, а в отдельных случаях может ее и повышать. Кроме того, это также позволяет на порядок снизить длительность

обучения модели за счет понижения размерности рассматриваемого массива признаков.

Полученные результаты показывают, что использование графика ремонтных работ также повышает точность прогноза. Для рассмотренных ансамблевых моделей снижение ошибки составило от 20 до 30%.

Нейронная сеть в данном случае оказалась менее эффективной по сравнению с ансамблями деревьев решений: ниже точность в целом извлекает меньше информации из графика ППР, что можно связать с недостаточным объемом имеющейся выборки данных.

Таким образом, потенциал повышения точности прогнозирования лежит в использовании дополнительных данных о техпроцессе рассматриваемого предприятия и в выборе подходящих методов.

Список источников

1. Lozinskaia A., Redkina A., Shenkman E. Electricity consumption forecasting for integrated power system with seasonal patterns // *Applied Econometrics*. 2020. Vol. 60. P. 5–25. <https://doi.org/10.22394/1993-7601-2020-60-5-25>.
2. Филиппова Т. А., Русина А. Г., Дронова Ю. В. Модели и методы прогнозирования электроэнергии и мощности при управлении режимами электроэнергетических систем: монография. Новосибирск: Изд-во НГТУ,

2009. 368 с.
3. Hernandez L., Baladron C., Aguiar J. M., Carro B., Sanchez-Esguevillas A. J., Lloret J., et al. A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings // *IEEE Communications Surveys & Tutorials*. 2014. Vol. 16. Iss. 3. P. 1460–1495. <https://doi.org/10.1109/SURV.2014.032014.00094>.
4. Deb Chirag, Zhang Fan, Yang Junjing, Lee Siew Eang, Shah Kwok Wei. A review on time series forecasting tech-

- niques for building energy consumption // *Renewable and Sustainable Energy Reviews*. 2017. Vol. 74. P. 902–924. <https://doi.org/10.1016/j.rser.2017.02.085>.
5. Hahn Heiko, Meyer-Nieberg Silja, Pickl Stefan. Electric load forecasting methods: tools for decision making // *European Journal of Operational Research*. 2009. Vol. 199. Iss. 3. P. 902–907. <https://doi.org/10.1016/j.ejor.2009.01.062>.
6. Cai M., Pipattanasomporn M., Rahman S. Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques // *Applied Energy*. 2019. Vol. 236. P. 1078–1088. <https://doi.org/10.1016/j.apenergy.2018.12.042>.
7. Qiu Xueheng, Zhang Le, Ren Ye, Suganthan P. N., Amaratunga G. Ensemble deep learning for regression and time series forecasting // *IEEE Symposium on Computational Intelligence in Ensemble Learning*. 2014. <https://doi.org/10.1109/CIEL.2014.7015739>.
8. Поляхов Н. Д., Приходько И. А., Ван Ефэн. Прогнозирование электропотребления на основе метода опорных векторов // *Известия Санкт-Петербургского государственного электротехнического университета*. 2014. № 10. С. 26–30.
9. Yildiz B., Bilbao J. I., Sproul A. B. A review and analysis of regression and machine learning models on commercial building electricity load forecasting // *Renewable and Sustainable Energy Reviews*. 2017. Vol. 73. P. 1104–1122. <https://doi.org/10.1016/j.rser.2017.02.023>.
10. Yang Jingfei, Stenzel J. Short-term load forecasting with increment regression tree // *Electric Power Systems Research*. 2006. Vol. 76. Iss. 9-10. P. 880–888. <https://doi.org/10.1016/j.epsr.2005.11.007>.
11. Matrenin P., Safaraliev M., Dmitriev S., Kokin S., Ghulomzoda A., Mitrofanov S. Medium-term load forecasting in isolated power systems based on ensemble machine learning models // *Energy Reports*. 2022. Vol. 8. P. 612–618. <https://doi.org/10.1016/j.egyr.2021.11.175>.
12. Matrenin P., Antonenkov D., Manusov V. Recurrent and ensemble models for short-term load forecasting of coal mining companies // *Ural-Siberian Smart Energy Conference*. 2021. <https://doi.org/10.1109/USSEC53120.2021.9655732>.
13. Neupane B., Perera K. S., Aung Zeyar, Woon Wei Lee. Artificial neural network-based electricity price forecasting for smart grid deployment // *IEEE International Conference on Computer Systems and Industrial Informatics*. 2012. <https://doi.org/10.1109/ICCSII.2012.6454392>.
14. Amarasinghe K., Marino D. L., Manic M. Deep neural networks for energy load forecasting // *IEEE 26th International Symposium on Industrial Electronics (Edinburgh, 19–21 June 2017)*. Edinburgh: IEEE, 2017. P. 1483–1488. <https://doi.org/10.1109/ISIE.2017.8001465>.
15. Jie Cai, Jiawei Luo, Shulin Wang, Sheng Yang. Feature selection in machine learning: a new perspective // *Neurocomputing*. 2018. Vol. 300. P. 70–79. <https://doi.org/10.1016/j.neucom.2017.11.077>.
16. Koprinska I., Rana M., Agelidis V. G. Correlation and instance based feature selection for electricity load forecasting // *Knowledge-Based Systems*. 2015. Vol. 82. P. 29–40. <https://doi.org/10.1016/j.knosys.2015.02.017>.
17. Bouktif S., Fiaz A., Ouni A., Serhani M. A. Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches // *Energies*. 2018. Vol. 11. Iss. 7. P. 1636. <https://doi.org/10.3390/en11071636>.
18. Huber P. J. Robust estimation of a location parameter // *The Annals of Mathematical Statistics*. 1964. Vol. 35. Iss. 1. P. 73–101. <https://doi.org/10.1214/aoms/1177703732>.
19. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., et al. Scikit-learn: machine learning in Python // *Journal of Machine Learning Research*. 2011. Vol. 12. Iss. 85. P. 2825–2830.
20. Ramachandran P., Zoph B., Le Quoc V. Searching for activation functions. 2017. <https://doi.org/10.48550/arXiv.1710.05941>.
21. Hoffer E., Hubara I., Soudry D. Train longer, generalize better: closing the generalization gap in large batch training of neural networks. 2017. <https://doi.org/10.48550/arXiv.1705.08741>. <https://doi.org/10.48550/arXiv.1705.08741>.
22. Khirirat S., Feyzmahdavian H. R., Johansson M. Mini-batch gradient descent: faster convergence under data sparsity // *IEEE 56th Annual Conference on Decision and Control (Melbourne, 12–15 December 2017)*. Melbourne: IEEE, 2017. P. 2880–2887. <https://doi.org/10.1109/CDC.2017.8264077>.
23. Linjordet T., Balog K. Impact of training dataset size on neural answer selection models // *Advances in Information Retrieval. Lecture Notes in Computer Science / eds. L. Azzopardi, B. Stein, N. Fuhr, P. Mayr, C. Hauff, D. Hiemstra*. Vol. 11437. Cham: Springer, 2019. P. 828–835. https://doi.org/10.1007/978-3-030-15712-8_59.

References

1. Lozinskaia A., Redkina A., Shenkman E. Electricity consumption forecasting for integrated power system with seasonal patterns. *Applied Econometrics*. 2020;60:5-25. <https://doi.org/10.22394/1993-7601-2020-60-5-25>.
2. Filippova T. A., Rusina A. G., Dronova Yu. V. *Models and methods of electric energy and power forecasting in power system operation management: monograph*. Novosibirsk: Novosibirsk State Technical University; 2009, 368 p. (In Russ.).
3. Hernandez L., Baladron C., Aguiar J. M., Carro B., Sanchez-Esguevillas A. J., Lloret J., et al. A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings. *IEEE Communications Surveys & Tutorials*. 2014;16(3):1460-1495. <https://doi.org/10.1109/SURV.2014.032014.00094>.
4. Deb Chirag, Zhang Fan, Yang Junjing, Lee Siew Eang, Shah Kwok Wei. A review on time series forecasting techniques for building energy consumption. *Renewable and*

- Sustainable Energy Reviews*. 2017;74:902-924. <https://doi.org/10.1016/j.rser.2017.02.085>.
5. Hahn Heiko, Meyer-Nieberg Silja, Pickl Stefan. Electric load forecasting methods: tools for decision making. *European Journal of Operational Research*. 2009;199(3):902-907. <https://doi.org/10.1016/j.ejor.2009.01.062>.
6. Cai M., Pipattanasomporn M., Rahman S. Day-ahead building-level load forecasts using deep learning vs traditional time-series techniques. *Applied Energy*. 2019;236:1078-1088. <https://doi.org/10.1016/j.apenergy.2018.12.042>.
7. Qiu Xueheng, Zhang Le, Ren Ye, Suganthan P. N., Amaratunga G. Ensemble deep learning for regression and time series forecasting. In: *IEEE Symposium on Computational Intelligence in Ensemble Learning*. 2014. <https://doi.org/10.1109/CIEL.2014.7015739>.
8. Polyakhov N. D., Prikhodko I. A., Efen Van. Electric load forecasting based on support vector machine optimized. *Izvestiya Sankt-Peterburgskogo gosudarstvennogo elektrotekhnicheskogo universiteta*. 2014;10:26-30. (In Russ.).
9. Yildiz B., Bilbao J. I., Sproul A. B. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*. 2017;73:1104-1122. <https://doi.org/10.1016/j.rser.2017.02.023>.
10. Yang Jingfei, Stenzel J. Short-term load forecasting with increment regression tree. *Electric Power Systems Research*. 2006;76(9-10):880-888. <https://doi.org/10.1016/j.epsr.2005.11.007>.
11. Matrenin P., Safaraliev M., Dmitriev S., Kokin S., Ghulomzoda A., Mitrofanov S. Medium-term load forecasting in isolated power systems based on ensemble machine learning models. *Energy Reports*. 2022;8:612-618. <https://doi.org/10.1016/j.egyr.2021.11.175>.
12. Matrenin P., Antonenkov D., Manusov V. Recurrent and ensemble models for short-term load forecasting of coal mining companies. In: *Ural-Siberian Smart Energy Conference*. 2021. <https://doi.org/10.1109/USSEC53120.2021.9655732>.
13. Neupane B., Perera K. S., Aung Zeyar, Woon Wei Lee. Artificial neural network-based electricity price forecasting for smart grid deployment. In: *IEEE International Conference on Computer Systems and Industrial Informatics*. 2012. <https://doi.org/10.1109/ICCSII.2012.6454392>.
14. Amarasinghe K., Marino D. L., Manic M. Deep neural networks for energy load forecasting. In: *IEEE 26th International Symposium on Industrial Electronics*. 19–21 June 2017, Edinburgh. Edinburgh: IEEE; 2017, p. 1483-1488. <https://doi.org/10.1109/ISIE.2017.8001465>.
15. Jie Cai, Jiawei Luo, Shulin Wang, Sheng Yang. Feature selection in machine learning: a new perspective. *Neurocomputing*. 2018;300:70-79. <https://doi.org/10.1016/j.neucom.2017.11.077>.
16. Koprinska I., Rana M., Agelidis V. G. Correlation and instance based feature selection for electricity load forecasting. *Knowledge-Based Systems*. 2015;82:29-40. <https://doi.org/10.1016/j.knosys.2015.02.017>.
17. Bouktif S., Fiaz A., Ouni A., Serhani M. A. Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches. *Energies*. 2018;11(7):1636. <https://doi.org/10.3390/en11071636>.
18. Huber P. J. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*. 1964;35(1):73-101. <https://doi.org/10.1214/aoms/1177703732>.
19. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., et al. Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*. 2011;12(85):2825-2830.
20. Ramachandran P., Zoph B., Le Quoc V. Searching for activation functions. 2017. <https://doi.org/10.48550/arXiv.1710.05941>.
21. Hoffer E., Hubara I., Soudry D. Train longer, generalize better: closing the generalization gap in large batch training of neural networks. 2017. <https://doi.org/10.48550/arXiv.1705.08741>.
22. Khirirat S., Feyzmahdavian H. R., Johansson M. Mini-batch gradient descent: faster convergence under data sparsity. In: *IEEE 56th Annual Conference on Decision and Control*. 12–15 December 2017, Melbourne. Melbourne: IEEE; 2017, p. 2880-2887. <https://doi.org/10.1109/CDC.2017.8264077>.
23. Linjordet T., Balog K. Impact of training dataset size on neural answer selection models. In: Azzopardi L., Stein B., Fuhr N., Mayr P., Hauff C., Hiemstra D. (eds.). *Advances in Information Retrieval. Lecture Notes in Computer Science*. Vol. 11437. Cham: Springer; 2019, p. 828-835. https://doi.org/10.1007/978-3-030-15712-8_59.

ИНФОРМАЦИЯ ОБ АВТОРАХ

Сергеев Никита Николаевич,
лаборант
межкафедральной Научно-исследовательской
лаборатории обработки, анализа и представления
данных в электроэнергетических системах,
Новосибирский государственный технический
университет,
630073, г. Новосибирск, пр-т К. Маркса, 20, Россия

INFORMATION ABOUT THE AUTHORS

Nikita N. Sergeev,
Laboratory Assistant
Interdepartmental Research Laboratory
for Processing, Analysis and Presentation of Data in
Power Systems,
Novosibirsk State Technical University,
20, K. Marks pr., Novosibirsk 630073, Russia

Матренин Павел Викторович,
кандидат технических наук,
старший научный сотрудник
межкафедральной Научно-исследовательской
лаборатории обработки, анализа и представления
данных в электроэнергетических системах,
Новосибирский государственный технический
университет,
630073, г. Новосибирск, пр-т К. Маркса, 20, Россия

Pavel V. Matrenin,
Cand. Sci. (Eng.),
Senior Researcher
Interdepartmental Research Laboratory
for Processing, Analysis and Presentation of Data in
Power Systems,
Novosibirsk State Technical University,
20, K. Marks pr., Novosibirsk 630073, Russia

Вклад авторов

Все авторы сделали эквивалентный вклад в подготовку публикации.

Contribution of the authors

The authors contributed equally to the preparation of the article.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Conflict of interests

The authors declare no conflicts of interests.

Все авторы прочитали и одобрили окончательный вариант рукописи.

The final manuscript has been read and approved by all the co-authors.

Информация о статье

Статья поступила в редакцию 23.06.2022; одобрена после рецензирования 05.08.2022; принята к публикации 21.09.2022.

Information about the article

The article was submitted 23.06.2022; approved after reviewing 05.08.2022; accepted for publication 21.09.2022.