

Прогнозирование потребления электроэнергии предприятиями народнохозяйственного комплекса в условиях неполноты информации

И. Д. Моргоев¹, А. Э. Дзгоев¹, Р. В. Ключев², А. Д. Моргоева¹

¹Северо-Кавказский горно-металлургический институт
(государственный технологический университет)
362021, Россия, Владикавказ, ул. Николаева, 44

²Московский политехнический университет
107023, Россия, Москва, ул. Б. Семеновская, 38

Аннотация. В работе рассмотрена проблема планирования спроса на электроэнергию для сбытовых организаций с помощью интеллектуального анализа данных. В силу того, что планирование объемов потребления открывает новые экономические возможности для предприятий при выходе на оптовый рынок электроэнергии, прогнозирование является необходимым экономическим рычагом для принятия оптимальных решений в процессе планирования и распределения ресурсов. Таким образом, целью проведенного исследования явилось получение достоверного прогноза потребления электроэнергии. Стоит отметить, что прогнозирование потребления электроэнергии позволит повысить эффективность принимаемых управленческих решений как для электросетевых компаний, так и для отдельных энергоемких потребителей (промышленных предприятий). В ходе исследования был применен комплекс методов научного познания, в том числе машинного обучения. В результате было построено несколько моделей машинного обучения, с помощью которых сделан прогноз потребления электроэнергии. Проведен сравнительный анализ результатов прогнозирования по метрикам качества: средней абсолютной ошибке прогноза и коэффициенту детерминации. Наилучшие значения указанных метрик были получены с помощью модели, основанной на алгоритме CatBoostRegressor. Стало быть, с целью прогнозирования электропотребления использование разработанной модели, на наш взгляд, будет наиболее целесообразно.

Ключевые слова: электроэнергетика, машинное обучение, регрессия, кластеризация, прогнозирование

Поступила 03.06.2022, одобрена после рецензирования 10.06.2022, принята к публикации 15.06.2022

Для цитирования: Моргоев И. Д., Дзгоев А. Э., Ключев Р. В., Моргоева А. Д. Прогнозирование потребления электроэнергии предприятиями народнохозяйственного комплекса в условиях неполноты информации // Известия Кабардино-Балкарского научного центра РАН. 2022. № 3 (107). С. 9–20. DOI: 10.35330/1991-6639-2022-3-107-9-20

MSC: 05-04

Analytical article

Forecasting the consumption of electricity by enterprises of the national economy complex in conditions of incomplete information

I.D. Morgoev¹, A.E. Dzgoev¹, R.V. Klyuev², A.D. Morgoeva¹

¹The North Caucasian Institute of Mining and Metallurgy
(State Technological University)
362011, Russia, Vladikavkaz, 44 Nikolaev street

²Moscow Polytechnic University
107023, Russia, Moscow, 38 B. Semenovskaya street

Abstract. The paper considers the problem of planning the demand for electricity for sales organizations using intellectual data analysis. Due to the fact that planning of consumption volumes opens up new economic opportunities for enterprises when entering the wholesale electricity market, forecasting is a necessary economic lever for making optimal decisions in the process of planning and allocating resources. Thus, the purpose of the study was to obtain a reliable forecast of electricity consumption. It should be noted that the forecasting of electricity consumption will improve the efficiency of management decisions for both electric grid companies and individual energy-intensive consumers (industrial enterprises). In the course of the study, a set of methods of scientific knowledge, including machine learning methods, was applied. As a result, several machine learning models were built, with the help of which a forecast of electricity consumption was made. A comparative analysis of the results of forecasting by quality metrics was carried out: the average absolute error of the forecast and the coefficient of determination. The best values of these metrics were obtained using a model based on the CatBoostRegressor algorithm. Therefore, in order to predict power consumption, the use of the developed model, in our opinion, will be most appropriate.

Key words: electric power industry, machine learning, regression, clustering, forecasting

Submitted 03.06.2022,

approved after reviewing 10.06.2022,

accepted for publication 15.06.2022

For citation. Morgoev I.D., Dzgoev A.E., Klyuev R.V., Morgoeva A.D. Forecasting the consumption of electricity by enterprises of the national economy complex in conditions of incomplete information. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2022. No. 3 (107). Pp. 9–20. DOI: 10.35330/1991-6639-2022-3-107-9-20

ВВЕДЕНИЕ

Как известно, электроэнергетика – основа народного хозяйства, так как ее успешное функционирование влияет на производственный потенциал государства и качество жизни населения. В связи с высокой важностью для страны этой областью хозяйства необходимо оптимально управлять. Прогнозирование как научно-аналитический этап процесса планирования или одна из функций «цикла управления» социально-экономическими системами [1] является одним из эффективных инструментов в управленческом процессе, позволяющим минимизировать перепроизводство и нехватку электроэнергии, что позитивно влияет на коммерческую рентабельность, экологическую обстановку и уровень жизни населения.

Эффективность при прогнозировании электропотребления, оцениваемая как соответствие требуемого и фактического результатов [1], достигается за счет решения минимаксной задачи, сформулированной следующим образом: при минимальном количестве ресурсов необходимо обеспечить потребителей всей необходимой энергией. Подобная задача очень часто возникает для электросетевых организаций и крупных промышленных предприятий.

Для сбытовых (электросетевых) организаций эта проблема обусловлена тем, что организации данного типа должны рассчитывать спрос на электроэнергию при ее генерации или покупке на оптовом рынке. Кроме того, подобный независимый расчет также может быть использован как фактор при выявлении коммерческих потерь электроэнергии, потому что в настоящее время одной из достаточно серьезных проблем на этапе передачи электроэнергии конечному потребителю для электросетевых компаний являются коммерческие потери электроэнергии [2].

Согласно данным Росстата¹, распределение электропотребления по секторам народного хозяйства Российской Федерации для электросетевых компаний за 2020 год выглядит, как показано на рисунке 1.

¹ Энергобаланс [Электронный ресурс]. Режим доступа: <https://rosstat.gov.ru/storage/mediabank/el-balans.xlsx> (дата обращения: 25.04.2022).

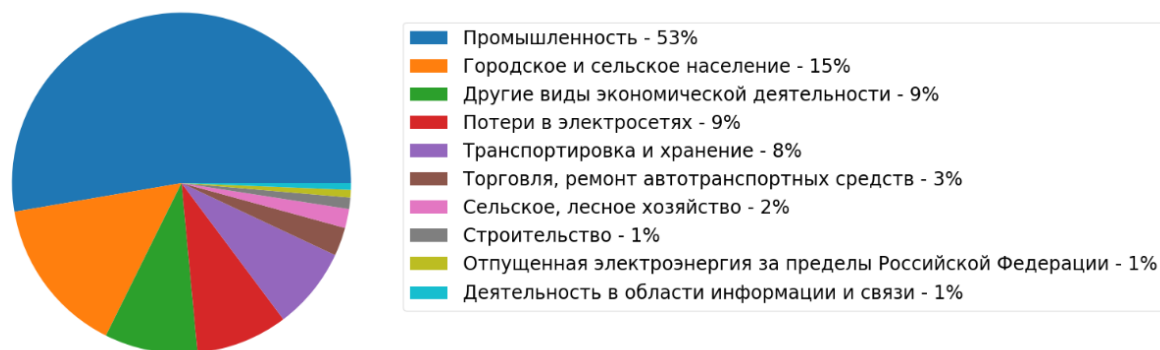


Рис. 1. Электропотребление по секторам народного хозяйства Российской Федерации за 2020 год

Fig. 1. Electricity consumption by sectors of the national economy of Russian Federation for 2020

Можно заметить, что наибольшее потребление электроэнергии приходится на промышленность и составляет 53 %. Соответственно, для электросетевых организаций прогнозирование потребления электрической энергии крупными предприятиями имеет высокую практическую значимость. Это объясняется тем, что данная категория потребителей электроэнергии является основной из-за большой энергоемкости применяемых на предприятиях технологических процессов.

В свою очередь для промышленных предприятий в условиях рыночных отношений при установлении цен на закупку электроэнергии особенно актуальной становится задача прогнозирования собственного электропотребления, так как это позволяет осуществлять закупку энергии по более выгодным тарифам, уменьшая тем самым производственные затраты. К экономически выгодным условиям², предлагаемым на оптовом рынке закупки электроэнергии, относятся, например, свободные двусторонние договоры, рынок на сутки вперед, балансирующий рынок.

В данном исследовании была поставлена цель – построить точный прогноз потребления электроэнергии по данным ее расхода предприятиями, относящимися к различным секторам экономики, с учетом интересов электросетевой компании (гарантирующего поставщика).

Гарантирующий поставщик³ – участник оптового и розничных рынков электрической энергии, в обязанности которого входит заключение договоров с любым обратившимся к нему потребителем, который расположен в границах зоны его деятельности.

Для достижения поставленной цели были обозначены следующие задачи:

1. Анализ литературных источников, содержащих результаты исследований в области прогнозирования потребления электроэнергии.
2. Сбор эмпирических данных для исследования.
3. Выявление факторов, влияющих на целевой результат – потребление электроэнергии.

²Федеральная розничная энергосбытовая компания: оптовый рынок электроэнергии и мощности [Электронный ресурс]. Режим доступа: <https://en-mart.com/optovyyj-rynok-elektroenergii-i-moshhnosti/> (дата обращения: 10.05.2022).

³Гарантирующий поставщик [Электронный ресурс]. Режим доступа: https://www.dvec.ru/electric_market/guaranteed_provider/ (дата обращения: 4.05.2022).

4. Построение моделей регрессии и кластеризации с помощью методов машинного обучения.

5. Прогнозирование электропотребления по данным предыдущих периодов.

6. Анализ и интерпретация результатов с выводами и обсуждением.

В процессе решения поставленных задач были получены результаты как теоретического, так и прикладного характера.

АНАЛИЗ ЛИТЕРАТУРНЫХ ИСТОЧНИКОВ

Решением проблемы прогнозирования электропотребления предприятия методами машинного обучения занимались многие отечественные и зарубежные ученые. С целью обобщения их опыта мы изучили некоторые работы теоретико-методологического и прикладного плана.

В статье Д. В. Антоненкова и П. В. Матренина [3] решается задача прогнозирования электропотребления горного предприятия. Авторы используют ретроспективные факторы почасового электропотребления за двое суток. В качестве методов машинного обучения используются следующие алгоритмы: случайный лес (Random Forest), адаптивный бустинг (AdaBoost), градиентный бустинг (XGBoost), многослойный перцептрон (MLP). В целом все методы продемонстрировали достаточно хорошие показатели для поставленной задачи и доказали свою эффективность.

В работе [4] Н. А. Серебрякова решается задача краткосрочного прогнозирования электропотребления гарантирующего поставщика. Для решения поставленной задачи выделены следующие факторы: циклические (дни недели, время года, праздничные дни), метеорологические (температура наружного воздуха, освещенность, скорость ветра) и случайные воздействия (плановые и аварийные отключения). В качестве методов использованы различные вариации нейросетевых алгоритмов.

В контексте решаемых задач также имеет смысл рассмотреть статью [5], в которой выполнена задача прогнозирования электропотребления офисного здания. В этом исследовании использовались следующие факторы: погодные (температура, влажность, скорость ветра, горизонтальная дальность видимости, температура точки росы) и временные данные (месяц, день, час, день недели, выходной и рабочий дни). Среди применяемых алгоритмов машинного обучения были рассмотрены: различные вариации линейной регрессии (Linear Regression, Lasso, Ridge), k ближайших соседей (KNeighborsRegressor), решающие деревья и леса (DecisionTreeRegressor, RandomForestRegressor), стохастический градиентный спуск (SGDRegressor), метод опорных векторов (LinearSVR), многослойный перцептрон (MLPRegressor), экстремальный градиентный бустинг (XGBRegressor). Лучшие результаты получились у модели, обученной методом XGBRegressor.

Аналогичные по направленности исследования есть и у зарубежных ученых. Так, в исследовании [6] построен набор моделей машинного обучения электропотребления обувного магазина. В качестве факторов были взяты: день недели, номер дня, номер недели, выходной/рабочий день, потребления за предыдущий день. В [6] были применены следующие алгоритмы машинного обучения: Linear Regression, Random Forest Regressor, Decision Tree Regressor, KNeighbors Regressor, LinearSVR. Лучший результат оказался у модели, обученной методом Random Forest Regressor.

Обобщая литературный обзор по рассматриваемой проблеме, можно отметить, что проблема, рассматриваемая в настоящей статье, актуальна, а машинное обучение является достаточно эффективным инструментом для ее решения.

ПРОГНОЗИРОВАНИЕ ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ

1. Сбор эмпирических данных для исследования. Представленное исследование проводилось по данным четырех предприятий Правобережного района республики Северная Осетия – Алания, два из которых относятся к промышленности, одно к сельскому хозяйству и еще одно к пищевой промышленности. Эти данные были предоставлены АО «Севкавказэнерго».

Исследование осуществлялось с использованием интерпретируемого высокоуровневого языка программирования Python версии 3.10.4, а также его библиотек: Numpy и Pandas – для обработки и манипуляций с данными, Matplotlib – для визуализации данных, Scikit-learn, XGBoost и CatBoost – для загрузки экземпляров моделей машинного обучения. Средой разработки был выбран Jupyter Notebook, так как этот инструмент позволяет эффективно осуществлять разработку и представление ее результатов в интерактивном виде.

Все исходные наборы данных являются достаточно большой выборкой и соответствуют некоторым характеристикам направления Big Data, таким как:

- постоянное увеличение объема, генерирование новых данных;
- генерация данных происходит в достаточно короткие промежутки времени⁴, можно сказать, в режиме «online». Этот процесс осуществляется в автоматизированных информационно-измерительных системах коммерческого учета электроэнергии (АИИС КУЭ).

Однако несмотря на то, что наборы данных можно отнести к большой выборке, прогнозируемую величину – потребление электроэнергии за некоторый период времени (получасовые периоды) – также можно представить как временной ряд и изобразить на графиках (рисунки 2–3), где по оси абсцисс – номер записи в выборке, а по оси ординат – объем потребленной электроэнергии, кВт·ч.

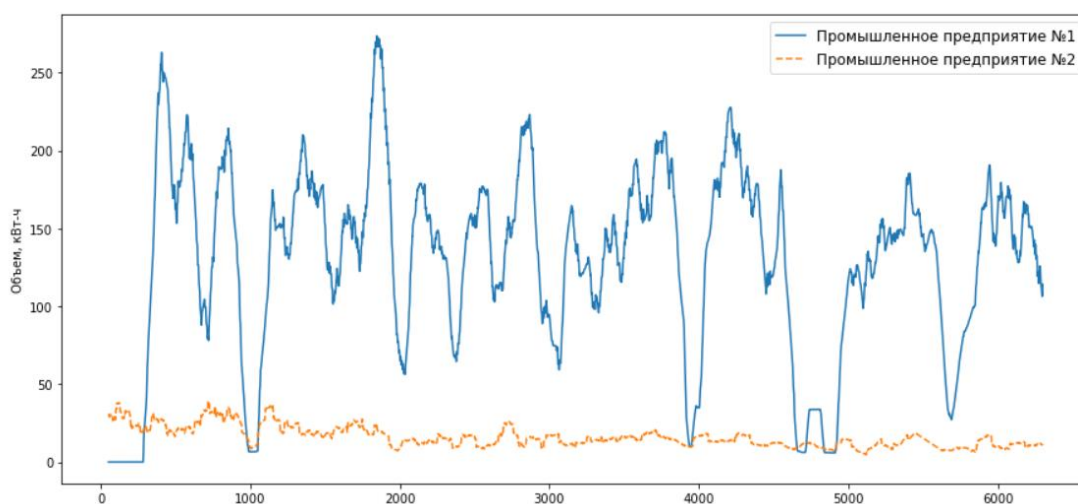


Рис. 2. Электропотребление промышленных предприятий

Fig. 2. Power consumption of industrial enterprises

⁴ IT словарь [Электронный ресурс]. Режим доступа: <https://science.involta.ru/search/glossary?q=big+data> (дата обращения: 15.05.2022).

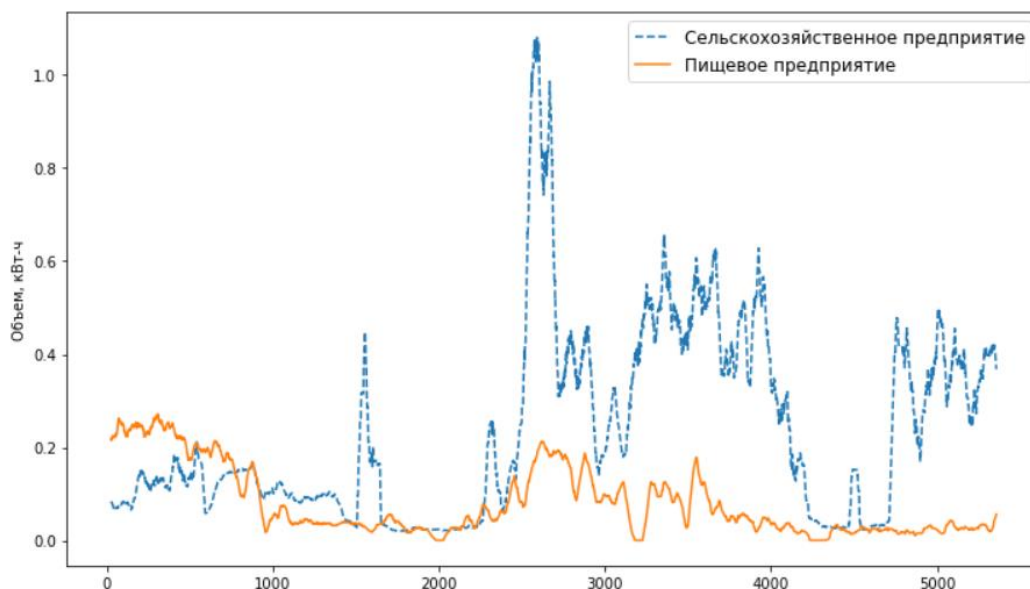


Рис. 3. Электропотребление сельскохозяйственного и пищевого предприятий

Fig. 3. Power consumption of agricultural and food enterprises

2. Выявление факторов. В качестве независимых переменных, способных описать целевую переменную, были выбраны факторы: температура окружающей среды, °С (X1), относительная влажность (%) на высоте 2 метра над поверхностью земли (X2), скорость ветра на высоте 10–12 метров над земной поверхностью, метры в секунду (X3), горизонтальная дальность видимости, километры (X4), температура точки росы, °С (X5), выходной (праздничный)/рабочий день недели, бинарный (X6), день недели (X7).

Признаки температура окружающей среды, относительная влажность, скорость ветра, горизонтальная дальность видимости, температура точки росы были взяты с электронного ресурса⁵, выходной (праздничный)/рабочий день недели – с электронного ресурса⁶.

В целях выявления линейной зависимости и проверки независимых переменных на интеркорреляцию между факторами была построена матрица корреляций, изображенная на рисунке 4 и рассчитанная по формуле парного коэффициента корреляции Пирсона (1), где $X_{1,i}$ – значения первого признака, $X_{2,i}$ – значения второго признака, \bar{X}_1 , \bar{X}_2 – средние значения обоих признаков [7].

$$r_{xy} = \frac{\sum_{i=1}^k (X_{1,i} - \bar{X}_1)(X_{2,i} - \bar{X}_2)}{\sqrt{\sum_{i=1}^k (X_{1,i} - \bar{X}_1)^2 \sum_{i=1}^k (X_{2,i} - \bar{X}_2)^2}}. \quad (1)$$

Рассмотрев матрицу корреляций (рис. 4), мы выявили достаточно значимую линейную зависимость (0,79) между признаками температуры окружающей среды и температуры точки росы. На основании этого факта было принято решение исключить фактор температуры точки росы, так как он более сложен в получении, чем температура окружающей среды.

⁵Архив погоды во Владикавказе (аэропорт) [Электронный ресурс]. Режим доступа: [https://rp5.ru/Архив_погоды_во_Владикавказе_\(аэропорт\)](https://rp5.ru/Архив_погоды_во_Владикавказе_(аэропорт)) (дата обращения: 14.04.2022).

⁶Производственный календарь [Электронный ресурс]. Режим доступа: <http://www.consultant.ru/law/ref/calendar/proizvodstvennyy/2022/> (дата обращения: 14.04.2022).

	X1	X2	X3	X4	X5	X6	X7
X1	1.000000	-0.546754	0.258664	0.397140	0.790303	-0.069329	0.063199
X2	-0.546754	1.000000	-0.339062	-0.589924	0.064959	0.012899	-0.024490
X3	0.258664	-0.339062	1.000000	0.155087	0.058396	-0.024379	-0.002359
X4	0.397140	-0.589924	0.155087	1.000000	0.055867	-0.022622	-0.152418
X5	0.790303	0.064959	0.058396	0.055867	1.000000	-0.068187	0.061213
X6	-0.069329	0.012899	-0.024379	-0.022622	-0.068187	1.000000	-0.000285
X7	0.063199	-0.024490	-0.002359	-0.152418	0.061213	-0.000285	1.000000

Рис. 4. Матрица корреляций

Fig. 4. Correlation matrix

Кроме того, в процессе предобработки числовых признаков обучающих наборов данных с целью приведения их к некоторой общей шкале (в диапазоне [0, 1]) без потери информации была проведена нормализация признаков посредством масштабирования по минимаксу по (2) [8]:

$$x_{\text{норм}}^{(i)} = \frac{x^{(i)} - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}, \tag{2}$$

где $x^{(i)}$ – индивидуальный образец, x_{min} – наименьшее значение в столбце признака, x_{max} – наибольшее значение в столбце признака.

3. Прогнозирование. Исходные наборы данных были разделены на обучающий и тестовые выборки в отношении 80/20. Создание и обучение моделей машинного обучения производилось по каждому из предприятий отдельно и затем оценивались по следующим метрикам:

1. Средняя абсолютная ошибка (MAE) – определяется по (3), где \hat{y}_i – прогнозируемое значение i -го наблюдения; y_i – истинное значение i -го наблюдения.

$$\text{MAE}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} * \sum_{i=0}^{n_{\text{samples}}-1} |y_i - \hat{y}_i|. \tag{3}$$

2. Коэффициент детерминации (R^2) – определяется на (4) [9].

$$R^2(y, \hat{y}) = \frac{\sum_{i=0}^n (y_i - \hat{y}_i)^2}{\sum_{i=0}^n (y_i - \hat{y}_i)^2 + \sum_{i=0}^n (y_i - \hat{y}_i)^2}. \tag{4}$$

Значения метрик качества каждой модели для предприятий приведены в таблице 1, где использованы следующие условные обозначения: А – промышленное предприятие № 1, В – промышленное предприятие № 2, С – пищевое предприятие, D – сельскохозяйственное предприятие.

После анализа полученных результатов был сделан вывод, что ни одна из моделей машинного обучения не является адекватной экспериментальным данным, о чем свидетельствуют значения метрик качества. Так, коэффициент детерминации (R^2) либо чуть выше нулевого значения, либо отрицательный, что подтверждает факт несоответствия аппроксимирующей зависимости экспериментальным данным.

Таблица 1

МЕТРИКИ КАЧЕСТВА МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ
 QUALITY METRICS FOR MACHINE LEARNING MODELS

Метрика качества Модель	A		B		C		D	
	MAE	r2	MAE	r2	MAE	r2	MAE	r2
Linear Regression	0.12	-0.04	0.04	-0.19	0.10	-3.16	0.12	-0.71
Lasso	0.11	-0.0002	0.04	-0.16	0.13	-4.64	0.11	-0.49
Ridge	0.12	-0.04	0.03	-0.19	0.1	-3.15	0.12	-0.72
KNeighborsRegressor	0.13	-0.27	0.039	-0.6	0.07	-1.58	0.13	-1.23
DecisionTreeRegressor	0.18	-1.34	0.05	-3.27	0.09	-3.23	0.16	-2.22
RandomForestRegressor	0.12	-0.003	0.037	-0.27	0.06	-1.27	0.12	-0.85
SGDRegressor	0.11	-0.01	0.04	-0.27	0.11	-3.25	0.12	-0.62
LinearSVR	0.11	-0.03	0.022	0.012	0.09	-2.27	0.14	-1.8
MLPRegressor	0.12	0.0012	0.047	-0.047	0.07	-1.44	0.12	-1.01
XGBRegressor	0.11	-0.04	0.041	-0.31	0.06	-1.23	0.12	-0.79
CatBostRegressor	0.12	-0.12	0.037	-0.29	0.06	-1.12	0.13	-0.98

На наш взгляд, подобный неудовлетворительный результат получен в связи с упущением некоторых важных факторов:

- загруженность предприятия – каждое из них имеет свой собственный заказ, во многом зависящий от коммерческой успешности выпускаемой продукции;
- технологический процесс производства – учет мощности приемников электроэнергии. Может изменяться в длительной перспективе, например, при перевооружении предприятия или сокращении производства.

Для проверки предположения о необходимости учета фактора «загруженность предприятия» (X8) была произведена кластеризация зависимой переменной на 5 кластеров методом k средних (KMeans). Результат этой кластеризации представляет собой искусственно созданный фактор загруженности предприятия (X8), состоящий из пяти уникальных значений. Признак X8 так же, как и все остальные, был нормализован по (2).

Фактор технологического процесса в общем случае, безусловно, необходимо учитывать, однако так как обучение производится по каждому из предприятий в отдельности, благодаря такому индивидуальному подходу этим признаком можно пренебречь ввиду невозможности получения как реальных, так и искусственных значений.

В целях выявления линейной зависимости между новым и старыми факторами построена матрица корреляций по (1), изображенная на рисунке 5. Сильной линейной зависимости между признаками не наблюдается, поэтому все факторы были использованы при моделировании.

	X1	X2	X3	X4	X6	X7	X8
X1	1.000000	-0.546754	0.258664	0.397140	-0.069329	0.063199	-0.121008
X2	-0.546754	1.000000	-0.339062	-0.589924	0.012899	-0.024490	0.019596
X3	0.258664	-0.339062	1.000000	0.155087	-0.024379	-0.002359	-0.013239
X4	0.397140	-0.589924	0.155087	1.000000	-0.022622	-0.152418	-0.017077
X6	-0.069329	0.012899	-0.024379	-0.022622	1.000000	-0.000285	-0.031117
X7	0.063199	-0.024490	-0.002359	-0.152418	-0.000285	1.000000	0.022617
X8	-0.121008	0.019596	-0.013239	-0.017077	-0.031117	0.022617	1.000000

Рис. 5. Матрица корреляций с учетом фактора «загруженность предприятия»

Fig. 5. Correlation matrix taking into account the factor «workload of the enterprise»

С учетом «нового» признака было проведено прогнозирование на тестовой выборке, результаты которого приведены в таблице 2, где предприятия условно обозначены аналогично таблице 1.

Таблица 2

МЕТРИКИ КАЧЕСТВА МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ С УЧЕТОМ ПРИЗНАКА ЗАГРУЖЕННОСТИ ПРЕДПРИЯТИЯ

QUALITY METRICS OF MACHINE LEARNING MODELS TAKING INTO ACCOUNT THE SIGN OF ENTERPRISE WORKLOAD

Метрика качества Модель	A		B		C		D	
	MAE	r2	MAE	r2	MAE	r2	MAE	r2
LinearRegression	0.11	0,0088	0,02	0.38	0.1	0.52	0.1	0.07
Lasso	0.11	-0.002	0.05	-3.8	0.16	-0.01	0.11	-4.41
Ridge	0.117	0.0117	0.02	0.38	0.1	0.52	0.1	0.07
KNeighborsRegressor	0.035	0.8546	0.02	0.69	0.02	0.69	0.028	0.82
DecisionTreeRegressor	0.036	0.8978	0.01	0.92	0.01	0.92	0.025	0.92
RandomForestRegressor	0.026	0.95	0.01	0.95	0.01	0.95	0.019	0.96
SGDRegressor	0.11	0.06	0.02	0.36	0.02	0.35	0.1	0.064
LinearSVR	0.10	0.09	0.02	0.35	0.02	0.35	0.078	-0.14
MLPRegressor	0.028	0.945	0.01	0.94	0.01	0.94	0.02	0.9525
XGBRegressor	0.027	0.949	0.02	0.94	0.02	0.94	0.02	0.9561
CatBostRegressor	0.027	0.948	0.01	0.95	0.01	0.95	0.02	0.9589

АНАЛИЗ И ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ

Подводя итоги моделирования с признаком загруженности предприятия, можно констатировать, что менее подходящими алгоритмами задачи данного исследования стали алгоритмы линейной регрессии (LinearRegression, Lasso, Ridge) и градиентного спуска (SGDRegressor). Наиболее релевантные результаты получены с помощью алгоритмов RandomForestRegressor (случайный лес), MLPRegressor, XGBRegressor и CatBostRegressor. Все примененные методы дали примерно одинаковый результат, однако лучшие показатели метрик качества оказались у модели, основанной на алгоритме CatBostRegressor.

Подводя итоги всего исследования, можно сделать следующие выводы:

1. Выявлены факторы, с учетом которых в моделях машинного обучения получены достаточно точные прогнозные данные для всех предприятий: температура, влажность, скорость ветра, горизонтальная дальность видимости, температура точки росы, выходной (праздничный)/рабочий день недели, день недели, загруженность предприятия.

2. Построены различные регрессионные модели машинного обучения. Лучшими из них для данной задачи являются RandomForestRegressor (случайный лес), MLPRegressor (нейронные сети), XGBRegressor и CatBoostRegressor.

3. Данное исследование наглядно демонстрирует необходимость учета факторов технологического процесса и загруженности предприятия при прогнозировании его электропотребления.

Таким образом, были достигнуты цель и задачи исследования: проведено прогнозирование потребления электроэнергии по данным расхода электроэнергии предприятиями, относящимися к различным секторам экономики, с учетом интересов электросетевой компании. Наиболее точные результаты прогнозирования были получены с помощью модели машинного обучения, основанной на алгоритме CatBoostRegressor. Важно заметить, что для повышения достоверности прогноза необходим персональный (по каждому предприятию) учет факторов, характеризующих технологический процесс предприятий. Как показало проведенное исследование, добавление в модель регрессии искусственного фактора загруженности предприятия значительно улучшило обобщающую способность алгоритма, позволив тем самым получить точный прогноз. Однако необходимо проведение дополнительных исследований характеристик, влияющих на величину потребления электрической энергии на предмет возможности дальнейшего использования в процессе прогнозирования. К таким характеристикам можно отнести данные о технологическом процессе, а также некоторые метеорологические данные, как, например, освещенность, долгота дня.

СПИСОК ЛИТЕРАТУРЫ

1. Волкова В. Н., Козлов В. Н. Системный анализ и принятие решений: Словарь-справочник: учебное пособие для вузов. Москва: Высшая школа, 2004. 616 с.

2. Моргоев И. Д., Дзгоев А. Э., Клюев Р. В. и др. Современные способы борьбы с коммерческими потерями в электроэнергетике // Энергетика будущего – цифровая трансформация: сборник трудов II Всероссийской научно-практической конференции. ЛГТУ: Липецк, 2021. С. 181–185.

3. Антоненков Д. В., Матренин П. В. Исследование ансамблевых и нейросетевых методов машинного обучения в задаче краткосрочного прогнозирования электропотребления горных предприятий // Электротехнические системы и комплексы. 2021. № 3(52). С. 57–65. DOI: 10.18503/2311-8318-2021-3(52)-57-65.

4. Серебряков Н. А. Анализ факторов, влияющих на совокупное электропотребление гарантирующего поставщика // Вестник Иркутского государственного технического университета. 2020. № 2(151). С. 366–381. DOI: 10.21285/1814-3520-2020-2-366-381.

5. Моргоева А. Д., Моргоев И. Д., Клюев Р. В., Ляшенко В. И. Прогнозирование нагрузки на электросеть как способ эффективного управления потреблением электрической энергии // Вести высших учебных заведений Черноземья. 2021. № 4(66). С. 39–51. DOI: 10.53015/18159958_2021_4_39.

6. Alfonso González-Briones, Sigeru Omatu, Mohd Saberi Mohamad. Machine Learning Models for Electricity Consumption Forecasting: A Review // 2nd International Conference on Computer Applications & Information Security (ICCAIS), IEEE. Riyadh, Saudi Arabia, 2019. 18851034. DOI: 10.1109/CAIS.2019.8769508.

7. Доугерти К. Введение в эконометрику: пер. с англ. Москва: ИНФРА-М, 1999. 402 с.

8. Раиша С., Мирджалили В. Python и машинное обучение: машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow 2. 3-е изд.: пер. с англ. СПб.: Диалектика, 2020. 848 с.

9. Элбон К. Машинное обучение с использованием Python. Сборник рецептов: пер. с англ. СПб.: БХВ-Петербург, 2019. 384 с.

Информация об авторах

Моргоев Ирбек Джабраилович, аспирант, Северо-Кавказский горно-металлургический институт (государственный технологический университет);

362011, Россия, Владикавказ, ул. Николаева, 44;

m.irbek@yandex.ru, ORCID: <https://orcid.org/0000-0003-4390-5662>

Дзгоев Алан Эдуардович, канд. техн. наук, доцент кафедры «Информационные технологии и системы», Северо-Кавказский горно-металлургический институт (государственный технологический университет);

362021, Россия, Владикавказ, ул. Николаева, 44;

dzgoev_alan@mail.ru, ORCID: <https://orcid.org/0000-0002-1314-6151>

Клюев Роман Владимирович, д-р техн. наук, профессор кафедры «Техники низких температур им. П. Л. Капицы», Московский политехнический университет;

107023, Россия, Москва, ул. Б. Семеновская, 38;

kluev-roman@rambler.ru, <https://orcid.org/0000-0003-3777-7203>

Моргоева Анжелика Джабраиловна, аспирант, Северо-Кавказский горно-металлургический институт (государственный технологический университет);

362011, Россия, Владикавказ, ул. Николаева, 44;

m.angelika-m@yandex.ru, ORCID: <https://orcid.org/0000-0003-2949-1993>

REFERENCES

1. Volkova V.N., Kozlov V.N. *Sistemnyj analiz i prinjatie reshenij: Slovar'-spravochnik* [System analysis and decision-making]. Moskva: Vvsshaja shkola. 2004. 616 p. (In Russian).

2. Morgoev I.D., Dzgoev A.E., Klyuev R.V. [et al.] Modern ways to combat commercial losses in the electric power industry. *Energetika budushchego – cifrovaya transformaciya: sbornik trudov II vsrossijskoj nauchno-prakticheskoi konferencii* [Energy of the future - digital transformation. Proceedings of the 2nd Scientific-practical conference]. Lipeck: LGTU. 2021. Pp. 181–185 (In Russian).

3. Antonenkov D.V., Matrenin P.V. Ensemble and neural network machine learning models for short-term load forecasting of open cast mining companies. *Electrotechnical systems and complexes*. 2021. No. 3(52). Pp. 57–65. DOI: 10.18503/2311-8318-2021-3(52)-57-65. (In Russian)

4. Serebryakov N.A. Analysis of factors affecting the electricity consumption of a delivery point cluster default provider. *Proceedings of Irkutsk State Technical University*. 2020. No. 2(151). Pp. 366–381. DOI: 10.21285/1814-3520-2020-2-366-381. (In Russian)

5. Morgoeva A.D., Morgoev I.D., Klyuev R.V., Lyashenko V.I. Forecasting the load on the power grid as a way to effectively manage the consumption of electrical energy. *News of Higher Educational Institutions of the Chernozem Region*. 2021. No. 4(66). Pp. 39–51. DOI: 10.53015/18159958_2021_4_39. (In Russian)

6. Alfonso González-Briones, Sigeru Omatu, Mohd Saberi Mohamad. Machine Learning Models for Electricity Consumption Forecasting: A Review. *2nd International Conference on Computer Applications & Information Security (ICCAIS), IEEE*. Riyadh, Saudi Arabia, 2019. 18851034. DOI: 10.1109/CAIS.2019.8769508.

7. Dougherty C. Introduction to econometrics. New York. Oxford University Press, 1999. 402 p.
8. Rashka S., Mirdzhalili V. Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2. Birmingham, Mumbai. Packt, 2020. 848 p.
9. Albon C. Machine Learning with Python Cookbook Practical Solutions from Preprocessing to Deep Learning. Beijing, Boston, Farnham, Sebastopol, Tokyo. O'Reilly, 2019. 384 p.

Information about the authors

Morgoev Irbek Dzhabrailovich, postgraduate, The North Caucasian Institute of Mining and Metallurgy (State Technological University);

362011, Russia, Vladikavkaz, 44 Nikolaev street;

m.irbek@yandex.ru, ORCID: <https://orcid.org/0000-0003-4390-5662>

Dzgoev Alan Eduardovich, Candidate of Technical Sciences, Associate Professor of the Department of Information Technologies and Systems, The North Caucasian Institute of Mining and Metallurgy (State Technological University);

362011, Russia, Vladikavkaz, 44 Nikolaev street;

dzgoev_alan@mail.ru, ORCID: <https://orcid.org/0000-0002-1314-6151>

Klyuev Roman Vladimirovich, Doctor of Technical Sciences, Professor of Low temperature engineering department named after P.L. Kapitsa, Moscow Polytechnic University;

107023, Russia, Moscow, 38 B. Semenovskaya street;

kluev-roman@rambler.ru, <https://orcid.org/0000-0003-3777-7203>

Morgoeva Anzhelika Dzhabrailovna, postgraduate, The North Caucasian Institute of Mining and Metallurgy (State Technological University);

362011, Russia, Vladikavkaz, 44 Nikolaev street;

m.angelika-m@yandex.ru, ORCID: <https://orcid.org/0000-0003-2949-1993>