

DOI: 10.12731/2227-930X-2024-14-1-290  
УДК 004.85



Научные обзоры

## ПРОГНОЗИРОВАНИЕ ПОТРЕБИТЕЛЬСКОЙ АКТИВНОСТИ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

*В.Д. Новиков, Р.М. Хамитов*

*В данной статье рассматривается прогнозирование потребительской активности, а в частности прогнозирование энергопотребления домохозяйств с использованием машинного обучения. Прогнозирование энергопотребления домохозяйств с использованием машинного обучения – это тема, которая затрагивает различные аспекты эффективного и экологичного использования электроэнергии. В статье рассматриваются различные методы и модели машинного обучения, которые могут быть применены для решения задачи прогнозирования.*

*В отдельную категорию выделено рассмотрение модели нейронных сетей такой как LSTM, дается ее описание, процесс обучения и использования, а также даны преимущества и недостатки данной модели. После чего на подготовленном датасете производится обучение модели для прогнозирования энергопотребления.*

**Ключевые слова:** *потребление энергии; обработка данных; машинное обучение; прогнозирование; нейронные сети; глубокое обучение*

**Для цитирования.** *Новиков В.Д., Хамитов Р.М. Прогнозирование потребительской активности с использованием методов машинного обучения // International Journal of Advanced Studies. 2024. Т. 14, № 1. С. 205-214. DOI: 10.12731/2227-930X-2024-14-1-290*

Scientific Reviews

## FORECASTING CONSUMER ACTIVITY USING MACHINE LEARNING METHODS

*V.D. Novikov, R.M. Khamitov*

*This article discusses forecasting consumer activity, and in particular forecasting household energy consumption using machine learning. Forecasting household energy consumption using machine learning is a topic that addresses various aspects of efficient and environmentally friendly use of electricity.*

*The article discusses various machine learning methods and models that can be applied to solve the forecasting problem. The consideration of a neural network model such as LSTM is highlighted in a separate category, its description, the learning and use process are given, as well as the advantages and disadvantages of this model are given. After that, a model is trained on the prepared dataset to predict energy consumption.*

**Keywords:** *energy consumption; data processing; machine learning; forecasting; neural networks; deep learning*

**For citation.** *Novikov V.D., Khamitov R.M. Forecasting Consumer Activity using Machine Learning Methods. International Journal of Advanced Studies, 2024, vol. 14, no. 1, pp. 205-214. DOI: 10.12731/2227-930X-2024-14-1-290*

### **Введение**

Энергопотребление домохозяйств является важным фактором, влияющим на баланс спроса и предложения на рынке электроэнергии. Точный и своевременный прогноз энергопотребления позволяет энергетическим компаниям оптимизировать свою работу, снизить затраты и уменьшить выбросы парниковых газов. Кроме того, прогнозирование энергопотребления может помочь

потребителям контролировать свой энергетический бюджет и повысить энергоэффективность.

Однако прогнозирование энергопотребления домохозяйств представляет собой сложную задачу, так как оно зависит от множества факторов, таких как погодные условия, сезонность, день недели, время суток, поведение и предпочтения потребителей, тип и состояние оборудования и т.д. [1]. Традиционные статистические методы, такие как авторегрессия, скользящее среднее или экспоненциальное сглаживание, могут не учитывать все эти факторы или требовать большого количества данных и времени для настройки и обновления моделей.

В последнее время все больше исследователей и практиков применяют методы машинного обучения для прогнозирования энергопотребления домохозяйств. Методы машинного обучения способны анализировать большие и сложные наборы данных, выявлять скрытые закономерности и зависимости, адаптироваться к изменяющимся условиям и обеспечивать высокую точность и надежность прогнозов [2].

### **Материалы и методы исследования**

В работе используются результаты зарубежных и отечественных научных исследований. Авторами применяются теоретические методы исследования, связанные с поиском и анализом информации для выявления связей и получения уникальных выводов.

### **Результаты и обсуждение**

Прежде всего стоит определить какие методы машинного обучения существуют, после чего можно перейти к деталям реализации.

Среди наиболее популярных и эффективных методов машинного обучения для прогнозирования энергопотребления домохозяйств можно выделить следующие:

1. Регрессионные модели. Регрессионные модели строят функциональную зависимость между энергопотреблением и одним

или несколькими объясняющими переменными, такими как температура, влажность, освещенность, давление и т.д. Регрессионные модели могут быть линейными или нелинейными, одно- или многомерными, статическими или динамическими [3]. Регрессионные модели обладают простотой, интерпретируемостью и низкой вычислительной сложностью, но могут быть чувствительны к выбросам, мультиколлинеарности и нелинейности данных.

2. Модели искусственных нейронных сетей. Модели искусственных нейронных сетей (ИНС) состоят из множества связанных узлов, называемых нейронами, которые имитируют работу биологического мозга [4]. ИНС способны аппроксимировать любую сложную нелинейную функцию, обучаясь на данных с помощью алгоритмов оптимизации, таких как градиентный спуск, обратное распространение ошибки или генетические алгоритмы. ИНС обладают высокой точностью, гибкостью и адаптивностью, но могут быть сложными, непрозрачными и требовательными к ресурсам.

3. Модели машинного обучения с учителем. Модели машинного обучения с учителем (МОСУ) используют размеченные данные, то есть данные, в которых известны как входные, так и выходные значения, для обучения алгоритма, который может предсказывать выходные значения для новых входных данных. МОСУ могут быть классификационными или регрессионными, в зависимости от типа выходной переменной. МОСУ обладают хорошей обобщающей способностью, масштабируемостью и разнообразием, но могут быть подвержены переобучению, недообучению и шуму данных [5].

4. Модели машинного обучения без учителя. Модели машинного обучения без учителя (МОБУ) используют неразмеченные данные, то есть данные, в которых известны только входные значения, для обнаружения скрытых структур, закономерностей или групп в данных. МОБУ могут быть кластеризационными, понижающими размерность, ассоциативными или аномальными, в зависимости от цели анализа [6]. МОБУ обладают способностью к самоорганизации, обнаружению новых знаний и обработке

больших объемов данных, но могут быть неустойчивыми, неоднозначными и сложными для интерпретации.

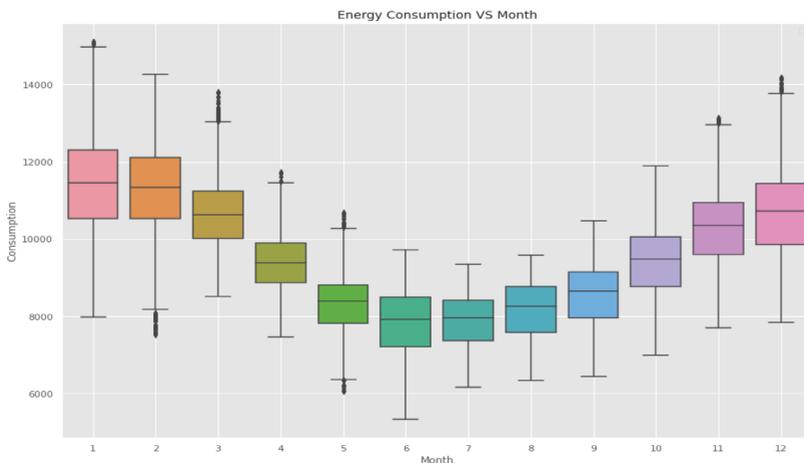
### *Прогнозирование потребления электрической энергии*

Для исследования прогнозирования энергопотребления домохозяйств был выбран датасет, в котором определялась величина потребления электрической энергии домохозяйствами за периоды с 2015 года по 2021 года. Учет потребления ввелся ежечасно, что позволяет более точно спрогнозировать конечный уровень энергопотребления.

Исследование было выполнено на языке программирования Python, использовалась версия языка программирования Python, такая как Python 3.12.0. В качестве среды программирования использовалась Jupyter Notebook.

Помимо прочего были использованы библиотеки Numpy и Pandas для вычислений и проведения манипуляций с данными. Matplotlib и Seaborn для визуализации данных. Scikit-learn для предобработки данных. Keras для обучения модели.

В начале исследования был проведен анализ исходного датасета. Данные потребления по месяцам представлены на рис. 1.



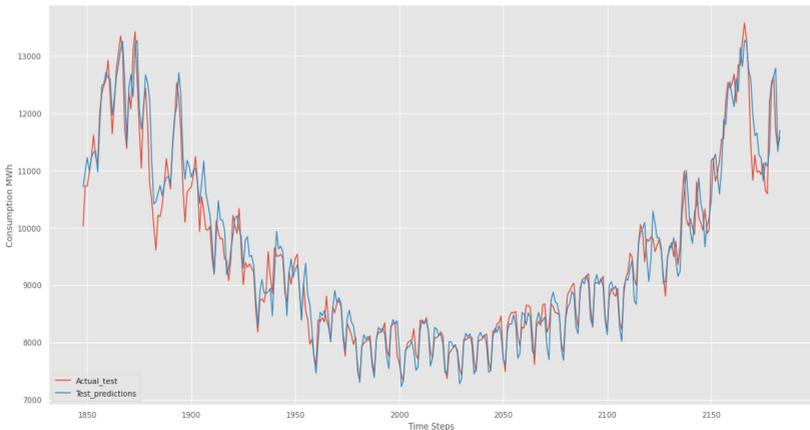
**Рис. 1.** Помесячное потребление электроэнергии

По месячному графику потребления электроэнергии, можно сделать вывод, что меньше всего потребление электроэнергии приходится на летние месяцы, это связано с тем, что люди меньше используют нагревательные приборы, которые они используют в зимние месяцы, как видно на графике, энергопотребление в зимний период существенно выше.

Следующим этапом была предобработка данных для машинного обучения.

Для прогнозирования потребления электроэнергии по дням, данные были обработаны встроенной функцией в Pandas, которая имеет название функция повторное выборки, а сам процесс называется передискретизация. Данная процедура позволила изменить частоту данных с почасовой на ежедневную.

Для модели LSTM необходимо привести исходные данные в диапазоне от 0 до 1, чтобы соответствовать масштабу слоя LSTM. Для этой цели была использована библиотека MinMaxScaler, которая позволяет нормализовать исходные данные. Таким образом используя данную библиотеку, данные были приведены к диапазону от 0 до 1.



**Рис. 2.** Прогнозирование энергопотребления с использованием тестовой выборки

Далее датасет был разделен на обучающую и тестовую выборку, было использовано соотношение 80:20. Обучение модели LSTM производилось с использованием обучающего набора и тестового набора для тестирования результатов в процессе обучения.

В результате при прогнозировании потребления электроэнергии были получены следующие значения, которые были представлены в виде графика на рис. 2.

По графику видно, что прогнозируемое значение энергопотребления почти не отличается от реального значение, средняя точность по результатам исследования составила 93,8 %, что позволяет говорить о точности обучения модели прогнозирования.

### **Заключение**

В результате проведенного исследования была разработана модель машинного обучения, которая основана на рекуррентной нейронной сети LSTM. Точность прогнозирования энергопотребления в процессе исследования составила 93,8%, что позволяет классифицировать полученные прогнозы энергопотребления как достоверные и считать полученную модель адекватной экспериментальным данным.

Помимо прочего, были выделены следующие выводы.

Во-первых, задача прогнозирования энергопотребления является актуальной для всех сфер общества, как для научного, с целью исследования применяемых методов прогнозирования, так и для компаний, позволяющая компаниям значительно сократить издержки по затратам на электроэнергию и внедрению различных видов энергосберегающих систем.

Во-вторых, прогнозирование энергопотребление довольно специфично и должно рассматриваться отдельно, учитывая различные факторы и анализа в зависимости от применяемое области, таким образом прогнозирование энергопотребления для предприятий будет отличаться от прогнозирования энергопотребления домохозяйств.

### ***Список литературы***

1. Васильев Г.В., Бердоносков В.Д. Методика по эффективному применению гибридных моделей нейронных сетей для прогнозирования энергопотребления // *Электротехнические системы и комплексы*. 2022. №4 (57). С. 88-95.
2. Моргоева А.Д., Моргоев И.Д. Прогнозирование потребления электрической энергии промышленным предприятием с помощью методов машинного обучения // *Известия ТПУ*. 2022. №7. С. 115-125.
3. Горбунова Е.Б. Нейросетевой подход к прогнозированию потребления энергоресурсов в городской среде // *Инженерный вестник Дона*. 2018. №4 (51). <http://ivdon.ru/ru/magazine/archive/n4y2018/5303>
4. Полуянович Н.К., Дубяго М.Н. Оценка воздействующих факторов и прогнозирование электропотребления в региональной энергосистеме с учетом режима ее эксплуатации // *Известия ЮФУ. Технические науки*. 2022. №2 (226). С. 31-46.
5. Ляндау Ю.В., Темирбулатов А.У. Обзор применения технологий искусственного интеллекта в электроэнергетической отрасли // *Инновации и инвестиции*. 2023. №8. С. 304-309.
6. Nurfaizi A., Hasanuddin M. Ticket Prediction using LSTM on a GLPI System // *International Journal of Open Information Technologies*. 2023. №7. <http://injoit.org/index.php/j1/article/view/1567>

### ***References***

1. Vasiliev G.V., Berdonosov V.D. Methodology for the effective application of the hybrid neural network models for the energy consumption forecasting (in Russian). *Elektrotekhnicheskie sistemy i komplekсы* [Electrotechnical systems and complexes], 2022, no. 4 (57), pp. 88-95.
2. Morgoeva A.D., Morgoev I.D. Forecasting of the electric energy consumption by the industrial enterprise by means of the machine learning methods. *Izvestiya TPU*, 2022, no. 7, pp. 115-125.

3. Gorbunova E.B. Neural network approach to forecasting energy resources consumption in urban environment. *Inzhenerny vestnik Dona*, 2018, no. 4 (51). <http://ivdon.ru/magazine/archive/n4y2018/5303>
4. Poluyanovich, N.K.; Dubyago, M.N. Estimation of the influencing factors and forecasting of the power consumption in the regional power system taking into account the mode of its operation. *Izvestiya YuFU. Tekhnicheskie nauki*, 2022, no. 2 (226), pp. 31-46.
5. Lyandau Yu.V., Temirbulatov A.U. Review of the application of artificial intelligence technologies in the electric power industry. *Innovatsii i investitsii* [Innovations and Investments], 2023, no. 8, pp. 304-309.
6. Nurfaizi A., Hasanuddin M. Ticket Prediction using LSTM on a GLPI System. *International Journal of Open Information Technologies*, 2023, no. 7. <http://injoit.org/index.php/j1/article/view/1567>

### **ДАННЫЕ ОБ АВТОРАХ**

**Новиков Вадим Денисович**, студент кафедры «Информационные технологии и интеллектуальные системы»  
ФГБОУ ВО «Казанский государственный энергетический университет»  
ул. Красносельская, 51, г. Казань, 420066, Российская Федерация  
[novikovschool@gmail.com](mailto:novikovschool@gmail.com)

**Хамитов Ренат Минзашарифович**, доцент кафедры «Информационные технологии и интеллектуальные системы», кандидат технических наук, доцент  
ФГБОУ ВО «Казанский государственный энергетический университет»  
ул. Красносельская, 51, г. Казань, 420066, Российская Федерация  
[hamitov@gmail.com](mailto:hamitov@gmail.com)

### **DATA ABOUT THE AUTHORS**

**Vadim D. Novikov**, student of the Department of Information Technologies and Intelligent Systems

*Kazan State Power Engineering University*

*51, Krasnoselskaya Str., Kazan, 420066, Russian Federation*

*novikovschool@gmail.com*

*ORCID: <https://orcid.org/0009-0006-8034-8956>*

**Renat M. Khamitov**, Associate Professor of the Department of Information Technologies and Intelligent Systems, Candidate of Technical Sciences

*Kazan State Power Engineering University*

*51, Krasnoselskaya Str., Kazan, 420066, Russian Federation*

*hamitov@gmail.com*

*SPIN-code: 7401-9166*

*ORCID: <https://orcid.org/0000-0002-9949-4404>*

*Scopus Author ID: 57222149321*

Поступила 28.02.2024

После рецензирования 10.03.2024

Принята 15.03.2024

Received 28.02.2024

Revised 10.03.2024

Accepted 15.03.2024