

ГЕНТОШ ЛЕСЯ

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-4957-1512>e-mail: [lesia.i.mochurad@lpnu.ua](mailto:lesia.i.mochurad@lpnu.ua)

ЛЕВКОВИЧ РОМАН

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0001-9393-714X>e-mail: [roman.y.levkovych@lpnu.ua](mailto:roman.y.levkovych@lpnu.ua)

## ЗАСТОСУВАННЯ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ

Прогнозування енергоспоживання є критично важливим для ефективного управління енергетичними ресурсами та зниження витрат. Незважаючи на те, що рекурентні нейронні мережі (RNN) традиційно використовуються для обробки часових рядів завдяки їх здатності моделювати часові залежності, згорткові нейронні мережі (CNN) можуть запропонувати значні переваги у швидкодії за рахунок паралельної обробки даних та здатності виявляти локальні шаблони у послідовностях. У даній статті ми проводимо аналіз продуктивності цих двох типів нейронних мереж, оцінюючи їх ефективність та точність у контексті прогнозування споживання електроенергії на прикладі CNN, часових згорткових мереж (TCN) та довготривалої короткочасної пам'яті (LSTM). У результаті встановлено, що згорткові нейронні мережі працюють на 30% точніше, та на 25% швидше.

Ключові слова: згорткові нейронні мережі, рекурентні нейронні мережі, споживання електроенергії, машинне навчання.

HENTOSH LESIA

Lviv Polytechnic National University

LEVKOVYCH ROMAN

Lviv Polytechnic National University

### USING CONVOLUTIONAL NEURAL NETWORKS TO PREDICT ELECTRICITY CONSUMPTION

Predicting energy consumption is critical for efficient energy management and cost reduction. Although recurrent neural networks (RNNs) have traditionally been used for time series processing due to their ability to model time dependencies, convolutional neural networks (CNNs) can offer significant performance advantages due to parallel data processing and the ability to detect local patterns in sequences. In this paper, we analyse the performance of these two types of neural networks, evaluating their efficiency and accuracy in the context of electricity consumption forecasting using CNNs, temporal convolutional networks (TCNs), and long short-term memory (LSTMs). The results of the study show that convolutional neural networks are 30% more accurate and 25% faster than other models. CNNs have an architecture that allows efficient processing of input data using parallel computing. This significantly increases the processing speed and makes it possible to identify local features in the data. This approach is especially useful when analysing large data sets where speed and accuracy are critical. TCNs combine the advantages of convolutional networks and traditional time series processing methods, which allows them to effectively model time dependencies. TCNs can achieve high forecasting accuracy due to their ability to account for both short-term and long-term dependencies in the data. LSTM-based models are also widely used for time series forecasting due to their ability to store information about previous states over long periods of time. However, compared to CNNs and TCNs, LSTMs can be less efficient in terms of processing speed, especially when dealing with large amounts of data. Our analysis shows that using CNNs for energy consumption forecasting provides a significant advantage in terms of speed and accuracy. This allows power system operators to better predict electricity consumption, reducing the risk of overloads and ensuring the stability of power grids. This approach also helps to optimise costs and improve environmental performance, which is an important aspect of modern energy management.

Keywords: convolutional neural network, recurrent neural network, electricity consumption, machine learning.

### Постановка задачі

Прогнозування споживання електроенергії є критично важливим завданням для забезпечення стабільності та ефективності енергетичних систем. Точні прогнози дозволяють операторам електричних мереж краще планувати виробництво та розподіл електроенергії, зменшуючи витрати та забезпечуючи надійність постачання. У світі зростаючого попиту на енергетичні ресурси та розвитку відновлюваних джерел енергії, таких як сонячні та вітрові електростанції, точне прогнозування споживання стає ще більш важливим. Це пов'язано з тим, що відновлювані джерела енергії зазвичай характеризуються високою варіативністю, що ускладнює управління енергетичними системами.

Існуючі методи прогнозування, зокрема традиційні статистичні підходи та деякі сучасні алгоритми машинного навчання, мають свої обмеження. Зокрема, LSTM мережі, наприклад, показали високу ефективність у багатьох випадках прогнозування часових рядів завдяки їх здатності зберігати довготривалі залежності в даних. Однак вони можуть бути обчислювально складними та потребувати значних ресурсів для тренування. Дослідження ефективних та обчислювально доступних методів прогнозування споживання електроенергії є надзвичайно актуальним, оскільки воно може відкрити нові можливості для підвищення ефективності та надійності енергетичних систем.

Метою даної роботи є дослідження ефективності та точності згорткових нейронних мереж (TCN та CNN) у порівнянні з LSTM мережами для прогнозування споживання електроенергії. Це дослідження спрямоване на виявлення найкращих підходів для одноваріативного та мультिवаріативного прогнозування, враховуючи точність прогнозів та швидкодію моделей.

Науковою новизною даної роботи є оцінка ефективності TCN, яка включає аналіз точності, швидкодії та розміру моделей. Це дозволяє оцінити не лише точність прогнозів, але й практичну застосовність моделей у реальних умовах, де обчислювальні ресурси можуть бути обмеженими.

### Аналіз останніх джерел

Нейромережі є сучасним підходом до розв'язування різних задач регресії, і прогнозування часових рядів не є винятком. Найпопулярнішим підходом серед нейромереж, який використовується у сучасних наукових статтях, є використання багатошарових перцептронів та LSTM. Проте, як вказано у [1], застосування багатошарових перцептронів є складним з кількох причин. По-перше, успішне впровадження цих моделей потребує ретельного підбору архітектури, що включає визначення оптимальної кількості шарів та нейронів у кожному з них, а також налаштування гіперпараметрів. По-друге, багатошарові перцептрони можуть потребувати значного часу для сходження під час тренування, особливо на великих наборах даних, що збільшує обчислювальні витрати та ускладнює процес оптимізації.

Прикладом таких досліджень є [2, 3], де автори досліджують прогнозування споживання електроенергії для Таїланду та Абу-Дабі, відповідно. Основними методами є саме статистичні підходи, та використання багатошарового перцептрону. І у обох роботах не вдалось знайти оптимальний розмір мережі, тому статистичні методи є точнішими на 10% і 15%, відповідно.

Саме тому набувають популярності інші типи архітектур нейромереж, а саме, зазвичай, обирають саме рекурентні мережі, оскільки вони розроблені для роботи з послідовними даними. Яскравим прикладом є стаття [4], де автор розглядає проблему прогнозування споживання електроенергії у розумному будинку саме за допомогою LSTM мережі.

Також, дана архітектура може бути модифікована, як-от у [5], де автори досліджують застосування Bidirectional LSTM для прогнозування даних. Особливістю даної модифікації є використання додаткових шарів LSTM, які обробляють інформацію у зворотному порядку. Це дозволяє використовувати додатку інформацію про контекст.

Варто зауважити, що окрім LSTM існують інші архітектури рекурентних нейромереж, зокрема:

- Gated Recurrent Unit (GRU): дана модель, як і LSTM, базується на двох воротах — ворота оновлення та ворота скидання, та використовує лише один внутрішній стан.
- Класична багатошарова рекурентна мережа: модель, яка зберігає інформацію про попередні елементи послідовності і використовує їх у наступних кроках.

Відповідно у роботах [6, 7] проводиться порівняльний аналіз застосування LSTM та GRU моделей для прогнозування споживання електроенергії у короткостроковій перспективі, в результаті досліджень встановлено, що обидві моделі показують однаковий результат з різницею в 1-2% точності у залежності від обраної метрики.

У статті [8] також проводиться порівняння LSTM та GRU моделей, проте тут досліджується прогнозування на короткий період (день і 3 дні) і на середній період (тиждень і 2 тижні), і в результаті підтверджено, що LSTM та GRU моделі показують приблизно однаковий результат для короткотермінового прогнозу, але у середньостроковій перспективі LSTM є точнішою моделлю.

У статті [9] розроблено модель згорткової нейронної мережі TCN, яка показала точніший результат для групи усталених вибірок для перевірки роботи рекурентних мереж: Sequential MNIST, Word-level LAMBADA, та інші, загалом на 9 вибірках з 11.

Відповідно, у роботі [10] розглядається застосування TCN для одноваріативного прогнозування споживання електроенергії у порівнянні з LSTM моделлю і встановлено, що TCN є на 10% точнішою та 30% швидшою.

Для прогнозування споживання електроенергії також використовують ансамблеві методи. Так, у роботі [11] розроблено ансамбль, який складається з групи нейромереж: LSTM, CNN та MLP у порівнянні з окремими моделями для прогнозування споживання електроенергії. А у [12] розроблено ансамбль LSTM моделей, де кожна модель відповідає певній погодній ситуації.

### Виклад основного матеріалу

Згорткові нейронні мережі є класом моделей глибинного навчання, які особливо підходять для аналізу візуальних зображень. Вони складаються з набору згорткових шарів та нелінійних активацій, які вивчають ієрархії просторових ознак з вхідних зображень. Кожен згортковий шар застосовує фільтри до вхідних даних і використовує операції згортки для ідентифікації та відображення значущих ознак, таких як краї, текстури та форми. У контексті прогнозування споживання енергії, CNN можуть бути використані шляхом застосування одномірних згортки та репрезентації кожної змінної як вхідного каналу.

TCN [9] є типом згорткових нейронних мереж, спеціалізованим на аналізі часових рядів. Його архітектура на використанні блоків, які складаються з повнзв'язних 1D згортки, а також згортки з розширенням, які дозволяють захоплювати зв'язки між більшими періодами у часі. Візуалізація цього блоку зображена на Рис. 1. Особливістю TCN є те, що розмір входу та виходу блоку однаковий, що робить його відмінним від звичайних згорткових мереж.

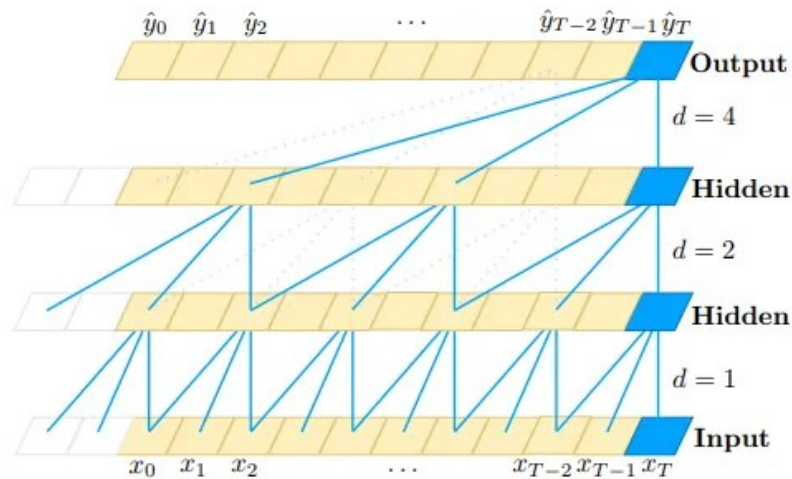


Рис. 1. Візуалізація одного блоку TCN мережі [9]

Мережі LSTM є видом рекурентних мереж, які розроблені, щоб впоратись з проблемою захоплення довгих залежностей у даних [13]. Основними складовими LSTM, зображення якої показано на Рис. , є "ворота", які регулюють потік інформації в мережі:

1. Ворота забування (Forget Gate): Ця частина вирішує, яку інформацію з попереднього кроку треба забути або проігнорувати. Вона визначає, яка частина попереднього стану LSTM не має враховуватися у поточному контексті.
2. Ворота входу (Input Gate): Цей механізм визначає, яку інформацію ми повинні оновити в пам'яті. Він враховує нові вхідні дані та вирішує, яка частина цих даних є важливою для збереження.
3. Ворота виходу (Output Gate): Після оновлення пам'яті ця частина LSTM вирішує, яка частина пам'яті буде використовуватися для генерації виходу.

Ці механізми дозволяють LSTM моделі ефективно працювати з довгостроковими залежностями в часових рядах, а також уникати проблем градієнтного зникнення, що часто виникають у звичайних рекурентних нейронних мережах.

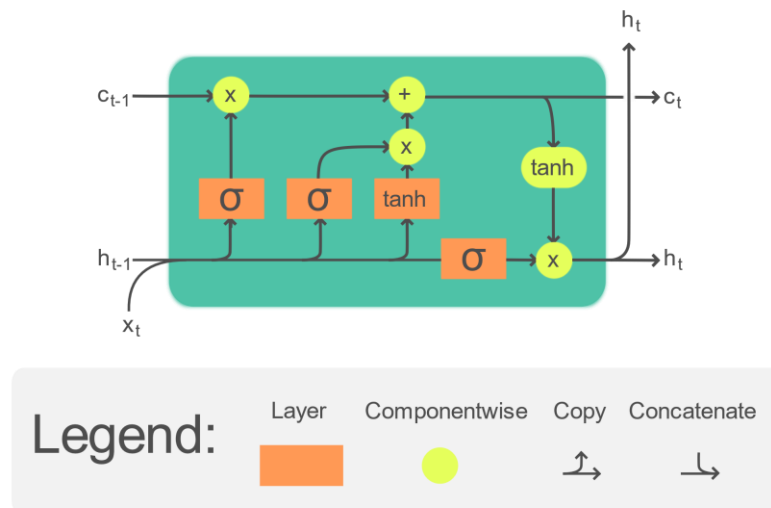


Рис. 2. Візуалізація будови LSTM блоку [13]

У ході досліджень ми використаємо MAE як метрику роботи мереж вона працює з власними значеннями, що дозволяє легко інтерпретувати значення. Також, дана метрика дає пряму оцінку середньої похибки, що дозволяє оцінити загальну точність прогнозів, що важливо для енергетиків, які потребують швидких та зрозумілих оцінок ефективності моделі.

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

Для аналізу оберемо вибірку споживання електроенергії 5567 домогосподарств у Лондоні, зібрану в період з листопада 2011 року до лютого 2014 року з інтервалом у півгодини [14]. Вибірка включає записи споживання електроенергії будинків різного соціально-економічного статусу та погодні дані, що включають такі змінні: температура, швидкість вітру, тиск, та словесний опис погоди (наприклад, «сонячно», «дощ», «туман»). Загалом вибірка містить 39 000 записів.

Для прогнозування загального споживання електроенергії дані всіх господарств були агреговані, утворюючи основну змінну. Оскільки погодні дані записувалися щогодини, для кожного півгодинного запису споживання використовувалася погодна інформація відповідного часу. Також, враховуючи, що змінні мають різний розмах, всі кількісні змінні були нормалізовані за допомогою мінмакс нормалізації для подальшого використання в моделі.

Для проведення експерименту зафіксуємо групу гіперпараметрів моделей:

1. Фіксований горизонт прогнозування у 25 записів, що аналогічно короткотерміновому прогнозуванню на найближчу добу, та на вхід мережі подаватимемо інформацію про споживання за попередню добу, що теж рівно 25 записів.
2. Кожна модель буде навчатись впродовж 50 епох.
3. Як оптимізатор використаємо Adam [15], з лінійною зміною швидкості навчання.

У процесі навчання перевіримо результат навчання з однієї змінної та кількох змінних: навчання на базі тільки історії споживання, та споживання, температури і атмосферного тиску. Архітектура CNN зображена на Рис.

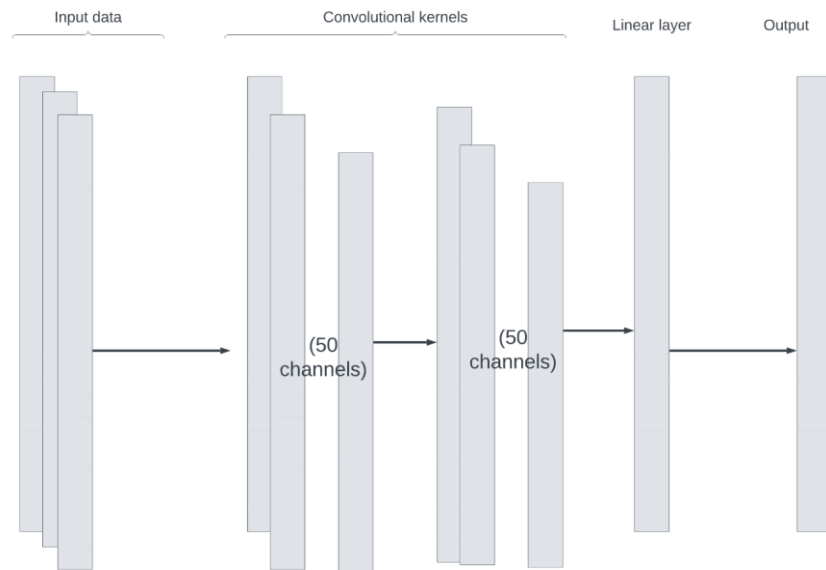


Рис. 3. Загальна архітектура CNN

В ході пошуку найкращої моделі спробуємо варіювати розмір ядра: 3 і 6 елементів. Архітектура TCN зображена на Рис.

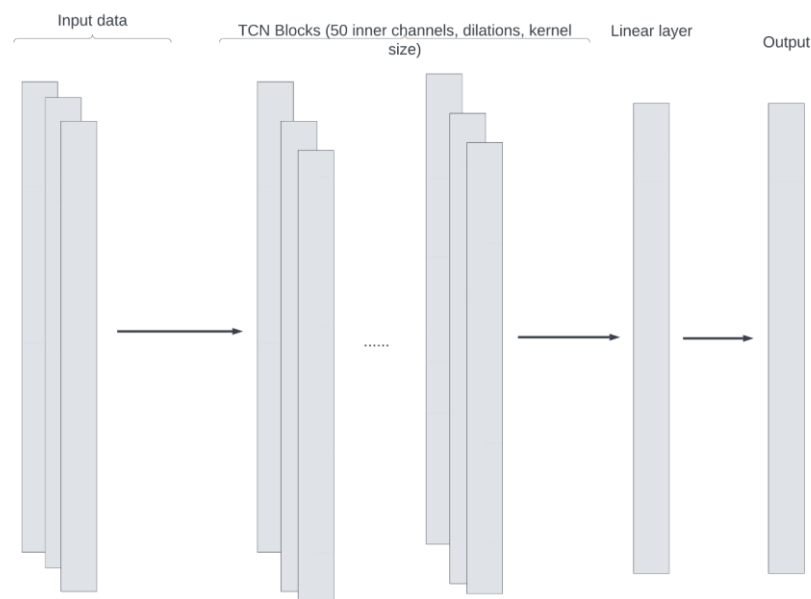


Рис. 4. Візуалізація архітектури TCN

Для пошуку оптимального розміру TCN мережі перевіримо такі параметри:

1. Враховуючи розмір вибірки, варто розглянути 1 або 2 блоки, щоб уникнути перенавчання.
2. У роботі [10] рекомендується збільшувати розширення на певний параметр, тому розглянемо

розширення розмірів: [1, 3, 5], [1, 3, 6] та [1, 2, 4].

3. Так як і у CNN мережі, встановимо розміри ядра згорток в 3 і 6.

У свою чергу, для LSTM будемо конфігурувати кількість прихованих шарів: 1 і 3, та кількість параметрів: 64 і 150.

У результаті навчання на обраній вибірці найкращі результати показали моделі, які записані в Таблиця 1.

Таблиця 1

**Параметри моделей, які показали найточніші результати**

Тип моделі	Блоки внутрішніх шарів	Кількість внутрішніх елементів	Кількість фільтрів	Розмір ядра	Розширення	MAE
LSTM	3	150	-	-	-	0.00876
CNN	2	-	50	3	-	0.00686
TCN	1	-	50	3	[1, 3, 5]	0.05332

Кожні 5 епох ми проводили перевірку моделей на валідаційній вибірці та отримали представлення метрики на рис. 5.

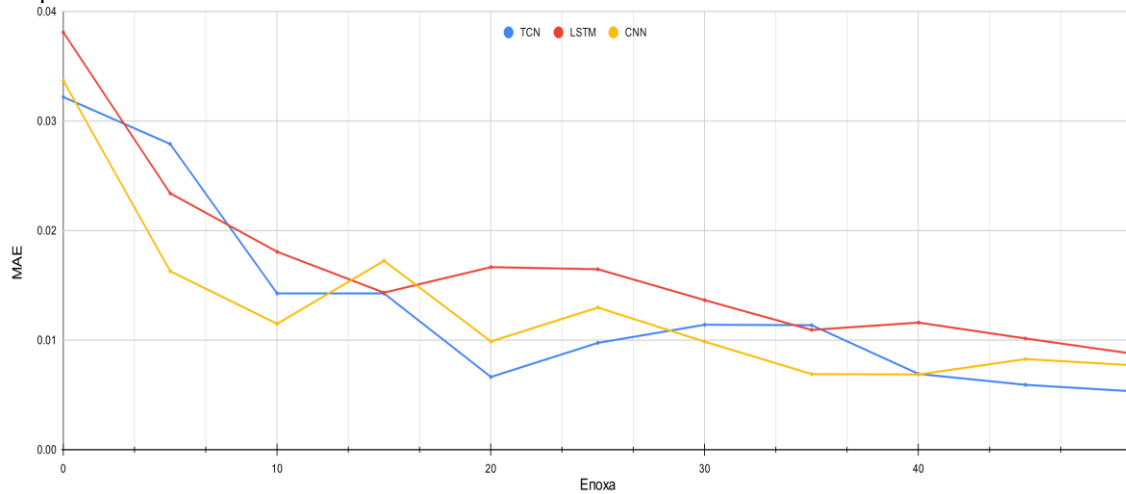


Рис. 5 Зміна метрики MAE на валідаційній вибірці

Також проведено оцінку швидкодії та розміру обраних моделей, які записані в Таблиця 2 і Таблиця 3, відповідно.

Таблиця 2

**Швидкодія обраних моделей**

Модель	Тривалість одного прогнозування (мс)	Середня тривалість однієї епохи (с)
LSTM	1.341	150
CNN	0.075	100
TCN	1.000	120

Таблиця 3

**Розмір обраних моделей**

Модель	Розмір (кВ)	Кількість параметрів
LSTM	380	96775
CNN	370	94275
TCN	185	46575

**Висновки**

Дослідження продемонструвало, що використання погодних даних при прогнозуванні електроенергії значно покращує точність моделей, враховуючи вплив зовнішніх факторів на споживання. Застосування CNN дозволило створити моделі, які працюють на 25% швидше та на 30% точніше у порівнянні з усталеними методами. Використання розширень у CNN значно покращило результати роботи моделей завдяки більш ефективній обробці часових рядів.

У рамках подальших досліджень пропонується розглянути два основних напрямки. Перший полягає у використанні даних як інтегральної частини вхідних даних для нейронних мереж, що може забезпечити значне покращення точності прогнозування за рахунок урахування часових відмінностей та сезонності в даних. Другий напрямок передбачає детальне дослідження використання архітектур трансформерів, які є усталеним підходом у задачах обробки природної мови, але ще не повсюдно застосовуються у роботі з часовими рядами.

## Література

1. Arnob S.S., M.S. Arefin A.I., Saber A.Y., Mamun K.A. Energy Demand Forecasting and Optimizing Electric Systems for Developing Countries. *IEEE Access*. 2023. vol. 11, IEEE. P. 39751–39775.
2. Chapagain K., Somsak K., Pisut K. Short-Term Electricity Demand Forecasting: Impact Analysis of Temperature for Thailand. *Energies*. 2020. vol. 13, no. 10. P. 1-29.
3. Luiz F., Afshin A. Short-term Forecasting of the Abu Dhabi Electricity Load Using Multiple Weather Variables. *Energy Procedia*. 2015. vol. 75. P. 3014-3026.
4. Durand D., Aguilar J., R-Moreno M.D. An Analysis of the Energy Consumption Forecasting Problem in Smart Buildings Using LSTM. *Sustainability*. 2022. vol.14, no. 20. P. 1-22.
5. Li Y., Wei H., Keying L., Xizheng Zh., Qin W. Short-term PV power prediction based on meteorological similarity days and SSA-BiLSTM. *Systems and Soft Computing*. 2024. vol. 6. P. 1-15.
6. Fazeel Abid, Muhammad Alam, Faten S. Alamri, Imran Siddique Multi-directional gated recurrent unit and convolutional neural network for load and energy forecasting: A novel hybridization. *AIMS Mathematics*. 2023. vol. 8, no. 9. P. 19993-20017.
7. Ribeiro A.M.N.C., P.R.X. do Carmo, I.R. Rodrigues, D. Sadok, T. Lynn, P.T. Endo Short-Term Firm-Level Energy-Consumption Forecasting for Energy-Intensive Manufacturing: A Comparison of Machine Learning and Deep Learning Models. *Algorithms*. 2020. vol. 13. P. 1-19.
8. Mahjoub S., Chrifi-Alaoui L., Marhic B., Delahoche L. Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks. *Sensors*. 2022. vol. 22, no. 11. P. 1-20.
9. Bai S., Kolter J.Z., Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*. 2019. P. 1-14.
10. Lara-Benítez P., Carranza-García M., Luna-Romera J.M., Riquelme J.C. Temporal Convolutional Networks Applied to Energy-Related Time Series Forecasting. *Appl. Sci*. 2020. vol. 10. P. 1-17.
11. Phyo P.P., Byun Y.-C. Hybrid Ensemble Deep Learning-Based Approach for Time Series Energy Prediction. *Symmetry*. 2021. vol. 13. P. 1-15.
12. Bedi J., Toshniwal D. Empirical Mode Decomposition Based Deep Learning for Electricity Demand Forecasting. *IEEE Access*. 2018. vol. 6. P. 49144-49156.
13. Long short-term memory. [https://en.wikipedia.org/wiki/Long\\_short-term\\_memory](https://en.wikipedia.org/wiki/Long_short-term_memory). (accessed on 10 January 2024).
14. Smart meters in London. <https://www.kaggle.com/datasets/jeanmidev/smart-meters-in-london>. (accessed on 10 January 2024).
15. Kingma D.P. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. 2014. P. 1-15.
16. SmartMeter Energy Consumption Data in London Households. <https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households> (accessed on 10 January 2024).