

О.П. Гожий, І.О. Калініна, В.О. Гожий, В.В. Димо

СИСТЕМНИЙ ПІДХІД ДО ПРОГНОЗУВАННЯ ПОПИТУ НА ЕЛЕКТРОЕНЕРГІЮ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Анотація. У статті досліджено системний підхід до вирішення проблеми прогнозування попиту на електроенергію в Україні на основі методів машинного навчання. Представлено послідовність етапів обробки даних при розв'язанні задачі прогнозування методами машинного навчання. Розглянуто методологію вирішення задачі прогнозування на часових рядах. Процес прогнозування складається з п'яти етапів. Перший етап включає збір, аналіз та інтерпретацію даних. Наступний етап включає процедури дослідження та підготовки даних. Третій етап - етап моделювання складається з трьох частин: підготовка набору даних для моделювання, вибір і навчання моделей і оцінка їх якості. Четвертий етап – етап прогнозування та розрахунок якісних показників прогнозів. На п'ятому етапі виконуються процедури підвищення ефективності обраної моделі прогнозування. На етапі моделювання використовувалися такі моделі: ARIMA, GAM, ANN і BSTS. Проведено аналіз моделей і на основі кожної моделі побудовано прогнози. Для побудованих моделей з найкращими показниками якості розраховано прогнозні значення. Прогнози порівнювали з даними валідаційної вибірки. Для вибору оптимальної моделі використовували такі показники: MAPE, MAE, MSE, RMSE. Найкращі результати показала модель BSTS.

Ключові слова: Forecasting, Electricity demand in Ukraine, Machine learning methods, Data processing, Methodology of forecasting, ARIMA, GAM, ANN, BSTS.

Вступ. Методи машинного навчання значно підвищують ефективність вирішення таких проблем машинного навчання, як класифікація, регресія та прогнозування. Але застосування кожного методу вимагає врахування особливостей набору даних, способу представлення даних, особливостей задачі, що розв'язується. Тому один із підходів до вирішення проблем прогнозування полягає в систематичному використанні кількох моделей прогнозування з подальшим вибором моделі, яка дає найкращі результати.

Сьогодні споживання енергії в усьому світі стрімко зростає через зростання попиту на електроенергію через постійно зростаюче населення планети, масштабний розвиток промисловості та технологій, підвищення рівня життя, масштабну індустріалізацію в країнах, що розвиваються, і необхідність підтримувати високі темпи сталого розвитку. Електроенергетика є основою економічного потенціалу будь-якої країни. Ця галузь відноситься до галузей критичної інфраструктури, які повинні дуже швидко реагувати

на зовнішні зміни, такі як стихійні лиха, військові дії, а також на зміну умов на ринку електроенергії.

Прогнозування попиту на електроенергію в Україні сьогодні є стратегічно важливим питанням, оскільки в умовах війни та постійних атак на енергетичну інфраструктуру необхідно оперативного розподіляти енергоресурси для задоволення потреб різних типів споживачів.

Враховуючи ці фактори, створення адекватних і точних моделей прогнозування попиту на електроенергію є необхідним і важливим для точного планування інвестицій у виробництво та розподіл електроенергії. Загальною складністю розробки якісних прогнозів є визначення достатньої кількості інформації для побудови прогнозних моделей. Якщо даних недостатньо, то прогноз буде неточним. Подібним чином, якщо інформація є неточною або надлишковою, попередня обробка даних і створення моделей для прогнозування буде складним. Тому виникає необхідність підвищення точності прогнозних моделей за рахунок використання сучасних ефективних методів і підходів.

В даній роботі досліджується системний підхід до вирішення завдань прогнозування. Для вирішення завдань прогнозування використовуються різноманітні методи і моделі машинного навчання, для реалізації програмних модулів використовується мова програмування *R*, та середовище розробки *RStudio*.

Постановка проблеми. Метою даної роботи є дослідження і розробка на основі системного підходу методології вирішення задач прогнозування за допомогою методів машинного навчання. А також дослідити задачу прогнозування попиту на електроенергію в Україні. Порівняти ефективність різних методів машинного навчання у вирішенні задачі прогнозування попиту на електроенергію.

Аналіз останніх досліджень. У роботі [1] наведено алгоритм прогнозування потреби в електроенергії, яка пов'язана з максимальним навантаженням в електромережі. Авторами обґрунтовано та застосовано метод СВР. Набір даних і параметри методу були налаштовані та оптимізовані гібридним методом. Такий підхід дозволяє зменшити загальну похибку прогнозу. Гібридний метод базується на поєднанні моделі нейронної мережі, методу ARIMA та модифікованого методу SVR.

У роботі [2] для обробки вихідних даних використано алгоритм штучної бджолиної колонії. Для прогнозування попиту на електроенергію алгоритм використовувався в поєднанні з ансамблевими моделями. Ряд незалежних вхідних змінних було використано для створення моделей однорідного ансамблю. Запропонована авторами модель ансамблю забезпечує більш точні прогнози. Автори [3] пропонують поєднати модель прямої нейронної мережі з архітектурою згорткової нейронної мережі для прогнозування попиту на електроенергію. Такий підхід виявився найефективнішим. Методи глибокого навчання також застосовувалися в [4]. У документі зроблено прогноз попиту на енергію на основі статистичних даних за попередні роки. Для більш глибокого вивчення даних використовувався метод кластерного аналізу. Навантаження було класифіковано за певними періодами та представлено у вигляді кластерів. Прогноз попиту на електроенергію оцінювали за допомогою нейромережових моделей і SVM.

У статті [5] представлено підхід до прогнозування попиту на електроенергію на основі гібридної моделі. Модель побудована за допомогою комбінації ARIMA та LSSVM. Результати прогнозування показують, що такий підхід до побудови моделі дозволяє враховувати аномальні значення в даних. У роботі [6] автори використали модель регресійного аналізу для прогнозування попиту на електроенергію в різних галузях промисловості.

Для оцінки пікового місячного попиту на електроенергію в [7] використовувалися наступні методи: метод ANFIS, спеціальний метод обробки даних та різні моделі нейронних мереж. У поєднанні із запропонованими моделями було виявлено, що ці методи краще підходять для визначення пікового попиту на електроенергію. У [8] він опублікував довгострокові прогнози попиту на електроенергію в Греції, а також використав зв'язок між часовими рядами та ефективними кількома критеріями. Модель оцінки витрат досліджено з використанням даних, зібраних між 1999 та 2013 роками. Вплив виробництва електроенергії в європейських країнах під час карантину досліджено в роботах [9,10].

У роботах [11,12] досліджено ефективність різних підходів і стратегій для прогнозування щоденного споживання енергії. Автори [13] досліджували підхід до прогнозування навантаження в електричній мережі за допомогою методів штучного інтелекту. Авторами [14] розроблено модель прогнозування попиту на електроенергію для житлових і комерційних будівель на основі ансамблевих методів. Розглянуто короткострокове прогнозування.

У [15] автори використовують моделі SVR у поєднанні з WOA, що включає навчання на основі еліти та хаотичної опозиції (ECWOA) для покращення результатів прогнозування. Результати експериментів показують, що врахування інформації про ціни на електроенергію призводить до підвищення точності прогнозування. У роботах [16,17] розглядаються підходи на основі алгоритмів машинного навчання для підвищення точності короткострокових прогнозів. Використовувалися такі методи: SVM, LSTM, SVR та ансамблеві структури.

У [18] для отримання прогнозних значень електроенергії використано сезонно скориговану регресію. У роботі [19] автори на конкретних прикладах продемонстрували перевагу LSTM над SVM у задачі прогнозування попиту на електроенергію. У [20–22] були досліджені різні архітектури нейронних мереж у поєднанні з евристичними алгоритмами для прогнозування попиту на електроенергію в різних країнах.

Методологія вирішення завдань прогнозування. Послідовність етапів обробки даних при розв'язанні прогнозної задачі наведена на рис. 1. Вона складається з наступних етапів: збір даних, дослідження та підготовка даних, побудова та навчання прогнозних моделей, вибір найкращих моделей та розрахунок прогнозів, оцінка та перевірка якості прогнозів. [27–28].

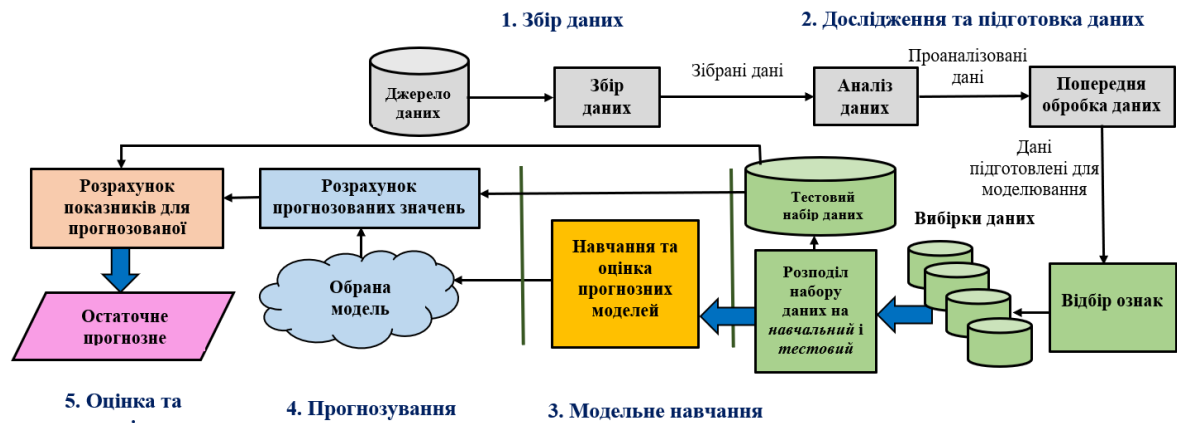


Рисунок 1 - Послідовність етапів вирішення задач машинного навчання

Інформаційна модель для вирішення задачі прогнозування за допомогою методів машинного навчання описана за допомогою наступного набору елементів:

$$ML_T = \{NData_T, NL_{Data}, NS_{Data}, MA_T, Mod_T, MP_T, Us_T\}$$

де $NData_T$ – набір наборів даних, які обробляються при вирішенні задачі машинного навчання; NL_{Data} – набір нелінійностей даних, які враховуються при вирішенні задачі машинного навчання; NS_{Data} – набір нестационарних процесів, які враховуються при вирішенні задачі машинного навчання; MA_T – набір методів аналізу та попередньої обробки масивів даних; Mod_T – набір методів побудови моделі та методів моделювання для вирішення задачі машинного навчання; MP_T – набір методів прогнозування на основі ймовірнісного статистичного аналізу з урахуванням нелінійності та нестационарності даних; Us_T - це набір невизначеностей під час вирішення задачі машинного навчання.

На основі розробленої інформаційної моделі створено опис етапів вирішення задач прогнозування та розроблено методологію вирішення задачі прогнозного моделювання на часових рядах, яка представлена на рисунку 2.

Методологія була розроблена на основі системного підходу. Методологія представлена у вигляді послідовності наступних етапів. Перший етап необхідний для збору, аналізу та інтерпретації вихідних даних. При завантаженні набір даних аналізується, визначається його структура та особливості окремих атрибутів. У результаті попередньої обробки даних набір готується для подальшого аналізу – інтелектуального аналізу даних.

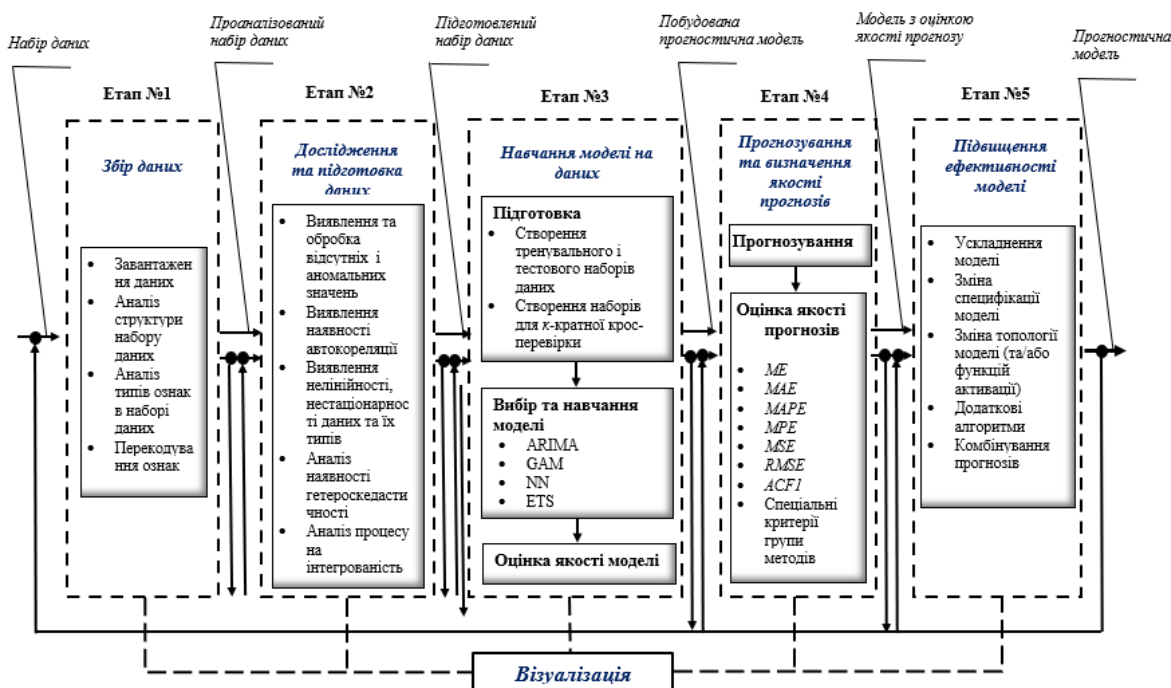


Рисунок 2 - Методологія вирішення задачі прогнозного моделювання на часових рядах

На *першому* етапі відбувається збір даних. *Другий* етап передбачає процедури дослідження та підготовки даних. Аналізується описова статистика для кожної змінної, ідентифікуються відсутні та аномальні значення, визначається рівень автокореляції, ідентифікуються нелінійність і нестационарність даних та їх типи, аналізується гетероскедастичність та аналізується процес на інтегрованість.

Третій етап - етап моделювання складається з трьох частин: підготовка набору даних для моделювання, вибір і навчання моделей і оцінка їх якості. Перед початком моделювання підготовлений набір даних розбивається на навчальну та тестову вибірки, а також створюються набори для перехресної перевірки. При виборі моделі алгоритми моделювання перевіряються на навчальній вибірці і вибирається найкращий за певними критеріями якості.

Четвертий етап — етап побудови прогнозів та оцінки їх якості. На *п'ятому* етапі виконуються процедури підвищення ефективності обраної моделі прогнозування. Для різних методів моделювання для підвищення якості можна використовувати наступні методи: ускладнення структури моделі, зміна її специфікацій, зміна топології моделі (та/або функцій активації), використання додаткових алгоритмів, комбінування прогнозних значень.

Важливою особливістю представленого методу є візуалізація. За допомогою візуалізації на кожному етапі можна скорегувати послідовність дій і повернутися до попередніх етапів. Етапи методу мають особливості, що відображають предметну область вирішення проблеми.

Прогнозування попиту на електроенергію. Набір даних. Данні для прогнозування попиту на електроенергію в Україні представлено на веб-ресурсі державного

оператора ринку електроенергії [23]. Інтервал часу спостереження охоплює період з 01.07.2019 по 04.10.2024. Набір містить погодинні дані про обсяги електроенергії, продажі електроенергії, попит на неї в МВт·год в електромережах України та ціни (рис. 3).

```
'data.frame': 69385 obs. of 10 variables:
 $ date      : chr "2019-07-01" "2019-07-01" "2019-07-01" "2019-07-01"
 $ hour      : int  1 2 3 4 5 6 7 8 9 10 ...
 $ energy_system : chr "Burshtyn peninsula" "Burshtyn peninsula" "Burshtyn peninsula" "Burshtyn peninsula"
 $ price     : num  939 939 939 939 939 939 939 939 2040 2040 ...
 $ amount_sales : num  313 241 259 153 153 ...
 $ amount_purchase : num  313 241 259 153 153 ...
 $ amount_sales_nominated : num  313 241 259 153 153 ...
 $ amount_purchase_nominated : num  364 334 320 320 317 ...
 $ demand    : num  -50.7 -93.3 -60.8 -167.9 -164.5 ...
 $ price_cap  : num  959 959 959 959 959 ...
```

Рисунок 3 - Структура набору даних «електрика»

Змінна з годинниковими позначками в день включається до змінної індексування дати та часу, а змінна попиту вибирається з набору як результуюча змінна. Зазначається, що змінна `energy_system` приймає лише три значення: півострів Бурштин, ОЕС України, ОЕС України (синхронізовано з системами ENTSO-E). Значення цієї змінної до 24 лютого 2022 року було розділено на дві окремі підсистеми: «ЄС України» та «Бурштинський енергоострів». З 24 лютого 2022 року українська енергосистема синхронізована з європейською енергосистемою ENTSO-E. Таким чином, до 24 лютого 2024 року в датасеті містяться дані попиту окремо для двох підсистем, а для розрахунку загальноукраїнських показників значення для цих підсистем підсумовуються (рис. 4).

```
> # Aggregation of the demand of individual power grid subsystems into one time series
> energy_demand <- data.frame(datetime = unique(data$datetime), demand=0)
> for(i in 1:length(energy_demand$datetime))
+ {
+   energy_demand$demand[i] = sum(data$demand[which(data$datetime==energy_demand$datetime[i])])
+ }
```

Рисунок 4 - Фрагмент коду для агрегування попиту окремих підсистем електромережі в один часовий ряд

Після перевірки в наборі даних було знайдено 5 відсутніх значень. Відсутні значення мають річний інтервал і реєструються в кінці березня або на початку квітня. Наявність розривів пов'язана з переходом годинників на літній час. Враховуючи, що загальний обсяг спостережень становить 69 385, для заповнення пропущених значень використано метод LOCF [24]. Після аналізу часових рядів частоту вибірки було зменшено. Погодинні спостереження попиту на електроенергію в наборі були агреговані в середньодобові.

Для побудови прогнозних моделей ми використали частину даних з 01.06.2022 по 04.10.2024, за винятком пікових ділянок ряду близько 24.02.2024 (рис. 5). На рисунку видно, що попит може приймати від'ємні значення – це відповідає ситуації, коли обсяги продажу електроенергії перевищують обсяги купівлі.

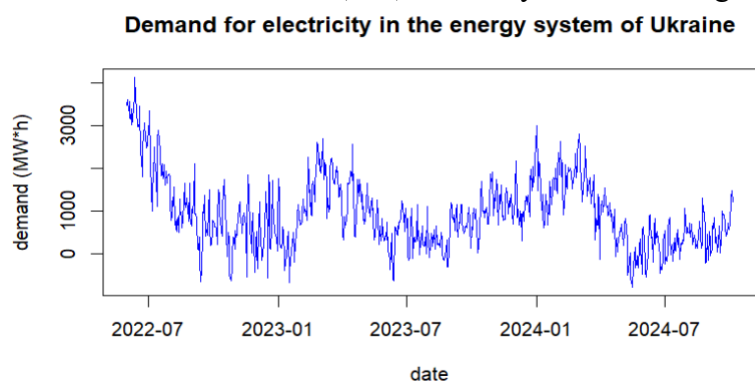


Рисунок 5 - Графік фрагмента часового ряду energy_demand

Важливою умовою побудови надійних прогностичних моделей є розуміння структури часових рядів. Декомпозиція ряду методом STL [25,26] дозволила визначити основні принципи моделювання. Гіпотетична нестационарність часового ряду підтверджується тестами Льюнг-Бокса на незалежність, розширеним тестом Дікі-Фуллера, тестом KPSS на рівень стаціонарності та тестом Філіпса-Перрона на одиничний корінь.

Часові ряди були перевірені на нелінійність за допомогою набору статистичних тестів. Результати випробувань підтвердили візуальну нелінійність ряду, оскільки р-значення менше 0,05. Гетероскедастичність ряду підтверджена тестом Маклеода-Лі (р-значення менше 0,05). В результаті тестів визначається кількість необхідних диференціацій і сезонних диференціацій для отримання стаціонарного часового ряду. Для перевірки часових рядів на автокореляцію були використані тести Дарбіна-Ватсона та Брейша-Готфрі. Для обох тестів отримані р-значення значно менші за 0,05, що підтверджує автокореляцію часового ряду. Графіки вибіркової автокореляційної функції (ACF) і часткової автокореляційної функції (PACF) на рисунку 8 відповідають очікуванням: автокореляційна функція монотонно зменшується зі збільшенням часового зсуву.

Результати попереднього аналізу підтверджують належність досліджуваного процесу до класу нелінійних і нестационарних.

Побудова та оцінка прогностичних моделей. Етап моделювання починається з поділу набору даних на дві частини: навчальну та тестову вибірки. Останні 14 спостережень (двотижневий діапазон) зберігаються як тестові спостереження, що відповідає горизонту прогнозу в 14 днів для короткострокового прогнозування (рис.6).

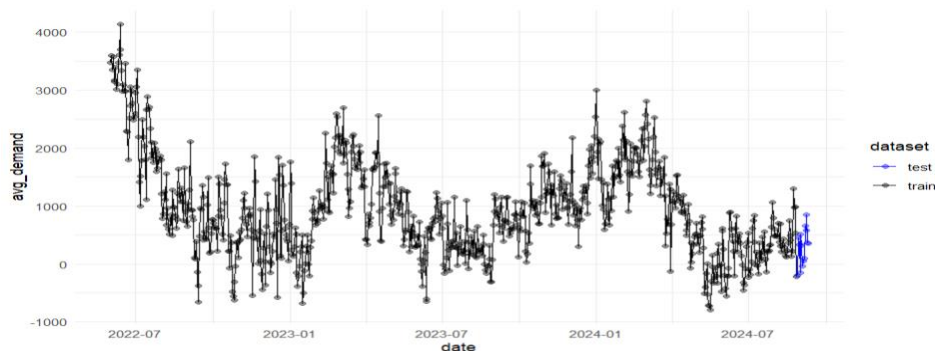


Рисунок 6 - Навчальні (чорний) та тестові (синій) дані щодо динаміки середньодобової потреби в електроенергії

На етапі моделювання в якості базових моделей прогнозу були використані наступні статистичні моделі: ARIMA, метод підгонки адитивних регресійних моделей (GAM), штучні нейронні мережі прямого поширення (NNAR), байєсовські структурні моделі часових рядів (BSTS). Вибір цих моделей обумовлений їх здатністю розпізнавати складні закономірності в рядах реального часу з урахуванням нелінійності та нестационарності досліджуваного процесу [27].

Прогнозування та оцінка результатів. Моделі з найкращими показниками якості були покладені в основу розрахунку прогнозних значень (табл. 1). Показники якості моделей визначали на дослідному зразку. Для вибору оптимальної моделі використовувалися такі показники: MAPE, MAE, MSE та RMSE.

На рисунку 7 показано візуалізацію прогнозних значень для часових рядів попиту на електроенергію в Україні, побудованих за допомогою моделі BSTS (M7). Чорна лінія представляє 50 навчальних даних, синя лінія представляє прогнозовані значення часового ряду. Жовті крапки на графіку – це дані досліджуваного зразка. Зелені пунктирні лінії обмежують 95% довірчий інтервал прогнозних значень.

Таблиця 1

Порівняльна таблиця критеріїв якості прогнозу для різних моделей

Types of models	RMSE	MAE	MAPE	MSE
ARIMA (1, 1, 2) (2, 0, 0) ₇	360.778	320.009	3.578	130160.77
GAM (annual and weekly seasonal components)	387.664	339.114	3.704	150283.38
BSTS (robust local linear trend + autoregressive component + weekly seasonality component)	289.527	215.634	1.002	83825.84
NNAR (n=15, Sigmoid, maxit=5000)	352.032	314.993	3.289	123926.34

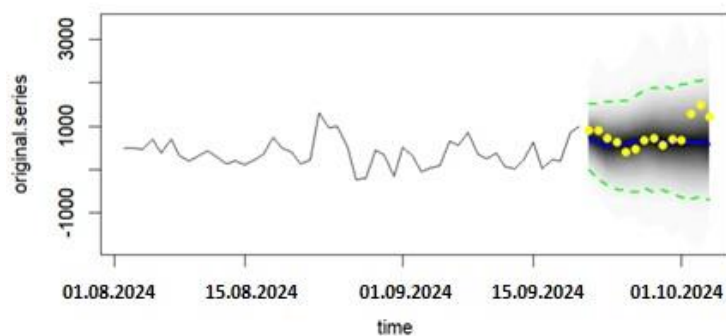


Рисунок 7 - Візуалізація прогнозних значень для фрагмента часового ряду

Представлені результати демонструють ефективність використання моделей BSTS для вирішення задач прогнозування. Подальше покращення результатів прогнозування можливе за рахунок використання комбінованих прогнозів [27, 28].

Висновки. У роботі представлено дослідження системного підходу до прогнозування. Була розглянута задача прогнозування попиту на електроенергію в Україні за допомогою методів машинного навчання. Дослідження проводили на основі даних 2019-2024 років. Розроблено та представлено послідовність етапів обробки даних при розв'язанні задачі прогнозування методами машинного навчання. Послідовність включає наступні кроки: збір даних, дослідження та підготовка даних, побудова та навчання моделей прогнозу, вибір найкращої моделі та розрахунок прогнозів, оцінка та перевірка якості прогнозів. Запропоновано загальну методику вирішення задач прогнозування часових рядів на основі методів машинного навчання. Рішення задачі прогнозування складається з п'яти етапів. На першому етапі здійснюється збір, аналіз та інтерпретація даних. На другому етапі проводяться дослідження та процедури підготовки даних. Третій етап - етап моделювання складається з трьох частин: підготовка набору даних для моделювання, вибір і навчання моделей і оцінка їх якості. Четвертий етап – етап прогнозування та визначення якості прогнозів. На п'ятому етапі виконуються процедури підвищення ефективності обраної моделі прогнозування. На етапі моделювання були використані такі моделі: ARIMA, GAM, ANN та BSTS. Було проведено детальний аналіз моделей і зроблено прогнози на основі кожної моделі. Для побудованих моделей з найкращими показниками якості розраховано прогнозні значення. Прогноз розроблявся на 2 тижні. Прогнози порівнювали з даними валідаційної вибірки. Для вибору оптимальної моделі та її оцінки використовувалися такі показники: MAPE, MAE, MSE, RMSE. Найкращі результати показала модель BSTS. Це підтверджує ефективність моделі BSTS при прогнозуванні на реальних даних.

ЛІТЕРАТУРА / REFERENCE

1. M. R. Kazemzadeh, A. Amjadian, T. Amraee. A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting. *Energy* 2020, 204, 117948.
2. J. Hao, X. Sun, Q. Feng. A Novel Ensemble Approach for the Forecasting of Energy Demand Based on the Artificial Bee Colony Algorithm. *Energies* 2020, 13, 550.
3. J. del Real, F. Dorado, J. Durán. Energy Demand Forecasting Using Deep Learning: Applications for the French Grid. *Energies* 2020, 13, 2242.
4. J. Bedi, D. Toshniwal. Deep learning framework to forecast electricity demand. *Appl. Energy* 2019, 238, pp. 1312–1326.
5. F. Kaytez. A hybrid approach based on autoregressive integrated moving average and least-square support vector machine for long-term forecasting of net electricity consumption. *Energy* 2020, 197, 117200.
6. S. Di Leo, P. Caramuta, P. Curci, C. Cosmi. Regression analysis for energy demand projection: An application to TIMES-Basilicata and TIMES-Italy energy models. *Energy* 2020, 196, 117058.

7. P. Ramsami, R. T. A. King. Neural Network Frameworks for Electricity Forecasting in Mauritius and Rodrigues Islands. In Proceedings of the 2021 IEEE PES/IAS PowerAfrica, Nairobi, Kenya, 2021, pp. 1–5.
8. D. Angelopoulos, Y. Siskos, J. Psarras. Disaggregating time series on multiple criteria for robust forecasting: The case of long-term electricity demand in Greece. *Eur. J. Oper. Res.* 2019, 275, pp. 252–265.
9. U. Şahin, S. Ballı, Y. Chen. Forecasting seasonal electricity generation in European countries under COVID-19-induced lockdown using fractional grey prediction models and machine learning methods. *Appl. Energy* 2021, 302, 117540.
10. R. Hou, S. Li, M. Wu, G. Ren, W. Gao, M. Khayatnezhad. Assessing of impact climate parameters on the gap between hydropower supply and electricity demand by RCPs scenarios and optimized ANN by the improved Pathfinder (IPF) algorithm. *Energy* 2021, 237, 121621.
11. Baba. Advanced AI-based techniques to predict daily energy consumption: A case study. *Expert Syst. Appl.* 2021, 184, 115508.
12. M. Pegalajar, L.G. B. Ruíz, M. P. Cuéllar, R. Rueda. Analysis and enhanced prediction of the Spanish Electricity Network through Big Data and Machine Learning techniques. *Int. J. Approx. Reason.* 2021, 133, pp. 48–59.
13. N. M. M. Bendaoud, N. Farah, S. Ben Ahmed. Applying load profiles propagation to machine learning based electrical energy forecasting. *Electr. Power Syst. Res.* 2022, 203, 107635.
14. R. Porteiro, L. Hernández-Callejo, S. Nesmachnow. Electricity demand forecasting in industrial and residential facilities using ensemble machine learning. *Rev. Fac. De Ing.* 2022, 102, pp. 9–25.
15. Y. Lu, G. Wang. A load forecasting model based on support vector regression with whale optimization algorithm. *Multimed. Tools Appl.* 2023, 82, pp. 9939–9959.
16. S. Li, X. Kong, L. Yue, C. Liu, M. A. Khan, Z. Yang, H. Zhang. Short-Term Electrical Load Forecasting Using Hybrid Model of Manta Ray Foraging Optimization and Support Vector Regression. *J. Clean. Prod.* 2023, 388, 135856.
17. J. Huang, M. Algahtani, S. Kaewunruen. Energy Forecasting in a Public Building: A Benchmarking Analysis on Long Short-Term Memory (LSTM), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost) Networks. *Appl. Sci.* 2022, 12, 9788.
18. C. E. Velasquez, M. Zocatelli, F. B. Estanislau, V. F. Castro. Analysis of time series models for Brazilian electricity demand forecasting. *Energy* 2022, 247, 123483.
19. F. Pallonetto, C. Jin, E. Mangina. Forecast electricity demand in commercial building with machine learning models to enable demand response programs. *Energy AI* 2022, 7, 100121.
20. E. C. May, A. Bassam, L. J. Ricalde, M. E. Soberanis, O. Oubram, O. M. Tzuc, A. Y. Alanis, A. Livas-García. Global sensitivity analysis for a real-time electricity market forecast by a machine learning approach: A case study of Mexico. *Int. J. Electr. Power Energy Syst.* 2022, 135, 107505.

21. W. J. Niu, Z. K. Feng, S. S. Li, H. J. Wu, J. Y. Wang. Short-term electricity load time series prediction by machine learning model via feature selection and parameter optimization using hybrid cooperation search algorithm. *Environ. Res. Lett.* 2021, 16, 055032.
22. R. Luzia, L. Rubio, C. E. Velasquez. Sensitivity analysis for forecasting Brazilian electricity demand using artificial neural networks and hybrid models based on Autoregressive Integrated Moving Average. *Energy* 2023, 274, 127365.
23. Energy Map. URL: <https://map.ua-energy.org/uk/resources/5a616fba-fbc9-4073-9532-9161592faca8/>
24. P. Bidyuk, I. Kalinina, A. Gozhyj. An Approach to Identifying and Filling Data Gaps in Machine Learning Procedures. *Lecture Notes on Data Engineering and Communications Technologies (Switzerland)*. 2022. Vol. 77, pp. 164-176.
25. R. J. Hyndman, G. Athanasopoulos. *Forecasting: Principles and Practice* 3rd ed. Edition. Publisher: OTexts. 2021, 442 p.
26. T. Aggarwal. Master the Power of Seasonal Decomposition of Time Series (STL): Unveiling the Essence of Time. 2023. URL: https://medium.com/@tushar_aggarwal/master-the-power-of-seasonal-decomposition-of-time-series-stl-unveiling-the-essence-of-time-26c19a910314
27. Kalinina., P. Bidyuk., A. Gozhyj, P. Malchenko. Combining Forecasts Based on Time Series Models in Machine Learning Tasks. *CEUR-WS*. 2023. Vol. 3426. Pp. 25-35. CEUR-WS.org/Vol-3426/paper2.pdf.
28. Kalinina, P. Bidyuk, A. Gozhyj. Construction of Forecast Models based on Bayesian Structural Time Series. *International Scientific and Technical Conference on Computer Sciences and Information Technologies. CSIT_2022*. 2022. Pp. 180-184. doi: 10.1109/CSIT56902.2022.10000484.

Received 31.03.2025.
Accepted 01.04.2025.

System approach to forecasting electricity demand based on machine learning

The article explores a systemic approach to solving the problem of forecasting electricity demand in Ukraine based on machine learning methods. The sequence of data processing stages when solving the forecasting problem using machine learning methods is presented. The methodology for solving the forecasting problem on time series is considered. The forecasting process consists of five stages. The first stage includes data collection, analysis and interpretation. The next stage includes data research and preparation procedures. The third stage - the modeling stage consists of three parts: preparation of a data set for modeling, selection and training of models and assessment of their quality. The fourth stage is the forecasting stage and calculation of quality indicators of forecasts. At the fifth stage, procedures are performed to increase the efficiency of the selected forecasting model. At the modeling stage, the following models were used: ARIMA, GAM, ANN and BSTS. The models were analyzed and forecasts were built based on each model. For the built models with the best quality indicators, forecast values were calculated. The forecasts were compared with the data of the validation sample. To select the optimal model, the following indicators were used: MAPE, MAE, MSE, RMSE. The BSTS model showed the best results.

Keywords: Forecasting, Electricity demand in Ukraine, Machine learning methods, Data processing, Methodology of forecasting, ARIMA, GAM, ANN, BSTS

Гожий Олександр Петрович – професор кафедри інтелектуальних інформаційних систем Чорноморського національного університету ім. Петра Могили, д.т.н., професор.

Калініна Ірина Олександрівна – Професор кафедри інтелектуальних інформаційних систем Чорноморського національного університету ім. Петра Могили, д.т.н., доцент.

Гожий Віктор Олександрович – старший викладач кафедри інтелектуальних інформаційних систем Чорноморського національного університету ім. Петра Могили, к.т.н..

Димо Валерій Володимирович – аспірант, кафедри інтелектуальних інформаційних систем Чорноморського національного університету ім. Петра Могили.

Gozhyj Oleksandr Petrovych - professor of the Department of Intellectual Information Systems of the Black Sea National University named after Petra Mohyly, D.of I.S., professor.

Kalinina Iryna Oleksandrivna - Professor of the Department of Intellectual Information Systems of the Black Sea National University named after Petra Mohyly, D.of I.S., associate professor.

Gozhyi Viktor Oleksandrovych – Senior Lecturer, Department of Intellectual Information Systems, Petro Mohyla Black Sea National University, Candidate of Technical Sciences.

Dymo Valeriy Volodymyrovych – Postgraduate Student, Department of Intellectual Information Systems, Petro Mohyla Black Sea National University.