

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ВАДИМА ГЕТЬМАНА**

**Навчально-науковий інститут  
«Інститут інформаційних технологій в економіці»**

**Кафедра математичного моделювання та статистики**

**Освітньо-професійна програма «Економічна кібернетика і Дата Сайнс»**

Галузь знань 05 «Соціальні та поведінкові науки»  
Спеціальність 051 «Економіка»

Форма навчання: очна (денна)

**КВАЛІФІКАЦІЙНА МАГІСТЕРСЬКА РОБОТА**

на тему «Моделювання динаміки споживання електроенергії в Україні»  
(назва теми)

здобувача Вікторової Ярослави Сергіївни  
(ПІБ, підпис)

Науковий керівник: кандидат економічних наук, доцент Ольга КАТУНІНА

\_\_\_\_\_  
(підпис)

**Робота допущена до захисту перед екзаменаційною комісією  
з атестації здобувачів вищої освіти (ЕК)**

Завідувач кафедри: к. ф.-м. н., професор  
Галина ВЕЛИКОІВАНЕНКО

\_\_\_\_\_  
(підпис)

**Київ 2024**

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ВАДИМА ГЕТЬМАНА**

**Навчально-науковий інститут  
«Інститут інформаційних технологій в економіці»**

**Кафедра математичного моделювання та статистики**

**Освітньо-професійна програма «Економічна кібернетика і Дата  
Сайнс»**

Галузь знань 05 «Соціальні та поведінкові науки»  
Спеціальність 051 «Економіка»

**ПОГОДЖЕНО**

Керівник проектної групи (гарант)  
освітньо-професійної програми  
Олена ПІСКУНОВА

*(підпис)*

2024 р.

**ЗАТВЕРДЖУЮ**

Завідувач кафедри математичного  
моделювання та статистики  
Галина ВЕЛИКОІВАНЕНКО

*(підпис)*

2024 р.

**ІНДИВІДУАЛЬНЕ ЗАВДАННЯ**

**здобувача вищої освіти**

Вікторової Ярослави Сергіївни

*(прізвище, ім'я, по батькові)*

очної (денної)

**форми навчання**

*очної (денної), заочної, дистанційної*

на підготовку кваліфікаційної магістерської роботи

**на тему**

«Моделювання динаміки споживання електроенергії в Україні»

Тему затверджено наказом ректора Університету від «19» вересня 2024 р. № 1610 /  
ст

Кваліфікаційна магістерська робота виконується на матеріалах  
сайту Energy Map <https://map.ua-energy.org> та інших відкритих джерел.

## План кваліфікаційної магістерської роботи

|          |                                                                                        |
|----------|----------------------------------------------------------------------------------------|
| Розділ 1 | Теоретичні засади аналізу та прогнозування розвитку електроенергетичного ринку України |
| Розділ 2 | Математичні моделі аналізу динаміки споживання електроенергії в Україні                |
| Розділ 3 | Аналіз та моделювання динаміки споживання електроенергії в Україні                     |

|                                                      |                                                                                                                                                                                                                                                                    |
|------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Об'єкт дослідження:                                  | рівень споживання електроенергії в Україні                                                                                                                                                                                                                         |
| Предмет дослідження:                                 | прогнозування рівня споживання електроенергії в Україні за допомогою методів та моделей прогнозування часових рядів                                                                                                                                                |
| Мета виконання кваліфікаційної магістерської роботи: | моделювання споживання електроенергії в Україні з використанням методів прогнозування часових рядів, зокрема методу експоненціального згладжування Holt-Winters, моделей ARIMA/SARIMA, а також LSTM. Визначення найефективнішого для підвищення точності прогнозу. |

**Конкретні завдання, які здобувач повинен виконати для досягнення поставленої мети:**

**У розділі 1:** 1) проаналізувати структуру та основні тенденції розвитку електроенергетичного ринку України; 2) дослідити напрями та особливості споживання електроенергії; 3) розглянути базові підходи до математичного моделювання та прогнозування споживання електроенергії.

**У розділі 2:** 1) дослідити інструментарій розвідувального аналізу часових рядів; 2) дослідити методи експоненціального згладжування, ARIMA/SARIMA та нейромережеві моделі для аналізу динаміки споживання електроенергії; 3) визначити систему метрик для оцінки точності моделей;

**У розділі 3** 1) виконати візуалізацію та препроцесинг даних про споживання електроенергії в Україні; 2) побудувати прогнози споживання електроенергії на основі розроблених моделей; 3) визначити кращу модель і побудувати прогноз; 4) оцінити практичну ефективність моделей для управління енергоспоживанням.

Завдання підготував  
науковий керівник

\_\_\_\_\_  
(підпис)

Ольга КАТУНІНА  
(ім'я, прізвище)

«\_\_» \_\_\_\_\_ 2024 р.

Завдання одержав  
здобувач

\_\_\_\_\_  
(підпис)

Ярослава ВІКТОРОВА  
(ініціали, прізвище)

«\_\_» \_\_\_\_\_ 2024 р.

## Реферат

Кваліфікаційна магістерська робота містить 78 сторінок, 4 таблиці, 26 рисунків, перелік використаних джерел з 44 найменувань, додаток.

### **«Моделювання динаміки споживання електроенергії в Україні»**

*Об'єктом дослідження є рівень споживання електроенергії в Україні.*

*Предметом дослідження є прогнозування рівня споживання електроенергії в Україні за допомогою методів та моделей прогнозування часових рядів.*

*Мета та завдання виконання дослідження.* Мета полягає у прогнозуванні споживання електроенергії в Україні з використанням методів прогнозування часових рядів, зокрема методу експоненціального згладжування Holt-Winters, моделей ARIMA/SARIMA, а також LSTM. Визначення найефективнішого для підвищення точності прогнозу.

Відповідно до поставленої мети визначені такі *завдання*:

- проаналізувати структуру та основні тенденції розвитку електроенергетичного ринку України;
- обґрунтувати вибір інструментарію математичного моделювання для дослідження динаміки споживання електроенергії в Україні;
- побудувати систему моделей для аналізу та прогнозування споживання електроенергії на базі моделей експоненціального згладжування, ARIMA - моделей та нейромережових моделей;
- оцінити прогностичну якість побудованих моделей за системою метрик, здійснити підбір гіперпараметрів для налаштування моделей прогнозування споживання електроенергії;
- вибрати найкращий модельний інструментарій та застосувати його для предиктивного моделювання обсягів споживання електроенергії в Україні.

*Теоретична, методична та практична значущість отриманих результатів.* Під час дослідження було систематизовано сучасні підходи до математичного моделювання та прогнозування споживання електроенергії, проаналізовано їх недоліки та переваги. Застосовано методи експоненціального згладжування Holt-Winters, ARIMA/SARIMA та нейромережові моделі LSTM для оцінки динаміки споживання електроенергії. Розроблено алгоритми для програмної реалізації запропонованих моделей, а коректність їх роботи перевірено на реальних статистичних даних енергетичного ринку України.

Практичні результати дослідження ти можуть бути використані для покращення управління енергоресурсами в Україні, розробки рекомендацій для операторів електроенергетичних систем та оптимізації планування навантажень.

Рік виконання кваліфікаційної магістерської роботи – 2024.

Рік захисту роботи – 2024.

*Ключові слова:* споживання електроенергії, часові ряди, прогнозування, ARIMA, LSTM, математичне моделювання, точність прогнозів.

**В і д г у к**  
про кваліфікаційну магістерську роботу  
здобувача навчально-наукового інституту "Інститут інформаційних  
технологій в економіці" Київського національного економічного  
університету імені Вадима Гетьмана  
освітньої програми "Економічна кібернетика і Дата Сайнс"  
**Вікторової Ярослави Сергіївни**  
на тему "Моделювання динаміки споживання електроенергії в Україні"

**1. Актуальність теми.** Питання, пов'язані з прогнозуванням споживання електроенергії, є актуальними, відіграють ключову роль у плануванні енергетичних систем, розподілі навантаження та забезпеченні ефективного управління енергетичними процесами. В умовах глобального переходу до відновлювальних джерел енергії та інтеграції сучасних технологій управління зростає затребуваність використання інструментів математичного моделювання для побудови прогнозів споживання електроенергії.

**2. Позитивні риси кваліфікаційної магістерської роботи.** Робота характеризується використанням низки моделей для аналізу та прогнозування часових рядів. Дослідження базується на реальних даних про споживання електроенергії в Україні, що забезпечує його прикладне спрямування.

**3. Наявність самостійних розробок автора.** У роботі проаналізовано структуру ринку споживання електроенергії в Україні, розглянуто методи математичного моделювання динаміки споживання, включаючи експоненціальне згладжування, ARIMA/SARIMA та нейромережеві моделі. Побудовано прогностичні моделі для визначення обсягів споживання електроенергії в Україні, здійснено їхнє порівняння за основними метриками.

**4. Цінність теоретичних висновків та практичних рекомендацій.** Запропоновані моделі можуть використовуватись для прийняття рішень у сфері енергетичного менеджменту, прогнозування та планування роботи енергетичних підприємств з метою підвищення ефективності використання енергоресурсів, зниження витрат та забезпечення стабільності енергопостачання.

**5. Наявність недоліків.** У дослідженні варто було б більш чітко систематизувати сучасні наукові підходи до моделювання динаміки споживання електроенергії; при дослідженні одного часового ряду обсягу споживання доцільно було б навести напрями та інструменти визначення інформативних ознак (Feature engineering) у структурі часового ряду для подальшого використання у моделях машинного навчання; доречним було б дослідити вплив факторів, що зумовлюють динаміку змінення обсягів споживання електроенергії в економіці України.

**6. Загальна оцінка кваліфікаційної магістерської роботи та її допущення до захисту перед ЕК.** Зважаючи на те, що кваліфікаційна магістерська робота здобувача Вікторової Ярослави Сергіївни на тему "Моделювання динаміки споживання електроенергії в Україні" є актуальною, дослідження є самостійним, усі поставлені завдання виконано, вважаю, що дана кваліфікаційна магістерська робота відповідає усім вимогам, рівень її виконання є задовільним, робота може бути допущена до захисту перед ЕК.

Науковий керівник: к.е.н, доцент, доцент кафедри математичного моделювання та статистики

\_\_\_\_\_ Катуніна О.С.

“\_\_” \_\_\_\_\_ 2024

## РЕЦЕНЗІЯ

на кваліфікаційну магістерську роботу  
здобувача навчально-наукового інституту "Інститут інформаційних технологій в економіці" Київського національного економічного університету імені Вадима Гетьмана  
освітньо-професійної програми "Економічна кібернетика і Дата Сайнс"  
Вікторової Ярослави Сергіївни  
на тему "Моделювання динаміки споживання електроенергії в Україні"

**Актуальність теми кваліфікаційної магістерської роботи і доцільність її розроблення.** У зв'язку зі стрімкими змінами економічної ситуації на глобальному та державному рівнях проведення аналізу та моделювання динаміки споживання електроенергії є актуальним. Електроенергетичний ринок України є предметом аналізу та вивчення багатьох дослідників. Прогнозування споживання електроенергії є важливим інструментом для ведення професійної діяльності диспетчерів ринку електроенергії, генеруючих та розподільчих компаній.

**Якість проведеного дослідження.** У роботі детально досліджено теоретичні аспекти моделювання динамічних процесів на сучасному ринку електроенергії. Математичний інструментарій моделювання та програмні засоби побудови прогнозів використані достатньо кваліфіковано. Якість моделей оцінено за системою метрик, що використовуються для порівняльного аналізу альтернативних інструментів моделювання у машинному навчанні.

**Позитивні риси кваліфікаційної магістерської роботи.** У роботі проведено детальний аналіз специфіки ринку електроенергії в Україні, систематизовано сучасні підходи до моделювання динамічних процесів на ринку споживання електроенергії, використано сучасний інструментарій для аналізу та прогнозування обсягів споживання електроенергії. Моделювання виконано на реальних даних електроенергетичного ринку України, усі аналітичні розрахунки, спостереження і висновки добре структуровані та проілюстровані.

**Зауваження.** У системі розглянутих у роботі модельних технологій дослідження динамічних тенденцій ринку споживання електроенергії було б доцільним додатково розглянути багатофакторні динамічні моделі, що дадуть можливість врахувати систему чинників, які впливають на обсяги споживаної електроенергії.

**Практична значимість висновків і рекомендацій.** Отримані результати та побудовані моделі для аналізу та прогнозування попиту на електроенергетичному ринку можуть використовуватися для диспетчеризації на ринку електроенергії, для розподілення виробництва електроенергії між генеруючими потужностями, побудови оптимальної стратегії для економічних суб'єктів електроенергетичного ринку.

### **Загальний висновок рецензента і оцінка кваліфікаційної магістерської роботи.**

Кваліфікаційна магістерська робота Вікторової Ярослави Сергіївни на тему "Моделювання динаміки споживання електроенергії в Україні" є актуальною, проведене дослідження є достатньо повним та обґрунтованим і відповідає вимогам, що висуваються до наукового дослідження у кваліфікаційних магістерських роботах. Вважаю, що кваліфікаційна магістерська робота Вікторової Ярослави Сергіївни може бути рекомендована до захисту в ДЕК і заслуговує позитивної оцінки.

### **Рецензент**

Місце роботи та посада рецензента

Науковий ступінь, учене звання (за наявності) \_\_\_\_\_  
(підпис, ПІБ)

Підпис засвідчую:

\_\_\_\_\_  
(посада, підпис)

Місце печатки організації, де працює рецензент

## Зміст

|                                                                                                                                                              |           |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>ВСТУП.....</b>                                                                                                                                            | <b>3</b>  |
| <b>РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ РОЗВИТКУ ЕЛЕКТРОЕНЕРГЕТИЧНОГО РИНКУ УКРАЇНИ.....</b>                                                 | <b>5</b>  |
| 1.1 Структура та тенденції розвитку електроенергетичного ринку України .....                                                                                 | 5         |
| 1.2 Напрями та особливості споживання електроенергії.....                                                                                                    | 8         |
| 1.3 Базові підходи до математичного моделювання та прогнозування споживання електроенергії .....                                                             | 13        |
| <b>РОЗДІЛ 2. МАТЕМАТИЧНІ МОДЕЛІ АНАЛІЗУ ДИНАМІКИ СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ В УКРАЇНІ .....</b>                                                               | <b>21</b> |
| 2.1 Інструментарій розвідувального аналізу одномірних часових рядів обсягів споживання електроенергії .....                                                  | 21        |
| 2.2 Методи експоненціального згладжування, авторегресійні та нейромережеві моделі аналізу динаміки споживання електроенергії .....                           | 29        |
| 2.3 Система метрик для оцінювання якості динамічних моделей споживання електроенергії .....                                                                  | 40        |
| <b>РОЗДІЛ 3. АНАЛІЗ ТА МОДЕЛЮВАННЯ ДИНАМІКИ СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ В УКРАЇНІ .....</b>                                                                    | <b>45</b> |
| 3.1 Аналіз результатів візуалізації та препроцесингу даних про споживання електроенергії в Україні .....                                                     | 45        |
| 3.2 Аналіз результатів моделювання споживання електроенергії на базі моделей експоненціального згладжування, ARIMA - моделей та нейромережевих моделей ..... | 58        |
| 3.3 Побудова та аналіз прогнозів рівня споживання електроенергії в Україні .....                                                                             | 61        |
| <b>ВИСНОВКИ .....</b>                                                                                                                                        | <b>64</b> |
| <b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....</b>                                                                                                                      | <b>66</b> |
| <b>ДОДАТОК А .....</b>                                                                                                                                       | <b>71</b> |

## ВСТУП

Задача прогнозування споживання електроенергії набуває особливої ваги в контексті планування роботи енергетичних систем, розподілу навантаження та ефективного управління енергетичними процесами.

Невірні прогнози у цій сфері можуть призвести до значних економічних втрат. Це пояснюється тим, що помилки в оцінці майбутнього споживання енергії викликають різноманітні проблеми. Якщо прогнози завищені, це веде до зайвих витрат палива всіх видів, що, в свою чергу, збільшує фінансові витрати на виробництво енергії.

З іншого боку, недооцінка потреби у споживанні енергії може спричинити дефіцит енергоресурсів, що погіршує якість енергопостачання і може призвести до перебоїв у подачі електрики споживачам. Тому підвищення точності прогнозування попиту на електроенергію є надзвичайно важливою та актуальною задачею, яка вимагає застосування сучасних та ефективних методів для її вирішення[1].

Чимало науковців займаються прогнозуванням попиту на електричну енергію. Серед таких досліджень можна відзначити роботи Егельйогу, Мохамеда і Гевена, Фунга і Туммали, Харріса і Лю, Лю, а також Лахані і Бумба. У їхніх працях використовуються різноманітні методи, включаючи регресійний аналіз і моделі на основі економічних показників. В деяких роботах, таких як у Макрідакіса і Вілрайта та Макрідакіса, Вілрайта і Хіндмана, представлені більш складні економіко-математичні моделі, що комбінують різні набори даних для покращення точності прогнозів.

Мета дослідження полягає у прогнозуванні споживання електроенергії в Україні з використанням методів прогнозування часових рядів, зокрема методу експоненціального згладжування Holt-Winters, моделей ARIMA/SARIMA, а також LSTM. Визначення найефективнішого для підвищення точності прогнозу.

Для досягнення поставленої мети необхідно вирішити наступні завдання:

1. Проаналізувати структуру та основні тенденції розвитку електроенергетичного ринку України.

2. Обґрунтувати вибір інструментарію математичного моделювання для дослідження динаміки споживання електроенергії в Україні.
  3. Побудувати систему моделей для аналізу та прогнозування споживання електроенергії на базі моделей експоненціального згладжування, ARIMA - моделей та нейромережових моделей.
  4. Оцінити прогностичну якість побудованих моделей за системою метрик, здійснити підбір гіперпараметрів для налаштування моделей прогнозування споживання електроенергії.
  5. Вибрати найкращий модельний інструментарій та застосувати його для предиктивного моделювання обсягів споживання електроенергії в Україні
- Об'єктом дослідження є рівень споживання електроенергії в Україні.

Предметом дослідження є прогнозування рівня споживання електроенергії в Україні за допомогою методів та моделей прогнозування часових рядів.

У роботі використані наступні методи дослідження: метод експоненціального згладжування Holt-Winters та ARIMA/SARIMA, LSTM. Для визначення адекватності моделей використано: числові характеристики, метрики оцінки точності для визначення адекватності моделей.

Інформаційною базою роботи є праці вітчизняних та зарубіжних науковців, наукові статті, статистичні дані з спеціалізованих інтернет-джерел та науково-методична література.

Робота складається зі вступу, трьох розділів, висновків та списку використаних джерел.

## **РОЗДІЛ 1**

### **ТЕОРЕТИЧНІ ЗАСАДИ АНЛІЗУ ТА ПРОГНОЗУВАННЯ РОЗВИТКУ ЕЛЕКТРОЕНЕРГЕТИЧНОГО РИНКУ УКРАЇНИ**

#### **1.1 Структура та тенденції розвитку електроенергетичного ринку України**

Споживання електроенергії є одним із ключових показників економічного розвитку будь-якої країни, зокрема України. Воно визначає рівень індустріалізації, соціального розвитку та забезпечення комфортних умов життя населення. В умовах сучасного світу, що характеризується стрімкою урбанізацією та технологічним прогресом, попит на електроенергію продовжує зростати [8, 9].

Електроенергія є важливим ресурсом для функціонування всіх галузей економіки: від важкої промисловості та транспорту до аграрного сектору і сфери послуг. Її раціональне використання безпосередньо впливає на конкурентоспроможність продукції, екологічну ситуацію та енергетичну безпеку держави. Водночас структура споживання електроенергії є відображенням економічної активності країни, зокрема частки промисловості, рівня автоматизації, урбанізації та соціального добробуту [19].

Енергетичний сектор України є важливою частиною економіки країни. Генерація електроенергії базується на атомних електростанціях, використанні вугілля, мазуту, природного газу, а також біологічного палива. Окрім цього, в Україні активно розвиваються відновлювальні джерела енергії, зокрема сонячні електростанції, вітрові турбіни та гідроелектростанції.

Приблизно 450 тисяч осіб працюють в енергетичній галузі, що становить близько 3% від загальної чисельності населення. За рівнем споживання енергії Україна посідає 28-ме місце у світі. Крім того, українські енергетичні компанії активно займаються експортом [35, 36].

Значний вплив на динаміку споживання електроенергії мають сезонні коливання, зокрема зростання попиту в зимові місяці через опалення та освітлення, а також у літній період через використання систем кондиціонування. Крім того, глобальні тренди, такі як перехід до електротранспорту, впровадження розумних мереж (smart grids) та відновлюваних джерел енергії, створюють додаткові виклики для традиційної енергетичної системи.

В Україні структура споживання електроенергії є різноманітною, охоплюючи промисловість, житлово-комунальний сектор, транспорт, будівництво, аграрний сектор та побутове споживання (рисунок 1.1).

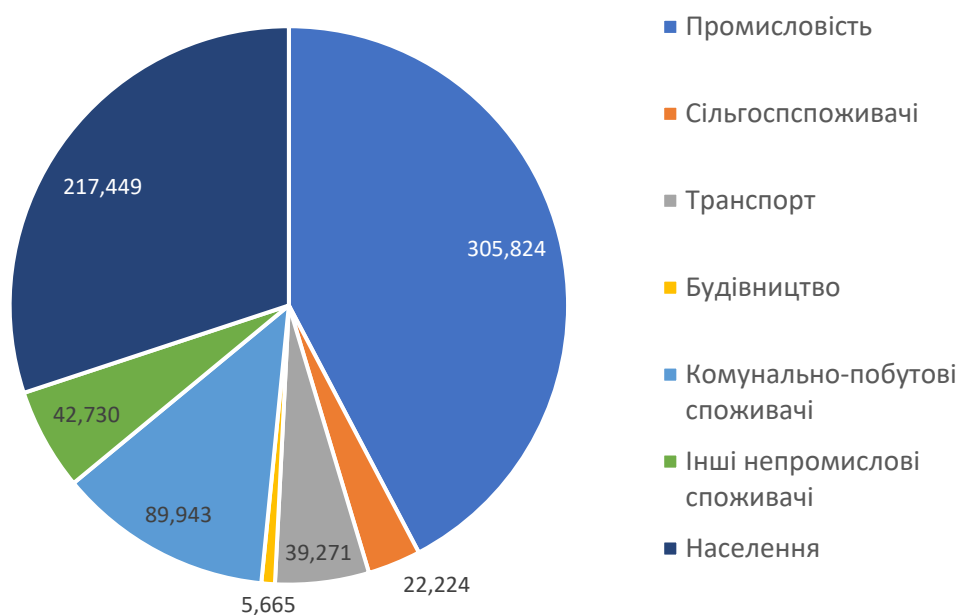


Рисунок 1.1 Споживання електроенергії за типом споживачів у млн кВт/год

*Джерело: розроблено автором на основі [37]*

Найбільшу частку займає промисловий сектор. Різні галузі промисловості та споживають значні обсяги електроенергії завдяки високій енергоємності виробничих процесів. Проте із впровадженням енергоефективних рішень споживання в цьому сегменті має тенденцію до зниження.

Другим за обсягами споживання є населення. У цьому сегменті спостерігається поступове зростання, зумовлене збільшенням кількості електроприладів у домогосподарствах та зростанням рівня урбанізації (рисунок 1.2).

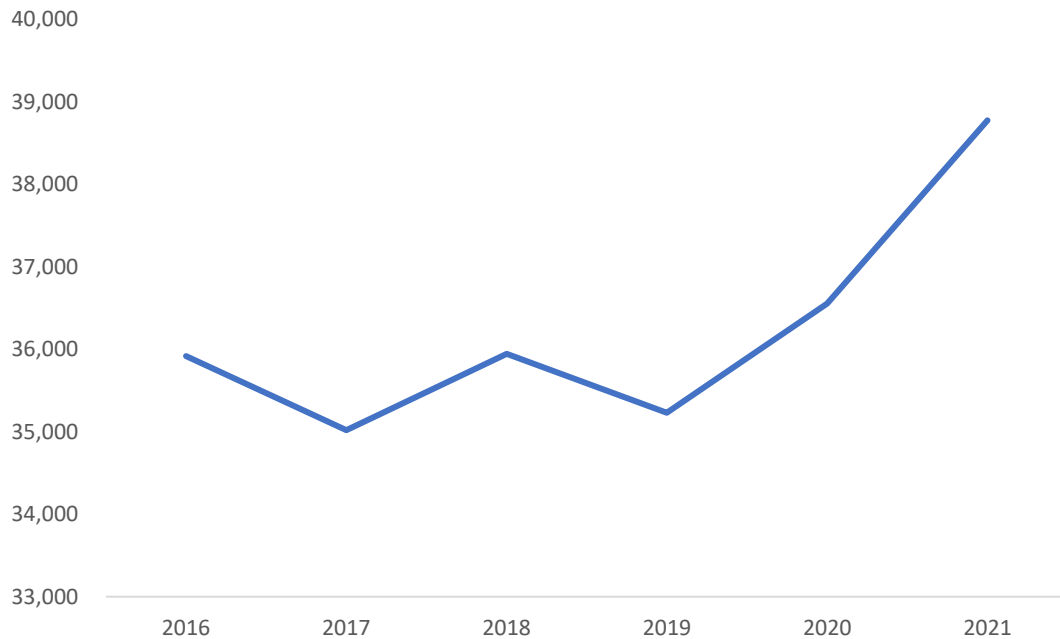


Рисунок 1.2 Споживання електроенергії населенням у млн кВт/год

*Джерело: розроблено автором на основі [37]*

Закриває трійку найбільших споживачів - комунально-побутові споживачі. Ця категорія споживачів включає установи та організації, що забезпечують життєдіяльність населення, а також побутові потреби громадян. До них належать: комунальні підприємства; соціальна інфраструктура (заклади освіти, охорони здоров'я, культури та спортивні об'єкти); громадське освітлення; малий бізнес та сфера послуг.

Транспортний сектор є одним із важливих, але порівняно невеликих споживачів електроенергії в Україні. Він охоплює різні види транспорту, які використовують електроенергію для своєї роботи. Основними компонентами транспортного сектору споживання є: електротранспорт у містах, залізничний

транспорт, інфраструктура транспорту. У великих містах, таких як Київ, Харків і Львів, ці види транспорту є основними споживачами електроенергії.

Іншими групами споживачі є сільськогосподарські споживачі, будівництво, і інші непромислові споживачі.

Категорія сільськогосподарські споживачі охоплює підприємства, що використовують електроенергію для виробництва сільськогосподарської продукції. Основні напрямки споживання включають: водопостачання та іригація, переробка сільгосппродукції, тваринництво, теплиці, обладнання для фермерських господарств.

Сільгоспсектор залежить від сезонності — найбільше навантаження припадає на період збирання врожаю та переробки продукції.

Будівельна галузь є однією з динамічних категорій споживачів електроенергії, особливо в період активної забудови. Основні потреби включають: освітлення будівельних майданчиків, електрообладнання, тимчасові споруди та приміщення, робота спеціалізованої техніки.

У будівництві споживання залежить від масштабу проєктів та інфраструктурної активності.

Для України, яка має значний потенціал у виробництві електроенергії завдяки природним ресурсам, включно з атомними, тепловими, гідро- та альтернативними джерелами енергії, забезпечення збалансованого розвитку енергетичного сектору є стратегічним завданням. Проте актуальними залишаються проблеми застарілої інфраструктури, нерівномірного розподілу потужностей між регіонами та необхідності інтеграції енергетичної системи з європейською.

## **1.2 Напрями та особливості дослідження споживання електроенергії**

Споживання електроенергії є важливим аспектом функціонування сучасного суспільства, що охоплює різні сфери життя — від побутових потреб до

роботи промислових підприємств та транспорту. Обсяги споживання змінюються в залежності від таких факторів, як тип споживачів, сезонні коливання, економічна активність та кліматичні умови.

Промисловість споживає значні обсяги енергії, що забезпечує стабільну підтримку виробничих процесів, а також варіюється залежно від економічної ситуації. Потреби населення, в свою чергу, значною мірою залежать від сезонних факторів, таких як опалення взимку та охолодження влітку.

Найбільший попит на електроенергію створюють енергоємні галузі, зокрема металургія, паливна та хімічна промисловість, які потребують значних обсягів енергії для забезпечення стабільного виробництва.

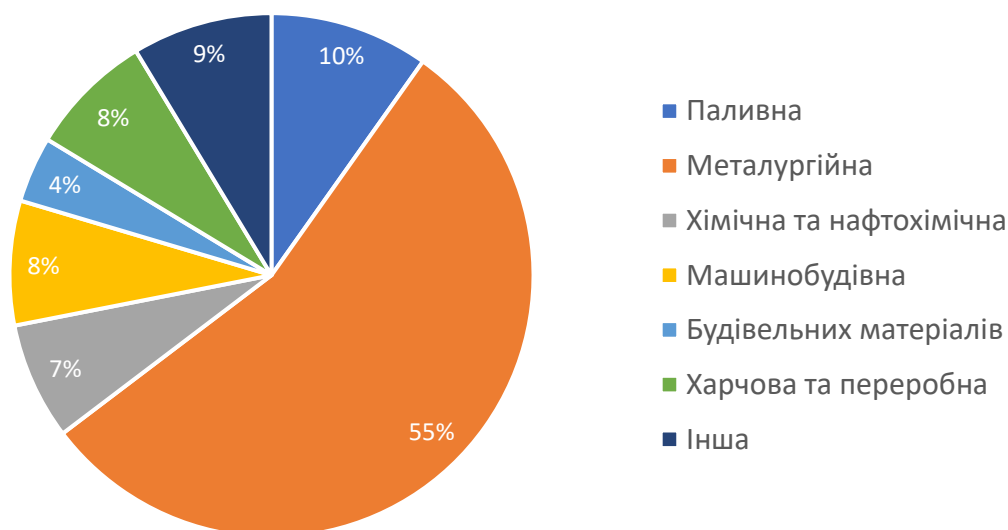


Рисунок 1.3 Споживання галузями промисловості, млн кВт\*год

*Джерело: розроблено автором на основі [37]*

На рисунку 1.3 відображено розподіл споживання електроенергії промисловістю України, де основну частку займає металургійна галузь – 55% загального споживання. Це пояснюється високою енергоємністю процесів, таких як виробництво сталі та обробка металів, які вимагають значних обсягів енергії.

Паливна промисловість, що включає видобуток і переробку енергетичних ресурсів (вугілля, газу, нафти), відповідає за 10% споживання. Хімічна та

нафтохімічна галузь, зокрема виробництво добрив і пластмас, споживає близько 9%, оскільки їхні виробничі процеси потребують стабільного енергозабезпечення.

Машинобудування (8%) та виробництво будівельних матеріалів (8%) також є значними споживачами. Перший сектор використовує енергію для механічної обробки й складання техніки, а другий – для виготовлення цементу, скла, цегли тощо. Харчова та переробна промисловість споживає 7% енергії, зокрема для підтримки температурних режимів і зберігання продукції.

Останні 4% припадають на інші галузі, що разом роблять незначний, але важливий внесок у загальне споживання. Такий розподіл підкреслює, що основними споживачами електроенергії в Україні є важка промисловість, зокрема металургія та паливний сектор. Проте обсяги їх споживання значною мірою залежать від сезонних змін, пов'язаних із кліматичними умовами, виробничими циклами та інших обставин. Зимовий та літній періоди мають суттєво різний вплив на обсяги енергоспоживання у різних галузях.

Зимова період характеризується збільшенням споживання електроенергії через необхідність обігріву виробничих приміщень і підтримання технологічних температур. Наприклад, металургійні підприємства, що використовують електродугові печі для плавлення сталі при температурах понад 1500 °С, потребують більше енергії в холодний період через підвищені втрати тепла. Аналогічно, хімічна промисловість, де процеси синтезу, як-от виробництво аміаку, вимагають нагрівання до 700 °С, взимку стикається з більшими енерговитратами.

Літній місяці, навпаки, часто супроводжується зниженням виробничої активності через планові відпустки працівників і зупинку частини підприємств, як це спостерігається в машинобудуванні. Водночас галузі, де потрібне охолодження, демонструють зростання енергоспоживання. Наприклад, харчова промисловість потребує значних обсягів енергії для роботи холодильних установок, які забезпечують зберігання продукції в умовах високих температур.

Таким чином, зимові місяці є періодом пікового навантаження через високі енерговитрати в енергоємних галузях, тоді як літній період поєднує в собі зменшення споживання в деяких секторах і зростання в інших, таких як

охолодження. Ці коливання потребують уважного планування енергозабезпечення для ефективного управління промисловим навантаженням.

Населення є значним споживачем електроенергії, і рівень споживання значною мірою залежить від кількох ключових чинників. Взимку споживання електроенергії зростає через потребу в опаленні, тоді як влітку спостерігається збільшення попиту через роботу кондиціонерів та інших охолоджувальних систем. Типи систем, які використовуються також визначають рівень споживання: електричні обігрівачі або кондиціонери споживають більше порівняно з іншими видами опалення. Отже, кліматичні умови, такі як тривалість зимових та літніх періодів, суттєво впливають на енергетичні навантаження, особливо в умовах крайніх температур.

Крім того, економічний добробут населення безпосередньо визначає обсяги споживання енергії. Вища купівельна спроможність дозволяє домогосподарствам більше інвестувати в побутову техніку та електроніку, що, у свою чергу, збільшує використання електроенергії. Також рівень урбанізації має значний вплив — міста, як правило, мають більшу щільність населення та високий рівень використання енергозалежних технологій. Вони потребують більше енергії для функціонування різноманітних систем, таких як освітлення, водопостачання та інших комунальних послуг.

Звички споживачів також відіграють важливу роль у формуванні рівня споживання. Активне використання електроприладів, особливо в години пік, може призводити до додаткового навантаження на енергосистему.

Таким чином, всі ці фактори взаємодіють між собою, утворюючи складну картину попиту на електричну енергію з боку населення.

Важливо зазначити, що енергоспоживання в комунально-побутовому секторі також має свої специфічні характеристики, оскільки цей сектор включає в себе широкий спектр установ, які забезпечують життєдіяльність міст, а також соціальну інфраструктуру.

Житлово-комунальні послуги, зокрема водопостачання та водовідведення, є стабільними споживачами електроенергії протягом року. Вони потребують

постійного енергопостачання для забезпечення життєдіяльності населення. Однак найвищі піки споживання зазвичай спостерігаються під час підготовки до зимового сезону, коли відбувається запуск систем опалення та підвищення навантаження на мережі.

Соціальна інфраструктура, включаючи лікарні, школи та адміністративні установи, споживає значну кількість енергії, зокрема для освітлення, опалення та забезпечення роботи різних систем, таких як ліфти та системи безпеки. Тут важливу роль відіграє оптимізація енергоспоживання, адже ці установи працюють безперервно, що вимагає стабільного енергопостачання.

Громадське освітлення, особливо в містах, також є одним з основних споживачів електричної енергії. Взимку, коли світловий день коротший, споживання енергії для освітлення вулиць та громадських місць зростає. Впровадження енергоефективних технологій, таких як світлодіодні лампи, здатне значно знизити витрати на електроенергію в цій сфері. Це не лише дозволяє знизити витрати, але й сприяє сталому розвитку міст.

Малий бізнес і сфера послуг, що включає магазини, кафе, ресторани та інші заклади, споживають електричну енергію для різноманітних потреб: освітлення, роботи побутових приладів та обігрівачів. Споживання електроенергії в цих секторах має виражену сезонність — влітку зростає використання кондиціонерів, а взимку — обігрівачів, що додатково навантажує енергетичні мережі.

Після розгляду споживання енергії комунально-побутовими споживачами важливо звернути увагу на ще один значний сектор, який також має великий вплив на енергетичний баланс країни — транспорт.

Протягом останніх років зростає обсяг споживання енергії в секторі транспорту через збільшення кількості перевезень, а також завдяки поступовій електрифікації деяких видів транспорту, таких як залізниця та міський електричний транспорт (трамваї, тролейбуси).

Найбільшим споживачем є залізничний транспорт. Залізничний сектор використовує електроенергію для руху потягів, функціонування інфраструктури (станції, сигнальні системи, освітлення) та для забезпечення комунікацій і

допоміжних систем. Велика частина енергоспоживання припадає на потяги дальнього сполучення, особливо в умовах високих обсягів перевезень вантажів, які є основною діяльністю залізниць. Залізничний транспорт є ефективнішим порівняно з іншими видами транспорту, але його енергетичне споживання залишається значним, особливо у пікові періоди перевезень.

Автомобільний транспорт також має свою частку в енергетичному балансі країни, з початком розвитку електричних автомобілів. Зростаючий інтерес до таких автомобілів стимулює необхідність модернізації інфраструктури для зарядки, що включає створення нових станцій для заряджання, що у свою чергу буде сприяти зростанню попиту на електричну енергію. Очікується, що споживання електричної енергії в цьому секторі поступово збільшуватиметься, особливо в містах.

Міський громадський транспорт, зокрема метро, трамваї, тролейбуси, і електрички, є значними споживачами електричної енергії. Їх споживання залежить від численних факторів, зокрема кількості маршрутів, їхньої протяжності, щільності попиту та технічного стану рухомого складу.

Загалом, однією з основних складових міського енергетичного споживання є метрополітен, особливо в великих містах. Наприклад, Київський метрополітен є одним з найбільших споживачів електроенергії серед міських транспортних систем, що безпосередньо пов'язано з великою кількістю перевезених пасажирів.

### **1.3 Базові підходи до математичного моделювання та прогнозування споживання електроенергії**

Прогнозування рівня споживання електроенергії є важливим компонентом енергетичного менеджменту, що забезпечує ефективне використання ресурсів, стабільність енергетичних систем і оптимізацію витрат. Оскільки умови попиту на енергію постійно змінюються, точне прогнозування допомагає уникати дефіциту чи надлишку енергії, що може призводити до значних фінансових втрат або аварій.

Це також сприяє плануванню розвитку енергетичної інфраструктури, включаючи модернізацію мереж, впровадження нових технологій та інтеграцію відновлюваних джерел енергії. Точні прогнози дозволяють не тільки оптимізувати виробництво та постачання електроенергії, але й знижують витрати на її транспортування, запобігаючи перевантаженням мереж і зменшуючи потребу у резервних потужностях. Крім того, на основі прогнозів можна більш ефективно планувати інвестиції в енергетичні проекти та забезпечити стійкість енергетичних систем у майбутньому [10].

Прогнозування поділяється на кілька видів залежно від тривалості періоду. Оперативні прогнози охоплюють періоди від кількох хвилин до кількох годин і використовуються для оперативного управління енергосистемами в реальному часі. Вони є надзвичайно важливими для підтримки балансу між попитом і пропозицією енергії в умовах швидких змін споживання, таких як зміни в погодних умовах або аварії в енергетичних мережах [6].

Неоперативні прогнози включають короткострокове, середньострокове, довгострокове та перспективне прогнозування. Короткострокове прогнозування (кілька годин або днів) враховує святкові періоди й інші тимчасові чинники, що можуть впливати на рівень попиту, а також на можливу зміну енергоспоживання в залежності від соціальних чи економічних подій. Середньострокове прогнозування (кілька днів чи місяців) охоплює сезонні коливання та соціально-економічні зміни, такі як зміни в промисловому виробництві, споживчих звичках або вплив нових технологій на споживання енергії. Довгострокове прогнозування (кілька років чи десятиліть) орієнтоване на інвестиційне планування та інфраструктурні проекти, а також на передбачення змін у структурі енергетичних потреб населення, промисловості та транспорту. Це включає також оцінку впливу відновлювальних джерел енергії, та змін в регулюванні енергетичних ринків. Перспективне прогнозування враховує інновації, розвиток технологій і потенційні непередбачувані зміни, забезпечуючи гнучкість і підготовку до майбутніх викликів, таких як зміни в кліматі, глобальні економічні тенденції та вплив нових енергетичних технологій [11].

Наївні методи прогнозування — це прості, але дієві способи створення базових моделей для прогнозування часових рядів. Вони не вимагають складних алгоритмів чи значних обчислювальних ресурсів, а базуються на простих статистичних або евристичних підходах. Використання таких моделей є важливим етапом не тільки, тому що вони забезпечують базове розуміння структури даних, а й для оцінки ефективності складніших моделей прогнозування, оскільки вони забезпечують точку відліку для порівняння.

Базові моделі використовують історичну інформацію для передбачення майбутнього. Наприклад, один із найпоширеніших методів — це використання середнього значення. В цьому випадку прогноз формується на основі припущення, що середній показник за минулий період залишиться стабільним у майбутньому. Такий підхід є ефективним для стаціонарних рядів, де значення не демонструють значних змін у часі.

$$\hat{y}_t = \frac{1}{n} \sum_{i=1}^n y_i \quad (1.1)$$

Інший популярний підхід базується на використанні останнього відомого значення ряду як прогнозу для наступного періоду. Цей метод передбачає, що значення в найближчому майбутньому будуть такими ж, як і останнє спостереження. Цей підхід особливо ефективний для короткострокових прогнозів, де дані демонструють стабільність або коли зміни є мінімальними. Однак цей метод може бути недієвим у випадках різких змін або при наявності трендів.

$$\hat{y}_{t+1} = y_t \quad (1.2)$$

Сезонні ряди потребують більш специфічного підходу. У цьому випадку прогноз ґрунтується на повторюваних паттернах, які спостерігалися в минулому. Наприклад, якщо дані демонструють сезонний цикл з періодом у рік, значення наступного року прогнозується на основі відповідних періодів попереднього року. Це дозволяє ефективно враховувати коливання, пов'язані з сезонністю.

$$\hat{y}_{t+m} = y_{t+m-k}, \quad (1.2)$$

де  $k$  — довжина сезону.

Метрики, такі як середня абсолютна похибка (MAE) або середня абсолютна відсоткова похибка (MAPE), використовуються для оцінки точності прогнозів, що дозволяє визначати ефективність методів у кількісному вираженні. Реалізація наївних методів є простою і може бути виконана за допомогою сучасних інструментів аналізу даних.

Хоча наївні методи є ефективними для початкового аналізу, вони мають обмеження. Вони не враховують складних трендів, змін у сезонності чи впливу зовнішніх факторів. Крім того, методи чутливі до шуму, який може спотворювати результати. Однак вони залишаються важливим інструментом для оцінки складніших моделей: якщо більш складний алгоритм не перевершує наївний прогноз, це свідчить про його неефективність або про необхідність перегляду моделі.

Наївні методи також добре підходять для випадків, коли дані є обмеженими або потрібен швидкий аналіз. Вони часто використовуються в бізнесі для оперативного прийняття рішень, де простота і швидкість мають вирішальне значення. Окрім цього, вони допомагають виявляти ключові особливості часового ряду, такі як тренди, сезонність або стабільність. Наприклад, систематичні похибки у прогнозах середнього значення можуть вказувати на наявність тренду, тоді як розбіжності у сезонних прогнозах можуть свідчити про зміну циклів.

Прогнозування, що базується на випадковому блуканні, є важливим доповненням до наївних методів прогнозування, особливо в умовах високої невизначеності, коли зміни в даних є непередбачуваними і залежать від випадкових факторів.

Випадкове блукання передбачає, що майбутнє значення є лише випадковим кроком від поточного, що може бути корисним для прогнозування ситуацій, де немає чітко вираженого тренду чи сезонності, а динаміка змін є випадковою.

$$y_t = y_{t-1} + \epsilon_t, \quad (1.2)$$

де  $y_t$  - значення в момент часу  $t$ ;

$y_{t-1}$  - значення в попередній момент часу;

$\epsilon_t$  - випадкова похибка, яка зазвичай має нормальний розподіл з нульовим середнім і певним стандартним відхиленням.

Основною рисою випадкового блуждання є відсутність передбачуваності майбутніх значень за допомогою простих статистичних моделей. Це означає, що для часових рядів, які демонструють випадкове блуждання, наївні методи можуть бути неефективними, оскільки дані не мають стабільної тенденції для побудови точних прогнозів.

Для перевірки наявності випадкового блуждання в часових рядах зазвичай застосовують тести на стаціонарність, оскільки такий процес є нестабільним, тобто його статистичні властивості, такі як середнє і дисперсія, змінюються з часом. Використовують також функції автокореляції, щоб виявити можливі циклічні зміни або кореляцію між попередніми і поточними значеннями.

Хоча цей процес не дає чітких передбачень для довгострокового прогнозу, він допомагає зрозуміти природу даних, де зміни визначаються випадковими факторами, і також слугує орієнтиром для оцінки складніших моделей.

Після застосування наївних методів наступним етапом є використання базових інструментів (рисунок 1.4) для більш детального аналізу та прогнозування часових рядів. До них відносяться статистичні методи, які дозволяють розкласти часовий ряд на компоненти, такі як тренд, сезонність та випадкові коливання. Одним із таких інструментів є моделі ARIMA, які дозволяють прогнозувати за допомогою комбінованих компонентів авторегресії та ковзного середнього. Ці методи значно покращують точність прогнозів порівняно з наївними підходами, адже враховують більш складні залежності у даних.

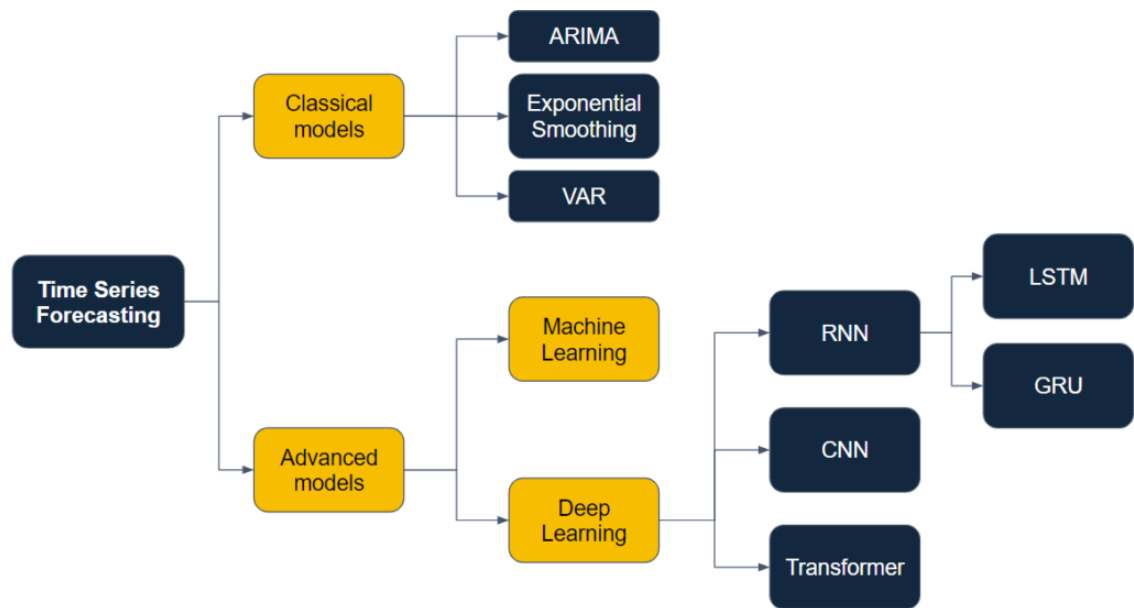


Рисунок 1.4 Базові інструменти аналізу часових рядів та прогнозування [39]

У даний час існує значна кількість наукових досліджень, присвячених моделюванню динамічних процесів на електроенергетичному ринку України. Розроблено низку ефективних моделей для прогнозування споживання електроенергії. Розробляються способи підбору та налаштування гіперпараметрів прогностичних моделей для збільшення точності прогнозування, використовується широка система метрик для оцінювання якості моделей, враховується велика кількість факторів, розробляються нові підходи та методи, вдосконалюються алгоритми розрахунків.

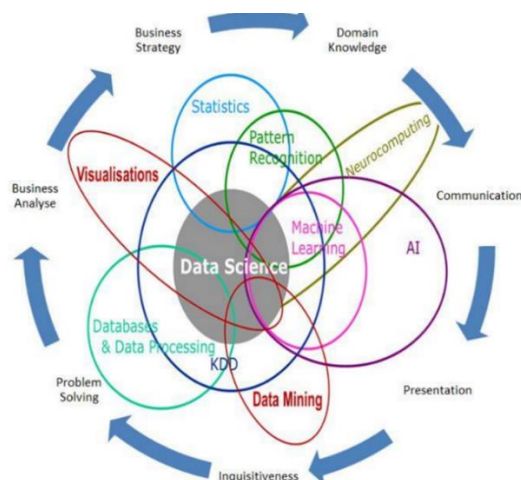


Рисунок 1.5 Складові та базові інструменти науки про дані Data Science [40]

З розвитком комп'ютерних технологій для прогнозування обсягів споживання електроенергії широко використовується інструментарій науки про дані Data Science (Рис. 1.5) та машинне навчання Machine Learning (Рис. 1.6).

Алгоритми машинного навчання мають здатність до самонавчання та можуть швидко знаходити оптимальні рішення поставлених завдань, що робить їх придатними для вирішення складних нелінійних задач прогнозування динаміки змінення рівня споживання електроенергії. Ці алгоритми охоплюють широкий спектр підходів, серед яких найбільш популярними є методи глибокого навчання, дерева рішень, регресійні моделі, кластеризація та байєсові методи.



Рисункок 1.6 Базові алгоритми Machine Learning [41]

На рисунку 1.6 представлено ієрархію алгоритмів машинного навчання, що наочно показує їх класифікацію та можливості застосування.

Інтеграція машинного навчання з класичними методами прогнозування дозволяє значно підвищити точність передбачень і більш ефективно враховувати

зміни в обсягах споживання. Використання таких технологій дає змогу створювати адаптивні моделі, здатні реагувати на швидкі зміни, та оптимізувати обробку і аналіз великих обсягів даних, які зростають з часом. Це дозволяє більш ефективно вирішувати завдання, які традиційні методи не можуть виконати так швидко або точно [38].

## РОЗДІЛ 2 МАТЕМАТИЧНІ МОДЕЛІ АНАЛІЗУ ДИНАМІКИ СПОЖИВАННЯ ЕКТРОЕНЕРГІЇ В УКРАЇНІ

### 2.1 Інструментарій розвідувального аналізу одномірних часових рядів обсягів споживання електроенергії

Економічні процеси зазвичай змінюються протягом певного періоду часу, і цей процес отримав назву економічної динаміки. Основною рисою економічної динаміки є те, що показники в майбутньому багато в чому залежать від їхніх значень у попередні періоди. Крім того, збільшення інтервалу часу між двома подіями може спричиняти значні зміни як в кількісних, так і в якісних характеристиках цих явищ. Сучасні методи аналізу часто базуються на виявленні моделей нелінійної поведінки часового ряду [21, 26].

Терміни "ряд динаміки" та "динамічний ряд" застосовуються для позначення часового ряду, який у англійській літературі відомий як "time series". Кожен елемент часового ряду, або його рівень, відповідає певному моменту часу. Рівні можуть мати як детерміновані, так і випадкові значення. Важливим аспектом часового ряду є послідовність його елементів, що не повинна змінюватися довільно. Іноді на один момент часу може припадати кілька значень різних показників, що призводить до утворення багатовимірного часового ряду [14].

Під поняттям «часовий ряд (time series)» розуміють набір значень певної змінної, що були зафіксовані через рівні проміжки часу. Якщо взяти за одиницю часу, наприклад, рік, квартал чи день, то спостереження  $x_1, \dots, x_n$  будуть відображати значення цієї змінної на момент  $t = 1, \dots, n$  [15, 16].

Часовий ряд можна описати як послідовність точок даних, які відображають події, що відбулися за певний рівний проміжок часу, де час виступає як незалежна змінна, а метою є прогнозування майбутніх тенденцій. Ці дані формують структуру, де кожен елемент містить числове значення та відповідну тимчасову позначку.

Часові ряди можуть бути застосовані до будь-якої змінної, що змінюється з часом. [17, 18].

Введення тимчасового компонента в часовий ряд є важливим моментом, що відрізняє його від звичайної (випадкової) вибірки статистичних даних. Основною особливістю часового ряду є те, що значення чітко пов'язані з конкретними моментами часу.

Дані часових рядів представляють собою набір вимірювань, отриманих на певних етапах часу. Якщо ці дані відобразити на графіку, час завжди буде однією з осей. Така інформація присутня скрізь, оскільки час є невід'ємною частиною всього, що підлягає спостереженню. З розвитком технологій і збільшенням використання датчиків та приладів, потік даних часових рядів став постійним і незмінним, оскільки ці системи активно генерують нові дані [20].

Дані часового ряду можна класифікувати на дві основні категорії: інтервальний часовий ряд — це послідовність значень, де кожен показник відображає зміну стану досліджуваного об'єкта протягом певного періоду часу, наприклад, року, кварталу, місяця або доби; моментний часовий ряд — це послідовність значень, що характеризують стан об'єкта на конкретний момент часу, наприклад, на початок кожного місяця або на перше вересня кожного року [5].

Аналіз часового ряду проводиться в багатьох різних контекстах, щоб вивчити, як змінюється певний показник протягом часу та здійснити прогноз його майбутніх значень. Для ефективного аналізу необхідно мати значну кількість даних, оскільки це забезпечує кілька важливих аспектів: послідовність і надійність (велика кількість точок даних дозволяє забезпечити стабільність аналізу та виявити тренди, сезонні коливання та інші патерни); репрезентативність вибірки (більший обсяг даних сприяє більш точному відображенню всієї популяції, що дозволяє робити більш надійні висновки про поведінку показника); фільтрація шуму (звдяки великій кількості даних можна ефективно відсіяти шуми та викиди, що дозволяє зосередитись на значущих закономірностях); врахування сезонності (великий набір даних дозволяє чітко виявити сезонні коливання і знайти стабільні залежності, що повторюються через певні проміжки часу) [7].

Аналіз часового ряду має дві основні мети: виявлення його структури або природи та прогнозування майбутніх значень, ґрунтуючись на поточних та попередніх вимірюваннях [17, 18].

Обидва завдання є взаємозалежними. Розв'язання першого завдання є необхідним для створення математичної моделі часового ряду, її правильної ідентифікації та формалізації. Така модель виступатиме як інструмент для аналізу часового ряду та основою для здійснення прогнозів з прийнятною похибкою.

Часовий ряд складається з чотирьох основних компонентів: тренду, сезонності, циклічності та випадкових змін. Ці елементи відображають різні закономірності, що можна виявити в даних, що спостерігаються, і вони є ключовими для аналізу та прогнозування. Виявлення цих складових дозволяє не тільки зрозуміти, як змінюються дані з часом, але й відрізняє часові ряди від інших типів даних, що спостерігаються, завдяки їхній структурованості та чітким періодичним патернам. [20, 25].

Тренд (рисунок 2.1) відображає загальні зміни в часовому ряді протягом тривалого періоду, вказуючи на його довгострокову тенденцію. Це може бути зростаючий, спадний або стаціонарний тренд. Оцінка лінійного тренду є статистичним методом, що дозволяє інтерпретувати дані шляхом побудови моделі, яка описує поведінку змінних у часі. Така оцінка допомагає встановити зв'язок між вимірюваннями і часом їх проведення, створюючи математичну модель, яка надає змогу передбачити майбутні зміни. Лінійна функція, яку отримують для опису тренду, дозволяє не лише зрозуміти, як змінюються дані, але й оцінити значущість відмінностей, що виникають у наборі даних під впливом різних факторів [22-24].

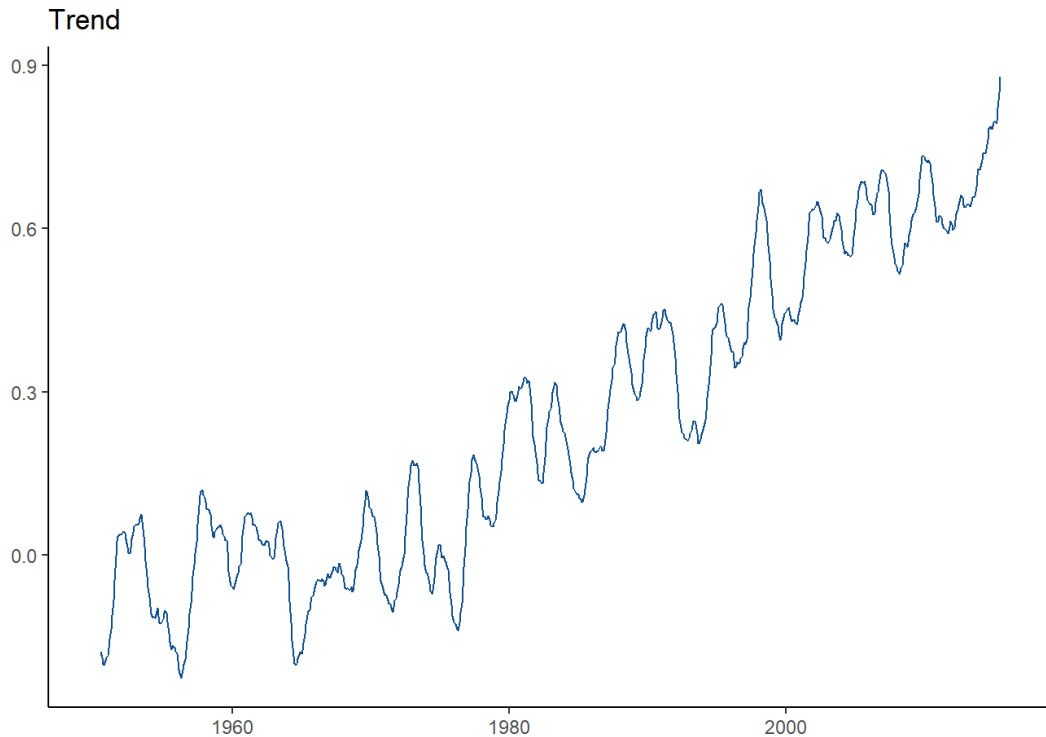


Рисунок 2.1 Приклад тренду [42]

Тренд найчастіше моделюється за допомогою лінійної регресії:

$$y_t = a + b \cdot t + \varepsilon_t \quad (2.1)$$

де  $y_t$  - значення часового ряду в момент часу  $t$ ;

$a$  - початкове значення (вільний член);

$b$  - коефіцієнт регресії, що показує нахил тренду;

$t$  - час;

$\varepsilon_t$  - залишкова компонента, що відображає випадкові відхилення.

Ця модель дозволяє не лише виявити основну тенденцію, але й прогнозувати майбутні значення ряду. Наприклад, для наступного періоду прогноз обчислюється як:

$$y_{t+1} = a + b \cdot (t + 1) \quad (2.2)$$

Якщо зміна даних має складніший характер, тренд може бути описаний нелінійною моделлю, наприклад, квадратичною:

$$y_t = a + b \cdot t + c \cdot t^2 + \varepsilon_t \quad (2.3)$$

де  $c$  враховує квадратичну залежність значень від часу.

Аналіз тренду є важливим етапом у дослідженні часових рядів, адже він дає змогу зрозуміти загальні тенденції, які відображають зміни показників у часі. Він слугує основою для прогнозування майбутньої поведінки показника, дозволяючи будувати моделі, які враховують минулі дані для передбачення майбутніх значень. Крім того, трендовий аналіз дає змогу оцінити вплив зовнішніх факторів, таких як сезонні коливання чи циклічні процеси, які можуть накладатися на основну тенденцію. Виявлення тренду дозволяє їх відокремити й отримати чисту картину змін.

Розуміння сезонності має вирішальне значення, оскільки сезонні впливи можуть суттєво спотворювати тренди й інші компоненти.

Сезонність у часових рядах відображає повторювані циклічні зміни, які пов'язані з певними періодами, такими як місяці, квартали чи роки. Вона проявляється як регулярні та прогнозовані закономірності, що виникають у даних через постійні інтервали часу менше року, наприклад, щотижня, щомісяця або щокварталу. Причинами сезонності можуть бути погодні умови, святкові періоди чи інші події, що регулярно впливають на показники.

Для цього використовуються дві основні моделі: адитивна, яка підходить для рядів із постійною амплітудою, та мультиплікативна, яка відображає залежність масштабів коливань від загального рівня даних. Кожна з цих моделей має власний підхід до розрахунку сезонного компонента, що дозволяє обрати найбільш відповідний метод для конкретного типу даних.

У адитивній моделі часових рядів сезонний компонент визначається як різниця між фактичними значеннями часового ряду  $Y$  та сумою інших компонентів — тренду ( $T$ ), циклічного компонента ( $C$ ) і нерегулярних змін ( $I$ ). Формула виглядає наступним чином:

$$S = Y - (T + C + I) \quad (2.4)$$

де  $S$  - сезонні значення;

$Y$  - фактичні дані;

$T$  - трендові значення;

$C$  - циклічні компоненти;

$I$  - нерегулярні (залишкові) компоненти.

У мультиплікативній моделі часових рядів сезонний компонент розраховується як частка фактичних значень і добутку інших компонентів, що виражається формулою:

$$S = \frac{Y}{T \cdot C \cdot I} \quad (2.5)$$

Ця модель передбачає, що компоненти взаємодіють між собою, створюючи множинний ефект. Мультиплікативний підхід використовується, коли вплив сезонності або інших компонентів змінюється залежно від масштабу часового ряду. Адитивна модель підходить для рядів із постійною амплітудою змін (сезонні коливання не залежать від загального рівня ряду). Мультиплікативна модель більш актуальна, коли масштаб сезонності змінюється разом із трендом (наприклад, відносне зростання або спад).

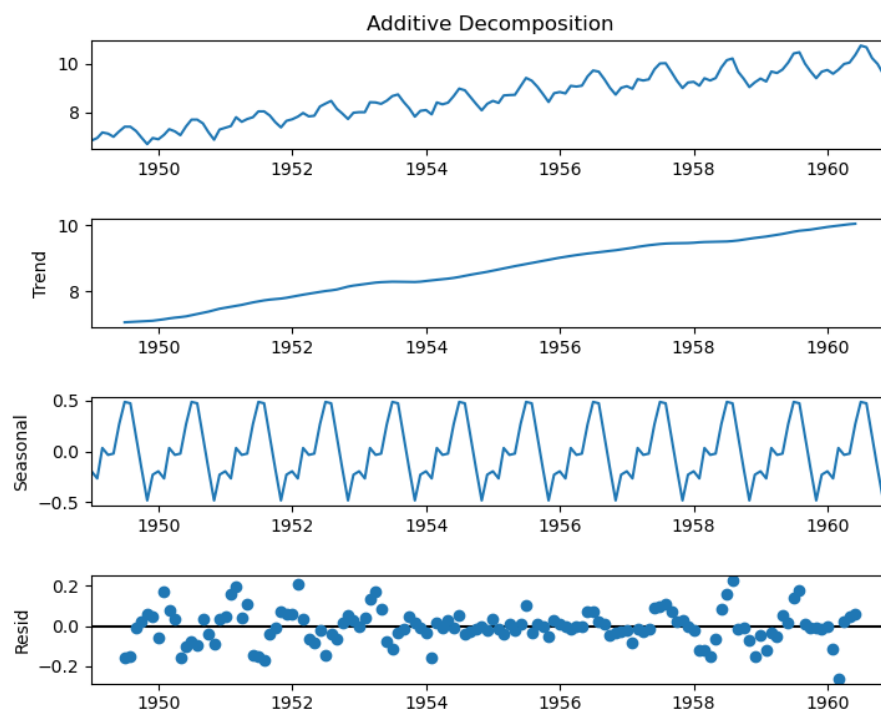


Рисунок 2.2 Приклад декомпозиції часового ряду [43]

Циклічність є одним із ключових компонентів часових рядів, який відображає тривалі зміни, пов'язані з фундаментальними факторами, такими як економічні чи бізнес-цикли. На відміну від сезонних змін, циклічність не має чітко визначеного періоду повторюваності, а її тривалість може коливатися від кількох років до десятиліть. Ці закономірності зазвичай обумовлені змінами в економічному середовищі, наприклад, діловими циклами, які включають етапи зростання, піків, спаду та відновлення.

Циклічні коливання характеризуються нерегулярністю та залежать від багатьох змінних, зокрема макроекономічних умов, політичних рішень чи природних явищ. Вони мають значення для аналізу довгострокових трендів, оскільки допомагають зрозуміти вплив економічного середовища на зміну показників і спрогнозувати можливі наслідки. Таким чином, циклічність відіграє важливу роль у моделюванні часових рядів, доповнюючи загальну картину їхньої поведінки.

Випадкові зміни у часових рядах являють собою флуктуації, які неможливо передбачити і які не сліднують систематичним закономірностям. Такі зміни часто викликаються випадковими факторами, похибками вимірювання або іншими невідомими впливами. Вони відрізняються від циклічних чи сезонних змін тим, що не мають регулярного повторення та не піддаються прогнозуванню.

Нерегулярні часові ряди відображають дані, які збираються в довільні моменти часу, наприклад, під час реєстрації кожної транзакції на фондовому ринку або фіксації подій, таких як зміни в рівні напруги чи сигнали про низький заряд батареї. У таких рядах інтервали часу можуть варіюватися, і це не дозволяє застосовувати стандартні часові зсуви для аналізу. Для таких даних важливо зберігати точний час кожного вимірювання, оскільки інтервали між спостереженнями можуть бути різними. Це дозволяє ефективно обробляти дані, що надходять із різних джерел, таких як герцові вимірювання або цифрові потоки, де неможливо ввести нульові значення.

Аналіз основних компонентів часового ряду дозволяє виявити його структуру та особливості, з'ясувати, як ці компоненти взаємодіють, і використовувати цю інформацію для точного прогнозування та прийняття обґрунтованих рішень. Це дає змогу зрозуміти динаміку ряду, а також глибше проникнути в процеси, що його формують, розпізнати ключові залежності та закономірності. За допомогою великої кількості даних можна підвищити надійність аналізу, виявити складні патерни та мінімізувати вплив випадкових коливань, що сприяє точнішому прогнозуванню майбутніх значень. Вивчення цих компонентів є критично важливим для кращого розуміння поведінки даних і для прогнозування на основі минулого досвіду.

На сьогоднішній день виділяють два принципово різних підходи до аналізу часових рядів: статистичний підхід (мета якого — виявлення закономірностей, таких як тренди, сезонність, циклічність, та побудова моделей, що пояснюють поведінку ряду), динамічний підхід (базується на дослідженні системи, що генерує часовий ряд, акцент робиться на моделюванні фізичних, економічних або інших процесів, що лежать в основі даних).

Обидва підходи часто доповнюють один одного, дозволяючи отримати глибші інсайти: статистичний аналіз забезпечує кількісну оцінку характеристик ряду, тоді як динамічний підхід дозволяє зрозуміти його причинно-наслідкову структуру. Залежно від мети дослідження, можна обирати один із підходів або комбінувати їх для досягнення більш точних і повних результатів.

Аналіз часових рядів є корисним інструментом для вивчення змін у даних, оскільки має кілька суттєвих переваг. По-перше, це дає змогу усунути викиди та зменшити шум у даних, що допомагає краще зрозуміти реальні закономірності й особливості змін. По-друге, завдяки застосуванню різноманітних моделей стає можливим глибше усвідомлення характеру змін у даних та оцінка їх значущості в реальному контексті. А також, цей підхід допомагає визначати тренди, сезонність, циклічні коливання та інші регулярні патерни, що у свою чергу дозволяє будувати прогнози, на основі яких можна оцінювати майбутні тенденції, такі як попит, продажі або поведінка клієнтів, що є корисними для стратегічного планування та прийняття рішень.

З іншого боку існує цілий ряд обмежень. По-перше, цей метод більше підходить для короткострокових прогнозів, адже з часом точність прогнозів знижується через вплив зовнішніх факторів, які важко врахувати. Крім того, для складних проектів або ситуацій, де результат залежить від численних змінних, аналіз часових рядів може виявитися недостатнім, оскільки він не завжди враховує всі зовнішні фактори. Важливо також зазначити, що точність аналізу значною мірою залежить від якості вхідних даних, тому неповні чи спотворені дані можуть серйозно вплинути на результат. Для виявлення стабільних закономірностей і трендів потрібно мати достатньо даних, і це може бути проблемою, якщо необхідної кількості точок немає або їх важко зібрати. І, звісно, навіть незначний шум чи викиди в даних можуть спотворити результати, тому важливо проводити ретельне очищення даних перед аналізом.

Отже, часовий ряд — це сукупність даних, які відображають динаміку певної змінної у часі. Він дозволяє проаналізувати, які чинники впливають на показник з періоду в період, і зрозуміти закономірності змін.

Попри свою ефективність для виявлення закономірностей і прогнозування, успішність такого аналізу значною мірою залежить від якості та обсягу даних, а також від поставлених цілей. Ретельно проведений аналіз допомагає глибше зрозуміти динаміку змін і приймати обґрунтовані рішення на основі історичних даних [26, 30].

## **2.2 Методи експоненціального згладжування, авторегресійні та нейромережеві моделі аналізу динаміки споживання електроенергії**

Одними з найпоширеніших підходів до моделювання часових рядів є авторегресійні моделі, які використовують математичний апарат для врахування залежностей між попередніми та поточними значеннями ряду. Зокрема, модель ARIMA (Autoregressive Integrated Moving Average) є універсальним інструментом,

який поєднує компоненти авторегресії, інтеграції та ковзного середнього. Завдяки цим елементам, ARIMA може ефективно працювати як із стаціонарними, так і з нестаціонарними рядами, трансформуючи останні у стаціонарну форму за допомогою диференціювання.

Однак багато процесів у реальному житті характеризуються чітко вираженими сезонними коливаннями, які базова модель ARIMA не враховує. Для розв'язання цієї проблеми було розроблено розширення SARIMA (Seasonal ARIMA), яке додає можливість моделювання сезонних патернів. SARIMA дозволяє ефективно аналізувати циклічні явища.

Авторегресійні моделі ARIMA та SARIMA відіграють важливу роль у сучасній аналітиці даних, адже вони дозволяють отримувати інформативні прогнози навіть для складних часових рядів.

Моделі ARIMA часто застосовуються для аналізу даних, які демонструють нестаціонарність. На початковому етапі для досягнення стаціонарності проводиться диференціювання — процес, що відповідає "інтегрованому" компоненту моделі. Цю операцію можна виконати один або кілька разів залежно від ступеня нестаціонарності даних. Будь-який часовий ряд, який немає сезонності, можна прогнозувати з використанням ARIMA-моделі [27-30].

Несезонні моделі ARIMA зазвичай позначаються як ARIMA ( $p$ ,  $d$ ,  $q$ ), де параметри  $p$ ,  $d$ ,  $q$  - представляють невід'ємні цілі числа. Зокрема:  $p$  - визначає порядок авторегресивного компонента (кількість часових лагів);  $d$  - відповідає кількості етапів диференціювання, необхідних для усунення нестаціонарності;  $q$  - позначає порядок моделі ковзного середнього [1];

Математичне представлення моделі виглядає наступним чином:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q}, \quad (2.6)$$

де  $y_t$  - спостереження в момент часу  $t$ ;

$\phi_1, \phi_2, \dots, \phi_p$  - параметри авторегресії (AR);

$\epsilon_t$  - випадкова помилка в момент часу  $t$ ;

$\theta_1, \theta_2, \dots, \theta_q$  - параметри ковзного середнього (MA);

$p$  - порядок авторегресії;  $q$  - порядок ковзного середнього.

Хоча модель ARIMA є ефективним інструментом для аналізу часових рядів, її застосування обмежується випадками, коли в даних відсутні чіткі сезонні компоненти. Однак багато реальних процесів мають яскраво виражену сезонність, наприклад періодичність споживання енергії. У таких випадках базова ARIMA не здатна належним чином моделювати сезонні патерни, що призводить до втрати точності прогнозування [31-33].

Для розширення можливостей ARIMA була розроблена модель SARIMA (Seasonal ARIMA), яка додає до базової структури механізми обробки сезонних даних. SARIMA враховує не лише короткострокові залежності, а й циклічні зміни через включення сезонних компонентів авторегресії, диференціювання та ковзного середнього. Це дозволяє моделі більш точно описувати процеси, де повторювані закономірності відіграють ключову роль.

Переходячи до опису SARIMA, важливо підкреслити її здатність комбінувати сильні сторони ARIMA з адаптивністю до сезонних коливань.

Сезонні моделі ARIMA мають формат:

$$ARIMA(p, d, q) (P, D, Q), \quad (2.7)$$

*де  $m$  - вказує на тривалість одного сезону;*

*а параметри:  $P$  – порядок авторегресії;*

*$D$  – диференціювання;*

*$Q$  – ковзного середнього для сезонної частини.*

Зовнішні змінні можуть бути інтегровані в модель через використання екзогенних регресорів. Сезонна ARIMA з екзогенними регресорами - SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors) дає можливість враховувати вплив зовнішніх факторів на модель, дозволяючи точніше прогнозувати за рахунок додавання додаткової інформації про змінні, що не входять безпосередньо в часовий ряд.

Модель SARIMAX записується як:

$$SARIMAX(p, d, q) (P, D, Q) \quad (2.8)$$

де  $p$  - порядок авторегресії (AR), тобто кількість попередніх значень, що використовуються для прогнозування;

$d$  - ступінь диференціювання, що визначає, скільки разів потрібно диференціювати часовий ряд для досягнення стаціонарності;

$q$  - порядок ковзного середнього (MA), що визначає кількість лагів помилок, які використовуються для побудови прогнозу;

$P$  - сезонний порядок авторегресії (AR), який враховує сезонні залежності в даних;

$D$  - сезонний порядок диференціювання, який вказує, скільки разів потрібно застосовувати диференціювання до сезонної складової;

$Q$  - сезонний порядок ковзного середнього (MA), що визначає, скільки попередніх сезонних помилок слід враховувати;

$s$  - кількість періодів у сезоні (наприклад,  $s = 12$  для даних з річною сезонністю при місячних вимірюваннях).

Математично модель виглядає так:

$$y_t = \beta_0 + \sum_{i=1}^k \beta_i X_{i,t} + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^Q \Theta_j \epsilon_{t-sj} + \sum_{i=1}^P \Phi_i y_{t-si}, \quad (2.9)$$

де  $y_t$  - значення часового ряду в момент часу  $t$ ;

$X_{i,t}$  - екзогенні змінні (зовнішні фактори, що впливають на часового ряду);

$\beta_0$  - константа, яка може бути інтерпретована як базовий рівень;

$\epsilon_t$  - залишковий термін або помилка на час  $t$ ;

$\phi_i$  - коефіцієнти авторегресії (AR) для лінійної залежності від попередніх значень ряду;

$\theta_j$  - коефіцієнти ковзного середнього (MA) для відображення шоків у часі;

$\Phi_i$  та  $\Theta_j$  - сезонні коефіцієнти;

$P$  та  $Q$  - порядки сезонної авторегресії та ковзного середнього.

Моделі ARIMA, SARIMA та SARIMAX — це потужні інструменти для прогнозування часових рядів, кожен з яких має свої особливості. ARIMA ідеально підходить для нестационарних рядів без сезонних коливань. Вона досить проста в реалізації, однак її головний недолік — неможливість обробляти сезонність. Крім того, для її коректного використання дані повинні бути стаціонарними або потребувати попередньої диференціації.

SARIMA є більш просунутим інструментом, оскільки дозволяє враховувати сезонність, що робить її підходящою для прогнозування даних з явними сезонними коливаннями. Однак налаштування цієї моделі може бути складнішим, а її чутливість до шуму або слабкої сезонної складової обмежує її ефективність у деяких випадках.

SARIMAX, на відміну від попередніх моделей, включає екзогенні змінні, тобто зовнішні фактори, які можуть впливати на часову серію. Це робить модель значно гнучкішою, дозволяючи врахувати додаткові впливи, але її ефективність напряду залежить від якості екзогенних даних та складності налаштувань.

Таким чином, вибір моделі залежить від конкретних особливостей даних: ARIMA підійде для простих рядів без сезонності, SARIMA — для сезонних серій, а SARIMAX стане кращим вибором, коли потрібно врахувати зовнішні фактори.

Іншим методом для побудови прогнозів є використання адаптивних підходів, серед яких значну популярність мають експоненціальні згладжування різних типів. Ці підходи зосереджені на наданні більшої ваги останнім спостереженням у часовому ряді, що дозволяє швидше реагувати на зміни в динаміці даних [2].

Основною перевагою таких методів є їх простота та здатність ефективно працювати з короткостроковими прогнозами, навіть за умов обмежених обчислювальних ресурсів. До основних видів належать просте експоненціальне згладжування для стаціонарних рядів, метод Холта для даних із трендом і метод Холта-Вінтерса, який враховує як тренд, так і сезонність.

Метод Хольта-Вінтерса є вдосконаленням методу Хольта, який дозволяє враховувати сезонні коливання в часових рядах. Додавання сезонної компоненти розширює модель, перетворюючи її на систему рівнянь із трьома коефіцієнтами згладжування:

$$\begin{aligned}\hat{y}_{t+d} &= (a_t + d \cdot b_t)c_{t-s+d}, \\ a_t &= \alpha_1 \frac{y_t}{c_{t-s}} + (1 - \alpha_1)(a_{t-1} - b_{t-1}), \\ b_t &= \alpha_2(a_t - a_{t-1}) + (1 - \alpha_2)b_{t-1},\end{aligned}\tag{2.10}$$

$$c_t = \alpha_3 \frac{y_t}{a_t} + (1 - \alpha_t) c_{t-s},$$

де  $\alpha_1, \alpha_2, \alpha_3$  - приймають значення з діапазону  $[0; 1]$ ;

$y$  - приймають значення з діапазону;

$a$  - згладжене значення спостереження;

$b$  - коефіцієнт тренду;

$c$  - індекс сезонності;

$t$  - індекс поточного спостереження [36].

Алгоритм побудови моделі розпочинається з обчислення коефіцієнтів  $\alpha_0$  та  $b_0$  визначають параметри лінійного тренду. Після цього сезонні компоненти  $c_t$  обчислюються за допомогою формули:

$$c_t = \frac{y_t}{a_0 + b_0 t}, \quad (2.11)$$

Залежно від природи сезонності часових рядів, використовуються адитивна або мультиплікативна моделі [37, 38].

Адитивна модель використовується, коли сезонний компонент часових рядів незалежний від рівня. Це означає, що сезонні коливання є стабільними та не змінюються в масштабі при зростанні чи зниженні рівня тренду, тобто сезонність має сталу амплітуду, яка не залежить від рівня ряду.

Основними формулами за допомогою, яких можна представити дану модель є:

1. Оновлення рівня (основна тенденція ряду на момент часу  $t$ :

$$L_t = \alpha(y_t - S_{t-m}) + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (2.12)$$

де  $y_t$  - фактичне значення ряду в момент часу  $t$ ;

$S_{t-m}$  - сезонний компонент для попереднього сезону;

$L_{t-1}$  - рівень тренду на попередній момент;

$T_{t-1}$  - тренд на попередній момент;  $\alpha$  - параметр згладжування рівня ( $0 < \alpha < 1$ ).

2. Оновлення тренду (швидкість змін рівня):

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (2.13)$$

де  $L_t - L_{t-1}$  - зміна рівня між моментами  $t$  і  $t-1$ ;

$\beta$ - параметр згладжування тренду ( $0 < \beta < 1$ ).

### 3. Оновлення сезонності:

$$S_t = \gamma(y_t - L_t) + (1 - \gamma)S_{t-m}, \quad (2.14)$$

де  $y_t - L_t$  - відхилення від рівня ряду в момент  $t$ ;

$\gamma$ - параметр згладжування тренду ( $0 < \gamma < 1$ ).

### 4. Прогноз:

$$\hat{y}_{t+h} = L_t + hT_t + S_{t+h-m}, \quad (2.15)$$

де:  $h$  - горизонт прогнозу;

$S_{t+h-m}$ - сезонний компонент, повторений із періодичністю  $m$ .

Отже, ключовими рисами адитивної моделі є сезонність, яка додається до тренду, і саме тому вона незалежна від рівня ряду, та те, що прогноз рівномірно змінюється відповідно до тренду і сезонності.

Мультиплікативна модель Холт-Вінтерса ідеально підходить для часових рядів, де сезонність змінюється пропорційно загальному рівню. У таких випадках сезонний компонент не є сталим: його величина зростає або зменшується залежно від рівня ряду.

У мультиплікативній моделі сезонний компонент масштабується відповідно до загального рівня  $L_t$  і тренду  $T_t$ . Це означає, що чим більший рівень, тим вищою є амплітуда сезонних коливань, а чим нижчий рівень, тим менш помітними стають сезонні коливання.

Мультиплікативна модель враховує залежність сезонного компонента від рівня, використовуючи множення замість додавання. Це дає змогу правильно моделювати часові ряди, де амплітуда змінюється разом із рівнем.

Формули мультиплікативної моделі:

Оновлення рівня  $L_t$ :

$$L_t = \alpha \frac{y_t}{S_{t-m}} + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (2.16)$$

де  $\frac{y_t}{S_{t-m}}$  - коригування фактичного значення ряду, щоб прибрати сезонний ефект;

Рівень обчислюється як усереднене співвідношення між фактичним значенням і сезонним компонентом.

Оновлення тренду  $T_t$ :

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad (2.17)$$

де як і в адитивній моделі, тренд відображає зміну рівня з часом.

Оновлення сезонності  $S_t$ :

$$S_t = \gamma \frac{y_t}{L_t} + (1 - \gamma)S_{t-m}, \quad (2.18)$$

де сезонний компонент визначається як співвідношення між фактичним значенням і рівнем  $L_t$ , враховуючи поточний сезон.

Прогноз ( $\hat{y}_{t+h}$ ):

$$\hat{y}_{t+h} = (L_t + hT_t) \cdot S_{t+h-m}, \quad (2.19)$$

де прогнозований рівень масштабується за допомогою сезонного компонента.

Ключовими рисами даної моделі є сезонність, яка масштабується разом із трендом та те, що вона підходить для часових рядів із змінною амплітудою.

Отже, адитивна та мультиплікативна моделі Холт-Вінтерса є ефективними методами прогнозування часових рядів, що враховують тренд, сезонність і рівень даних. Адитивна модель підходить для рядів зі сталою сезонністю, яка не змінюється залежно від рівня, тоді як мультиплікативна модель ефективно моделює сезонність, пропорційну тренду. Вибір моделі залежить від характеру даних: для стабільної сезонності використовують адитивний підхід, а для змінної — мультиплікативний. Обидві моделі є надійними інструментами для аналізу рядів із сезонними та трендовими компонентами.

Штучні нейронні мережі є методом прогнозування, який ґрунтується на математичному моделюванні роботи нейронів. Цей підхід дозволяє враховувати

складні нелінійні залежності між прогнозованою змінною та її вхідними параметрами (предикторами) [12, 13].

Нейронна мережа являє собою структуру, побудовану з "нейронів", організованих у кілька шарів. Нижній шар відповідає за вхідні дані, верхній — за вихідні значення, а між ними можуть розташовуватись приховані шари з "прихованими нейронами", які обробляють і трансформують інформацію. Така організація дозволяє моделювати складні залежності між вхідними та вихідними даними.

Найпростіші нейронні мережі не мають прихованих шарів і фактично є аналогом лінійної регресії. Вхідні змінні пов'язані з коефіцієнтами, що називаються "вагами", які використовуються для обчислення прогнозів шляхом лінійної комбінації вхідних даних. Значення ваг визначаються в процесі навчання моделі за допомогою алгоритму оптимізації, який мінімізує функцію витрат. Додавання проміжного шару з прихованими нейронами перетворює нейронну мережу на нелінійну модель.

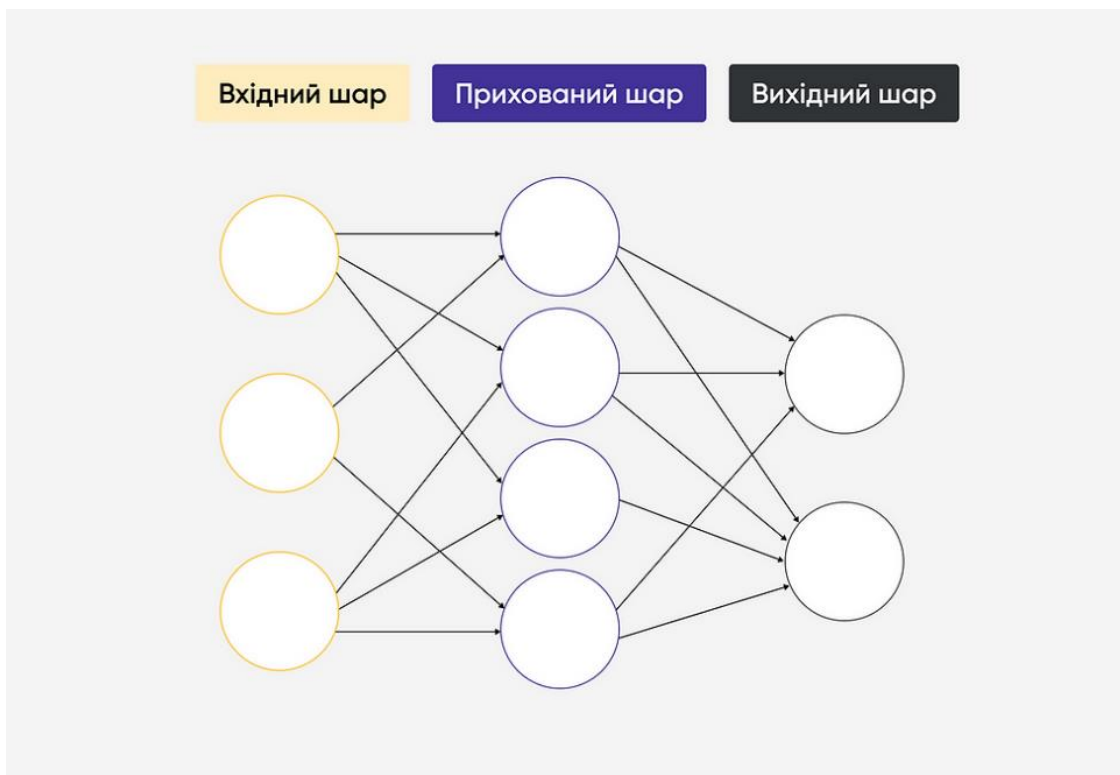


Рисунок 1.3 Приклад нейронної мережі [44]

На рисунку 2.3 зображена структура відома як багатошарова мережа прямого поширення, де кожен шар вузлів отримує дані від попереднього шару. Виходи вузлів одного шару слугують вхідними даними для наступного. Кожен вузол обробляє вхідні дані шляхом обчислення зваженої лінійної комбінації, після чого результат піддається трансформації за допомогою нелінійної функції. Наприклад, вхідні дані до прихованого нейрона  $j$  спочатку лінійно об'єднуються для отримання проміжного результату:

$$z_j = b_j + \sum_{i=1}^n w_{i,j} x_i \quad (2.20)$$

де  $z_i$  - проміжний результат для нейрона  $j$  перед застосуванням нелінійної активаційної функції;

$b_i$  - зміщення (*bias*) для нейрона  $j$ . Це додаткова вагова величина, яка дозволяє моделі краще узгоджуватись із даними;

$w_{i,j}$  - вага, яка з'єднує вхід  $i$  з нейроном  $j$ ;

$x_i$  - вхідне значення для вузла  $i$ ;

$n$  - кількість вхідних значень.

На прихованому шарі результат обробляється за допомогою нелінійної функції активації, наприклад, сигмоїди. Це дозволяє передати дані на наступний шар, зменшуючи вплив екстремальних вхідних значень, що підвищує стійкість мережі на більш високому рівні:

$$s(z) = \frac{1}{1 + e^{-z}} \quad (2.21)$$

де  $z$  - вхідне значення, яке є результатом зваженої лінійної комбінації (формула з попереднього кроку);

$e^{-z}$  - Експоненціальна функція, яка залежить від  $z$ ;

$s(z)$  - результат застосування сигмоїдної функції, що інтерпретується як ймовірність або активація нейрона  $j$ ;

Параметри  $b_j$  та  $w_i$  залежать від навчальних даних. Значення ваг  $w_i$  зазвичай обмежуються, щоб уникнути їх надмірного збільшення. Цей процес контролюється

так званим "параметром регуляризації", який часто задають зі значенням 0,1. На початковому етапі ваги ініціалізуються випадковими значеннями, які згодом оновлюються на основі навчальних даних. Оскільки в прогнозах, створених нейронною мережею, присутній елемент випадковості, модель тренують кілька разів, використовуючи різні початкові значення, а потім результати усереднюють. Кількість прихованих шарів та вузлів у кожному з них задається заздалегідь.

Головною перевагою нейронних мереж у порівнянні з іншими нелінійними моделями є їхня здатність гнучко адаптуватися до будь-яких нелінійних залежностей без необхідності попереднього припущення про природу процесу генерації даних [41].

Після того як нейронна мережа набуває нелінійності завдяки додаванню прихованих шарів, вона здатна моделювати складніші залежності. Однак для задач, де необхідно обробляти послідовні дані або враховувати довготривалі залежності, стандартні нейронні мережі з прихованими шарами можуть бути неефективними. У таких випадках на допомогу приходять більш спеціалізовані моделі, як-от Long Short-Term Memory (LSTM) [34].

LSTM є розширенням рекурентних нейронних мереж (RNN), яке дозволяє зберігати важливу інформацію на довший період, вирішуючи проблему затухання градієнтів, що виникає у традиційних RNN при навчанні з великими послідовностями даних. Ця модель особливо корисна для роботи з часовими рядами та іншими типами даних, де важливі взаємозв'язки між елементами, що відносяться до різних часів або етапів. Завдяки своїй архітектурі, LSTM здатен ефективно зберігати і передавати важливу інформацію протягом довгих проміжків часу, що робить його ідеальним для задач, де необхідно вловлювати залежності на великих інтервалах.

Основними елементами архітектури LSTM є: осередок пам'яті (зберігає інформацію протягом тривалого часу, відповідає за передачу важливих даних через послідовність); гейтова система (керує потоком інформації в осередку пам'яті) і складається з вхідного гейту (визначає, яку нову інформацію записати до осередку пам'яті), гейту забування (вирішує, яку інформацію слід "забути") та вихідного

гейту (регулює, яку інформацію передати як вихід поточного стану); стан осередку (шлях, яким інформація може передаватися майже без змін через мережу).

Кожен гейт використовує сигмоїдну функцію активації для обчислення ймовірностей (значень від 0 до 1), які визначають, наскільки певна інформація буде пропущена. Основні рівняння, які використовуються для побудови моделі є:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.22)$$

Гейт забування  $f_t$  обчислюється через сигмоїдну функцію активації, що залежить від попереднього прихованого стану  $h_{t-1}$  та вхідних даних  $x_t$ .

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.23)$$

Гейт входу  $i_t$  також обчислюється за допомогою сигмоїдної функції і контролює, яку нову інформацію потрібно записати до осередку пам'яті.

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.24)$$

Потенційний стан осередку пам'яті  $\tilde{C}_t$  обчислюється за допомогою тангенс-гіперболічної функції, що визначає, яку нову інформацію додати до стану.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (2.25)$$

Оновлений стан осередку пам'яті  $C_t$  обчислюється як комбінація старого стану та нової інформації, зваженої за допомогою гейту входу.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.26)$$

Вихідний гейт  $o_t$  контролює, яка частина інформації з осередку пам'яті буде передана у вихідну змінну.

LSTM має кілька важливих переваг, що робить її потужним інструментом для роботи з послідовними даними. Завдяки своїй унікальній архітектурі, LSTM може зберігати інформацію протягом тривалого часу, що критично важливо для аналізу послідовностей, де важливі далекі залежності. Ця модель також дуже адаптивна і може бути використана для різних завдань. Крім того, завдяки спеціальним механізмам, вона менш вразлива до проблеми затухання градієнтів, з якою часто стикаються традиційні рекурентні мережі.

У результаті LSTM є дуже гнучкою і здатною ефективно моделювати складні та довготривалі залежності в даних, що робить її ідеальним вибором для багатьох завдань, зокрема для моделювання часових рядів.

### **2.3 Система метрик для оцінювання якості динамічних моделей споживання електроенергії**

У процесі прогнозування та вибору початкових параметрів моделі постає завдання оцінювання точності передбачень. Таке оцінювання є важливим як для порівняння різних моделей між собою, так і для визначення надійності отриманого прогнозу, тобто потрібно оцінити, наскільки вони відповідають фактичним спостереженням. Існує багато методів, які дозволяють оцінити точність прогнозування, але для розрахунку будь-якого з них необхідно знати величину похибки прогнозу на кожному етапі.

Перевірка адекватності моделі може включати декілька етапів. Перш за все, порівнюються прогнози, отримані від моделі, з фактичними значеннями часового ряду. Якщо прогнози добре узгоджуються з реальними спостереженнями, можна стверджувати, що модель є адекватною

Оцінка залишків є важливим етапом при перевірці адекватності моделі прогнозування. Теоретично припускається наявність випадкової складової. Залишкові значення визначаються як різниця між фактичними спостереженнями та прогнозами.

Похибка прогнозу на один крок уперед у момент часу  $t$  визначається як:

$$e_t = X_t - \hat{X}_t \quad (2.27)$$

де  $X_t$  - це фактичне значення вхідної послідовності в момент часу  $t$ ;

$\hat{X}_t$  - прогнозоване значення на цей момент, отримане на попередньому кроці.

Оцінка адекватності моделі прогнозування часових рядів також може бути здійснена за допомогою різних числових критеріїв.

Однією з найпоширеніших метрик для оцінки точності прогнозів є корінь середньоквадратичної помилки (RMSE), який дає більш інтуїтивно зрозуміле значення помилки в оригінальних одиницях вимірювання. Корінь середньоквадратичної помилки дає інтуїтивно зрозуміле значення помилки в оригінальних одиницях вимірювання та враховує величину помилок, що дозволяє краще відобразити реальну ефективність моделі. Це важливо в задачах, де великі відхилення можуть бути критичними для точного прогнозування.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (X_t - \hat{X}_t)^2}, \quad (2.28)$$

де  $X_t$  - це фактичне значення вхідної послідовності в момент часу  $t$ ;

$\hat{X}_t$  - прогнозоване значення на цей момент, отримане на попередньому кроці;

$n$  - кількість спостережень.

RMSE є корисним показником для оцінки точності моделей, оскільки він дає уявлення про середнє відхилення прогнозів від фактичних значень. Підвищена чутливість до великих помилок дозволяє моделі враховувати значні відхилення, що робить її ефективною для виявлення аномалій у даних. Це важливо, коли навіть незначні зміни можуть мати великий вплив на точність прогнозу [18].

Оскільки RMSE вимірюється в тих самих одиницях, що й вихідні дані, він надає зручний інструмент для порівняння різних моделей. Проте, цей показник не завжди точно відображає точність моделі в умовах неоднорідних або масштабованих даних, оскільки його значення можуть варіюватися залежно від розміру вихідних величин. Крім того, припущення про нормальний розподіл помилок може обмежити точність оцінки, якщо розподіл даних суттєво відрізняється.

Загалом, менші значення RMSE вказують на кращу точність моделі, проте для комплексних даних доцільно використовувати додаткові метрики для оцінки ефективності прогнозів.

Наступним популярним критерієм є MAE - це середня абсолютна помилка, яка вимірює середнє абсолютне відхилення прогнозованих значень від фактичних значень.

Менші значення MAE також свідчать про кращу адекватність моделі.

$$MAE = \frac{1}{n} \sum_{t=1}^n |X_t - \hat{X}_t|, \quad (2.29)$$

де  $X_t$  - це фактичне значення вхідної послідовності в момент часу  $t$ ;

$\hat{X}_t$  - прогнозоване значення на цей момент, отримане на попередньому кроці;

$n$  - кількість спостережень.

У метриці MAE (середня абсолютна помилка) помилка підноситься до абсолютного значення, що робить її менш чутливою до великих відхилень порівняно з іншими метриками, які підносять помилку до квадрату. Це дозволяє отримувати більш стабільні оцінки, оскільки великі аномалії не мають такого сильного впливу на результат. Окрім того, MAE вимірюється в одиницях, що відповідають вихідним даним, що робить її дуже зрозумілою і зручною для інтерпретації. Оскільки ця метрика не підсилює великі помилки, вона дає більш прямолінійну оцінку середнього рівня помилок.

Проте, важливо зазначити, що MAE може бути менш чутливою до незначних відхилень. Це означає, що модель може не виявляти малі зміни або тренди, що може стати проблемою в задачах, де важливі навіть невеликі корективи. Однак одна з головних переваг MAE полягає в тому, що її можна використовувати для оптимізації моделей. При такій оптимізації мета полягає в мінімізації абсолютних помилок, що є особливо корисним для задач, де важливо уникати великих відхилень у прогнозах, гарантуючи стабільність і точність.

Ще одним критерієм для оцінки помилки прогнозування часового ряду можна використовувати середню абсолютну процентну помилку (Mean Absolute Percentage Error, MAPE).

MPE (Mean Percentage Error): це середнє відсоткове відхилення прогнозованих значень від реальних, що показує, наскільки в середньому прогнози відрізняються від фактичних даних у відсотках:

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (2.30)$$

де  $y_i$  - значення вихідного ряду;

$\hat{y}_i$  - прогнозовані значення;

$n$  - кількість спостережень.

Значення MPE, близьке до нуля, свідчить про хорошу адекватність моделі. Ця метрика має декілька суттєвих переваг, зокрема здатність вимірювати помилки в відсоткових термінах, що дає змогу порівнювати точність прогнозів між різними моделями або часовими рядами, навіть якщо їх масштаби різні. Це особливо корисно, коли йдеться про дані з різними одиницями вимірювання або величинами. Відсоткові відхилення також простіше інтерпретувати, що робить результат більш зрозумілим для більшості користувачів.

Однак, попри ці переваги, MPE має певні обмеження. Вона може бути надмірно чутливою до великих відхилень, зокрема коли фактичні значення значно менші за прогнозовані. У таких випадках великі відсоткові помилки можуть суттєво спотворити результат, особливо коли дані мають високу варіативність. Крім того, коли фактичні значення наближаються до нуля або дуже малі, MPE може стати нестабільною. В таких випадках навіть незначні відхилення можуть призвести до значних відсоткових помилок, що робить оцінку точності менш надійною.

Отже, правильний вибір критеріїв оцінки моделей є критично важливим для вимірювання точності прогнозів та вимагає комплексного підходу, оскільки кожен з них має свої сильні та слабкі сторони. Критерії, такі як RMSE, MAPE та MAE, є ефективними інструментами для оцінки точності прогнозів однієї й тієї ж послідовності при зміні параметрів моделі або застосуванні різних моделей, але їхня чутливість до різних аспектів даних (наприклад, великих відхилень або малих значень) вимагає уважного підходу до вибору відповідного показника.

Загалом, комбінування кількох критеріїв дозволяє отримати більш обґрунтовану та точну оцінку моделі, що допомагає вибрати найбільш адекватний підхід для прогнозування в конкретних умовах.

## РОЗДІЛ 3

### АНАЛІЗ ТА МОДЕЛЮВАННЯ ДИНАМІКИ СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ В УКРАЇНІ

#### 3.1 Аналіз результатів візуалізації та препроцесингу даних про споживання електроенергії в Україні

Звітність про обсяги споживання електричної енергії за попередні роки надає важливу інформацію для побудови моделей на основі часових рядів. Аналізуючи ці дані, можна виявити закономірності та тенденції у споживанні електроенергії, що допомагає здійснити прогнозування на майбутнє.

За основу були взяті дані про споживання електроенергії в Україні за період з 31 січня 2016 року по 31 січня 2022 року, що включає рівно 6 років щомісячних значень. Споживання подається у млн кВт/год [37].

Набір даних складається з 73 спостережень, які рівномірно розподілені в часі з місячною частотою. Основними показниками в аналізі є значення споживання електроенергії (стовпець `consumption`) та відповідні дати спостережень (стовпець `Date`) [3].

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 73 entries, 0 to 72
Data columns (total 2 columns):
#   Column          Non-Null Count  Dtype
---  -
0   consumption     73 non-null    float64
1   Date            73 non-null    datetime64[ns]
dtypes: datetime64[ns](1), float64(1)
memory usage: 1.3 KB
(None,
  consumption      Date
0    11055.116 2016-01-31
1    10406.871 2016-02-29
2    10257.894 2016-03-31
3     9281.558 2016-04-30
4     8842.304 2016-05-31)
```

Рисунок 3.1 Загальна інформація про дані

*Джерело: розроблено автором*

Перевіряємо ряд на пропуски та отримуємо результат зображений на рисунку 3.2, який свідчить, що пропуски відсутні.

```
Кількість пропущених значень у кожному стовпці:
consumption      0
Date             0
dtype: int64
```

Пропущених значень немає.

Рисунок 3.2 Перевірка на відсутність значень

*Джерело: розроблено автором*

Для отримання загального уявлення про структуру даних було обчислено наступні статистичні показники (рисунок 3.3):

Середнє значення споживання, яке дорівнює 10,067 млн кВт/год, показує типовий обсяг споживання електроенергії за місяць протягом усього досліджуваного періоду. Воно є свідчить про значну інтенсивність споживання в електроенергії в Україні.

Медіанне значення - 9,859 млн кВт/год, вказує, що половина спостережень має споживання нижче цього рівня, а інша половина — вище.

Стандартне відхилення свідчить про варіативність даних щодо середнього значення. Значення 932.27 вказує на відносно значний розкид спостережень, що може бути зумовлено сезонними коливаннями, економічними змінами чи іншими факторами.

```
mean_consumption': 10067.01912328767,
median_consumption': 9859.4,
std_dev_consumption': 932.2697266247712,
min_consumption': 8720.154,
max_consumption': 12062.8,
date_range': (Timestamp('2016-01-31 00:00:00'),
Timestamp('2022-01-31 00:00:00'))}
```

Рисунок 3.3 Основні статистичні показники

*Джерело: розроблено автором*

Дані демонструють значну варіативність споживання, що коливається в межах від 8,720 до 12,062 млн кВт/год. Такий розкид може бути пов'язаний із сезонними змінами або іншими зовнішніми факторами. Порівняння середнього та медіанного значень показує, що розподіл споживання близький до симетричного.

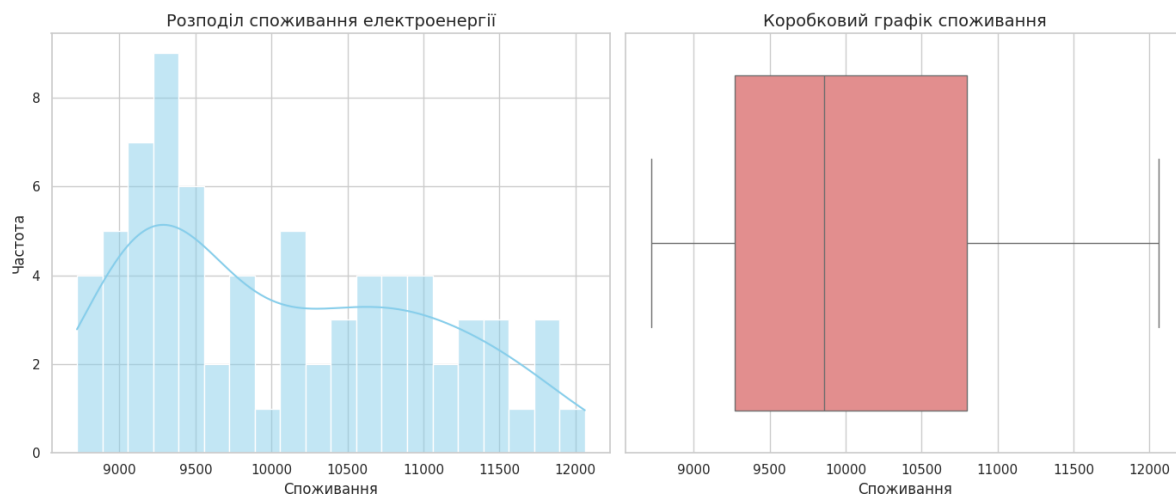


Рисунок 3.4 Візуалізація розподілу даних

*Джерело: розроблено автором*

З рисунку 3.4 бачимо, що спостерігається характерна хвиля з двома піками, що свідчить про наявність двох основних періодів, коли споживання досягає максимальних значень і показує наявність певної сезонності споживання. Коробковий графік не показує значних викидів, оскільки жодне із значень не вийшов за межі "вусів" коробки, що свідчить про відносно стабільне споживання протягом періоду дослідження.

Наступним кроком є візуалізація наших даних. Будуємо графік залежності споживання від дати.

З рис. 3.5 бачимо, що дані демонструють періодичні коливання, що може свідчити про наявність сезонності. Найбільші значення споживання, ймовірно, припадають на певні періоди (наприклад, зима і літо).

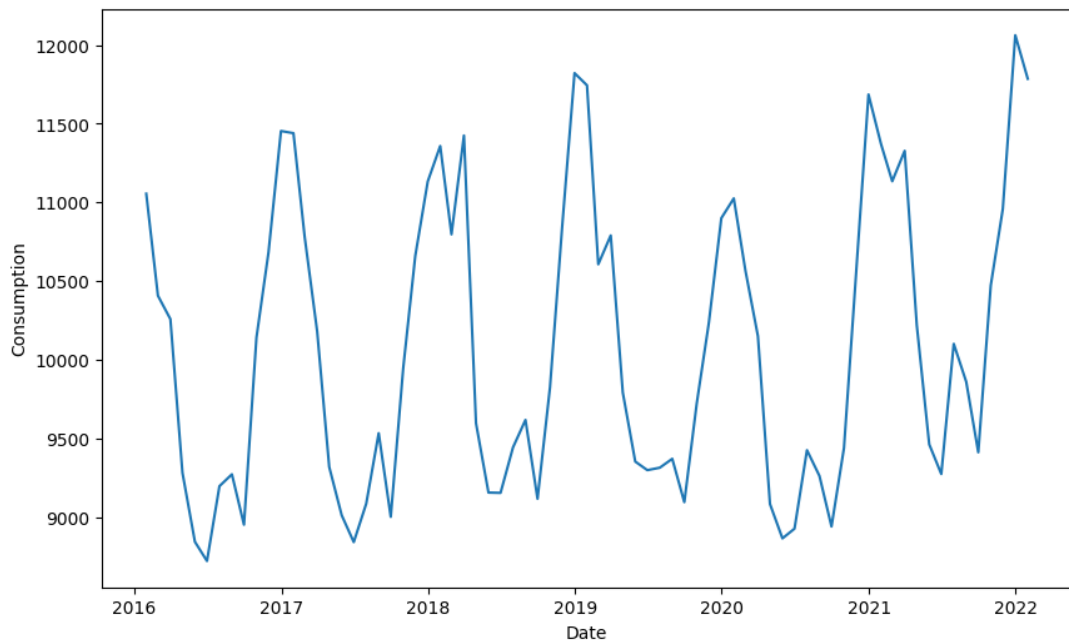


Рисунок 3.5 Візуалізація загального споживання електроенергії

*Джерело: розроблено автором*

Для перевірки стаціонарності часового ряду було використано тест Дікі-Фуллера (ADF).

**p-value: 0.932184883073237**

Рисунок 3.6 тест Дікі-Фуллера

*Джерело: розроблено автором*

Результати тесту зображені на рисунку 3.6 показали, що  $p\text{-value} = 0.93$ . Оскільки  $p\text{-value}$  більше 0.05, ми не відкидаємо нульову гіпотезу про наявність одиничного кореня в ряді, що вказує на те, що ряд є нестационарним. У такому випадку існує ймовірність наявності тренду в даних.

Оскільки, для моделювання будемо використовувати AutoARIMA та Holt-Winters, то потреби у диференціюванні ряду немає. Оскільки, AutoARIMA автоматично визначить, чи потрібно застосувати його чи ні, що дозволить моделі коректно обробити тренд та сезонність. Модель Holt-Winters, в свою чергу, сама ефективно справляється з трендом і сезонними коливаннями без необхідності попереднього диференціювання, оскільки вона вбудовано враховує ці компоненти

при побудові прогнозів. Для LSTM доцільно виконати попереднє оброблення даних (наприклад, нормалізацію або стандартизацію), для побудови ефективного прогнозу.

Наступним етапом, стало використання функції ACF на часовому ряді, що допомагає аналізувати його автокореляцію і виявляти можливі залежності між спостереженнями у часі.

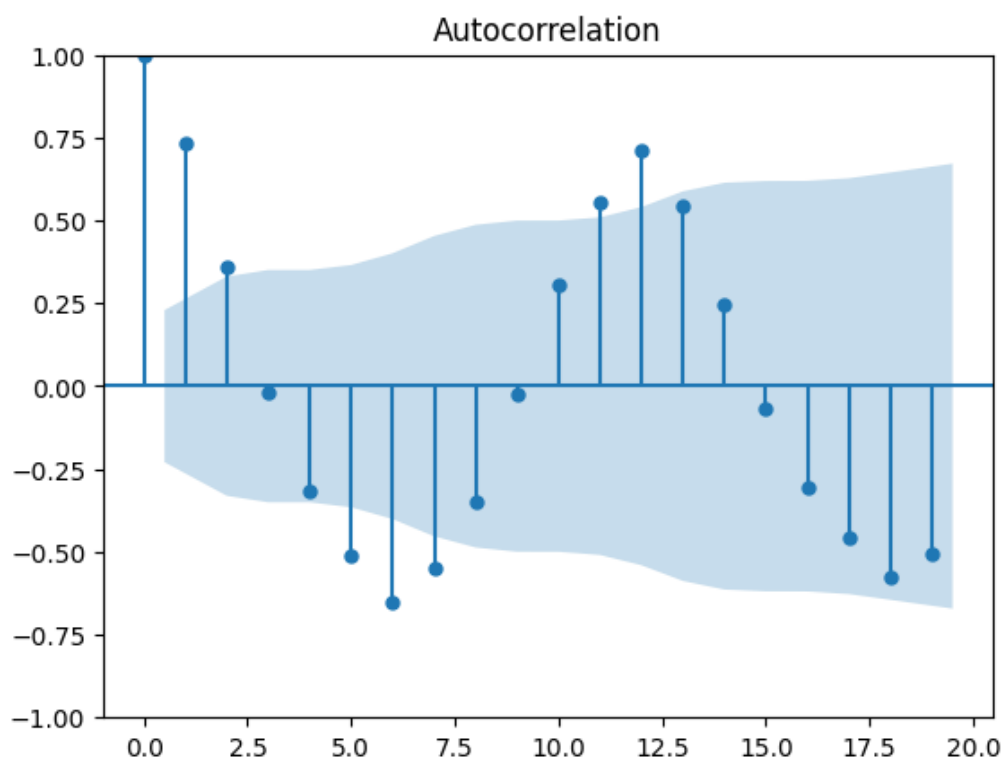


Рисунок 3.7 Результати функції ACF

*Джерело: розроблено автором*

Результати автокореляційної функції (ACF) з наведеними лагами на рис. 3.7 можна пояснити за такими критеріями: залежність у даних (наявність значущих автокореляцій на початкових лагах (1, 2, 3) свідчить про те, що дані мають залежність між поточними та попередніми значеннями); можлива сезонність (помітні піки автокореляції на лагу ~6 та ~12 можуть свідчити про сезонні тенденції, а саме піврічна або річна сезонність).

Використаємо також метод сезонної декомпозиції для розкладання часового ряду на його основні компоненти: тренд, сезонність і залишкові значення.

На основі графіків поданих на рис 3.8 можна зробити наступні висновки.

Графік тренду показує поступове зростання значень, що свідчить про наявність тренду в часі. Цей тренд не є різким, але є чітким і показує, що споживання електроенергії поступово збільшується, ймовірно через фактори, як-от зростання населення, індустріалізація чи інші економічні причини.

На графіку сезонної складової видно явну сезонність. Періодичні коливання мають схожу амплітуду та форму, що свідчить про адитивну природу.

Графік залишків показує випадкові коливання навколо нульової лінії без явних трендів чи сезонних циклів. Це означає, що після виділення компонентів тренду та сезонності, залишки не містять структурних патернів і є випадковими. Це позитивний сигнал, оскільки це вказує на те, що більшість варіацій у даних були описані через тренд і сезонність, і залишки не містять неврахованих впливів.

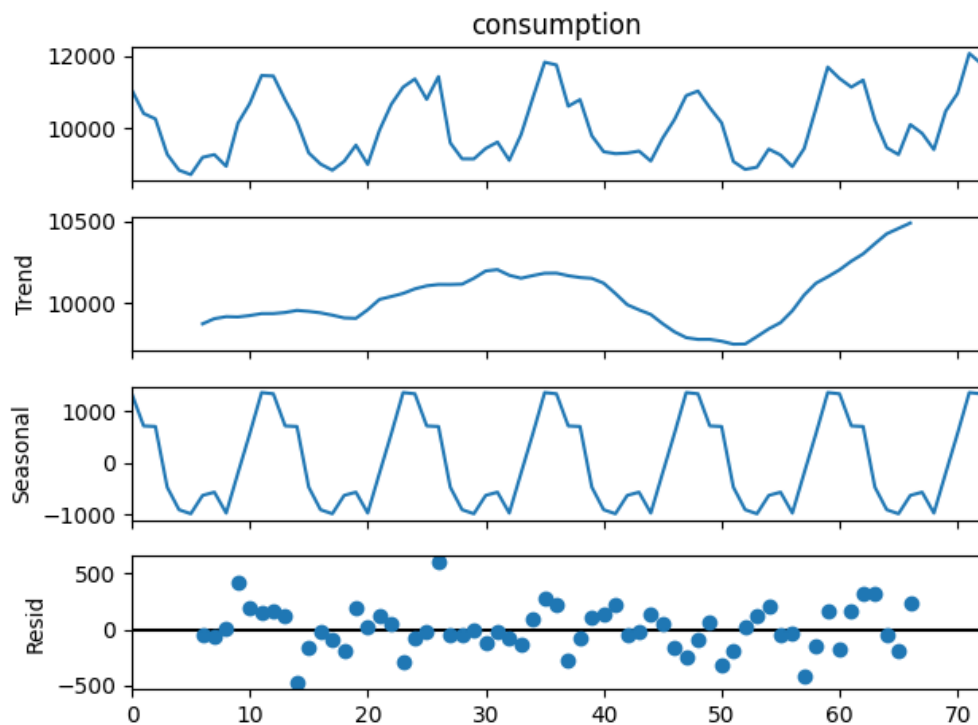


Рисунок 3.8 Результати методу сезонної декомпозиції

*Джерело: розроблено автором*

Для подальшого моделювання даних тренд та сезонність не будуть видалятися, оскільки обрані моделі, AutoARIMA, Holt-Winters та LSTM, здатні самостійно враховувати і налаштовувати ці компоненти під час побудови прогнозів.

Після аналізу часового ряду переходимо до прогнозування. Для початку розділимо дані на тренувальну та тестову вибірку, у пропорції 80%/20%, де 80% відведено для навчальної вибірки

Для порівняння прогнозованих значень застосуємо різні методи прогнозування. Почнемо з побудови моделі ARIMA. У процесі прогнозування часових рядів для підбору параметрів моделі була використана функція `pm.auto_arma()`. Її налаштування забезпечують автоматизацію вибору оптимальних параметрів моделі та врахування сезонності [4].

```
model = pm.auto_arma(train, seasonal=True, m=12, stepwise=True, trace=True, suppress_warnings=True)
```

### Рисунок 3.9 Побудова моделі AutoARIMA

*Джерело: розроблено автором*

На рисунку 3.9 наведено ключові налаштування моделі: включений параметр сезонності, який активує врахування сезонних коливань; періоду сезонності дорівнює 12 і визначає тривалість одного сезонного циклу. Використовується покроковий алгоритм для пошуку оптимальних параметрів. Такий підхід значно зменшує час обчислення, порівняно з повним перебором усіх можливих комбінацій параметрів.

У результаті отримуємо найкращу модель:

$$\text{SARIMAX}(1, 0, 0) \times (1, 0, [1], 12), \quad (3.1)$$

де  $p = 1$  - перший автокореляційний компонент;

$d = 0$  - дані стаціонарні (без диференціювання);

$q = 0$  - відсутність MA-компонентів;

$P = 1$  - перший сезонний AR-компонент;

$D = 0$  - без сезонного диференціювання;

$Q = [1]$  - перший сезонний MA-компонент;

$m = 12$  - сезонність кожні 12 періодів.

На рисунку 3.10 наведено параметри найкращої моделі.

|           | coef      | std err  | z     | P> z  | [0.025   | 0.975]   |
|-----------|-----------|----------|-------|-------|----------|----------|
| intercept | 2448.3181 | 945.466  | 2.590 | 0.010 | 595.239  | 4301.397 |
| ar.L1     | 0.4691    | 0.107    | 4.404 | 0.000 | 0.260    | 0.678    |
| ar.S.L12  | 0.5338    | 0.167    | 3.201 | 0.001 | 0.207    | 0.861    |
| ma.S.L12  | 0.2618    | 0.260    | 1.009 | 0.313 | -0.247   | 0.770    |
| sigma2    | 1.424e+05 | 2.75e+04 | 5.168 | 0.000 | 8.84e+04 | 1.96e+05 |

Рисунок 3.10 Параметри найкращої моделі SARIMAX

*Джерело: розроблено автором*

На основі параметрів, отриманих із моделі SARIMAX сформулюємо її аналітичний вигляд:

$$y_t = 2448.3181 + 0.4691y_{t-1} + 0.5338y_{t-12} + 0.2618\epsilon_{t-12} + \epsilon_t \quad (3.2)$$

де:  $c = 2448.31$  — константа, що враховує базове споживання електроенергії;  $\phi_1 = 0.46$  — коефіцієнт авторегресії, який показує, як сильно поточне споживання електроенергії залежить від значення споживання в попередньому періоді ( $y_{t-1}$ ). Якщо значення  $\phi_1$  велике, це означає, що споживання в поточному періоді сильно залежить від попереднього періоду. У даному випадку коефіцієнт вказує на помірну залежність: тобто поточне споживання зростає приблизно на 47% від рівня попереднього періоду;  $\phi_{12} = 0.53$  — це коефіцієнт сезонної авторегресії, який вказує на вплив споживання електроенергії в тому ж періоді попереднього року ( $y_{t-12}$ ). В цьому випадку він свідчить про помірний, але все ж значний вплив сезонних коливань — якщо споживання в попередньому році було високим у певному періоді, то в поточному році воно також буде відповідно високим. Це означає, що річна сезонність (наприклад, зимові та літні піки споживання) є важливим фактором для прогнозування поточного споживання. В залежності від пори року, споживання електроенергії може значно коливатись, і тому попереднє значення споживання в тому ж місяці минулого року використовується для корекції прогнозу;  $\theta_{12} = 0.26$  — це параметр, який враховує шум у попередньому сезонному періоді ( $\epsilon_{t-12}$ ). Він показує, як сильно залишки попереднього сезону

можуть впливати на поточний прогноз. В цьому випадку він означає, що помилки, допущені в прогнозах за попередній сезон, мають вплив на поточний прогноз, але цей вплив не є дуже сильним (тобто шум не настільки великий, щоб повністю змінювати прогноз). Це також може вказувати на те, що модель враховує минулі неточності для кращого коригування майбутніх прогнозів;  $\theta\epsilon_t$  — залишки моделі.

На рис. 3.11 наведено візуалізацію прогнозу. Графік ілюструє розподіл фактичних та прогнозованих значень споживання енергії на навчальній і тестовій вибірках, а також показує, наскільки точно модель відображає динаміку часових рядів.

Загалом модель ефективно врахувала загальну тенденцію та сезонні коливання споживання енергії. Однак у деяких періодах прогноз відрізняється від фактичних значень, що може бути спричинено нерегулярними сплесками або зниженнями, які модель не змогла передбачити.



Рисунок 3.11 Візуалізація прогнозних значень моделі SARIMAX

*Джерело: розроблено автором*

Перейдемо до наступного прогнозування. Метод, який був використаний — метод експоненціального згладжування Хольта-Вінтера (Holt-Winters) [2].

```
model = ExponentialSmoothing(
    train['consumption'],
    trend='mul',
    seasonal='add',
    seasonal_periods=12
).fit(optimized=True)
```

Рисунок 3.12 Побудова моделі Holt-Winters

*Джерело: розроблено автором*

На рисунку 3.12 наведено налаштування використані для моделювання: тип тренду встановлений як мультиплікативний, тобто тренд зростає або зменшується пропорційно до значень ряду; сезонність моделі задається як адитивна (на основі проведеного попереднього аналізу). Це вказує, що сезонні коливання додаються до основного рівня ряду, і їхній вплив є постійним незалежно від масштабу ряду; період сезонності дорівнює 12. Це означає, що модель враховує сезонні коливання, які повторюються кожні 12 періодів; увімкнений оптимізований автоматичний підбір параметрів моделі. Він забезпечує автоматичний вибір оптимальних значень коефіцієнтів для підвищення точності прогнозу.

Коефіцієнти моделі:  
 Рівень ( $\alpha$ ): 0.4293  
 Тренд ( $\beta$ ): N/A  
 Сезонність ( $\gamma$ ): 0.0001  
 Початковий рівень: 9881.5724  
 Початковий тренд: N/A

Рисунок 3.13 Параметри моделі Holt-Winters

*Джерело: розроблено автором*

На рисунку 3.13 наведено отримані коефіцієнти параметрів отримані для даної моделі. Отже, аналітичний вигляд моделі з врахуванням отриманих коефіцієнтів має вигляд:

$$\hat{y}_{t+h} = (L_t \cdot T_t^h) + S_{t+h-12k}, \quad (3.3)$$

де:  $L_t$  — рівень у момент часу  $t$ ;  $T_t$  — тренд, який у даній моделі не враховується;  $y_t$  — фактичне значення ряду у момент часу  $t$ ;  $h$  — кількість періодів прогнозування;  $m = 12$  — кількість сезонних періодів;  $S_t$  — сезонність, яка оновлюється за формулою:

$$L_t = 0.4293 \cdot \frac{y_t}{T_{t-1}} + (1 - 0.4293) \cdot (L_{t-1} \cdot T_{t-1}), \quad (3.4)$$

$$S_t = 0.0001 \cdot (y_t - L_t) + (1 - 0.0001) \cdot S_{t-12}, \quad (3.5)$$

Розглянемо детальніше значення коефіцієнтів:  $\alpha = 0.4293$  — це коефіцієнт для параметра рівня, який визначає, наскільки сильно нові спостереження впливають на поточний прогноз. Отримане значення вказує на те, що поточне значення прогнозу на 42.93% залежить від поточного спостереження, а решта (57.07%) — від попередніх значень прогнозу. Це вказує на достатньо високий ваговий вплив останнього спостереження; Тренд відсутній у цій моделі, тобто припускається, що зміна значень часового ряду з часом не враховується;  $\gamma = 0.0001$  — коефіцієнт вказує на вплив сезонних коливань на прогноз. Поточне значення вказує на слабкий вплив сезонних змін на поточний прогноз.



Рисунок 3.14 Візуалізація прогнозних значень моделі Holt-Winters

Джерело: розроблено автором

На рисунку 3.14 зображено результати прогнозу за допомогою моделі Holt-Winters. Аналіз дозволяє оцінити, наскільки добре модель враховує сезонність, тенденції та рівень ряду, а також порівняти реальні та прогнозовані дані.

Модель вдало враховує річну сезонність, що помітно на періодичних коливаннях пунктирної лінії. Використання автоматичного підбору параметрів оптимізує результати, забезпечуючи найкраще наближення до реальних даних.

В цілому прогноз добре збігається з тестовими даними, хоча в окремих точках можливі незначні відхилення.

Третім варіантом моделі обрано модель LSTM (Long Short-Term Memory). Модель налаштовувалася автоматично з використанням бібліотеки Keras Tuner, що дозволяє здійснювати оптимізацію гіперпараметрів для досягнення найкращих результатів.

Структура моделі виглядає наступним чином: LSTM шар, де кількість нейронів визначається за допомогою гіперпараметра, який може варіюватися в діапазоні від 32 до 128 з кроком 32. Це дозволяє моделі адаптуватися до різних рівнів складності задачі; використовувалась функція активації ReLU (Rectified Linear Unit), яка є стандартною для багатьох нейронних мереж завдяки своїй здатності прискорювати навчання; вхідний шар моделі приймає вхідні дані розміром, що відповідає кількості часових кроків у послідовності (12 місяців для річної сезонності) та одиничному каналу (одна ознака - споживання енергії); вихідний шар має один нейрон, оскільки модель прогнозує одночасно одне значення.

Налаштування параметрів виконано за допомогою методу Keras Tuner. Для цього було застосовано алгоритм, який автоматично обирає найкращі значення гіперпараметрів на основі мінімізації функції втрат.

| Layer (type)  | Output Shape | Param # |
|---------------|--------------|---------|
| lstm (LSTM)   | (None, 64)   | 16,896  |
| dense (Dense) | (None, 1)    | 65      |

Total params: 16,961 (66.25 KB)  
Trainable params: 16,961 (66.25 KB)  
Non-trainable params: 0 (0.00 B)

Рисунок 3.15 Архітектура моделі LSTM

Джерело: розроблено автором

На рисунку 3.15 зображено архітектуру моделі LSTM. Ця модель має лише треновані параметри, що означає, що всі параметри підлягають оптимізації під час навчання. У процесі навчання модель отримала оптимальні ваги, що дозволяють передбачати споживання електроенергії на основі історичних даних.

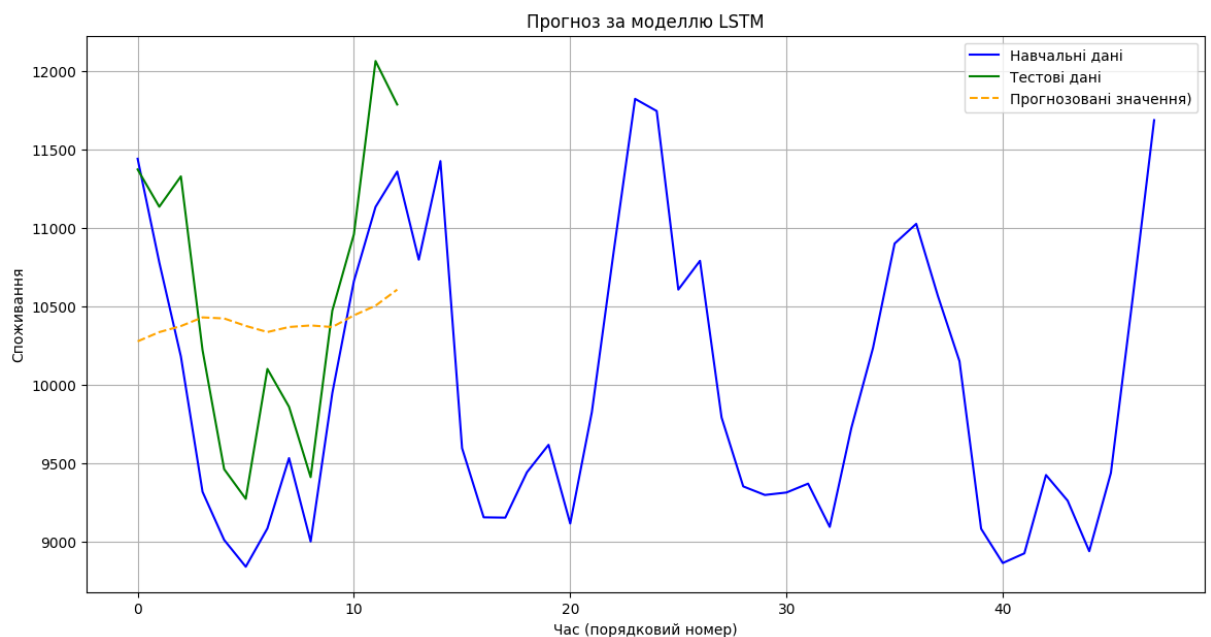


Рисунок 3.16 Візуалізація прогнозних значень моделі LSTM

Джерело: розроблено автором

Результати прогнозування за допомогою моделі наведені на рисунку 3.16. Загальний тренд графіка демонструє, як модель LSTM намагається відтворити споживання, спираючись на навчальні дані. Прогнози слідують основним тенденціям, зокрема зростанням та спадам, але є деякі розбіжності. Хоча прогнози

не є ідеальними, вони дають корисну інформацію для розуміння загальних тенденцій у споживанні.

### **3.2 Аналіз результатів моделювання споживання електроенергії на базі моделей експоненціального згладжування, ARIMA - моделей та нейромережевих моделей**

Після виконання прогнозування за допомогою кількох моделей важливо обрати найкращу, використовуючи різні критерії.

Проаналізуємо стандартні відхилення залишків (Residual sd) для оцінка точності моделі. Стандартне відхилення залишків дає уявлення про середнє відхилення прогнозованих значень від фактичних. Чим менше це значення, тим точніше модель. Високе стандартне відхилення може свідчити про те, що модель не змогла точно передбачити поведінку даних. При наявності декількох моделей воно може допомогти порівняти, яка з моделей дає найкраще узгодження з фактичними даними.

У таблиці 3.1 наведено результати тесту для всіх моделей:

Таблиця 3.1 Результати Residual sd для всі моделей

| Type         | Residual sd |
|--------------|-------------|
| SARIMAX      | 538.8556    |
| Holt-Winters | 249.7881    |
| LSTM         | 843.2926    |

*Джерело: розроблено автором*

Отже, найкращий результат показала модель Holt-Winters, вона має найнижче значення стандартного відхилення залишків - 249.7881. На другому місці

- SARIMAX модель, яка має значення, яке дорівнює 538.8556. Останнє місце посідає LSTM модель з найгіршим значенням - 843.2926.

Далі необхідно перевірити залишки кожної моделі на предмет того, чи є вони «білим шумом», тобто випадково розподіленими. Для цього застосовується тест Ljung-Box, який дає значення p-value. Якщо p-value менше визначеного рівня значущості (наприклад, 0.05), це означає, що гіпотеза про відсутність автокореляції відхиляється. У випадку, коли p-value більше цього рівня, відсутні достатні підстави для відхилення гіпотези про незалежність залишків.

Таблиця 3.2 Результати тесту Ljung-Box для моделей

| Type         | lb_pvalue |
|--------------|-----------|
| SARIMAX      | 0.03      |
| Holt-Winters | 0.89      |
| LSTM         | 0.0009    |

*Джерело: розроблено автором*

Отже, згідно до результатів наведених у таблиці 3.2 після проведення тесту Ljung-Box для всіх моделей, 2 моделі (SARIMAX та LSTM) мають нижчі значення p-value за рівень значущості 0.05, а тому не пройшли тест на наявність білого шуму. У свою чергу модель Holt-Winters показала хороший результат і на даний момент є найкращим варіантом для побудови прогнозу.

Для подальшого аналізу застосуємо числові критерії, а саме критерії RMSE, MAE, MPE. У таблиці 3.3 наведено результати отриманих для всіх моделей:

Таблиця 3.3 Результати критеріїв для всіх моделей

|              | RMSE | MAE | MPE   |
|--------------|------|-----|-------|
| SARIMAX      | 1032 | 880 | 7.93% |
| Holt-Winters | 646  | 596 | 5.64% |
| LSTM         | 890  | 763 | 2.03% |

*Джерело: розроблено автором*

Розглянемо детально результати трьох моделей (SARIMAX, Holt-Winters та LSTM) для прогнозування. Вони представлені за допомогою трьох різних метрик: RMSE (Root Mean Squared Error), MAE (Mean Absolute Error) та MPE (Mean Percentage Error). Ось як можна інтерпретувати ці результати.

RMSE (Root Mean Squared Error): SARIMAX має найбільше значення, що вказує на те, що модель має найвищі середні квадратичні відхилення між прогнозами та фактичними даними, тобто її прогнози менш точні в порівнянні з іншими моделями; Holt-Winters має найменше значення, що означає, що ця модель дала найбільш точні прогнози з усіх трьох моделей; LSTM займає проміжне місце, що вказує на помірну точність моделі порівняно з іншими.

MAE (Mean Absolute Error): SARIMAX має найбільше значення, що означає, що ця модель має найбільше середнє абсолютне відхилення прогнозів від фактичних значень; Holt-Winters знову показує найменше значення, що підтверджує, що ця модель є найбільш точною за середнім розміром помилок серед трьох моделей; LSTM займає середнє місце, що свідчить про помірну точність в порівнянні з іншими двома моделями.

MPE (Mean Percentage Error): SARIMAX має найвищий результат, що означає, що відносні помилки прогнозів цієї моделі найбільші. Тобто її прогнози в середньому відрізняються на 7.93% від фактичних значень. Holt-Winters має середнє значення 5.65%, що показує, що її відносні помилки менші, ніж у SARIMAX. LSTM має найменше значення 2.03%, що вказує на те, що ця модель дає найбільш точні прогнози з точки зору відсоткових помилок.

Отже, Holt-Winters є найкращою моделлю серед трьох за всіма метриками. Це вказує на те, що ця модель дає найточніші прогнози з найменшими помилками. LSTM демонструє добрі результати, займаючи друге місце за точністю, з меншими відсотковими помилками. SARIMAX виявляється найменш точною моделлю з найбільшими помилками, як за абсолютними величинами (MAE), так і за відносними помилками (MPE). Вона також має найбільше квадратичне відхилення (RMSE), що свідчить про найбільші відхилення в її прогнозах.

На підставі аналізу різних критеріїв оцінки адекватності моделей для прогнозування рівня споживання, ми можемо зробити наступні висновки.

Модель Holt-Winters є найбільш підходящою для даного завдання, оскільки вона продемонструвала найкращі результати за точністю прогнозів, а також мала стабільні залишки, що підтверджується високим p-value за тестом Ljung-Box. Вона не лише забезпечує найменші значення RMSE та MAE, а й продемонструвала найменше стандартне відхилення залишків, що робить її надійною і ефективною для прогнозування.

Модель LSTM, хоч і продемонструвала низький рівень відносних помилок, не змогла перевершити Holt-Winters за точністю прогнозів, і наявність автокореляції в залишках вказує на потенціал для подальшого вдосконалення.

Модель SARIMAX не змогла забезпечити достатньо точних прогнозів і має серйозні обмеження, що підтверджується її високими значеннями RMSE та MAE, а також виявленою автокореляцією в залишках.

Отже, Holt-Winters є найбільш ефективною моделлю для цього завдання, і її варто використовувати для подальших прогнозів, оскільки вона забезпечує стабільні та точні результати з мінімальними помилками та відсутністю структурних недоліків у залишках.

### **3.3 Побудова та аналіз прогнозів рівня споживання електроенергії в Україні**

Розглянемо моделювання динаміки споживання електроенергії з використанням найкращої моделі, визначеної на основі порівняння кількох підходів.

Моделювання базується на часових рядах, які відображають зміну рівня споживання електроенергії протягом часу. Застосування найкращої моделі, обраної

після аналізу залишків та порівняння метрик точності, забезпечує більш високу достовірність прогнозів та дозволяє зробити висновки щодо майбутніх тенденцій.

Побудуємо прогноз на майбутнє за допомогою моделі Holt-Winters та відобразимо результати на графіку.

У таблиці 3.4 наведено результати прогнозу на майбутній період споживання електроенергії у млн кВт/год.

Таблиця 3.4 Результати прогнозу споживання електроенергії у млн кВт/год

| Date       | Consumption |
|------------|-------------|
| 2022-02-28 | 10511       |
| 2022-03-31 | 10586       |
| 2022-04-30 | 9367        |
| 2022-05-31 | 8991        |
| 2022-06-30 | 8937        |
| 2022-07-31 | 9141        |
| 2022-08-31 | 9369        |
| 2022-09-30 | 8915        |
| 2022-10-31 | 9856        |
| 2022-11-30 | 10605       |
| 2022-12-31 | 11341       |
| 2023-01-31 | 11383       |

*Джерело: розроблено автором*

На рисунку 3.17 представлено результати моделювання споживання електроенергії, а саме історичні дані та прогноз на наступні 12 місяців.



Рисунок 3.17 Моделювання споживання електроенергії

*Джерело: розроблено автором*

З результатів зображених у таблиці 3.4 та рисунку 3.17 демонструє, що модель ефективно прогнозує споживання електроенергії, враховуючи сезонність і загальну динаміку. Цей прогноз може слугувати основою для планування ресурсів, оцінки потреб в енергетиці чи управління енергетичними процесами протягом наступного року.

## ВИСНОВКИ

Прогнозування рівня споживання електроенергії є одним із найважливіших завдань для енергетичних компаній, системних операторів та регуляторів. Це дозволяє не лише ефективно розподіляти енергетичні ресурси, але й забезпечувати безперервність та надійність постачання електроенергії споживачам. Завдяки точному прогнозуванню стає можливим оптимізувати роботу енергетичної інфраструктури, зменшити витрати на її експлуатацію, уникнути перевантаження мереж та мінімізувати ризики дефіциту енергії. Окрім цього, прогнозування споживання електроенергії слугує базою для довгострокового стратегічного планування, сприяє розвитку екологічно чистих технологій та сприяє підвищенню ефективності управління енергетичним сектором у цілому.

У даній роботі було проведено дослідження рівня споживання електроенергії та застосовано різні методи прогнозування, зокрема такі моделі, як: SARIMAX, Holt-Winters та LSTM. Порівняння цих методів показало, що найкращим для прогнозування динаміки споживання є модель Holt-Winters, яка враховує наявний тренд та сезонність.

Отримані моделі прогнозування оцінені за допомогою різних критеріїв, таких як середня квадратична помилка (RMSE), середня абсолютна помилка (MAE) та середня відносна помилка (MAPE). Також були застосовані методи перевірки залишків на наявність білого шуму та їх стандартних відхилень. Модель Holt-Winters показала найкращі значення по всіх критеріях, що підтверджує її ефективність.

У свою чергу отримані результати підтверджують ефективність використання методів прогнозування на основі часових рядів для прогнозування споживання електроенергії.

Для покращення точності прогнозування в майбутньому можуть бути враховані додаткові змінні, такі як погодні умови або економічні показники, що впливають на попит на електроенергію. Використання більш складних моделей глибинного навчання, а також інтеграція ансамблевих методів може дозволити

досягти кращих результатів. Розширення періоду навчання моделей та адаптивне коригування прогнозів в реальному часі дозволить враховувати зміни в умовах споживання.

Завдяки прогнозуванню можна також вдосконалити технології енергоспоживання, оптимізувати тарифи на електроенергію та сприяти зниженню витрат на виробництво і постачання енергії. Важливим кроком є також інтеграція прогнозних моделей з автоматизованими системами управління енергоспоживанням для оперативного реагування на зміни в попиті.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Kalimoldayev M., Marinich T. Analysis of modern approaches for the prediction of electric energy consumption. Open Engineering. 2020. Vol. 1, no. 10. P. 350—361. URL: <https://doi.org/10.1515/eng-2020-0028> (date of accessed: 29.11.2024)
2. Alon I. Time Series Model Forecasting of Boot Using Holt, Winter and Decomposition Method. Journal of Industrial Safety Engineering. 2015. Vol. 2, no. 2. P. 23—31. URL: [https://www.researchgate.net/publication/327847645\\_Time\\_Series\\_Model\\_Forecasting\\_of\\_Boot\\_Using\\_HoltWinter\\_and\\_Decomposition\\_Method/link/5ba9273d92851ca9ed2255a1/download](https://www.researchgate.net/publication/327847645_Time_Series_Model_Forecasting_of_Boot_Using_HoltWinter_and_Decomposition_Method/link/5ba9273d92851ca9ed2255a1/download) (date of accessed: 29.11.2024)
3. Pandas is a fiscally sponsored project of NumFOCUS. Pandas. URL: <https://pandas.pydata.org/> (date of accessed: 29.11.2024)
4. Python & NumPy utilities. Keras. URL: <https://keras.io> (date of accessed: 29.11.2024)
5. Прогнозування.Wiki ТНТУ. URL: <https://wiki.tntu.edu.ua/Прогнозування> (дата звернення: 29.11.2024).
6. Electricity demand forecasting with hybrid classical statistical and machine learning algorithms: Case study of Ukraine / T. González Grandón et al. Applied Energy. 2024. Vol. 355. P. 122249. URL: <https://doi.org/10.1016/j.apenergy.2023.122249> (date of access: 19.12.2024).
7. Економіко-математичні методи і моделі: практикум / Білик Т. О., Великоіваненко Г. І., Водзянова Н. К., Волощук С. Д., Матвійчук А. В., Савіна С. С., Шатарська І. Ф. – К. : КНЕУ, 2014.– 220 с
8. Isaac Kofi Nti, Moses Teimeh, Owusu Nyarko-Boateng, Adebayo Felix Adekoya. Electricity load forecasting: a systematic review. URL: <https://jesit.springeropen.com/articles/10.1186/s43067-020-00021-8> (date of accessed: 29.11.2024)
9. Роман, Грищенко. (2022). Прогнозування споживання електричної енергії електротехнічних комплексів міської електричної мережі. International

- Science Journal of Engineering & Agriculture. URL: [https://www.researchgate.net/publication/362421377\\_PROGNOZUVANNA\\_SP\\_OZIVANNA\\_ELEKTRICNOI\\_ENERGII\\_ELEKTROTEHNICNIH\\_KOMPLEKSIV\\_MISKOI\\_ELEKTRICNOI\\_MEREZI](https://www.researchgate.net/publication/362421377_PROGNOZUVANNA_SP_OZIVANNA_ELEKTRICNOI_ENERGII_ELEKTROTEHNICNIH_KOMPLEKSIV_MISKOI_ELEKTRICNOI_MEREZI) (date of accessed: 29.11.2024)
10. Gassar, A. Cha, S. Energy prediction techniques for large-scale buildings towards a sustainable built environment: A review. *Energy Build.* 2020. Vol. 8 no 12. P. 224-238.
  11. Калінчик В.П., Буравльова М.Т, Калінчик В.В., Скосирев В.Г. Прогнозування показників енергоспоживання, генерації і вартості отриманої енергії. URL: <https://doi.org/10.32838/2663-5941/2020.2-1/38> (date of accessed: 29.11.2024)
  12. Живогляд О. С. Прогнозування електроспоживання на базі нейронної мережі. URL: <https://ela.kpi.ua/server/api/core/bitstreams/3b8ef16f-3c44-49ef-8d8a-7a6447118850/content> (date of accessed: 29.11.2024)
  13. D. S. Zaruba, M. Y. Shvets, і Y. V. Khokhlov, «Машинне навчання для прогнозування споживання та генерації електроенергії», *Мікросист., Електрон. та Акуст.*, т. 24, вип. 6, с. 17–21, Груд 2019. URL: <https://doi.org/10.20535/2523-4455.2019.24.6.186996> (дата звернення: 29.11.2024)
  14. Солодка Д. М. Поняття часового ряду, його пам'яті в технічному аналізі. *Наука і молодь. Гуманітарна серія.* 2012. С. 35–38.
  15. What is Time Series? Why we Need Time Series. *EDUCBA*. URL: <https://www.educba.com/what-is-time-series/> (date of accessed: 29.11.2024).
  16. Time Series Analysis and Forecasting. Data-Driven Insights. *Analytics Vidhya*. URL: [https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-to-time-series-analysis/#What\\_Is\\_Time\\_Series\\_Analysis](https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-to-time-series-analysis/#What_Is_Time_Series_Analysis) (accessed: 29.11.2024).
  17. Часові ряди та що з ними робити. Гайд для початківців. *Dou.ua*. URL: <https://dou.ua/forums/topic/40751/> (дата звернення: 29.11.2024).
  18. Makridakis S., Spiliotis E., Assimakopoulos V. Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*. 2018. Vol. 13, no.

3. P. e0194889. URL: <https://doi.org/10.1371/journal.pone.0194889> (date of access: 29.11.2024).
19. Kalimoldayev M., Marinich T. Analysis of modern approaches for the prediction of electric energy consumption. *Open Engineering*. 2020. Vol. 1, no. 10. P. 350—361. URL: <https://doi.org/10.1515/eng-2020-0028> (date of access: 30.11.2024).
20. Hayes A. What Is a Time Series and How Is It Used to Analyze Data. Investopedia. URL: <https://www.investopedia.com/terms/t/timeseries.asp> (accessed: 30.11.2024)
21. Gupta A. (Complete Guide) Time Series Analysis: Types & Examples. Simplilearn. Simplilearn.com. URL: <https://www.simplilearn.com/tutorials/statistics-tutorial/what-is-time-series-analysis> (accessed: 30.11.2024).
22. What is Time Series Data. Definition, Examples, Types & Uses. InfluxData. URL: <https://www.influxdata.com/what-is-time-series-data/> (accessed: 30.11.2024).
23. Time Series Analysis: Definition, Types, Techniques, and When It's Used. Tableau. URL: <https://www.tableau.com/learn/articles/time-series-analysis#definition> (accessed: 30.11.2024).
24. Importance of Time Series Analysis in Data Science. Customer-Centric Business. URL: <https://customerthink.com/importance-of-time-series-analysis-in-data-science/> (accessed: 30.11.2024).
25. Gouriéroux C., Monfort A., Gallo G. *Time Series and Dynamic Models*. Cambridge University Press, 2010.
26. Time Series Analysis for Better Decision Making in Business. Outsource Market Research Services | Research Optimus. URL: <https://www.researchoptimus.com/article/what-is-time-series-analysis.php> (accessed: 30.11.2024).
27. Vignesh Arumugam, Vijayalakshmi Natarajan. Time Series Modeling and Forecasting Using Autoregressive Integrated Moving Average and Seasonal Autoregressive Integrated Moving Average Models. *Instrumentation Measure Métrologie* Vol. 22, No. 4, August, 2023, pp. 161-168 URL: <https://doi.org/10.18280/i2m.220404> (date of accessed: 30.11.2024)

28. Kontopoulou, V.I., Panagopoulos, A.D., Kakkos, I., Matsopoulos, G.K. A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. Future Internet 2023. URL: <https://doi.org/10.3390/fi15080255> (date of accessed: 30.11.2024)
29. ARIMA for Time Series Forecasting: A Complete Guide. URL: <https://www.datacamp.com/tutorial/arima> (date of accessed: 30.11.2024)
30. Cools M., Moons E. Investigating the Variability in Daily Traffic Counts through use of ARIMAX and SARIMAX Models: Assessing the Effect of Holidays on Two Site Locations. Transportation Research Record: Journal of the Transportation Research Board Journal Indexing and Metrics. 2009. Vol. 4, no. 6. P. 57—66. URL: <https://doi.org/10.3141/2136-07> (date of access: 30.11.2024).
31. Alon I. Time Series Model Forecasting of Boot Using Holt, Winter and Decomposition Method. Journal of Industrial Safety Engineering. 2015. Vol. 2, no. 2. P. 23—31. URL: [https://www.researchgate.net/publication/327847645\\_Time\\_Series\\_Model\\_Forecasting\\_of\\_Boot\\_Using\\_HoltWinter\\_and\\_Decomposition\\_Method/link/5ba9273d92851ca9ed2255a1/download](https://www.researchgate.net/publication/327847645_Time_Series_Model_Forecasting_of_Boot_Using_HoltWinter_and_Decomposition_Method/link/5ba9273d92851ca9ed2255a1/download) (date of access: 30.11.2024).
32. Hyndman R., Athanasopoulos G. Forecasting: principles and practice : підручник. 2nd ed. Australia : Melbourne, 2018. 292 p. URL: <https://otexts.com/fpp2/> (date of access: 30.11.2024).
33. Spall J. Bayesian Analysis of Time Series and Dynamic Models : підручник. 978th ed. Лондон : CRC Press, 1988. 576 p.
34. Ana Paula Miranda Diniz, Patrick Marques Ciarelli, Evandro Ottoni Teatini Salles, Klaus Fabian Coco. Use of deep neural networks for clogging detection in the Submerged Entry Nozzle of the continuous casting. URL: <https://doi.org/10.1016/j.eswa.2023.121963> (date of access: 30.11.2024)
35. Електроенергетика України: стан і перспективи. URL: <https://blog.youcontrol.market/ieliektroienierghietika-ukrayini-stan-i-pierspiektivi/> (date of access: 30.11.2024)

- 36.Динаміка і структура споживання електроенергії в Україні. URL: <https://uaea.com.ua/dysp/ee-cons.html> (date of access: 30.11.2024)
37. Energy Map. Обсяг споживання електроенергії за групами споживачів URL: <https://map.ua-energy.org/uk/resources/742384e9-83c7-44f4-8c0b-0cd74b56561b/> (date of access: 30.11.2024)
- 38.Peixeiro M. Time Series Forecasting in Python, 2022.
- 39.Andres D. Classification of Advanced Time Series Forecasting Methods and when to use them. Medium. URL: <https://medium.datadriveninvestor.com/classification-of-advanced-time-series-forecasting-methods-and-when-to-use-them-51709e683b48> (date of access: 01.12.2024).
- 40.Cognitive Data Science - Bridging Minds and Technology. IndiaAI. URL: <https://indiaai.gov.in/article/cognitive-data-science-bridging-minds-and-technology> (date of access: 01.12.2024).
- 41.Machine learning algorithms. Machine Learning For All. URL: <https://mlforall.wordpress.com/2019/09/24/machine-learning-algorithms/> (date of access: 01.12.2024).
- 42.Chapter 1 Introduction to Time Series | Time Series for Beginners. *Home | Bookdown*. URL: <https://bookdown.org/JakeEsprabens/431-Time-Series/introduction-to-time-series.html> (date of access: 01.12.2024).
- 43.Нейромереживо. Аналіз та прогнозування часових рядів. *Друкарня*. URL: <https://drukarnia.com.ua/articles/analiz-ta-prognozuvannya-chasovikh-ryadiv-H6gg0> (дата звернення: 01.12.2024).
- 44.Мещерякова К. Нейромережа: що це, як працює, найкращі приклади. *High Bar Journal*. URL: <https://journal.gen.tech/post/sho-take-neiromerezhi> (дата звернення: 01.12.2024).

## ДОДАТОК А

### Перелік використаних бібліотек

```
from google.colab import files
import numpy as np
import pmdarima as pm
import matplotlib.pyplot as plt
import math
from sklearn.metrics import mean_squared_error, mean_absolute_error
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from statsmodels.stats.diagnostic import acorr_ljungbox
```

### Імпорт датасету у середовище Google Colab

```
uploaded = files.upload()
data = pd.read_excel('/content/All_types.xlsx', sheet_name='Дані')
data.head()
```

### Обробка даних

```
# Перевірка наявності пропущених значень у даних
missing_values = data.isnull().sum()

# Виведення інформації про пропущені значення
print("Кількість пропущених значень у кожному стовпці:")
print(missing_values)

# Перевірка, чи є пропуски взагалі
if missing_values.sum() > 0:
    print("\nПропущені значення знайдено.")
else:
    print("\nПропущених значень немає.")
```

```
# Візуалізація даних
plt.figure(figsize=(10, 6))
plt.plot(data.consumption)
plt.title('Time Series Data')
plt.xlabel('Date')
plt.ylabel('Consumption')
plt.show()
```

```
# Графічний аналіз автокореляції (ACF)
plot_acf(data.consumption)
plt.show()
```

```
# Сезонна декомпозиція часових рядів
```

```
decomposition = seasonal_decompose(data.consumption, model='additive',
period=12)
decomposition.plot()
plt.show()
```

```
# Ljung-Box Test
ljung_box = acorr_ljungbox(data.consumption, lags=12)
print(ljung_box)
```

## SARIMAX

```
# Підготовка даних
data['Date'] = pd.to_datetime(data['Date'])
data.set_index('Date', inplace=True)
data['consumption'] = pd.to_numeric(data['consumption'], errors='coerce')
y = data['consumption']
```

```
# Розділення на навчальну та тестову вибірку
train_size = int(len(y) * 0.8) # 80% для навчальної вибірки
train, test = y[:train_size], y[train_size:]
```

```
# Побудова моделі AutoARIMA на навчальній вибірці
model = pm.auto_arima(train, seasonal=True, m=12, stepwise=True, trace=True,
suppress_warnings=True)
```

```
# Підсумок моделі
print("Параметри моделі:")
print(model.summary())
```

```
# Прогноз на тестових даних
forecast = model.predict(n_periods=len(test))
```

```
# Оцінка моделі: RMSE, MAE, MAPE
rmse = math.sqrt(mean_squared_error(test, forecast))
mae = mean_absolute_error(test, forecast)
mpe = np.mean((test - forecast) / test) * 100

print(f"RMSE: {rmse}")
print(f"MAE: {mae}")
print(f"MPE: {mpe}")
```

```
# Формування таблиці результатів
results = pd.DataFrame({
    'Date': test.index,
    'Actual': test.values,
    'Forecast': forecast
})
```

```

}))
results.set_index('Date', inplace=True)
print("\nПерші 10 рядків результатів:")
print(results.head(10))

# Візуалізація результатів прогнозування
plt.figure(figsize=(12, 6))
plt.plot(train.index, train, label='Навчальні дані', color='blue')
plt.plot(test.index, test, label='Фактичні тестові дані', color='green')
plt.plot(test.index, forecast, label='Прогнозовані дані', color='orange',
         linestyle='--')
plt.title('Прогноз за моделлю ARIMA')
plt.xlabel('Дата')
plt.ylabel('Споживання')
plt.legend()
plt.grid(True)
plt.show()

```

```

# Обчислення залишків
residuals = test - forecast
# Ljung-Box тест
ljung_box_test = acorr_ljungbox(residuals, lags=[12], return_df=True)
print("Ljung-Box Test Results:")
print(ljung_box_test)

```

```

# Обчислення стандартного відхилення залишків (Residual SD)
residual_sd = np.std(residuals)
print(f"Standard Deviation of Residuals (Residual SD): {residual_sd:.4f}")

```

## Holt-Winters

```

# Підготовка даних
data['Date'] = pd.to_datetime(data['Date'])
data.set_index('Date', inplace=True)
data['consumption'] = pd.to_numeric(data['consumption'], errors='coerce')

# Розділення на тренувальну та тестову вибірки
train_size = int(len(data) * 0.8)
train, test = data.iloc[:train_size], data.iloc[train_size:]

# Побудова та автоматичне налаштування моделі Хольта-Вінтера
model = ExponentialSmoothing(
    train['consumption'],
    trend='mul',
    seasonal='add',
    seasonal_periods=12
).fit(optimized=True) # Використання автоматичного налаштування

```

```

# Прогноз на тестову вибірку
test_pred = model.forecast(len(test))

# Оцінка якості моделі
mae = mean_absolute_error(test['consumption'], test_pred)
rmse = mean_squared_error(test['consumption'], test_pred, squared=False)
mpe = np.mean((test['consumption'] - test_pred) / test['consumption']) * 100

print("Показники якості моделі на тестових даних:")
print(f"MAE: {mae}")
print(f"RMSE: {rmse}")
print(f"MPE: {mpe}")

# Параметри моделі
print("Аналітичний вигляд моделі:")
print(f"Тип тренду: {model.model.trend}")
print(f"Тип сезонності: {model.model.seasonal}")
print(f"Кількість сезонних періодів: {model.model.seasonal_periods}")

print("\nКоефіцієнти моделі:")
print(f"Рівень ( $\alpha$ ): {model.params['smoothing_level']:.4f}")
print(f"Тренд ( $\beta$ ): {model.params.get('smoothing_slope', 'N/A')}")
print(f"Сезонність ( $\gamma$ ): {model.params['smoothing_seasonal']:.4f}")
print(f"Початковий рівень: {model.params['initial_level']:.4f}")
print(f"Початковий тренд: {model.params.get('initial_slope', 'N/A')}")

# Візуалізація результатів прогнозу
plt.figure(figsize=(12, 6))
plt.plot(train.index, train['consumption'], label='Навчальні дані',
color='blue')
plt.plot(test.index, test['consumption'], label='Тестові дані', color='green')
plt.plot(test.index, test_pred, label='Прогнозовані значення', color='orange',
linestyle='--')
plt.title('Прогноз за моделлю Хольта-Вінтерса')
plt.xlabel('Дата')
plt.ylabel('Споживання')
plt.legend()
plt.grid(True)
plt.show()

# Обчислення залишків
residuals = test['consumption'].values - test_pred.values
# Ljung-Box тест
ljung_box_test = acorr_ljungbox(residuals, lags=[12], return_df=True)
print("Ljung-Box Test Results:")
print(ljung_box_test)

```

```

# Обчислення стандартного відхилення залишків (Residual SD)
residual_sd = np.std(residuals)

print(f"Standard Deviation of Residuals (Residual SD): {residual_sd:.4f}")

# Прогноз на наступні 12 місяців
forecast = model.forecast(36)

# Створення таблиці з прогнозом на 12 місяців вперед
forecast_df = pd.DataFrame({
    'Прогноз (12 місяців вперед)': forecast
}, index=pd.date_range(start=data.index[-1] + pd.Timedelta(days=1),
periods=12, freq='M'))

# Виведення таблиці
print("\nПрогноз на наступні 12 місяців:")
print(forecast_df)

# Візуалізація результату
forecast_df['Дата'] = forecast_df.index
original_data = data['consumption']

# Створення графіка
plt.figure(figsize=(10, 6))

# Побудова графіка для оригінальних даних
plt.plot(original_data.index, original_data, label='Оригінальні дані',
color='blue')

# Побудова графіка для прогнозу
plt.plot(forecast_df['Дата'], forecast_df['Прогноз (12 місяців вперед)'],
label='Прогноз (12 місяців)', color='red', linestyle='--')

# Додавання підписів до осей
plt.xlabel('Дата')
plt.ylabel('Споживання')

# Додавання заголовка та легенди
plt.title('Прогноз на наступні 12 місяців')
plt.legend()

# Відображення графіка
plt.grid(True)
plt.show()

```

## LSTM

```

# Підготовка даних
data.set_index('Date', inplace=True)

```

```
# Масштабування даних (Min-Max нормалізація)
scaler = MinMaxScaler(feature_range=(0, 1))
data['consumption'] = scaler.fit_transform(data[['consumption']])
```

```
# Функція для створення послідовностей
def create_sequences(data, time_steps=12):
    X, y = [], []
    for i in range(len(data) - time_steps):
        X.append(data[i:i + time_steps])
        y.append(data[i + time_steps])
    return np.array(X), np.array(y)
```

```
# Створення тренувальних і тестових вибірок
time_steps = 12 # Наприклад, 12 місяців для річної сезонності
X, y = create_sequences(data['consumption'].values, time_steps)
split_index = int(len(X) * 0.8)
X_train, y_train = X[:split_index], y[:split_index]
X_test, y_test = X[split_index:], y[split_index:]
```

```
# Перетворення даних у формат (samples, time_steps, features)
X_train = X_train.reshape((X_train.shape[0], X_train.shape[1], 1))
X_test = X_test.reshape((X_test.shape[0], X_test.shape[1], 1))
```

```
# Функція для створення моделі LSTM з оптимізацією параметрів
def build_model(hp):
    model = Sequential()

    # Оптимізація кількості нейронів у LSTM шарі
    model.add(LSTM(units=hp.Int('units', min_value=32, max_value=128,
step=32),
                    activation='relu',
                    input_shape=(X_train.shape[1], 1)))

    # Додавання додаткових шарів (опціонально)
    if hp.Boolean('add_dense'):
        model.add(Dense(units=hp.Int('dense_units', min_value=32,
max_value=128, step=32), activation='relu'))

    model.add(Dense(1)) # Вихідний шар
    model.compile(optimizer=Adam(learning_rate=hp.Float('learning_rate',
min_value=1e-5, max_value=1e-2, sampling='LOG')),
                  loss='mean_squared_error')

    return model
```

```
# Створення і налаштування Tuner для пошуку кращих гіперпараметрів
```

```
tuner = kt.Hyperband(build_model, objective='val_loss', max_epochs=10,
factor=3, directory='my_dir', project_name='lstm_tuning')
```

```
# Виконання пошуку кращих гіперпараметрів
tuner.search(X_train, y_train, epochs=10, validation_data=(X_test, y_test))
```

```
# Отримання найкращої моделі
best_model = tuner.get_best_models(num_models=1)[0]
```

```
# Виведення аналітичного вигляду моделі
print("\nАрхітектура найкращої моделі LSTM:")
best_model.summary()
```

```
# Прогнозування на тестових даних
y_pred_scaled = best_model.predict(X_test)
y_pred = scaler.inverse_transform(y_pred_scaled)
y_test_actual = scaler.inverse_transform(y_test.reshape(-1, 1))
```

```
# Оцінка моделі
from sklearn.metrics import mean_squared_error, mean_absolute_error
mse = mean_squared_error(y_test_actual, y_pred)
rmse = np.sqrt(mse)
mae = mean_absolute_error(y_test_actual, y_pred)
mpe = np.mean((y_test_actual - y_pred) / y_test_actual) * 100

print(f"Root Mean Squared Error: {rmse}")
print(f"Mean Absolute Error: {mae}")
print(f"Mean Percentage Error: {mpe}")
```

```
# Таблиця результатів прогнозування
results = pd.DataFrame({
    'Actual': y_test_actual.flatten(),
    'Predicted': y_pred.flatten()
})
print("\nПерші 10 результатів прогнозування:")
print(results.head(10))
```

```
# Візуалізація результатів прогнозування
plt.figure(figsize=(14, 7))
```

```
# Навчальні дані
plt.plot(range(len(y[:split_index])),
scaler.inverse_transform(y[:split_index].reshape(-1, 1)),
label='Навчальні дані', color='blue')
```

```

# Тестові дані
plt.plot(range(len(y[split_index:])), y_test_actual,
         label='Тестові дані', color='green')

# Прогнозовані дані
plt.plot(range(len(y[split_index:])), y_pred,
         label='Прогнозовані значення', color='orange', linestyle='--')

# Налаштування графіка
plt.title('Прогноз за моделлю LSTM')
plt.xlabel('Час (порядковий номер)')
plt.ylabel('Споживання')
plt.legend()
plt.grid(True)
plt.show()

```

```

# Обчислення залишків (резидуалів)
residuals = y_test_actual.flatten() - y_pred.flatten()

```

```

# Ljung-Box тест
ljung_box_test = acorr_ljungbox(residuals, lags=[12], return_df=True)
print("Ljung-Box Test Results:")
print(ljung_box_test)

```

```

# Обчислення стандартного відхилення залишків (Residual SD)
residual_sd = np.std(residuals)

print(f"Standard Deviation of Residuals (Residual SD): {residual_sd:.4f}")

```



## Звіт подібності

### метадані

Заголовок

Вікторова\_Катуніна\_плагіат

Автор

Науковий керівник / Експерт

Вікторова

Катуніна

підрозділ

кафедра математичного моделювання та статистики

### Тривога

У цьому розділі ви знайдете інформацію щодо текстових спотворень. Ці спотворення в тексті можуть говорити про МОЖЛИВІ маніпуляції в тексті. Спотворення в тексті можуть мати навмисний характер, але частіше характер технічних помилок при конвертації документа та його збереженні, тому ми рекомендуємо вам підходити до аналізу цього модуля відповідально. У разі виникнення запитань, просимо звертатися до нашої служби підтримки.

|                        |                                                                                   |    |
|------------------------|-----------------------------------------------------------------------------------|----|
| Заміна букв            |  | 3  |
| Інтервали              |  | 0  |
| Мікропробіли           |  | 8  |
| Білі знаки             |  | 0  |
| Парафрази (SmartMarks) |  | 25 |

### Обсяг знайдених подібностей

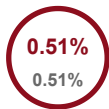
Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.



КП 1

25

Довжина фрази для коефіцієнта подібності 2



КП 2

12042

Кількість слів



КЦ

92196

Кількість символів

### Подібності за списком джерел

Нижче наведений список джерел. В цьому списку є джерела із різних баз даних. Колір тексту означає в якому джерелі він був знайдений. Ці джерела і значення Коефіцієнту Подібності не відображають прямого плагіату. Необхідно відкрити кожне джерело і проаналізувати зміст і правильність оформлення джерела.

#### 10 найдовших фраз

Колір тексту

| ПОРЯДКОВИЙ<br>НОМЕР | НАЗВА ТА АДРЕСА ДЖЕРЕЛА URL (НАЗВА БАЗИ)                                                                                                                                                                                     | КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ<br>(ФРАГМЕНТІВ) |        |
|---------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------|--------|
| 1                   | <a href="https://ir.kneu.edu.ua/bitstreams/151df6cf-1099-4401-925c-5a0fe3423590/download">https://ir.kneu.edu.ua/bitstreams/151df6cf-1099-4401-925c-5a0fe3423590/download</a>                                                | 32                                        | 0.27 % |
| 2                   | <a href="https://ir.kneu.edu.ua/bitstreams/151df6cf-1099-4401-925c-5a0fe3423590/download">https://ir.kneu.edu.ua/bitstreams/151df6cf-1099-4401-925c-5a0fe3423590/download</a>                                                | 29                                        | 0.24 % |
| 3                   | <a href="#">Прогнозування попиту на електроенергію за допомогою аналізу часових рядів</a><br>12/16/2022<br>Kyiv National Economic University named after Vadym Hetman KNEU (кафедра математичного моделювання та статистики) | 24                                        | 0.20 % |
| 4                   | <a href="http://ebib.pp.ua/vremennoy-ryad-vidyi-vremennyih-ryadov-7564.html">http://ebib.pp.ua/vremennoy-ryad-vidyi-vremennyih-ryadov-7564.html</a>                                                                          | 17                                        | 0.14 % |

|    |                                                                                                                                                                                                                              |    |        |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|--------|
| 5  | <a href="#">Прогнозування попиту на електроенергію за допомогою аналізу часових рядів</a><br>12/16/2022<br>Kyiv National Economic University named after Vadym Hetman KNEU (кафедра математичного моделювання та статистики) | 16 | 0.13 % |
| 6  | <a href="#">Клименко Горна плагіат</a><br>5/29/2024<br>Kyiv National Economic University named after Vadym Hetman KNEU (кафедра математичного моделювання та статистики)                                                     | 16 | 0.13 % |
| 7  | <a href="https://elibrary.kdpu.edu.ua/bitstream/123456789/6918/1/Magistr_Potapenko_lm-17.pdf">https://elibrary.kdpu.edu.ua/bitstream/123456789/6918/1/Magistr_Potapenko_lm-17.pdf</a>                                        | 14 | 0.12 % |
| 8  | <a href="#">Прогнозування попиту на електроенергію за допомогою аналізу часових рядів</a><br>12/16/2022<br>Kyiv National Economic University named after Vadym Hetman KNEU (кафедра математичного моделювання та статистики) | 14 | 0.12 % |
| 9  | <a href="#">Прогнозування попиту на електроенергію за допомогою аналізу часових рядів</a><br>12/16/2022<br>Kyiv National Economic University named after Vadym Hetman KNEU (кафедра математичного моделювання та статистики) | 13 | 0.11 % |
| 10 | ФАЕТ_151_AKIK_2023_Стельмах_БО<br>7/10/2024<br>Ukrainian national aviation university (Ukrainian national aviation university)                                                                                               | 13 | 0.11 % |

з бази даних RefBooks (0.00 %)

| ПОРЯДКОВИЙ НОМЕР | ЗАГОЛОВОК | КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ) |
|------------------|-----------|----------------------------------------|
|------------------|-----------|----------------------------------------|

з домашньої бази даних (1.29 %)

| ПОРЯДКОВИЙ НОМЕР | ЗАГОЛОВОК                                                                                                                                                                                                                    | КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ) |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|
| 1                | <a href="#">Прогнозування попиту на електроенергію за допомогою аналізу часових рядів</a><br>12/16/2022<br>Kyiv National Economic University named after Vadym Hetman KNEU (кафедра математичного моделювання та статистики) | 99 (8) 0.82 %                          |
| 2                | <a href="#">Клименко Горна плагіат</a><br>5/29/2024<br>Kyiv National Economic University named after Vadym Hetman KNEU (кафедра математичного моделювання та статистики)                                                     | 37 (3) 0.31 %                          |
| 3                | <a href="#">Прогнозування цін на пальне в Україні</a><br>12/16/2022<br>Kyiv National Economic University named after Vadym Hetman KNEU (кафедра математичного моделювання та статистики)                                     | 19 (2) 0.16 %                          |

з програми обміну базами даних (0.57 %)

| ПОРЯДКОВИЙ НОМЕР | ЗАГОЛОВОК                                                                                                                                                                         | КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ) |
|------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|
| 1                | ФАЕТ_151_AKIK_2023_Стельмах_БО<br>7/10/2024<br>Ukrainian national aviation university (Ukrainian national aviation university)                                                    | 25 (3) 0.21 %                          |
| 2                | ПРОГНОЗУВАННЯ СПОЖИВАННЯ ЕНЕРГІЇ: ПОРІВНЯЛЬНИЙ АНАЛІЗ МОДЕЛЕЙ ЧАСОВИХ РЯДІВ<br>12/16/2020<br>Sumy State University (Кафедра прикладної математики та моделювання складних систем) | 24 (3) 0.20 %                          |

|   |                                                                                                                                                                         |        |        |
|---|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------|
| 3 | МОДЕЛЮВАННЯ ВПЛИВУ СОЦІАЛЬНИХ МЕРЕЖ НА ДИНАМІКУ РИНКУ КРИПТОВАЛЮТ<br>5/28/2021<br>Sumsy State University (Кафедра прикладної математики та моделювання складних систем) | 11 (1) | 0.09 % |
| 4 | 2023_61130000_Khasyshyn_Nadiia_Mykhailivna_118007<br>11/21/2024<br>National University "Lviv Politechnika" (National University Lviv Politechnika)                      | 9 (1)  | 0.07 % |

### з Інтернету (1.19 %)

| ПОРЯДКОВИЙ<br>НОМЕР | ДЖЕРЕЛО URL                                                                                                                                                                           | КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ<br>(ФРАГМЕНТІВ) |        |
|---------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------|--------|
| 1                   | <a href="https://ir.kneu.edu.ua/bitstreams/151df6cf-1099-4401-925c-5a0fe3423590/download">https://ir.kneu.edu.ua/bitstreams/151df6cf-1099-4401-925c-5a0fe3423590/download</a>         | 66 (3)                                    | 0.55 % |
| 2                   | <a href="https://elibrary.kdpu.edu.ua/bitstream/123456789/6918/1/Magistr_Potapenko_lm-17.pdf">https://elibrary.kdpu.edu.ua/bitstream/123456789/6918/1/Magistr_Potapenko_lm-17.pdf</a> | 39 (4)                                    | 0.32 % |
| 3                   | <a href="http://ebib.pp.ua/vremennoy-ryad-vidyi-vremennyih-ryadov-7564.html">http://ebib.pp.ua/vremennoy-ryad-vidyi-vremennyih-ryadov-7564.html</a>                                   | 17 (1)                                    | 0.14 % |
| 4                   | <a href="https://essuir.sumdu.edu.ua/bitstream-download/123456789/94448/1/Psarov_mag_rob.pdf">https://essuir.sumdu.edu.ua/bitstream-download/123456789/94448/1/Psarov_mag_rob.pdf</a> | 11 (1)                                    | 0.09 % |
| 5                   | <a href="https://bpci.kiev.ua/rada/aref/autoref_Metelitsa.pdf">https://bpci.kiev.ua/rada/aref/autoref_Metelitsa.pdf</a>                                                               | 10 (1)                                    | 0.08 % |

### Список прийнятих фрагментів (немає прийнятих фрагментів)

| ПОРЯДКОВИЙ НОМЕР | ЗМІСТ | КІЛЬКІСТЬ ОДНАКОВИХ СЛІВ (ФРАГМЕНТІВ) |
|------------------|-------|---------------------------------------|
|------------------|-------|---------------------------------------|

# ПРОГНОЗУВАННЯ СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ В УКРАЇНІ

**Вікторова Ярослава Сергіївна**

здобувач вищої освіти інституту інформаційних технологій в економіці  
*Київський національний економічний університет ім. В. Гетьмана, Україна*

**Науковий керівник: Катуніна Ольга Сергіївна**

канд. екон. наук, доцент, доцент кафедри математичного моделювання та  
статистики

*Київський національний економічний університет ім. В. Гетьмана, Україна*

Поточний стан енергосистеми України демонструє нестабільність енергопостачання внаслідок пов'язаного з військовими діями значного пошкодження енергетичної інфраструктури. Понад 70% енергетичних об'єктів, зокрема атомні та теплові електростанції, зазнали серйозних пошкоджень, що ускладнює забезпечення стабільного постачання електроенергії [1]. Це створює високий ризик виникнення дефіциту енергетичних потужностей, особливо в зимовий період. Відновлення пошкоджених об'єктів потребує значних інвестицій та часу, що підвищує навантаження на державний бюджет і знижує здатність системи реагувати на критичні ситуації.

Інтеграція української енергосистеми до європейської мережі ENTSO-E стала важливим кроком у напрямку стабільності енергопостачання. З 2022 року було здійснено синхронізацію з європейськими енергосистемами, що дозволило підвищити надійність енергопостачання в Україні. Однак, існуючі потужності між Україною та Європою не здатні забезпечити повний обсяг імпорту електроенергії в умовах пікового попиту [2]. Підключення до європейської мережі створює додаткові можливості для збільшення імпорту, але обмежена пропускна здатність та відсутність інфраструктури, яка дозволяє оптимізувати цей процес, залишається значною проблемою для української енергетики.

Війна значно вплинула на відновлювальну енергетику України. Більше 14% сонячних електростанцій були пошкоджені, а понад 950 МВт потужностей знаходяться на окупованих територіях. Це призвело до значного зниження

виробництва електроенергії з відновлювальних джерел. Зокрема, станції на півдні та сході України зазнали руйнувань через обстріли, що ускладнило відновлення та подальший розвиток цієї галузі. Тому відновлення інфраструктури та фінансова стабільність стають критичними для майбутнього розвитку ВДЕ в Україні [3]. Однією з головних проблем є діючі цінові обмеження на ринку електроенергії, які суттєво ускладнюють імпорт електрики. Встановлені нижчі граничні ціни в Україні значно знижують економічну вигоду від імпорту, оскільки ціни на електроенергію на європейському ринку часто значно вищі. Це призводить до того, що постачання електроенергії з-за кордону не є вигідним для інвесторів та компаній, що займаються імпортом, навіть у випадку гострої потреби в електриці. За словами експертів, скасування цінових обмежень могло б істотно покращити ситуацію та забезпечити енергетичну стабільність, особливо в зимовий період, коли попит на електроенергію перевищує можливості внутрішнього виробництва [4]. Таким чином, для досягнення енергетичної безпеки та стабільності національної енергосистеми необхідно запровадити комплексні заходи, спрямовані на відновлення та модернізацію інфраструктури, збільшення частки відновлювальних джерел енергії та зміну регуляторної політики, зокрема в частині цінової політики на імпорт електроенергії. Важливим є використання сучасних методів моделювання, які дозволяють прогнозувати динаміку попиту.

Система моделей для прогнозування споживання електроенергії була створена на основі аналізу часових рядів із використанням щомісячних даних про споживання (у млн кВт·год). Спочатку здійснено підготовку даних, що включала нормалізацію, перевірку на пропуски, аналіз сезонності та трендів. На цьому етапі були застосовані графічні методи та оцінки автокореляційних функцій. Далі були обрані три основні моделі для прогнозування: SARIMAX, яка дозволяє враховувати сезонні коливання та зовнішні фактори, Хольта-Вінтерса, що ефективно фокусується на сезонності та трендах, а також LSTM, яка використовує нейронні мережі для моделювання довгострокових залежностей. Моделі були оцінені за метриками RMSE (корінь середньоквадратичної

помилки), MAE (середня абсолютна помилка) та MPE (середній відсоток помилки). Дані були поділені на тренувальні та тестові набори для проведення валідації, що дозволило оцінити ефективність моделей на нових, раніше невідомих даних [5].

Аналіз показав, що серед усіх застосованих моделей найкращі результати продемонструвала модель Хольта-Вінтерса, яка показала високу точність при використанні щомісячних даних. Модель ефективно врахувала сезонність і тренди, що дозволило здійснити точний прогноз споживання електроенергії на горизонті 6-12 місяців. Модель SARIMAX використовувалася для аналізу сезонних змін у споживанні електроенергії, зокрема для виявлення повторюваних патернів у даних. Незважаючи на це, модель продемонструвала обмежену точність при прогнозуванні значних відхилень, оскільки чутливо реагувала на зміни структури даних, спричинені позасезонними факторами. Модель LSTM показала значний потенціал для довгострокового прогнозування завдяки здатності моделювати нелінійні залежності в часових рядах. Проте її ефективність була обмежена через недостатній обсяг даних для навчання, що вплинуло на точність у короткостроковому горизонті. Отримані результати дозволяють рекомендувати інтеграцію різних методів прогнозування для забезпечення високої точності моделей, які можуть бути використані для адаптації до різних викликів у роботі енергосистеми України, зокрема в періоди дефіциту, економічних потрясінь та сезонних коливань попиту.

### **Список використаних джерел:**

1. В Україні зруйновані, пошкоджені або захоплені понад 70% усієї генерації - експерт. URL: <https://www.ukrinform.ua/rubric-economy/3861654-v-ukraini-zrujnovani-poskodzeni-abo-zahopleni-ponad-70-usiei-generacii-ekspert.html> (дата звернення: 07.12.2024)
2. Українську енергосистему синхронізовано з європейською енергомережею. URL: <https://energysecurityua.org/ua/novyny/ukrainska-enerhosystema-zavershyla-protses-synkhronizatsii-z-ievropeyskoiu-enerhomerezheiu/> (дата звернення: 07.12.2024).

3. Ситуація у галузі. URL: <https://www.aseu.solar/about-6> (дата звернення: 07.12.2024).

4. Цінові обмеження стримують імпорт електроенергії з ЄС, – Харченко. URL: <https://www.obozrevatel.com/ukr/ekonomika-glavnaya/tsinovi-obmezhennya-strimuyut-import-elektroenergii-z-es-harchenko.htm> (дата звернення: 07.12.2024).

5. Peixeiro M. Time Series Forecasting in Python, 2022.



IN 82027  
від 13.12.2024

# СЕРТИФІКАТ

учасника конференції



ВІКТОРОВА ЯРОСЛАВА СЕРГІЙІВНА

ВЗЯВ(-ЛА) УЧАСТЬ У VIII МІЖНАРОДНІЙ СТУДЕНТСЬКІЙ НАУКОВІЙ КОНФЕРЕНЦІЇ  
АКТУАЛЬНІ ПИТАННЯ ТА ПЕРСПЕКТИВИ  
ПРОВЕДЕННЯ НАУКОВИХ ДОСЛІДЖЕНЬ

13 ГРУДНЯ 2024 РІК • М. ДНІПРО, УКРАЇНА

В рамках участі було опубліковано тези доповіді учасника на тему:

ПРОГНОЗУВАННЯ СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ В  
УКРАЇНІ

ДИРЕКТОР МОЛОДІЖНОЇ НАУКОВОЇ ЛІГИ  
ГОЛОВА ОПТКОМІТЕТУ КОНФЕРЕНЦІЇ  
ІГОР КОРЕНЮК

