

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ПРИРОДОКОРИСТУВАННЯ
ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

КВАЛІФІКАЦІЙНА РОБОТА

другого (магістерського) рівня вищої освіти

на тему: «Підвищення ефективності процесу узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання»

Виконав: студент групи Іт-61

Спеціальності 126 «Інформаційні системи та технології»

(шифр і назва)

Тхір Андрій Дмитрович

(Прізвище та ініціали)

Керівник: к.т.н., в.о. доцента Татомир А.В.

(Прізвище та ініціали)

Рецензент: к.т.н., доцент Бабич М.І.

(Прізвище та ініціали)

ДУБЛЯНИ-2024

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ПРИРОДОКОРИСТУВАННЯ
ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Другий (магістерський) рівень вищої освіти
Спеціальність 126 «Інформаційні системи та технології»

«ЗАТВЕРДЖУЮ»

Завідувач кафедри _____

д.т.н., проф. А.М. Тригуба

« ____ » _____ 2023 р.

ЗАВДАННЯ

на кваліфікаційну роботу студенту

Тхіру Андрію Дмитровичу

1. Тема роботи: «Підвищення ефективності процесу узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання»

Керівник роботи Татомир Андрій Володимирович, в.о. доцента
затверджені наказом по університету від 30.06.2023 року № 137/к-с.

2. Строк подання студентом роботи 10.01.2024 р.

3. Вихідні дані до роботи: дані автоматизованого машинного навчання; методика дослідження моделей машинного навчання.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які необхідно розробити) _____

Вступ.

1. Аналіз стану питання в теорії та практиці.

2. Вибір засобів та підготовка даних для узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання.

3. Результати розробки моделей прогнозування обсягів резервного виробництва електроенергії автономними джерелами на основі використанні бібліотеки Fedot.

4. Охорона праці та безпека у надзвичайних ситуаціях.

5. Визначення ефективності запропонованих розробок.

Висновки та пропозиції.

Список використаної літератури.

5. Перелік ілюстраційного матеріалу (з точним зазначенням обов'язкових слайдів): аналіз стану питання в теорії та практиці; вибір засобів та підготовка даних для узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання; результати розробки моделей прогнозування обсягів резервного виробництва електроенергії автономними джерелами на основі використанні бібліотеки Fedot; економічна ефективність.

6. Консультанти з розділів:

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1, 2, 3, 5	<i>Татомир А.В., к.т.н., в.о. доцента кафедри інформаційних технологій</i>		
4	<i>Городецький І.М., доцент кафедри фізики, інженерної механіки та безпеки виробництва</i>		

7. Дата видачі завдання

30 червня 2023 р.

Календарний план

№ з/п	Назва етапів кваліфікаційної роботи	Терміни виконання етапів роботи	Примітка
1	<i>Написання першого розділу</i>	<i>28.04-20.05.23</i>	
2	<i>Виконання другого розділу та аркушів ілюстраційного матеріалу до нього</i>	<i>21.05-14.08.23</i>	
3.	<i>Виконання третього розділу та аркушів ілюстраційного матеріалу до нього</i>	<i>15.08-10.11.23</i>	
4.	<i>Написання розділу «Охорона праці та безпека у надзвичайних ситуаціях»</i>	<i>11.11-30.11.23</i>	
5.	<i>Оцінення ефективності запропонованої системи</i>	<i>01.12-20.12.23</i>	
6.	<i>Завершення оформлення розрахунково-пояснювальної записки та аркушів ілюстраційного матеріалу</i>	<i>21-31.12.23</i>	

Студент _____ Тхір А.Д.
(підпис)

Керівник роботи _____ Татомир А.В.
(підпис)

УДК 621.31:006.6:004.9

Підвищення ефективності процесу узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання.

Тхір А.Д. Кафедра інформаційних технологій – Дубляни, ЛНУП, 2024.

Кваліфікаційна робота: 71 с. текст. част., 26 рис., 3 табл., 10 арк. ілюстраційного матеріалу, 42 джерел.

Виконано аналіз стану предметної області та узгодження обсягів виробництва електроенергії. Обґрунтовано доцільність використання машинного навчання для процесів узгодження обсягів резервного виробництва електроенергії. Проведено аналіз моделей машинного навчання для процесів узгодження обсягів резервного виробництва електроенергії. Подано інструментарій автоматичного машинного навчання AutoML. Обґрунтування доцільності узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі AutoML.

Подано можливі сценарії машинного навчання. Виконано аналіз існуючих AutoML фреймворків та обґрунтування його вибору. Подано особливості автоматичного машинного навчання для часових рядів. Проаналізовано Open-Source фреймворк Fedot для AutoML.

Підготовлено дані для узгодження обсягів резервного виробництва електроенергії автономними джерелами.

Виконано аналіз часових рядів обсягів резервного виробництва електроенергії. Подано результати прогнозування обсягів виробництва електроенергії генератором та вітроустановками. Наведено результати визначення раціональної моделі прогнозування обсягів виробництва електроенергії.

Розроблено заходи із охорони праці та визначено ефективність від розробок.

ЗМІСТ

ВСТУП	7
РОЗДІЛ 1. АНАЛІЗ СТАНУ ПИТАННЯ В ТЕОРІЇ ТА ПРАКТИЦІ.....	9
1.1. Аналіз стану предметної області та узгодження обсягів виробництва електроенергії	9
1.2. Доцільність використання машинного навчання для процесів узгодження обсягів резервного виробництва електроенергії.....	13
1.3. Аналіз моделей машинного навчання для процесів узгодження обсягів резервного виробництва електроенергії	17
1.4. Інструментарій автоматичного машинного навчання automl.....	20
1.5. Обґрунтування доцільності узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання.....	23
РОЗДІЛ 2. ВИБІР ЗАСОБІВ ТА ПІДГОТОВКА ДАНИХ ДЛЯ УЗГОДЖЕННЯ ОБСЯГІВ РЕЗЕРВНОГО ВИРОБНИЦТВА ЕЛЕКТРОЕНЕРГІЇ АВТОНОМНИМИ ДЖЕРЕЛАМИ НА ОСНОВІ АВТОМАТИЗОВАНОГО МАШИННОГО НАВЧАННЯ	25
2.1. Можливі сценарії машинного навчання	25
2.2. Аналіз існуючих automl фреймворків та обґрунтування його вибору	27
2.3. Автоматичне машинне навчання для часових рядів	31
2.4. Open-source фреймворк fedot для автоматичного машинного навчання (AutoML)	33
2.5. Підготовка даних для узгодження обсягів резервного виробництва електроенергії автономними джерелами	35
РОЗДІЛ 3. РЕЗУЛЬТАТИ РОЗРОБКИ МОДЕЛЕЙ ПРОГНОЗУВАННЯ ОБСЯГІВ РЕЗЕРВНОГО ВИРОБНИЦТВА ЕЛЕКТРОЕНЕРГІЇ АВТОНОМНИМИ ДЖЕРЕЛАМИ НА ОСНОВІ ВИКОРИСТАННІ БІБЛІОТЕКИ FEDOT.....	39
3.1. Аналіз часових рядів обсягів резервного виробництва електроенергії.....	39

3.2. Результати прогнозування обсягів виробництва електроенергії генератором.....	44
3.3. Результати прогнозування обсягів виробництва електроенергії вітроустановками	49
3.3. Результати визначення раціональної моделі прогнозування обсягів виробництва електроенергії	51
РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	54
4.1. Небезпечні і шкідливі виробничі чинники під час створення моделей машинного навчання.....	54
4.2. Заходи щодо попередження небезпечних і шкідливих виробничих чинників.....	55
4.3. Розробка логічно-імітаційної моделі процесу виникнення травм під час монтажу інтелектуальної інформаційної системи узгодження обсягів резервного виробництва електроенергії автономними джерелами	56
4.4. Розробка заходів із забезпечення безпеки під час надзвичайних ситуацій	59
Розділ 5. ВИЗНАЧЕННЯ ЕФЕКТИВНОСТІ ВІД ВИКОРИСТАННЯ МОДЕЛІ МАШИННОГО НАВЧАННЯ ДЛЯ УЗГОДЖЕННЯ ОБСЯГІВ РЕЗЕРВНОГО ВИРОБНИЦТВА ЕЛЕКТРОЕНЕРГІЇ АВТОНОМНИМИ ДЖЕРЕЛАМИ.....	61
ВИСНОВКИ І ПРОПОЗИЦІЇ.....	63
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	67

ВСТУП

У сучасному світі зростання виробництва та використання електроенергії відіграють ключову роль у забезпеченні життєвого комфорту, розвитку технологій та підтримці сталих економічних та екологічних стандартів. Попит на електроенергію постійно зростає, що ставить перед сучасними енергетичними системами виклики, пов'язані із забезпеченням стабільності постачання електроенергії та зменшенням впливу на довкілля. Однією з ключових стратегій вирішення цих проблем є збільшення обсягів резервного виробництва електроенергії за допомогою автономних джерел, та використання сучасних технологій для досягнення цієї мети.

Автоматизоване машинне навчання (Machine Learning) стає все більш актуальним і потужним інструментом у вирішенні складних завдань, пов'язаних з управлінням та оптимізацією енергетичних систем. Використання методів машинного навчання дозволяє покращити прогнозування попиту на електроенергію, оптимізувати виробництво та розподіл електроенергії та розробити ефективні стратегії збільшення обсягів резервного виробництва.

Наша кваліфікаційна робота спрямована на вивчення та дослідження питань, пов'язаних із підвищенням ефективності процесу прогнозування обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання. Робота включає в себе аналіз підходів, а також розглядає автоматизоване машинного навчання та розвиває методи та технології для підвищення ефективності використання автономних джерел електроенергії та оптимізації їх роботи на основі машинного навчання.

Існує потреба у розробці та дослідженні інноваційних підходів до вирішення проблеми підвищення ефективності резервного виробництва електроенергії, з орієнтацією на стійкість та зменшення негативного впливу на навколишнє середовище. Досягнення цього може внести вагомий внесок у підвищення стійкості постачання електроенергії та забезпечення потреб суспільства в енергії.

У роботі будуть розроблені загальні принципи та підходи до вирішення даної проблеми, проведено аналіз існуючих досліджень та методів, розроблені конкретні моделі та рішення, спрямовані на підвищення ефективності виробництва резервної електроенергії з використанням автономних джерел та методів автоматизованого машинного навчання.

Отже, узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання має практичну цінність та забезпечує підвищення ефективності процесу управління енергосистемами та зменшення витрат електроенергії.

Об'єктом дослідження є методи автоматизованого машинного навчання для узгодження обсягів резервного виробництва електроенергії автономними джерелами.

Предмет дослідження є вплив характеристик систем резервного виробництва електроенергії на вибір ефективних методів автоматизованого машинного навчання.

Метою даної кваліфікаційної роботи є обґрунтування доцільності та ефективності використання автоматизованого машинного навчання для виконання процесу узгодження обсягів резервного виробництва електроенергії автономними джерелами.

РОЗДІЛ 1.

АНАЛІЗ СТАНУ ПИТАННЯ В ТЕОРІЇ ТА ПРАКТИЦІ

1.1. Аналіз стану предметної області та узгодження обсягів виробництва електроенергії

Електроенергія має важливе значення для розвитку людства. Її використання зростає в різних сферах, таких як житлові та комерційні будівлі, промисловість, медицина, транспорт, громадське освітлення, робототехніка та машини, електростанції, холодильне обладнання та обладнання для кондиціонування повітря тощо [26]. Структура виробництва електроенергії в Україні представлена на рис. 1.1.

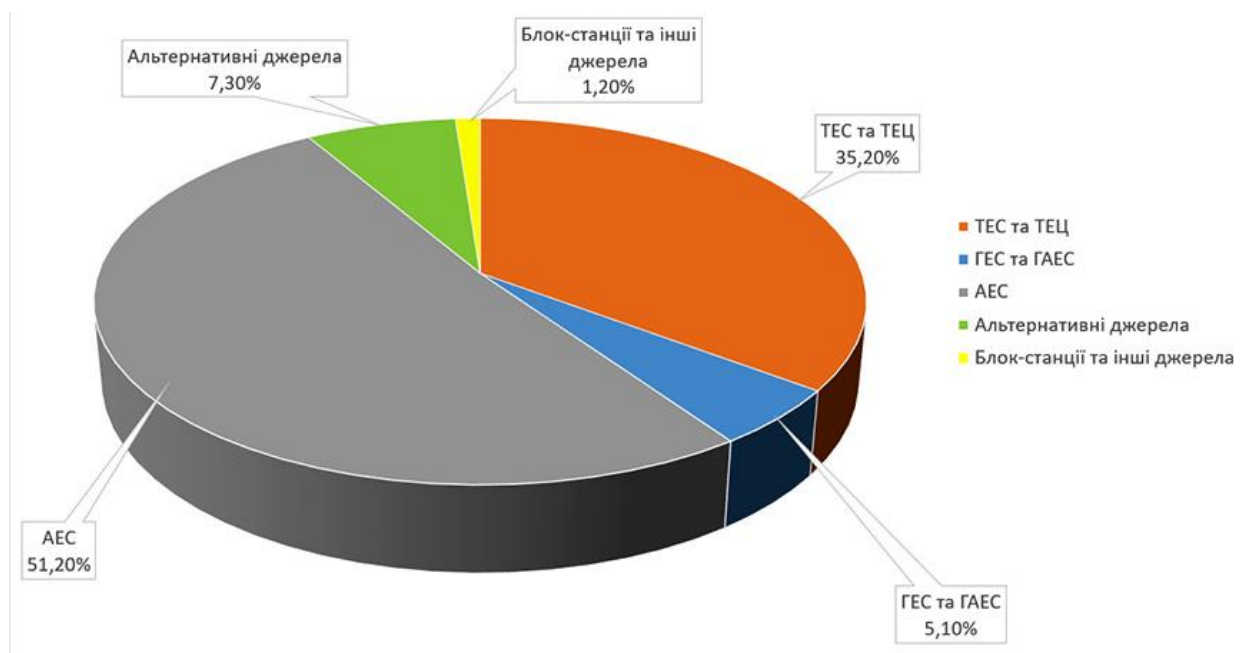


Рисунок 1.1 – Структура виробництва електроенергії в Україні

Енергія, отримана від електричної системи, присутня в різних секторах, і часто машини можуть працювати 24 години на добу безперервно завдяки електриці. Електроенергія не накопичується і потребує стабільної електричної системи з постійним балансом між виробництвом, передачею та попитом [27].

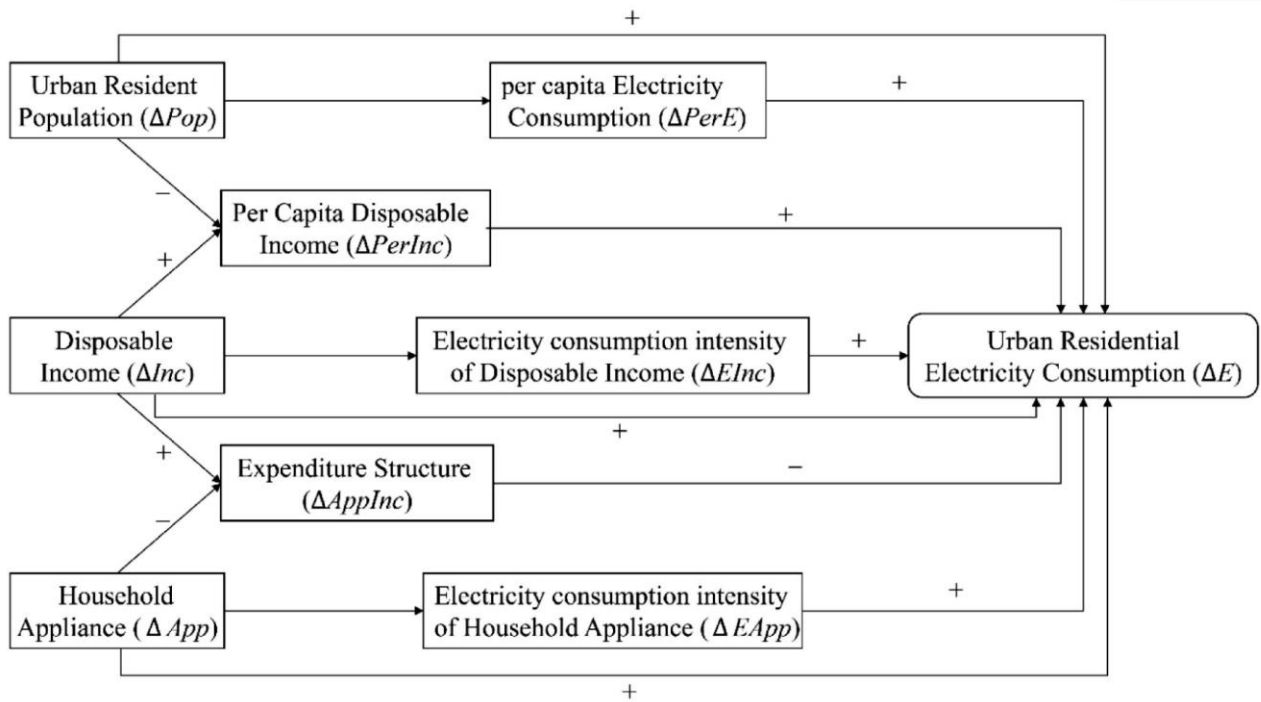


Рисунок 1.2 – Шлях впливу різних факторів на споживання електроенергії міськими будинками [27]

Існує багато типів установок для виробництва електроенергії, таких як тепла енергія, яка перетворюється на електрику, гідроелектрична енергія, яка перетворюється за допомогою гравітаційного потенціалу або кінетичної енергії від джерела гідроенергії, фотоелектрична енергія, яка перетворюється з сонячної енергії за допомогою сонячної енергії. клітини, енергія вітру яких перетворюється з механічної енергії за допомогою вітрових турбін [29].

У роботі [29] подано функцію Watt-Var (рис. 1.3), яка дозволяє інвертору генерувати реактивну потужність у напрямку, протилежному активній потужності.

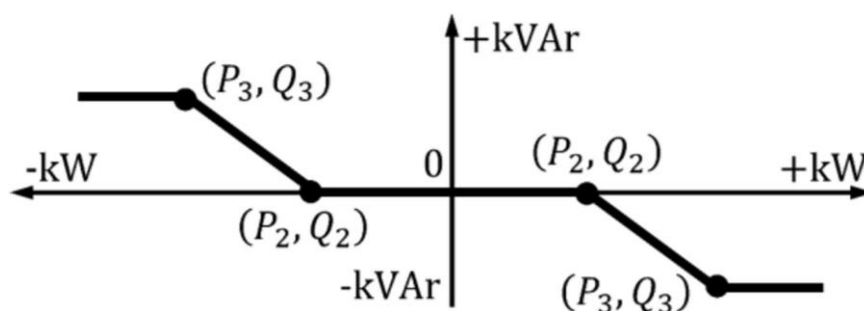


Рисунок 1.3 – Функція інвертора Watt-Var ($P_i, Q_i, i = 1, \dots, 4$, налаштування функції, встановлені енергокомпанією).

Існує діапазон активної потужності, де інвертор працює при одиничному коефіцієнті потужності, тобто нульовій реактивній потужності. Зауважте, що реактивна потужність інвертора повинна відповідати мінімальним вимогам IEEE Std 1547 – 2018. Тобто інвертор повинен мати реактивну потужність щонайменше 44% від номінальної. Інші функції можна перевірити за допомогою нашого методу за таким же підходом.

Розглянемо сонячно-батарейну систему і побутову техніку одного будинку (рис. 1.4). При цьому може використовуватися лінійна оптимізаційна модель HEMS, яку можна знайти в літературі. При цьому є деякі обмеження, що стосуються інвертора та обміну електроенергією House-Grid, оскільки на них безпосередньо впливають вимоги комунального підприємства. Обмеження EV, EWH та AC.

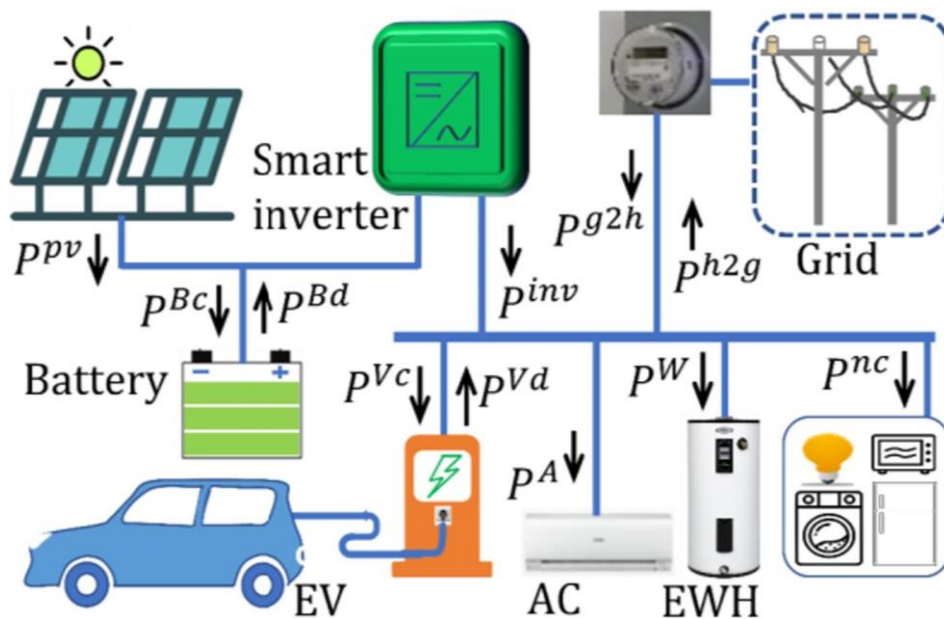


Рисунок 1.4 – Схема підключення побутової техніки:

EV – електромобіль, AC – кондиціонер; EWH – електричний водонагрівач

Споживання енергії зростає з аналізом валового внутрішнього продукту (ВВП) на душу населення. Багатші країни зазвичай споживають більше енергії на людину, ніж країни, що розвиваються. Таким чином, ці економічні відносини також базуються на тому, скільки енергії країна може виробити та

скільки енергії їй потрібно буде виробити в найближчі роки, тому країни повинні планувати та інвестувати в інфраструктуру, необхідну для виробництва енергії.

Задачі оптимізації формулюються та розв'язуються для 24-годинного горизонту часу з кроком 5 хв ($T = 288, \Delta T = 5/60$ год).

Ціни на електроенергію: ціна продажу (енергія від будинку до мережі) (λ) з 30-хвилинними кроками, як на поточному японському ринку електроенергії [42], і двоетапна ціна покупки (енергія від мережі до будинку) (C) рекомендована компанією Tokyo Electric Power Company Holdings, Inc.

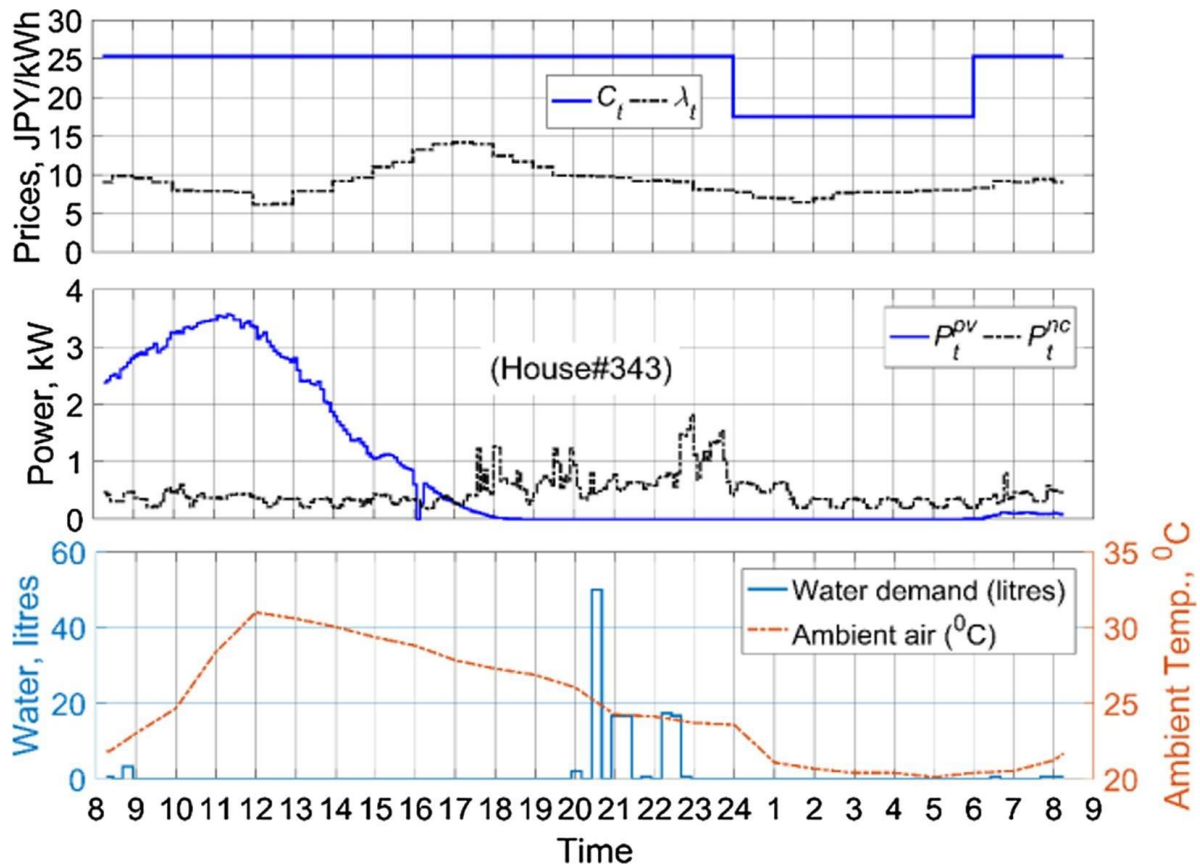


Рисунок 1.5 – Приклад даних будинку №343 [42]

Енергетичні підприємства або електростанції залежать від розрахункового попиту на електроенергію для задоволення навантаження, підключеного до мережі [30].

Що стосується горизонту прогнозу, короткострокове прогнозування навантаження (STLF) може бути корисним для додатків у реальному часі, таких як керування генеруючими установками електроенергії [31]. Середньостроковий прогноз може бути використаний для бюджетування ресурсів, а довгостроковий – для планування розширення мережі енергосистеми. Однак попит на електроенергію залежить від погодних умов, таких як температура, швидкість вітру, опади та інших факторів, і попит змінюється залежно від дня, сезону, торгівлі та промислової діяльності, як-от години пік, звичайні години, робочі дні, вихідні дні, свята та ін. Таким чином, сезонність, тренд, шум, викиди та інші аспекти присутні під час вивчення попиту на електроенергію, що вимагає від енергопостачальних компаній методів точного прогнозування навантаження для зменшення витрат виробництва електроенергії та витрат [32].

Дані часових рядів електроенергії зазвичай мають нелінійну поведінку, що означає випадкові та періодичні компоненти, вбудовані в ряди, такі як час, сезонні події, поведінка економічної діяльності, помилки вимірювання (викиди), відсутні значення та шум [33]. Ці компоненти є невід’ємними для даних часових рядів, і кожен набір даних міститиме різні характеристики. Потім потрібні етапи попередньої обробки та перетворення даних, щоб забезпечити відповідні вхідні дані для моделей прогнозування. З належною методологією моделі машинного навчання (ML) можуть впоратися з такими негариздами, точно передбачаючи наступні кроки, які необхідно зробити в завданнях прогнозування [34].

1.2. Доцільність використання машинного навчання для процесів узгодження обсягів резервного виробництва електроенергії

Машинне навчання (ML) – це підобласть штучного інтелекту (ШІ), яка надає прикладні моделі класифікації, регресії та прогнозування на основі

досвіду. Досвід отримується від експериментально вимірних історичних даних. Такими даними можуть бути текст, аудіо, зображення, відео, дослідження в галузі медицини, вимірювання датчиків і споживання електроенергії, серед інших доступних джерел даних [35].

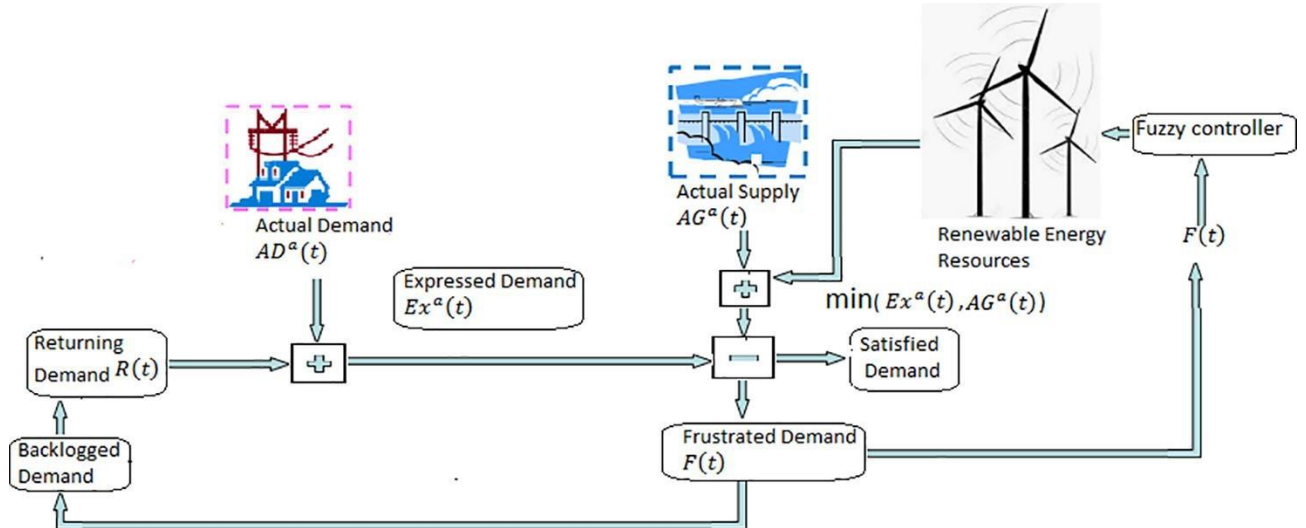


Рисунок 1.6 – Замкнена система реагування на попит електроенергії [35]

На рис. 1.6 показана схема реагування на попит, тобто як реагує система контуру керування зворотним зв'язком для моделювання контролера у разі збурень. Контролер працює як замкнута система для постійного моніторингу енергосистеми та здатності вживати належних заходів для усунення балансу між моделлю попиту та постачання. Забезпечує механізм оптимальної диспетчеризації електроенергії у разі невидимих збоїв.

Моделі ML можуть вчитися на досвіді та використовувати ці знання для відтворення деяких завдань, таких як прогнозування споживання енергії на тиждень наперед, прогнозування погодних умов, запобігання збоїв у машинах, які працюють, тощо. За горизонт прогнозування навантаження приймають кількість кроків вперед (годин, діб тощо), які потрібно передбачити. Це можна визначити в чотирьох різних категоріях [35]:

- 1) дуже короткострокове прогнозування навантаження (VSTLF);
- 2) короткострокове прогнозування навантаження STLF;
- 3) середньострокове прогнозування навантаження (MTLF);
- 4) довгострокове прогнозування навантаження (LTLF).

Граничні горизонти для цих чотирьох категорій становлять один день, два тижні та три роки відповідно.

Окрім прогресу в додатках штучного інтелекту та машинного навчання, ще існує багато проблем, які потрібно подолати. Навчання для деяких моделей ML іноді є неефективним і поверхневим, що вимагає більших обчислювальних зусиль і здатності до узагальнення [36].

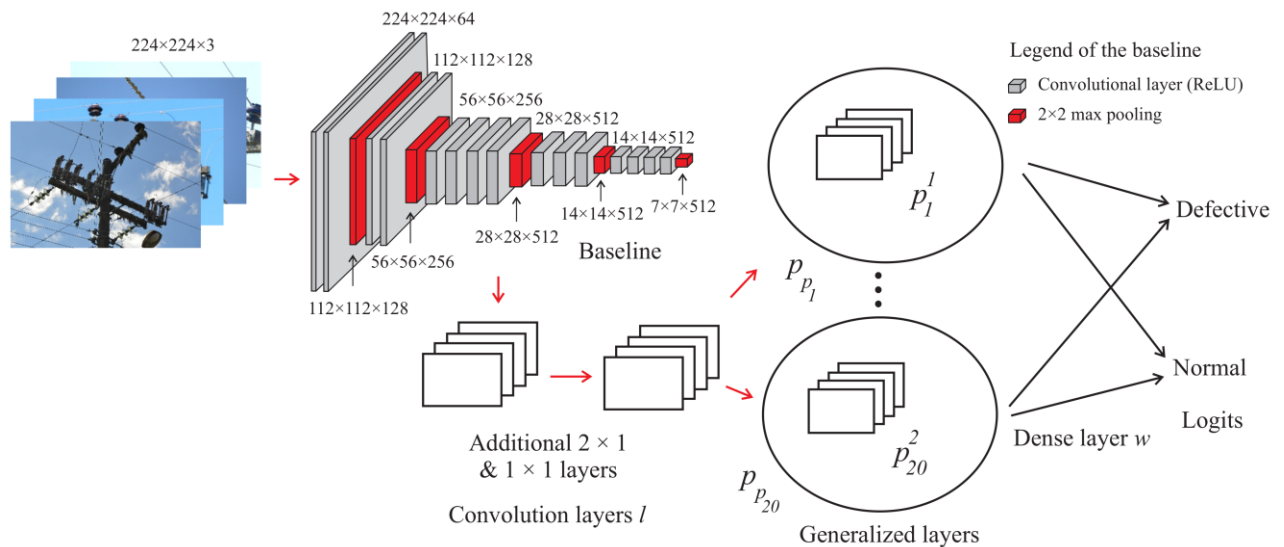


Рисунок 1.7 – Архітектура Semi-ProtoPNet [36]

Моделі ML не мають автоматично налаштованих параметрів, званих гіперпараметрами, які безпосередньо впливають на їх продуктивність. Деякі нещодавні дослідження, такі як представлене в роботі [37], показали, що комбінування різних індивідуальних моделей машинного навчання може дати кращий результат порівняно з окремими моделями, якщо вони побудовані належним чином.

Набір даних, який використовується як вхідні дані в запропонованій моделі, складається з трьохмісячних вимірювань швидкості вітру. Вимірювання проводять кожні 10 хвилин на висоті 95 метрів від рівня землі, що відповідає вітрогенератору з діаметром лопаті 100 м. Вимірювальне обладнання встановлено для моніторингу вітрового ресурсу вітрової електростанції потужністю приблизно 150 МВт, розташованої в штаті Ріо-Гранді-ду-Норті, у

північно-східному регіоні Бразилії, а точніше в місті Паразіньо. Вимірювання проводяться з березня по травень 2020 року (рис. 1.8).

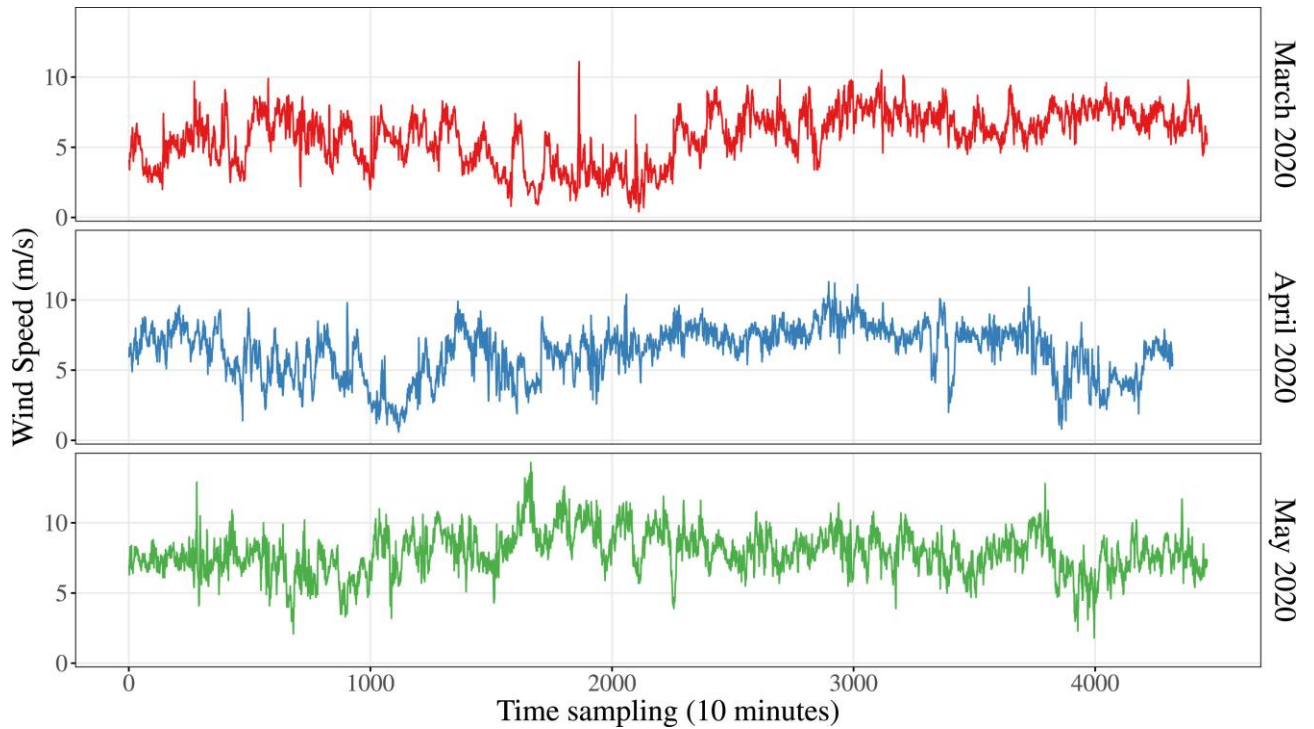


Рисунок 1.8 – Набори даних за березень, квітень і травень 2020 року відповідно [37]

Часовий ряд складається з 4464 проб за березень і травень 2020 року та 4320 проб за квітень 2020 року з періодом вибірки 10 хвилин, як показано на рис. 1.8. Використання даних за різні місяці є важливим, оскільки додаткові дані дозволяють моделям вивчити максимальну кількість характеристик даних. Крім того, він має свої кліматичні характеристики кожного місяця і, таким чином, дає можливість оцінити модель у різних кліматичних ситуаціях.

На рис. 1.9 представлено Робочий процес навчання ансамблю STACK.

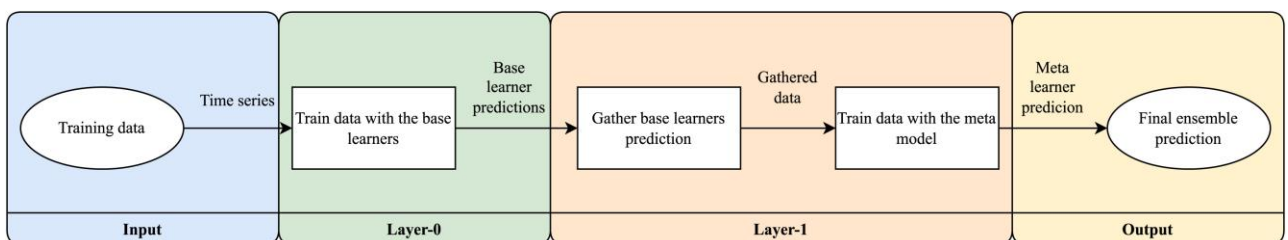


Рисунок 1.9 – Робочий процес навчання ансамблю STACK

1.3. Аналіз моделей машинного навчання для процесів узгодження обсягів резервного виробництва електроенергії

Ансамблеві моделі виявилися кращими, ніж найгірший індивідуальний предиктор, а в деяких випадках і кращий, ніж найкращий індивідуальний предиктор. Wang та ін. у роботі [38] застосували адаптивний метод декомпозиції на основі Variation Mode Decomposition (VMD) і SampEn для декомпозиції даних і eXtreme Gradient Boosting (XGBoost) для короткострокового прогнозування навантаження, який показав кращі результати, ніж будь-який індивідуальний алгоритм навчання. У цій роботі використовуються два різні набори даних для дослідження моделі ансамблю та методів декомпозиції для VSTLF на один день вперед.

Доведено, що на поведінку промислових споживачів у споживанні електроенергії впливають відповідні фактори, такі як режим роботи, температура навколишнього середовища та правила календаря (рис. 1.10).

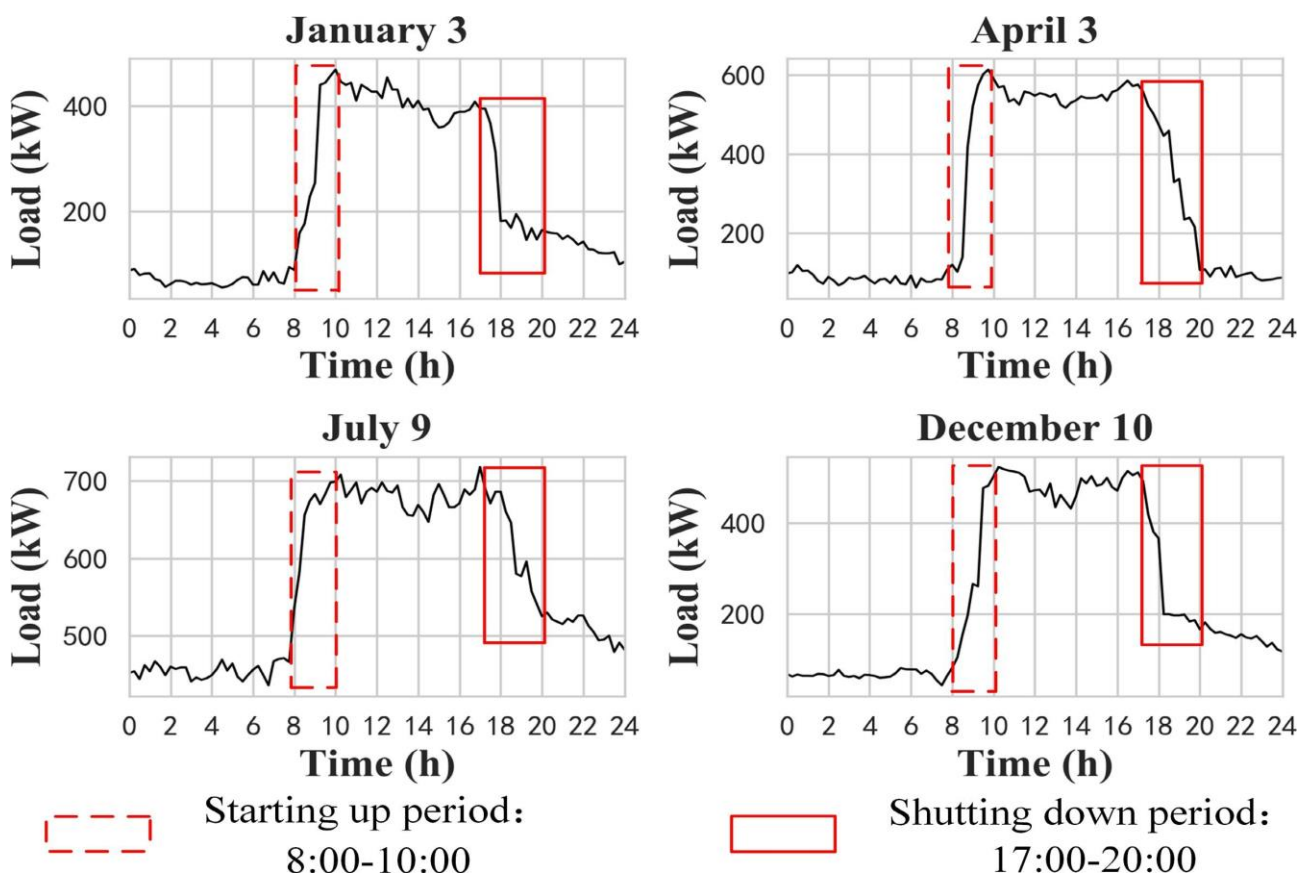


Рисунок 1.10 – Період пуску та зупинки виробничого обладнання [38]

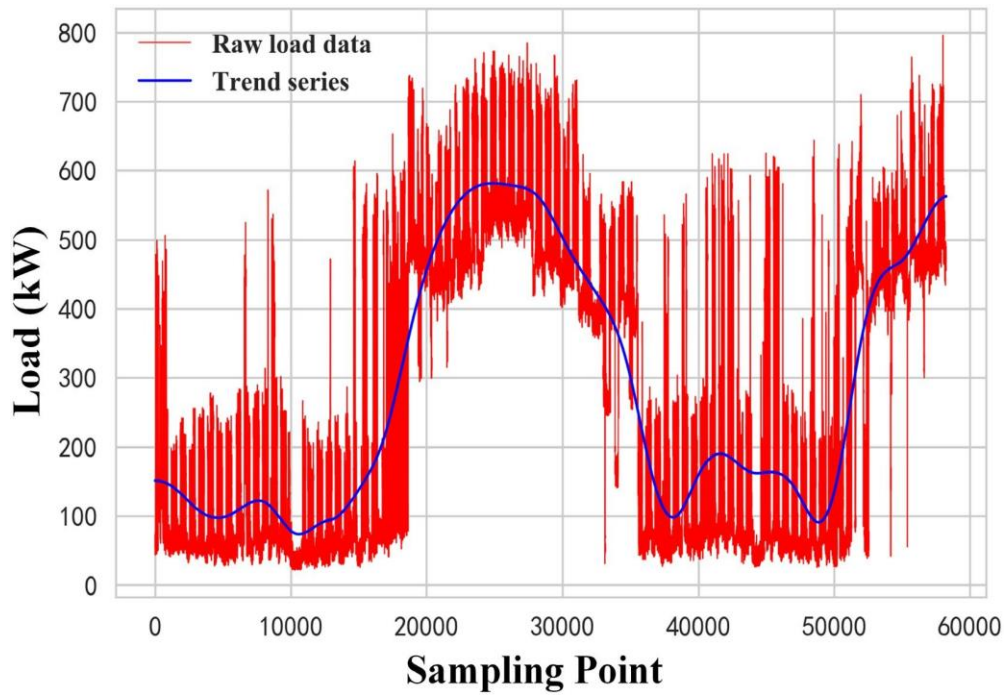


Рисунок 1.11 – Необроблені дані про навантаження та ряди тенденцій

Проілюстровано на прикладі декомпозиції даних про навантаження промислових споживачів під час експерименту. Представлено ряд трендів необроблених даних про навантаження витягується за допомогою EMD (рис. 1.11).

У результаті проведених досліджень отримані криві прогнозу навантаження всіх моделей (рис. 1.12).

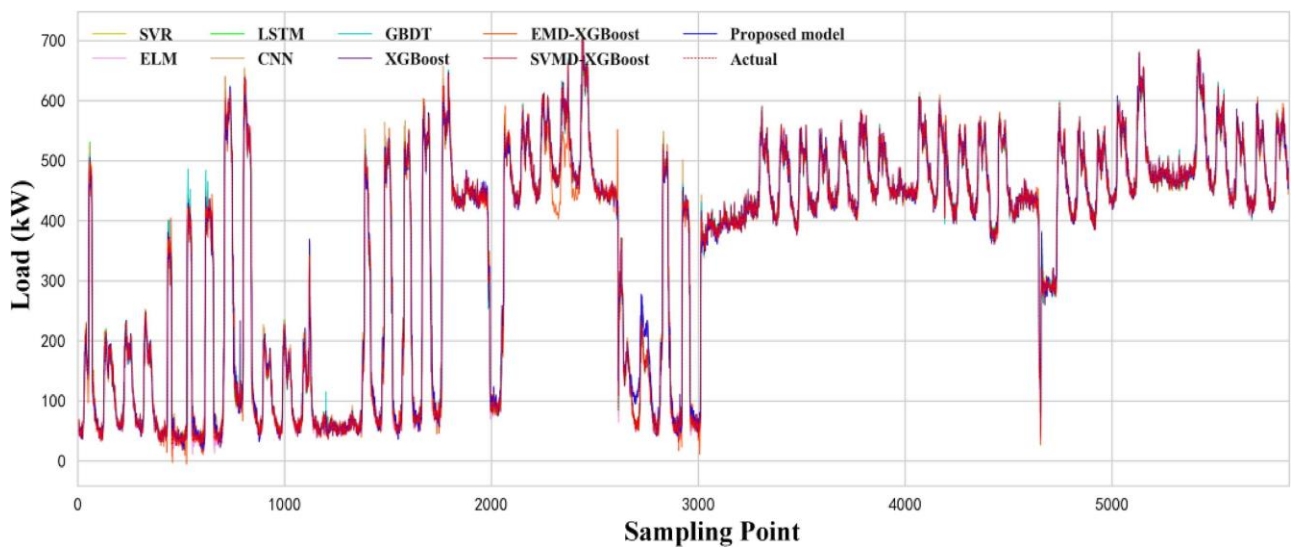


Рисунок 1.12 – Крива прогнозу навантаження всіх моделей протягом 01.04.2019 та 31.05.2019

Основні внески цього дослідження підсумовані таким чином:

1) Моделі ML, включно з градієнтним регресором (GBR), XGBoost, k-Nearest Neighbors (kNN) і регресією опорного вектора (SVR), розроблені та оцінені як в окремих, так і в комбінованих моделях для VSTLF.

2) Нова гібридна структура на основі GBR, XGBoost, kNN і SVR, інтегрована з методами декомпозиції сигналу, такими як декомпозиція сезонного тренду з використанням локально оціненого згладжування діаграми розсіювання, LOESS (STL), емпірична декомпозиція моди (EMD), емпірична мода декомпозиції Ensemble (EEMD), Повний ансамблевий емпіричний режим декомпозиції з адаптивним шумом (CEEMDAN) і емпіричне вейвлет-перетворення (EWT) розроблено для підвищення точності класичних структур ML у VSTLF.

3) Налаштування гіперпараметрів моделей ML, інтегрованих з методами декомпозиції, базувалося на методі автоматизованого машинного навчання.

4) Розширена оцінка на основі двох наборів даних, одного від Національного оператора електричної системи (ONS), а іншого від Незалежних системних операторів Нової Англії (ISO-NE), виконується для демонстрації потенціалу запропонованого ML, інтегрованого з методами декомпозиції в VSTLF. Були розглянуті етапи попередньої обробки: очищення даних, розробка функцій і обробка викидів.

5) Для оцінки моделей було прийнято такі загальноприйняті показники для VSTLF, як середньоквадратична помилка (RMSE), середня абсолютна помилка (MAE) і середня абсолютна відсоткова помилка (MAPE) із 10-кратною перехресною перевіркою. Статистичні тести продемонстрували, що запропоновані моделі забезпечили кращу продуктивність за допомогою декомпозиції сигналів порівняно з моделями ML без процесу декомпозиції сигналів часових рядів у VSTLF.

Експерименти VSTLF були розділені на два набори даних, ONS та ISO-NE. Мета полягала в тому, щоб оцінити, наскільки добре регресійні моделі

можуть передбачити попит на електроенергію між різними алгоритмами ML з методами декомпозиції та без них. Щоб досягти найнижчих значень показників продуктивності RMSE та MAPE для всіх оцінюваних моделей ML, для вибору найкращих гіперпараметрів використовувався інструментарій автоматичного машинного навчання (AutoML).

1.4. Інструментарій автоматичного машинного навчання AutoML

AutoML – це API-сервіс TAO Toolkit, який автоматично вибирає гіперпараметри глибокого навчання для вибраної моделі та набору даних.

Automated Machine Learning надає методи та процеси, щоб зробити машинне навчання доступним для експертів, які не займаються машинним навчанням, щоб підвищити ефективність машинного навчання та прискорити дослідження машинного навчання.

За останні роки машинне навчання (ML) досягло значних успіхів, і на нього покладається дедалі більша кількість дисциплін. Однак цей успіх значною мірою покладається на те, що фахівці з машинного навчання людини зможуть виконувати такі завдання:

- ✓ Попередня обробка та очищення даних.
- ✓ Виберіть і побудуйте відповідні елементи.
- ✓ Виберіть відповідне модельне сімейство.
- ✓ Оптимізація гіперпараметрів моделі.
- ✓ Проектувати топологію нейронних мереж (якщо використовується глибоке навчання).
- ✓ Моделі машинного навчання після обробки.
- ✓ Критично аналізувати отримані результати.

Оскільки складність цих завдань часто виходить за рамки тих, хто не є фахівцем з машинного навчання, швидке зростання програм машинного навчання спричинило попит на готові методи машинного навчання, які можна

використовувати легко й без експертних знань. Ми називаємо отриману область досліджень, спрямовану на прогресивну автоматизацію машинного навчання, *AutoML*.

Дослідження в галузі автоматизованого машинного навчання є дуже різноманітними та пропонують пакети та методи, орієнтовані як на дослідників, так і на кінцевих користувачів.

Протягом останніх років було розроблено кілька готових пакетів, які забезпечують автоматизоване машинне навчання:

AutoWEKA – це підхід для одночасного вибору алгоритму машинного навчання та його гіперпараметрів; у поєднанні з пакетом *WEKA* він автоматично дає хороші моделі для широкого спектру наборів даних.

Auto-sklearn – це розширення *AutoWEKA*, що використовує бібліотеку Python *scikit-learn*, яка замінює звичайні класифікатори та регресори *scikit-learn*.

Auto-PyTorch базується на структурі глибокого навчання *PyTorch* і спільно оптимізує гіперпараметри та нейронну архітектуру.

Інші добре відомі пакети *AutoML* включають:

AutoGluon – це багатосаровий підхід стекування різноманітних моделей машинного навчання.

H2O AutoML забезпечує автоматичний вибір моделі та групування для платформи машинного навчання та аналізу даних *H2O*.

MLBoX – це бібліотека *AutoML* із трьома компонентами: попередня обробка, оптимізація та прогнозування.

TRIO – це помічник із обробки даних, який оптимізує конвеєри машинного навчання за допомогою генетичного програмування.

TransmogriAI – це бібліотека *AutoML*, яка працює на основі *Spark*.

Дослідники стверджують, що продуктивність певної техніки залежить як від фундаментальної якості алгоритму, так і від деталей його налаштування, і що іноді важко зрозуміти, чи дана техніка справді краща, чи просто краще налаштована. Щоб покращити ситуацію, Bergstra et al. запропоновані

результати звітності, отримані шляхом налаштування всіх алгоритмів за допомогою того самого інструментарію оптимізації гіперпараметрів. Доповідь Скаллі та ін. ICLR'18 «Прокляття переможця» наводить аргументи в тому ж напрямку та наводить нещодавні приклади, у яких правильна оптимізація гіперпараметрів базових ліній покращилася порівняно з останніми сучасними результатами та нещодавно запропонованими методами.

Оптимізація гіперпараметрів і конфігурація алгоритмів забезпечують методи автоматизації стомлюючого, трудомісткого та схильного до помилок процесу налаштування гіперпараметрів для нових поточних завдань. Відомі, наприклад, надаємо пакети для оптимізації гіперпараметрів:

- ✓ SMAC3 – повторна реалізація алгоритму SMAC на Python;
- ✓ ДЕНВ – Differential Evolution у поєднанні з HyperBand.

Поле пошуку архітектури стосується проблеми пошуку добре продуктивної архітектури глибокої нейронної мережі. Автоматизований пошук архітектури може значно пришвидшити розробку нових програм глибокого навчання, оскільки розробникам не потрібно точно оцінювати різні архітектури.

Пакети для пошуку архітектури та оптимізації гіперпараметрів для глибокого навчання включають:

- ✓ Auto-PyTorch
- ✓ АвтоКерас
- ✓ Талос.

ТАО Toolkit API надає інтерфейс блокнота Jupyter, щоб спробувати функцію AutoML.

Кожен запуск AutoML містить кілька навчальних експериментів. Наприкінці запуску AutoML ви можете отримати доступ до конфігурації, яка містить гіперпараметри найкращої моделі серед кількох експериментів, а також файл двійкової ваги для розгортання моделі у вашій програмі.

AutoML підтримує всі моделі ТАО, крім моделі MAL – Auto Labeling.

1.5. Обґрунтування доцільності узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання

Забезпечення сталого та надійного постачання електроенергії є завданням для сучасних енергетичних систем. У разі виникнення аварійних ситуацій, включаючи природні катастрофи, технічні збої або інші непередбачені обставини, резервне виробництво стає критичним місцем для забезпечення безперебійного постачання електроенергії. Використання автономних джерел енергії, таких як сонячні панелі, вітрові турбіни, акумуляторні системи тощо, дозволяє зменшити залежність від традиційних джерел електроенергії та підвищити надійність систем в умовах аварій.

Зростаюча обуреність щодо зміни клімату та погіршення екологічної ситуації змушує шукати більш екологічно чисті та стабільні рішення у сфері енергетики. Використання автономних джерел енергії, які базуються на відновлювальних джерелах, сприяє зменшенню викидів парникових газів та інших забруднювачів довкілля. Це виконання виконання міжнародних зобов'язань щодо скорочення викидів та збереження природних ресурсів.

Впровадження автономних джерел енергії та система автоматизованого машинного навчання може призвести до значних економічних видів для регіонів та підприємств. Таким чином, витрати на побудову та експлуатацію таких систем можуть бути значно меншими в порівнянні з традиційними електростанціями на вугіллі чи газі. Додатково, автоматизоване машинне навчання дозволяє оптимізувати виробництво та споживання електроенергії, що зменшує витрати та збільшує загальну ефективність системи.

Розвиток та впровадження автоматизованого машинного навчання в галузі енергетики відкриває широкий простір для інновацій та новаторських рішень. Це сприяє розвитку нових технологій, методів та підходів, які можуть бути використані не лише в енергетичній галузі, але й в інших сферах. Такий

підхід стимулює розвиток галузі та сприяє появі нових можливостей для покращення сучасного енергетичного сектора.

Отже, тема кваліфікаційної роботи «Підвищення ефективності процесу узгодження обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання» є достатньо своєчасною та має практичну і теоретичну цінність та забезпечує розвиток цифровізації у енергетиці.

РОЗДІЛ 2.

ВИБІР ЗАСОБІВ ТА ПІДГОТОВКА ДАНИХ ДЛЯ УЗГОДЖЕННЯ ОБСЯГІВ РЕЗЕРВНОГО ВИРОБНИЦТВА ЕЛЕКТРОЕНЕРГІЇ АВТОНОМНИМИ ДЖЕРЕЛАМИ НА ОСНОВІ АВТОМАТИЗОВАНОГО МАШИННОГО НАВЧАННЯ

2.1. Можливі сценарії машинного навчання

Для того, щоб розв'язати будь-яку задачу за допомогою методів машинного навчання, потрібно пройти багато кроків починаючи від очищення даних, підготовки датасету, вибору найбільш інформативних ознак і перетворення ознакового простору до добору моделі машинного навчання і налаштування її гіперпараметрів. Таку послідовність часто представляють у вигляді пайплайна (рис. 2.1).

Однак навіть на роботу з лінійними пайплайнами (рис. 2.1, а), для підбору їхньої структури та параметрів можуть знадобитися дні, а іноді й тижні. У процесі розв'язання складних завдань пайплайни набувають іншої структури. Для підвищення якості моделювання часто використовують ансамблеві методи (стекінг), що об'єднують кілька моделей (рис. 2.1, б), або роблять пайплайн розгалуженим. В останньому випадку задіюють різні методи попереднього опрацювання для варіативних моделей ML, що навчаються на різних частинах датасету (рис. 2.1, в).

Автоматизована підготовка пайплайну – це переважно завдання комбінаторної оптимізації або пошуку найкращого поєднання можливих факторів, зокрема множини обчислювальних блоків. У цьому разі пайплайн описують у вигляді спрямованого ациклічного графа (directed acyclic graph, DAG), який може бути трансформований у графік обчислень, а ефективність визначають за допомогою цільових функцій, що чисельно оцінюють якість, складність, стійкість та інші властивості отриманої моделі.

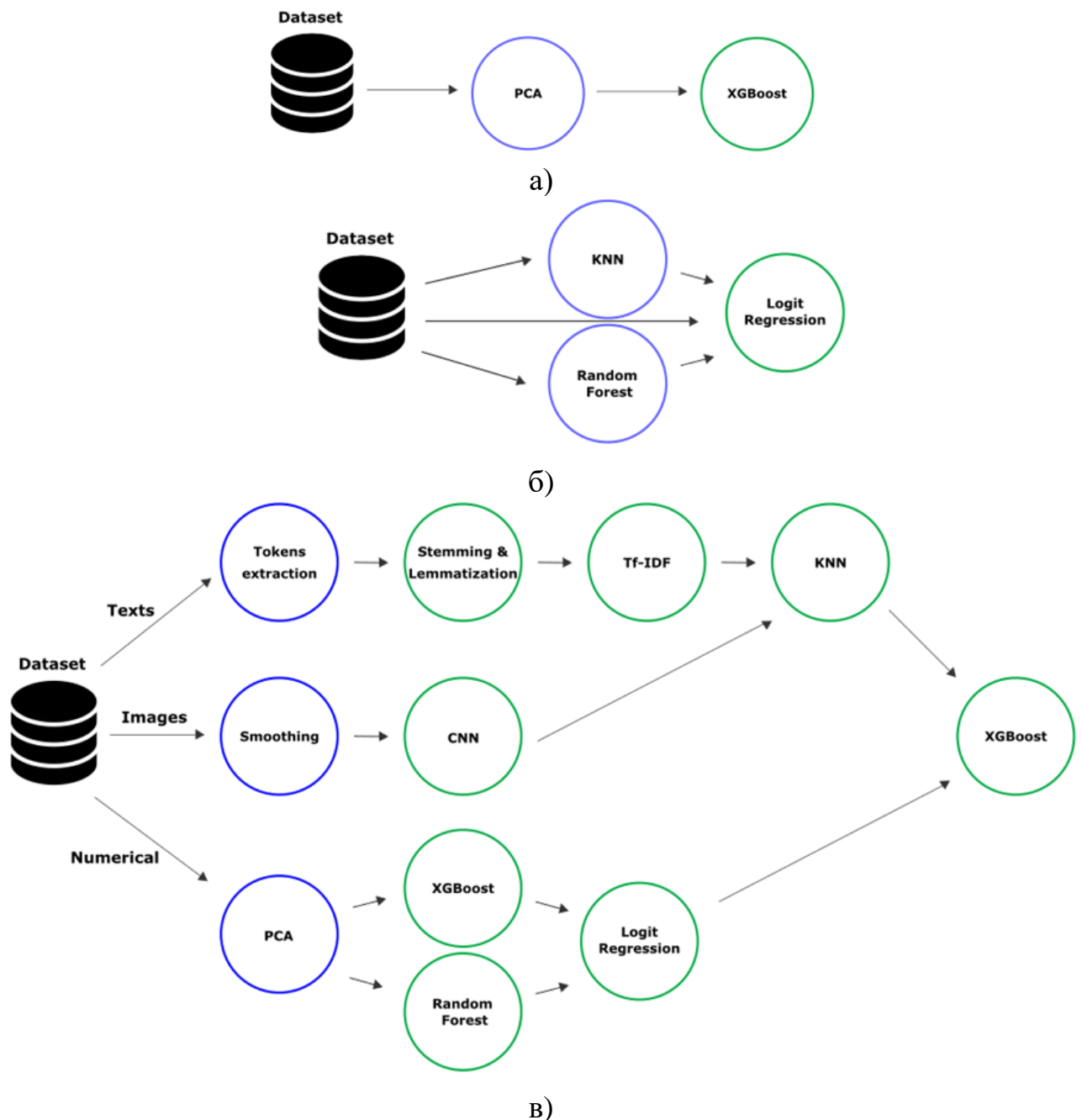


Рисунок 2.1 – Можливі сценарії машинного навчання

Моделі зі структурою типу (рис. 2.1, б) і (рис. 2.1, в) фактично стають композитними, тобто містять у собі різні алгоритми машинного навчання. Наприклад, можна об'єднати байєсівську мережу і згорткову для передбачення на мультимодальних даних. А з композитними моделями та пайплайнами машинного навчання можна працювати за допомогою методів і технологій AutoML.

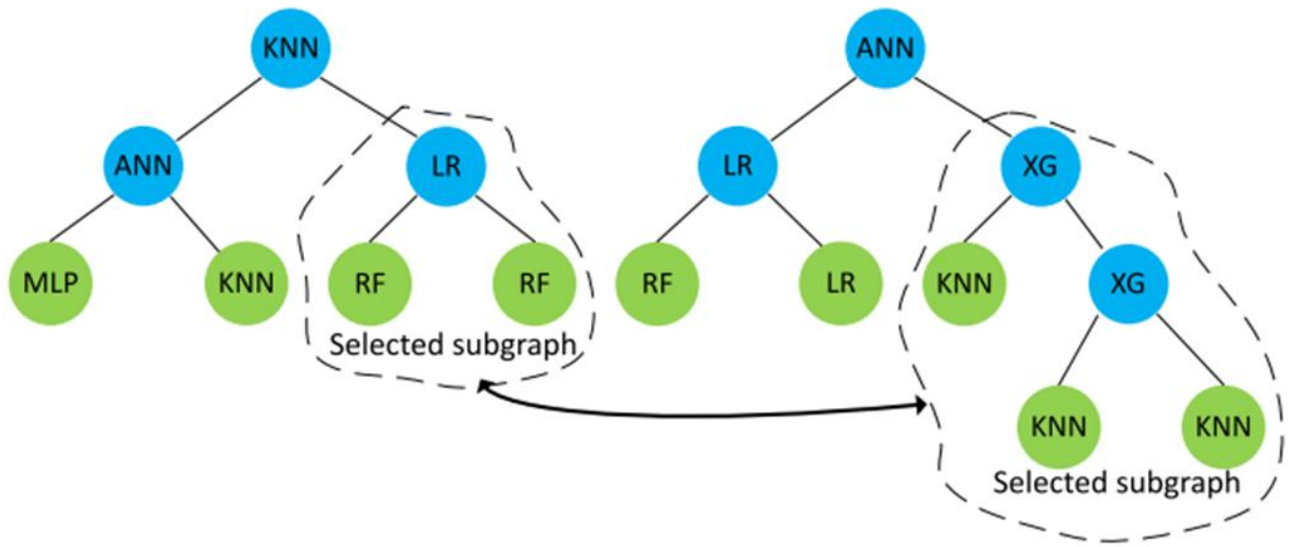
2.2. Аналіз існуючих AutoML фреймворків та обґрунтування його вибору

Найпримітивніший метод розв'язання цього завдання – випадковий пошук (random search) з оцінкою поєднань блоків. Досконаліший підхід – мета-евристичні алгоритми оптимізації, до яких належать ройові та еволюційні (останні можна реалізувати за допомогою фреймворків TPOT і FEDOT). Такі алгоритми повинні мати спеціалізовані оператори кросовера, мутації та селекції для застосування до особин, що описуються графом (зазвичай деревом), працювати з багатокритеріальною цільовою функцією, включати додаткові процедури для створення стійких і не схильних до перенавчання пайплайнів (наприклад, регуляризацію).

Оператори кросовера і мутації можуть бути реалізовані класичним способом – у вигляді кросовера піддерев (subtree crossover), коли обираються дві батьківські особини, які обмінюються випадковими частинами своїх графів. Але це – не єдина можлива реалізація, є і більш семантично складні варіанти (наприклад, one-point crossover). Мутація на деревах також передбачає реалізацію, що включають випадкову зміну моделі (або обчислювального блоку) у випадковому вузлі графа на підходящий варіант із пулу моделей, видалення випадкового вузла, а також випадкове додавання піддерев.

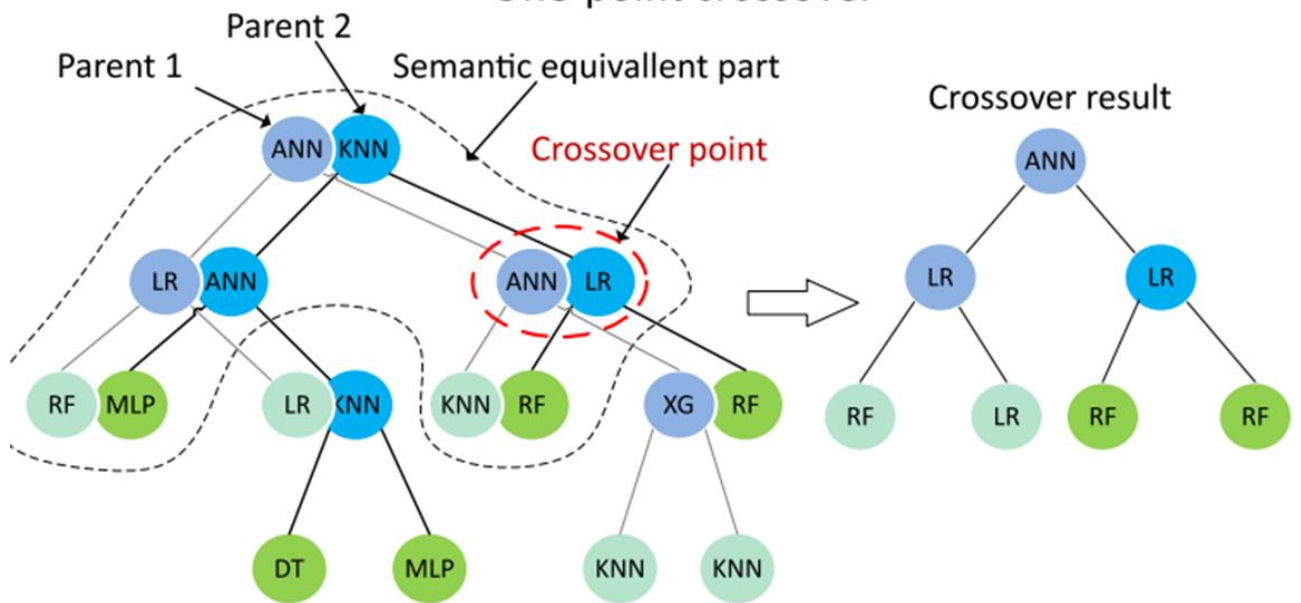
В ідеальному випадку, AutoML дає змогу виключити експерта-аналітика з процесу розробки, експлуатації та впровадження моделі. Однак поки що домогтися цього повною мірою складно – більша частина AutoML фреймворків підтримує розв'язання окремих завдань автоматичного машинного навчання (налагодження гіперпараметрів, вибір ознак тощо) у межах фіксованих пайплайнів і тільки для деяких типів даних (табл. 2.1). Наведене порівняння, зрозуміло, не претендує на повноту – його виконували на основі аналізу відкритої документації та прикладів, а стан справ у сфері AutoML може швидко змінюватися.

Subtree crossover



a)

One-point crossover



б)

Рисунок 2.2 – Пайплайни машинного навчання у вигляді кросовера: а) прості піддерева (subtree crossover); б) семантично складні (one-point crossover)

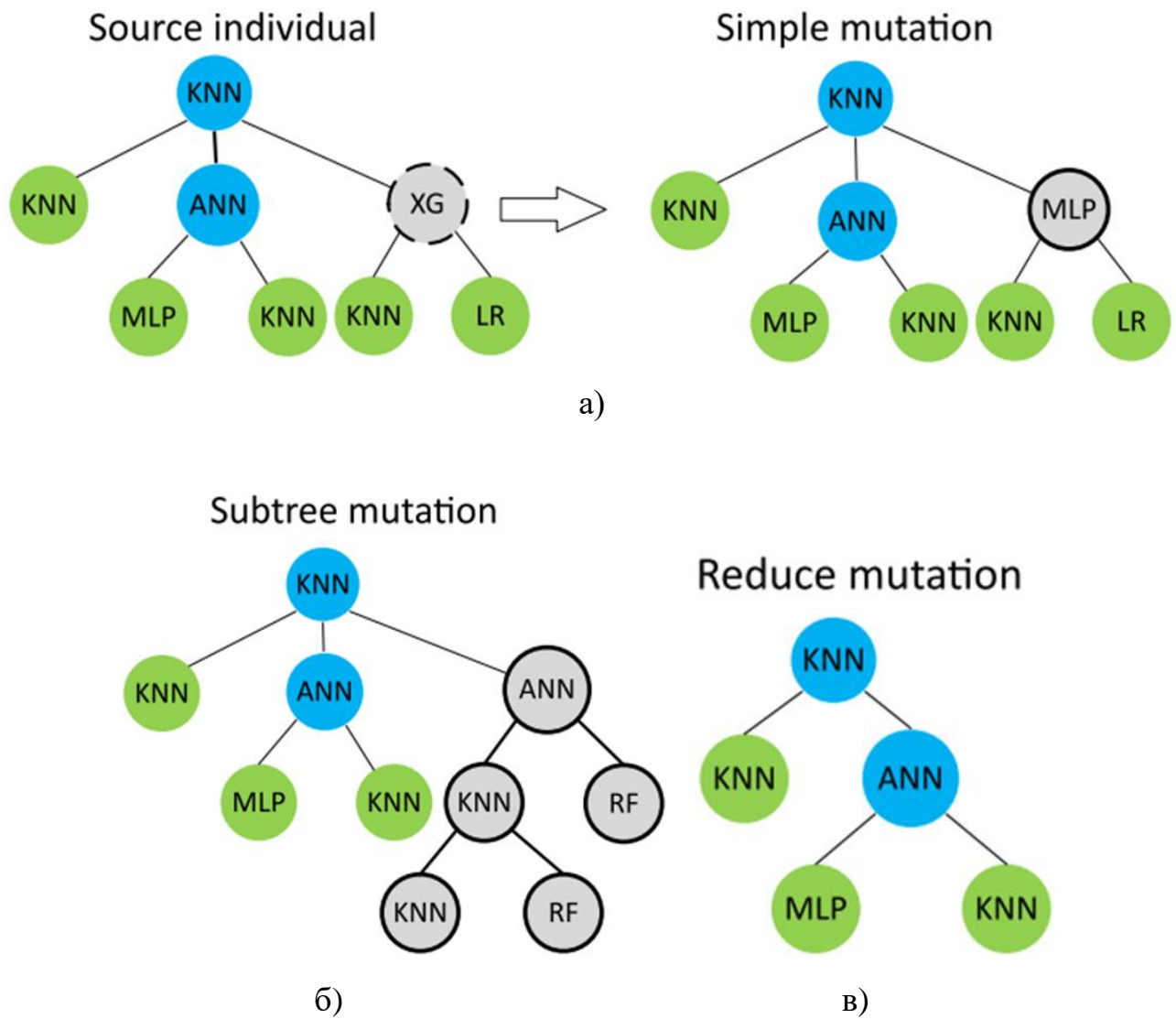


Рисунок 2.3 – Пайплайни машинного навчання у вигляді мутації: а) випадкова зміна моделі (або обчислювального блоку); б) видалення випадкового вузла; в) випадкове додавання піддерев

Таблиця 2.1 – Характеристики існуючих AutoML фреймворків

Назва	Тип пайплайну	Метод оптимізації	Вхідні дані	Масштабування	Додаткові особливості
TPOT	Variable	GP	Tabular	Multiprocessing, Rapids	Code generation

H2O	Fixed	Grid Search	Tabular, Texts	Hybrid	–
Auto Sklearn	Fixed	SMAC	Tabular	–	–
ATM	Fixed	BTB	Tabular	Hybrid	–
FEDOT	Variable	GP + hyperopt	Tabular, Timeseries, Images, Texts	–	Composite pipelines
Auto Gluon	Fixed	Fixed Defaults	Tabular, Images, Texts	–	NAS, AWS integration
LAMA	Fixed	Optuna	Tabular, Texts	ML Space	Profiling
NNI	Fixed	Bayes	Tabular, Images	Hybrid, Kubernetes	NAS, WebUI

Представлена таблиця 2.1 містить інформацію про різні інструменти автоматизації машинного навчання (AutoML). Кожен інструмент має свої унікальні характеристики та особливості, такі як методи оптимізації, типи пайплайнів, підтримувані типи даних та додаткові можливості. Вибір конкретного інструменту може залежати від конкретних вимог та властивостей задачі машинного навчання.

Окремим завданням у галузі машинного навчання є автоматична генерація пайплайнів. Існує кілька фреймворків, які надають подібні функції, серед open-source, наприклад це TPOT, AutoGluon, MLJAR або H2O. Такі AutoML фреймворки розв'язують задачу оптимізації завдяки побудові такого пайплайну, що дає кінцевий прогноз з найменшою (серед усіх розглянутих рішень) помилкою. Здебільшого структуру пайплайна зафіксовано, і добирають тільки гіперпараметри, але деякі фреймворки здатні отримувати як розв'язок моделі довільної структури. Це оптимізаційне завдання (знаходження пайплайна довільної структури) вирішується зазвичай за допомогою еволюційних алгоритмів, приклад фреймворки TPOT та FEDOT.

Існують також і пропрієтарні SaaS-рішення, як-от DataRobot, GoogleAutoTables, Amazon SageMaker, які допомагають не тільки автоматизувати ML експерименти, а й надають можливості AutoML.

Як правило, AutoML бібліотеки та сервіси успішно вирішують дві найпопулярніші задачі в машинному навчанні це класифікація та регресія на табличних даних. Рідше підтримуються завдання, пов'язані з опрацюванням зображень, тексту та прогнозування часових рядів. У нашій роботі ми розглядаємо можливості автоматичного машинного навчання у завданні прогнозування часових рядів.

2.3. Автоматичне машинне навчання для часових рядів

Незважаючи на те, що завдання прогнозування є доволі затребуваним у науці та бізнесі, більша частина open-source бібліотек автоматичного машинного навчання не надають можливостей для формування пайплайнів для завдання прогнозування часових рядів. Причин тому може бути кілька, одна з яких - складність в адаптації поточної функціональності бібліотеки для прогнозування рядів без перероблення інструментарію для інших завдань (класифікації та регресії).

Враховуючи те, що опрацювання часових рядів відрізняється від звичного набору дій під час розв'язання задачі регресії. Відмінності починаються вже з розбиття вихідної вибірки. Так, наприклад, перемішувати дані у випадковому порядку для валідації моделі часового ряду немає потреби. По-іншому для часових рядів формуються свої ознаки. На вихідному ряді, як правило, моделі машинного навчання не навчають, а його потрібно перевести в інше подання. Впроваджувати такі конструкції в уже наявний AutoML проєкт зі своїм legacy буває проблематично. Це пов'язано із тим що розробники часто відмовляються від часових рядів і концентруються на опрацюванні конкретних типів даних, лише табличних, або лише тексту.

Інший спосіб полягає у адаптуванні функціональності готового AutoML-інструменту для завдання прогнозування. На цю роль добре підходять регресійні моделі. Однак, судячи з деяких прикладів використання open-source версії, користувачеві доведеться взяти на себе завдання первинної обробки вихідних рядів, наприклад, витяг ознак. Для повноцінної роботи такого урізаного інструментарію може не вистачати.

Хоча завдання прогнозування поведінки одновимірного масиву виглядає тривіальним, існує безліч інструментів, які можна використовувати під час роботи з часовими рядами. Наприклад, є можливість будувати інтерпретовані ансамблі з моделей (щоб одна модель відтворювала високочастотну складову часового ряду, друга – низькочастотну, а третя – об'єднувала їхні прогнози). Існує потреба і виконанні налаштування гіперпараметрів у пайплайнах для часових рядів. Використовуються екзогенні (допоміжні) часові ряди. Застосовуються специфічні методи попереднього оброблення (від згладжування ковзним середнім до перетворення Бокса-Кокса). Також слід мати можливість застосовувати in-sample і out-of-sample прогнозування;

Врахувати всі перераховані можливості в одному фреймворку, і, при цьому, не обмежуватися тільки часовими рядами, це є досить складне завдання.

2.4. Open-source фреймворк FEDOT для автоматичного машинного навчання (AutoML)

Заслуговує на увагу поняття про конвеєри для проблем машинного навчання. Конвеєр – це ациклічний орієнтований граф. У термінах FEDOT цей граф називається ланцюгом, або складеною моделлю, або конвеєром (рис. 2.3).



Рисунок 2.4 – Схема використання FEDOT

Основні абстракції, які використовує FEDOT:

1. Операція – це дія, яка виконується над даними. Це може бути дія для попередньої обробки даних (нормалізація, стандартизація, заповнення прогалів) або модель машинного навчання, яка дає прогнози;
2. Вузол – це контейнер, у якому розміщується операція. В одному вузлі може бути лише одна операція. Первинний вузол приймає лише необроблені дані, а вторинний вузол використовує вихід з вузлів попереднього рівня як предиктори;
3. Ланцюг або конвеєр – це ациклічний орієнтований граф, який складається з вузлів. Конвеєри машинного навчання у FEDOT реалізовані через клас Chain (рис. 2.5).

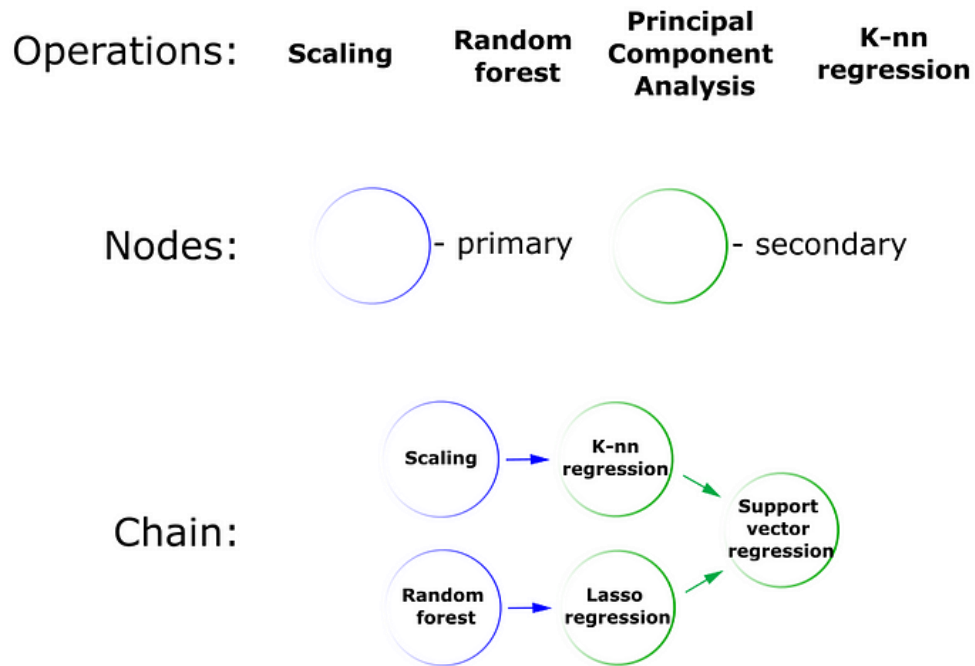


Рисунок 2.5 – Операції, вузли та ланцюжки (конвеєри) у фреймворку FEDOT

У структуру такого конвеєра можна вставити як моделі машинного навчання, так і класичні моделі, такі як авторегресія (AR) для часових рядів. Варто сказати, що заслуговує на увагу метод SSA, який використовує окремі перетворення. Цей підхід також використовується в одній із версій бібліотеки H2O. Застосування майже всіх моделей машинного навчання часових рядів полягає в побудові таких матриць (рис. 2.6).

Розглянемо цей спосіб перетворення рядів докладніше. Часовий ряд – це послідовність значень, де наступні значення зазвичай залежать від попередніх. Отже, ми можемо використовувати поточні та попередні елементи часового ряду, щоб зробити прогноз. Наприклад, потрібно передбачити ряд на один елемент вперед, використовуючи поточне та одне попереднє значення.

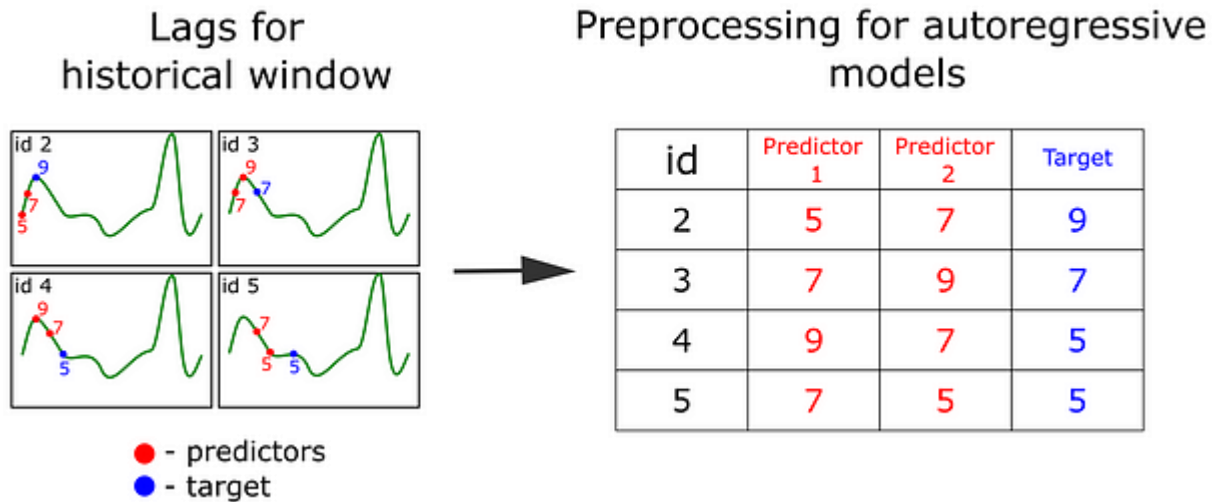


Рисунок 2.6 – Побудова матриць для моделей машинного навчання часових рядів

Таке перетворення називається «перетворенням із запізненням» часових рядів. У FEDOT є окрема операцію «відстає». Важливим гіперпараметром є розмір ковзного вікна, який визначає, скільки попередніх значень будемо використовувати як предиктори.

Будь-яку модель машинного навчання можна використовувати як прогностичну модель. Існує кілька конкретних моделей для прогнозування часових рядів у FEDOT (таких як AR та ARIMA). Крім того, були додані спеціальні методи попередньої обробки часових рядів, як згладжування ковзного середнього або згладжування за Гауссом.

2.5. Підготовка даних для узгодження обсягів резервного виробництва електроенергії автономними джерелами

Набір даних для узгодження обсягів резервного виробництва електроенергії автономними джерелами представляється багатовимірний часовий ряд, що відображає середньодобове вироблення електроенергії вітряною установкою та середньодобове вироблення електроенергії за допомогою дизельного генератора. Обидва показники вимірюються в кВт·год.

Крім дати і часу у нашому наборі є наступні дані, що описуються такими атрибутами:

1. diesel_fuel_kWh – середньодобове вироблення електроенергії за допомогою дизельного генератора (кВт·год);
2. wind_power_kWh – середньодобове вироблення електроенергії вітряною установкою (кВт·год).

Робота вітряного електрогенератора залежить від швидкості вітру, і якщо вона знижується, то для підтримки вироблення електроенергії на достатньому рівні використовується дизельний генератор. Звідси і простежується поведінка часових рядів – коли вироблення електроенергії на вітряку падає, на дизельному генераторі воно зростає, і навпаки. Варто також зауважити, що тимчасові ряди мають пропуски. Нам необхідно прогнозувати навантаження на дизельний генератор на 14 днів наперед. Таким чином, цільова змінна – вироблення енергії дизельним генератором та вітроенергетичною установкою.

Завантаження, підготовку та роботу із заданим набором даних здійснювали із використанням блокноту Jupyter Notebook. Для цього виконуємо імпорт необхідних для роботи бібліотек та функцій (рис. 2.7).

```
In [1]: # Additional imports
import os
import numpy as np
import pandas as pd
import seaborn as sns
from sklearn.metrics import mean_squared_error, mean_absolute_error
# Imports for creating plots
import matplotlib.pyplot as plt
from pylab import rcParams
rcParams['figure.figsize'] = 15, 7
import warnings
warnings.filterwarnings('ignore')
# fedot api
from fedot.api.main import Fedot
# Tasks to solve
from fedot.core.repository.tasks import Task, TaskTypeEnum, TsForecastingParams
# Input data for fit and predict
from fedot.core.data.data import InputData
# Data types in FEDOT
from fedot.core.repository.dataset_types import DataTypesEnum
# Pipeline and nodes
from fedot.core.pipelines.pipeline import Pipeline
from fedot.core.pipelines.node import PrimaryNode, SecondaryNode
# Data
from fedot.core.data.data import InputData
from fedot.core.data.data_split import train_test_data_setup
from fedot.core.repository.dataset_types import DataTypesEnum
# Tasks
from fedot.core.repository.tasks import Task, TaskTypeEnum, TsForecastingParams
# Metric
from sklearn.metrics import mean_absolute_error
```

Рисунок 2.7 – Імпорт необхідних для роботи бібліотек та функцій у блокнот Jupyter Notebook

Наша програма для прогнозування часових рядів базується на використанні бібліотеки Fedot та моделей машинного навчання. Нами імпортовано бібліотек os, numpy, pandas, seaborn (основні бібліотеки для роботи з даними та візуалізації), matplotlib (для створення графіків), warnings (для управління попередженнями), Fedot (основна бібліотека для розв'язання задач машинного навчання за допомогою автоматизованого машинного навчання AutoML). Представлений код (рис. 2.7) забезпечує імпорти для роботи з даними часових рядів, визначає параметри задачі та використовує метрики для оцінки результатів.

Наступним кроком є завантаження даних за допомогою функції read_csv. Після цього ми можемо переглянути перші п'ять рядків даних за допомогою функції dataset.head() (рис. 2.8).

```
In [4]: # Результат
df.head()

Out[4]:
```

	datetime	diesel_fuel_kWh	wind_power_kWh
0	2022-05-01	545.70	225.72
1	2022-05-02	558.45	212.76
2	2022-05-03	651.10	54.00
3	2022-05-04	580.55	47.52
4	2022-05-05	616.25	114.48

Рисунок 3.8 – Результати завантаження даних за допомогою функції read_csv та перегляду перших п'яти рядків даних за допомогою функції dataset.head()

Для виявлення та обробки пропущених значень, аномалій та інших проблем у даних ми використовували функції Pandas і NumPy. Першим кроком було визначення типу даних кожного поля (рис. 2.9).

```
In [7]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 397 entries, 0 to 396
Data columns (total 3 columns):
#   Column          Non-Null Count  Dtype
---  ---          -
0   datetime        397 non-null   datetime64[ns]
1   diesel_fuel_kWh  397 non-null   float64
2   wind_power_kWh  397 non-null   float64
dtypes: datetime64[ns](1), float64(2)
memory usage: 9.4 KB
```

Рисунок 2.9 – Результати визначення типу даних кожного поля

Дані містять інформацію про середньодобове вироблення електроенергії за допомогою дизельного генератора та вітряних електростанцій протягом 397 днів (рис. 2.10).

```
In [6]: df.describe()
```

```
Out[6]:
```

	datetime	diesel_fuel_kWh	wind_power_kWh
count	397	397.000000	397.000000
mean	2022-11-14 18:22:40.201511424	522.709320	316.752846
min	2022-05-01 00:00:00	-0.850000	-1.080000
25%	2022-08-08 00:00:00	317.050000	-1.080000
50%	2022-11-15 00:00:00	549.100000	184.680000
75%	2023-02-22 00:00:00	779.450000	528.120000
max	2023-05-31 00:00:00	1135.600000	1792.800000
std	NaN	303.843858	386.785012

Рисунок 2.10 – Результати визначення статистичних характеристик даних про середньодобове вироблення електроенергії дизельними генераторами та вітряних електростанцій

Середнє значення вироблення електроенергії дизельними генераторами становить 2452,71 кВт·год, а вітряними електростанціями – 316,75 кВт·год. Мінімальне значення вироблення електроенергії дизельними генераторами становить 0,85 кВт·год, а вітряними електростанціями -1,08 кВт·год. Це означає, що в один з днів вироблення електроенергії дизельними генераторами було мінусове, що є недопустимим.

Максимальне значення вироблення електроенергії дизельними генераторами становить 1135,6 кВт·год, а вітряними електростанціями – 1792,8 кВт·год. Стандартне відхилення вироблення електроенергії дизельними генераторами становить 303,84 кВт·год, а вітряними електростанціями – 386,79 кВт·год. Це означає, що більшість спостережень вироблення електроенергії дизельними генераторами знаходяться в межах 303,84 кВт·год від середнього значення, а більшість спостережень вироблення електроенергії вітряними електростанціями знаходяться в межах 386,79 кВт·год від середнього значення.

РОЗДІЛ 3.

РЕЗУЛЬТАТИ РОЗРОБКИ МОДЕЛЕЙ ПРОГНОЗУВАННЯ ОБСЯГІВ РЕЗЕРВНОГО ВИРОБНИЦТВА ЕЛЕКТРОЕНЕРГІЇ АВТОНОМНИМИ ДЖЕРЕЛАМИ НА ОСНОВІ ВИКОРИСТАННІ БІБЛІОТЕКИ FEDOT

3.1. Аналіз часових рядів обсягів резервного виробництва електроенергії

Для проведення аналізу часових рядів зміни обсягів резервного виробництва електроенергії застосовується бібліотека Fedot, спрямована на автоматизоване машинне навчання.

Процес аналізу часових рядів обсягів резервного виробництва електроенергії може включати такі етапи:

1. Завантаження та підготовка даних – використання бібліотеки Pandas для завантаження даних з відповідного джерела.
2. Перевірка та обробка пропущених значень, якщо такі є.
3. Перетворення даних у часовий ряд.

```
In [8]: def plot_dual_axis_time_series(df):
        """ Function for dual-axis time series visualization """

        fig, ax1 = plt.subplots()

        # Plot the first variable on the left y-axis
        line1, = ax1.plot(df['datetime'], df['diesel_fuel_kwh'], label='Diesel fuel', color='green', linewidth=2)
        ax1.set_xlabel('Datetime')
        ax1.set_ylabel('Diesel fuel, kWh', color='green')
        ax1.tick_params('y', colors='green')

        # Create a second y-axis to plot the second variable
        ax2 = ax1.twinx()
        line2, = ax2.plot(df['datetime'], df['wind_power_kwh'], label='Wind power', color='blue', linewidth=2)
        ax2.set_ylabel('Wind power, kWh', color='blue')
        ax2.tick_params('y', colors='blue')

        # Add grid
        ax1.grid(True)

        # Combine Legends from both axes
        lines = [line1, line2]
        ax1.legend(lines, [line.get_label() for line in lines], loc='upper left')

        plt.title('Electricity Generation Comparison')
        plt.show()

        # Call the dual-axis plot function
        plot_dual_axis_time_series(df)
```

Рисунок 3.1 – Код для візуалізації часового ряду обсягів резервного виробництва електроенергії

4. Візуалізація часового ряду – використання бібліотек Matplotlib або Seaborn для візуалізації часового ряду.
5. Аналіз тренду, сезонності та інших особливостей.

Графік візуалізації часового ряду обсягів резервного виробництва електроенергії представлено на рис. 3.2.

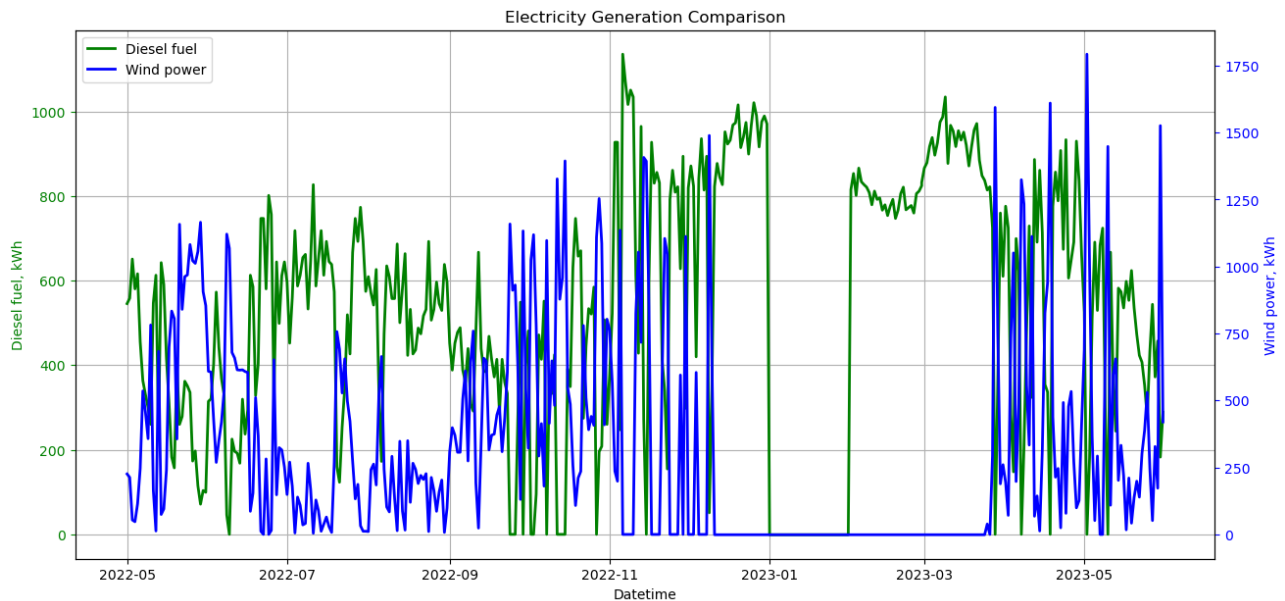


Рисунок 3.2 – Графік візуалізації часового ряду обсягів резервного виробництва електроенергії

Наявно два часові ряди, з яких перший – це середньодобове виробництво електроенергії за допомогою дизельного генератора (кВт.год), а другий часовий ряд – це виробництво електроенергії за допомогою вітрогенератора (енергія вітру).

Робота вітрогенератора залежить від сили вітру, а коли відсутній вітер, для підтримання виробництва електроенергії на достатньому рівні використовується дизель-генератор. Звідси можна простежити поведінку часових рядів. Коли виробництво електроенергії на вітрогенераторі падає, воно зростає на дизель-генераторі, і навпаки. Також варто зазначити, що часовий ряд має розриви.

Нами написано код, що використовує бібліотеки Matplotlib та Seaborn для візуалізації розподілу обсягів виробництва електроенергії вітроенергетичною установкою та генератором (рис. 3.3-3.6).

```
In [9]: plt.figure(figsize=(20,8))

plt.subplot(1,2,1)
plt.title('Schedule of electricity production by wind pover installations')
sns.histplot(df.wind_power_kWh)

plt.subplot(1,2,2)
plt.title('Spread electricity production by wind pove installations ')
sns.boxplot(y=df.wind_power_kWh)

plt.show()
```

Рисунок 3.3 – Код візуалізації розподілу обсягів виробництва електроенергії вітроенергетичною установкою

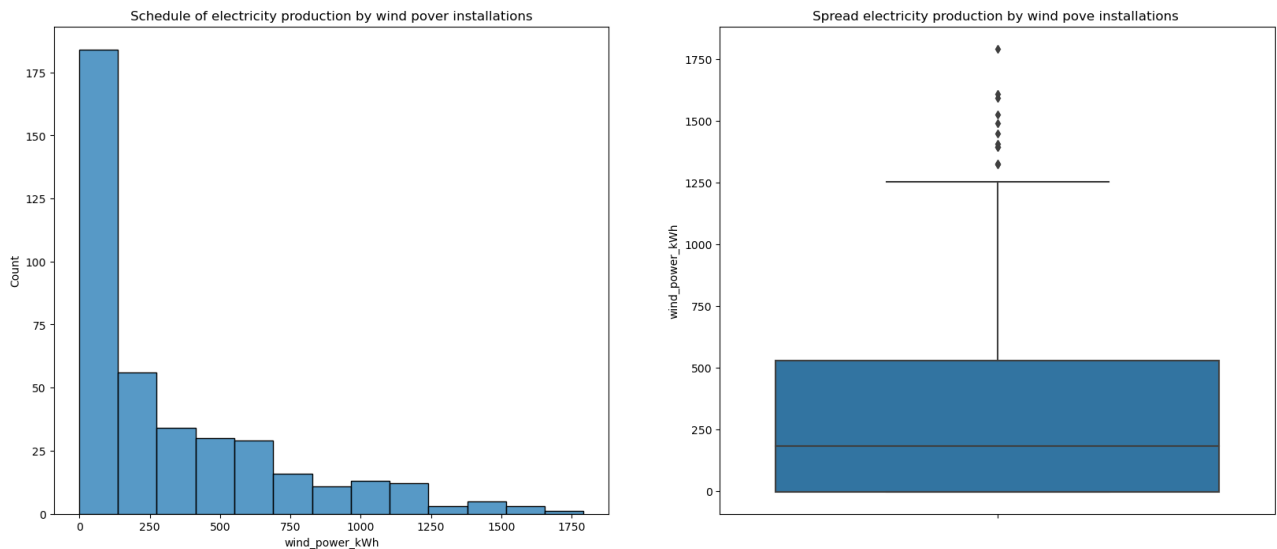


Рисунок 3.4 – Графік розподілу обсягів виробництва електроенергії вітроенергетичною установкою

```
In [10]: plt.figure(figsize=(20,8))

plt.subplot(1,2,1)
plt.title('Schedule of electricity generation by generators')
sns.histplot(df.diesel_fuel_kWh)

plt.subplot(1,2,2)
plt.title('Electricity generation by generators Spread')
sns.boxplot(y=df.diesel_fuel_kWh)

plt.show()
```

Рисунок 3.5– Код візуалізації розподілу обсягів виробництва електроенергії генератором

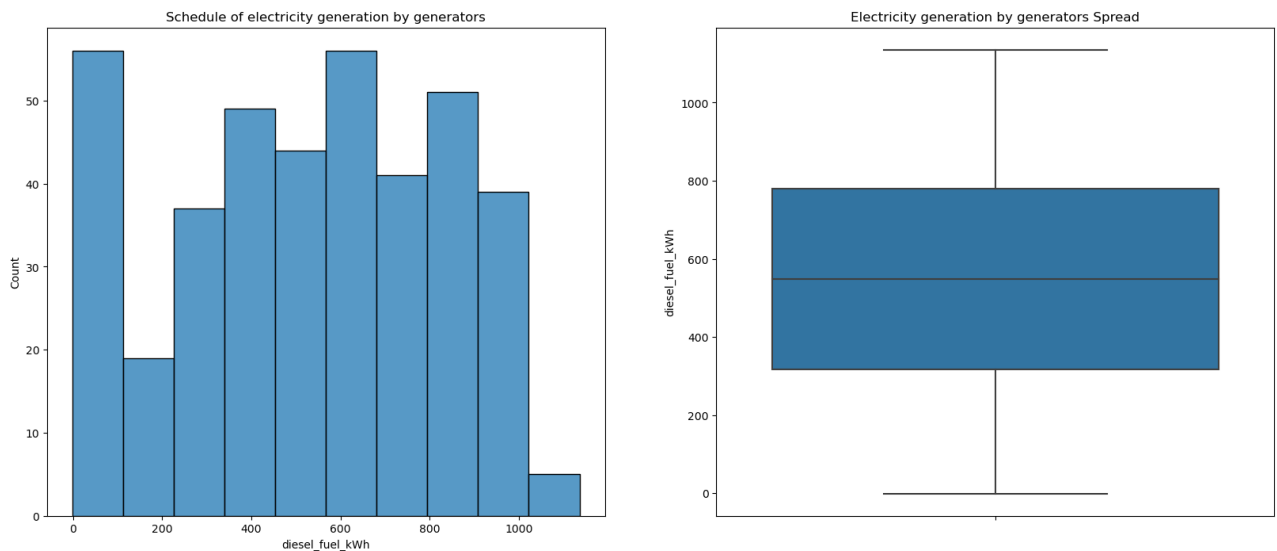


Рисунок 3.6 – Графік розподілу обсягів виробництва електроенергії генератором

Побудовані графіки дають можливість виконати аналіз розподілів обсягів виробництва електроенергії вітроенергетичною установкою та генератором за допомогою гістограм та boxplot.

Отримані гістограми вказують на розподіл значень. Зокрема, обсяг виробництва електроенергії вітроенергетичною установкою розподілений за законом Вейбулла, де понад 175 значень припадає на інтервал 0...125 кВт.год. Водночас обсяг виробництва електроенергії генератором описується рівномірним законом розподілу, де обсяги виробництва електроенергії генератором коливаються в межах 300...800 кВт.год. Графік із boxplot дозволяє визначити основні характеристики розподілу.

У нас є завдання підготувати модель для прогнозування обсягів виробництва електроенергії вітроенергетичною установкою на 7 днів вперед (для часового ряду `diesel_fuel_kWh` (червона крива)). Для цього слід позбутися пропусків. У FEDOT використовуються дві групи методів для відновлення значень у часових рядах:

- ✓ прості методи, такі як лінійна інтерполяція та фільтри Савицького-Голі;

✓ методи, що базуються на композитних моделях структури (просунуте заповнення пропусків).

Тепер заповнимо пропуски в часових рядах значеннями потужності вітрогенератора. Для цього ми зіставимо значення двох часових рядів і розв'яжемо задачу парної регресії.

```
In [12]: forecast_length = 14

In [13]: # Wrap data into InputData
task = Task(TaskTypesEnum.ts_forecasting,
            TsForecastingParams(forecast_length=forecast_length))

# Get time series from dataframe
traffic = np.array(df['diesel_fuel_kWh']) # Use the correct column name
input_data = InputData(idx=np.arange(0, len(traffic)),
                      features=traffic,
                      target=traffic,
                      task=task,
                      data_type=DataTypesEnum.ts)

# Split data into train and test
train_input, predict_input = train_test_data_setup(input_data)
```

Рисунок 3.7 – Код для вказування довжини прогнозу та розділення даних на навчальні та тестові вибірки

Наступний код задає `forecast_length = 14` довжину прогнозу, тобто кількість часових точок, на які ви хочете зробити прогноз. Після цього створюється об'єкт завдання (`Task`) для прогнозування часових рядів. Об'єкт `InputData` визначає дані, які використовуються для навчання та прогнозування моделі. У цьому випадку використовуються індекси часу, значення часового ряду (`features` та `target`) і вказується тип даних (`DataTypesEnum.ts`). Виконується розділення даних на навчальний та тестовий набори (рис. 3.8).

```
In [14]: print("Train data:")
print("Index range:", train_input.idx[0], "-", train_input.idx[-1])
print("Features shape:", train_input.features.shape)
print("Target shape:", train_input.target.shape)

print("\nTest data:")
print("Index range:", predict_input.idx[0], "-", predict_input.idx[-1])
print("Features shape:", predict_input.features.shape)
print("Target shape:", predict_input.target.shape)

Train data:
Index range: 0 - 382
Features shape: (383,)
Target shape: (383,)

Test data:
Index range: 383 - 396
Features shape: (383,)
Target shape: (14,)
```

Рисунок 3.8 – Код для виведення результатів розділення даних на навчальні та тестові вибірки

Використовується функція `train_test_data_setup` для розділення даних на навчальний (`train_input`) та тестовий (`predict_input`) набори.

3.2. Результати прогнозування обсягів виробництва електроенергії генератором

Наш часовий ряд досить простий для інтерпретації. Тому ми можемо використати наші знання про циклічність, щоб встановити лаги вручну. Пропонуємо наступні лаги: 140 елементів та 7 елементів (рис. 3.9).

```
In [16]: # Initialise pipeline for making forecasting
pipeline = get_two_branches_pipeline(lagged_param_1=60,
                                   lagged_param_2=14)

# Fit
pipeline.fit(train_input)

# Make forecast
output = pipeline.predict(predict_input)
forecast = np.ravel(np.array(output.predict))

2023-12-29 17:20:28,391 - From C:\Users\trian\AppData\Roaming\Python\Python311\site-packages\keras\src\losses.py:2976: The name
tf.losses.sparse_softmax_cross_entropy is deprecated. Please use tf.compat.v1.losses.sparse_softmax_cross_entropy instead.
```

Рисунок 3.9 – Код для встановлення лаг

Насамперед виконаємо прогноз обсягів виробництва електроенергії генератором використовуючи навчену модель. Задаємо прогноз на 7 елементів вперед (рис. 3.10).

```
In [15]: def get_two_branches_pipeline(lagged_param_1=10, lagged_param_2=10,
                                     model_1='ridge', model_2='ridge', model_3='ridge'):
    """
    Pipeline looking like this
    lagged -> model_1 \
                    \ model_3 -> final forecast
                    /
    lagged -> model_2 /
    """
    # First Level
    node_lagged_1 = PrimaryNode('lagged')
    node_lagged_1.parameters = {'window_size': lagged_param_1}
    node_lagged_2 = PrimaryNode('lagged')
    node_lagged_2.parameters = {'window_size': lagged_param_2}

    # Second Level
    node_second_1 = SecondaryNode(model_1, nodes_from=[node_lagged_1])
    node_second_2 = SecondaryNode(model_2, nodes_from=[node_lagged_2])

    # Third Level - root node
    node_final = SecondaryNode(model_3, nodes_from=[node_second_1, node_second_2])
    pipeline = Pipeline(node_final)

    return pipeline
```

Рисунок 3.10 – Код навчання моделі прогнозування за допомогою пайплайна з двох гілок (branches) з лагів та різних моделей для кожної гілки

Написаний нами код (рис. 3.10) виконує задачу прогнозування за допомогою пайплайна, який складається з двох гілок (branches) з лагів та різних моделей для кожної гілки.

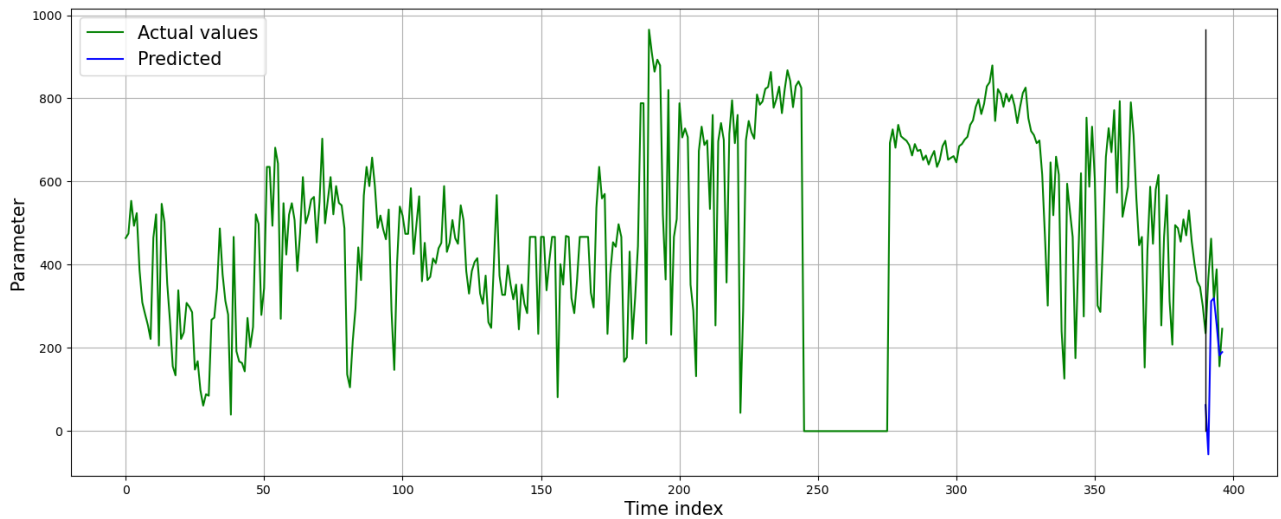


Рисунок 3.11 – Графік із результатами прогнозу обсягів виробництва електроенергії генератором використовуючи навчену модель

У результаті здійснюється прогноз за допомогою навченої моделі на тестових даних (`predict_input`) (рис. 3.11). Результат прогнозу знаходиться в змінній `forecast`. Цей код видає прогнозовані значення для заданого пайплайна та вхідних даних. Встановлено, що середня абсолютна похибка становить 137.9.

Спробуємо знайти оптимальні параметри за допомогою алгоритму налаштування гіперпараметрів у вузлах. Нами використано клас `PipelineTuner` (і клас `TunerBuilder`) для налаштування гіперпараметрів у вузлах конвеєра. Ці класи дозволяють налаштовувати гіперпараметри у вузлах за допомогою одночасного тюнінгу (рис. 3.12).

```
In [19]: from fedot.core.pipelines.tuning.tuner_builder import TunerBuilder
from golem.core.tuning.simultaneous import SimultaneousTuner
from fedot.core.repository.quality_metrics_repository import \
    MetricsRepository, RegressionMetricsEnum

init_pipeline = get_two_branches_pipeline()

# Start tuning all nodes in the pipeline
tuner = TunerBuilder(train_input.task)\
    .with_tuner(SimultaneousTuner)\
    .with_metric(RegressionMetricsEnum.MAE)\
    .with_iterations(50) \
    .build(train_input)
tuned_pipeline = tuner.tune(init_pipeline)

#final fit
tuned_pipeline.fit(train_input)
print('Pipeline tuned')
```

Рисунок 3.12 – Код навчання моделі прогнозування із налаштуванням гіперпараметрів у вузлах

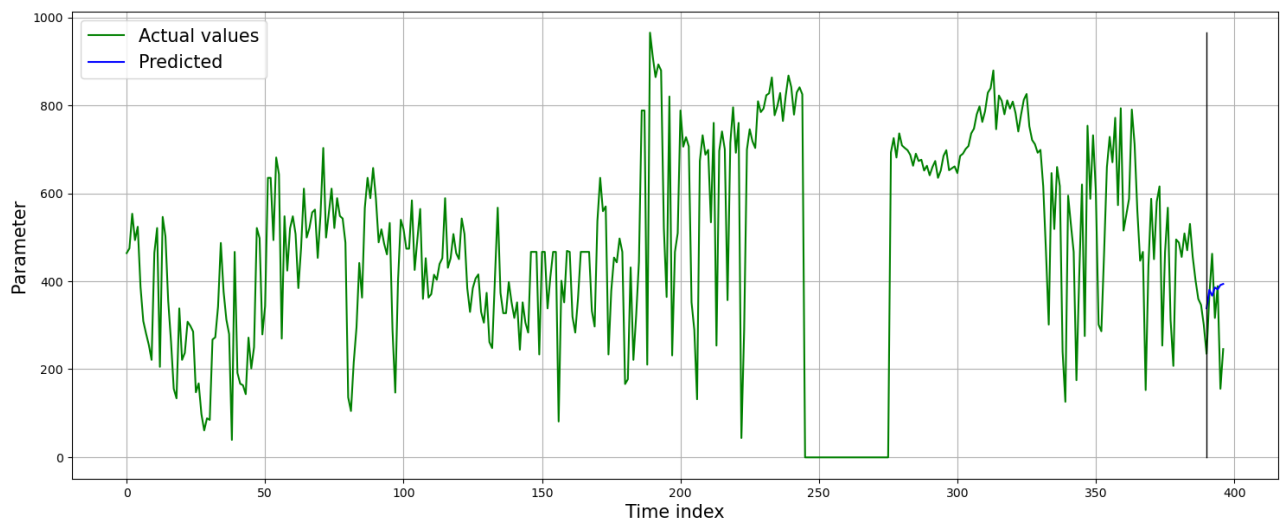


Рисунок 3.13 – Графік із результатами прогнозу обсягів виробництва електроенергії генератором із налаштуванням гіперпараметрів у вузлах

Встановлено, що середня абсолютна похибка становить Середня абсолютна похибка: 95.99.

Для прогнозування часових рядів, як і для інших задач, ми можемо використовувати API фреймворку FEDOT (простий підхід), або будувати моделі на основі внутрішніх методів ядра (просунутий підхід).

Використовуємо простий підхід (приклад API) для навчання моделі (рис. 3.14).

```
In [24]: # Define parameters
task_parameters = TsForecastingParams(forecast_length=forecast_length)

# Init model for the time series forecasting
model = Fedot(problem='ts_forecasting', task_params=task_parameters, timeout=1)

# Run AutoML model design in the same way
pipeline = model.fit(features=train_input)
```

Рисунок 3.14 – Код навчання моделі прогнозування із використанням простого підходу (приклад API)

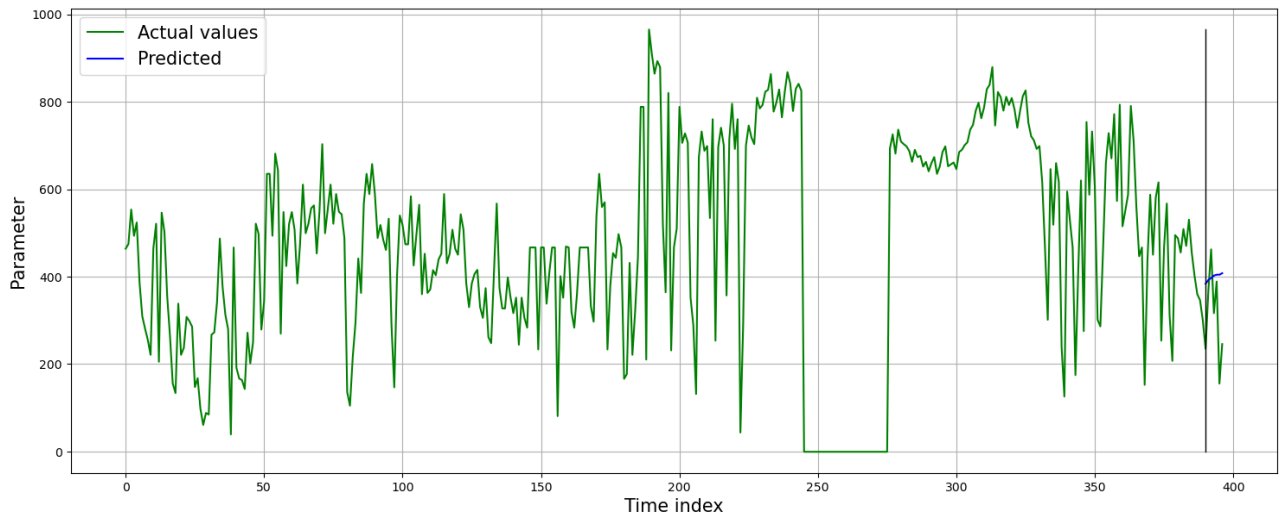


Рисунок 3.15 – Графік із результатами прогнозу обсягів виробництва електроенергії генератором із використанням простого підходу (приклад API)

Встановлено, що середня абсолютна похибка становить 107.4.

У Fedot за створення конвеєрів відповідає еволюційний алгоритм. Цей еволюційний алгоритм генерує початкову популяцію, що складається з конвеєрів, а потім застосовує оператори мутації, відбору та кросовера для пошуку оптимального рішення. У той же час, існує дуже велика кількість модифікацій еволюційного алгоритму. Наприклад, можна використовувати тільки один оператор мутації, або використовувати певну схему. Розглянемо конкретний приклад, де у Fedot існує декілька операторів мутації, а саме

Задамо метрику для оптимізації – MAE (рис. 3.16). Запускаємо алгоритм для пошуку оптимального рішення.

```

In [33]: metric_function = MetricsRepository().metric_by_id(RegressionMetricsEnum.MAE)
builder = ComposerBuilder(task=task). \
    with_optimizer_params(parameters=optimiser_parameters). \
    with_requirements(composer_requirements). \
    with_metrics(metric_function).with_initial_pipelines([init_pipeline])
composer = builder.build()

obtained_pipeline = composer.compose_pipeline(data=train_input)

In [34]: # Fit pipeline on the entire dataset
obtained_pipeline.fit(train_input)

# Make prediction
obtained_output = obtained_pipeline.predict(predict_input)
obtained_forecast = np.ravel(np.array(obtained_output.predict))

plot_results(actual_time_series = traffic,
             predicted_values = obtained_forecast,
             len_train_data = len(traffic)-forecast_length)

# Print MAE metric
print(f'Mean absolute error: {mean_absolute_error(predict_input.target, obtained_forecast):.3f}')

```

Рисунок 3.16 – Код навчання моделі прогнозування із використанням
КОМПОЗИТОРА

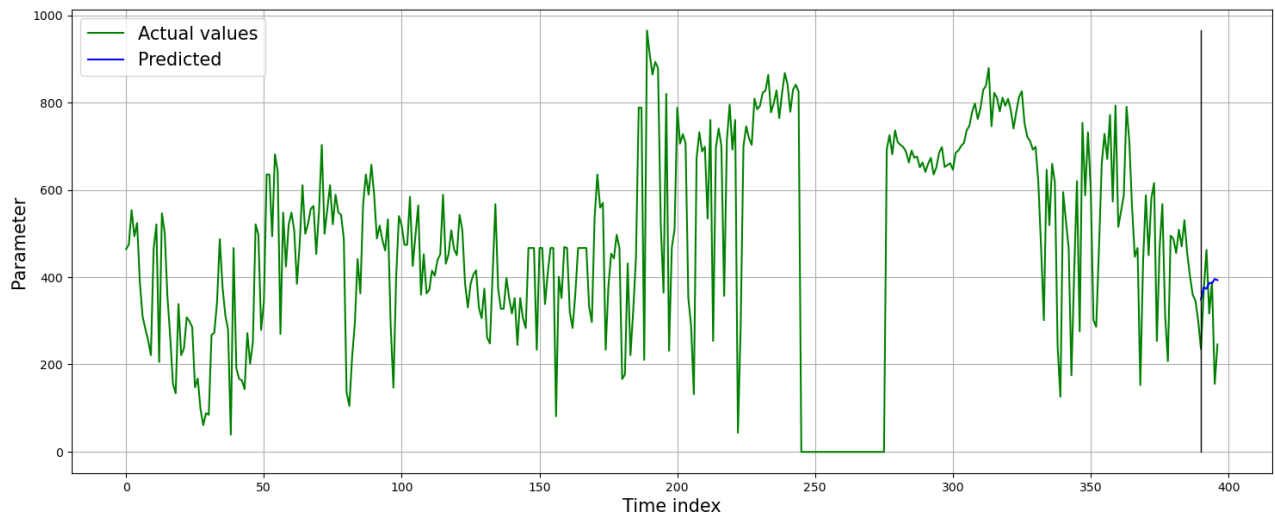


Рисунок 3.17 – Графік із результатами прогнозу обсягів виробництва
електроенергії генератором із використанням композитора

Встановлено, що середня абсолютна похибка становить 96.07. Як видно з метрики MAE, нам вдалося покращити точність прогнозування за допомогою композитора.

3.3. Результати прогнозування обсягів виробництва електроенергії вітроустановками

Із використанням аналогічних підходів та із вище поданим кодом виконано прогнозування обсягів виробництва електроенергії вітроустановками. У результаті здійснюється прогноз за допомогою навченої моделі на тестових даних (`predict_input`) (рис. 3.18).

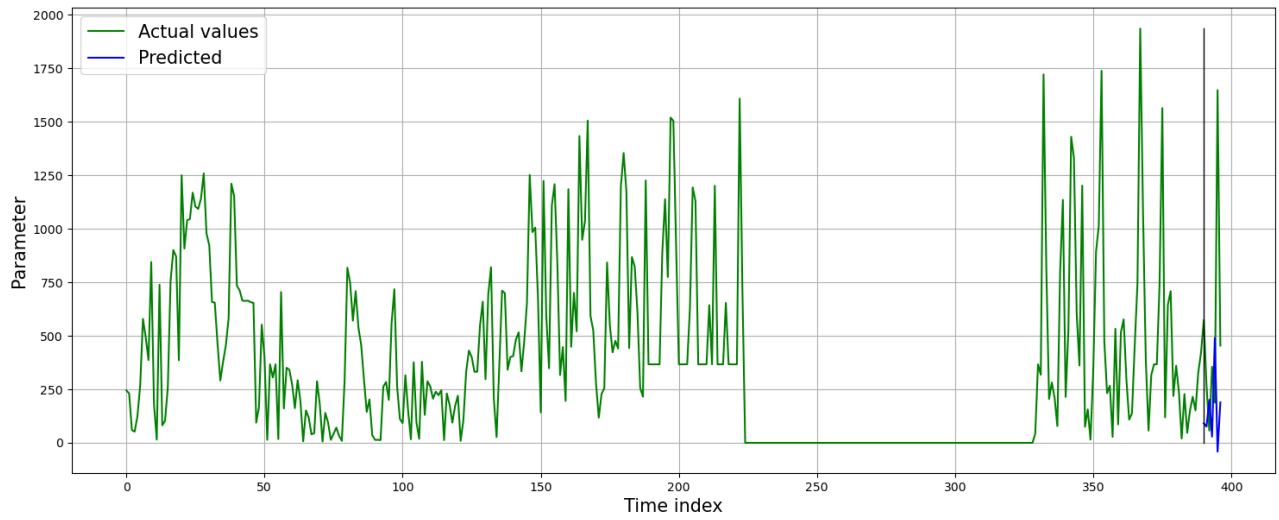


Рисунок 3.18 – Графік із результатами прогнозу обсягів виробництва електроенергії вітроустановками використовуючи навчену модель

Встановлено, що середня абсолютна похибка становить 406.12.

Знаходимо оптимальні параметри за допомогою алгоритму налаштування гіперпараметрів у вузлах. Нами використано клас `PipelineTuner` (і клас `TunerBuilder`) для налаштування гіперпараметрів у вузлах конвеєра. Ці класи дозволяють налаштовувати гіперпараметри у вузлах за допомогою одночасного тюнінгу (рис. 3.19).

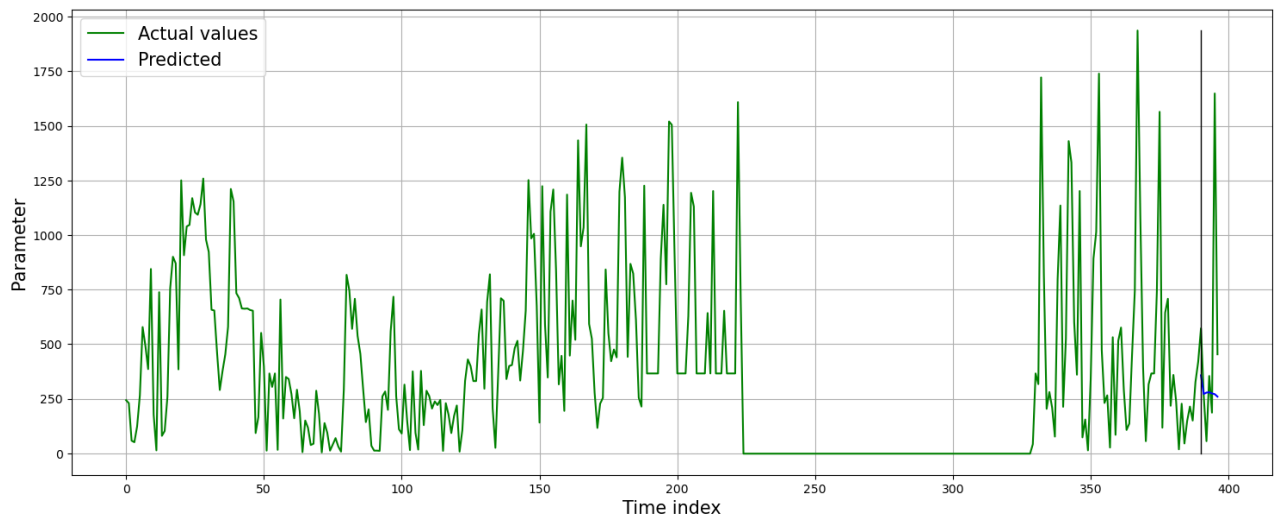


Рисунок 3.19 – Графік із результатами прогнозу обсягів виробництва електроенергії вітроустановками із налаштуванням гіперпараметрів у вузлах

Встановлено, що середня абсолютна похибка становить Середня абсолютна похибка – 311.0.

Використовуємо простий підхід (приклад API) для навчання моделі (рис. 3.20).

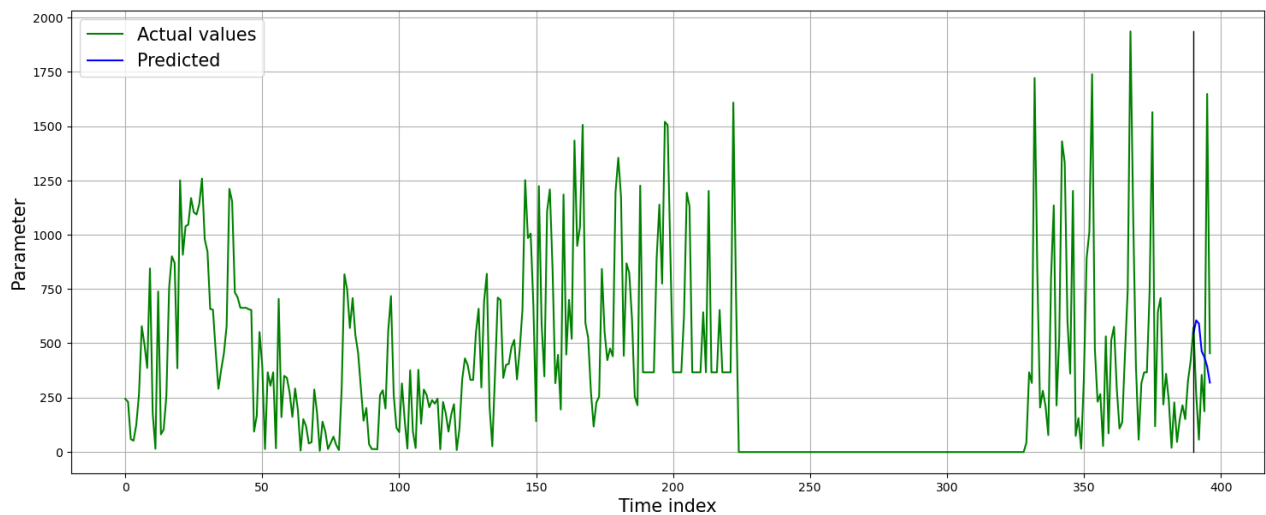


Рисунок 3.20 – Графік із результатами прогнозу обсягів виробництва електроенергії вітроустановками із використанням простого підходу API

Встановлено, що середня абсолютна похибка становить 309.2.

Розглянемо конкретний приклад, де у Fedot існує декілька операторів мутації, а саме навчання моделі прогнозування із використанням композитора (рис. 3.21).

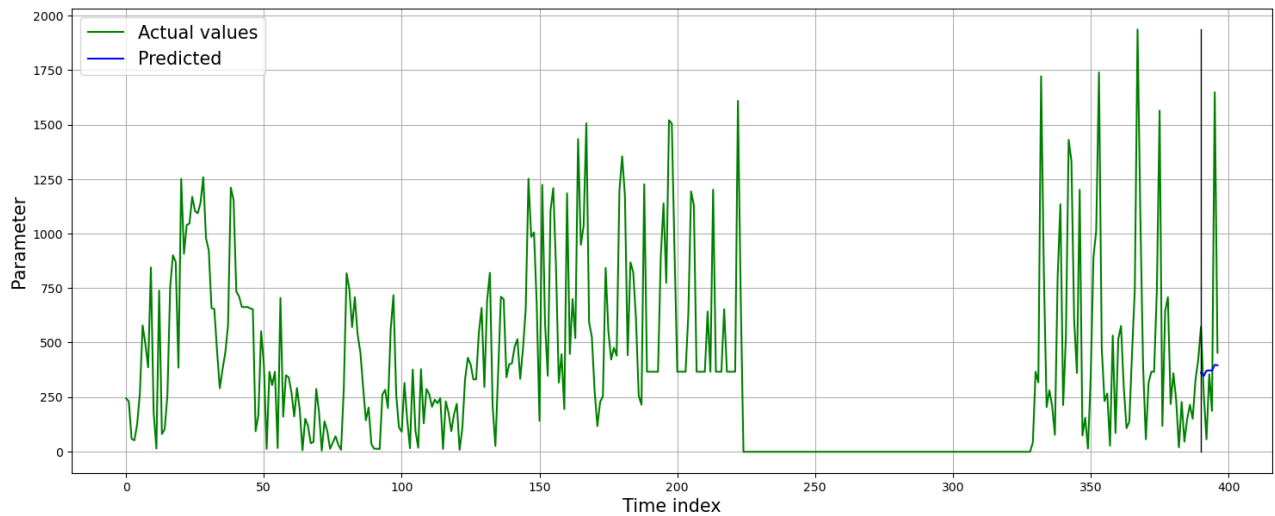


Рисунок 3.21 – Графік із результатами прогнозу обсягів виробництва електроенергії вітроустановками із використанням композитора

Встановлено, що середня абсолютна похибка становить 302.4. Як видно з метрики MAE, нам вдалося покращити точність прогнозування за допомогою композитора.

3.3. Результати визначення раціональної моделі прогнозування обсягів виробництва електроенергії

На підставі розроблених моделей машинного навчання для прогнозу обсягів виробництва електроенергії генераторами та вітроустановками встановлено їх показники, які подано у табл. 3.1.

Таблиця 3.1 – Результати визначення показників розроблених моделей машинного навчання для прогнозу обсягів виробництва електроенергії генераторами та вітроустановками

Назва моделі	Показники моделі			
	RMSE	MAE	MAPE	SMAPE
Прогноз обсягів виробництва електроенергії генераторами				
Fedot із пайплайномі із двох гілок (branches)	190.530	137.982	42.346%	62.687%
Fedot із налаштуванням гіперпараметрів у вузлах конвеєра	120.979	95.999	43.440%	30.962%
Fedot із використанням підходу API	179.543	155.613	67.945%	43.959%
Fedot із використанням композитора	178.932	149.198	66.056%	42.436%
Прогноз обсягів виробництва електроенергії вітроустановками				
Fedot із пайплайномі із двох гілок (branches)	698.612	486.142	118.414%	130.422%
Fedot із налаштуванням гіперпараметрів у вузлах конвеєра	540.149	311.008	90.230%	63.008%
Fedot із використанням підходу API	536.317	317.747	103.937%	64.200%
Fedot із використанням композитора	497.781	347.021	157.361%	69.771%

Найкращою моделлю для прогнозу обсягів виробництва електроенергії генераторами є та, яка використовує налаштування гіперпараметрів у вузлах конвеєра. Ця модель досягає найменших значень Root Mean Squared Error

(RMSE) та Mean Absolute Error (MAE), відповідно 120.979 та 95.999. Крім того, величини Mean Absolute Percentage Error (MAPE) та Symmetric Mean Absolute Percentage Error (SMAPE) становлять 43.440% та 30.962%, відображаючи високий рівень точності прогнозу.

Для прогнозу обсягів виробництва електроенергії вітроустановками, також модель із налаштуванням гіперпараметрів у вузлах конвеєра виявляється найкращою. Вона досягає найменших значень RMSE та MAE – 540.149 та 311.008 відповідно, а також відзначається MAPE та SMAPE на рівні 90.230% та 63.008%.

Обрані показники свідчать про високу точність та ефективність обраних моделей для обох типів прогнозів.

РОЗДІЛ 4.

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Небезпечні і шкідливі виробничі чинники під час створення моделей машинного навчання

Створення моделей машинного навчання може бути пов'язане з рядом небезпечних і шкідливих виробничих чинників. Ці чинники можуть призвести до травм, захворювань і навіть смерті працівників (табл. 4.1).

Таблиця 4.1 – Небезпечні та шкідливі виробничі чинники

Фізичні	електричний струм, шум, вібрація, підвищена температура та запиленість
Хімічні	відсутні
Біологічні	відсутні
Психофізіологічні	перевантаження, монотонність праці, стрес

До фізичних небезпечних і шкідливих виробничих чинників, пов'язаних із створенням моделей машинного навчання, належать:

✓ Електричний струм. Працівники, які працюють з обладнанням для машинного навчання, можуть бути піддані дії електричного струму. Це може призвести до електротравм, у тому числі до летальних випадків.

✓ Фізичні фактори. Працівники, які працюють з обладнанням для машинного навчання, можуть бути піддані дії фізичних факторів, таких як шум, вібрація, підвищена температура та запиленість. Ці фактори можуть призвести до захворювань органів слуху, опорно-рухового апарату, дихальної системи та інших захворювань.

До психофізіологічних небезпечних і шкідливих виробничих чинників, пов'язаних із створенням моделей машинного навчання, належать:

✓ **Перевантаження праці.** Працівники, які працюють над створенням моделей машинного навчання, можуть бути піддані перевантаженню праці. Це може призвести до перевтоми, зниження працездатності та інших захворювань.

✓ **Монотонність праці.** Працівники, які працюють над створенням моделей машинного навчання, можуть бути піддані монотонності праці. Це може призвести до зниження працездатності, погіршення зору та інших захворювань.

✓ **Стрес.** Працівники, які працюють над створенням моделей машинного навчання, можуть бути піддані стресу. Це може призвести до нервових розладів, захворювань серцево-судинної системи та інших захворювань.

4.2. Заходи щодо попередження небезпечних і шкідливих виробничих чинників

Для попередження небезпечних і шкідливих виробничих чинників під час створення моделей машинного навчання необхідно виконати ряд заходів. Забезпечити належне утримання обладнання та робочих місць. Устаткування для машинного навчання та робочі місця повинні регулярно перевірятися та утримуватися в належному стані. Це допоможе запобігти виникненню електротравм, отруєнь, опіків, травм опорно-рухового апарату та інших травм.

Запровадити заходи щодо захисту працівників від дії фізичних факторів. Працівники, які працюють з обладнанням для машинного навчання, повинні бути забезпечені засобами індивідуального захисту (ЗІЗ). Це допоможе запобігти дії шуму, вібрації, підвищеної температури та запиленості.

Забезпечити раціональну організацію праці. Працівники, які працюють над створенням моделей машинного навчання, повинні дотримуватися норм тривалості робочого часу, робочого тижня та інших норм трудового

законодавства. Це допоможе запобігти перевантаженню праці, монотонності праці та стресу.

Створення моделей машинного навчання може бути пов'язане з рядом небезпечних і шкідливих виробничих чинників. Ці чинники можуть призвести до травм, захворювань і навіть смерті працівників. Для попередження небезпечних і шкідливих виробничих чинників необхідно забезпечити належне утримання обладнання та робочих місць, запровадити заходи щодо захисту працівників від дії фізичних факторів та забезпечити раціональну організацію праці.

4.3. Розробка логічно-імітаційної моделі процесу виникнення травм під час монтажу інтелектуальної інформаційної системи узгодження обсягів резервного виробництва електроенергії автономними джерелами

Для побудови логіко-імітаційної моделі процесу, формування і виникнення аварії та травми під час монтажу інтелектуальної інформаційної системи узгодження обсягів резервного виробництва електроенергії автономними джерелами складемо список базових подій. Вони лежатимуть у основі даної моделі. Кожному пункту списку присвоюємо певне значення ймовірності виникнення. Нижче подано сам список:

- | | |
|--|---------------|
| 1. Стан контролю з охорони праці | $P_1 = 0,2;$ |
| 2. Несерйозне відношення до проходження ТО інструменту | $P_2 = 0,1;$ |
| 3. Відсутність комплектуючих установки..... | $P_3 = 0,2;$ |
| 4. Невисока міцність | $P_4 = 0,03;$ |

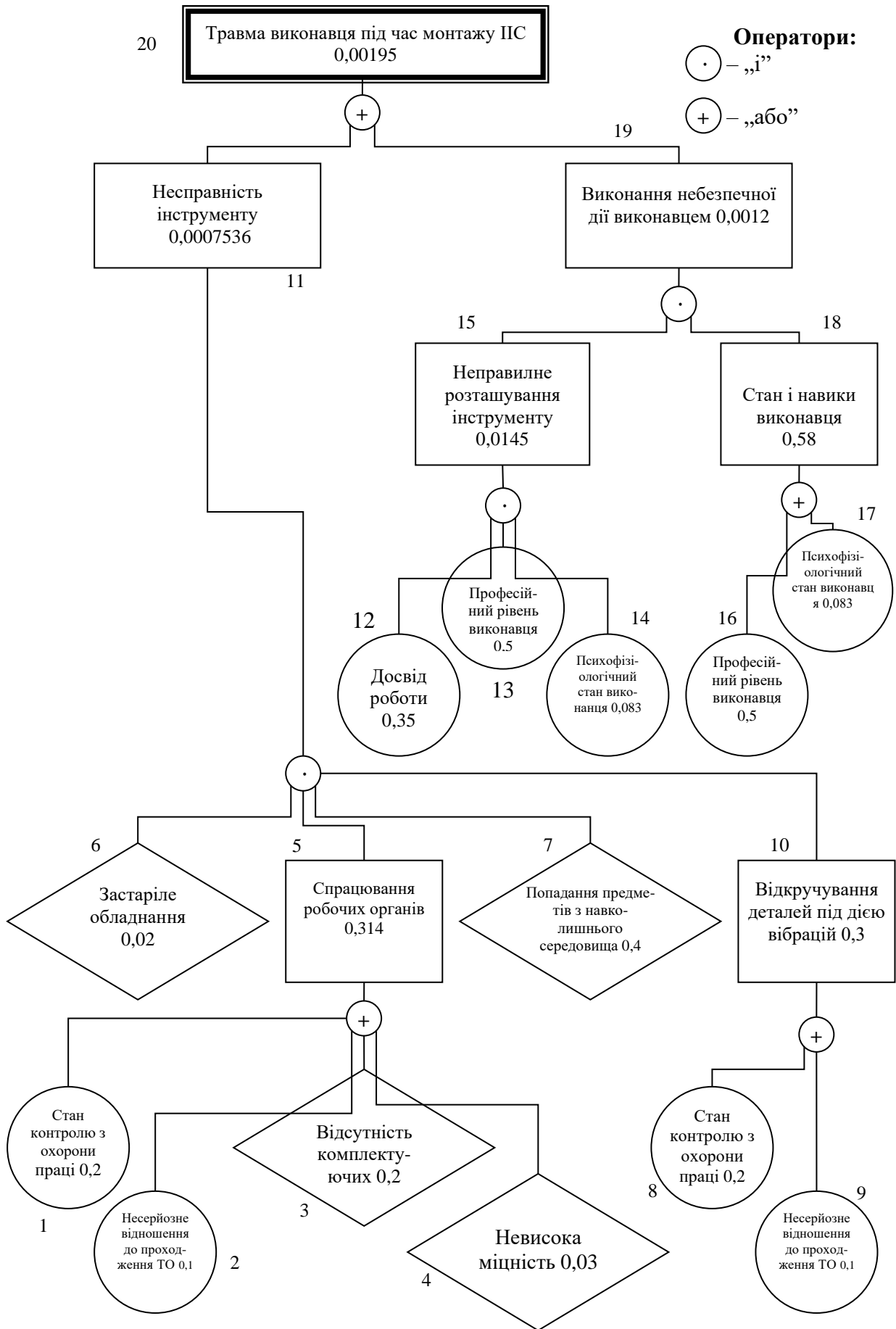


Рис. 4.1. Логіко-імітаційна модель процесу формування та виникнення аварії та травми під час монтажу інтелектуальної інформаційної системи узгодження обсягів резервного виробництва електроенергії автономними джерелами

1. Використання застарілого обладнання..... $P_6 = 0,02$;
2. Попадання сторонніх предметів $P_7 = 0,4$;
3. Досвід роботи виконавця $P_{12} = 0,35$.
4. Професійний рівень виконавця $P_{13} = 0,5$;
5. Психофізіологічний стан виконавця..... $P_{14} = 0,083$;

На основі даного списку будуємо матрицю логічних взаємозв'язків між окремими пунктами, графічне представлення якої зображено на рис. 4.1.

Розрахуємо ймовірності виникнення подій, що входять у дану логіко-імітаційну модель процесу монтажу інтелектуальної інформаційної системи узгодження обсягів резервного виробництва електроенергії автономними джерелами (на прикладі ймовірності отримання травми виконавця).

Ймовірність виникнення події P_5 визначаємо наступним чином:

$$P_5 = 0,2 + 0,1 + 0,2 + 0,003 - 0,2 \cdot 0,1 - 0,2 \cdot 0,03 - 0,2 \cdot 0,03 - 0,1 \cdot 0,2 - 0,1 \cdot 0,03 - 0,2 \cdot 0,03 + 0,2 \cdot 0,1 \cdot 0,2 + 0,1 \cdot 0,2 \cdot 0,03 + 0,2 \cdot 0,1 \cdot 0,2 + 0,2 \cdot 0,1 \cdot 0,03 - 0,2 \cdot 0,1 \cdot 0,2 \cdot 0,03 = 0,314$$

Ймовірність виникнення події P_{10} визначаємо так:

$$P_{10} = 0,2 + 0,1 = 0,3.$$

Ймовірність виникнення події P_{11} визначаємо:

$$P_{11} = 0,02 \cdot 0,314 \cdot 0,4 \cdot 0,3 = 0,00075.$$

Ймовірність виникнення події P_{15} визначаємо наступним чином:

$$P_{15} = 0,35 \cdot 0,5 \cdot 0,083 = 0,0145.$$

Ймовірність події P_{18} :

$$P_{18} = 0,5 + 0,083 = 0,58.$$

Ймовірність події P_{19} :

$$P_{19} = 0,0145 \cdot 0,083 = 0,0012.$$

Ймовірність події P_{20} :

$$P_{20} = 0,00075 + 0,0012 = 0,00195.$$

Ймовірність травми рівна ймовірності виникнення аварії, бо остання можлива лише за умови монтажу автоматизованої системи управління енергоспоживанням людиною.

Логіко-імітаційні моделі аварій і травм допомагають зменшити ймовірність виникнення аварійних та травмонебезпечних ситуацій. Якщо необхідно оцінити рівень небезпеки будь-якого робочого місця, слід уважно вивчити і побудувати логічні моделі можливих небезпечних ситуацій, які охоплюють як стан обладнання і самого робочого місця, так і поведінку працюючого і обчислити ймовірність виникнення травми.

Після аналізу результатів моделювання ймовірність виникнення травми можна звести до дуже малої величини – достатньо зменшити вплив ймовірностей вихідних факторів, які до неї призводять.

4.4. Розробка заходів із забезпечення безпеки під час надзвичайних ситуацій

Заходи щодо безпеки у надзвичайних ситуаціях – це заходи, які вживаються для запобігання, підготовки до, реагування на та відновлення після надзвичайних ситуацій. Ці заходи спрямовані на захист людей, майна та навколишнього середовища від потенційних збитків.

Розробка заходів щодо безпеки у надзвичайних ситуаціях передбачає кілька етапів. Насамперед виконується оцінка ризиків. На цьому етапі проводиться оцінка потенційних ризиків, які можуть виникнути внаслідок надзвичайних ситуацій. Оцінка ризиків включає в себе аналіз потенційних загроз, ймовірності їх виникнення та можливих наслідків.

Після цього виконується розробка планів. На основі результатів оцінки ризиків розробляються плани заходів щодо безпеки у надзвичайних ситуаціях. Плани заходів включають в себе заходи щодо запобігання, підготовки до, реагування на та відновлення після надзвичайних ситуацій.

Важливим етапом є впровадження планів. Розроблені плани заходів щодо безпеки у надзвичайних ситуаціях повинні бути впроваджені в практику. Це означає, що всі зацікавлені сторони, включаючи персонал, повинні бути проінформовані про плани та навчені їх виконувати.

Завершальним етапом є відстеження та оцінка. Плани заходів щодо безпеки у надзвичайних ситуаціях повинні регулярно відстежуватися та оцінюватися. Це допоможе забезпечити їх ефективність та внести необхідні зміни.

Заходи щодо безпеки у надзвичайних ситуаціях є важливим елементом будь-якої організації. Вони допомагають захистити людей, майно та навколишнє середовище від потенційних збитків, а також зменшити негативні наслідки надзвичайних ситуацій.

РОЗДІЛ 5.
ВИЗНАЧЕННЯ ЕФЕКТИВНОСТІ ВІД ВИКОРИСТАННЯ МОДЕЛІ
МАШИННОГО НАВЧАННЯ ДЛЯ УЗГОДЖЕННЯ ОБСЯГІВ
РЕЗЕРВНОГО ВИРОБНИЦТВА ЕЛЕКТРОЕНЕРГІЇ АВТОНОМНИМИ
ДЖЕРЕЛАМИ

Ефективність від використання моделі машинного навчання для узгодження обсягів резервного виробництва електроенергії автономними джерелами визначається як співвідношення між вигодами, отриманими від використання моделі, та витратами на її розробку та впровадження.

Ефективність від використання моделі машинного навчання для узгодження обсягів резервного виробництва електроенергії автономними джерелами визначається за наступною формулою:

$$E = (B - C) / C, \quad (5.1)$$

де E – ефективність;

B – вигоди від використання моделі;

C – витрати на розробку та впровадження моделі.

Припустимо, що витрати на збирання та підготовку даних склали 10 тис. грн, витрати на навчання моделі склали 200 тис. грн, а витрати на впровадження моделі склали 30 тис. грн. Вигоди від використання моделі оцінюються в 300 тис. грн на рік.

У цьому випадку ефективність від використання моделі машинного навчання буде становити:

$$E = (300 - 240) / 240 = 0,25.$$

Тобто, ефективність становить 25%. Це означає, що кожні 1 грн, витрачені на розробку та впровадження моделі, приносять 0,25 грн вигод.

Загалом, використання моделі машинного навчання може бути ефективним способом узгодження обсягів резервного виробництва електроенергії автономними джерелами. Однак, для того, щоб досягти високої ефективності, необхідно ретельно підібрати підходи до машинного навчання та забезпечити якість даних.

ВИСНОВКИ І ПРОПОЗИЦІЇ

У сучасних системах електропостачання зростає роль автономних джерел енергії, таких як вітрові електростанції. Вони характеризуються високою непередбачуваністю виробництва електроенергії, що може призвести до дисбалансу між попитом і пропозицією та, як наслідок, до аварій на енергосистемі. Для забезпечення стабільної роботи енергосистеми необхідно узгоджувати обсяги виробництва електроенергії автономними джерелами енергії з обсягами споживання. Це завдання є складним і трудомістким, оскільки вимагає постійного моніторингу та аналізу великої кількості даних.

Наша кваліфікаційна робота спрямована на вивчення та дослідження питань, пов'язаних із підвищенням ефективності процесу прогнозування обсягів резервного виробництва електроенергії автономними джерелами на основі автоматизованого машинного навчання. Робота включає в себе аналіз підходів, а також розглядає автоматизоване машинного навчання та розвиває методи та технології для підвищення ефективності використання автономних джерел електроенергії та оптимізації їх роботи на основі машинного навчання.

Виконано аналіз стану предметної області та узгодження обсягів виробництва електроенергії. Встановлено, що стосовно енергозабезпечення об'єктів виконують прогнози, які стосуються короткострокового прогнозування навантаження (STLF), що є корисним для додатків у реальному часі, таких як керування генеруючими установками електроенергії. Середньостроковий прогноз може виконуватися для бюджетування ресурсів, а довгостроковий – для планування розширення мережі енергосистеми.

Дослідження в галузі автоматизованого машинного навчання є дуже різноманітними та пропонують пакети та методи, орієнтовані як на дослідників, так і на кінцевих користувачів. Нами проаналізовано можливі сценарії машинного навчання. Моделі зі структурою типу (рис. 2.1, б) і (рис. 2.1, в) фактично стають композитними, тобто містять у собі різні алгоритми машинного навчання. Наприклад, можна об'єднати байєсівську мережу і

згорткову для передбачення на мультимодальних даних. А з композитними моделями та пайплайнами машинного навчання можна працювати за допомогою методів і технологій AutoML.

Провели аналіз існуючих AutoML фреймворків та здійснили обґрунтування його вибору. Представлена таблиця 2.1 містить інформацію про різні інструменти автоматизації машинного навчання (AutoML). Кожен інструмент має свої унікальні характеристики та особливості, такі як методи оптимізації, типи пайплайнів, підтримувані типи даних та додаткові можливості. Вибір конкретного інструменту може залежати від конкретних вимог та властивостей задачі машинного навчання. Оптимізаційні завдання (знаходження пайплайна довільної структури) вирішується зазвичай за допомогою еволюційних алгоритмів, приклад фреймворки TPOT та FEDOT. У нашій роботі вибрано для дослідження фреймворк FEDOT.

Заслуговує на увагу поняття про конвеєри для проблем машинного навчання Конвеєр – це ациклічний орієнтований граф. У термінах FEDOT цей граф називається ланцюгом, або складеною моделлю, або конвеєром (рис. 2.3). Часовий ряд – це послідовність значень, де наступні значення зазвичай залежать від попередніх. Отже, ми можемо використовувати поточні та попередні елементи часового ряду, щоб зробити прогноз. Якщо потрібно передбачити ряд на один елемент вперед, використовуючи поточне та одне попереднє значення. Існує кілька конкретних моделей для прогнозування часових рядів у FEDOT (таких як AR та ARIMA).

Набір даних для узгодження обсягів резервного виробництва електроенергії автономними джерелами представляється багатовимірний часовий ряд, що відображає середньодобове вироблення електроенергії вітряною установкою та середньодобове вироблення електроенергії за допомогою дизельного генератора. Обидва показники вимірюються в кВт·год. Нами визначено статистичні характеристики даних про середньодобове вироблення електроенергії дизельними генераторами та вітряними електростанціями. Середнє значення вироблення електроенергії дизельними

генераторами становить 2452,71 кВт·год, а вітряними електростанціями – 316,75 кВт·год. Мінімальне значення вироблення електроенергії дизельними генераторами становить 0,85 кВт·год, а вітряними електростанціями 1,08 кВт·год.

Для проведення аналізу часових рядів зміни обсягів резервного виробництва електроенергії застосовується бібліотека Fedot, спрямована на автоматизоване машинне навчання.

У нас є завдання підготувати модель для прогнозування обсягів виробництва електроенергії вітроенергетичною установкою на 7 днів вперед (для часового ряду `diesel_fuel_kWh` (червона крива)). Для цього слід позбутися пропусків. У FEDOT використовуються дві групи методів для відновлення значень у часових рядах:

Нами виконано прогнозування обсягів виробництва електроенергії генератором та вітроенергетичними установками. Для цього виконували дослідження та порівняли 4 сценарії використання Fedot: 1) із пайплайномі із двох гілок (`branches`); 2) із налаштуванням гіперпараметрів у вузлах конвеєра; 3) із використанням підходу API; 4) із використанням композитора.

Найкращою моделлю для прогнозу обсягів виробництва електроенергії генераторами є та, яка використовує налаштування гіперпараметрів у вузлах конвеєра. Ця модель досягає найменших значень Root Mean Squared Error (RMSE) та Mean Absolute Error (MAE), відповідно 120.979 та 95.999. Крім того, величини Mean Absolute Percentage Error (MAPE) та Symmetric Mean Absolute Percentage Error (SMAPE) становлять 43.440% та 30.962%, відображаючи високий рівень точності прогнозу.

Для прогнозу обсягів виробництва електроенергії вітроустановками, також модель із налаштуванням гіперпараметрів у вузлах конвеєра виявляється найкращою. Вона досягає найменших значень RMSE та MAE – 540.149 та 311.008 відповідно, а також відзначається MAPE та SMAPE на рівні 90.230% та 63.008%.

Обрані показники свідчать про високу точність та ефективність обраних моделей для обох типів прогнозів.

Нами розроблено заходи, які забезпечать організувати розробку моделей машинного навчання із дотриманням всіх вимог охорони праці та безпеки у надзвичайних ситуаціях.

Ефективність від використання моделі машинного навчання для узгодження обсягів резервного виробництва електроенергії автономними джерелами визначається як співвідношення між вигодами, отриманими від використання моделі, та витратами на її розробку та впровадження. На підставі проведених розрахунків встановлено, що ефективність становить 25%. Це означає, що кожні 1 грн, витрачені на розробку та впровадження моделі, приносять 0,25 грн вигод.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Боярчук В.М., Тригуба А.М., Луб П.М., Фтома О.В., Лут М.Т., Батечко Н.Г., Волошин С.М. Енергетичний менеджмент і аудит в агропромисловому комплексі. Підручник для студентів вищих навчальних закладів. К. ЦП «Компринт», 2015. 630с.
2. Введення в машинне навчання за допомогою Python и Scikit-Learn. URL: <https://habr.com/ua/company/mlclass/blog/247751/> (дата звернення: 20.05.2022).
3. Жидецький В.Ц., Джигирей В.С., Мельников О.В. Основи охорони праці. Підручник. Вид. 5-е, доповнене. Львів: Афіша, 2012. 350с.
4. Карпа Д. М., Цмоць І. Г., Опотьяк Ю. В. Нейромережеві засоби прогнозування споживання енергоресурсів. Науковий вісник НЛТУ України. 2018, т. 28, № 5. С. 140–146.
5. Класифікація в Python з Scikit-Learn та Pandas. URL: <https://stackabuse.com/classification-in-python-with-scikit-learn-and-pandas/> (дата звернення: 17.05.2022).
6. Лехман С.Д., Рублев В.І., Рябцев Б.І. Запобігання аварійності і травматизму у сільському господарстві. К.: Урожай, 1993. 267 с.
7. Лутц М. Программирование на Python. I том. СПб.: Символ-плюс, 2015. 992 с.
8. Марк Саммерфилд. Программирование на Python 3. Подробное руководство. Пер. с англ. СПб.: Символ-Плюс, 2013. 608 с.
9. Навчання нейромережі з учителем, без вчителя, з підкріпленням – у чому відмінність? URL: <https://neurohive.io/ru/osnovy-data-science/obuchenie-s-uchitelem-bez-uchitelja-s-podkrepleniem/>(дата звернення: 22.10.2022).
10. Огляд методів класифікації у машинному навчанні за допомогою Scikit-Learn. URL: <https://tproger.ru/translations/scikit-learn-in-python/s://stackabuse.com/classification-in-python-with-scikit-learn-and-pandas/> (дата звернення: 10.05.2022).

11. Плескач В.Л., Рогушина Ю.В., Кустова Н.П. Інформаційні технології та системи. К.: Книга, 2004. 519 с.
12. Прохоренок Н.А. Python 3 и PyQt. Разработка приложений. СПб.: БХВ-Петербург, 2012. 704 с.
13. Tryhuba A., Ivanyshyn V., Chaban V., Mushenyk I., Zharikova O. Computer model of resource demand planning for dairy farms. *Independent Journal of Management & Production (Special Edition ISE, S&P)*. 2021. 12(3), pp. 138-149. URL: <http://www.ijmp.jor.br/index.php/ijmp/article/view/1531/1971> (Last accessed: 21.10.2023).
14. Tryhuba, A., Boyarchuk, V., Tryhuba, I., Ftoma, O., Padyuka, R., Rudynets, M. Forecasting the Risk of the Resource Demand for Dairy Farms Basing on Machine Learning. *Proceedings of the 2nd International Workshop on Modern Machine Learning Technologies and Data Science (MoMLeT+DS 2020)*. 2020. I. P. 327-340.
15. Цмоць І. Г., Роман В. І. Вдосконалення методу групування енергетичних даних у системі багаторівневого управління енергоефективністю економіки регіону. *Науковий вісник НЛТУ України*. 2019, т. 29, № 1. С. 116–120.
16. Boyarchuk V., Ivanyshyn V., Tryhuba A., Zasada M., Hutsol T., Tatomyr A., Tryhuba I., Nurek T., Glowacki Sz., Brys A. Substantiation of the configuration of agricultural power supply systems using wind energy based on computer simulation. *Monograph*. – Warszawa: 2020 – 126 p.
17. Koval N., Tryhuba A., Kondysiuk I., Tryhuba I., Boiarchuk O., Rudynets M., Grabovets V., Onyshchuk V., Forecasting the Fund of Time for Performance of Works in Hybrid Projects Using Machine Training Technologies. *Proceedings of the 3rd International Workshop on Modern Machine Learning Technologies and Data Science Workshop. Proc. 3rd International Workshop (MoMLeT&DS 2021)*. Volume I: Main Conference. Lviv-Shatsk, Ukraine, June 5-6, 2021. pp.196-206.

18. Ratushny R., Tryhuba A., Bashynsky O., Ptashnyk V. Development and usage of a computer model of evaluating the scenarios of projects for the creation of fire fighting systems of rural communities. XI-th International Scientific and Practical Conference on Electronics (ELIT-2019). 2019. P. 34-39. URL: <https://ieeexplore.ieee.org/document/8892320> (Last accessed: 16.10.2023).
19. Syrotyuk V., Syrotyuk S., Ptashnyk V., Tryhuba A., Baranovych S., Gielzecki J. A hybrid system with intelligent control for the processes of resource and energy supply of a greenhouse complex with application of energy renewable sources. *Przegląd elektrotechniczny* 96 (7), 149-152.
20. Hutsol T., Glowacki S., Tryhuba A. Current Trends of Biohydrogen Production from Biomass – Green Hydrogen. Monograph. Warsaw: 2021. 102 p.
21. Tryhuba A., Bashynsky O., Garasymchuk I., Gorbovy O., Vilchinska D. Research of the variable natural potential of the wind and energy energy in the northern strip of the Ukrainian carpathians. *E3S Web of Conferences* 154, 06002
22. Tryhuba A., Ratushny R., Bashynsky O., Shcherbachenko O. Identification of firefighting system configuration of rural settlements. *Fire and Environmental Safety Engineering. MATEC Web Conf. FESE 2018.* 247. doi: <https://doi.org/10.1051/matecconf/201824700035>
23. Tryhuba A., Zachko O., Grabovets V., Berladyn O., Pavlova I., Rudynets M. Examining the effect of production conditions at territorial logistic systems of milk harvesting on the parameters of a fleet of specialized road tanks. *Eastern-European Journal of Enterprise Technologies.* 2018. 5(3). P. 59-70. URL: [http://nbuv.gov.ua/UJRN/Vejpte_2018_5\(3\)__7](http://nbuv.gov.ua/UJRN/Vejpte_2018_5(3)__7). (Last accessed: 17.10.2023).
24. Tryhuba, A., Kondysiuk, I., Tryhuba, I., Boiarchuk, O., Tatomyr, A., Intellectual information system for formation of portfolio projects of motor transport enterprises. *CEUR Workshop Proceedings, 2022, 3109*, pp. 44–52.
25. Tryhuba, A., Malanchuk, O., Tryhuba, I. Prediction of the Duration of Inpatient Treatment of Diabetes in Children Based on Neural Networks. *CEUR Workshop Proceedings, 2023, 3426*, pp. 122–135.

26. Sarkodie S.A., Adams S. Electricity access, human development index, governance and income inequality in Sub-Saharan Africa. *Energy Rep*, 6 (2020), pp. 455-466.
27. Hao P., Yin S., Wang D., Wang J. Exploring the influencing factors of urban residential electricity consumption in China. *Energy Sustain Dev*, 72 (2023), pp. 278-289.
28. Wabukala B.M., Bergland O., Rudaheranwa N., Watundu S., Adaramola M.S., Ngoma M., Rwaheru A.A. Unbundling barriers to electricity security in Uganda: A review. *Energy Strategy Rev*, 44 (2022), Article 100984.
29. Dao V., Ishii H., Takenobu Y., Yoshizawa S., Hayashi Y. Intensive quadratic programming approach for home energy management systems with power utility requirements. *Int J Electr Power Energy Syst*, 115 (2020), Article 105473
30. Fan G.-F., Zhang L.-Z., Yu M., Hong W.-C., Dong S.-Q. Applications of random forest in multivariable response surface for short-term load forecasting. *Int J Electr Power Energy Syst*, 139 (2022), Article 108073.
31. Stefenon S.F., Seman L.O., Mariani V.C., Coelho L.d.S. Aggregating prophet and seasonal trend decomposition for time series forecasting of Italian electricity spot prices. *Energies*, 16 (3) (2023), p. 1371.
32. Neto N.F., Stefenon S.F., Meyer L.H., Ovejero R.G., Leithardt V.R.Q. Fault prediction based on leakage current in contaminated insulators using enhanced time series forecasting models. *Sensors*, 22 (16) (2022), p. 6121.
33. Branco N.W., Cavalca M.S.M., Stefenon S.F., Leithardt V.R.Q. Wavelet LSTM for fault forecasting in electrical power grids. *Sensors*, 22 (21) (2022), p. 8323.
34. Stefenon S.F., Corso M.P., Nied A., Perez F.L., Yow K.-C., Gonzalez G.V., Leithardt V.R.Q. Classification of insulators using neural network based on computer vision. *IET Gener, Transm Distrib*, 16 (6) (2021), pp. 1096-1107.
35. Ali M., Adnan M., Tariq M. Optimum control strategies for short term load forecasting in smart grids. *Int J Electr Power Energy Syst*, 113 (2019), pp. 792-806.

36. Stefenon S.F., Singh G., Yow K.-C., Cimatti A. Semi-ProtoPNet deep neural network for the classification of defective power grid distribution structures. *Sensors*, 22 (13) (2022), p. 4859.
37. Silva R.G., Moreno S.R., Ribeiro M.H.D.M., Larcher J.H.K., Mariani V. C., dos Santos Coelho L. Multi-step short-term wind speed forecasting based on multi-stage decomposition coupled with stacking-ensemble learning approach. *Int J Electr Power Energy Syst*, 143 (2022), Article 108504.
38. Wang Y., Sun S., Chen X., Zeng X., Kong Y., Chen J., Guo Y., Wang T. Short-term load forecasting of industrial customers based on SVM and XGBoost. *Int J Electr Power Energy Syst*, 129 (2021), Article 106830.
39. Zhang Q., Zhang J. Short-term load forecasting method based on EWT and IDBSCAN. *J Electr Eng Technol*, 15 (2) (2020), pp. 635-644.
40. Sankalpa C., Kittipiyakul S., Laitrakun S. Forecasting short-term electricity load using valiyeard ensemble learning. *Energies*, 15 (22) (2022), p. 8567
41. Duan Y. A novel interval energy-forecasting method for sustainable building management based on deep learning. *Sustainability*, 14 (14) (2022), p. 8584.
42. Japan Electric Power eXchange (JEPX) [accessed on November 16, 2017]. <http://www.jepx.org/english>