

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: **«Інтервальне прогнозування індивідуального
споживання електроенергії домогосподарствами із
використанням рекурентних нейронних мереж»**

Виконав: студент групи Іт-61

Спеціальності 126 «Інформаційні системи та
технології»

(шифр і назва)

Малетич Іван Ігорович

(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

Малетичу Івану Ігоровичу

1. Тема роботи: «Інтервальне прогнозування індивідуального споживання електроенергії домогосподарствами із використанням рекурентних нейронних мереж»

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 для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами; методика дослідження рекурентних нейронних мереж.

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

Вступ.

1. Аналіз стану предметної області та завдання кваліфікаційної роботи.

2. Вибір засобів та використанням рекурентних нейронних мереж для прогнозування індивідуального споживання електроенергії домогосподарствами.

3. Результати обґрунтування моделей із рекурентними нейронними мережами для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами.

4. Охорона праці та безпека у надзвичайних ситуаціях.

5. Визначення ефективності запропонованої інтелектуальної інформаційної системи.

Висновки та пропозиції.

Список використаної літератури.

5. Перелік ілюстраційного матеріалу (з точним зазначенням обов'язкових слайдів): аналіз стану предметної області та завдання кваліфікаційної роботи; результати вибору засобів та використання рекурентних нейронних мереж для прогнозування індивідуального споживання електроенергії домогосподарствами; результати обґрунтування архітектури моделей із рекурентними нейронними мережами для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами; результати визначення точності моделей; 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>	
7.	<i>Завершення роботи в цілому</i>	<i>01-10.01.24</i>	

Студент _____ Малетич І.І.
(підпис)

Керівник роботи _____ Татомир А.В.
(підпис)

УДК 004.83.004 : 621.311.32

Інтервальне прогнозування індивідуального споживання електроенергії домогосподарствами із використанням рекурентних нейронних мереж.

Малетич І.І. Кафедра інформаційних технологій – Дубляни, ЛНУП, 2024.

Кваліфікаційна робота: 76 с. текст. част., 14 рис., 3 табл., 10 арк. ілюстраційного матеріалу, 66 джерел.

Виконано аналіз підходів до прогнозування індивідуального споживання електроенергії домогосподарствами, способів збору та обробки даних для обґрунтування моделей прогнозування споживання електроенергії індивідуальними споживачами та структури моделі прогнозування споживання електроенергії домогосподарствами. Здійснено вибір засобів та використанням рекурентних нейронних мереж для прогнозування індивідуального споживання електроенергії домогосподарствами. Подано особливості прогнозування індивідуального споживання електроенергії домогосподарствами. Описано методи LSTM із одновимірними вхідними даними і вихідним вектором та LSTM «кодувальник-декодувальник» із одновимірним вхідними даними. Здійснено вибір засобів для створення моделей.

Здійснено підготовку даних для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами. Проведено аналіз даних для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами. Розроблено моделі LSTM із одновимірними вхідними даними та результати прогнозування індивідуального споживання електроенергії та LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними та результати прогнозування індивідуального споживання електроенергії. Здійснено порівняння LSTM моделей для прогнозування індивідуального споживання електроенергії.

Розроблено заходи із охорони праці та розраховано ефективність від використання моделі LSTM.

ЗМІСТ

ВСТУП	7
РОЗДІЛ 1. АНАЛІЗ СТАНУ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ЗАВДАННЯ КВАЛІФІКАЦІЙНОЇ РОБОТИ.....	9
1.1. Аналіз підходів до прогнозування індивідуального споживання електроенергії домогосподарствами	9
1.2. Аналіз способів збору та обробки даних для обґрунтування моделей прогнозування споживання електроенергії індивідуальними споживачами.....	13
1.3. Структура моделі прогнозування споживання електроенергії домогосподарствами	19
1.4. Аналіз схем визначення прогнозованих значень енергоспоживання індивідуальними споживачами.....	22
1.5. Обґрунтування доцільності інтервального прогнозування індивідуального споживання електроенергії домогосподарствами з використанням рекурентних нейронних мереж.....	23
РОЗДІЛ 2. ВИБІР ЗАСОБІВ ТА ВИКОРИСТАННЯМ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ ІНДИВІДУАЛЬНОГО СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ ДОМОГОСПОДАРСТВАМИ	25
2.1. Особливості прогнозування індивідуального споживання електроенергії домогосподарствами	25
2.2. LSTM із одновимірними вхідними даними і вихідним вектором.....	28
2.3. LSTM «кодувальник-декодувальник» із одновимірним вхідними даними .	30
2.4. Вибір засобів для створення моделей	31
РОЗДІЛ 3. РЕЗУЛЬТАТИ ОБґРУНТУВАННЯ МОДЕЛЕЙ ІЗ РЕКУРЕНТНИМИ НЕЙРОННИМИ МЕРЕЖАМИ ДЛЯ ІНТЕРВАЛЬНОГО ПРОГНОЗУВАННЯ ІНДИВІДУАЛЬНОГО СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ ДОМОГОСПОДАРСТВАМИ	34
3.1. Підготовка даних для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами	34

3.2. Аналіз даних для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами	37
3.3. Розробка моделі LSTM із одновимірними вхідними даними та результати прогнозування індивідуального споживання електроенергії	44
3.4. Розробка моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними та результати прогнозування індивідуального споживання електроенергії	47
3.5. Порівняння LSTM моделей для прогнозування індивідуального споживання електроенергії	52
РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	54
4.1. Небезпечні та шкідливі виробничі чинники під час розробки нейромереж	54
4.2. Заходи щодо запобігання небезпечним та шкідливим виробничим чинникам	55
4.3. Розробка логічно-імітаційної моделі процесу виникнення травм під час розробки нейромереж.....	56
4.4. Розробка заходів із безпеки під час надзвичайних ситуацій	60
РОЗДІЛ 5. ВИЗНАЧЕННЯ ЕКОНОМІЧНОЇ ЕФЕКТИВНОСТІ ВІД ВИКОРИСТАННЯ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ ІНДИВІДУАЛЬНОГО СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ ДОМОГОСПОДАРСТВАМИ	61
ВИСНОВКИ І ПРОПОЗИЦІЇ.....	64
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	69

ВСТУП

У сучасному світі інформаційні технології та зростаючі вимоги до енергоефективності, виробництва та споживання електроенергії набувають все більшого значення. Особливо актуальним стає завдання точного та передбачливого прогнозування індивідуального споживання електроенергії домогосподарствами, щоб забезпечити стабільну роботу енергосистеми та зменшити витрати на виробництво та постачання електроенергії [1].

У цьому контексті використання рекурентних нейронних мереж виявляється потужним інструментом для розв'язання задачі інтервального прогнозування споживання електроенергії [28]. Вони можуть аналізувати складні та динамічні відносини між специфічними факторами, такими як погода, часові зміни, а також індивідуальні споживацькі звички, що забезпечує отримання інтервальних прогнозів, які враховують особливості споживання електроенергії і забезпечують точність.

Ця кваліфікаційна робота спрямована на дослідження та розробку методів та моделей інтервального прогнозування індивідуального споживання електроенергії домогосподарствами з використанням рекурентних нейронних мереж. Вона відкриває можливість покращення управління енергосистемами, зменшення витрат та оптимізації виробництва електроенергії, сприяючи сталому розвитку та підвищенню якості життя населення.

У цьому контексті подальше вивчення інтервального прогнозування індивідуального споживання електроенергії з використанням рекурентних нейронних мереж має великий науковий та практичний інтерес, і ця робота спрямована на внесення важливого внеску у зазначену область досліджень.

Отже, розробка моделей інтервального прогнозування індивідуального споживання електроенергії домогосподарствами із використанням рекурентних нейронних мереж має практичну цінність та забезпечує покращення управління енергосистемами та зменшення витрат на виробництво та постачання електроенергії.

Об'єктом дослідження є рекурентні нейронні мережеві моделі інтервального прогнозування індивідуального споживання електроенергії домогосподарствами.

Предмет дослідження є вплив характеристик індивідуальних домогосподарств на архітектуру рекурентних нейронних мережевих моделей інтервального прогнозування індивідуального споживання електроенергії.

Метою даної кваліфікаційної роботи є підвищення ефективності прогнозування індивідуального споживання електроенергії домогосподарствами завдяки розробці моделі із використанням рекурентних нейронних мереж (RNN).

РОЗДІЛ 1.

АНАЛІЗ СТАНУ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ЗАВДАННЯ КВАЛІФІКАЦІЙНОЇ РОБОТИ

1.1. Аналіз підходів до прогнозування індивідуального споживання електроенергії домогосподарствами

Традиційно прогнозування споживання електроенергії в житловому секторі було зосереджено на прогнозуванні сукупного погодинного споживання електроенергії сотнями чи тисячами домогосподарств. Прогнози використовувалися як фундаментальний інструмент для належного управління існуючою централізованою енергосистемою, оскільки здатність передбачати періоди високого попиту дозволяє вмикати додаткові електростанції завчасно, щоб уникнути масового дефіциту електроенергії [35]. Ці агреговані методи прогнозування також широко використовувалися для підтримки попиту на електроенергію на оптових ринках електроенергії.

Протягом останніх десятиліть усвідомлення необхідності значного скорочення викидів парникових газів зросло, і відбувається масовий перехід до розподілених відновлюваних джерел енергії (ВДЕ). У цій новій парадигмі розподіленої системи електроенергії великі електростанції замінюються багатьма меншими.

Побутові клієнти залишають пасивну роль «тільки платників» і починають відігравати центральну роль, роблячи внески як генератори та розумно регулюючи періоди, які вони споживають. Таким чином, щоб гарантувати належну синхронізацію цієї більш складної електричної мережі, що складається з цих нових генераторів із великими коливаннями, та оптимізувати енергоефективність усієї системи, існує необхідність точного прогнозування споживання електроенергії в житловому секторі на індивідуальному або дезагрегованому рівні. Це вимагає більш точних моделей

прогнозування, здатних передбачити як виробництво ВДЕ, так і споживання електроенергії на індивідуальному рівні або на рівні громади [21].

При прогнозуванні споживання електроенергії житловим сектором важливим фактором впливу є поведінка споживачів. Це пояснюється тим, що індивідуальне споживання електроенергії значною мірою залежить від процесу прийняття рішень людьми, який має притаманний стохастичний характер [33]. При прогнозуванні споживання електроенергії на агрегованому рівні мінливість кожного споживача збалансована, а ефект поведінки користувача послаблюється. Однак при розробці індивідуальних або дезагатованих моделей компонент поведінки кожного клієнта набуває значення, що призводить до значної мінливості в моделях споживання енергії клієнтами. Отже, на агрегованому рівні ослаблення ефекту поведінки користувача спричиняє значно менші помилки прогнозування, ніж на індивідуальному рівні.

Ще одна проблема, пов'язана з прогнозуванням індивідуального споживання електроенергії, пов'язана зі збором даних і доступом до них (рис. 1.1).

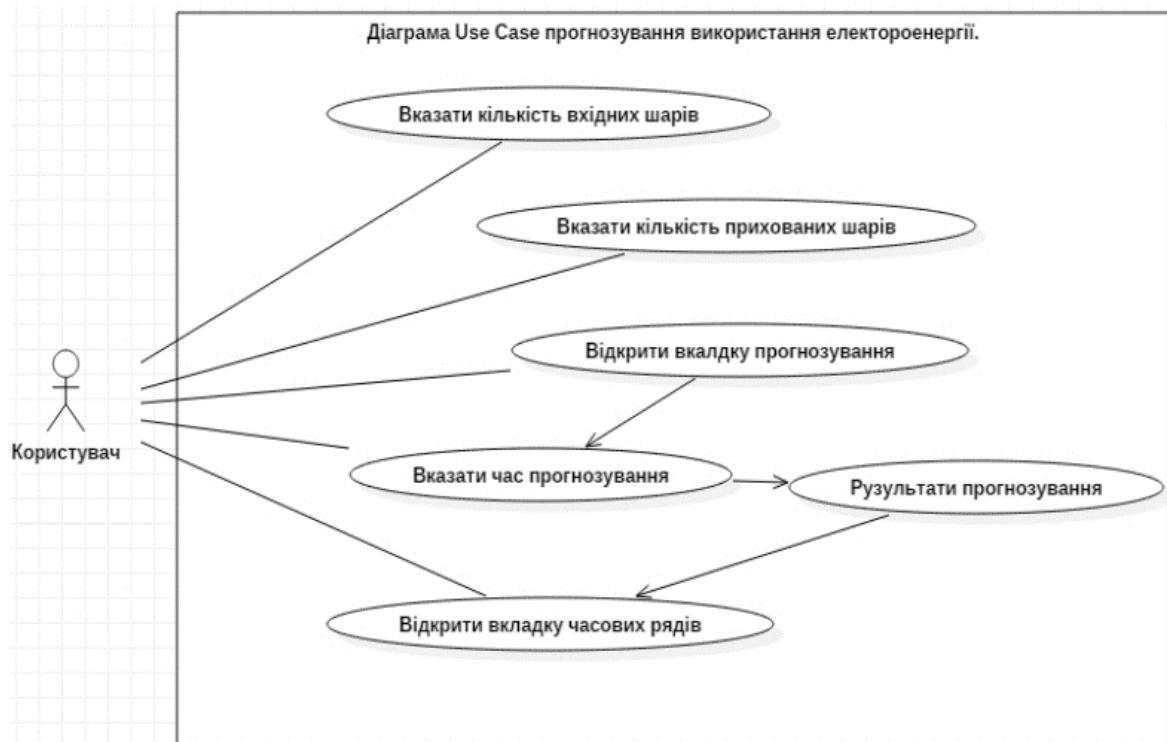


Рисунок 1.1 – Діаграма прецедентів прогнозуванням індивідуального споживання електроенергії

У багатьох дослідницьких роботах для прогнозування індивідуального споживання електроенергії використовувалися субліційні дані споживання приладів або результати опитувань споживачів. Однак такі дані важко отримати, працюючи в реальних бізнес-сценаріях із великою кількістю житлових домогосподарств [37]. Наприклад, збір щоденних графіків використання, заповнених користувачами, абсолютно неможливий, коли кількість клієнтів значна. Крім того, отримати доступ до цього типу персональних даних навряд чи можливо на практиці через проблеми конфіденційності даних. Коли дані надходять від побутових приладів або домашніх енергетичних систем, хоча їх проникнення на ринок швидко зростає, розгортання та встановлення апаратних компонентів все ще є складним завданням під час масового впровадження. Це головні причини, чому виникає потреба використання даних інтелектуальних лічильників, доповнених даними про місцеву погоду, які можна отримати через двосторонні угоди з компаніями-трейдерами електроенергії та веб-службами прогнозу погоди.

Аналіз даних, що надходять із інтелектуальних лічильників, вимагає надійних методів обробки відсутніх даних і викидів, оскільки вони часто присутні в наборах даних. У окремих дослідницьких роботах це питання детально не розглядалося, оскільки більшість із них не розглядали реальні дані, що надходять із системи електроенергії. Однією потреб є внесок у заповнення цієї прогалини шляхом розробки інноваційної комбінації методів кластеризації, процедур класифікації та моделей прогнозування на основі машинного навчання, які здатні досягти плавного управління цими невідповідностями даних.

Нещодавно розроблений підхід поєднує прогнозування моделі поведінки з моделлю прогнозування споживання електроенергії для окремих користувачів. Мета полягає в тому, щоб спрогнозувати споживання електроенергії на годину на один день вперед для великої кількості житлових домогосподарств (без агрегування), враховуючи значний вплив стохастичної частини поведінки споживачів. Він починається з кластеризації для визначення

кластерів поведінки, після чого реалізується модель класифікації XGBoost (рис. 1.2) для прогнозування кластера поведінки на добу наперед.

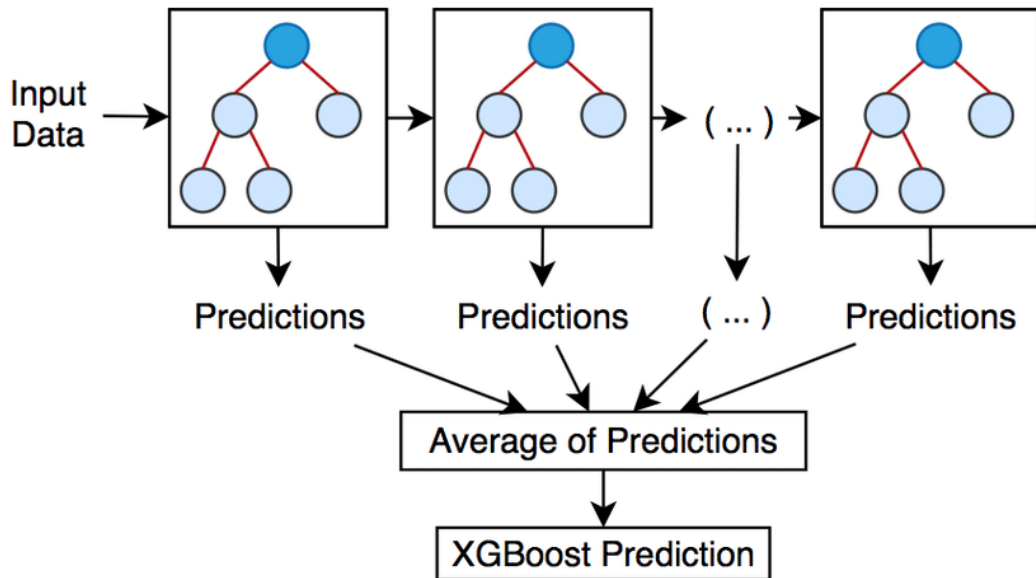


Рисунок 1.2 – Архітектура моделі класифікації XGBoost

Цей передбачуваний шаблон поведінки користувача стає основним входом для індивідуальної моделі прогнозування споживання електроенергії, заснованої на методах штучної нейронної мережі (АНМ), і дозволяє більш точно фіксувати стохастичний компонент поведінки.

Використання моделі поведінки, заснованої на методах класифікації часових рядів, зазвичай ускладнює інтерпретацію її точності. Звичайні двійкові метрики для оцінки точності часових рядів не дають повного розуміння при роботі з моделями щоденного споживання, що є основним елементом класифікації в нашому випадку.

Комбінований підхід, який підтримується новою розробленою метрикою, спрямований на створення основи для масштабного аналізу, орієнтованого на користувача, у секторі електроенергетики. Очікується, що такі збагачені методи, керовані даними, будуть дуже прибутковими для операторів електромереж, агрегаторів реагування на попит або компаній, що торгують електроенергією, у їхніх портфоліо послуг для кінцевих користувачів.

1.2. Аналіз способів збору та обробки даних для обґрунтування моделей прогнозування споживання електроенергії індивідуальними споживачами

Для аналізу енергетичної поведінки побутових приладів розроблено різні імовірнісні моделі на основі даних. У роботі [65] використовували дані про споживання офісної техніки в офісній будівлі, щоб отримати уявлення про поведінку мешканців. Наукова праця [52] передбачає стохастичне моделювання споживання енергії малою побутовою технікою та класифікував її на чотири групи: аудіовізуальна, комп'ютерна, кухонна та інша техніка.

У роботі [64] представлено метод ансамблю для прогнозування погодинного споживання електроенергії (рис. 1.3), використовуючи дані з ручних пристроїв (таких як посудомийні машини, теплові насоси, телевізори).

У роботі [39] розглянути тисячі профілів споживання електроенергії були створені з комбінацій моментів активації від різних побутових приладів.

Феррандес-Пастор та інші у роботі [30] представили дизайн вбудованих апаратних компонентів, здатних дезагрегувати споживання електроенергії домогосподарствами. Чоу та Тран [28] використовували дані, зібрані з обладнання, встановленого в експериментальній будівлі, щоб оцінити ефективність різних методів машинного навчання. Водночас Хак та ін. [34] використовували дані про споживання електроприладів (холодильник, духовка, фен, освітлення, кондиціонер, ноутбук, водонагрівач, телевізор, праска, сушарка для білизни) для прогнозування споживання енергії.

Інші дослідження використовували результати опитувань мешканців, щоб посилити свої моделі. Aerts та ін. [17] розробили імовірнісну модель для визначення типових моделей зайнятості на основі опитування про час використання.

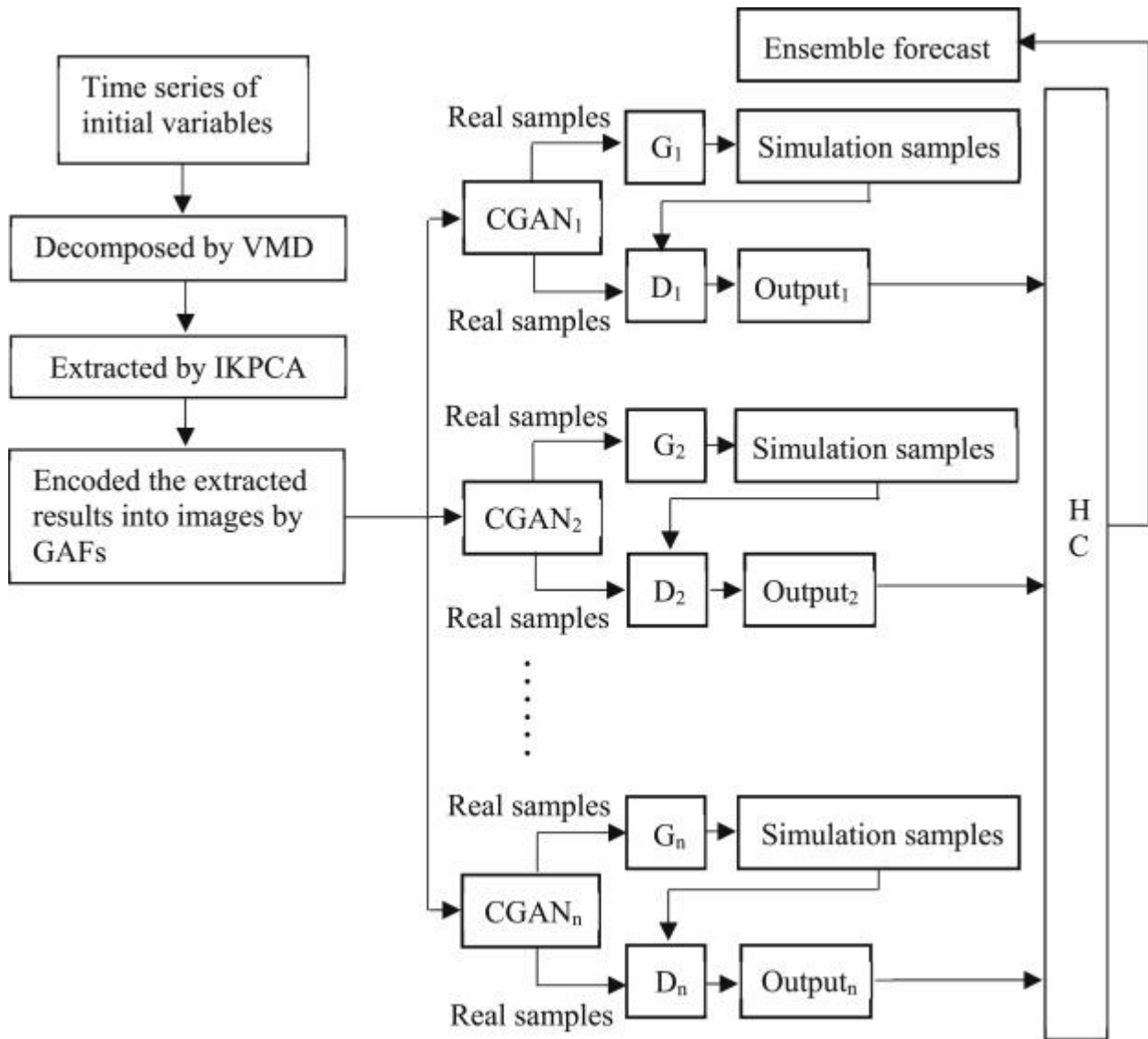


Рисунок 1.3 – Принципи для ансамблю CGAN [64]

У роботі [41] Liisberg та ін. використовували приховані моделі Маркова в поєднанні з опитуваннями, щоб охарактеризувати поведінку мешканців у житловому будинку (рис. 1.4).

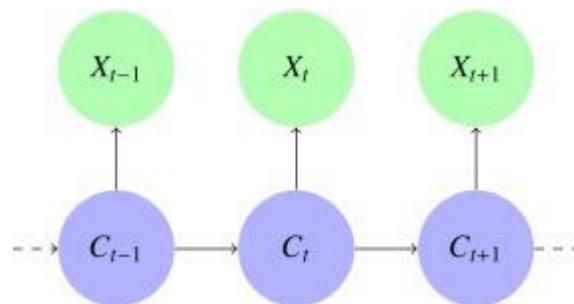


Рисунок 1.4 – Орієнтований граф прихованої Марковської моделі [41]

Паласіос-Гарсія та ін. [48] представили висхідну стохастичну модель попиту на електроенергію для приладів опалення та охолодження (рис. 1.5). Вони заявили, що профіль заповнюваності є наріжним каменем моделі, підкресливши, що споживання енергії в житловому секторі сильно пов'язане з діяльністю людей.

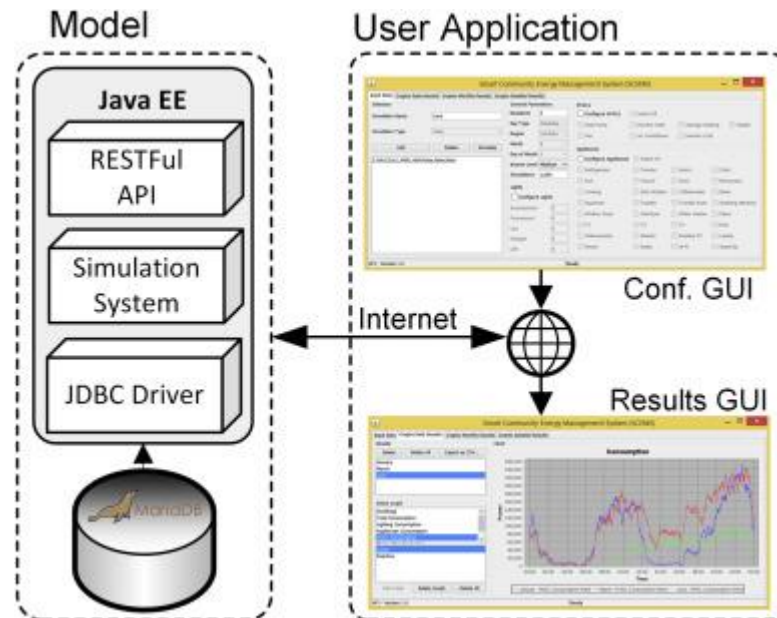


Рисунок 1.5 – Логічна реалізація системи моделювання та розроблений GUI [48]

Вони заповнювали свою модель інформацією, зібраною в щоденнику, який люди мали заповнювати з частотою 10 хвилин і включати щоденні дії.

Усі згадані автори використовували дані із систем підрахунків, здатних вимірювати споживання електроенергії побутовими приладами в житлових домогосподарствах. Деякі з них також використовували результати опитувань клієнтів, щоб покращити свої моделі прогнозування. Однак такі дані важко отримати під час роботи в реальних сценаріях, коли дані з підсистем вимірювання недоступні. У цьому випадку фіскальні смарт-лічильники стають цінним джерелом даних, тут є лише один показник загального споживання електроенергії домогосподарством.

Багато дослідників зосередили свою діяльність на отриманні моделей споживання електроенергії та виконанні сегментації споживачів,

використовуючи ці дані в алгоритмах кластеризації [19; 20; 42]. Вилучення типових моделей споживання допомогло операторам системи розподілу електроенергії (DSO) або роздрібним продавцям електроенергії розробити інноваційні структури тарифів, спрямовані на ефективне профілювання споживання, як зазначено у роботах [29; 27; 45]. У роботі [18] було виконано комплексний огляд доступних методів кластеризації часових рядів та їх застосування.

Окрім характеристики енергетичної поведінки, протягом останніх років також інтенсивно аналізувалося прогнозування споживання електроенергії побутовими споживачами на добу наперед. Детальні огляди [44; 62; 63] показали, що коли дані надходять виключно з інтелектуальних фіскальних лічильників, найпоширенішим підходом є імовірнісні моделі, керовані даними, і що ШНМ стали переважати серед цих методів.

Чоу та Тран [28] провели комплексне дослідження використання різних моделей прогнозування споживання на основі даних на основі часових рядів споживання енергії житловими домогосподарствами (рис. 1.6), знову дійшовши висновку, що найефективнішою моделлю на основі даних є ШНМ.

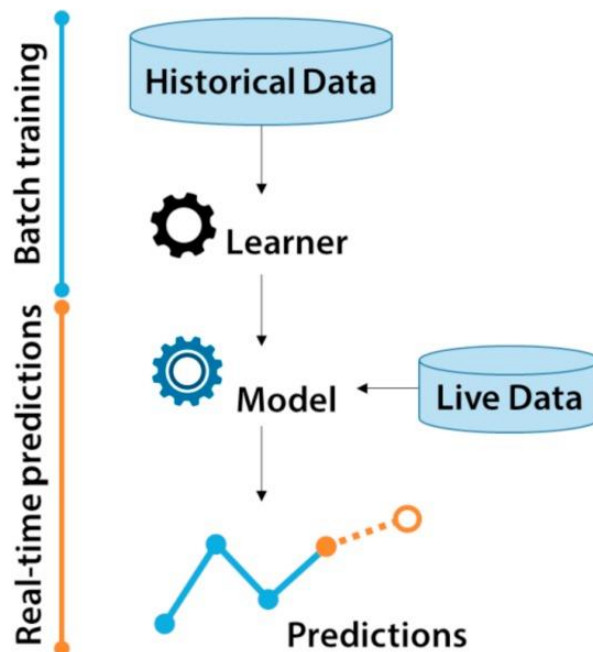


Рисунок 1.6 – Структура єдиних моделей прогнозування [28]

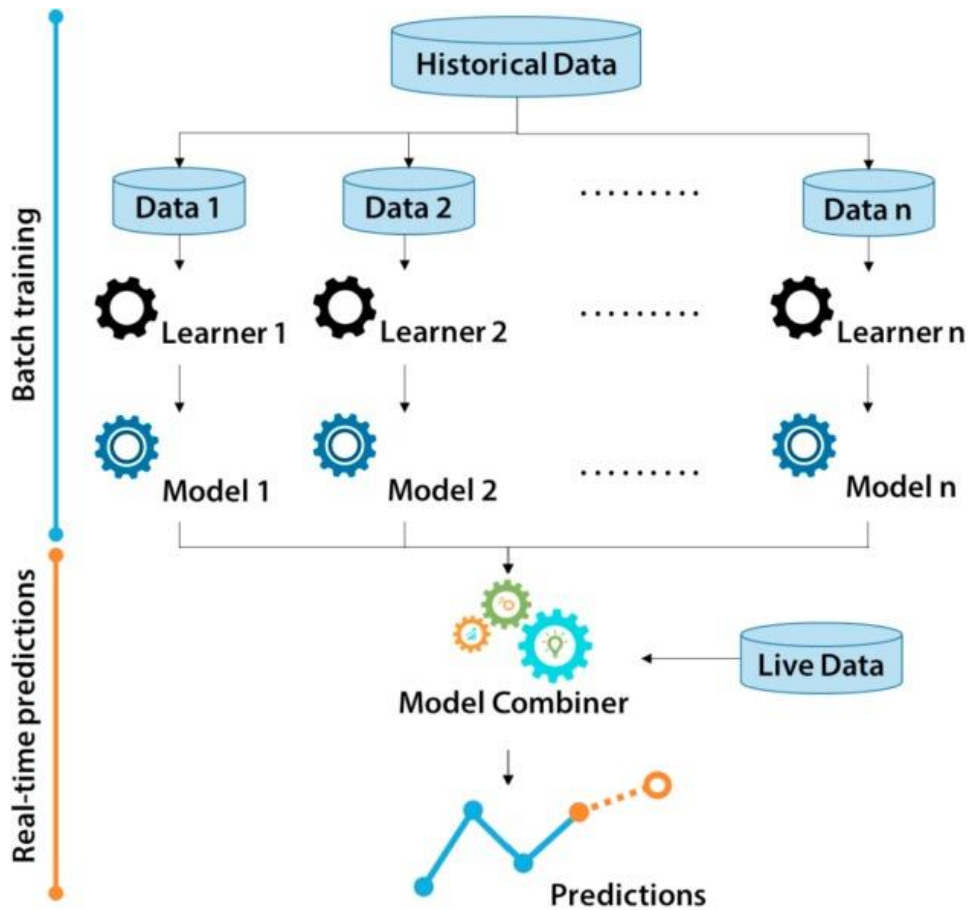


Рисунок 1.7 – Каркас моделей ансамблю [28]

Причиною високої продуктивності ШНМ, які використовуються для прогнозування споживання електроенергії, є їх здатність фіксувати нелінійні зв'язки між вхідними даними та цільовими параметрами [40; 46].

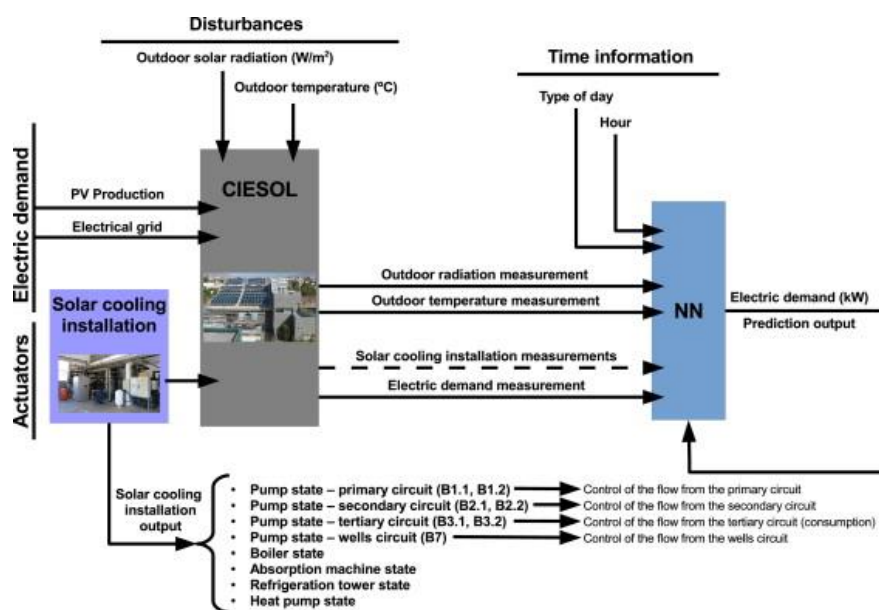


Рисунок 1.8 – Прийнятий підхід на основі ШНМ [46]

Тим не менш, усі зазначені моделі, що керуються даними, включаючи ШНМ, мають певні обмеження в охопленні нестабільних, але дуже періодичних енергетичних звичок побутових споживачів, коли єдиними доступними даними є споживання електроенергії, календарна інформація та дані про зовнішню погоду. Ці набори даних містять інформацію про продуктивність користувачів протягом дня, залежності від клімату, а також тижневі, щоденні та погодинні залежності. Однак немає способу оцінки поведінки користувача, не пов'язаного з цими сезонними залежностями. Якщо клієнт вирішує вмикати пральну машину двічі на тиждень, коли повертається додому, ШНМ або модель на основі регресії не можуть зробити висновок про таку періодичну поведінку за допомогою цих вхідних даних (рис. 1.9). Необхідно включити принаймні одну нову екзогенну змінну, яка може мати прямий чи непрямий зв'язок із цією поведінкою користувача.

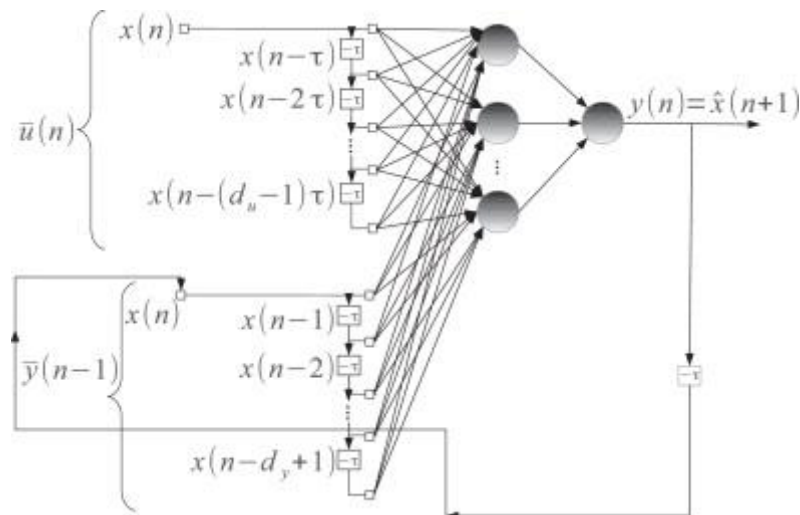


Рисунок 1.9 – Архітектура нелінійної авторегресії з екзогенними входами (NARX) [46]

У дослідницьких роботах [47; 32; 62] було виявлено, що хорошою процедурою для визначення цього нового входу є визначення щоденного профілю споживання електроенергії, який є найбільш ймовірно, відбудеться в 24-годинному горизонті прогнозування за допомогою комбінації методів кластеризації, щоденного вилучення ознак і методів машинного навчання. Ця процедура є значним покращенням у порівнянні з раніше згаданими

методологіями прогнозування споживання електроенергії. Зазначена стратегія побудована на основі того, що поведінка людини неодноразово підкоряється певним розпорядкам дня, про що має свідчити подібність у їхніх кривих споживання електроенергії.

1.3. Структура моделі прогнозування споживання електроенергії домогосподарствами

Щоб оцінити моделі прогнозування споживання електроенергії домогосподарствами слід порівняти точність результатів погодинного прогнозування споживання із і без включення нової запропонованої функції поведінки. Для цього проводиться аналіз, порівнюючи результати двох моделей.

Перша модель складається з ШНМ, навчена прогнозувати щогодинне споживання електроенергії, використовуючи 24-годинний період від попереднього дня. А друга модель складається з моделі споживання, до якої вводиться додатковий вхід – передбачення моделі поведінки.

Перша модель описана схемою на рис. 1.10, за винятком нижньої рамки, яка представляє модель поведінки (пов'язана пунктирною стрілкою). Друга пов'язана модель є повною діаграмою на рис. 1.10.

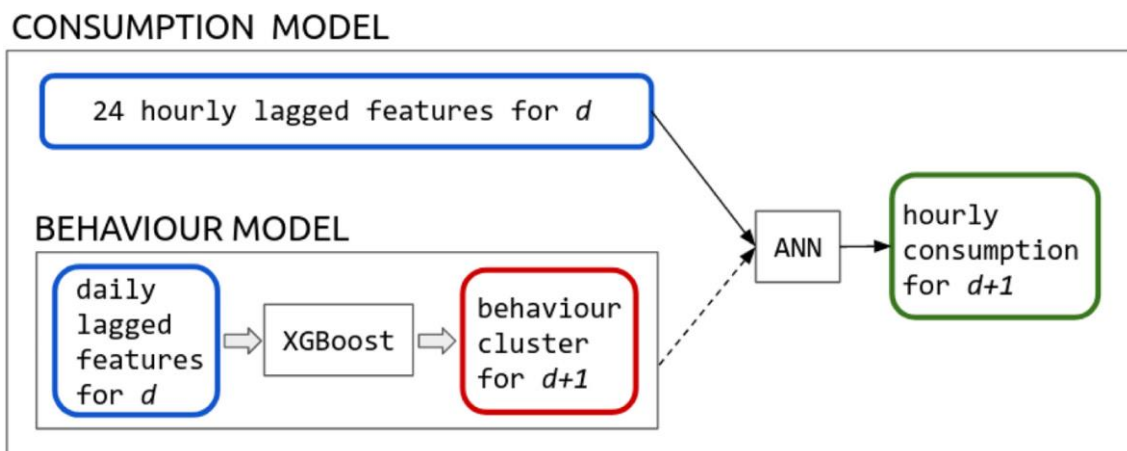


Рисунок 1.10 – Узагальнена структура моделі прогнозування [66]

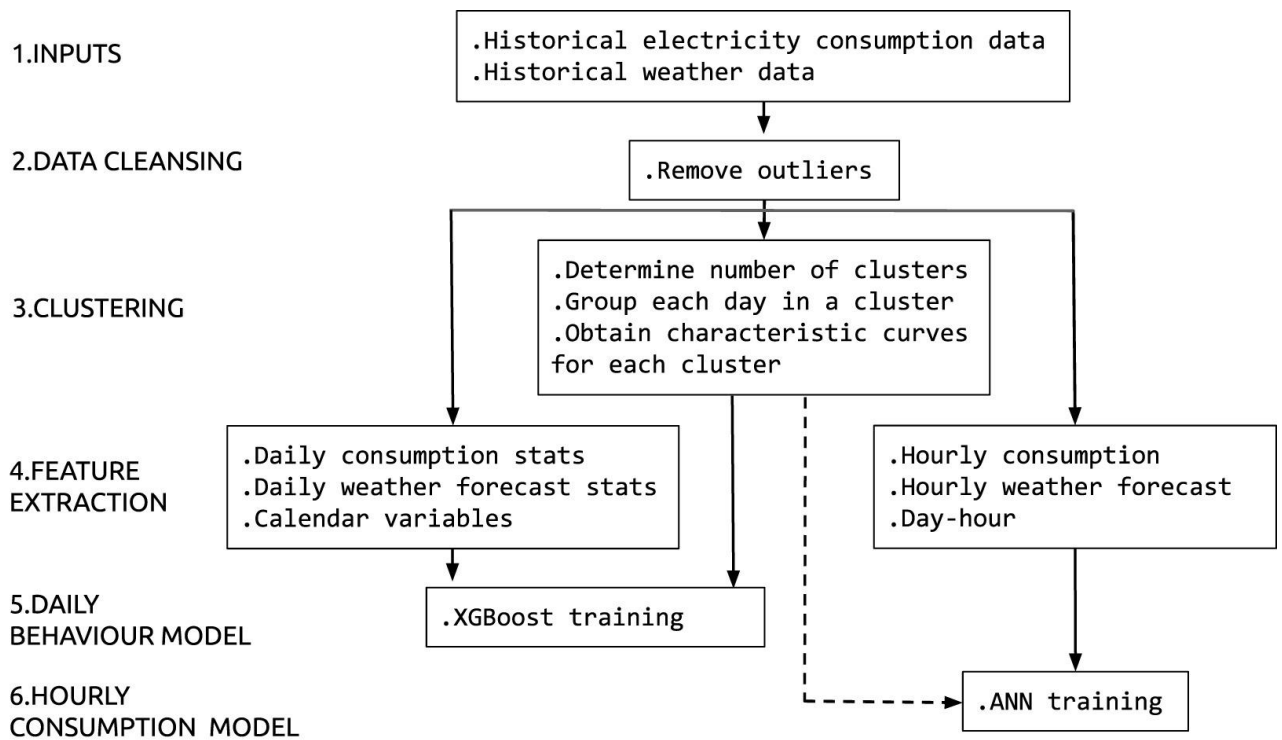


Рисунок 1.11 – Процес навчання моделі [66]

На рис. 1.11 пунктирна стрілка показує функції, які не враховуються під час запуску лише моделі споживання.

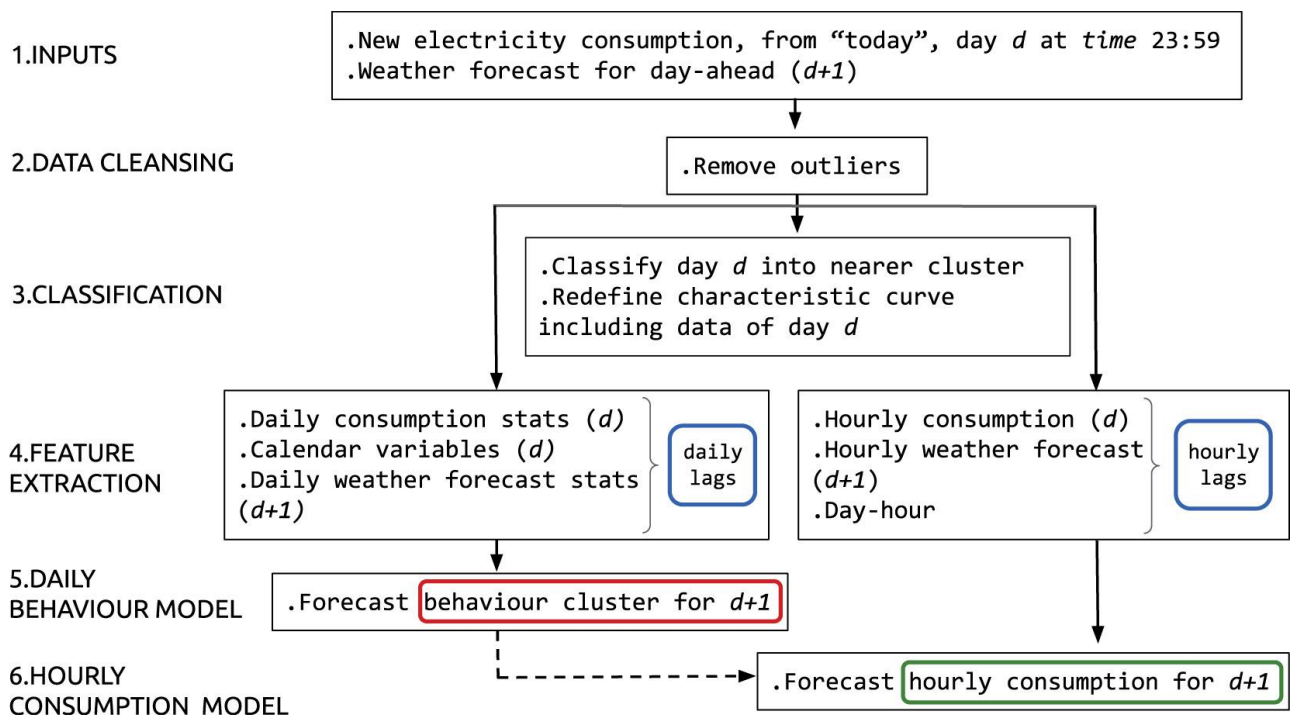
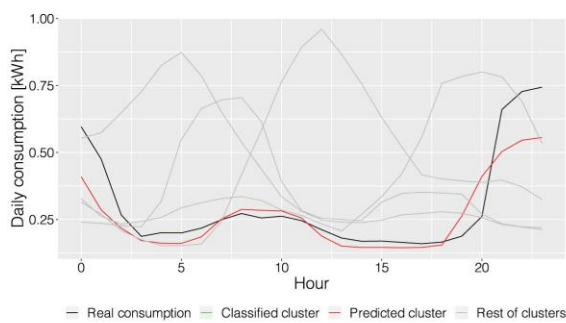


Рисунок 1.12 – Процес прогнозування енергоспоживання на добу вперед [66]

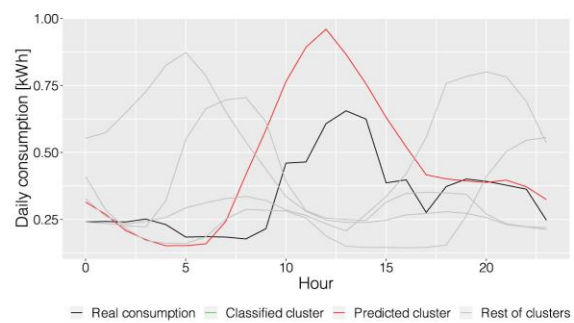
Останній етап детально показаний на рис. 1.12.

Для того, щоб працювати в повсякденному сценарії реального часу, запропонована методологія, яка складається з процесу робочого процесу, розділена на два етапи: перший – етап навчання (рис. 1.11), який використовує історичні дані та другий – етап прогнозування на добу вперед (рис. 1.12).

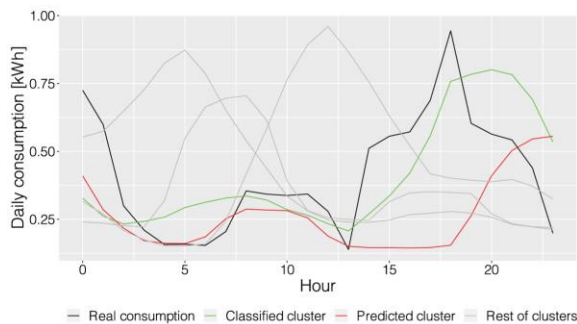
Пунктирні стрілки на рисунках показують, як відокремити модель поведінки від моделі споживання. Залежно від доступності обчислювальної обробки та шуканої точності етап навчання може виконуватися щорічно, сезонно або щомісяця. З іншого боку, етап прогнозування повинен виконуватися щодня (рис. 1.13), оскільки вимоги до обчислень набагато нижчі.



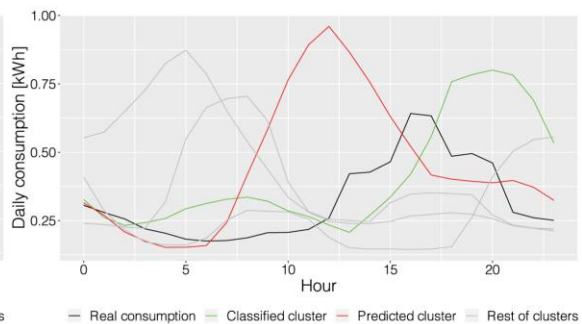
(a) case $EDA = 0$ (perfectly assigned). Predicted and classified curves are overlapped.



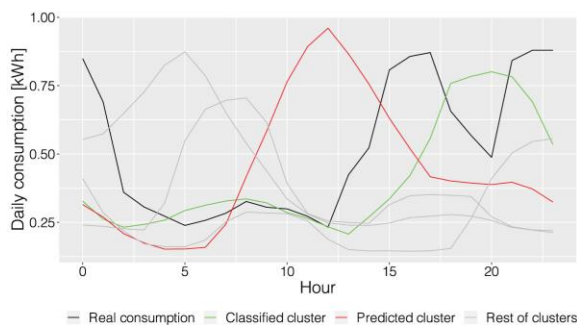
(b) case $EDA = 0$ (perfectly assigned). Predicted and classified curves are overlapped.



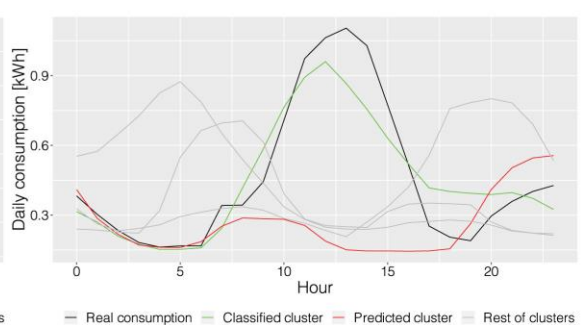
(c) case $0 < EDA \leq 1$ (correctly assigned).



(d) case $0 < EDA \leq 1$ (correctly assigned).



(e) case $1 < EDA$ (incorrectly assigned).



(f) case $1 < EDA$ (incorrectly assigned).

Рисунок 1.13 – Добові криві споживання для одного користувача
Весь процес виконується окремо для кожного клієнта.

1.4. Аналіз схем визначення прогнозованих значень енергоспоживання індивідуальними споживачами

Запропонована методологія має на меті визначити гнучкість будь-якого побутового споживання електроенергії на основі наступних кроків:

- 1) **обробка даних**, яка включає збір історичних даних загального споживання електроенергії зі звичайного інтелектуального лічильника з низькою частотою вибірки і розробка додаткових функцій;
- 2) **дезагрегація** даних, зокрема визначення стану ON–OFF приладів, які загалом мають більшу гнучкість;
- 3) **прогноз** використання техніки в короткостроковій перспективі;
- 4) **характеристика гнучкості** різних приладів.

Існує два альтернативні методи виконання етапу дезагрегації: NILM, описаний на рис. 1.14 та ІЛ, що представлений на рис. 1.15.

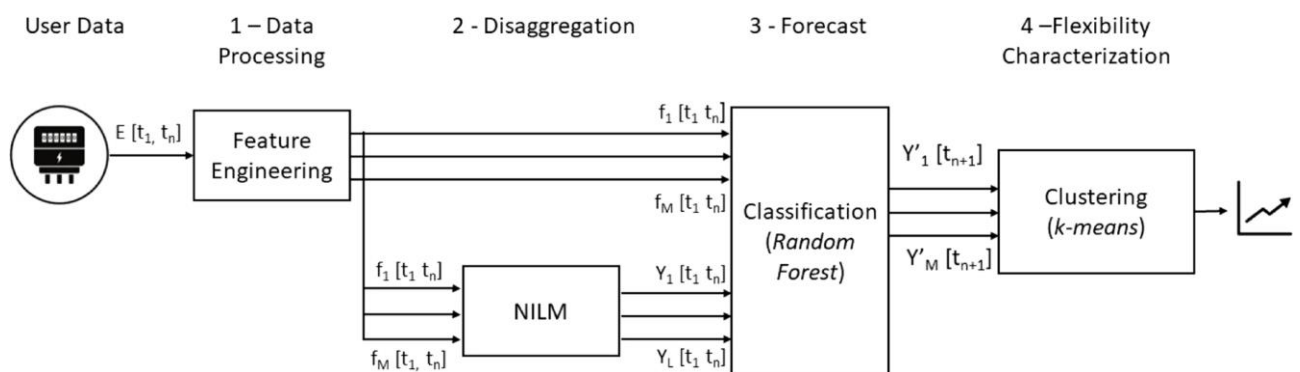


Рисунок 1.14 – Схема визначення прогнозованих значень енергоспоживання індивідуальними споживачами із NILM

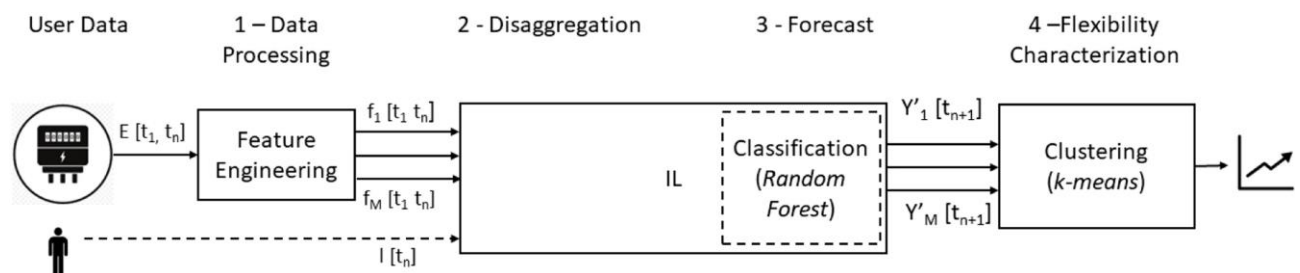


Рисунок 1.14 – Схема визначення прогнозованих значень енергоспоживання індивідуальними споживачами із ІЛ

Різні кроки докладно описані в наступних підрозділах для кожного з альтернативних методів.

1.5. Обґрунтування доцільності інтервального прогнозування індивідуального споживання електроенергії домогосподарствами з використанням рекурентних нейронних мереж

У сучасному світі, де суспільство поставити перед собою завдання зниження споживання енергоресурсів та підвищення ефективності використання електроенергії, інтервальне прогнозування індивідуального споживання електроенергії домогосподарствами стає надзвичайно важливою задачею. Обґрунтовуючи доцільність виконання кваліфікаційної роботи на дану тему, розглянемо декілька ключових аспектів:

1) Енергоефективність та стійкість постачання електроенергії. Оптимізація та забезпечення індивідуального споживання дозволяє зменшити навантаження на енергосистеми та підвищити стійкість постачання електроенергії, зменшуючи ризик перебоїв у постачанні.

2) Зменшення витрат електроенергії. Інтервальне прогнозування дозволить домогосподарствам оптимізувати своє споживання, що призведе до зменшення витрат електроенергії і, відповідно, зниження рахунків за електроенергію.

3) Збереження природних ресурсів і зменшення викидів. Зменшення споживання електроенергії сприятиме збереженню природних ресурсів та зменшення викидів CO₂, що є актуальним завданням у боротьбі з кліматичними змінами.

4) Споживча свідомість і контроль. Розробка інтервальних прогнозів споживання електроенергії надати домогосподарствам більше інформації про

їхні звички та можливості для зменшення споживання, сприяючи формуванню більшої екологічної свідомості.

5) Розвиток інформаційних технологій. Використання рекурентних нейронних мереж та інших методів аналізу даних є актуальним та інноваційним напрямком досліджень, які допоможуть розвинути навички та знання у сфері штучного інтелекту та машинного навчання.

6) Співробітництво з енергетичними компаніями. Результати дослідження можуть бути корисними для енергетичних компаній, які можуть використовувати інтервальні прогнози для оптимізації постачання та планування енергопостачання.

Усі ці аспекти об'єднуються важливою темою, яка є великим потенціалом для покращення якості життя, зменшення негативного впливу на довкілля та створення більш стійких та ефективних енергетичних систем. Таким чином, виконання кваліфікаційної роботи з цієї теми є кроком у сприянні сталому розвитку та розвитку інноваційних інформаційних технологій у сфері енергетики.

РОЗДІЛ 2.

ВИБІР ЗАСОБІВ ТА ВИКОРИСТАННЯМ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ ІНДИВІДУАЛЬНОГО СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ ДОМОГОСПОДАРСТВАМИ

2.1. Особливості прогнозування індивідуального споживання електроенергії домогосподарствами

Для прогнозування індивідуального споживання електроенергії домогосподарствами слід використовувати моделі, що базуються на часових рядах і передбачають вирішення задачі прогнозування. Зокрема, у нашій роботі будемо використовувати глибоке навчання на мові Python. При цьому для прогнозування часових рядів будемо використовувати рекурентні нейронні мережі, які базуються на одновимірних та багатовимірних моделях, а також на використанні точкового та інтервального прогнозування.

Отже наша робота базується на інтервальному прогнозуванні часових рядів за допомогою рекурентних нейронних мереж з довгою короткостроковою пам'яттю (LSTM), що забезпечує виконання прогнозу споживання електроенергії. Базовою одиницею в прихованому шарі мережі LSTM є блок пам'яті, що замінює приховані блоки в «традиційній» RNN (рис. 2.1).

Блок пам'яті містить одну або більше комірок пам'яті і пару адаптивних мультиплікативних блоків стробування, які керують входом і виходом для всіх комірок у блоці. Блоки пам'яті дозволяють коміркам мати спільні вентиля (за умови, що завдання це дозволяє), таким чином зменшуючи кількість адаптивних параметрів. Кожна комірка пам'яті має в своїй основі рекурентно самозв'язаний лінійний блок, який називається «Карусель постійних помилок» (СЕС), активацію якого ми називаємо станом комірки. СЕС вирішує проблему зникаючої помилки – за відсутності нового входу або при надходженні до комірки сигналів про помилку, локальний зворотний потік помилок у СЕС залишається постійним, не зростає і не спадає. СЕС захищений як від прямої

активації, так і від зворотного потоку помилки за допомогою вхідних і вихідних вентилів відповідно.

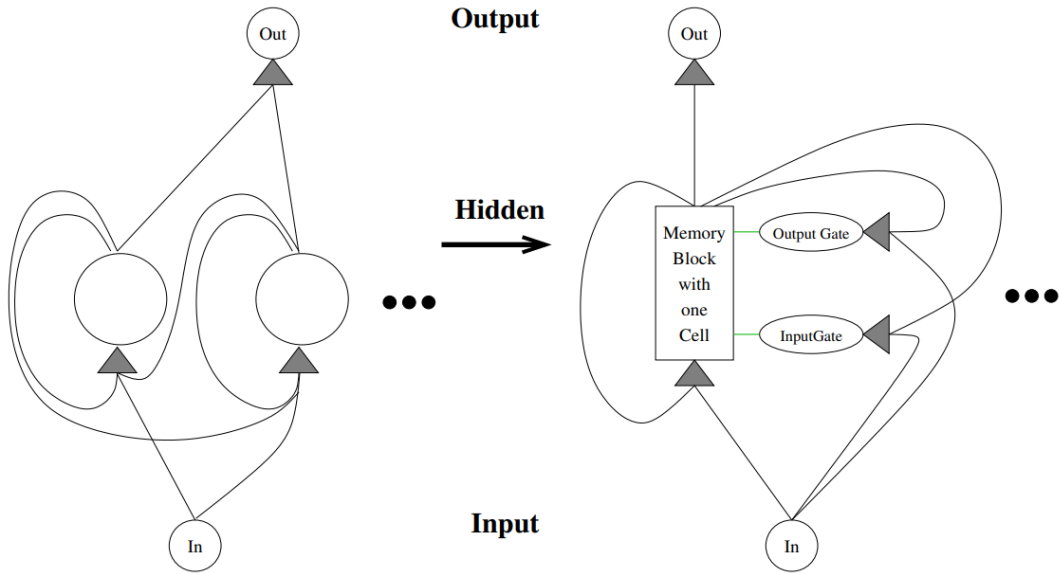


Рисунок 2.1 – ШНМ з одним повністю рекурентним прихованим шаром (ліворуч) та LSTM-мережа з блоками пам'яті у прихованому шарі (показано лише один праворуч)

Коли вентилі закриті (активація біля нуля), нерелевантні входи та шум не потрапляють до комірки, і стан комірки не впливає на решту мережі. На рис. 2.2 показано блок пам'яті з однією коміркою.

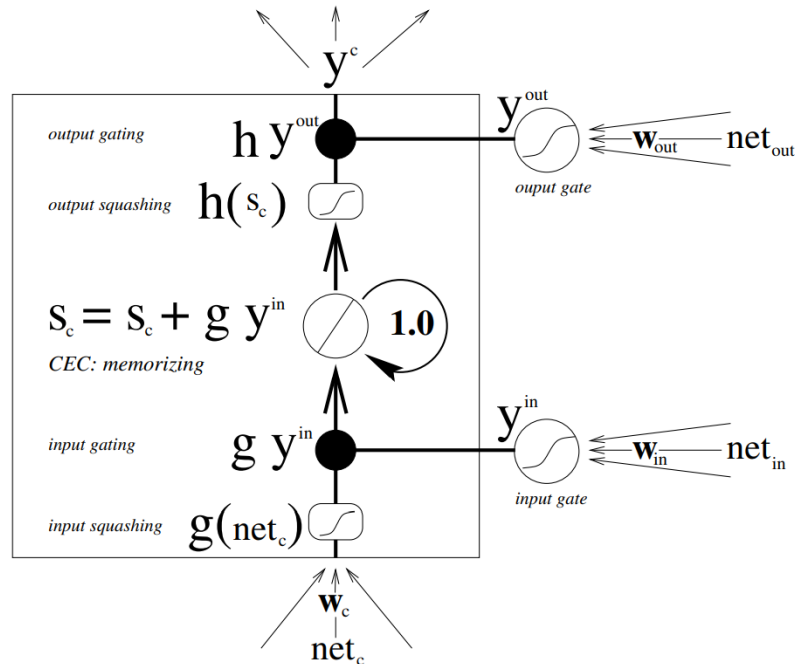


Рисунок 2.2 – Традиційна комірka LSTM має лінійний блок з рекурентним самопідключенням з вагою 1.0 (CEC)

Вхідні та вихідні вентиля регулюють доступ до читання та запису в комірку, стан якої позначається S_c . Функція g замикає вхід комірки; h замикає вихід комірки.

Стан комірки S_c , оновлюється на основі її поточного стану та трьох джерел вхідних даних net_c подається на вхід самої комірки, а net_{in} та net_{out} – на вхідні та вихідні вентиля.

Ми розглядаємо дискретні кроки часу $t = 1, 2, \dots$. Один крок включає оновлення всіх одиниць (прямий прохід) та обчислення сигналів помилки для всіх ваг (зворотний прохід). Активація вхідного вентиля y^{in} та вихідного вентиля y^{out} та обчислюються наступним чином:

$$\begin{aligned} net_{out_j}(t) &= \sum_m \omega_{out_j} m y^m(t-1); \\ y^{out_j}(t) &= f_{out_j}(net_{out_j}(t)) \end{aligned}, \quad (2.1)$$

$$\begin{aligned} net_{in_j}(t) &= \sum_m \omega_{in_j} m y^m(t-1); \\ y^{in_j}(t) &= f_{in_j}(net_{in_j}(t)) \end{aligned}, \quad (2.2)$$

У цій роботі j позначає блоки пам'яті; v позначає комірки пам'яті в блоці j (з S_j комірок), так що c_j позначає v -ту комірку j -го блоку пам'яті; w_{lm} є вагою на з'єднанні від блоку m до блоку l . Індекс m варіюється між усіма блоками-джерелами, як визначено топологією мережі (якщо активація блоку-джерела $y_m(t-1)$ відноситься до блоку-входу, то замість нього використовується поточний зовнішній вхід $y_m(t)$). Для вентилів f є логістичним сигмоїдом (з діапазоном $[0, 1]$):

$$f_x = \frac{1}{1 + e^{-x}}, \quad (2.3)$$

Вхідні дані для самої комірки мають вигляд:

$$net_{c_j^v}(t) = \sum_m \omega_{c_j^v} m y^m(t-1), \quad (2.4)$$

Яка обробляється g центрованою логістичною сигмоїдною функцією з діапазоном $[-2,2]$ (якщо не вказано інакше):

$$g(x) = \frac{4}{1 + e^{-x}} - 2, \quad (2.5)$$

Внутрішній стан комірки пам'яті $S_c(i)$ обчислюється шляхом додавання згорнутого вхідного сигналу зі стробуванням до стану на останньому кроці $S_c(t - 1)$.

Зі зростанням використання інтелектуальних лічильників електроенергії у домогосподарствах та широкого впровадження різних технологій їх виробництва, таких як сонячні батареї та інші, з'являється багато даних про споживання електроенергії. Ці дані можуть бути представлені як багатовимірний часовий ряд і використовуватися не тільки для моделювання, а й прогнозування майбутнього споживання електроенергії.

На відміну від інших алгоритмів машинного навчання, рекурентні нейронні мережі з довгою короткочасною пам'яттю (LSTM) здатні автоматично виявляти ознаки з часових послідовностей, обробляти багатовимірні дані, а також виводити послідовності змінної довжини, завдяки чому їх можна використовувати для інтервального прогнозування.

У нашій роботі пропонується використовувати для дослідження наступні рекурентні нейронні мережі з довгою короткочасною пам'яттю (LSTM):

- ✓ із одновимірними вхідними даними і вихідним вектором;
- ✓ «кодувальник-декодувальник» із одновимірним вхідними даними.

2.2. LSTM із одновимірними вхідними даними і вихідним вектором

LSTM на основі одновимірних вхідних даних і вихідного вектора – це тип LSTM, який приймає одновимірний вектор вхідних даних і генерує

одновимірний вектор вихідного даних. Цей тип LSTM може бути використаний для різних завдань, таких як класифікація, регресія та розпізнавання образів.

Конструкція LSTM на основі одновимірних вхідних даних і вихідного вектора аналогічна звичайній LSTM. Основна відмінність полягає в тому, що вхідний вектор має один розмір, а вихідний вектор також має один розмір.

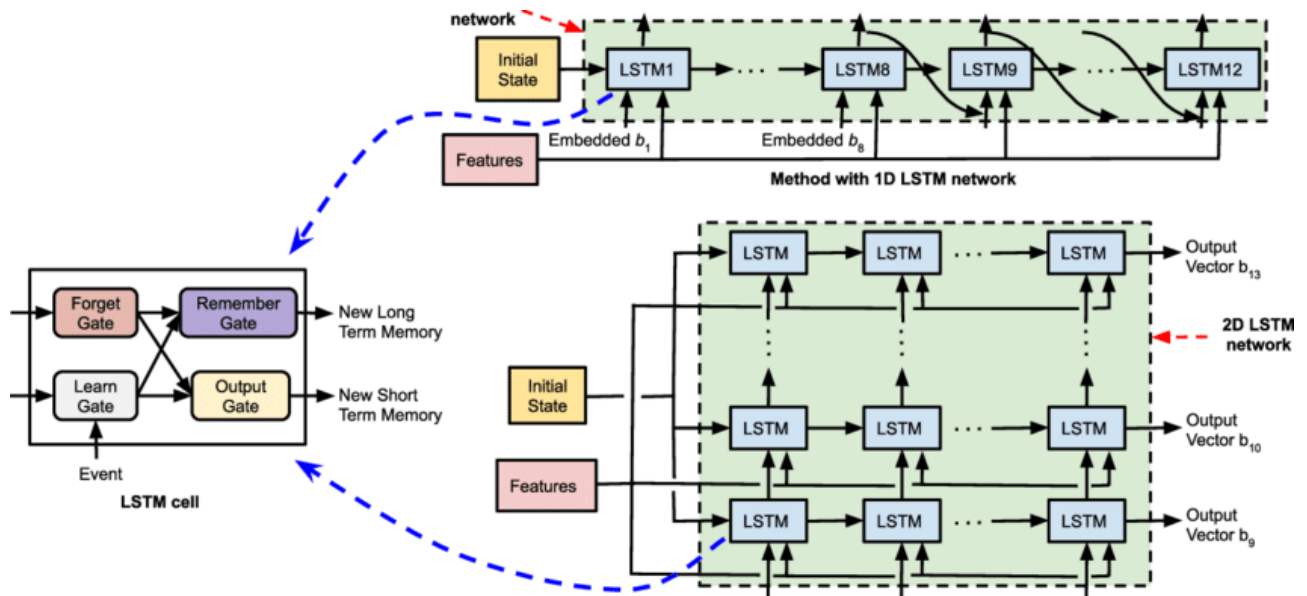


Рисунок 2.3 – Структура комірки LSTM, метод з 1D мережею LSTM та метод з 2D мережею LSTM

Наприклад, якщо вхідний вектор має розмір 10, то він буде представляти 10 різних значень. Вихідний вектор також буде мати розмір 10, і він буде представляти 10 різних значень.

LSTM на основі одновимірних вхідних даних і вихідного вектора може бути навчена за допомогою стандартної процедури навчання LSTM. Для цього використовується набір даних, який містить вхідні дані та бажані вихідні дані. LSTM навчається мінімізувати похибку між фактичними вихідними даними та бажаними вихідними даними.

Кожен запис набору даних містить одновимірний вектор вхідних даних, який представляє характеристики об'єкта. LSTM на основі одновимірних вхідних даних і вихідного вектора – це потужний інструмент, який можна використовувати для розв'язання завдання прогнозування індивідуального

споживання електроенергії домогосподарствами. Цей тип LSTM простий у реалізації та навчанні, і він може бути ефективним для вирішення багатьох завдань.

2.3. LSTM «кодувальник-декодувальник» із одновимірними даними

LSTM типу «кодувальник-декодувальник» на основі одновимірних вхідних даних – це тип LSTM, який складається з двох частин, а саме кодувальника та декодувальника. Кодувальник перетворює вхідні дані в представлений вектор, а декодер перетворює представлений вектор у вихідні дані (рис. 2.4).

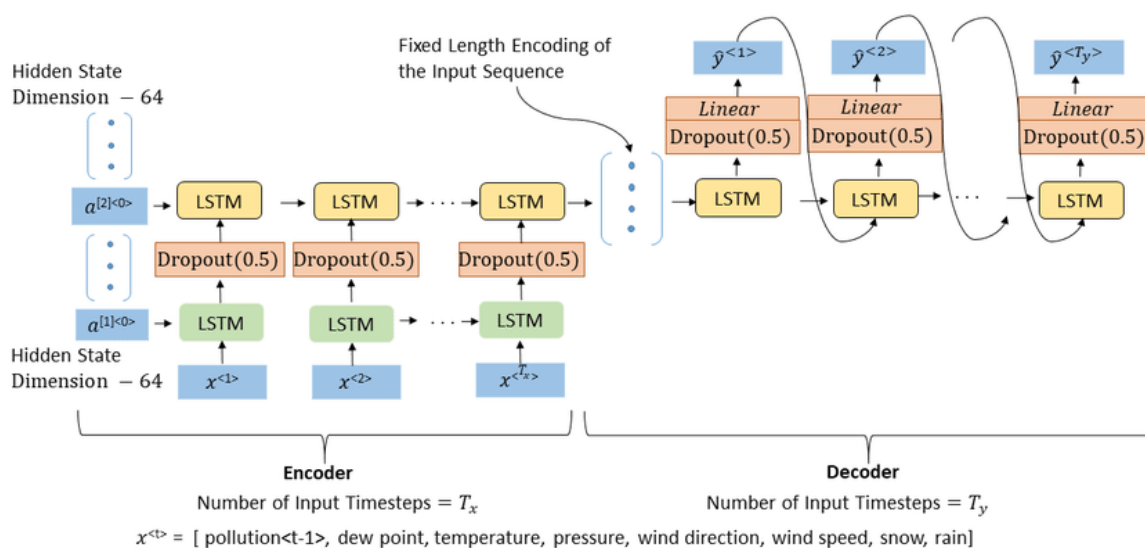


Рисунок 2.4 – Модель кодера-декодера з використанням стекованих LSTM для кодування та одного рівня LSTM для декодування

Кодувальник складається з одного або декількох шарів LSTM. Вхідні дані подаються на кодувальник послідовно, і кодувальник використовує свої LSTM-шари для створення представленого вектора.

Декодер складається з одного або декількох шарів LSTM. Представлений вектор подається на декодер, і декодер використовує свої LSTM-шари для створення вихідного вектора.

LSTM типу «кодувальник-декодувальник» на основі одновимірних вхідних даних можна використовувати для різних завдань, таких як:

✓ Переклад – кодувальник перетворює вхідні дані на мову оригіналу, а декодер перетворює представлений вектор на мову перекладу.

✓ Стиснення – кодувальник перетворює вхідні дані в представлений вектор, який є меншим за розміром, ніж вхідні дані. Декодер потім перетворює представлений вектор у вихідні дані, які є точними копіями вхідних даних.

✓ Реконструкція – кодувальник перетворює вхідні дані в представлений вектор. Декодер потім перетворює представлений вектор у вихідні дані, які є реконструйованими вхідними даними.

LSTM типу «кодувальник-декодувальник» на основі одновимірних вхідних даних – це потужний інструмент, який можна використовувати для вирішення завдання прогнозування індивідуального споживання електроенергії домогосподарствами. Цей тип LSTM простий у реалізації та навчанні, і він може бути ефективним для вирішення багатьох завдань.

2.4. Вибір засобів для створення моделей

Для виконання нашого проекту «Обґрунтування ефективної рекурентної нейронної мережевої моделі для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами» передбачалося, що будуть встановлені такі Python-бібліотеки:

- ✓ SciPy з версією Python 3;
- ✓ нейромережева бібліотека Keras (версія 2.2 або вище) з низькорівневою бібліотекою TensorFlow або Theano;
- ✓ Scikit-Learn;
- ✓ Pandas;
- ✓ NumPy;
- ✓ Matplotlib.

Після цього здійснено налаштування середовища Python. Для роботи над проектом використовували Jupyter де і налаштовували параметри середовища Python для машинного та глибокого навчання. Jupyter використовується для розробки та тестування нейронних мереж. Він дозволяє легко створювати інтерактивні блокноти, які містять код, текст, графіки та інші елементи.

SciPy – це бібліотека Python для наукового обчислення. Вона включає в себе набори даних, функції, алгоритми та інші інструменти, які використовуються в різних галузях науки та інженерії. SciPy є важливою бібліотекою для машинного та глибокого навчання, оскільки вона надає широкий спектр функцій для обробки даних, статистики, математики та ін. SciPy є основною бібліотекою для наукового обчислення, яка використовується в проекті. Вона надає широкий спектр функцій для обробки даних, статистики, математики та ін., які необхідні для виконання різних завдань, пов'язаних з прогнозуванням споживання електроенергії.

Keras – це високорівнева нейромережева бібліотека Python. Вона побудована на основі низькорівневих бібліотек TensorFlow або Theano. Keras спрощує розробку та навчання нейронних мереж, оскільки вона надає простий і інтуїтивний інтерфейс. Keras використовується для розробки та навчання нейронних мереж. Вона є високорівневою бібліотекою, яка спрощує розробку нейронних мереж, оскільки вона надає простий і інтуїтивний інтерфейс.

Scikit-Learn – це бібліотека Python для машинного навчання. Вона включає в себе набори даних, алгоритми та інші інструменти, які використовуються для різних завдань машинного навчання, таких як класифікація, регресія, кластеризація та ін. Scikit-Learn є важливою бібліотекою для машинного та глибокого навчання, оскільки вона надає широкий спектр алгоритмів машинного навчання, які можна використовувати для різних завдань. Scikit-Learn використовується для оцінки ефективності нейронних мереж. Вона включає в себе набори даних, алгоритми та інші інструменти, які використовуються для різних завдань машинного навчання, таких як класифікація, регресія, кластеризація та ін.

Pandas – це бібліотека Python для роботи з аналітичними даними. Вона надає широкий спектр функцій для обробки, аналізу та візуалізації даних. Pandas є важливою бібліотекою для машинного та глибокого навчання, оскільки вона дозволяє легко працювати з даними, які використовуються для навчання нейронних мереж. Pandas використовується для роботи з даними, які використовуються для навчання нейронних мереж. Вона надає широкий спектр функцій для обробки, аналізу та візуалізації даних.

NumPy – це бібліотека Python для чисельних обчислень. Вона надає швидкий і ефективний доступ до даних. NumPy є важливою бібліотекою для машинного та глибокого навчання, оскільки вона використовується для зберігання та обробки даних, які використовуються для навчання нейронних мереж. NumPy використовується для зберігання та обробки даних, які використовуються для навчання нейронних мереж. Вона є швидким і ефективним способом доступу до даних.

Matplotlib – це бібліотека Python для візуалізації даних. Вона надає широкий спектр функцій для створення графіків, діаграм та інших візуалізацій. Matplotlib є важливою бібліотекою для машинного та глибокого навчання, оскільки вона дозволяє легко візуалізувати результати навчання нейронних мереж. Matplotlib використовується для візуалізації результатів навчання нейронних мереж. Вона надає широкий спектр функцій для створення графіків, діаграм та інших візуалізацій.

РОЗДІЛ 3.

РЕЗУЛЬТАТИ ОБҐРУНТУВАННЯ МОДЕЛЕЙ ІЗ РЕКУРЕНТНИМИ НЕЙРОННИМИ МЕРЕЖАМИ ДЛЯ ІНТЕРВАЛЬНОГО ПРОГНОЗУВАННЯ ІНДИВІДУАЛЬНОГО СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ ДОМОГОСПОДАРСТВАМИ

3.1. Підготовка даних для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами

Набір даних для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами являє собою багатовимірний часовий ряд, що відображає споживання електроенергії однієї сім'ї протягом декількох років. Нами було зібрано дані в період з грудня 2019 року по листопад 2023 року. Виміри споживання електроенергії виконувалися розумним лічильником кожну хвилину. Усі інтелектуальні функції приладу обліку забезпечуються вбудованим радіомодулем. Цей пристрій реєструє імпульси, що генеруються електролічильником, і зберігає їх в енергонезалежній пам'яті. Радіомодулі відправляють дані лічильників до ресурсних компаній через безпечні бездротові канали зв'язку (мобільний зв'язок, Wi-Fi або 3G-з'єднання). Різні моделі інтелектуальних електролічильників можуть мати відмінності у функціоналі. Проте всі вони діють за однією основною ідеєю: зчитують дані з приладу обліку та передають їх на сервер.

Крім дати і часу набір наші дані складаються із наступних атрибутів:

1. `global_active_power` – загальна активна потужність, споживана житловим будинком (вимірюється в кіловатах);
2. `global_reactive_power` – загальна реактивна потужність, споживана житловим будинком (вимірюється в кіловатах);
3. `voltage` – середня напруга (вимірюється у вольтах);
4. `global_intensity` – середнє значення сили струму (вимірюється в амперах);

5. `sub_metering_1` – активна енергія споживання кухонної кімнати (вимірюється у ват-годинах активної енергії);
6. `sub_metering_2` – активна енергія споживання пральні (вимірюється у ват-годинах активної енергії);
7. `sub_metering_3` – активна енергія споживання систем клімат-контролю (вимірюється у ват-годинах активної енергії).

Активна і реактивна енергія відносяться до технічних деталей змінного струму (активна потужність це і є споживана потужність, за яку ми, власне, платимо за лічильником). Змінна `sub_metering` може бути створена шляхом віднімання суми трьох визначених `sub_metering`-змінних від загальної споживаної енергії в такий спосіб:

$$\text{sub_metering_remainder} = (\text{global_active_power} * 1000 / 60) - (\text{sub_metering_1} + \text{sub_metering_2} + \text{sub_metering_3}) \quad (3.1)$$

Завантаження та підготовка набору даних виконували у блокноті Jupyter Notebook. Для цього виконуємо імпорт потрібних функцій та бібліотек, що показано на рис. 3.1.

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
from sklearn.model_selection import train_test_split
from pandas import read_csv
from matplotlib import pyplot as plt
from matplotlib import pyplot
```

Рис. 3.1. Імпорт потрібних функцій та бібліотек у блокнот Jupyter Notebook

Бібліотеки Pandas і NumPy використовуватимемо для завантаження даних, їх очищення, виконання статистичних обчислень та аналізу. Бібліотеки Seaborn і Matplotlib дозволяють створювати різноманітні графіки, діаграми та інші візуалізації для дослідження даних та виявлення закономірностей. Функція `train_test_split` використовується для підготовки наборів даних для навчання та тестування моделей машинного навчання, а бібліотека `scikit-learn` надає широкий спектр алгоритмів машинного навчання.

Після цього виконуємо завантаження наших даних, використовуючи функцію `read_csv` та `dataset.head()` (рис. 3.2).

```
In [11]: dataset.head()
```

```
Out[11]:
```

	Global_active_power	Global_reactive_power	Voltage	Global_intensity	Sub_metering_1	Sub_metering_2	Sub_metering_3	sub_metering_4
datetime								
2019-12-16 17:24:00	3.87872	0.38456	227.7948	16.928	0.0	0.92	15.64	48.085335
2019-12-16 17:25:00	4.93120	0.40112	226.6211	21.160	0.0	0.92	14.72	66.546669
2019-12-16 17:26:00	4.94408	0.45816	226.2913	21.160	0.0	1.84	15.64	64.921341
2019-12-16 17:27:00	4.95696	0.46184	226.7278	21.160	0.0	0.92	15.64	66.056000
2019-12-16 17:28:00	3.37272	0.48576	228.6096	14.536	0.0	0.92	15.64	39.652000

Рисунок 3.2 – Результати завантаження та виведення масиву початкових даних споживання електроенергії домогосподарствами

Нами використано функції Pandas і NumPy для виявлення та обробки пропущених значень, аномалій та інших проблем у даних. Насамперед проаналізовано інформацію про тип даних (рис. 3.3).

```
In [13]: dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 2075259 entries, 2019-12-16 17:24:00 to 2023-11-26 21:02:00
Data columns (total 8 columns):
#   Column                Dtype
---  -
0   Global_active_power    float64
1   Global_reactive_power  float64
2   Voltage                float64
3   Global_intensity       float64
4   Sub_metering_1         float64
5   Sub_metering_2         float64
6   Sub_metering_3         float64
7   sub_metering_4         float32
dtypes: float32(1), float64(7)
memory usage: 134.6 MB
```

Рисунок 3.3 – Результати аналізу типу даних споживання електроенергії домогосподарствами

Закінчуються дані у кінці листопада 2023 року, найближча остання субота була 20 листопада. З урахуванням цього на валідацію відводиться 46 тижнів.

```
In [17]: dataset.describe()
```

```
Out[17]:
```

	Global_active_power	Global_reactive_power	Voltage	Global_intensity	Sub_metering_1	Sub_metering_2	Sub_metering_3	sub_metering_4
count	2.049280e+06	2.049280e+06	2.049280e+06	2.049280e+06	2.049280e+06	2.049280e+06	2.049280e+06	2.049280e+06
mean	1.004286e+00	1.212402e-01	2.336147e+02	4.257539e+00	1.032169e+00	1.194638e+00	5.941772e+00	8.569520e+00
std	9.727106e-01	1.104675e-01	3.142787e+00	4.088845e+00	5.660789e+00	5.356264e+00	7.762182e+00	8.812889e+00
min	6.992000e-02	0.000000e+00	2.165040e+02	1.840000e-01	0.000000e+00	0.000000e+00	0.000000e+00	-2.207998e+00
25%	2.833600e-01	4.704000e-02	2.318203e+02	1.288000e+00	0.000000e+00	0.000000e+00	0.000000e+00	3.496000e+00
50%	5.538400e-01	9.800000e-02	2.337797e+02	2.392000e+00	0.000000e+00	0.000000e+00	9.200000e-01	5.060000e+00
75%	1.405760e+00	1.901200e-01	2.356033e+02	5.888000e+00	0.000000e+00	9.200000e-01	1.564000e+01	9.537334e+00
max	1.023224e+01	1.362200e+00	2.465255e+02	4.452800e+01	8.096000e+01	7.360000e+01	2.852000e+01	1.148467e+02

Рисунок 3.4 – Результати оцінення статистичних характеристик даних споживання електроенергії домогосподарствами

Стосовно атрибуту «Global_active_power», то він має середнє значення близько 1,0, але також має значне стандартне відхилення, що вказує на мінливість у споживанні активної потужності. Атрибут «Global_reactive_power» має відносно невелике середнє значення та стандартне відхилення, що вказує на більш стабільне споживання реактивної потужності. «Voltage» має середнє значення близько 233 В і відносно невелике стандартне відхилення, що вказує на стабільність напруги. «Global_intensity» має середнє значення близько 4,2 А, але також має значне стандартне відхилення, що вказує на мінливість у загальній інтенсивності струму. «Sub_metering_1», «Sub_metering_2», «Sub_metering_3» мають значні відмінності у 25%, 50% та 75% процентилях, що вказує на різні закономірності споживання електроенергії. «Sub_metering_4» має негативне мінімальне значення, що може потребувати подальшого дослідження.

3.2. Аналіз даних для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами

Візуалізація даних виконувалася завдяки створенню графіків та діаграм за допомогою бібліотек Seaborn або Matplotlib, щоб оцінити зв'язки між змінними та виявити закономірності. Насамперед було побудовано гістограму зміни загальної активної потужності, споживаної житловим будинком (кВт) (рис. 3.5).

Побудована гістограма показує розподіл значень змінної Global_active_power. Середнє значення змінної становить 1,00 кВт, а середньоквадратичне відхилення - 0,97 кВт.

На основі графіка можна сказати, що більшість значень змінної припадає на інтервал від 0,5 до 1,5 кВт. Значення змінної відхиляються від середнього значення не більше ніж на два середньоквадратичних відхилення. Є кілька значень, які значно перевищують середньоквадратичне відхилення.

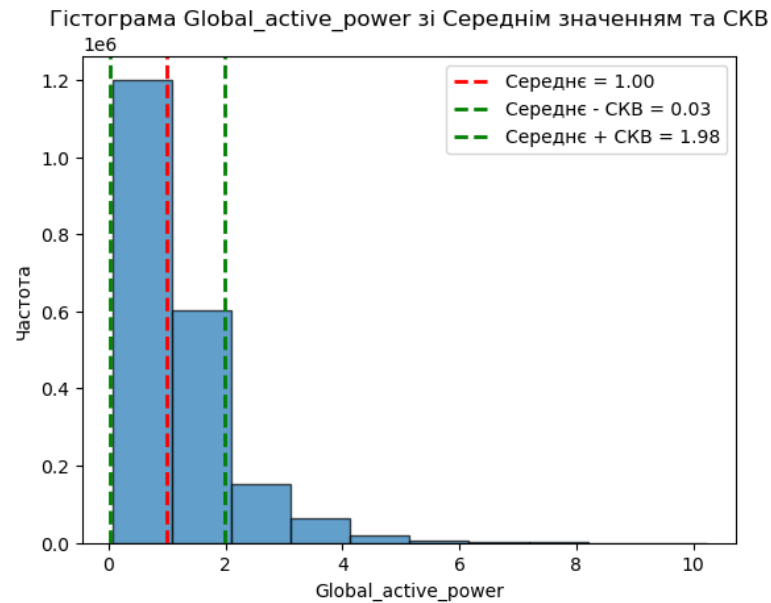


Рисунок 3.5 – Гістограма зміни загальної активної потужності, споживаної житловим будинком (кВт)

Отримані результати можна використовувати для подальшого аналізу змінної «Global_active_power». Наприклад, можна провести дослідження, щоб з'ясувати, чому деякі значення змінної значно перевищують середньоквадратичне відхилення. Загалом, гістограма показує, що розподіл значень змінної Global_active_power є відносно стійким. Однак, є кілька значень, які значно відхиляються від середнього значення, і їх слід розглянути окремо.

Нами аналізувалися закономірності спостережень у часі, що являють багатовимірний часовий ряд, і найкращий спосіб зрозуміти цей часовий ряд – це створити лінійні графіки. Ми можемо почати зі створення окремого лінійного графіка для кожної з восьми змінних.

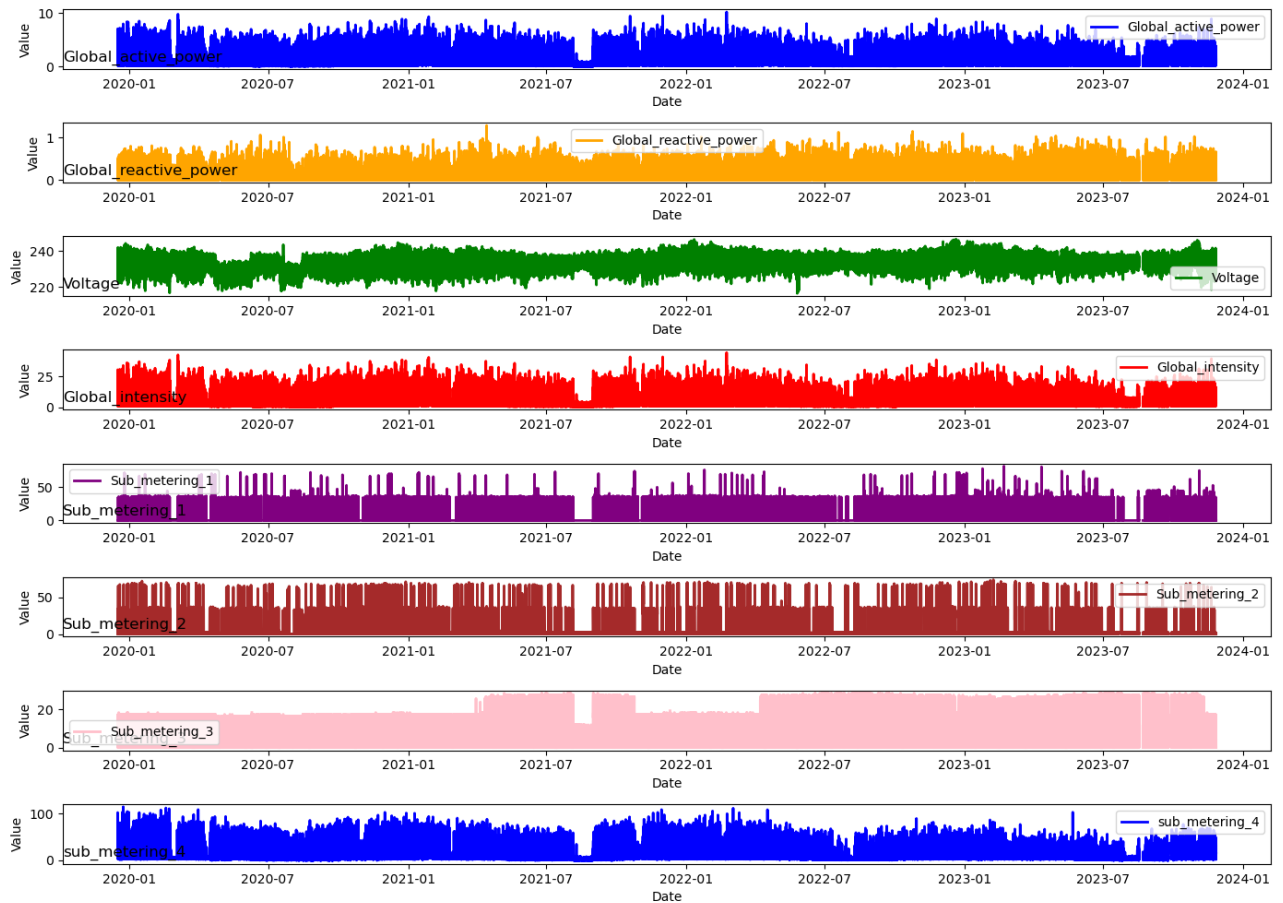


Рисунок 3.6 – Тенденції зміни складових індивідуального споживання електроенергії домогосподарствами

Запуск створеного коду створює єдине зображення з вісьмома підсхемами, по одній для кожної змінної (рис. 3.6). Встановлено, що в «Sub_metering_3» (активна енергія споживання систем клімат-контролю, кВАр) відбувалося щось аномальне, що може не відповідати спекотним чи холодним рокам. Можливо, встановили нову систему кондиціонування повітря. Встановлено що, атрибут «Sub_metering_4», зменшується з часом або демонструє тенденцію до зниження, можливо, збігаючись із значним збільшенням, яке спостерігається наприкінці для «Sub_metering_3». Ці спостереження підсилюють необхідність дотримуватися тимчасового порядку підпоследовностей цих даних під час підгонки та оцінки будь-якої моделі.

Побудуємо графік активної потужності для кожного року, щоб побачити, чи є спільні закономірності за роки. Перший 2019 рік має менше ніж один місяць даних, тому його відобразити не будемо на графіка (3.7).

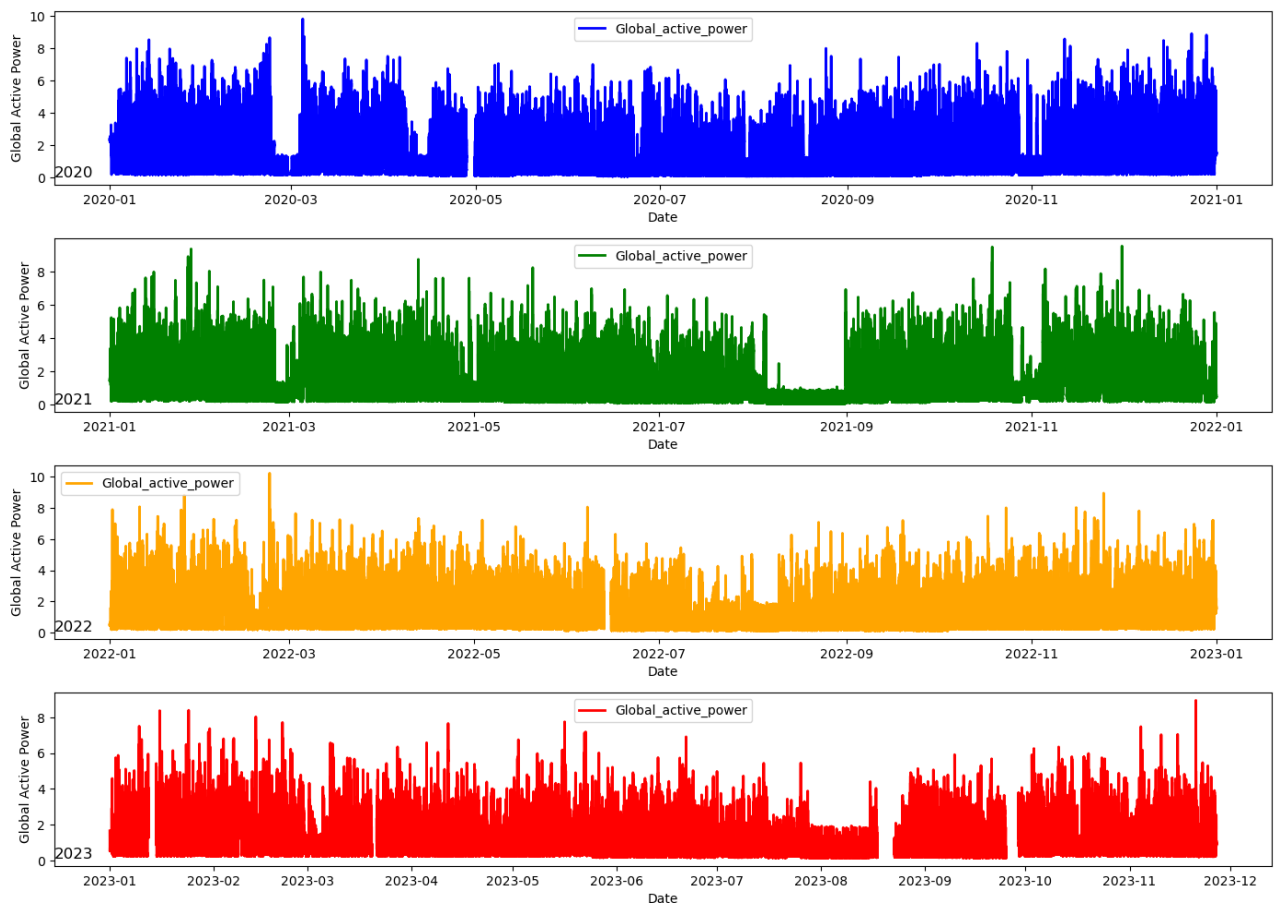


Рисунок 3.7 – Тенденції зміни складових індивідуального споживання електроенергії домогосподарствами

Спостерігається деякі загальні закономірності протягом багатьох років, наприклад, у лютому-березні та серпні-вересні, де помітно спостерігається зниження споживання електроенергії домогосподарством. Також спостерігається тенденція до зниження споживання електроенергії домогосподарством впродовж літніх місяців (середина року) та зростання споживання в зимові місяці. Вони можуть вказувати на річну сезонну структуру споживання. Також є кілька фрагментів відсутніх даних принаймні на першому, третьому та четвертому графіках.

Нами розглянуто місячне індивідуальне споживання електроенергії домогосподарствами впродовж 2023 року (рис. 3.8).

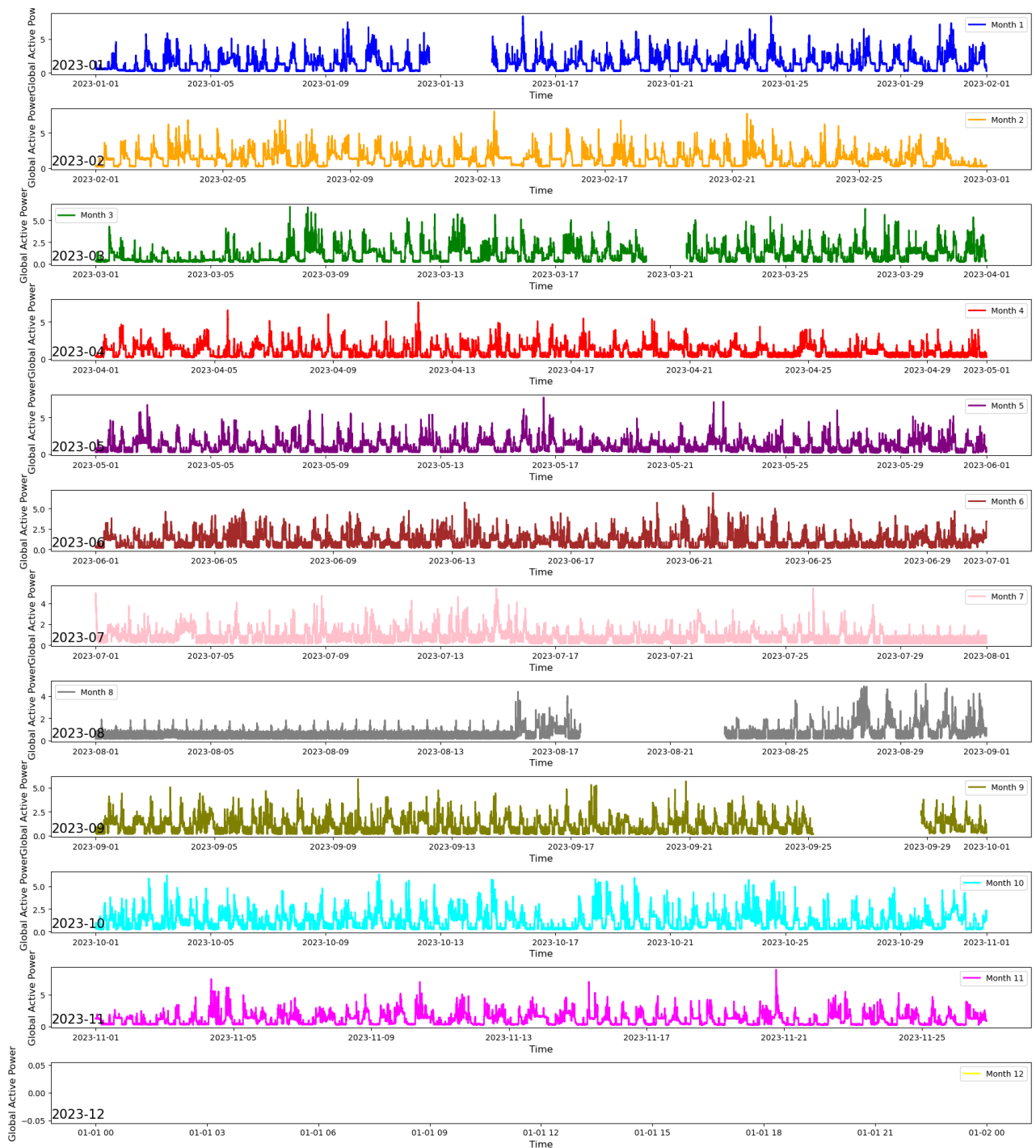


Рисунок 3.7 – Тенденції зміни місячного індивідуального споживання електроенергії домогосподарством впродовж 2023 року

Отримані тенденції зміни місячного індивідуального споживання електроенергії домогосподарством впродовж 2023 року дають можливість виявити загальні структури за місяці, наприклад щоденні та тижневі моделі.

Існує спільність протягом днів; наприклад, багато днів споживання починається рано вранці, приблизно о 6-7 ранку. У деякі дні спостерігається

зниження споживання в середині дня, що може мати сенс, якщо більшість мешканців знаходяться поза домом.

У деякі дні ми бачимо значне споживання протягом ночі, що в північній півкулі січень може збігатися з використанням системи опалення. Пора року, зокрема сезон і погода, яку він приносить, буде важливим фактором у моделюванні цих даних, як і очікувалося.

Враховуючи це, нами виконано розподіли даних часових рядів, які характеризують індивідуальне споживання електроенергії домогосподарством (рис. 3.8).

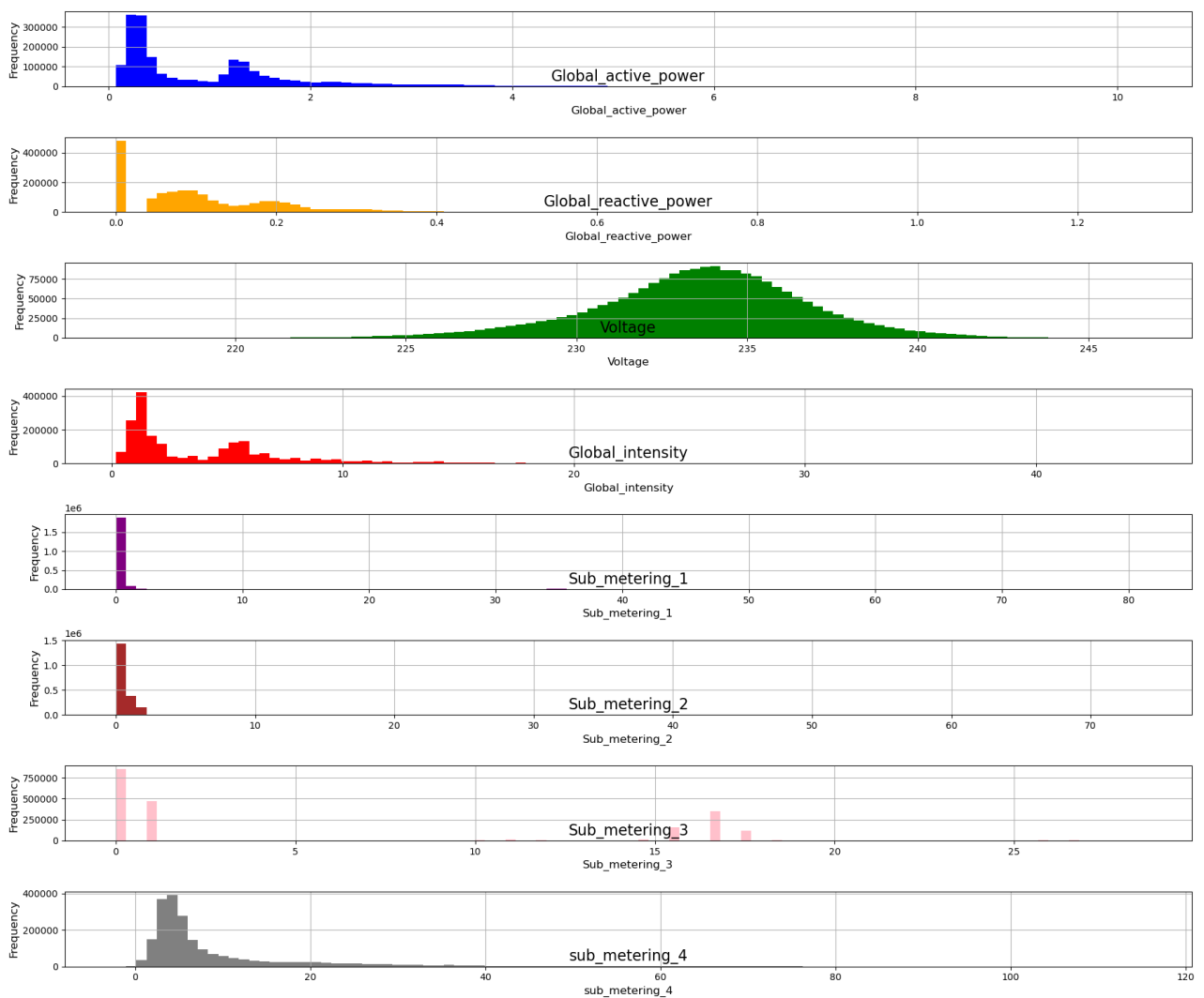


Рисунок 3.8 – Розподіли даних часових рядів, які характеризують індивідуальне споживання електроенергії домогосподарством

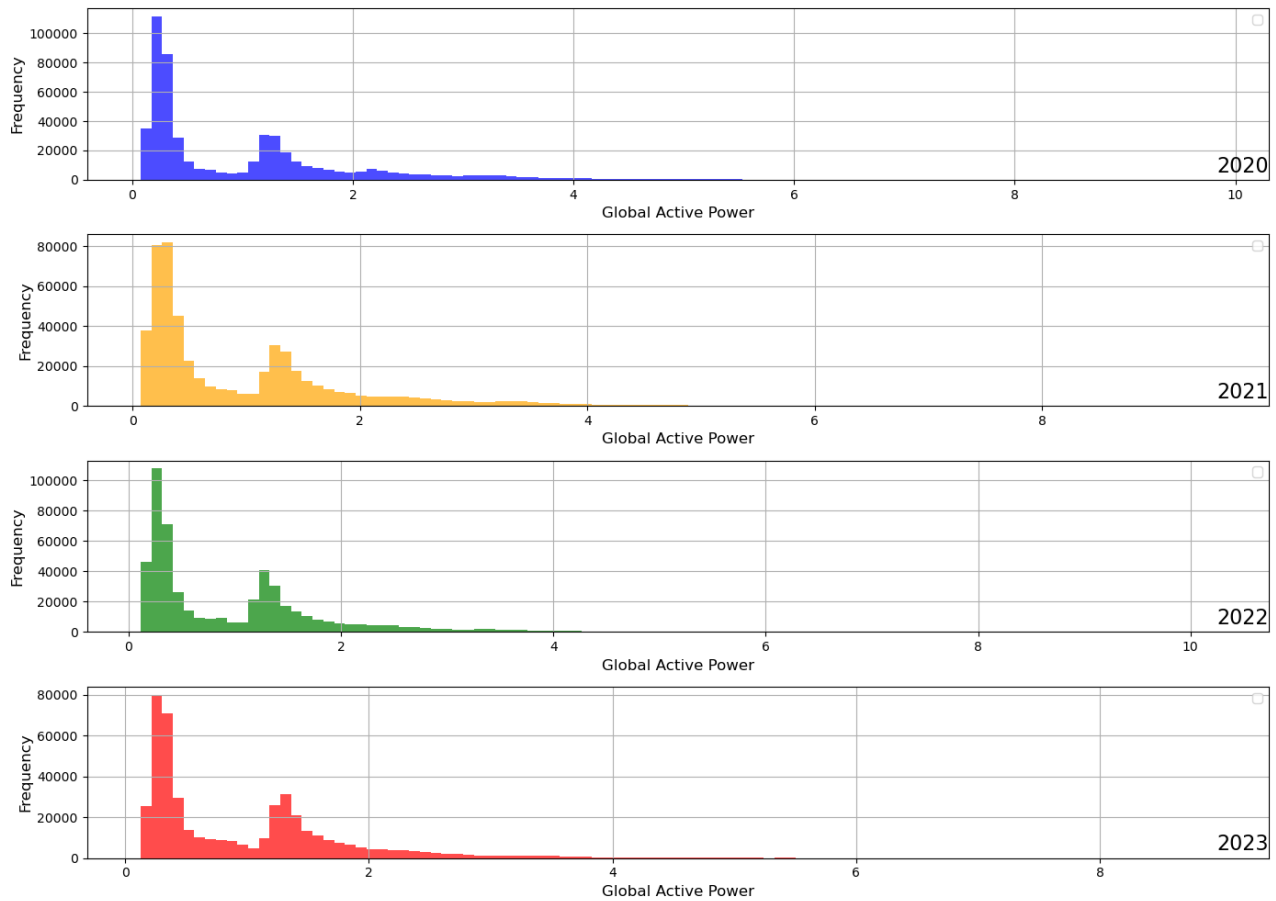


Рисунок 3.9 – Розподіли змінної «Global_active_power», яка характеризує індивідуальне споживання електроенергії домогосподарством, впродовж окремих років

Встановлено, що розподіл споживання активної потужності за досліджувані роки виглядає дуже схожим. Розподіл дійсно бімодальний з одним піком близько 0,3 кВт і, можливо, іншим близько 1,3 кВт. Існує довгий хвіст розподілу до зростання споживання електроенергії. Це може відкрити двері для ідей дискретизації даних і поділу їх на пік 1, пік 2 або довгий хвіст. Ці групи або кластери для використання в день або годину можуть бути корисними для розробки прогнозної моделі.

Означене завдання відноситься до завдань інтервального прогнозування часових рядів. При цьому модель, на вхід якої подається більше однієї ознаки, називається моделлю на основі багатовимірних вхідних даних (або багатовимірною моделлю). Вона може використовуватися для виконання інтервального прогнозування.

3.3. Розробка моделі LSTM із одновимірними вхідними даними та результати прогнозування індивідуального споживання електроенергії

Нами розроблена проста модель LSTM із одновимірними вхідними даними, яка трансформує загальне щотижневе енергоспоживання в денні послідовності та виконує прогноз енергоспоживання на наступний тиждень у вигляді вихідного вектора. Саме ця модель буде основою для виконання складніших моделей, які розроблено та описано нижче. Кількість попередніх днів, використовуваних як вхідні дані, визначає одновимірну послідовність даних, яку буде подано на вхід моделі LSTM із одновимірними вхідними даними для вивчення і виділення ознак. Архітектура моделі LSTM із одновимірними вхідними даними представлена на рис. 3.10.

```
Model: "sequential_1"
```

Layer (type)	Output Shape	Param #
lstm_1 (LSTM)	(None, 50)	11600
dense_1 (Dense)	(None, 1)	51

```

Total params: 11651 (45.51 KB)
Trainable params: 11651 (45.51 KB)
Non-trainable params: 0 (0.00 Byte)

```

Рисунок 3.10 – Архітектура моделі LSTM із одновимірними вхідними даними

Була використана архітектура LSTM для роботи з одновимірними часовими рядами (рис. 3.10). Модель складалася зі шару LSTM з 50 нейронами та вихідного Dense шару з одним нейроном. Модель компілювалася з функцією втрати `mean_squared_error` та оптимізатором Adam.

Для аналізу і прогнозування індивідуального споживання електроенергії домогосподарствами було розроблено модель, використовуючи дані, які були представлені у вигляді часових рядів. Для цього був використаний датасет, який включав ознаки, такі як `Global_active_power`, `Global_reactive_power`,

Voltage, Global_intensity, Sub_metering_1, Sub_metering_2, Sub_metering_3, Sub_metering_4.

```
In [24]: import numpy as np
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
from keras.models import Sequential
from keras.layers import LSTM, Dense
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt

# Зчитуємо датасет
df = pd.read_csv('F:\Dipl_2023\Dipl_mag_126\Малетич\household_power_consumption_days.csv', parse_dates=['datetime'], index_col='datetime')

# Вибираємо ознаки для вхідних даних і вихідний вектор
features = ['Global_active_power', 'Global_reactive_power', 'Voltage', 'Global_intensity', 'Sub_metering_1', 'Sub_metering_2', 'Sub_metering_3', 'Sub_metering_4']
target = 'sub_metering_4'

# Виділяємо дані для тренування
X = df[features].values
y = df[target].values

# Масштабуємо дані
scaler_X = MinMaxScaler(feature_range=(0, 1))
X_scaled = scaler_X.fit_transform(X.reshape(-1, len(features)))

scaler_y = MinMaxScaler(feature_range=(0, 1))
y_scaled = scaler_y.fit_transform(y.reshape(-1, 1))
```

Рисунок 3.11 – Код імпорту бібліотек, зчитування та підготовки даних для навчання моделі LSTM із одновимірними вхідними даними

Дані були масштабовані за допомогою Min-Max масштабування для забезпечення оптимальності роботи нейронної мережі LSTM. Датасет був розділений на тренувальний та тестовий набори для ефективно оцінки моделі (рис. 3.12).

```
# Розділяємо дані на тренувальний та тестовий набори
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y_scaled, test_size=0.2, random_state=42)

# Змінюємо форму даних для LSTM
X_train = X_train.reshape((X_train.shape[0], 1, X_train.shape[1]))
X_test = X_test.reshape((X_test.shape[0], 1, X_test.shape[1]))

# Будуємо LSTM модель
model = Sequential()
model.add(LSTM(50, input_shape=(X_train.shape[1], X_train.shape[2])))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')

# Тренуємо модель
history = model.fit(X_train, y_train, epochs=10, batch_size=32, validation_data=(X_test, y_test), verbose=1)

# Дійсна передбачення
y_pred = model.predict(X_test)

# Зворотнє масштабування передбачень та дійсних значень
y_pred_inv = scaler_y.inverse_transform(y_pred)
y_test_inv = scaler_y.inverse_transform(y_test)

# Виводимо середньоквадратичну помилку (MSE)
mse = mean_squared_error(y_test_inv, y_pred_inv)
print('Mean Squared Error:', mse)
```

Рисунок 3.12 – Код навчання та оцінки моделі LSTM із одновимірними вхідними даними

Модель була навчена протягом 50 епох з пакетом розміром 32. Під час тренування використовувались дані з тренувального набору, і результати валідації були оцінені на тестовому наборі. Для оцінки ефективності моделі використовувалася середньоквадратична помилка (MSE).

Після навчання моделі була проведена оцінка на тестовому наборі. Отримана середньоквадратична помилка дозволила оцінити точність прогнозів.

Останнім кроком було розроблення коду для виведення графіків втрат під час тренування та валідації моделі (рис. 3.13).

```
# Графіки метрик
plt.plot(history.history['loss'], label='Train Loss')
plt.plot(history.history['val_loss'], label='Test Loss')
plt.title('Model Loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()
plt.show()

# Графік прогнозу та дійсних значень
plt.plot(y_test_inv, label='True Values')
plt.plot(y_pred_inv, label='Predicted Values')
plt.title('True vs Predicted Values')
plt.xlabel('Data Point')
plt.ylabel('sub_metering_4')
plt.legend()
plt.show()
```

Рисунок 3.13 – Код для виведення графіків втрат під час тренування та валідації моделі LSTM із одновимірними вхідними даними

На графіках будуть відображені тенденції зміни втрат під час тренування та валідації моделі, а також порівняння фактичних та прогнозованих значень для оцінки точності моделі.

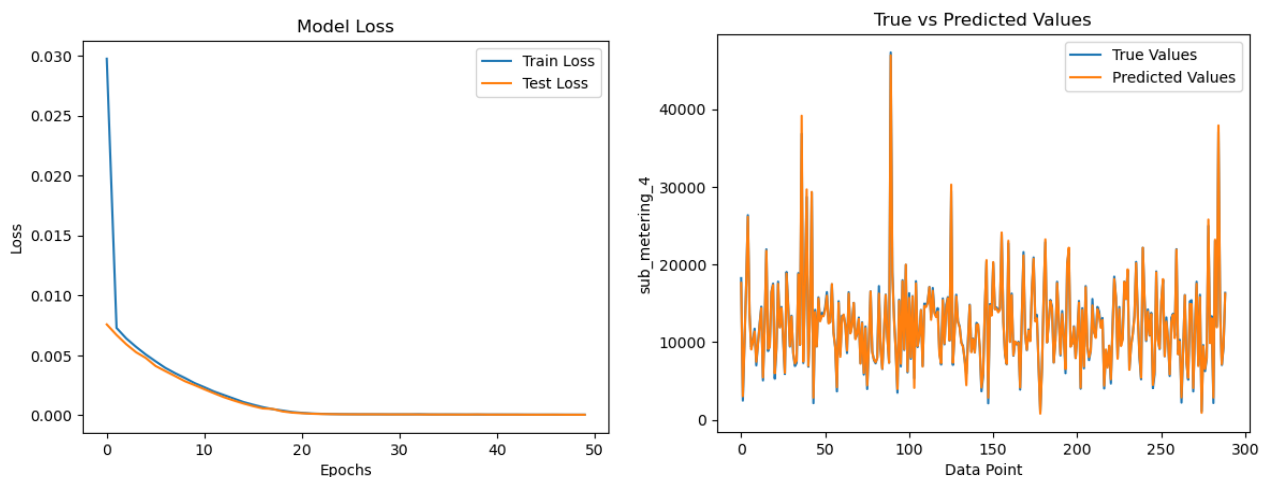


Рисунок 3.14 – Графіки втрат під час тренування та валідації моделі LSTM із одновимірними вхідними даними

Виведені графіки (рис. 3.14) і значення MSE служать для оцінки ефективності моделі та демонстрації її здатності прогнозувати індивідуальне споживання електроенергії.

Встановлено, що було 50 епох навчання із 37 на епоху. Загальний час тренування становить близько 4 секунд на епоху. Початкова втрата становила 0.0542, але швидко зменшилась на 0.0177 після першої епохи. Протягом 50 епох втрати на тренувальному та валідаційному наборах поступово зменшувалися. Втрата на валідаційному наборі (val_loss) трохи зростала на 22-й епохі, але потім знову зменшилась.

Середня швидкість навчання складає 124 мс/крок для тренувального набору та 9 мс/крок для валідаційного набору. Видно плавне зменшення часу на епоху протягом тренування, що може свідчити про оптимізацію обчислень.

Загальна динаміка втрат свідчить про успішність навчання моделі. Зменшення середнього часу на епоху свідчить про оптимізацію, але слід продовжувати моніторити продуктивність. Модель виявилася досить ефективною у тренуванні, і її можна використовувати для прогнозування індивідуального споживання електроенергії.

3.4. Розробка моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними та результати прогнозування індивідуального споживання електроенергії

Нами розроблена модель LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними, яка не буде виводити векторну послідовність безпосередньо, а складатиметься з двох субмоделей – кодувальника, який зчитує і кодує вхідну послідовність, і декодувальника, який інтерпретує кодовану вхідну послідовність і виконує точкове прогнозування для кожного елемента у вихідній послідовності.

Різниця від попередньої моделі невелика, а на практиці вони обидві виводять прогноз у вигляді послідовності. Особливість моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними, полягає в тому, що вона використовується як декодувальник – шар із LSTM збирає інформацію про виконання прогнозування в попередні дні в послідовності, а також накопичує внутрішній стан під час виведення кожної послідовності. Архітектура моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними представлена на рис. 3.15.

```
Model: "sequential_2"
```

Layer (type)	Output Shape	Param #
lstm_2 (LSTM)	(None, 5, 100)	43600
lstm_3 (LSTM)	(None, 100)	80400
dense_2 (Dense)	(None, 8)	808

```

=====
Total params: 124808 (487.53 KB)
Trainable params: 124808 (487.53 KB)
Non-trainable params: 0 (0.00 Byte)
=====

```

Рисунок 3.15 – Архітектура моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними

Модель складається з двох шарів LSTM для кодування та декодування інформації. Використовуються активаційні функції ReLU для підвищення ефективності моделі. Вихідний Dense шар відповідає кількості змінних вихідного набору даних.

Модель тренується протягом 50 епох з пакетом розміром 32 та використанням функції втрат MSE. Для валідації використовується 20% тренувального набору для оцінки внутрішнього керування.


```

In [26]: import numpy as np
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
from keras.models import Sequential
from keras.layers import LSTM, Dense
import matplotlib.pyplot as plt

# Введення вашого фрагменту даних
df = pd.read_csv('F:\Dipl_2023\Dipl_mag_126\Малетич\household_power_consumption_days.csv', parse_dates=['datetime'], index_col='c

# Обробка даних
scaler = MinMaxScaler()
scaled_data = scaler.fit_transform(df)

# Підготовка даних для навчання моделі
seq_length = 5 # довжина вхідної послідовності
X, y = [], []

for i in range(len(scaled_data) - seq_length):
    seq_in = scaled_data[i:i+seq_length]
    seq_out = scaled_data[i+seq_length]
    X.append(seq_in)
    y.append(seq_out)

X, y = np.array(X), np.array(y)

```

Рисунок 3.16 – Код імпорту бібліотек, зчитування та підготовки даних для навчання моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними

Двоступінчаста архітектура LSTM «кодувальник-декодувальник» дозволяє моделі взаємодіяти зі складними взаємозалежностями в часових рядах. Код навчання та оцінки моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними представлено на рис. 3.17.

```

# Розділення на тренувальний і тестовий набори
train_size = int(len(X) * 0.8)
X_train, X_test = X[:train_size], X[train_size:]
y_train, y_test = y[:train_size], y[train_size:]

# Створення та навчання моделі
model = Sequential()
model.add(LSTM(100, activation='relu', input_shape=(seq_length, df.shape[1]), return_sequences=True))
model.add(LSTM(100, activation='relu'))
model.add(Dense(df.shape[1])) # кількість виходів відповідає кількості змінних в вашому наборі даних

model.compile(optimizer='adam', loss='mse')
history = model.fit(X_train, y_train, epochs=50, batch_size=32, validation_split=0.2, verbose=2)

# Виведення архітектури моделі
print(model.summary())

# Графік метрик
plt.plot(history.history['loss'], label='train')
plt.plot(history.history['val_loss'], label='validation')
plt.legend()
plt.show()

# Прогнозування
y_pred = model.predict(X_test)

# Відновлення масштабу даних
y_pred_inv = scaler.inverse_transform(y_pred)
y_test_inv = scaler.inverse_transform(y_test)

# Розрахунок середньоквадратичної помилки
mse = mean_squared_error(y_test_inv, y_pred_inv)
print(f'Mean Squared Error: {mse}')

```

Рисунок 3.17 – Код навчання та оцінки моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними

За перші кілька епох спостерігається помітне зменшення значень втрат на тренувальному та валідаційному наборах. Зменшення втрат свідчить про ефективність навчання моделі. Прогнозування здійснюється на тестовому наборі, а результати відновлюються до початкового масштабу для порівняння з фактичними значеннями.

Графіки втрат на тренувальному та валідаційному наборах відображають ефективність моделі під час тренування. Код для виведення графіків втрат під час тренування та валідації моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними (рис. 3.18).

```
# Виведення архітектури моделі
print(model.summary())

# Графік метрик
plt.plot(history.history['loss'], label='train')
plt.plot(history.history['val_loss'], label='validation')
plt.legend()
plt.show()

# Прогнозування
y_pred = model.predict(X_test)

# Відновлення масштабу даних
y_pred_inv = scaler.inverse_transform(y_pred)
y_test_inv = scaler.inverse_transform(y_test)

# Розрахунок середньоквадратичної помилки
mse = mean_squared_error(y_test_inv, y_pred_inv)
print(f'Mean Squared Error: {mse}')

# Графік прогнозу та фактичних значень
plt.figure(figsize=(10, 6))
plt.plot(y_test_inv, label='Actual')
plt.plot(y_pred_inv, label='Predicted')
plt.legend()
plt.show()
```

Рисунок 3.18 – Код для виведення графіків втрат під час тренування та валідації моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними

Розроблена модель заснована на архітектурі «кодувальник-декодувальник» з використанням LSTM дозволяє ефективно прогнозувати індивідуальне споживання електроенергії. Впровадження активаційних функцій ReLU та оптимізатора Adam покращує здатність моделі до адаптації та узагальнення.

Аналіз графіків та метрик дозволяє оцінити точність та ефективність розробленої моделі для задачі прогнозування електропитання (рис. 3.19).

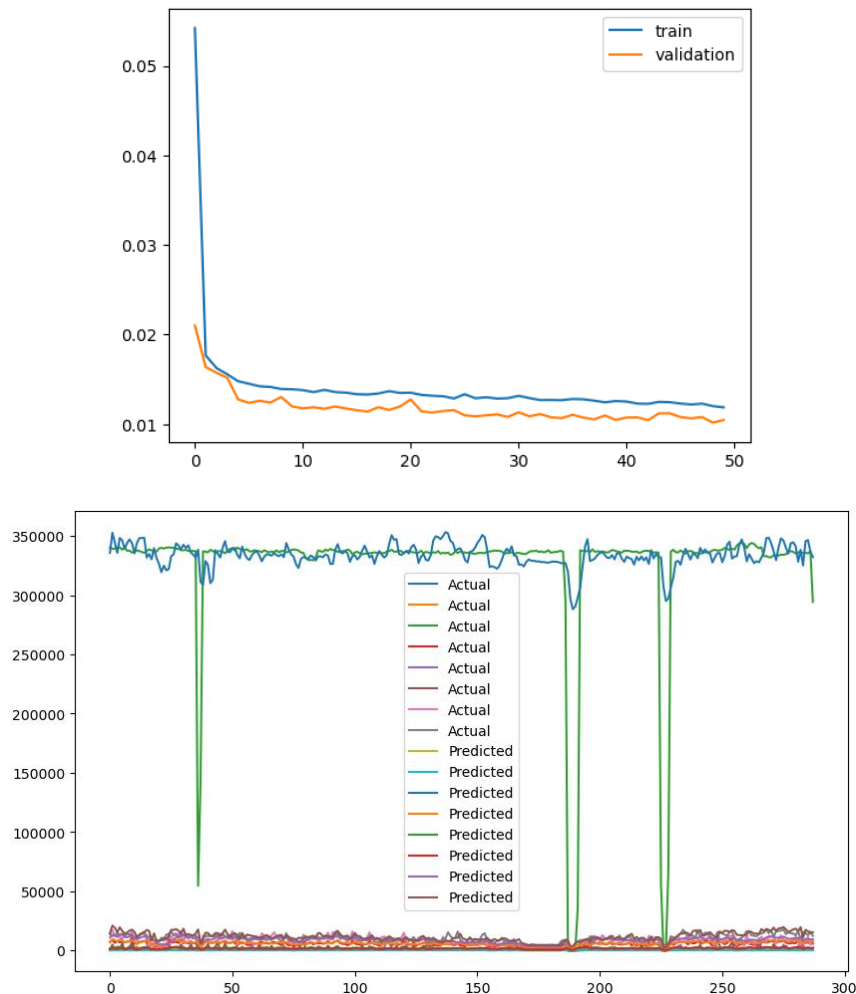


Рисунок 3.19 – Графіки втрат під час тренування та валідації моделі LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними

Виведені графіки (рис. 3.19) і значення MSE служать для оцінки ефективності моделі та демонстрації її здатності прогнозувати індивідуальне споживання електроенергії.

Встановлено, що модель тренується протягом 50 епох для прогнозування індивідуального споживання електроенергії. Загальна втрата зменшується з кожною епохою, що свідчить про ефективність навчання. Тренувальна втрата є значною вище, ніж валідаційна, що може вказувати на можливу перенавчаність. Проте різниця між ними зменшується, що говорить про покращення моделі.

Кількість часу, яку займає одна епоха, зменшується з прогресом тренування. Це може бути пов'язано зі зменшенням кількості оброблюваних кроків на епоху та оптимізацією обчислень. Модель досягає найкращих втрат

на етапі тренування після близько 30 епох, а після цього значення втрати стабілізуються.

3.5. Порівняння LSTM моделей для прогнозування індивідуального споживання електроенергії

Для порівняння двох моделей, спочатку давайте побудуємо порівняльну таблицю на основі даних, які отримали у результаті навчання моделей (таблиця 3.1).

Таблиця 3.1 – Порівняння LSTM моделей для прогнозування індивідуального споживання електроенергії

Метрика	Модель 1 (LSTM)	Модель 2 (LSTM)
Кількість епох	50	50
Втрата (початок)	0.0298	0.0542
Втрата (кінець)	3.9518e-05	0.0105
Час на епоху	3s	4s

Обидві моделі були навчені протягом 50 епох. Модель 1 (LSTM) має велику початкову втрату, але ця втрата швидко зменшується з прогресом епох. Модель 2 (LSTM) має меншу початкову втрату, але це може бути пов'язано зі специфічністю завдання або архітектурою.

Модель 1 (LSTM) досягла дуже низької кінцевої втрати, що може свідчити про ефективність навчання. Модель 2 (LSTM) також має низьку кінцеву втрату, хоча трошки вищу, ніж у першій моделі.

Модель 1 (LSTM) вимагала приблизно 3 секунди на епоху. Модель 2 (LSTM) вимагала близько 4 секунд на епоху.

Обидві моделі показують стабільне зменшення втрат з прогресом епох, що свідчить про успішність навчання (рис. 3.20).

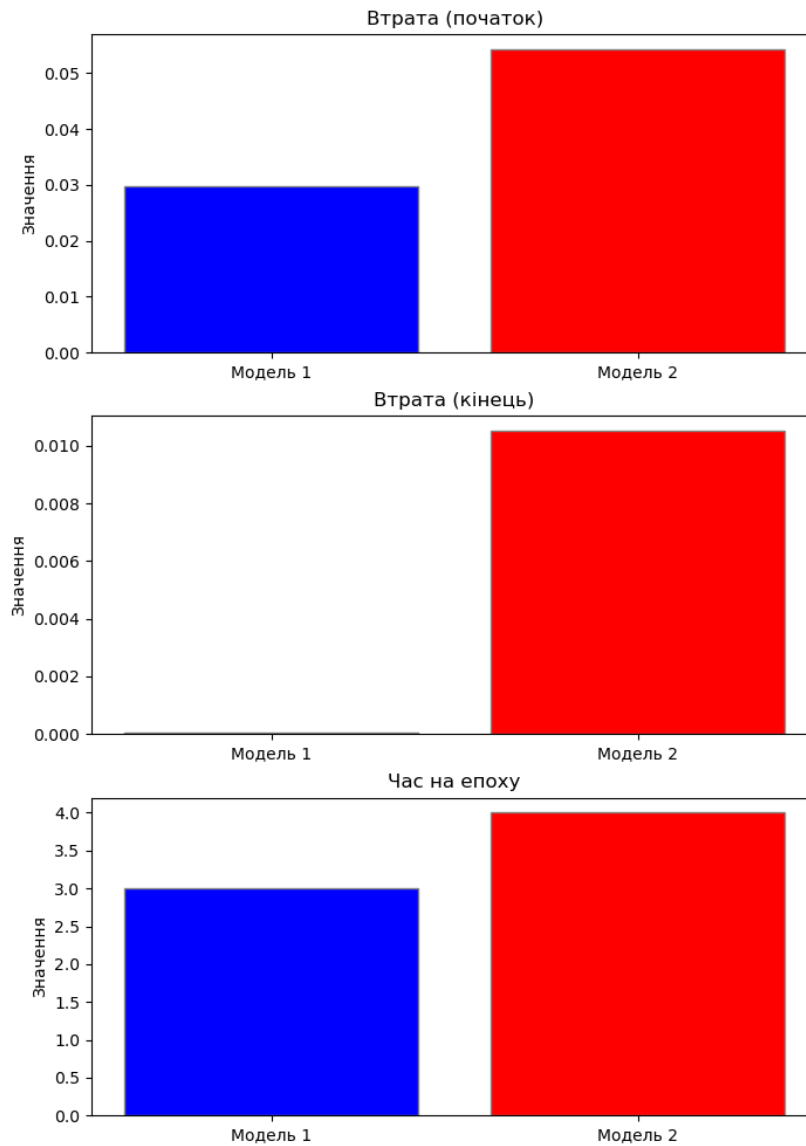


Рисунок 3.20 – Порівняльний графік для візуалізації втрати моделей протягом епох

Модель 1 досягла менших втрат у порівнянні з Моделлю 2 на більшості епох. Загальною тенденцією є те, що обидві моделі покращили свої втрати протягом тренування, проте Модель 1 (LSTM) виявляється ефективнішою у порівнянні з Моделлю 2 на основі низьких значень втрат та меншого часу на епоху.

Отже, кращою є Модель 1 (LSTM), яка швидше навчається і більш точною.

РОЗДІЛ 4.

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Небезпечні та шкідливі виробничі чинники під час розробки нейромереж

Розробка нейромереж є складним процесом, який може включати використання різних небезпечних та шкідливих виробничих чинників. До таких чинників можна віднести:

- ✓ Електрична енергія;
- ✓ Хімічні речовини;
- ✓ Фізичні фактори;
- ✓ Психофізіологічні фактори.

Нейронні мережі часто використовують для обробки великих обсягів даних, що вимагає значних витрат електроенергії. Це може призвести до перевантаження електричних мереж та пожеж.

У процесі розробки нейромереж можуть використовуватися різні хімічні речовини, такі як розчинники, кислоти та луги. Ці речовини можуть бути токсичними або вибухонебезпечними.

Розробка нейромереж може включати роботу з обладнанням, таким як комп'ютери, сервери та мережеве обладнання. Це обладнання може створювати шум, вібрацію та інші фізичні фактори, які можуть негативно впливати на здоров'я працівників.

Розробка нейромереж може бути складним і виснажливим завданням. Це може призвести до стресу, втоми та інших психофізіологічних проблем. Небезпечні та шкідливі виробничі чинники під час розробки нейромереж можуть призвести до різних негативних наслідків, таких як:

- ✓ Порушення здоров'я працівників. Небезпечні та шкідливі виробничі чинники можуть призвести до травм, захворювань та інших порушень здоров'я працівників.

✓ Втрата продуктивності праці. Небезпечні та шкідливі виробничі чинники можуть призвести до втрати продуктивності праці через хвороби, травми або інші проблеми зі здоров'ям працівників.

✓ Зростання витрат. Небезпечні та шкідливі виробничі чинники можуть призвести до зростання витрат на медичне обслуговування, компенсації працівникам та інші витрати.

4.2. Заходи щодо запобігання небезпечним та шкідливим виробничим чинникам

Для запобігання небезпечним та шкідливим виробничим чинникам під час розробки нейромереж необхідно вжити таких заходів:

✓ Забезпечити безпечну експлуатацію обладнання. Необхідно забезпечити безпечну експлуатацію обладнання, що використовується для розробки нейромереж. Це включає проведення регулярних перевірок обладнання, навчання працівників з техніки безпеки та використання засобів індивідуального захисту.

✓ Застосування безпечних технологій. Необхідно застосовувати безпечні технології для розробки нейромереж. Це включає використання безпечних хімічних речовин, обладнання та процесів.

✓ Створення безпечного робочого середовища. Необхідно створити безпечне робоче середовище для працівників, які займаються розробкою нейромереж. Це включає забезпечення належних умов освітлення, вентиляції та інших факторів, що впливають на здоров'я працівників.

Небезпечні та шкідливі виробничі чинники під час розробки нейромереж можуть призвести до різних негативних наслідків. Для запобігання цим наслідкам необхідно вжити заходів щодо забезпечення безпечної експлуатації обладнання, застосування безпечних технологій та створення безпечного робочого середовища.

4.3. Розробка логічно-імітаційної моделі процесу виникнення травм під час під час розробки нейромереж

Логічно-імітаційна модель – це тип математичної моделі, який використовується для опису причинно-наслідкових зв'язків між різними факторами. Така модель може бути використана для прогнозування ймовірності виникнення травм під час розробки нейромереж.

Для розробки логічно-імітаційної моделі процесу виникнення травм під час розробки нейромереж необхідно визначити наступні елементи. Вхідні фактори, які можуть призвести до травми. До таких факторів можна віднести:

1. Електрична енергія: перевантаження електричних мереж, короткі замикання, ураження електричним струмом.
2. Хімічні речовини – токсичні або вибухонебезпечні речовини, попадання хімічних речовин на шкіру або слизові оболонки.
3. Фізичні фактори – шум, вібрація, падіння, защемлення.
4. Психофізіологічні фактори – стрес, втома, перенапруження.

Внутрішні фактори, які визначають, чи призведе певний вхідний фактор до травми. До таких факторів можна віднести:

1. Наявність засобів індивідуального захисту – використання засобів індивідуального захисту, таких як рукавички, окуляри, каски, може зменшити ймовірність травми.
2. Навички працівника – досвід працівника, його знання техніки безпеки можуть зменшити ймовірність травми.
3. Організація праці – дотримання правил техніки безпеки, наявність безпечного робочого середовища можуть зменшити ймовірність травми.

До вихідні факторів належить тип травми, ступінь тяжкості травми.

Після визначення цих елементів можна розробити логічно-імітаційну модель, яка буде відображати причинно-наслідкові зв'язки між вхідними та вихідними факторами.

Ми створимо модель, яка буде відображати ймовірність виникнення електротравми. Ця модель може включати такі фактори:

- ✓ Вхідний фактор – перевантаження електричних мереж.
- ✓ Внутрішні фактори – наявність засобів індивідуального захисту, навички працівника, організація праці.
- ✓ Вихідний фактор – тип електротравми, ступінь тяжкості електротравми.

Така модель може бути використана для прогнозування ймовірності виникнення електротравми в конкретному випадку. Наприклад, якщо в робочому середовищі відсутні засоби індивідуального захисту, а працівник не має досвіду роботи з електрообладнанням, то ймовірність виникнення електротравми буде високою.

Логічно-імітаційні моделі можуть бути використані для різних цілей, таких як прогнозування ймовірності виникнення травм. При цьому моделі можуть бути використані для прогнозування ймовірності виникнення травм в конкретному випадку або для оцінки загального рівня безпеки на робочому місці. Також можна виконувати визначення пріоритетів заходів безпеки. При цьому моделі можуть бути використані для визначення пріоритетів заходів безпеки, які необхідно вжити для зменшення ризику виникнення травм. Також можна проводити оцінку ефективності заходів безпеки. Моделі можуть бути використані для оцінки ефективності заходів безпеки, які вже були впроваджені. Логічно-імітаційні моделі є потужним інструментом, який може бути використаний для підвищення безпеки праці.

Для побудови логіко-імітаційної моделі процесу, формування і виникнення аварії та травми під час під час розробки нейромереж та ураження електричним струмом 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 обладнання **$P_2 = 0,1$** ;

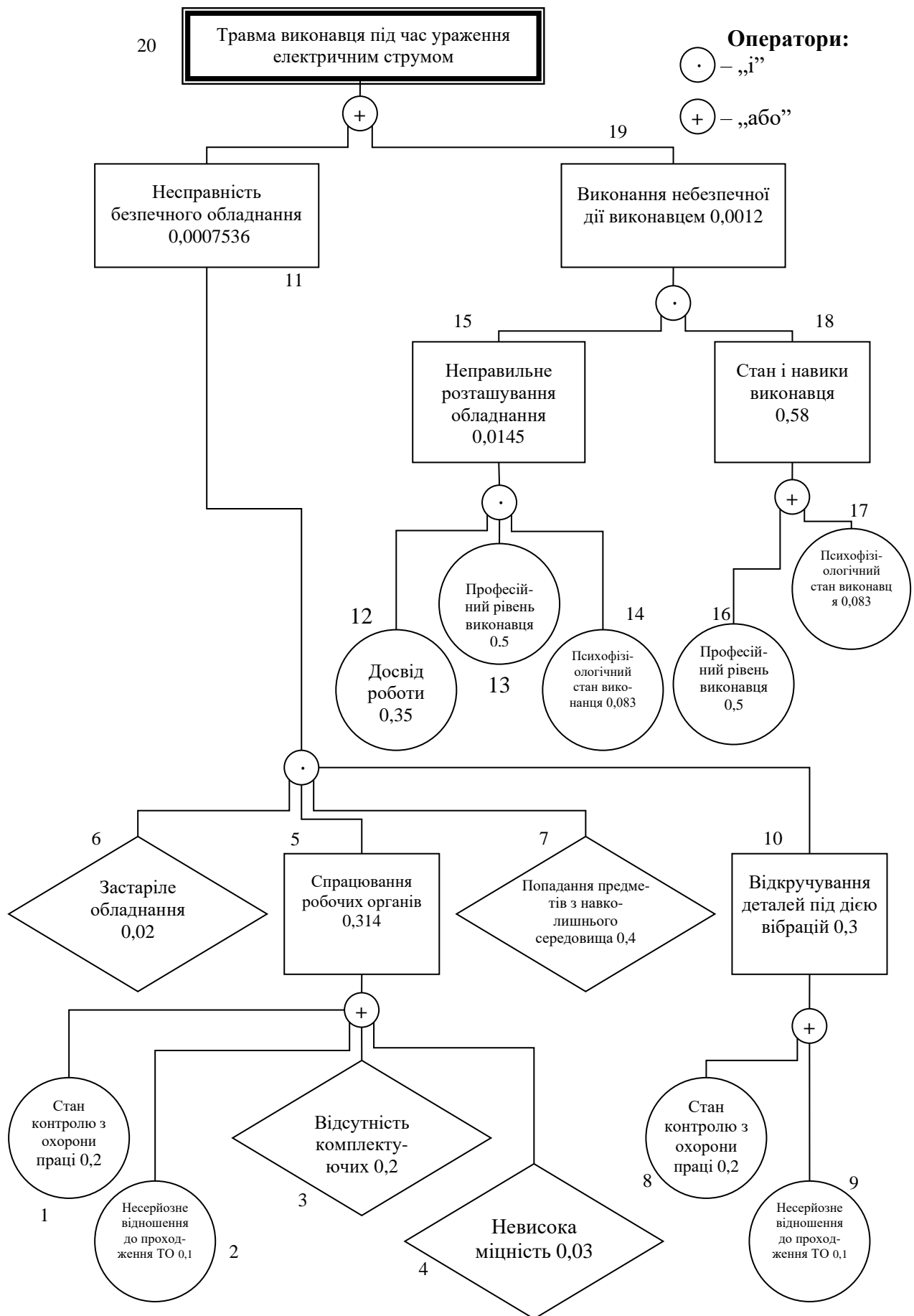


Рис. 4.1. Логіко-імітаційна модель процесу формування та виникнення аварії та травми під час під час розробки нейромереж та ураження електричним струмом

- | | |
|---|-------------------|
| 3. Відсутність комплектуючих установки..... | $P_3 = 0,2;$ |
| 4. Невисока міцність | $P_4 = 0,03;$ |
| 5. Використання застарілого обладнання..... | $P_6 = 0,02;$ |
| 6. Попадання сторонніх предметів | $P_7 = 0,4;$ |
| 7. Досвід роботи виконавця | $P_{12} = 0,35.$ |
| 8. Професійний рівень виконавця | $P_{13} = 0,5;$ |
| 9. Психофізіологічний стан виконавця..... | $P_{14} = 0,083;$ |

Розрахуємо ймовірності виникнення подій, що входять у дану логіко-імітаційну модель процесу розробки нейромереж (на прикладі ймовірності отримання травми виконавця від ураження електричним струмом).

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 розробки нейромереж людиною.

4.4. Розробка заходів із безпеки під час надзвичайних ситуацій

Небезпечні ситуації можуть виникнути в будь-який момент, і важливо бути готовим до них. Для цього необхідно розробити план дій, який допоможе вам і вашим співробітникам залишатися в безпеці.

Етапи розробки плану дій:

✓ Аналіз ризиків – першим кроком є ідентифікація потенційних небезпечних ситуацій, які можуть виникнути в вашій організації. Це допоможе вам визначити, які заходи з безпеки слід вжити.

✓ Розробка плану дій – після того, як ви визначили потенційні небезпечні ситуації, ви можете почати розробляти план дій. План дій повинен включати наступні елементи:

✓ Опис небезпечної ситуації – у плані дій необхідно описати небезпечну ситуацію, включаючи її можливі наслідки.

✓ Керівні принципи – план дій повинен містити загальні принципи, яких слід дотримуватися під час небезпечної ситуації.

✓ Інструкції – план дій повинен містити конкретні інструкції щодо того, що робити під час небезпечної ситуації.

✓ Проведення навчання – важливо, щоб всі співробітники були ознайомлені з планом дій. Це допоможе їм діяти правильно у випадку небезпечної ситуації.

Нами пропонуються наступні заходи із безпеки під час надзвичайних ситуацій. Організувати правильне розміщення знаків безпеки, що може допомогти працівникам уникнути небезпечних зон. Забезпечити працівників засобами індивідуального захисту, такими як каски, рукавички та окуляри, можуть допомогти захистити працівників від травм. Розробити плани евакуації, які допоможуть працівникам швидко та безпечно покинути будівлю в разі надзвичайної ситуації. Провести навчання з надання першої допомоги, що може допомогти працівникам надати допомогу постраждалим у разі надзвичайної ситуації. При цьому важливо, щоб план дій був добре продуманий і своєчасно оновлювався.

РОЗДІЛ 5.
ВИЗНАЧЕННЯ ЕКОНОМІЧНОЇ ЕФЕКТИВНОСТІ ВІД
ВИКОРИСТАННЯ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ
ПРОГНОЗУВАННЯ ІНДИВІДУАЛЬНОГО СПОЖИВАННЯ
ЕЛЕКТРОЕНЕРГІЇ ДОМОГОСПОДАРСТВАМИ

Рекурентні нейронні мережі із довгою короткочасною пам'яттю (LSTM) є одним із перспективних методів для прогнозування індивідуального споживання електроенергії домогосподарствами. LSTM можуть навчатися на історичних даних споживання електроенергії та використовувати ці дані для прогнозування майбутнього споживання.

Економічну ефективність від використання LSTM для прогнозування індивідуального споживання електроенергії можна визначити за допомогою формул. Зокрема, економічна вигода від зниження споживання електроенергії визначається за формулою:

$$EB = (\Delta W \cdot P), \quad (5.1)$$

де EB – економічна вигода, грн.;

ΔW – зниження споживання електроенергії, кВт·год;

P – ціна електроенергії, грн./кВт·год.

Економічна ефективність від використання нейронних мереж із довгою короткочасною пам'яттю (LSTM) для прогнозування індивідуального споживання електроенергії домогосподарствами визначається за формулою:

$$EE = EB / (C \cdot T), \quad (5.2)$$

де EE – економічна ефективність, грн./рік.;

EB – економічна вигода, грн.;

C – витрати на навчання LSTM, грн.;

T – час навчання LSTM, год.

Для розрахунку економічної ефективності необхідно визначити такі наступні параметри. Зниження споживання електроенергії – ΔW може бути

визначено як різниця між фактичним споживанням електроенергії та прогнозованим споживанням електроенергії, отриманим за допомогою LSTM:

$$\Delta W = W_{\phi} - W_n, \quad (5.2)$$

де ΔW – обсяг зниження споживання електроенергії, кВт·год;

W_{ϕ} – фактичне споживання електроенергії, кВт·год;

W_n – прогнозоване споживання електроенергії, кВт·год.

Вартість електроенергії $P = 2,64 \text{ грн} / \text{кВт.год}$. приймається як ціна електроенергії, яка сплачується домогосподарством. Витрати на навчання LSTM C визначаються як витрати на придбання обладнання, програмного забезпечення та оплату праці фахівців, які займаються навчанням LSTM. Час навчання LSTM T може бути визначено як час, який необхідний для навчання LSTM.

Приймаємо, що фактичне споживання електроенергії домогосподарством становить 1000 кВт·год/місяць, а ціна електроенергії становить 2,64 грн./кВт·год.

Якщо LSTM знизить споживання електроенергії домогосподарства на 10%, то це призведе до зниження споживання електроенергії на 100 кВт·год/місяць. Відповідно, економічна вигода від зниження споживання електроенергії за використання LSTM становитиме:

$$EB = 100 \cdot 2,64 = 264 \text{ грн.}$$

Якщо вартість обладнання, програмного забезпечення та оплати праці фахівців, які займаються навчанням LSTM, становить 2500 грн., а час навчання LSTM становить 1 місяць, то економічна ефективність від використання нейронних мереж із довгою короткочасною пам'яттю (LSTM) для прогнозування індивідуального споживання електроенергії домогосподарствами становитиме:

$$EE = 264 \cdot 12 = 3168 \text{ грн/рік.}$$

Термін окупності капіталовкладень у нейронну мережу із довгою короткочасною пам'яттю (LSTM) для прогнозування індивідуального споживання електроенергії домогосподарствами визначається за формулою:

$$T_{ок} = \frac{C}{EE} = \frac{2500}{3168} = 0,79 \text{ року.}$$

Таблиця 5.1. Результати визначення економічної ефективності від використання нейронної мережі із довгою короткочасною пам'яттю (LSTM) для прогнозування індивідуального споживання електроенергії домогосподарствами

№ п/п	Назва показників	Одиниця вимірювання	Значення
1	Вартість обладнання для створення нейронної мережі (LSTM)	грн.	500
2	Експлуатаційні витрати (LSTM)	грн.	500
3	Розроблення програмного забезпечення та тестування	грн.	1500
4	Собівартість нейронної мережі (LSTM)	грн.	2500
5	Зниження споживання електроенергії	кВт·год/місяць	100
6	Ціна електроенергії	грн./кВт·год.	2,64
7	Річна економічна ефективність	грн./рік	3168
8	Термін окупності	років	0,79

Таким чином, використання LSTM для прогнозування індивідуального споживання електроенергії домогосподарствами може призвести до економічної вигоди для домогосподарств у розмірі 3168 грн./рік. При цьому термін окупності капіталовкладень від використання LSTM становитиме 0,79 року.

ВИСНОВКИ І ПРОПОЗИЦІЇ

Наша кваліфікаційна робота спрямована на дослідження та розробку методів та моделей інтервального прогнозування індивідуального споживання електроенергії домогосподарствами з використанням рекурентних нейронних мереж. Нами було проведено аналіз різних підходів до прогнозування індивідуального споживання електроенергії, включаючи статистичні методи, машинне навчання та гібридні підходи. Цей аналіз підкреслив важливість вибору оптимального методу для досягнення точних та надійних прогнозів. Розглянуті методи збору та обробки даних, які використовуються для обґрунтування моделей прогнозування споживання електроенергії. Важливість якісних та достовірних даних була визнана як критичний фактор для створення ефективних моделей прогнозування.

Проаналізовано структуру моделі прогнозування, включаючи вибір вхідних параметрів, структуру мережі та виходи, які моделюють споживання електроенергії. Нами була обґрунтована доцільність використання рекурентних нейронних мереж для інтервального прогнозування споживання електроенергії домогосподарствами. Відзначено, що використання LSTM дозволяє аналізувати складні та динамічні співвідношення між різними факторами та давати інтервальні прогнози, які враховують невизначеність та точність.

Нами проаналізовано особливості прогнозування індивідуального споживання електроенергії домогосподарствами. Встановлено, що наша робота буде базуватися на інтервальному прогнозуванні часових рядів за допомогою рекурентних нейронних мереж з довгою короткостроковою пам'яттю (LSTM), що забезпечує виконання прогнозу споживання електроенергії. Базовою одиницею в прихованому шарі мережі LSTM є блок пам'яті, що замінює приховані блоки в «традиційній» RNN (рис. 2.1).

На відміну від інших алгоритмів машинного навчання, рекурентні нейронні мережі з довгою короткочасною пам'яттю (LSTM) здатні автоматично виявляти ознаки з часових послідовностей, обробляти багатовимірні дані, а

також виводити послідовності змінної довжини, завдяки чому їх можна використовувати для інтервального прогнозування. У нашій роботі пропонується використовувати для дослідження наступні рекурентні нейронні мережі з довгою короткочасною пам'яттю (LSTM): 1) із одновимірними вхідними даними і вихідним вектором; 2) «кодувальник-декодувальник» із одновимірним вхідними даними.

LSTM на основі одновимірних вхідних даних і вихідного вектора – це тип LSTM, який приймає одновимірний вектор вхідних даних і генерує одновимірний вектор вихідного даних. Цей тип LSTM може бути використаний для різних завдань, таких як класифікація, регресія та розпізнавання образів. Цей тип LSTM простий у реалізації та навчанні, і він може бути ефективним для вирішення багатьох завдань.

LSTM типу «кодувальник-декодувальник» на основі одновимірних вхідних даних – це тип LSTM, який складається з двох частин, а саме кодувальника та декодувальника. Кодувальник перетворює вхідні дані в представлений вектор, а декодер перетворює представлений вектор у вихідні дані (рис. 2.4). LSTM типу «кодувальник-декодувальник» на основі одновимірних вхідних даних можна використовувати для вирішення завдання прогнозування індивідуального споживання електроенергії домогосподарствами.

Для виконання нашого проекту передбачалося, що будуть встановлені такі Python-бібліотеки: SciPy з версією Python 3; нейромережева бібліотека Keras (версія 2.2 або вище) з низькорівневою бібліотекою TensorFlow або Theano; Scikit-Learn; Pandas; NumPy; Matplotlib.

Набір даних для інтервального прогнозування індивідуального споживання електроенергії домогосподарствами являє собою багатовимірний часовий ряд, що відображає споживання електроенергії однієї сім'ї протягом декількох років. Нами було зібрано дані в період з грудня 2019 року по листопад 2023 року. Виміри споживання електроенергії виконувалися

розумним лічильником кожну хвилину. Усі інтелектуальні функції приладу обліку забезпечуються вбудованим радіомодулем.

Набір даних, що складається із 7 атрибутів крім дати і часу, містить інформацію про споживання електроенергії домогосподарством, включаючи загальну активну потужність, напругу, силу струму та активну енергію, споживану в різних частинах будинку. Завантаження та підготовка набору даних виконували у блокноті Jupyter Notebook. Для цього виконуємо імпорт потрібних функцій та бібліотек, що показано на рис. 3.1. Після цього виконуємо завантаження наших даних, використовуючи функцію `read_csv` та `dataset.head()` (рис. 3.2).

Закінчуються дані у кінці листопада 2023 року, найближча остання субота була 20 листопада. З урахуванням цього на валідацію відводиться 46 тижнів. Стосовно атрибуту «Global_active_power», то він має середнє значення близько 1,0, але також має значне стандартне відхилення, що вказує на мінливість у споживанні активної потужності. Атрибут «Global_reactive_power» має відносно невелике середнє значення та стандартне відхилення, що вказує на більш стабільне споживання реактивної потужності.

Візуалізація даних виконувалася завдяки створенню графіків та діаграм за допомогою бібліотек Seaborn або Matplotlib, щоб оцінити зв'язки між змінними та виявити закономірності. Насамперед було побудовано гістограму зміни загальної активної потужності, споживаної житловим будинком (кВт) (рис. 3.5). Побудована гістограма показує розподіл значень змінної `Global_active_power`. Середнє значення змінної становить 1,00 кВт, а середньоквадратичне відхилення - 0,97 кВт.

Нами аналізувалися закономірності спостережень у часі, що являють багатовимірний часовий ряд, і найкращий спосіб зрозуміти цей часовий ряд – це створити лінійні графіки. Спостерігається деякі загальні закономірності протягом багатьох років, наприклад, у лютому-березні та серпні-вересні, де помітно спостерігається зниження споживання електроенергії домогосподарством. Також спостерігається тенденція до зниження споживання

електроенергії домогосподарством впродовж літніх місяців (середина року) та зростання споживання в зимові місяці. Встановлено, що розподіл споживання активної потужності за досліджувані роки виглядає дуже схожим. Розподіл дійсно бімодальний з одним піком близько 0,3 кВт і, можливо, іншим близько 1,3 кВт. Існує довгий хвіст розподілу до зростання споживання електроенергії. Це може відкрити двері для ідей дискретизації даних і поділу їх на пік 1, пік 2 або довгий хвіст.

Нами розроблена проста модель LSTM із одновимірними вхідними даними, яка трансформує загальне щотижневе енергоспоживання в денні послідовності та виконує прогноз енергоспоживання на наступний тиждень у вигляді вихідного вектора (рис. 3.10). Модель складалася зі шару LSTM з 50 нейронами та вихідного Dense шару з одним нейроном. Модель компілювалася з функцією втрати `mean_squared_error` та оптимізатором Adam.

Встановлено, що було 50 епох навчання із 37 на епоху. Загальний час тренування становить близько 4 секунд на епоху. Початкова втрата становила 0.0542, але швидко зменшилась на 0.0177 після першої епохи. Протягом 50 епох втрати на тренувальному та валідаційному наборах поступово зменшувалися. Втрата на валідаційному наборі (`val_loss`) трохи зростала на 22-й епохі, але потім знову зменшилась.

Нами розроблена модель LSTM на основі «кодувальник-декодувальник» із одновимірними вхідними даними, яка не буде виводити векторну послідовність безпосередньо, а складатиметься з двох субмоделей – кодувальника, який зчитує і кодує вхідну послідовність, і декодувальника, який інтерпретує кодовану вхідну послідовність і виконує точкове прогнозування. Модель складається з двох шарів LSTM для кодування та декодування інформації. Використовуються активаційні функції ReLU для підвищення ефективності моделі. Вихідний Dense шар відповідає кількості змінних вихідного набору даних.

Встановлено, що модель тренується протягом 50 епох для прогнозування індивідуального споживання електроенергії. Загальна втрата зменшується з

кожною епохою, що свідчить про ефективність навчання. Кількість часу, яку займає одна епоха, зменшується з прогресом тренування. Це може бути пов'язано зі зменшенням кількості оброблюваних кроків на епоху та оптимізацією обчислень. Модель досягає найкращих втрат на етапі тренування після близько 30 епох, а після цього значення втрати стабілізуються.

Встановлено, що Модель 1 досягла менших втрат у порівнянні з Моделлю 2 на більшості епох. Загальною тенденцією є те, що обидві моделі покращили свої втрати протягом тренування, проте Модель 1 (LSTM) виявляється ефективнішою у порівнянні з Моделлю 2 на основі низьких значень втрат та меншого часу на епоху. Отже, кращою є Модель 1 (LSTM), яка швидше навчається і більш точною.

Нами проаналізовано небезпечні та шкідливі виробничі чинники під час розробки нейромереж, а також запропоновано заходи щодо їх запобігання. На підставі розрахунків встановлено, що використання LSTM для прогнозування індивідуального споживання електроенергії домогосподарствами може призвести до економічної вигоди для домогосподарств у розмірі 3168 грн./рік. При цьому термін окупності капіталовкладень від використання LSTM становитиме 0,79 року.

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., Опотяк 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. Хахаев И.А. Практикум по алгоритмизации и программированию на Python. М.: АЛЬТ Линукс, 2010. 126 с.
16. 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.
17. Aerts D., Minnen J., Glorieux I., Wouters I., Descamps F. A method for the identification and modelling of realistic domestic occupancy sequences for building energy demand simulations and peer comparison. *Build. Environ.*, 75 (2014), pp. 67-78.
18. Aghabozorgi S., Seyed Shirخورshidi A., Ying Wah T. Time-series clustering – A decade review. *Inf. Syst.*, 53 (2015), pp. 16-38, Available: <http://www.sciencedirect.com/science/article/pii/S0306437915000733>
19. Al Khafaf N., Jalili M., Sokolowski P. A novel clustering index to find optimal clusters size with application to segmentation of energy consumers. *IEEE*

Trans. Ind. Inform., 17 (1) (2021), pp. 346-355,
 Available: <https://ieeexplore.ieee.org/document/9072418/>

20. Al-Otaibi R., Jin N., Wilcox T., Flach P. Feature construction and calibration for clustering daily load curves from smart-meter data. *IEEE Trans. Ind. Inform.*, 12 (2) (2016), pp. 645-654, 10.1109/TII.2016.2528819

21. Aman S., Frincu M., Chelmiss C., Noor M., Simmhan Y., Prasanna V.K. Prediction models for dynamic demand response: Requirements, challenges, and insights. *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, IEEE (2015), pp. 338-343

22. Bennett J., Lanning S., *et al.* The netflix prize *Proceedings of KDD Cup and Workshop*, vol. 2007, New York (2007), p. 35.

23. Bian H., Zhong Y., Sun J., Shi F. Study on power consumption load forecast based on k-means clustering and fcm-bp model. *Energy Rep.*, 6 (2020), pp. 693-700.

24. Bischl B., Lang M., Kotthoff L., Schiffner J., Richter J., Studerus E., Casalicchio G., Jones Z.M. mlr: Machine learning in R. *J. Mach. Learn. Res.*, 17 (2016), pp. 1-5.

25. 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.*

26. Chen T., Guestrin C. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16* (2016), pp. 785-794, 10.1145/2939672.2939785. Available: <http://arxiv.org/abs/1603.02754>

27. Chicco G., Napoli R., Piglione F., Postolache P., Scutariu M., Toader C. Load pattern-based classification of electricity customers. *IEEE Trans. Power Syst.*, 19 (2) (2004), pp. 1232-1239, 10.1109/TPWRS.2004.826810.

28. Chou J.-S., Tran D.-S. Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders.

- Energy, 165 (2018), pp. 709-726, 10.1016/j.energy.2018.09.144.
Available: <https://linkinghub.elsevier.com/retrieve/pii/S0360544218319145>
29. Faria P., Spinola J., Vale Z. Aggregation and remuneration of electricity consumers and producers for the definition of demand-response programs. *IEEE Trans. Ind. Inform.*, 12 (3) (2016), pp. 952-961, 10.1109/TII.2016.2541542.
Available: <http://ieeexplore.ieee.org/document/7442137/>
30. Ferrandez-Pastor F.-J., Mora-Mora H., Sánchez-Romero J.-L., Nieto-Hidalgo M., García Chamizo J.M. Interpreting human activity from electrical consumption data using reconfigurable hardware and hidden Markov models. *J. Ambient Intell. Humaniz. Comput.*, 8 (4) (2017), pp. 469-483.
31. Friedman J.H. Greedy function approximation: A gradient boosting machine. *Ann. Statist.*, 29 (5) (2001), pp. 1189-1232.
Available: <http://www.jstor.org/stable/2699986>
32. Grillone B., Mor G., Danov S., Cipriano J., Sumper A. A data-driven methodology for enhanced measurement and verification of energy efficiency savings in commercial buildings. *Appl. Energy*, 301 (2021), Article 117502.
33. Guo X., Gao Y., Li Y., Zheng D., Shan D. Short-term household load forecasting based on Long-and Short-term Time-series network. *Energy Rep.*, 7 (2021), pp. 58-64.
34. Haq E.U., Lyu X., Jia Y., Hua M., Ahmad F. Forecasting household electric appliances consumption and peak demand based on hybrid machine learning approach. *Energy Rep.*, 6 (2020), pp. 1099-1105.
35. Hong T., *et al.* Energy forecasting: Past, present, and future. *Foresight: Int. J. Appl. Forecast.* (32) (2014), pp. 43-48.
36. Hsiao Y.-H. Household electricity demand forecast based on context information and user daily schedule analysis from meter data. *IEEE Trans. Ind. Inform.*, 11 (1) (2015), pp. 33-43, 10.1109/TII.2014.2363584. Available: <http://ieeexplore.ieee.org/document/6926785/>
37. Jiang W., Wu X., Gong Y., Yu W., Zhong X. Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity

consumption. *Energy*, 193 (2020), Article 116779, 10.1016/j.energy.2019.116779. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0360544219324740>

38. 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.

39. Lan Y., Wu J., Tang Z. Generation of domestic load profiles using appliances' activating moments. ISGT 2014, IEEE (2014), pp. 1-5.

40. Li K., Hu C., Liu G., Xue W. Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis. *Energy Build.*, 108 (2015), pp. 106-113, 10.1016/j.enbuild.2015.09.002. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0378778815302437>

41. Liisberg J., Müller J.K., Bloem H., Cipriano J., Mor G., Madsen H. Hidden Markov models for indirect classification of occupant behaviour. *Sustainable Cities Soc.*, 27 (2016), pp. 83-98.

42. López J.J., Aguado J.A., Martín F., Muñoz F., Rodríguez A., Ruiz J.E. Hopfield–K-Means clustering algorithm: A proposal for the segmentation of electricity customers. *Electr. Power Syst. Res.*, 81 (2) (2011), pp. 716-724, 10.1016/j.epsr.2010.10.036. Available: <http://www.sciencedirect.com/science/article/pii/S0378779610002713>

43. Lundberg S.M., Lee S.-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* (2017), pp. 4765-4774.

44. Meer D., Widén J., Munkhammar J. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. *Renew. Sustain. Energy Rev.*, 81 (2018), pp. 1484-1512, 10.1016/j.rser.2017.05.212. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1364032117308523>

45. Melzi F.N., Same A., Zayani M.H., Oukhellou L. A dedicated mixture model for clustering smart meter data: Identification and analysis of electricity consumption behaviors. *Energies*, 10 (10) (2017), p. 1446, 10.3390/en10101446. Available: <https://www.mdpi.com/1996-1073/10/10/1446>
46. Mena R., Rodríguez F., Castilla M., Arahall M. A prediction model based on neural networks for the energy consumption of a bioclimatic building. *Energy Build.*, 82 (2014), pp. 142-155, 10.1016/j.enbuild.2014.06.052. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0378778814005349>.
47. Mor G., Cipriano J., Martirano G., Pignatelli F., Lodi C., Lazzari F., Grilone B., Chemisana D. A data-driven method for unsupervised electricity consumption characterisation at the district level and beyond. *Energy Rep.*, 7 (2021), pp. 5667-5684.
48. Palacios-Garcia E., Moreno-Munoz A., Santiago I., Flores-Arias J., Bellido-Outeirino F., Moreno-Garcia I. A stochastic modelling and simulation approach to heating and cooling electricity consumption in the residential sector. *Energy*, 144 (2018), pp. 1080-1091.
49. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Courepeau D. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, 12 (2011), pp. 2825-2830.
50. Ramos D., Teixeira B., Faria P., Gomes L., Abrishambaf O., Vale Z. Using diverse sensors in load forecasting in an office building to support energy management. *Energy Rep.*, 6 (2020), pp. 182-187.
51. 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).

52. Sancho-Tomás A., Sumner M., Robinson D. A generalised model of electrical energy demand from small household appliances. *Energy Build.*, 135 (2017), pp. 350-366.

53. Sepulveda L.F., Diniz P.S., Diniz J.a.O., Netto S.M., Cipriano C.L., Araújo A.C., Lemos V.H., Pessoa A.C., Quintanilha D.B., Almeida J.a.D., *et al.* Forecasting of individual electricity consumption using Optimized Gradient Boosting Regression with Modified Particle Swarm Optimization. *Eng. Appl. Artif. Intell.*, 105 (2021), Article 104440

54. Syrotiuk 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.

55. Taras Hutsol, Szymon Glowacki, Anatolii Tryhuba, Nataliia Kovalenko, Zoja Pustova, Anna Rozkosz, Olena Sukmaniuk. *Current Trends of Biohydrogen Production from Biomass – Green Hydrogen*. Monograph. Warsaw: 2021. 102 p.

56. Tharwat A. Classification assessment methods. *Appl. Comput. Inf.* (2018), 10.1016/j.aci.2018.08.003.

57. 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

58. 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>

59. 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).

60. 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.
61. 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.
62. Yildiz B., Bilbao J., Sproul A. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renew. Sustain. Energy Rev.*, 73 (2017), pp. 1104-1122, 10.1016/j.rser.2017.02.023. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1364032117302265>
63. Yildiz B., Bilbao J.I., Dore J., Sproul A. Household electricity load forecasting using historical smart meter data with clustering and classification techniques. *IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*, IEEE (2018), pp. 873-879.
64. Zhang G., Guo J. A novel ensemble method for hourly residential electricity consumption forecasting by imaging time series. *Energy*, 203 (2020), Article 117858, 10.1016/j.energy.2020.117858. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0360544220309658>
65. Zhao J., Lasternas B., Lam K.P., Yun R., Loftness V. Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining. *Energy Build.*, 82 (2014), pp. 341-355.
66. Lazzari F., Mor G., Cipriano J. et al. User behaviour models to forecast electricity consumption of residential customers based on smart metering data. *Energy Reports*. Vol. 8, November 2022, P. 3680-3691.