

УДК 004.8, 004.94, 621.31

DOI: <https://doi.org/10.36994/2788-5518-2024-01-07-05>

ВИКОРИСТАННЯ МАШИННОГО НАВЧАННЯ ТА МЕРЕЖЕВИХ НАБОРІВ ДАНИХ ДЛЯ МОДЕЛЮВАННЯ ЕНЕРГОСИСТЕМ

Войтех Д. В., Відкритий міжнародний університет розвитку людини «Україна», Київ, Україна. <https://orcid.org/0009-0003-8997-5495>, d.voitekh@gmail.com

Тимошенко А. Г., к.т.н., доц., Відкритий міжнародний університет розвитку людини «Україна», Київ, Україна. <https://orcid.org/0000-0003-0954-3186>, timoshag@i.ua

Анотація. Світова енергетика останнім часом зазнає суттєвих змін, у тому числі завдяки тенденціям переходу на відновлювані джерела електроенергії, потребам у більш розподіленій генерації та підвищенні надійності електромереж. Ця стаття презентує дослідження основних завдань у галузі моделювання енергомереж, для вирішення яких доцільне використання методів машинного навчання. У ній розглядається, як машинне навчання може сприяти оптимізації функціонування мереж і відігравати ключову роль у модернізації енергетики та забезпеченні більш стабільного процесу переходу до децентралізованої архітектури енергосистеми.

Стаття охоплює широкий спектр питань, пов'язаних із застосуванням машинного навчання в енергосистемах, включно з аналізом задач оптимального розподілу потужності, оцінкою стану системи, прогнозуванням попиту та генерації тощо. Також досліджуються різні типи даних, що використовуються для навчання моделей, і сформуовано практичні поради щодо їх обробки. Наведено сценарії застосування й категоризацію різних методів машинного навчання, такі як навчання з учителем, без учителя та навчання з підкріплення. Крім того, розглядаються гібридні підходи, які поєднують машинне навчання з класичними фізичними моделями, забезпечуючи при цьому більшу ефективність і стабільність.

Наостанок у статті наводиться аналіз одного з найбільших мережеских наборів даних від GridKit, у рамках якого проведено моделювання загальноєвропейської високовольтної мережі щодо особливостей її структури та зв'язків. На основі результатів також виділено перспективні напрями майбутніх досліджень із використанням методів машинного навчання, таких як великі мовні моделі та графові нейронні мережі. Загалом ця стаття є багатостороннім аналізом важливих задач моделювання енергетичної галузі, актуальних наборів даних і моделей машинного навчання та може використовуватися для подальшого вдосконалення інтелектуальних енергетичних систем.

Ключові слова: машинне навчання, аналіз мережеских даних, електричні мережі, інтелектуальні енергосистеми, моделювання енергетичних процесів, прогнозування попиту та пропозиції, оптимізація мереж.

APPLICATIONS OF MACHINE LEARNING AND NETWORK DATASETS FOR POWER SYSTEMS MODELING

Dmytro Voitekh, Open International University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0009-0003-8997-5495>, d.voitekh@gmail.com

Anatolii Tymoshenko, Ph.D., Ass. Prof., Open International University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0003-0954-3186>, timoshag@i.ua

Abstract. Global energy has undergone significant changes in recent decades, including trends toward the transition to renewable energy sources, the need for more distributed generation, and the enhancement of grid reliability. This article presents a study of the main tasks in the field of energy network modeling, for which the use of machine learning methods is advisable. It examines how machine learning can contribute to optimizing network operations and play a key role in the modernization of energy systems, ensuring a more stable transition process to a decentralized energy system architecture.

The article covers a wide range of issues related to the application of machine learning in energy systems, including the analysis of optimal power distribution tasks, system state estimation, demand and generation forecasting, among others. It also explores the various types of data used for training models and provides practical advice on their processing. Scenarios for the application and categorization of different machine learning methods, such as supervised learning, unsupervised learning, and reinforcement learning, are presented. Additionally, hybrid approaches that combine machine learning with classical physical models, thereby ensuring greater efficiency and stability, are discussed.

Finally, the article provides an analysis of one of the largest network datasets from GridKit, within the framework of which the modeling of the European high-voltage network regarding its structural features and connections is carried out. Based on the results, promising directions for future research using machine learning methods, such as large language models and graph neural networks, are highlighted. Overall, this article is a comprehensive analysis of important tasks in energy modeling, relevant datasets, and machine learning models and can be used to further enhance intelligent energy systems.

Key words: machine learning, power grid data analysis, power grids, smart power systems, energy process modeling, demand and supply forecasting, network optimization.

Вступ

Перед людством постає ряд складних викликів, серед яких зменшення впливу на екологію та прискорення процесу децентралізації енергетики [1]. Окремо варто виділити проблему інтеграції численних невеликих генераторів на основі відновлюваних джерел енергії, таких як сонячні панелі й вітряки, та нових класів споживачів, таких як електромобілі, що вносять додаткову невизначеність в енергобаланс і обумовлюють потребу в нових підходах до управління енергосистемою [1, 2].

Класичні методи аналізу й моделювання енергосистем вже сьогодні на практиці виявляються недостатньо ефективними, оскільки відповідні моделі зазвичай не враховують значну кількість сучасних динамічних факторів [6]. Наприклад, уже зараз відновлювані джерела електроенергії становлять суттєву частку генерації електроенергії, їх генерація є мінливою і залежить, зокрема, від погодних умов [1, 6]. Водночас можливості щодо контролю та моніторингу самої мережі є досить обмеженими, тому постає потреба в більш комплексних інструментах моделювання, які дадуть змогу врахувати вищезгадані особливості й забезпечити стабільність параметрів електромережі [1, 2]. Також є очевидним, що в перспективі з подальшим розширенням мережі та розвитком відновлюваної генерації ця потреба буде лише посилюватися [1, 2].

Технології у сфері штучного інтелекту і, зокрема, машинного навчання швидко розвиваються завдяки регулярній появі у різних галузях нових, великих наборів даних, а також періодичному вдосконаленню спеціалізованих обчислювальних систем [2]. Останні досягнення машинного навчання великою мірою обумовлені появою і успішним впровадженням нових архітектур моделей глибинного навчання (deep learning), серед них трансформерні нейромережі (transformer neural networks) та графові нейромережі (graph neural networks, GNN) [2]. Уже зараз стрімкий розвиток цієї сфери суттєво впливає і на вищезгадані виклики, пов'язані з енергетикою, а моделі на базі машинного навчання можуть доповнити або навіть замінити традиційні методи [2].

Проте, незважаючи на всі перспективи застосування машинного навчання в моделюванні енергосистем, часто таким дослідженням не вистачає аналізу доцільності його використання для конкретної задачі [2, 6]. Наприклад, складні архітектури глибоких нейромереж, що потребують для навчання великої кількості даних, не можуть бути ефективно адаптовані до завдань енергетики, де таких вибірок може не бути в загальному доступі [2]. З іншого боку, впровадження подібних моделей для симуляції і прогнозування показників енергосистеми на великому об'ємі даних матиме багато перешкод, а також ресурсних і часових обмежень [2].

Виконання досліджень

Модель енергосистеми. Енергосистема є складною багаторівневою структурою, що об'єднує передавальну мережу, завданням якої є передача електроенергії на великі відстані за допомогою високовольтних ліній, і мережу розподілу, яка доставляє електроенергію до кінцевих споживачів [1]. У класичних енергосистемах 20-го століття електроенергія передавалася в одному напрямку, від великих централізованих генераторів електроенергії до споживачів [1]. Проте з ростом кількості приватних електростанцій, у тому числі на основі відновлюваних джерел живлення, усе більшу частку становлять двоспрямовані потоки електроенергії [1]. Це спричиняє ряд проблем і викликів, що негативно впливають на ефективність роботи електромереж [1, 2].

Ключовою вимогою безпечного функціонування енергосистеми є чітке дотримання балансу виробництва та споживання електроенергії, що, зокрема, забезпечується сталою частотою (50 або 60 Гц залежно від країни, в Україні стандартною мережевою частотою є 50 Гц). Будь-який дисбаланс призводить до зміни частоти: вона падає за нестачі генерації і зростає в разі надлишку [1]. Підтримання енергобалансу досягається за допомогою планування й маневрування потужністю контрольованих генераторів, таких як теплові (вугільні та газові) електростанції [1, 2].

Топологія енергосистеми описує фізичне розташування та з'єднання всіх її компонентів і зазвичай надається у вигляді орієнтованого або неорієнтованого графа. Цей граф складається з вузлів (генератори, навантаження, підстанції) і ребер (лінії електропередачі), що їх з'єднують. У загальному випадку модель енергосистеми можна описати у вигляді системи рівнянь, якими встановлюються взаємозв'язки між компонентами та їх вплив [3]. Також варто зазначити, що модель мережі розподілу суттєво відрізняється від передавальної мережі тим, що має меншу зв'язність, більшу мінливість, великі значення відношення опору до реактивного опору (відношення R/X) та дисбаланс фаз через

асиметричність навантаження. Відмінності між мережами обумовлюють необхідність у різних підходах до їх аналізу. Наприклад, через велике значення R/X опір ліній у мережах розподілу не можна ігнорувати, на відміну від систем передачі [3].

Традиційні підходи засновані на складних фізичних моделях і великій кількості факторів [1, 2]. Проте через масштаб сучасних енергосистем часто для прискорення моделювання можуть застосовуватися деякі апроксимації. Такі спрощені моделі можуть давати задовільні результати для значної частини задач, але іноді призводять до погіршення якості [2]. Тож методи машинного навчання можуть використовуватися як потужні апроксиматори, що мають більш прогнозований час відгуку та кращі можливості адаптації до змін у поведінці системи, ніж класичні методи [2]. Далі наведено огляд основних задач аналізу енергосистем і можливостей використання для їх вирішення підходів машинного навчання:

Задача потоків потужності (power flow problem). Ця задача є однією з найважливіших для аналізу усталеного режиму роботи енергосистем. Спочатку визначається стан системи в нормальних умовах роботи, зокрема величини напруги й фазового кута в усіх вузлах системи. На основі цих даних можуть бути аналітично розраховані інші величини, такі як активна та реактивна потужність, втрати на кожній лінії тощо. Ця інформація є надзвичайно важливою для операторів енергосистем, оскільки вона дає їм змогу розуміти поведінку системи за різних умов експлуатації, виявляти потенційні проблеми та вживати заходів для підтримання стабільності. Задача аналізу розподілу потужності зазвичай описується у вигляді системи нелінійних рівнянь [4]:

$$\begin{aligned} P_i &= V_i \sum_{j \in N} V_j (G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)), i \in N, \\ Q_i &= V_i \sum_{j \in N} V_j (G_{ij} \sin(\theta_i - \theta_j) - B_{ij} \cos(\theta_i - \theta_j)), i \in N, \end{aligned} \quad (1)$$

де N – множина усіх вузлів системи. $P_i = P_i^G - P_i^L$ та $Q_i = Q_i^G - Q_i^L$ є активними та реактивними потужностями вузла i відповідно. Верхні індекси G і L використовуються для позначення потужності генерації (generation) та навантаження (load) відповідно. V_i та θ_i є величиною напруги та фазовим кутом вузла i . G_{ij} та B_{ij} є провідністю та сприйнятливостю лінії між вузлами i та j – дійсна та уявна частина комплексної провідності лінії Y_{ij} відповідно.

У більш загальному випадку рівняння розподілу потоків потужності можна виразити таким нелінійним рівнянням [4]:

$$f(x, p, A) = 0, \quad (2)$$

де x – вектор невідомих системи, p – вектор параметрів ліній (опори і реактивні опори), A є матрицею зв'язків, яка описує топологію мережі. Оскільки аналітичного рішення для даної задачі не існує, зазвичай вона вирішується за допомогою ітераційних методів, наприклад методом Ньютона. У термінах машинного навчання, якщо враховувати, що ця модель описує усталені процеси, вона може бути виражена так [4]:

$$y = F(x) + \varepsilon, \quad (3)$$

де x та y є вхідним на вихідним вектором відповідно, ε – помилка (або шум), F – модель, задана у вигляді функції.

Задача оптимального розподілу потужності (optimal power flow). Ця задача заснована на деяких умовах попередньої та є поширеною серед операторів енергосистем. Формулюється як задача оптимізації з обмеженнями [5]:

$$\min_{x,u} c(x,u),$$

$$\text{за умови що: } g_i(x,u) = 0 (1 \leq i \leq n),$$

$$h_j(x,u) \leq 0 (1 \leq j \leq m), \quad (4)$$

де $c(\cdot)$ – цільова функція, яка визначається для конкретної задачі (наприклад, максимізація стабільності мережі, мінімізація операційних витрат тощо), x – вектор змінних, що описує стан системи (генерацію та споживання вузлів), u – вектор змінних керування, таких як цільові значення параметрів генераторів, а $g(\cdot)$ та $h(\cdot)$ – лінійні або нелінійні рівняннями та нерівності, які задають обмеження енергосистеми. Якщо застосувати цей тип задачі до вищезгаданої проблеми потоків потужності, то система нерівностей передбачатиме потрапляння значення напруги та фазового кута у визначені робочі діапазони й умову, що комплексна потужність кожної лінії передачі не перевищує її максимальні значення.

Задача оптимального розподілу потужності є NP-повною [5]. Тому зазвичай використовуються спрощені підходи для її розв'язку, наприклад лінійна апроксимація або ж метод «чорної скрині», що може використовувати моделі машинного навчання [2, 5]. На відміну від підходу «чорної скрині»,

у деяких останніх дослідженнях машинне навчання комбінується з фізичними моделями, правилами і класичними інструментами для аналізу електромереж, разом утворюючи так звану фізично обґрунтовану модель машинного навчання або гібридну модель [6]. Перевага цієї концепції полягає у тому, що для навчання такої моделі зазвичай потрібно значно менше даних, оскільки модель від самого початку має потужний сигнал у вигляді формалізованих обмежень і правил мережі. Таким чином досягається оптимальне співвідношення використання наявних знань про систему та її властивості, а також автоматичне вивчення її закономірностей на основі даних [6]. Принцип роботи даних методів узагальнено у вигляді схеми на рис. 1.



Рис. 1. Фізична модель на основі системи рівнянь і правил та альтернативи з використанням машинного навчання: «чорна скриня» та гібридна модель

Задача оцінки стану (state estimation). Є надзвичайно важливою для забезпечення ситуаційної обізнаності в енергосистемі та має на меті передбачення її найбільш імовірного стану, що зазвичай представлений величинами напруг і фазових кутів вузлів. Для вирішення цієї задачі потрібні історичні дані відповідних вимірювань та актуальна топологічна модель, включно з параметрами компонентів. Основною метою задачі є надання точного оцінювання стану системи в реальному часі (позначається як x) який згоджується з доступними метриками, насамперед величинами струму й напруги. Класична задача формулюється так [7]:

$$\tilde{x}(z) = \underset{x}{\operatorname{argmin}} (z - h(x))^2, \quad (5)$$

де z – вектор вимірювань, $h(\cdot)$ – нелінійна функція, що описує зв'язок між вектором стану системи x та вимірюваннями z . Оптимальна оцінка стану $\tilde{x}(z)$ обчислюється на основі мінімізації відхилення z від $h(x)$.

Для класичних способів розв'язання цієї задачі, наприклад для методу зважених найменших квадратів, потрібна значна кількість даних щодо вимірювань і точна структура моделі мережі, що часто є неможливим. І навіть за цих умов під час застосування таких методів можна стикнутися з проблемами щодо збіжності й надмірною чутливістю до викидів у даних [2, 7]. Зі свого боку, підходи машинного навчання показують перспективні результати в різних аспектах вирішення задачі оцінки стану енергосистеми, таких як виявлення помилкових даних, автоматизована ідентифікація топології, генерація синтетичних даних, рекомендація схожих за поведінкою і параметрами компонентів системи [2].

Джерела даних. Для забезпечення різних функцій в енергосистемі використовується широкий спектр джерел даних, таких як системи диспетчерського управління та збору даних (supervisory control and data acquisition, SCADA), метеорологічні дані, а також мережеві вимірювальні пристрої, наприклад пристрої вимірювання комплексної амплітуди (phasor management unit, PMU) та розумні лічильники [2]. Крім того, існують численні джерела даних, які надаються операторами енергосистем та їх об'єднаннями, такими як ENTSO-E [8]. Ці дані містять історичні часові ряди (попит і пропозиція, динамічні ціни на електроенергію, генерація відновлюваних джерел живлення), а також параметри генераторів, мережевих ліній та основних споживачів. Основні типи джерел даних узагальнено на рис. 2.



Рис. 2. Джерела даних, що використовуються для керування й моделювання енергосистем

Ці дані можуть бути використані для побудови моделей машинного навчання для різних задач [2, 8]. Проте потрібно враховувати, що в кожного джерела даних є свої особливості й обмеження (наприклад, різна частота та моменти вимірювань). Зокрема, покази з PMU, підключеного до мережі, можуть приходити з інтервалом, який вимірюється приблизно в мілісекундах, дані SCADA – у секундах, дані з розумних лічильників абонентів – у хвилинах, дані про погоду – у годинах тощо [8]. Тому для ефективного використання цих даних під час моделювання потрібно виконувати стандартизацію даних, що може передбачати зниження частоти вибірки даних із високою частотою вимірювань до частоти вибірки з нижчою частотою або навпаки [2, 8]. Цю операцію можна виконувати як за допомогою простої лінійної інтерполяції, так і використовуючи більш складні методи, у тому числі на основі машинного навчання, наприклад, застосовуючи для синтезу даних генеративні змагальні мережі (generative adversarial networks, GANs) [9].

Основні приклади використання машинного навчання. Машинне навчання застосовується для широкого спектра прикладних задач, пов'язаних із різними аспектами моделювання енергосистеми [2]. Основні сімейства методів машинного навчання, прикладів їх використання та релевантних алгоритмів зведено у схему, яку зображено на рис. 3.

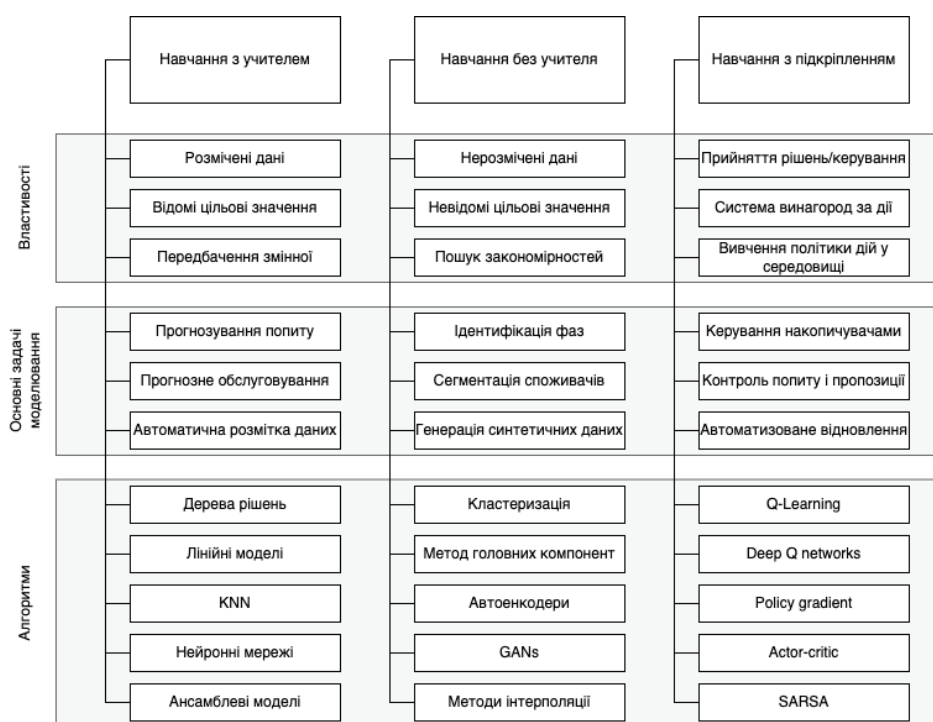


Рис. 3. Категоризація методів машинного навчання відповідно до різних задач моделювання енергосистеми

Далі наведено опис найбільш популярних сценаріїв, для яких або вже використовують, або можуть у перспективі почати застосовувати методи машинного навчання.

Для **навчання з учителем (supervised learning)** двома основними типами моделей є класифікація і регресія. У контексті електромереж такі моделі є найбільш дослідженими та широко розповсюдженими і, наприклад, використовуються для прогнозування (forecasting) локальних ринкових цін на електроенергію або стану вузлів системи на основі часових рядів [2]. Ці прогнози надають

важливу інформацію і використовуються для прийняття рішень операторами енергосистем та власниками енергетичних об'єктів. Наприклад, точне прогнозування генерації відновлюваної енергії допомагає операторам більш ефективно керувати сегментами мережі з великою часткою розподіленої генерації, зменшуючи експлуатаційні витрати й ризики виникнення дисбалансу [2]. Регресійні моделі відрізняються горизонтом планування, що може становити як секунди або хвилини (для моделювання у режимі наближеному до реального часу), так і місяці або роки (для стратегічного планування й аналізу ризиків) [10]. Іншим прикладом є прогнозне обслуговування (predictive maintenance) [2]. Використовуючи історичні дані показів стану мережевого обладнання, такі як деградація ізоляції у трансформаторах та історія виявлених несправностей, можна навчити модель для передбачення виходу обладнання з ладу, що дає змогу операторам вчасно вживати превентивні заходи й уникати аварійних ситуацій [2].

Навчання без учителя (unsupervised learning) є дуже перспективним сімейством методів для застосування в енергосистемах, оскільки не потребує дороговартісних і зазвичай наявних лише в невеликій кількості розмічених даних та охоплює широкий спектр популярних задач, таких як кластеризація (clustering), зменшення розмірності (dimensionality reduction), виявлення аномалій (anomaly detection), генерація синтетичних даних (synthetic data generation) [2]. У рамках енергосистем прикладом їх застосування є ідентифікація фазових підключень – для ефективного розміщення нових розподілених потужностей у мережі розподілу (наприклад, побутові та комерційні сонячні електростанції) важливо правильно ідентифікувати фазове підключення кожного споживача [11]. Для вирішення цієї задачі можна застосувати метод кластеризації, щоб згрупувати вузли системи щодо використання тих чи інших комбінацій фаз на основі показів з їх розумних лічильників або інших мережевих приладів [2, 11]. Іншим прикладом є сегментація споживачів електроенергії – автоматизоване визначення профілів мережевого навантаження (наприклад, комутовані або регульовані прилади), а також їх закономірностей споживання [11]. Розв'язання цієї задачі дає змогу енергетичним компаніям розробляти ефективніші стратегії залучення нових і заохочення наявних споживачів, враховуючи їх енергетичні потреби. Також таким чином оператори можуть більш ефективно реагувати на зміну попиту за різних умов, включно з погодою, соціальними подіями та святами [11].

Методи **навчання з підкріпленням (reinforcement learning)** є перспективними для задач прийняття рішень і керування енергетичними системами, оскільки дають змогу вивчити оптимальну політику і спланувати потрібну послідовність дій у заданому середовищі з метою досягнення бажаного результату [2]. Прикладом застосування може бути автоматизована система керування мережевими накопичувачами електроенергії в умовах динамічного ціноутворення електроенергії, як-от Tesla Autobidder [12]. Метою такої системи є максимізація доходів унаслідок оптимальних дій щодо збереження та продажу оператору енергосистеми електроенергії залежно від поточного балансу попиту та пропозиції. Водночас ці агенти мають враховувати ряд обмежень щодо експлуатації накопичувача, наприклад потребу уникнення глибоких розрядів [12]. Іншим цікавим прикладом є автоматизація оператора енергосистеми за допомогою моделі на основі навчання з підкріпленням [2, 12]. Метою агента на основі подібної моделі є підтримка стабільної напруги в мережі й уникнення аварійних відключень. Для цього він може виконувати дії щодо регулювання та комутації компонентів системи [13].

Аналіз загальноєвропейського мережевого набору даних від GridKit. У загальному доступі існує велика кількість наборів даних, що описують топологію електромереж різних країн і регіонів, параметри мережевих об'єктів та історію їх зміни [3]. Такі набори зазвичай створюються на основі даних, що надаються національними операторами та міжнародними енергетичними агентствами й організаціями. Для доповнення цих даних також часто використовуються відкриті картографічні ресурси, такі як OpenStreetMap, що містять географічні координати й інші додаткові атрибути енергетичних об'єктів [8].

GridKit є одним із популярних інструментів, що використовує вищезгаданий підхід, і на його основі регулярно створюються великі мережеві вибірки [14]. Прикладом є набір даних європейської та північноамериканської високовольтних мереж станом на 2016 р., який базується, зокрема, на інформації, що регулярно публікується асоціацією ENTSO-E (Європейською мережею операторів систем передачі електроенергії) [8]. Саме європейська складова цієї вибірки використовувалася в рамках цього дослідження для аналізу й моделювання.

Вибірка відкрита для завантаження через ресурс Zenodo та надається у вигляді архіву з двома CSV файлами – вузлами й ребрами мережі відповідно [14]. Файли містять значну кількість метаданих, таких як тип вузла, оператор, номінальні показники кожного компонента (частота, напруга, потужність), їх географічне розташування тощо.

У рамках аналізу основним завданням було дослідити структуру, особливості даних, а також їх потенціал для моделювання, зокрема, використовуючи машинне навчання. Обробка даних проводилася за допомогою мови програмування Python, бібліотеки для роботи з табличними даними Pandas і бібліотеки для побудови й аналізу графів Networkx. Попередньо дані було очищено (видалено дублі-

кати, нормалізовано типи вузлів і залишено лише найбільший компонент зв'язності) та перетворено їх на networkx-граф. Усього граф складається з 13 871 вузла та 18 804 ребер. Як зображено на рис. 4, у вибірці виокремлено 3 основні типи вузлів: 1) вузлова точка – усього 6330 елементів, це умовна назва для місць з'єднання кількох ліній електропередачі (без трансформаторного обладнання); 2) підстанція – усього 6395 елементів, вузли, де відбувається трансформація та розподіл електроенергії; 3) генератор – усього 1146 елементів, у цих вузлах відбувається генерація електричної енергії різних джерел (теплових, атомних, відновлюваних тощо).

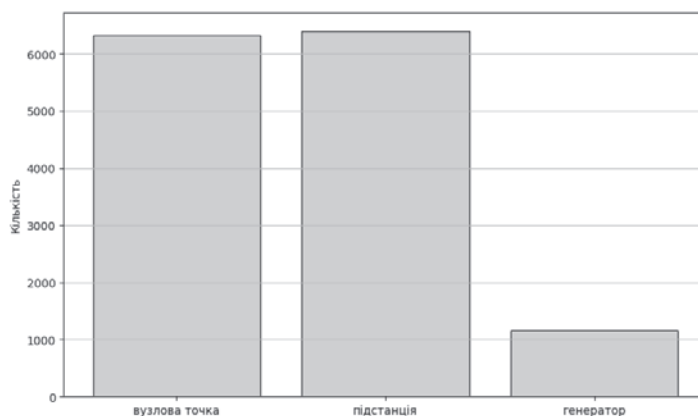


Рис. 4. Гістограма кількості вузлів різних типів у вибірці GridKit

Для аналізу структури та властивостей мережевої моделі, відповідно до даних, за допомогою функції `curve_fit` бібліотеки `scikit-learn` було виконано автоматизований підбір параметрів степеневого розподілу (power law). Також було розраховано степені вузлів (кількість зв'язків), що продемонструвало їх підпорядкування раніше отриманому степеневому закону. Це можна побачити на рис. 5. Причина цієї закономірності полягає в тому, що мережі (і не тільки енергетичні) зазвичай розширюються за принципом приєднання нових вузлів до вже наявних вузлів із високим ступенем. Тож завдяки цій властивості вузли в мережах поділяються на незначну кількість елементів зі значними степенями і велику кількість вузлів із малими степенями.

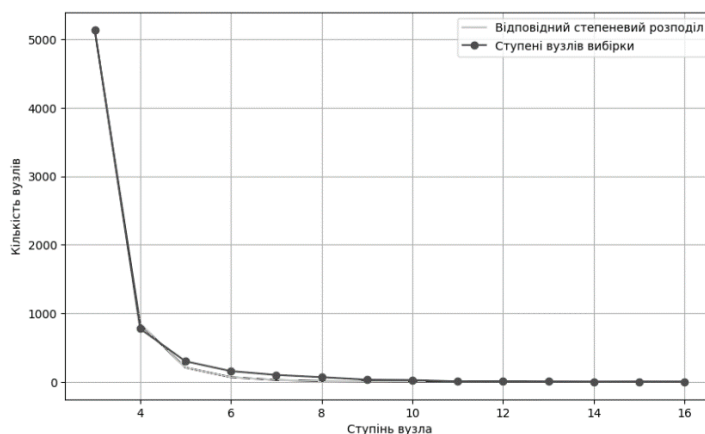


Рис. 5. Розподіл степенів вузлів графа та його зв'язок зі степеневим законом

Крім того, за допомогою засобів `Networkx` було розраховано одну з найважливіших метрик оцінки впливу кожного з вузлів на мережу – центральність за посередництвом (*betweenness centrality*). Чисельно ця метрика виражає відношення кількості найкоротших шляхів, які проходять через цей вузол, до загальної кількості найкоротших шляхів у мережі. Значення перебуває у межах від 0 до 1, де 0 означає, що вершина не має впливу на будь-які інші, а 1 означає, що вершина задіяна в найкоротших шляхах між усіма вершинами. Для наочності на рис. 6 наведено український сегмент мережі з основними вузлами, їх типами, а також величинами центральності. Чим більший вузол на рисунку, тим більше його значення центральності.

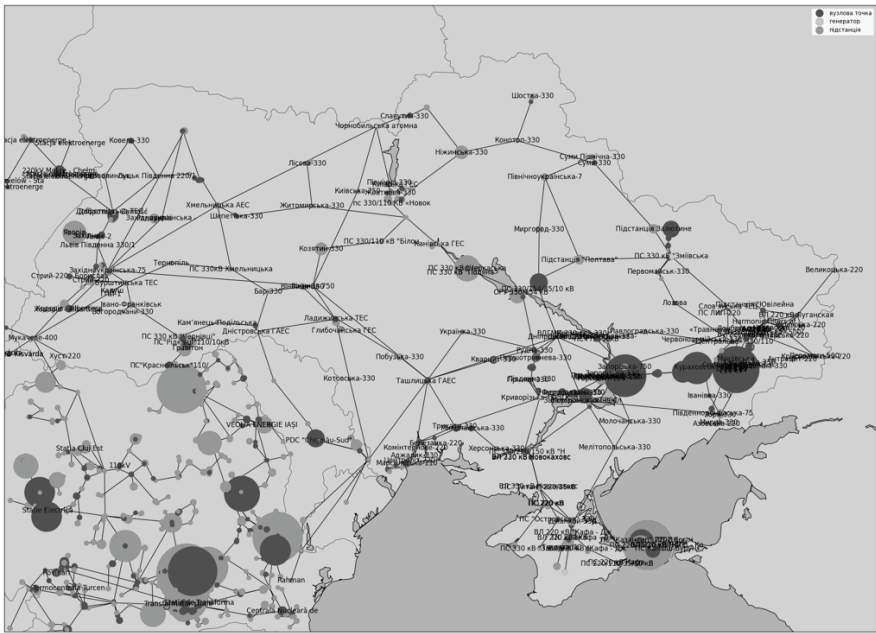


Рис. 6. Український сегмент мережі вибірки GridKit з розрахованими коефіцієнтами центральності

Із цього зображення можна перекоонатися, що найбільш важливими є вузли, які забезпечують транскордонну передачу електроенергії, а також розташовані в щільних сегментах мережі. На рис. 7 наведено частотний розподіл коефіцієнтів центральності за посередництвом по всій європейській вибірці в логарифмічному масштабі. Він також свідчить про те, що вплив вузлів у мережі підпорядковується ступеневому закону.

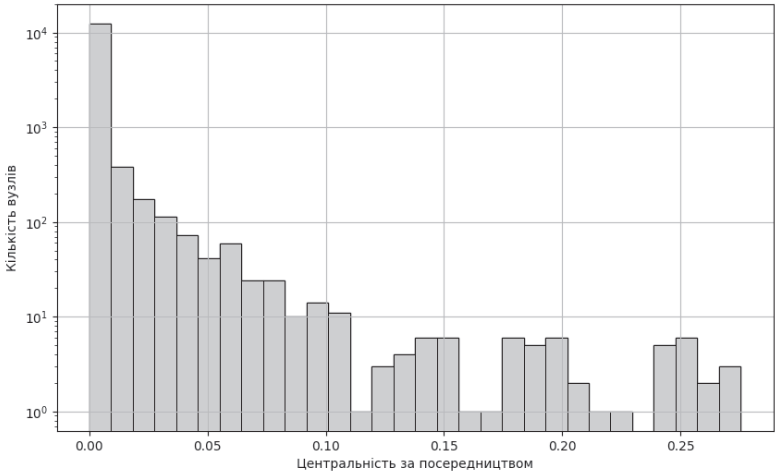


Рис. 7. Розподіл значень коефіцієнтів центральності за посередництвом

Вузли європейської мережі з найбільшими значеннями коефіцієнтів центральності за посередництвом наведено в табл. 1. Серед них є елементи всіх типів, проте переважно це вузлові точки.

Таблиця 1

Вузли європейської високовольтної мережі з найбільшою центральністю

Мережевий вузол	Тип	Коефіцієнт
Redipuglia (Італія)	вузлова точка	0.275585
Dugale (Італія)	підстанція	0.270770
RTP Divača (Словенія)	підстанція	0.266539
Ostiglia;Centrale termoelettrica di Ostiglia (Італія)	генератор	0.259387

Продовження таблиці 1

Udine Ovest – Cordignano (Італія)	вузлова точка	0.252068
Sandrigo (Італія)	підстанція	0.250508
Udine Ovest (Італія)	вузлова точка	0.248413
Poste électrique de Albertville (Франція)	вузлова точка	0.243978
Superphénix – Rondissone-Albertville (Франція)	вузлова точка	0.242329
TS Melina (Хорватія)	вузлова точка	0.240040

Крім центральності за посередництвом, також аналізувалися інші розповсюджені метрики важливості вузлів, такі як степенева центральність (degree centrality) та алгоритм PageRank. Було проведено кореляційний аналіз розподілів коефіцієнтів вузлів, отриманих за допомогою різних методів, а результати наведено в табл. 2. Ці значення демонструють, що степенева центральність і PageRank дають схожі результати, оскільки однаково ґрунтуються на кількості зв'язків вузлів. Водночас центральність за посередництвом має суттєво інший розподіл через те, що використовує складнішу оцінку, засновану на прохідності вузлів.

Таблиця 2
Кореляція різних метрик центральності

	Центральність за посередництвом	Центральність за степенем	PageRank
Центральність за посередництвом	–	0.2957	0.2665
Центральність за степенем	–	–	0.9540
PageRank	–	–	–

Цей аналіз мережевої вибірки даних також адресує задачу симуляції відмов вузлів, використовуючи різні підходи (відключення випадкового вузла або найбільш важливих вузлів за раніше згаданими критеріями центральності), у тому числі, як швидко мережа розпадається на енергетичні острови зі збільшенням кількості відключень відповідно до цих методів порівняно з випадковими відмовами. Тобто ця мережева модель даних може використовуватися для широкого спектра як класичних задач, так і таких, де доцільне використання машинного навчання.

Перспективні напрями досліджень. Крім згаданих раніше методів машинного навчання, які вже стали популярними й отримали реальні приклади застосування в енергетичній сфері, останнім часом стрімкого розвитку зазнав напрям, пов'язаний із моделюванням природної мови (natural language processing). Насамперед це обумовлено появою великих мовних моделей (large language models, LLMs), як загальних, так і адаптованих до конкретної предметної сфери, що дали змогу адресувати складні виклики, які не вдавалося ефективно вирішити раніше [15]. Водночас, на відміну від таких галузей, як охорона здоров'я, юриспруденція або екологія, в енергетиці подібні моделі досі не є розповсюдженими [15]. Останні дослідження виділяють декілька основних сценаріїв застосування мовних моделей для аналізу енергосистеми. Ключовий із них – це побудова, оновлення й пошук за графом знань електромережі. Ідея полягає у генерації графової структури на основі правил, закономірностей, а також типових відхилень, що притаманні енергомережам, і асоціація з елементами цього графа реальних вузлів і частин енергосистеми з урахуванням їх історичної поведінки [15]. Цей механізм зв'язку вузлів з елементами графа знань зазвичай реалізується за допомогою отримання ембедінгів (embeddings) – семантичних векторних репрезентацій, що генеруються нейромережами (у тому числі мовними моделями), для яких можна розраховувати певну обрану метрику схожості й відповідно до отриманого значення визначати приналежність об'єктів до певних категорій та один до одного, аналізувати зв'язки між ними [15]. Крім того, важливою функцією використання мовних моделей є можливість їх впровадження як асистента, основним завданням є оркестрація пошуку за вищезгаданим графом і надання текстових пояснень за результатами аналізу. Це може спростити роботу, зробити результати моделювання більш зрозумілими і зменшити кількість складних дій у різних інструментах, які потрібно виконувати оператору енергосистеми, замінивши більшість із них на взаємодію з асистентом за допомогою природної мови. Основним викликом у реалізації такого підходу є адаптація загальних мовних моделей до енергетичного домену, оскільки загальні моделі самі по собі не мають достатнього контексту, щоб якісно виконувати енергетичні завдання, і через складнощі отримання великої кількості спеціалізованих даних, які можна було б використати для донавчання моделей [2, 15].

Як зазначалося раніше, оскільки електромережі зазвичай наводяться у вигляді графової структури, ще одним важливим напрямом є моделювання за допомогою графових нейронних мереж. Завдяки своїм можливостям щодо якісного вивчення властивостей, закономірностей і зв'язків вузлів у графах, цей тип моделей усе частіше стає об'єктом досліджень, пов'язаних з енергосистемами [2, 16]. Напри-

клад, в одному з останніх досліджень пропонується баєсівська модель машинного навчання, компонентом якої є графова нейромережа для прогнозування потоків потужності трансформаторів, що з'єднують звичайні й високовольтні лінії електропередачі [16]. Завдяки ймовірнісній складовій цієї моделі вона ефективніше розподіляє увагу щодо впливу різних вузлів мережі та враховує широкий спектр сигналів, як-от погодні умови, баланс енергосистеми, взаємодія із суміжними вузлами. Крім того, векторні представлення вузлів, отримані з графових нейромереж, мають семантичні властивості, на основі яких можна проводити різноманітні типи аналізу (кластеризація, оцінка схожості тощо) [16].

Висновки

Застосування різноманітних методів машинного надає потенціал для більш ефективного моделювання й управління енергосистемами, а також ефективного вирішення ключових задач цієї предметної сфери. До того ж використання машинного навчання часто забезпечує кращу, ніж у класичних методах, пропускну здатність і якість результатів, завдяки властивостям моделей навчатися на великих обсягах даних, регулярно оновлюватися і враховувати більшу кількість різних факторів.

Попри поточні успіхи у впровадженні машинного навчання, існує ряд серйозних викликів, зокрема потреба у великих масивах даних для навчання та застосування складних моделей до специфічних завдань енергетики. Також важливим напрямом є поєднання цих методів з наявними фізичними моделями для досягнення більшої точності результатів та потенційного зменшення обсягу даних, потрібних для навчання.

Проведений аналіз одного з найбільших мережевих наборів даних від GridKit у поєднанні з сформованими на основі цього напрямками подальших досліджень може бути використаний для подальшого вдосконалення методів машинного навчання в контексті їх адаптації до енергетичної сфери, ефективнішого застосування апріорних знань про досліджувану задачу, а також упровадження новітніх підходів, які добре зарекомендували себе для схожих проблем в інших галузях. Усе це разом має на меті допомогти забезпечити стабільну роботу енергосистеми та швидше реагування на зміни в мережі.

Література

1. Sinsel, S.R., Riemke, R.L., & Hoffmann, V.H. (2020). Challenges and solution technologies for the integration of variable renewable energy sources – a review. *Renewable Energy*, 145, 2271–2285. DOI: <https://doi.org/10.1016/j.renene.2019.06.147>.
2. Lee, S., Seon, J., Hwang, B., Kim, S., Sun, Y., & Kim, J. (2024). Recent Trends and Issues of Energy Management Systems Using Machine Learning. *Energies*, 17, 624. DOI: <https://doi.org/10.3390/en17030624>.
3. Shen, T., Li, Y., & Xiang, J. (2018). A Graph-Based Power Flow Method for Balanced Distribution Systems. *Energies*, 11, 511. DOI: <https://doi.org/10.3390/en11030511>.
4. Huang, W.-T., & Yang, W.-C. (2011). Power flow analysis of a grid-connected high-voltage microgrid with various distributed resources. In *Proceedings of the 2011 International Conference on Materials and Applications for Communications and Electronics (MACE)*. DOI: <https://doi.org/10.1109/MACE.2011.5987225>.
5. ScienceDirect. Optimal power flow. [Електронний ресурс]. Режим доступу: <https://www.sciencedirect.com/topics/engineering/optimal-power-flow>.
6. Stiasny, J., & Chatzivasileiadis, S. (2023). Physics-informed neural networks for time-domain simulations: Accuracy, computational cost, and flexibility. *Electric Power Systems Research*, 224, Article 109748. DOI: <https://doi.org/10.1016/j.epsr.2023.109748>.
7. Vijaychandra, J., Prasad, B. R. V., Darapureddi, V. K., Rao, B. V., & Knypinski, Ł. (2023). A review of distribution system state estimation methods and their applications in power systems. *Electronics*, 12 (3), 603. DOI: <https://doi.org/10.3390/electronics12030603>.
8. Ledesma, P., Arredondo, F., & Moreno, M. Á. (2023). Adapting ENTSO-E network data for power system simulation. In *2023 IEEE International Conference on Environment and Electrical Engineering and 2023 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)* (pp. 1–5). DOI: <https://doi.org/10.1109/EEEIC/ICPSEurope57605.2023.10194624>.
9. Li, J., Chen, Z., Cheng, L., & Liu, X. (2022). Energy data generation with Wasserstein deep convolutional generative adversarial networks. *Energy*, 257, 124694. DOI: <https://doi.org/10.1016/j.energy.2022.124694>.
10. Benti, N.E., Chaka, M.D., & Semie, A.G. (2023). Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects. *Sustainability*, 15 (7), 7087. DOI: <https://doi.org/10.3390/su15097087>.
11. Zhou, L., Zhang, Y., Liu, S., Li, K., Li, C., Yi, Y., & Tang, J. (2020). Consumer phase identification in low-voltage distribution network considering vacant users. *International Journal of Electrical Power & Energy Systems*, 121, 106079. DOI: <https://doi.org/10.1016/j.ijepes.2020.106079>.
12. Tesla. Autobidder. [Електронний ресурс]. Режим доступу: <https://www.tesla.com/support/energy/tesla-software/autobidder>.
13. Cao, D., et al. (2020). Reinforcement learning and its applications in modern power and energy systems: A review. *Journal of Modern Power Systems and Clean Energy*, 8 (6), 1029–1042. DOI: <https://doi.org/10.35833/MPCE.2020.000552>.
14. GridKit: European and North-American extracts. [Електронний ресурс]. Режим доступу: <https://zenodo.org/records/47317>.
15. Lin, C., et al. (2023). Knowledge graph completion for power grid main equipment using pretrained language models. In D. S. Huang, P. Premaratne, B. Jin, B. Qu, K. H. Jo, & A. Hussain (Eds.), *Advanced Intelligent Computing Technology and Applications. ICIC 2023. Lecture Notes in Computer Science* (Vol. 14089). Springer. DOI: https://doi.org/10.1007/978-981-99-4752-2_68.
16. Beinert, D., Holzhüter, C., Thomas, J.M., & Vogt, S. (2023). Power flow forecasts at transmission grid nodes using graph neural networks. *Energy and AI*, 14, 100262. DOI: <https://doi.org/10.1016/j.egyai.2023.100262>.

References

1. Sinsel, S.R., Riemke, R.L., & Hoffmann, V.H. (2020). Challenges and solution technologies for the integration of variable renewable energy sources – a review. *Renewable Energy*, 145, 2271–2285. DOI: <https://doi.org/10.1016/j.renene.2019.06.147>.
2. Lee, S., Seon, J., Hwang, B., Kim, S., Sun, Y., & Kim, J. (2024). Recent Trends and Issues of Energy Management Systems Using Machine Learning. *Energies*, 17, 624. DOI: <https://doi.org/10.3390/en17030624>.
3. Shen, T., Li, Y., & Xiang, J. (2018). A Graph-Based Power Flow Method for Balanced Distribution Systems. *Energies*, 11, 511. DOI: <https://doi.org/10.3390/en11030511>.
4. Huang, W.-T., & Yang, W.-C. (2011). Power flow analysis of a grid-connected high-voltage microgrid with various distributed resources. In *Proceedings of the 2011 International Conference on Materials and Applications for Communications and Electronics (MACE)*. DOI: <https://doi.org/10.1109/MACE.2011.5987225>.
5. ScienceDirect. Optimal power flow. Retrieved from <https://www.sciencedirect.com/topics/engineering/optimal-power-flow>.
6. Stiasny, J., & Chatzivasileiadis, S. (2023). Physics-informed neural networks for time-domain simulations: Accuracy, computational cost, and flexibility. *Electric Power Systems Research*, 224, Article 109748. DOI: <https://doi.org/10.1016/j.epsr.2023.109748>.
7. Vijaychandra, J., Prasad, B.R.V., Darapureddi, V.K., Rao, B.V., & Knypirski, Ł. (2023). A review of distribution system state estimation methods and their applications in power systems. *Electronics*, 12 (3), 603. DOI: <https://doi.org/10.3390/electronics12030603>.
8. Ledesma, P., Arredondo, F., & Moreno, M.Á. (2023). Adapting ENTSO-E network data for power system simulation. In *2023 IEEE International Conference on Environment and Electrical Engineering and 2023 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)* (pp. 1–5). DOI: <https://doi.org/10.1109/EEEIC/ICPSEurope57605.2023.10194624>.
9. Li, J., Chen, Z., Cheng, L., & Liu, X. (2022). Energy data generation with Wasserstein deep convolutional generative adversarial networks. *Energy*, 257, 124694. DOI: <https://doi.org/10.1016/j.energy.2022.124694>.
10. Benti, N.E., Chaka, M.D., & Semie, A.G. (2023). Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects. *Sustainability*, 15 (7), 7087. DOI: <https://doi.org/10.3390/su15097087>.
11. Zhou, L., Zhang, Y., Liu, S., Li, K., Li, C., Yi, Y., & Tang, J. (2020). Consumer phase identification in low-voltage distribution network considering vacant users. *International Journal of Electrical Power & Energy Systems*, 121, 106079. <https://doi.org/10.1016/j.ijepes.2020.106079>.
12. Tesla. Autobidder. Retrieved from <https://www.tesla.com/support/energy/tesla-software/autobidder>.
13. Cao, D., et al. (2020). Reinforcement learning and its applications in modern power and energy systems: A review. *Journal of Modern Power Systems and Clean Energy*, 8 (6), 1029–1042. DOI: <https://doi.org/10.35833/MPCE.2020.000552>.
14. GridKit: European and North-American extracts. Retrieved from <https://zenodo.org/records/47317>.
15. Lin, C., et al. (2023). Knowledge graph completion for power grid main equipment using pretrained language models. In D. S. Huang, P. Premaratne, B. Jin, B. Qu, K. H. Jo, & A. Hussain (Eds.), *Advanced Intelligent Computing Technology and Applications. ICIC 2023. Lecture Notes in Computer Science* (Vol. 14089). Springer. DOI: https://doi.org/10.1007/978-981-99-4752-2_68.
16. Beinert, D., Holzhüter, C., Thomas, J. M., & Vogt, S. (2023). Power flow forecasts at transmission grid nodes using graph neural networks. *Energy and AI*, 14, 100262. DOI: <https://doi.org/10.1016/j.egyai.2023.100262>.