

УДК 004.9

## МЕТОДИКА ПРОГНОЗУВАННЯ ОБСЯГІВ СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ НА БАЗІ МЕТОДІВ МАШИННОГО НАВЧАННЯ

Бесчастний Владислав Ігорович

д.т.н., проф. каф. ІС Заболкі Костянтин Вячеславович  
Національний університет «Одеська політехніка», УКРАЇНА

**АНОТАЦІЯ.** Прогнозування обсягів споживання електроенергії є дуже важливою задачею для енергетичних компаній та державних регуляторів. Прогнозування обсягів споживання електроенергії є важливим інструментом для планування виробництва електроенергії. Енергетичні компанії повинні забезпечити достатні обсяги електроенергії для задоволення попиту на різних рівнях споживачів, включаючи побутові, комерційні та промислові підприємства. Якщо енергетична компанія не зможе забезпечити достатні обсяги електроенергії для виконання запитів споживачів, то це може призвести до відключення електропостачання, що може призвести до серйозних наслідків для споживачів.

**Вступ.** Прогнозування обсягів споживання електроенергії є дуже важливою задачею для енергетичних компаній та державних регуляторів. Крім того, прогнозування обсягів споживання електроенергії також є важливим для забезпечення надійності енергопостачання. Прогнозування обсягів споживання електроенергії дозволяє енергетичним компаніям ефективно використовувати свої ресурси. Наприклад, якщо прогноуються великі обсяги споживання електроенергії, енергетичні компанії можуть збільшити виробництво електроенергії або використовувати додаткові ресурси для забезпечення високого рівня енергопостачання. Прогнозування обсягів споживання електроенергії є критично важливим для розвитку відновлюваної енергетики, оскільки забезпечення стабільного та надійного енергопостачання залежить від того, наскільки точні прогнози споживання електроенергії.

**Мета роботи.** Розробка та дослідження методики прогнозування обсягів споживання електроенергії на базі методів машинного навчання для зменшення похибки прогнозування.

**Основна частина роботи.** Для того, щоб отримати максимально точну модель прогнозування обсягів споживання електроенергії, необхідно спочатку виконати ряд операцій, які були розроблені для того, щоб підготуватися до етапу створення будь-якої якісної моделі прогнозування. Однак перш ніж переходити до прогнозного моделювання, важливо добре розібратися в ознаках. Моделі машинного навчання отримують знання з набору даних, що складається із зібраних спостережень, і, використовуючи ці спостереження як основу, роблять висновки про нові вхідні значення.

1. Набір даних та розвідувальний аналіз: Проводиться збір даних після чого відбувається розвідувальний аналіз. Розвідувальний аналіз даних, який іноді називають скорочено EDA – це процес, який має на меті провести аналіз даних з різних позицій. Це метод, який демонструє, як правильно проводити аналіз даних. Він використовує різноманітні аналітичні підходи для покращення розуміння даних.

2. Попередня обробка даних: Для того, щоб отримати найкращий можливий рівень продуктивності моделі, необхідно визначити численні характеристики даних. При застосуванні методів машинного навчання в будь-якій галузі досліджень, формат даних, змінні, які вибираються для їх представлення, і тип даних, що використовуються – це все фактори, які відіграють важливу і значну роль.

3. Навчання моделі: Процес навчання та перевірки наших моделей є важливим кроком у досягненні хорошого результату в нашому прагненні до машинного навчання. Перший етап полягає в тому, щоб розділити дані на навчальні та тестові набори даних.

4. Тестування та оцінка моделі: Оцінювання – це один з останніх етапів побудови моделі прогнозування. Для прийняття рішень на наступних етапах процесу важливо знати, наскільки добре працює наша модель. Перш ніж запускати новостворену модель у масове виробництво, як

і у випадку з будь-чим іншим, ми хочемо переконатися, що розуміємо, наскільки добре вона працює під час тестування.

5. Випуск: Після завершення розробки та тестування інтерфейс випускається у продакшн. Протягом післярелізного супроводу здійснюється моніторинг та збір даних щодо взаємодії користувачів з продуктом, а також проводяться виправлення помилок та вдосконалення інтерфейсу на основі отриманих відгуків.

Використовуючи статистичну термінологію, ми можемо пояснити, що середньоквадратична похибка – це міра, яка являє собою розрахунок стандартного відхилення оцінених похибок від лінії регресії. Середньоквадратична похибка, або RMSE – це статистична міра, яка показує, наскільки сильно здогадки відхиляються від найкращої моделі.

Скоріше, це лише показує, що модель потребує певного вдосконалення. Ми все одно можемо отримати на її основі результат, навіть якщо оцінка низька, якщо є значущі змінні, а це означає, що зміни в цих факторах означатимуть зміну цільової змінної.

Середня абсолютна похибка, також скорочено MAR, є середнім значенням абсолютної величини помилок, тоді як середня квадратична похибка, скорочено MSE, є середнім значенням квадратів помилок[1].

Коли йдеться про завдання класифікації, ми використовуємо різні метрики для оцінювання, зокрема матриці плутанини, точність, логарифмічні функції втрат та інші.

Попереднє встановлення етапів розвідувального аналізу даних, а також визначення методів обробки набору даних та можливих оцінок моделей прогнозування дозволяє сформулювати методіку прогнозування обсягів споживання електроенергії на базі методів машинного навчання. Блок-схема роботи методіки наведена на рис. 1.

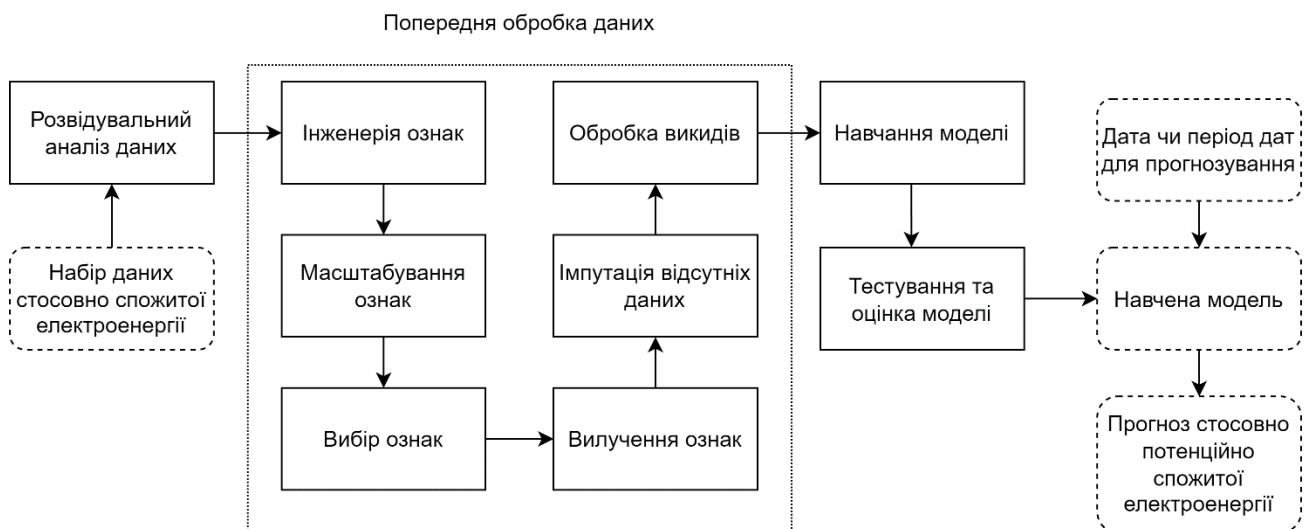


Рисунок 1 – Загальна ідея роботи методіки

Вхідними даними запропонованої методіки є набір історичних показників спожитої електроенергії протягом певного періоду часу. Після завантаження набору даних відбувається процес розвідувального аналізу, а також кроки, направлені на попередню обробку даних. Кожен етап роботи методіки був розглянутий вище [2].

Розвідувальний аналіз та попередня обробка даних дозволяє сформувати набір даних, який далі використовується для навчання та тестування моделі машинного навчання.

Після закінчення процесу навчання, готова модель може бути використана для отримання прогнозів стосовно споживання електроенергії на певну дату або діапазон дат.

Під час розробки методіки були встановлені чинники, що мають вплив на загальну ефективність моделі прогнозування. Такими чинниками є якість вхідного набору даних, процес відбору ознак, а також задача вибору належної моделі машинного навчання.

Було встановлено, що важливим етапом запропонованої методики є процес розвідувального аналізу даних. Були встановлені етапи розвідувального аналізу та цілі, які досягаються в процесі розвідувального аналізу.

Крім того було визначено, що після процесу розвідувального аналізу необхідно провести заходи з попередньої обробки даних для забезпечення належної ефективності роботи методики [3].

До таких етапів попередньої обробки можна віднести процес інжинірингу ознак, вибір належних та змістовних ознак в наборі даних, визначення стратегії роботи з відсутніми даними, а також ідентифікація та видалення викидів в даних [4].

Нажаль, під час війни більшість відкритих джерел даних були закриті (це окремо стосується питання видобутку та споживання електроенергії, зважаючи на події кінця 2022 року), тому для дослідження методики не вдалось сформувати належний набір даних стосовно споживання електроенергії в Україні [5].

PJM Interconnection LLC (PJM) є регіональною організацією з передачі електроенергії (RTO) у Сполучених Штатах. Вона є частиною мережі Eastern Interconnection, що управляє системою передачі електроенергії, яка обслуговує всі або частину штатів Делавер, Іллінойс, Індіана, Кентуккі, Меріленд, Мічиган, НьюДжерсі, Північна Кароліна, Огайо, Пенсильванія, Теннессі, Вірджинія, Західна Вірджинія та Округ Колумбія [6].

Набір даних містить погодинні дані про споживання електроенергії взяті з веб-сайту PJM і вказані в мегаватах (МВт). Регіони змінювалися з роками, тому дані можуть з'являтися лише за певні дати в кожному регіоні.

**Висновки.** Дослідження роботи моделей машинного навчання показало, що глибинна нейронна мережа показує кращі результати на тестовому наборі. Однак, на відміну від Prophet і XGBoost, вона не може прогнозувати дані поза вибіркою, оскільки прогнозовані значення починають збігатися. Для глибинної нейронної мережі значення середньої абсолютної похибки склало 24.85, що є менше показника базового рівня на 4.2%.

### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Forecasting Solar Power Generation on the basis of Predictive and Corrective Maintenance Activities. URL: <https://arxiv.org/pdf/2205.08109>.
2. Neural Network Middle-Term Probabilistic Forecasting of Daily Power Consumption. URL: <https://arxiv.org/pdf/2006.16388>.
3. A Deep Learning Framework for Short-term Power Load Forecasting. URL: <https://arxiv.org/pdf/1711.11519>.
4. A Multi-Scheme Ensemble Using Cooperative Soft-Gating With Application to Power Forecasting for Renewable Energy Generation. URL: <https://arxiv.org/pdf/1803.06344>.
5. Transfer Learning in the Field of Renewable Energies -- A Transfer Learning Framework Providing Power Forecasts Throughout the Lifecycle of Wind Farms After Initial Connection to the Electrical Grid. URL: <https://arxiv.org/pdf/1906.01168>.
6. A Powerful Modelling Framework for Nowcasting and Forecasting COVID-19 and Other Diseases. URL: <https://arxiv.org/pdf/1912.05965>.

### METHODOLOGY FOR CREATING A HYBRID RECOMMENDATION SYSTEM TO SUPPORT UI/UX DESIGNERS

Vladyslav

Ph.D., prof. Caf. IS Kostiantyn Zashcholkin  
Odesa Polytechnic National University, UKRAINE

**ANNOTATION.** Forecasting electricity consumption is a very important task for energy companies and government regulators. Forecasting electricity consumption is an important tool for planning electricity production. Energy companies must provide sufficient amounts of electricity to meet demand at various levels of consumers, including residential, commercial and industrial. If the power company is unable to provide sufficient amounts of electricity to meet the demands of consumers, it may lead to power outages, which may lead to serious consequences for consumers.