

---

Курсова  
студента КН-3  
Телегіна Нікити

**ПРОГНОЗУВАННЯ  
ЕНЕРГОСПОЖИВАННЯ  
БУДІВЕЛЬ ЗА ДОПОМОГОЮ  
МАШИННОГО НАВЧАННЯ**

2025



Обстріл енергетики, 12 вересня  
2025

# МЕТА ДОСЛІДЖЕННЯ

1. розглянути теоретичні засади прогнозування часових рядів
2. Проаналізувати існуючу методологію прогнозування енергоспоживання.
3. створити власну модель, що зможе покращити попередні результати.

# ПОСТАНОВКА ЗАДАЧІ

На основі вибірки даних енергоспоживання певного міста, сформулювати погодинний прогноз енерговитрат на наступний місяць.



Чикаго, США (2011)

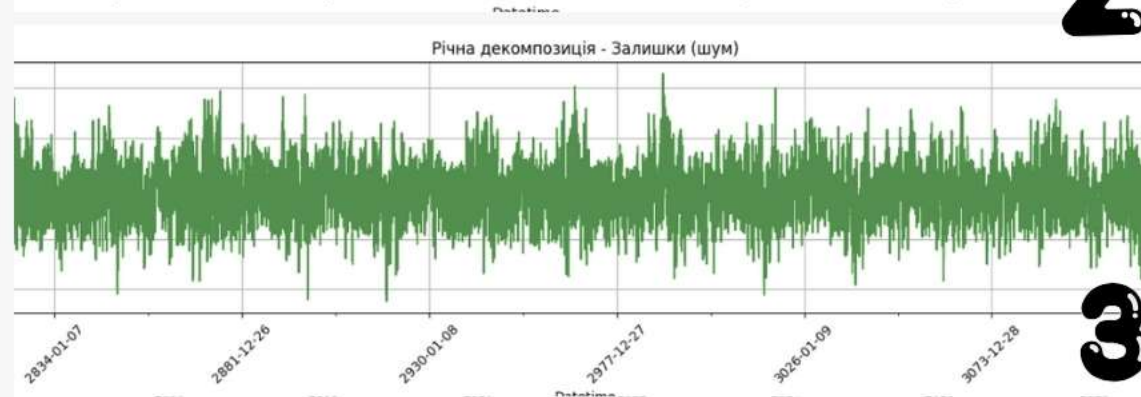
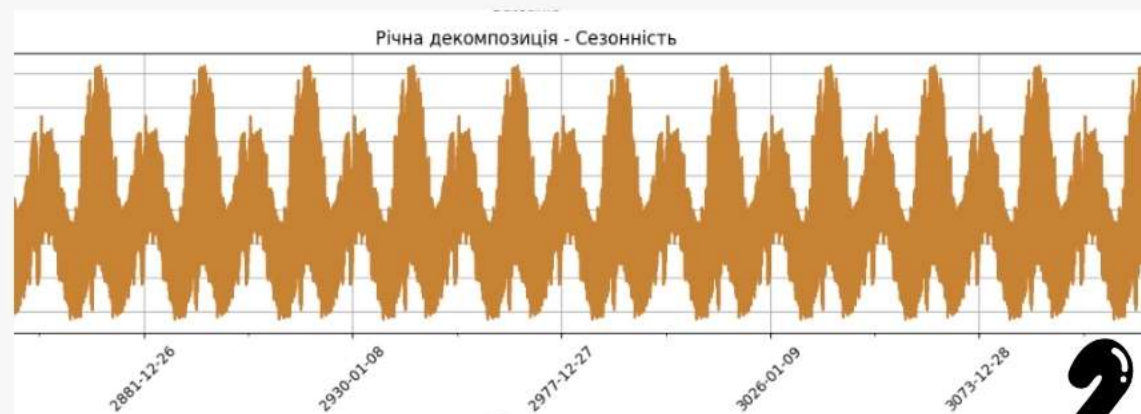
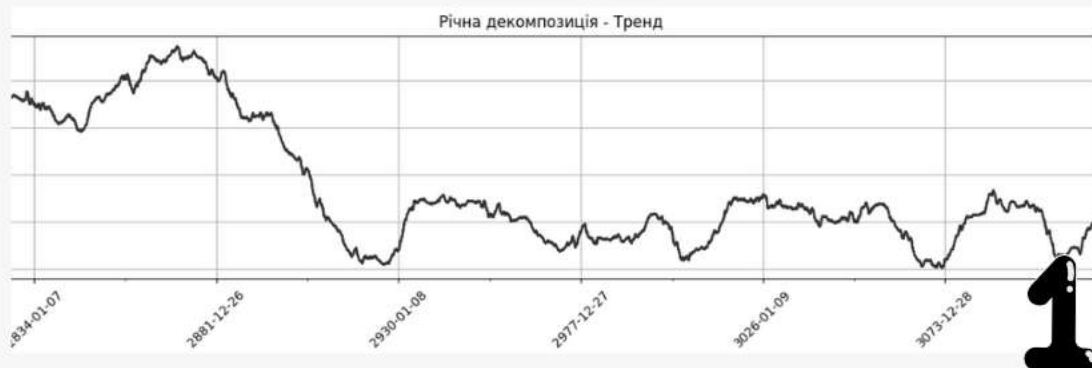
| 1  | Datetime, COMED_MW          |
|----|-----------------------------|
| 2  | 2011-12-31 01:00:00, 9970.0 |
| 3  | 2011-12-31 02:00:00, 9428.0 |
| 4  | 2011-12-31 03:00:00, 9059.0 |
| 5  | 2011-12-31 04:00:00, 8817.0 |
| 6  | 2011-12-31 05:00:00, 8743.0 |
| 7  | 2011-12-31 06:00:00, 8735.0 |
| 8  | 2011-12-31 07:00:00, 8993.0 |
| 9  | 2011-12-31 08:00:00, 9363.0 |
| 10 | 2011-12-31 09:00:00, 9545.0 |
| 11 | 2011-12-31 10:00:00, 9676.0 |
| 12 | 2011-12-31 11:00:00, 9937.0 |

# ПРОГНОЗУВАННЯ ЕНЕРГОСПОЖИВАННЯ БУДІВЕЛЬ ЯК ЧАСОВОГО РЯДУ

Часовий ряд - це ряд спостережень  $y_1, y_2, \dots, y_n$  випадкової величини ( $t$ ), утворених у послідовні моменти часу  $t_1, t_2, \dots, t_n$ .

## ОСНОВНІ КОМПОНЕНТИ ЧАСОВОГО РЯДУ:

1. **ТРЕНД** (ЗАГАЛЬНА ТЕНДЕНЦІЯ)
2. **СЕЗОННІСТЬ** (ПРИРОДНІ ЗАКОНОМІРНОСТІ)
3. **ЦИКЛІЧНІСТЬ** (ЛЮДСЬКІ ЦИКЛИ)
4. **ШУМ** (ВИКИДИ В ДАНИХ)

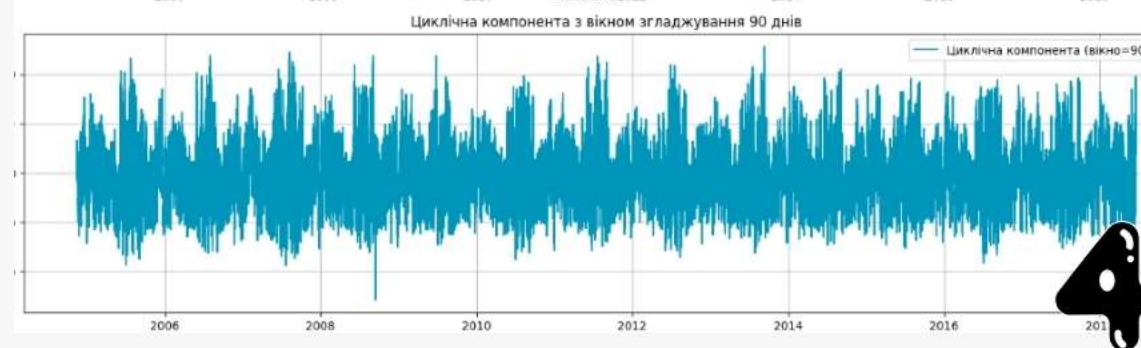
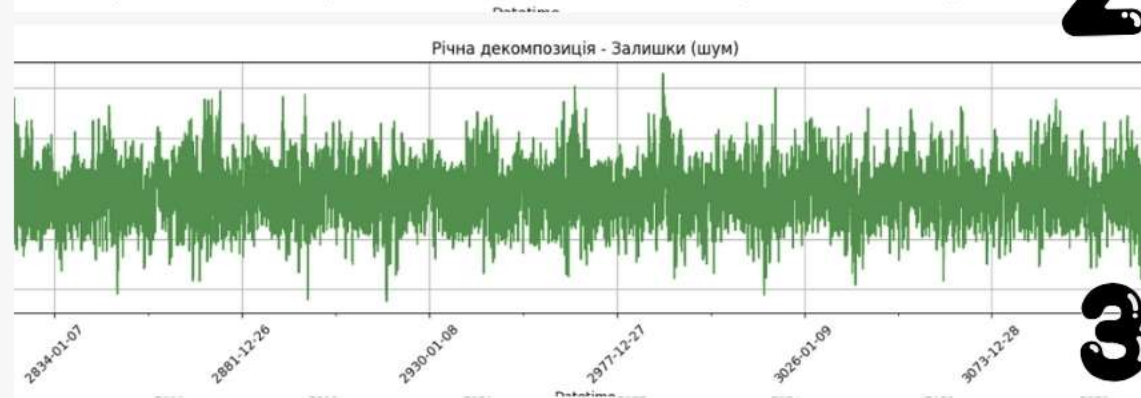
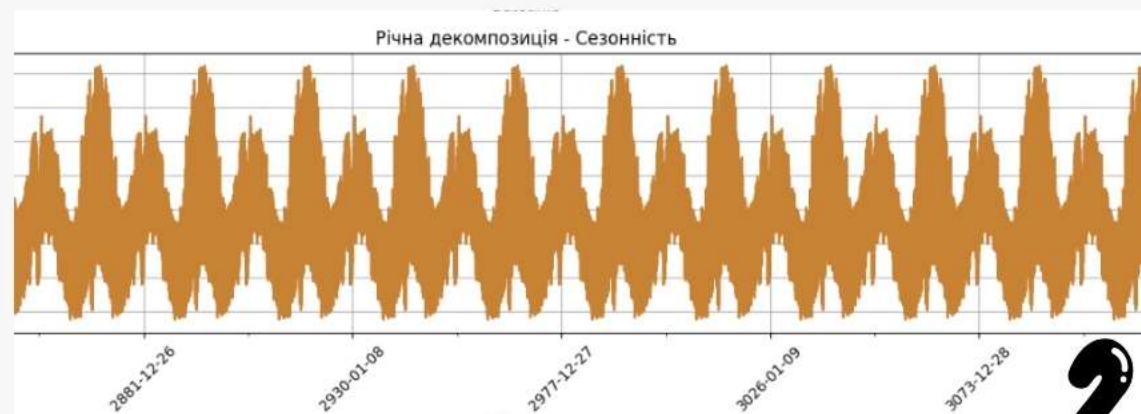
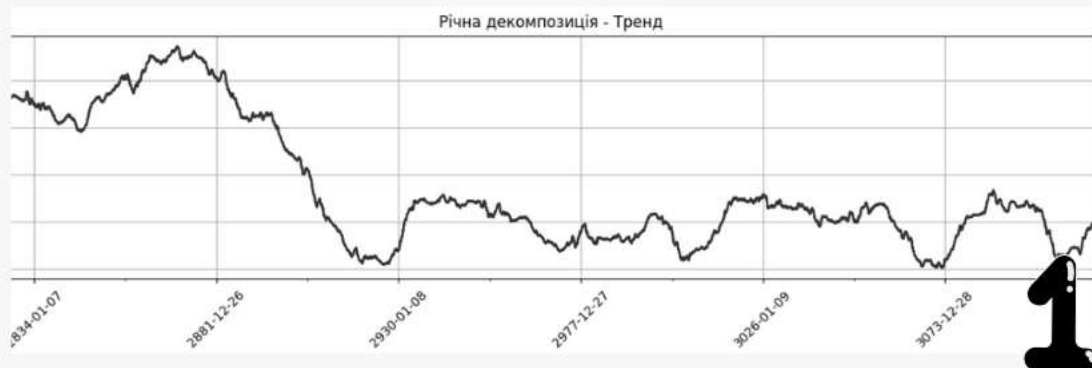


# ПРОГНОЗУВАННЯ ЕНЕРГОСПОЖИВАННЯ БУДІВЕЛЬ ЯК ЧАСОВОГО РЯДУ

Часовий ряд - це ряд спостережень  $y_1, y_2, \dots, y_n$  випадкової величини ( $t$ ), утворених у послідовні моменти часу  $t_1, t_2, \dots, t_n$ .

## ОСНОВНІ КОМПОНЕНТИ ЧАСОВОГО РЯДУ:

1. **ТРЕНД** (ЗАГАЛЬНА ТЕНДЕНЦІЯ)
2. **СЕЗОННІСТЬ** (ПРИРОДНІ ЗАКОНОМІРНОСТІ)
3. **ЦИКЛІЧНІСТЬ** (ЛЮДСЬКІ ЦИКЛИ)
4. **ШУМ** (ВИКИДИ В ДАНИХ)



# ЩО МОЖЕ ВПЛИНУТИ НА ЕНЕРГОСПОЖИВАННЯ МІСТА?

## ПРИРОДНІ УМОВИ

температура,  
вологість,  
сезонність клімату

## ЛЮДСЬКИЙ ВПЛИВ

робоча діяльність,  
забезпечення  
потреб

## ДЕРЖАВА

державні свята,  
регулювання роботи  
виробництв



Нічний Чикаго

## ЗОВНІШНІ ФАКТОРИ

## ЦІНА ЕНЕРГОРЕСУРСІВ

“  
Як тоді можна  
прогнозувати  
енергоспоживання?  
”



# МЕТОДОЛОГІЯ ПРОГНОЗУВАННЯ

## СТАТИСТИЧНІ МОДЕЛІ

Математичні моделі, які використовують статистичні припущення для прогнозування майбутніх значень часових рядів

**ПРИКЛАДИ МОДЕЛЕЙ:**  
ARIMA, SARIMA,  
SARIMAX

## МОДЕЛІ НА ОСНОВІ ДЕРЕВ РІШЕНЬ

Алгоритми машинного навчання, які приймають рішення шляхом послідовного розгалуження даних за певними правилами

**ПРИКЛАДИ МОДЕЛЕЙ:**  
RANDOM FOREST, XGBOOST,  
LIGHTGBM

## НЕЙРОННІ МОДЕЛІ

Обчислювальні системи, які складаються з нейронів, здатних навчатися розпізнавати складні патерни в даних.

**ПРИКЛАДИ МОДЕЛЕЙ:**  
GRU, LSTM



# СТАТИСТИЧНІ МОДЕЛІ

**ОБРАНІ МОДЕЛІ:** ARIMA, SARIMA, SARIMAX

## РЕЗУЛЬТАТИ

ARIMA і SARIMA мали лише 2 параметри, але через врахування сезонності SARIMA дає відчутне покращення. SARIMAX з 5 параметрами демонструє кращі результати, а найскладніша модель SARIMAX+ з кліматичними параметрами використовує 16 параметрів і досягає найвищої точності.

|                      | ARIMA  | SARIMA | SARIMAX | SARIMAX+ |
|----------------------|--------|--------|---------|----------|
| Кількість параметрів | 2      | 2      | 5       | 16       |
| MAPE                 | 14.88% | 12.56% | 11.80%  | 11.23%   |

**S** - SEASONAL

**A** - AUTO

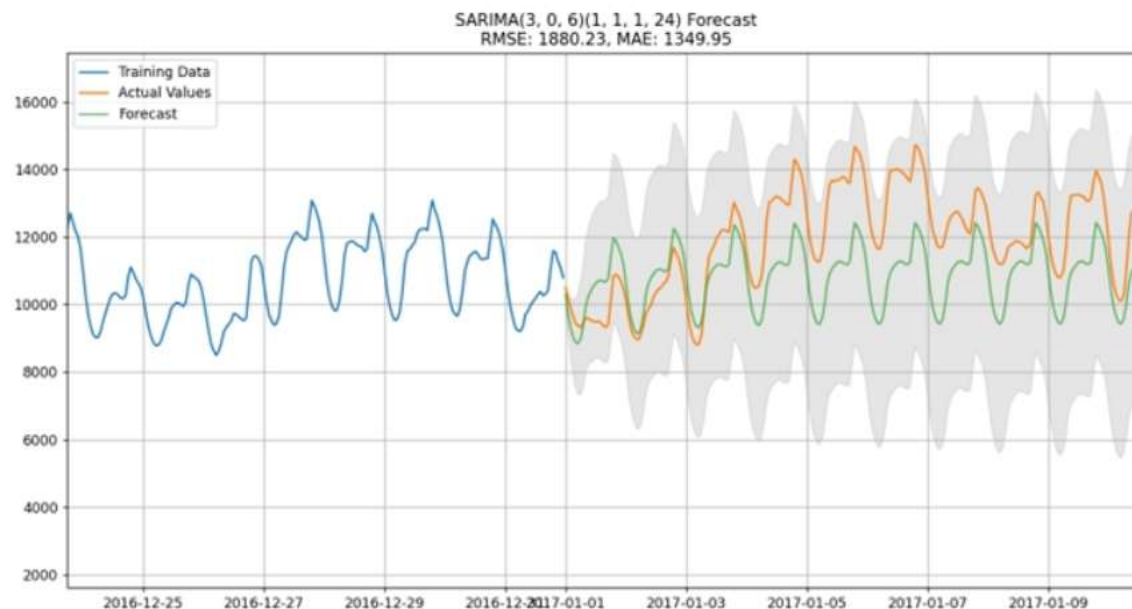
**R** - REGRESSIVE

**I** - INTEGRATED

**M** - MOVING

**A** - AVERAGE

**X** - WITH EXOGENOUS VARIABLES



# МОДЕЛІ НА ОСНОВІ ДЕРЕВ РІШЕНЬ

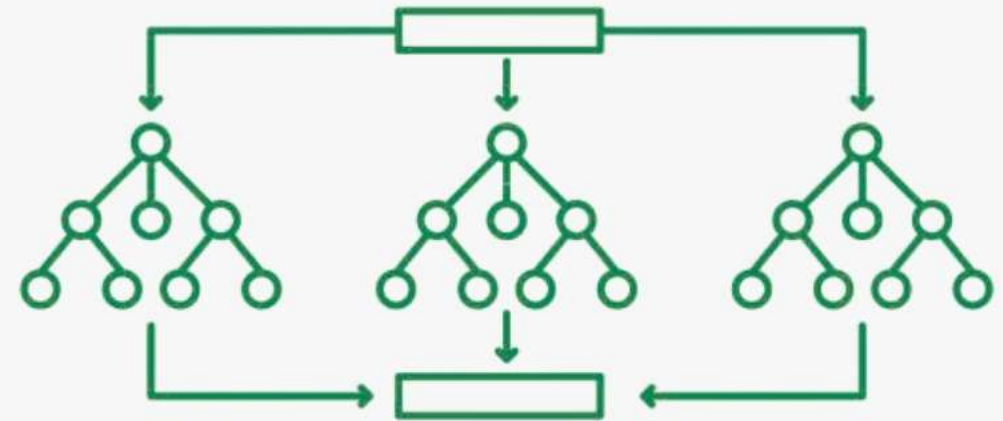
RANDOM FOREST, XGBOOST, LIGHTGBM

## РЕЗУЛЬТАТИ

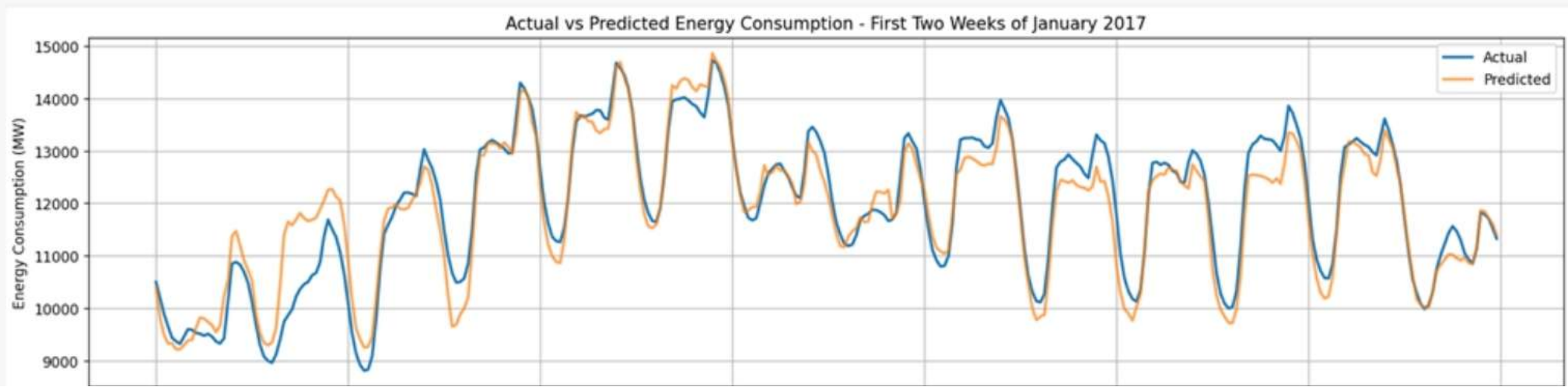
**Random Forest** - найнижчий  $R^2$  (0,8611),  
найкраща продуктивність взимку.

**XGBoost** - найкращий результат ( $R^2=0,8846$ ),  
найнижча похибка влітку й весною.

**LightGBM** - хороший збалансований  
результат: середній  $R^2$  (0,8705)



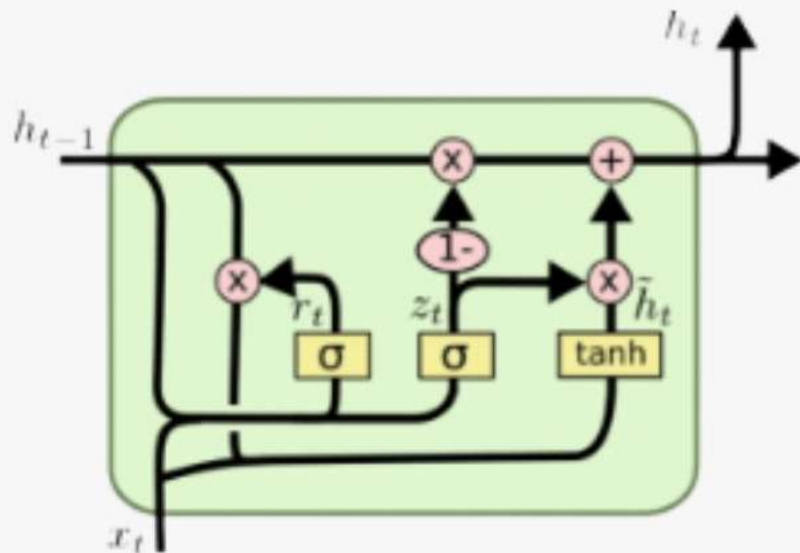
На рисунку результат роботи XGboost моделі



# НЕЙРОННІ МОДЕЛІ

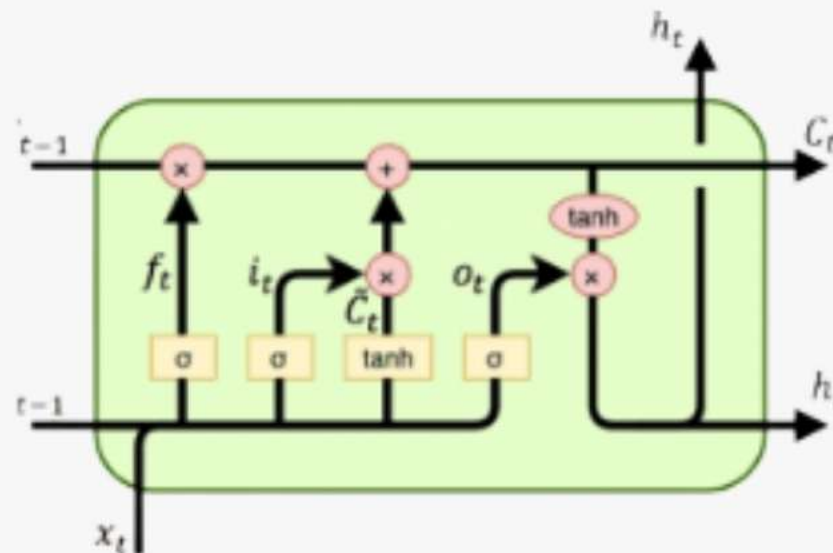
## GRU МОДЕЛЬ

рекурентна нейронна мережа зі спрощеною архітектурою воріт, що ефективно обробляє послідовні дані з меншою кількістю параметрів.



## LSTM МОДЕЛЬ

потужна рекурентна нейронна мережа з механізмом довготривалої пам'яті, здатна вивчати залежності в даних на великих часових проміжках.



На малюнках  
відповідні  
структури нейронів  
моделей.  
Джерело

# РЕЗУЛЬТАТИ

|             | Linear Regression | Random Forest  | XGBoost        | LightGBM       | GRU              | LSTM            |
|-------------|-------------------|----------------|----------------|----------------|------------------|-----------------|
| R2          | 0.7012            | 0.8611         | 0.8846         | 0.8705         | 0.7935           | 0.8293          |
| Summer RMSE | 1422.34 (11.29%)  | 983.60 (7.81%) | 621.46 (4.93%) | 949.38 (7.54%) | 1357.23 (10.78%) | 1214.23 (9.64%) |
| Spring RMSE | 897.15 (8.66%)    | 628.11 (6.06%) | 506.96 (4.89%) | 542.75 (5.24%) | 661.64 (6.39%)   | 612.41 (5.91%)  |
| Winter RMSE | 1011.26 (8.66%)   | 753.17 (6.45%) | 739.56 (6.33%) | 795.20 (6.81%) | 815.10 (6.98%)   | 762.27 (6.53%)  |

Найкраща модель - **XGBoost**,  
найслабша - звичайна **лінійна**  
**регресія**.

На основі отриманих результатів було вирішено  
побудувати **гібридну модель**, що поєднає сильні  
сторони XGBoost та LSTM моделей.

# ГІБРИДНА МОДЕЛЬ

Основа: XGBoost та LSTM моделі

## RIDGE РЕГРЕСІЯ

Для об'єднання прогнозів було застосовано мета-модель на основі Ridge регресії, що визначила оптимальні ваги для кожної моделі, мінімізуючи помилку прогнозу



## РЕАЛІЗОВАНІ ПОКРАЩЕННЯ:

- Двонаправлений LSTM (Bidirectional LSTM)
- Механізм уваги (Attention mechanism)
- Регуляризація L2
- Dropout (0.3)
- Рання зупинка (Early stopping)
- Оптимізація параметрів алгоритмом (Particle Swarm Optimization, PSO)

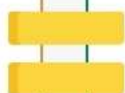
## LSTM МОДЕЛЬ

RMSE = 702.81 MW  
MAE = 497.08 MW  
 $R^2 = 0.8768$



## XGBOOST МОДЕЛЬ

RMSE = 647.28 MW  
MAE = 443.16 MW  
 $R^2 = 0.8955$



**GIBRID MODEL: RMSE = 615.36 MW,  
MAE = 424.82 MW,  $R^2 = 0.9055$ .**

Гібридна модель покращила всі метрики, зокрема зниження RMSE на 5% порівняно з XGBoost та на 12.4% порівняно з LSTM до покращень.

---

Потенційне

використання:

- «Розумний дім»
- Оптимізація електрогенерації
- Побудова превентивних механізмів втрат електрогенерації

**ДЯКУЮ ЗА  
УВАГУ**

Посилання на GitHub:

[https://github.com/nikita-ukma/buildings\\_consumption.git](https://github.com/nikita-ukma/buildings_consumption.git)

Джерела:

[RNN, LSTM & GRU](#)

---