

ОПТИМІЗАЦІЯ НАВЧАННЯ LSTM-МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ СОНЯЧНОЇ ГЕНЕРАЦІЇ З ВИКОРИСТАННЯМ БАЄСІВСЬКОГО ПІДХОДУ

Зінкевич П. О., Балюта С. М., Жуков Д. Ю., Корнієнко В. І.
Національний університет харчових технологій, Київ, Україна
E-mail: zinkevychpro@nuft.edu.ua

Куєвда Ю. В.
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна

Optimization of LSTM Network Training for Solar Generation Forecasting Using a Bayesian Approach

A methodology for training deep neural networks based on LSTM models using Bayesian optimization has been developed. Within the framework of the research, the optimization of the LSTM network training parameters was performed for the task of short-term forecasting of electricity generation by solar power plants.

Інтенсивна інтеграція штучного інтелекту та нейронних мереж у різні галузі діяльності, зокрема в енергетичний сектор, відкриває нові можливості для підвищення ефективності виробництва та управління енергією. Сонячна енергія є одним із провідних джерел відновлюваної енергії, що перетворюється в електроенергію за допомогою фотоелектричних станцій (ФЕС).

Одним із ключових викликів для енергетичного ринку України є забезпечення надійного короткострокового прогнозування генерації електроенергії. Ефективне короткострокове прогнозування дозволяє оптимізувати управління виробництвом, розподілом і зберіганням електроенергії, що сприяє підвищенню надійності та сталості функціонування енергетичної інфраструктури.

На сьогоднішній день у сфері прогнозування електричної енергії застосовуються фізичні, статистичні та інтелектуальні методи. Серед інтелектуальних підходів особливу популярність здобули моделі штучних нейронних мереж із довгою короткочасною пам'яттю (LSTM), які демонструють високу ефективність при роботі з часовими рядами [1].

Метою цього дослідження є навчання штучних нейронних мереж типу LSTM для прогнозування сонячної генерації з метою наближення до оптимальних гіперпараметрів та підбору варіантів навчання за допомогою байєсівської оптимізації. Побудова моделей прогнозування здійснювалася у середовищі MATLAB 2022b із використанням інструменту Experiment Manager.

Об'єктом дослідження є ФЕС в Житомирській області з встановленою потужністю 130 кВт. Навчання моделей проводилося з використанням погодинних вимірів кожного дня (24 виміри на добу) зафіксованих протягом

2024 року (8760 виміри). Дані вимірювань містять вхідні дані: сонячна інсоляція (кВт/м²), швидкість вітру (м/с) та температура навколишнього середовища; а також вихідні дані — потужність, електрична енергія, вироблена ФЕС (кВт*год).

Навчання моделі LSTM в середовищі MATLAB передбачає виконання декількох етапів: запуск інструменту Experiment Manager; створення функції навчання з навчальними даними; налаштування гіперпараметрів; налаштування параметрів байєсівської оптимізації; навчання моделі [2].

Першим етапом дослідження є запуск інструменту Experiment Manager. Experiment Manager завантажує проєкт, який складається з опису, таблиці гіперпараметрів, функції налаштування, набору метрик для оцінки результатів експерименту та додаткові параметри для обмеження тривалості експерименту.

Другим етапом дослідження є розробка функції навчання, яка працює з навчальними даними. Ця функція визначає складові, необхідні для проведення експерименту, зокрема: вхідні дані, архітектуру нейронної мережі, параметри та алгоритм процесу навчання. У рамках дослідження для навчання було використано оптимізаційні алгоритми *adam*, *rmsprop* та *sgdm*.

Третім етапом дослідження є налаштування гіперпараметрів моделі. Для досягнення високої продуктивності LSTM-мереж необхідна ретельна оптимізація широкого спектра гіперпараметрів. У цьому контексті байєсівська оптимізація виступає ефективною стратегією пошуку оптимальних комбінацій параметрів у межах експерименту. Гіперпараметри істотно впливають не лише на точність моделі, але й на обчислювальні витрати — зокрема, час навчання та споживання пам'яті. Інструмент *Experiment Manager* автоматично здійснює пошук найефективніших конфігурацій, використовуючи результати попередніх запусків для генерації нових комбінацій параметрів.

Для дослідження використовуються наступні гіперпараметри: *LSTMDepth*, який вказує на кількість шарів LSTM в мережі (діапазон від 1 до 3); *NumHiddenUnits*, що визначає кількість прихованих одиниць або кількість інформації, яка зберігається на кожному кроці часу (діапазон від 50 до 300); *InitialLearnRate*, який визначає початкову швидкість навчання (діапазон від 0,001 до 0,1).

Четвертим етапом дослідження є налаштування параметрів байєсівської оптимізації, де вказується тривалість експерименту, максимальний час (у секундах) і максимальна кількість спроб для запуску (3 спроби).

П'ятий етап дослідження полягає в запуску навчального експерименту для задачі регресії. У цьому процесі Experiment Manager автоматично виконує підбір найефективнішої комбінації гіперпараметрів відповідно до заданого критерію оптимізації. Кожна нова ітерація експерименту використовує оновлену комбінацію гіперпараметрів, сформовану на основі результатів попередніх запусків. Для оцінки якості моделі у кожному випробуванні обчислюється середньоквадратична похибка (RMSE). Після завершення експерименту Experiment Manager зберігає отримані результати з можливістю подальшого експорту до робочого простору MATLAB для аналізу, верифікації та тестування побудованої моделі. Результати дослідження наведено в табл. 1.

Таблиця 1. Оцінка продуктивності алгоритму навчання

Спроба для запуску	Алгоритм навчання	Гіперпараметр			Час навчання	RMSE
		LSTMDepth	NumHiddenUnits	InitialLearnRate		
1	adam	1	174	0,0079	00:07:56	2,82
2		3	58	0,0030	00:11:22	2,18
3		3	153	0,0098	00:14:46	2,82
1	rmsprop	3	60	0,0057	00:11:19	3,38
2		3	219	0,0074	00:13:22	5,04
3		2	54	0,0010	00:08:53	0,36
1	sgdm	1	279	0,0010	00:06:44	8,78
2		3	294	0,0073	00:13:22	13,9
3		1	123	0,0081	00:06:43	8,61

Аналіз результатів прогнозування (Табл. 1) показав, що найкращих результатів у контексті точності прогнозування вдалося досягти за допомогою алгоритму навчання *rmsprop*, при якому середньоквадратична помилка склала $RMSE = 0,36$, а час навчання — 00:08:53. Оптимальна конфігурація мережі передбачала використання двох шарів LSTM (гіперпараметр LSTMDepth), 54 прихованих одиниць (гіперпараметр NumHiddenUnits), що визначають обсяг інформації, яку мережа зберігає на кожному часовому кроці, а також початкову швидкість навчання (InitialLearnRate) = 0,001.

Зазначимо, що збільшення кількості прихованих одиниць може призводити до перенавчання моделі, збільшення часу її навчання та зростання похибки прогнозування. Надто низьке значення швидкості навчання спричиняє повільну збіжність, тоді як надто високе — може призвести до нестабільності процесу навчання або отримання неоптимальних результатів. Вибір оптимального значення швидкості навчання є залежним як від характеристик вхідних даних, так і від архітектури нейронної мережі.

У межах проведеного дослідження було виконано оптимізацію параметрів навчання LSTM-мережі для задачі короткострокового прогнозування генерації електричної енергії сонячними електростанціями. Результати експериментів продемонстрували, що алгоритм *rmsprop* забезпечує найвищу ефективність при прогнозуванні.

Література

1. Konstantinou M., Peratikou S., Charalambides A. G. (2021) Solar Photovoltaic Forecasting of Power Output Using LSTM Networks, *Atmosphere*, 12(1):124. URL : <https://doi.org/10.3390/atmos12010124>
2. Jozefowicz R., Zaremba W., Sutskever I. (2015) An Empirical Exploration of Recurrent Network Architectures, *Proceedings of the 32nd International Conference on Machine Learning*, 2342–2350.