

## Аналіз використання алгоритму оптимізації сірих вовків у гібридних нейронних мережах для фінансового прогнозування

Максим Жураковський, студент<sup>1</sup>(ORCID: 0009-0004-5024-8685)  
Сергій Позднякович, студент<sup>1</sup>(ORCID: 0009-0006-7765-8339)

<sup>1</sup> Київський національний університет будівництва і архітектури, Київ, Україна

### АНОТАЦІЯ

Робота присвячена аналізу ефективності застосування алгоритму оптимізації зграї сірих вовків при використанні гібридних нейронних мережах для фінансового прогнозування. На основі зробленого аналізу було проведено тестові експерименти та зроблено висновки про можливість такої системи.

*Ключові слова:* гібридні нейронні мережі, оптимізація, Grey Wolf Optimizer, прогнозування, фінанси, CNN, LSTM.

### 1. ВСТУП

Фінансове прогнозування є складною задачею, що вимагає використання ефективних методів машинного навчання та оптимізації. Гібридні нейронні мережі, що поєднують класичні архітектури з алгоритмами оптимізації, можуть підвищити точність прогнозів. Один із таких алгоритмів – алгоритм оптимізації сірих вовків (Grey Wolf Optimizer, GWO), який імітує поведінку зграї вовків під час полювання. Однак застосування GWO в гібридних нейронних мережах має певні труднощі, які потребують детального аналізу.

### 2. МЕТА РОБОТИ

Метою роботи є дослідження можливості використання алгоритму зграї сірих вовків для оптимізації прогнозування за допомогою гібридних нейромереж у фінансовій сфері; створення тестової системи для експериментів з даними.

### 3. АЛГОРИТМ ОПТИМІЗАЦІЇ GWO

Натхненням для алгоритму стали, як можна зрозуміти з його назви, сірі вовки. Ці тварини знаходяться на вершині харчового ланцюгу, віддають перевагу житті у зграї (від п'яти до дванадцяти особин в середньому) та мають соціальну домінуючу ієрархію, що найбільш приваблює дослідників.

Соціальна ієрархія для математичної моделі алгоритму використовується таким чином: найкраще рішення визначається, як альфа ( $\alpha$ ), друге та третє за якістю відповідно стають бета ( $\beta$ ) та дельта ( $\delta$ ), а всі інші – омегами ( $\omega$ ).

При полюванні вовки оточують свою жертву. Для математичного представлення цієї поведінки використовуються наступні рівняння:

$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)| \quad (1)$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D}, \quad (2)$$

де  $t$  – номер ітерації,  $\vec{A}$  та  $\vec{C}$  – вектори коефіцієнтів,  $\vec{X}_p$  – вектор позиції жертви, а  $\vec{X}$  – вектор позиції вовка.

Вектори  $\vec{A}$  та  $\vec{C}$  розраховуються наступним чином:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (3)$$

$$\vec{C} = 2 \cdot \vec{r}_2, \quad (4)$$

де  $\vec{a}$  лінійно зменшується від 2 до 0, а  $\vec{r}_1$  та  $\vec{r}_2$  – випадкові вектори в  $[0;1]$ .

Відповідно описаним рівнянням вовк може змінювати свою позицію ( $X$ ,  $Y$ ), щоб отримувати ідеальну для досягнення жертви ( $X^*$ ,  $Y^*$ ). При налаштуванні  $\vec{A}$  та  $\vec{C}$  вовк досягатиме різних місць для ліпшого полювання в колі, чий центр – жертва. Випадкові вектори дозволяють займати будь-яку позицію посеред цього кола.

В абстрактному просторі позиція жертви спочатку невідома, тому, для математичної стимуляції, дано, що в нас є три найліпші позиції для трьох вовків ( $\alpha$ ,  $\beta$ ,  $\delta$ ). Тепер останні агенти мають налаштувати свою позицію так, щоб знаходитися у відповідності з центром полювання.

Пошуковий агент оновлює свою позицію у відповідності до альфи, бети та дельти в  $n$ -вимірному просторі.

Коли жертва перестає рухатись, вовки її атакують. Для того, щоб змодельовати математично цю ситуацію, значення  $\vec{a}$  зменшується з часом.  $\vec{A}$  також зменшується за своєю формулою. Якщо узагальнити, він знаходиться у проміжку  $[-2a;2a]$ .

### 4. ГІБРИДНА НЕЙРОННА МЕРЕЖА CNN+LSTM

Гібридні нейронні мережі – система, яка складається з декількох різних типів нейронних мереж, генетичних алгоритмів, статистичних моделей або методів штучного інтелекту.

Згортова нейронна мережа (CNN) – це регуляризований тип прямої нейронної мережі, яка вивчає функції самостійно за допомогою оптимізації фільтра (або ядра). [1]

Згортова нейронна мережа складається з вхідного рівня, прихованих шарів і вихідного рівня. У CNN приховані шари включають одну або кілька частин, які виконують згортки. Зазвичай це включає шар, який виконує скалярний добуток ядра згортки на вхідну матрицю шару. Цей продукт зазвичай є внутрішнім продуктом Фробеніуса, а його функція активації зазвичай ReLU. Коли ядро згортки ковзає вздовж вхідної матриці для шару, операція згортки генерує карту ознак, яка, у свою чергу, сприяє введенню наступного шару. Після цього йдуть інші рівні, такі як шари об'єднання, повністю підключені шари та шари нормалізації [1] (рис. 1).

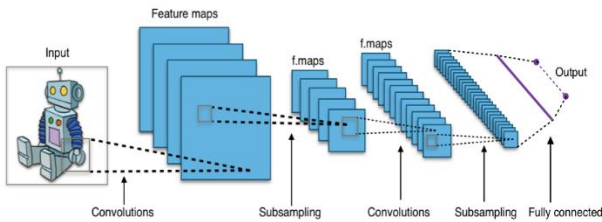


Рисунок 1 Типова архітектура CNN [2]

Long short-term memory (LSTM) – це тип рекурентної нейронної мережі (RNN), спрямований на пом'якшення проблеми зникнення градієнта, з якою зазвичай стикаються традиційні RNN. Його відносна нечутливість до довжини розриву є його перевагою перед іншими методами навчання послідовності. [3]

Блок LSTM зазвичай складається з осередка та трьох воріт: вхідних, вихідних і забуття. Осередок запам'ятовує значення протягом довільних інтервалів часу, а ворота регулюють потік інформації в клітинку та з неї. Ворота забуття вирішують, яку інформацію відкинути з попереднього стану, зіставляючи попередній стан і поточний вхід зі значенням від 0 до 1. Округлене значення до 1 означає збереження інформації, а значення 0 означає відкидання. Вхідні ворота вирішують, які фрагменти нової інформації зберігати в поточному стані комірки, використовуючи ту саму систему, що й ворота забуття. Вихідні ворота контролюють, які частини інформації в поточному стані комірки виводити, призначаючи значення від 0 до 1 інформації, враховуючи попередній і поточний стани (рис. 2). [3]

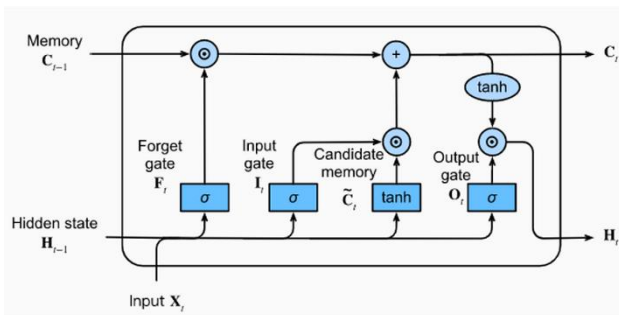


Рисунок 2 Типова архітектура LSTM [4]

Згорткові нейронні мережі використовуються для отримання ознак з часових рядів, а потім LSTM обробляє залежності між ними. CNN допомагають виявити локальні закономірності даних, такі як патерни зміни цін, а LSTM моделі враховують довгострокові залежності.

Гібрид добре справляється із спотвореними фінансовими даними та може використовувати історичні дані для захоплення важливих закономірностей.

## 5. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

Для навчання моделей був обраний набір даних часового ряду з щоденними акціями [5]. Він був перевірений на наявність нульових значень та очищений від викидів.

Потім було створено та розглянуто результати виконання прогнозів чистим гібридом нейромереж, без використання оптимізації. Тренування моделі проходило на даних, розбитих у пропорції 4:1 на тренувальні та тестові.

Після навчання протягом 10 епох було отримано значення loss при запуску моделі на тестових даних та зроблено прогнозування.

Робота з оптимізованим за допомогою алгоритму GWO гібридом проходила також само, але спочатку було знайдено найліпші гіперпараметри для налаштування нейромереж за допомогою цього самого алгоритму.

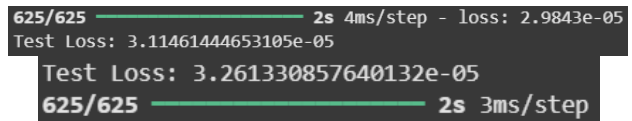


Рисунок 3 Значення Loss для чистого та оптимізованого гібрида

Після проведення експериментів для чистої гібридної моделі та оптимізованої за допомогою GWO, можна побачити, що було отримано краще значення loss для другої моделі, але потрачений час на оптимізацію та навчання є завеликим, з чого можна зробити висновок, що розглянута комбінація оптимізації та гібрида показує гарний результат, але її використання при динаміці змін фінансових даних є скрутним.

## 6. ВИСНОВКИ

- 1) Було створено гібридну нейронну мережу, проведено аналіз методів оцінки якості оптимізації гібридних нейронних мереж, описано та використано алгоритм оптимізації GWO, за допомогою якого налаштовано гіперпараметри для нейромережі.
- 2) Були зібрані та оброблені дані для навчання та отримано прогнози за цими даними.
- 3) Було порівняно чисту гібридну модель з її оптимізованим варіантом та зроблено висновки.

## Список літератури

- [1] Матеріал з вільного сайту «Wikipedia» [Електронний ресурс] Режим доступу: [https://en.wikipedia.org/wiki/Convolutional\\_neural\\_network\\_k](https://en.wikipedia.org/wiki/Convolutional_neural_network_k)
- [2] Boonyuen, K., Kaewprapha, P., Weesakul, U., Srivihok, P. (2019). Convolutional Neural Network Inception-v3: A Machine Learning Approach for Leveling Short-Range Rainfall Forecast Model from Satellite Image. In: Tan, Y., Shi, Y., Niu, B. (eds) Advances in Swarm Intelligence. ICSI 2019. Lecture Notes in Computer Science(), vol 11656. Springer, Cham.
- [3] Матеріал з вільного сайту «Wikipedia» [Електронний ресурс] Режим доступу: [https://en.wikipedia.org/wiki/Long\\_short-term\\_memory](https://en.wikipedia.org/wiki/Long_short-term_memory)
- [4] Матеріал з вільного сайту «Dive Into Deep Learning» [Електронний ресурс] Режим доступу: [https://d2l.ai/chapter\\_recurrent-modern/lstm.html](https://d2l.ai/chapter_recurrent-modern/lstm.html)
- [5] Матеріал з вільного сайту «Hugging Face» [Електронний ресурс] Режим доступу: <https://huggingface.co/datasets/chuyin0321/timeseries-daily-stocks>

і Робота виконана під керівництвом канд. техн. наук., доц., доц. кафедри ІТ Олени Горди