



**PROCEEDINGS OF THE  
I INTERNATIONAL SCIENTIFIC  
AND THEORETICAL CONFERENCE**

SCIENTIFIC REVIEW OF  
THE ACTUAL EVENTS,  
ACHIEVEMENTS AND  
PROBLEMS

01.12.2023

BERLIN

FEDERAL REPUBLIC OF GERMANY

*with the proceedings of the*

I International Scientific and Theoretical Conference

**Scientific review of the  
actual events, achievements  
and problems**

**01.12.2023**

Berlin, Federal Republic of Germany

**Berlin, 2023**

## SECTION 24.

### PHARMACY AND PHARMACOTHERAPY

**Дорош Анна Сергіївна**

здобувач вищої освіти фармацевтичного факультету  
Національний медичний університет імені О.О. Богомольця, Україна

**Науковий керівник: Пушкарьова Ярослава Миколаївна** 

канд. хім. наук, доцент, доцент кафедри аналітичної, фізичної та колоїдної хімії  
Національний медичний університет імені О.О. Богомольця, Україна

## ПРОГНОЗУВАННЯ ЛІПОФІЛЬНОСТІ МОЛЕКУЛ ЛІКАРСЬКИХ ЗАСОБІВ ЗА ДОПОМОГОЮ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ

***Анотація.** Ліпофільність обов'язково враховується та оцінюється на етапі моделювання нових лікарських засобів, оскільки є однією із складових успішного застосування досліджуваної молекули як потенційного лікарського препарату. Досліджено можливість застосування штучної нейронної мережі прямого поширення сигналу та каскадної нейронної мережі для прогнозування ліпофільності молекул лікарських засобів за різними наборами молекулярних дескрипторів. Показано, що алгоритми штучних нейронних мереж є ефективними засобами вирішення такого роду задач прогнозування.*

**Вступ.** Відкриття ліків – це складний і важкий процес. Щоб визначити, чи буде запропонована молекула успішним лікарським засобом, необхідно оцінити її ефективність і можливість лікування певних захворювань. На початковому етапі експериментальних досліджень із виявлення нових потенційних лікарських засобів необхідно перевірити велику кількість хімічних молекул, щоб виявити сполуки, які демонструють певну біологічну активність [1–3].

Ліпофільність важлива для прогнозування успіху молекули як лікарського препарату, оскільки є показником метаболізму, фармакокінетики та фармакодинаміки [1]. Експериментальне визначення ліпофільності вимагає кількісного визначення концентрації сполуки, присутньої у неводній та водній фазах. Є різні підходи щодо визначення ліпофільності, але, зазвичай, вони довготривалі та трудомісткі, кожен зі своїми недоліками та перевагами [4, 5].

**Мета дослідження** – оцінити ефективність алгоритмів штучної нейронної мережі прямого поширення сигналу та каскадної нейронної мережі щодо прогнозування ліпофільності молекул лікарських засобів із застосуванням різних наборів фізико-хімічних параметрів.

**Методи дослідження.** Реалізацію штучних нейронних мереж виконано із застосуванням програмного комплексу Matlab R2022b.

Штучні нейронні мережі прямого поширення сигналу, як і будь-яка інша штучна нейронна мережа, складається з шарів нейронів. Сигнали (тобто вхідні вектори чи вхідна інформація) надходять на вхідний рівень нейронів. Потім вони надходять та обробляються нейронами прихованого шару. Виходи (кінцевий результат) надають нейрони вихідного

шару. Таким чином, сигнали рухаються тільки в прямому напрямку. Кожен наступний шар має з'єднання з попереднім шаром [6–8].

Каскадна нейрона мережа є різновидом мережі прямого поширення сигналу, але з однією відмінністю. Каскадна мережа має додаткові підключення від входу до кожного рівня та від кожного рівня до всіх наступних рівнів [6–8].

**Масиви даних.** У даній роботі для оцінки ефективності алгоритмів штучних нейронних мереж щодо прогнозування ліпофільності молекул лікарських засобів використали масив даних [9], що містить 34 лікарські сполуки, що характеризуються такими фізичними та фізико-хімічними параметрами: енергія нульової точки; теплова енергія; ентальпія; вільна енергія; енергія активації; теплоємність; ентропія; енергія Гартрі-Фока; стеричний фактор; найвища зайнята молекулярна орбітальна енергія; найнижча незайнята молекулярна орбітальна енергія; молекулярна рефракція; коефіцієнт розподілу. Також для 34 досліджених лікарських засобів відомі експериментально визначені значення ліпофільності ( $\log P$ ) [9].

**Результати.** Табл. 1 містить інформацію щодо архітектури реалізованих у даній роботі алгоритмів штучних нейронних мереж [10].

Таблиця 1

**Параметри навчання нейронної мережі прямого поширення сигналу та каскадної нейронної мережі**

Тип параметру	Опис
Метод навчання	Левенберга-Марквардта, алгоритм градієнтної оптимізації
Мінімізована функція похибки	Середня квадратична похибка
Кількість прихованих нейронів	n–m
Функції активації	Сигмоїдальна / лінійна
Метод ініціалізації	Алгоритм Нгуена-Відроу

Для кожного нижче описаного набору дескрипторів знаходилося оптимальне число прихованих нейронів та оцінювалася прогностична сила нейронних мереж. Оптимальне число прихованих нейронів оцінювалося за допомогою навчання нейронних мереж на усіх 34 лікарських сполуках та оцінкою середньоквадратичної похибки (оптимальне значення прихованих нейронів це число, при якому отримане найменше значення середньоквадратичної похибки). Тобто оцінювали як якісно нейронні мережі «вивчили» подану інформацію та відтворили її. Для оцінки прогностичної сили досліджуваних лікарських сполуки випадковим чином поділили на навчальну вибірку (75%) та тестову вибірку (25%).

1) *Прогнозування ліпофільності із застосуванням 6 дескрипторів (енергія нульової точки; теплова енергія; ентальпія; вільна енергія; енергія активації; теплоємність)*

Встановлено, що оптимальна кількість нейронів прихованого шару для нейронної мережі прямого поширення сигналу складає 9, а для каскадної нейронної мережі – 7.

Рис. 1 та 2 представляють кореляцію між прогнозованими значеннями ліпофільності та експериментальними значеннями ліпофільності у випадку навчання нейронних мереж зі встановленим оптимальним значенням прихованих нейронів: для нейронної мережі прямого поширення сигналу  $R = 0,965$  та для каскадної мережі  $R = 0,962$ .

При оцінці прогностичної сили запропонованих моделей отримані такі результати: для нейронної мережі прямого поширення сигналу  $R = 0,958$ ; для каскадної нейронної мережі  $R = 0,874$ .

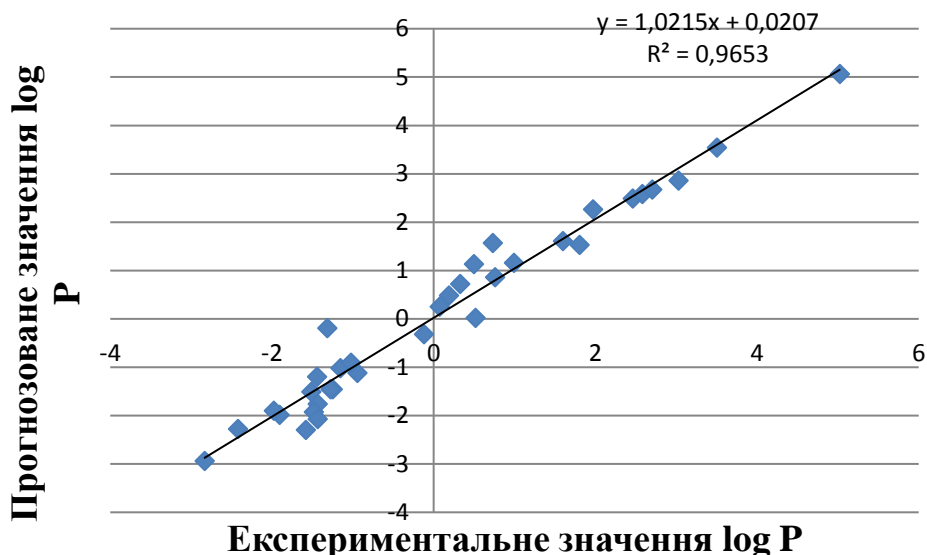


Рис. 1. Кореляція між прогнозованими та експериментальними значеннями ліпофільності для нейронної мережі прямого поширення сигналу для усіх 34 лікарських сполук із застосуванням 6 дескрипторів

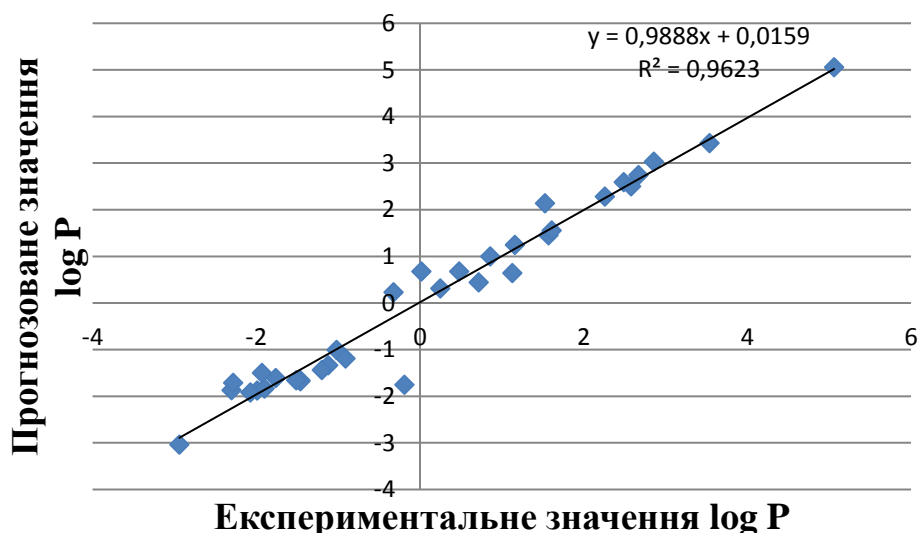


Рис. 2. Кореляція між прогнозованими та експериментальними значеннями ліпофільності для каскадної нейронної мережі для усіх 34 лікарських сполук із застосуванням 6 дескрипторів

2) Прогнозування ліпофільності із застосуванням 7 дескрипторів (ентропія; енергія Гартрі-Фока; стеричний фактор; найвища зайнята молекулярна орбітальна енергія; найнижча незайнята молекулярна орбітальна енергія; молекулярна рефракція; коефіцієнт розподілу)

Встановлено, що оптимальна кількість нейронів прихованого шару для нейронної мережі прямого поширення сигналу складає 4, а для каскадної нейронної мережі – 7.

Рис. 3 та 4 представляють кореляцію між прогнозованими значеннями ліпофільності та експериментальними значеннями ліпофільності у випадку навчання нейронних мереж із встановленим оптимальним значенням прихованих нейронів: для нейронної мережі прямого поширення сигналу  $R = 0,985$  та для каскадної мережі  $R = 0,980$ .

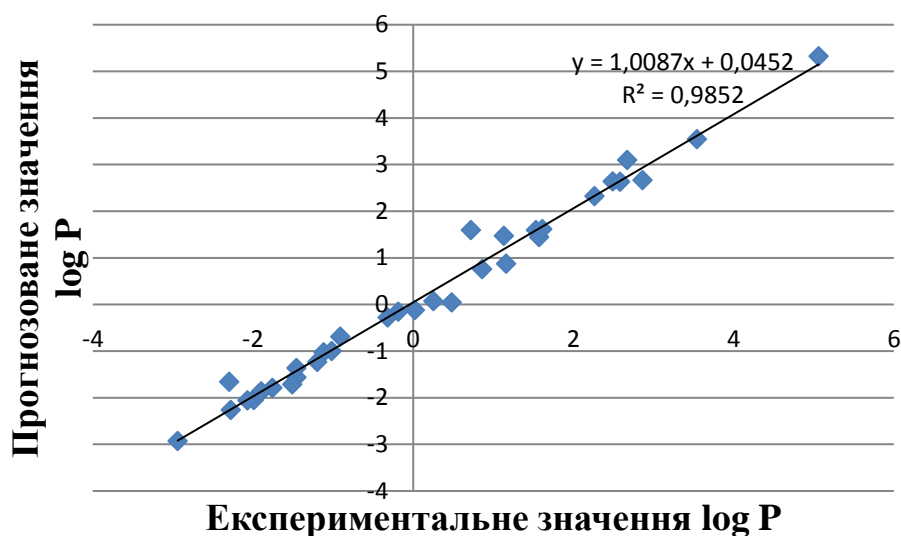


Рис. 3. Кореляція між прогнозованими та експериментальними значеннями ліпофільності для нейронної мережі прямого поширення сигналу для усіх 34 лікарських сполук із застосуванням 7 дескрипторів

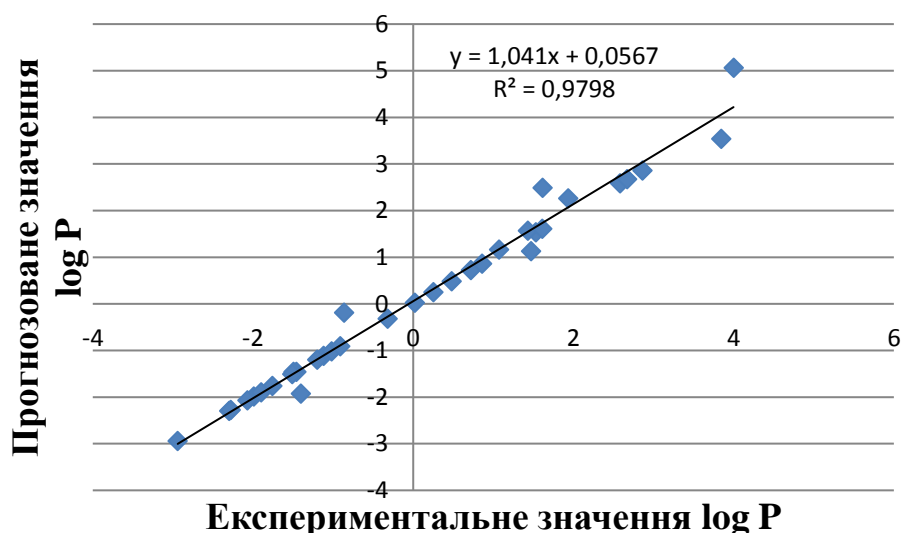


Рис. 4. Кореляція між прогнозованими та експериментальними значеннями ліпофільності для каскадної нейронної мережі для усіх 34 лікарських сполук із застосуванням 7 дескрипторів

При оцінці прогностичної сили запропонованих моделей отримані такі результати: для нейронної мережі прямого поширення сигналу  $R = 0,983$ ; для каскадної нейронної мережі  $R = 0,982$ .

3) Прогнозування ліпофільності із застосуванням 2 дескрипторів (коефіцієнту розподілу та енергії нульової точки)

Встановлено, що оптимальна кількість нейронів прихованого шару для нейронної мережі прямого поширення сигналу складає 5, а для каскадної нейронної мережі – 6.

Рис. 5 та 6 представляють кореляцію між прогнозованими значеннями ліпофільності та експериментальними значеннями ліпофільності у випадку навчання нейронних мереж із встановленим оптимальним значенням прихованих нейронів: для нейронної мережі прямого поширення сигналу  $R = 0,988$  та для каскадної мережі  $R = 0,992$ .

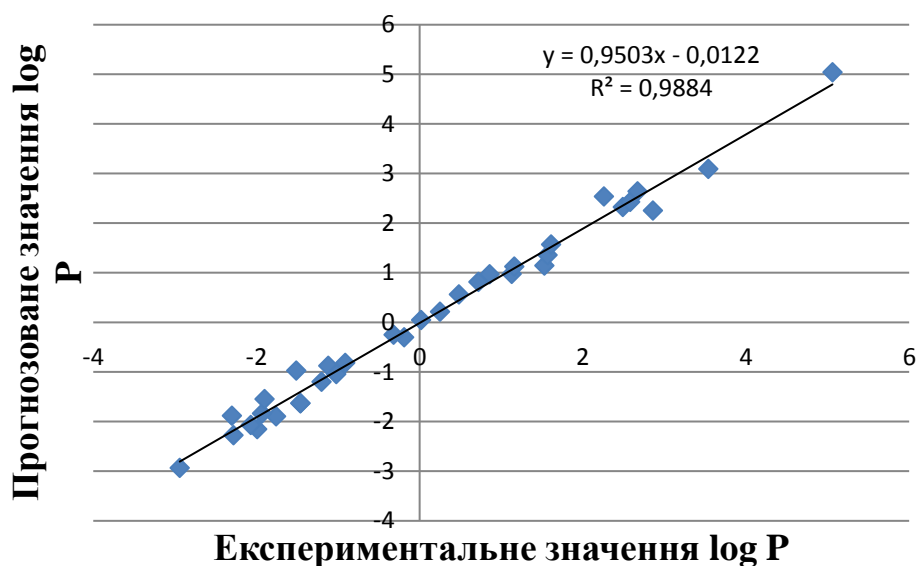


Рис. 5. Кореляція між прогнозованими та експериментальними значеннями ліпофільності для нейронної мережі прямого поширення сигналу для усіх 34 лікарських сполук із застосуванням 2 дескрипторів

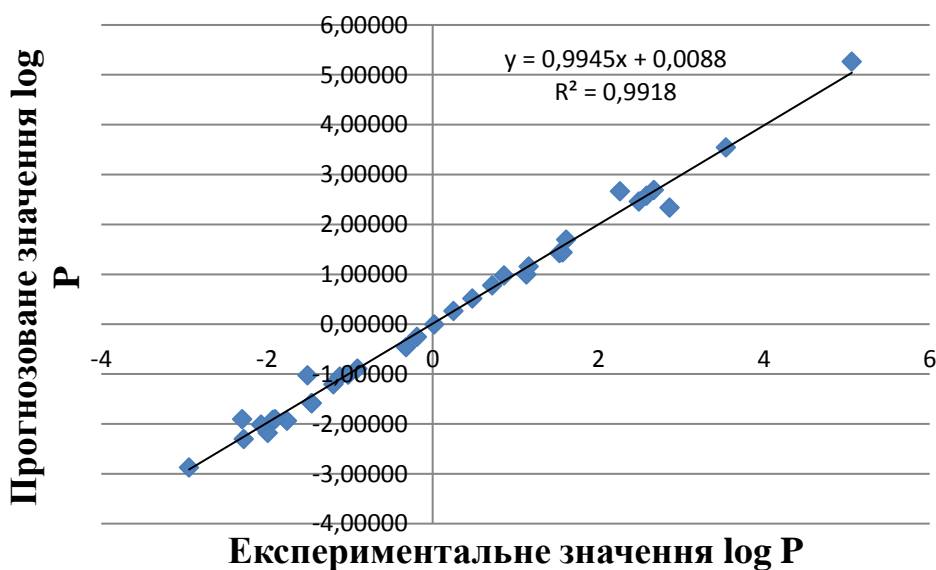


Рис. 6. Кореляція між прогнозованими та експериментальними значеннями ліпофільності для каскадної нейронної мережі для усіх 34 лікарських сполук із застосуванням 2 дескрипторів

При оцінці прогностичної сили запропонованих моделей отримані такі результати: для нейронної мережі прямого поширення сигналу  $R = 0,988$ ; для каскадної нейронної мережі  $R = 0,991$ .

**Висновки.** Алгоритми штучної нейронної мережі прямого поширення сигналу та каскадної нейронної мережі характеризуються високою ефективністю прогнозування ліпофільності молекул лікарських засобів із застосуванням різних наборів фізико-хімічних параметрів. Проведене дослідження прогнозування ліпофільності дозволяє зробити висновок, що більш точні результати отримані із застосуванням каскадної нейронної мережі та лише двох дескрипторів (коефіцієнту розподілу та енергії нульової точки).

**Список використаних джерел:**

1. Datta, R., Das, D., & Das, S. (2021). Efficient lipophilicity prediction of molecules employing deep-learning models. *Chemometrics and Intelligent Laboratory Systems*, 213, 104309.
2. Ghosh, S., Nie, A., An, J., & Huang, Z. (2006). Structure-based virtual screening of chemical libraries for drug discovery. *Current opinion in chemical biology*, 10(3), 194-202.
3. Walters, W. P. (2018). Virtual chemical libraries: miniperspective. *Journal of medicinal chemistry*, 62(3), 1116-1124.
4. OECD. (1995). Partition Coefficient (n-Octanol/Water), Shake Flask Method, No. 107. *OCED*, 1(July), 1-4.
5. Hartmann, T., & Schmitt, J. (2004). Lipophilicity–beyond octanol/water: a short comparison of modern technologies. *Drug Discovery Today: Technologies*, 1(4), 431-439.
6. Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis*, 22(5), 717-727.
7. Mutihac, L., & Mutihac, R. (2008). Mining in chemometrics. *Analytica Chimica Acta*, 612(1), 1-18.
8. Pushkarova, Y., Zaitseva, G., & Al Saker, M. (2022, September). Prediction of Toxicity of Phenols Using Artificial Neural Networks. In *2022 12th International Conference on Advanced Computer Information Technologies (ACIT)* (pp. 493-496). IEEE.
9. Ibrahim, A. A., Abd-Alrazzak, A. Y., AbdalRazaq, E. A., Sulliman, E. A., & Shamil, T. (2020). Theoretical Prediction of Lipophilicity for Some Drugs Compounds. *Oriental J. of Chem*, 36(1), 114-119.
10. Pushkarova, Y., & Kholin, Y. (2012). The classification of solvents based on solvatochromic characteristics: the choice of optimal parameters for artificial neural networks. *Open Chemistry*, 10(4), 1318-1327.