

ПОРІВНЯЛЬНИЙ АНАЛІЗ РЕЗУЛЬТАТІВ ЗАСТОСУВАННЯ РІЗНИХ АРХІТЕКТУР НЕЙРОННИХ МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ ІНФЕКЦІЙНИХ ЗАХВОРЮВАНЬ

У даній роботі представлено результати прогнозування часових рядів із використанням методів машинного навчання та мобільних технологій. Основними результатами цієї роботи є порівняння ефективності різних архітектур нейронних мереж у рамках прогнозування поширення інфекційних захворювань в Україні за період з грудня 2016 року по січень 2024 року.

Для досягнення мети було вирішено такі завдання: проаналізовано поточний стан прогнозування часових рядів; проаналізовано існуючі аналоги систем; обрано необхідні архітектури нейронних мереж, як один із методів машинного навчання; проаналізовано та нормалізовано набір даних про поширення інфекційних захворювань; проведено тестування. Розроблена система може бути ефективним інструментом для прийняття раціональних управлінських рішень щодо забезпечення епідеміологічного благополуччя та біобезпеки населення.

Ключові слова: Прогнозування, часові ряди, методи машинного навчання, нейронні мережі.

Серед надзвичайних ситуацій особливий негативний вплив на здоров'я людини надають війни та військові конфлікти. Війни завжди супроводжуються спалахами інфекційних хвороб, підвищенням рівня захворюваності та смертності від інфекцій. Історія війн продемонструвала, що втрати людських життів від інфекцій часто перевищували втрати від бойових травм. Незважаючи на досягнення сучасної медицини, інфекції продовжують супроводжувати війни у 21 столітті. Тому є важливою задача прогнозування розповсюдження інфекційних захворювань для можливості кращої підготовки до можливих спалахів захворювань [1].

У ході дослідження було проведено систематичний збір даних про поширення інфекційних захворювань в Україні з грудня 2016 року по січень 2024 року. Щомісячні звіти отримували з національних баз даних охорони здоров'я, а саме Центру громадського здоров'я МОЗ України. Дані, що збирає Центр громадського здоров'я, наведено згідно зі звітною формою №1; у звітах наведено абсолютні цифри та інтенсивні показники на 100 000 населення. Дані наведено у порівнянні за певний проміжок часу за поточний і минулий роки, а також за повний рік (12 місяців) [2]. Для врахування коливань чисельності населення протягом досліджуваного періоду дані були нормалізовані за оцінками чисельності населення Державної служби статистики України. Абсолютні значення були нормалізовані для відображення кількості захворювань за поточний місяць на 1000000 населення України за поточний рік. Для нормалізації кількості захворювань було використано наступну формулу:

$$\frac{\text{абс. кількість захворювань за поточний місяць} * 1\,000\,000}{\text{кількість населення за поточний рік}} \quad (1)$$

Ця нормалізація дозволила провести аналіз на 1000000 населення, полегшивши порівняння за різні місяці та роки. У випадках, коли з'являлися дублікати звітів або суперечливі дані, застосовувався консенсусний підхід, віддаючи пріоритет місяцям із найбільшою кількістю зареєстрованих випадків, щоб забезпечити точність і надійність набору даних. Ця ретельна обробка даних була важливою для забезпечення чіткого та точного представлення тенденцій інфекційних захворювань в Україні протягом зазначеного періоду часу. Всього в наборі даних присутні 57 інфекційних хвороб, зокрема холера, черевний тиф, паратиф А, В, С, шигельоз та інші.

Було побудовано та проаналізовано графіки захворювань з нормалізованими значеннями. Нижче наведено приклад графіків захворювань (рисунк 1).

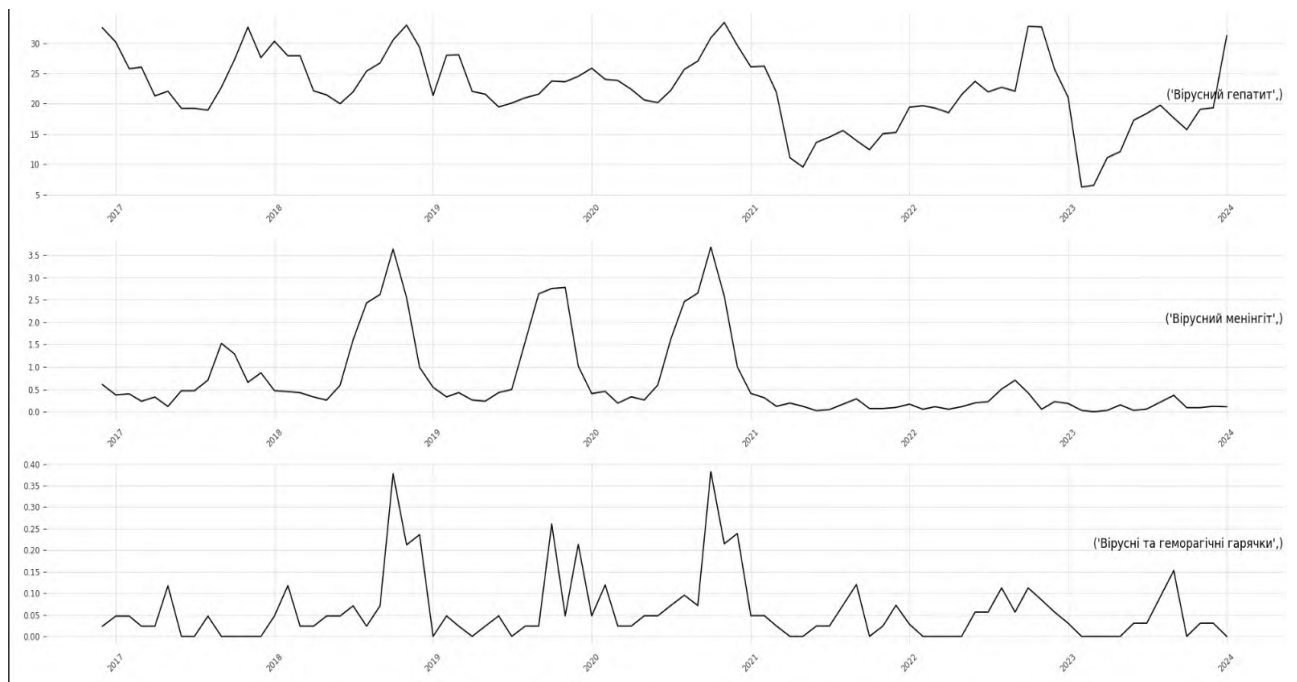


Рис. 1. Графіки розповсюдження інфекційних захворювань згідно датасету.
Джерело: сформовано автором на основі [2]

Серед 57 вибраних інфекційних захворювань спостерігалось збільшення кількості випадків у 22 випадках після початку війни в Україні (24.02.2022), наприклад, випадки лямбліозу із середнім показником до війни 4,53 збільшилися до 10,06, що становить збільшення на 121,91%, кишкові протозойні захворювання із середнім показником до війни 5,06 збільшилися до 11,19, що становить збільшення на 120,9%.

Загалом спостерігалось зменшення кількості випадків у 35 інфекційних захворюваннях із середнім зменшенням кількості випадків на 35,52%, та у 22 захворюваннях із середнім збільшенням кількості випадків на 62,38%.

З огляду на ці цифри, вплив надзвичайних ситуацій очевидний. Масштабна агресія Росії в Україні призвела до значної надзвичайної ситуації, яка вплинула на всі аспекти життя та спричинила додаткові надзвичайні ситуації, такі як повені та посухи внаслідок пошкодження Каховської гідроелектростанції. Інші ризики включають потенційне забруднення радіонуклідами внаслідок аварій на Чорнобильській та Запорізькій АЕС, лісові пожежі на Кінбурнській косі, які спустошили понад 1500 гектарів лісу, та викиди хімічних речовин, включаючи хлор та аміак. У контексті цих надзвичайних ситуацій динаміка та прояви епідемій інфекційних захворювань можуть відрізнятися від їх природного розвитку. Ключовими факторами ризику, що впливають на епідемічну ситуацію на тлі каскаду надзвичайних ситуацій, спричинених війною, є значна міграція населення, переповненість бомбосховищ та міграційних шляхів, підвищений рівень стресу, що призводить до підвищеної вразливості до інфекцій, порушення водо- та енергопостачання, різке зростання популяцій гризунів та пов'язані з цим епізоотії, забруднення харчових продуктів, стікання різних хімічних речовин у водойми, затоплення природних екосистем, активація механізмів передачі інфекційних агентів, збільшення кількості безпритульних тварин та їхньої взаємодії з дикими тваринами, а також забруднення територій внаслідок ракетних та артилерійських ударів та вибухів наземних мін [1].

У ході дослідження для прогнозування та аналізу часових рядів було проаналізовано наступні архітектури нейронних мереж: RNN, Block RNN, N-BEATS, TCN, Transformer Model, N-HiTS.

Оскільки нейронна мережа з архітектурою Transformer Model є передовою у вирішенні проблеми прогнозування часових рядів, її було обрано основною для дослідження [3]. На рисунку 2 зображено архітектуру нейронної мережі Transformer Model.

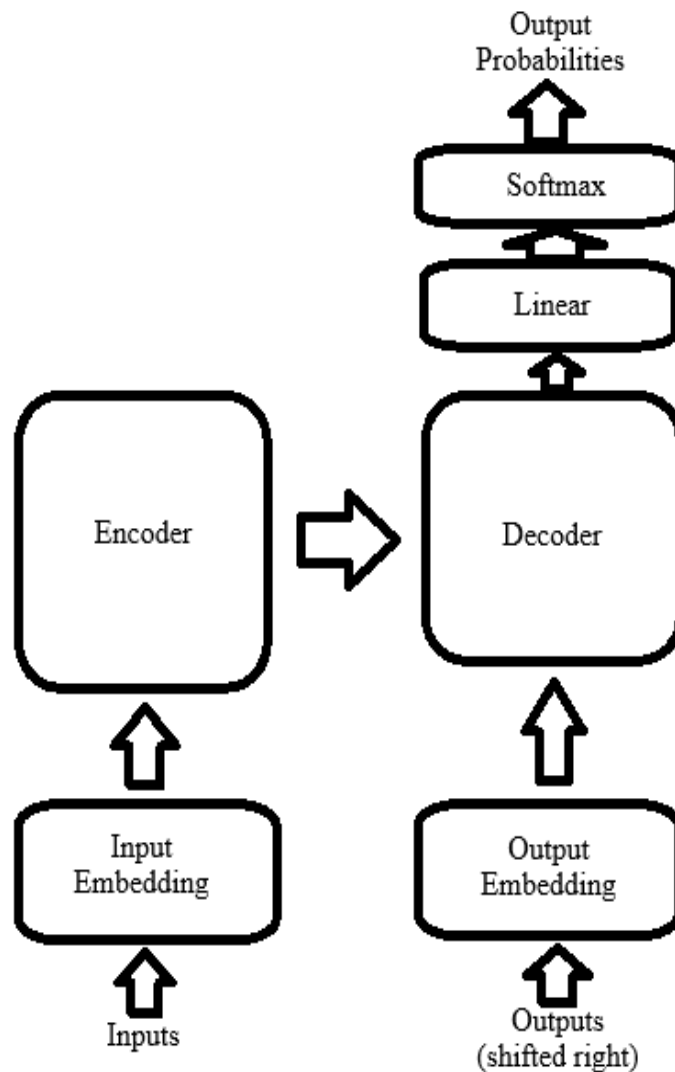


Рис. 2. Архітектура нейронної мережі Transformer Model.

Джерело: сформовано автором

Нормалізований набір даних про випадки інфекційних захворювань в Україні за період з грудня 2016 року по січень 2024 року послужив основою для навчання нейронних мереж. Ці дані також були нормалізовані в межах від 0 до 1 для покращення навчання Transformer Model. Нейронна мережа була навчена двом окремим підходам: по-перше, з використанням всіх часових рядів, що охоплюють усі інфекційні захворювання, присутні в наборі даних; і по-друге, зосередження на одному часовому ряді для конкретного інфекційного захворювання.

Для обох конфігурацій навчання модель використовувала історичні дані за попередні 24 місяці для прогнозування випадків на наступні шість місяців. Ці часові рамки мали на меті вловити тенденції та закономірності, які могли б підвищити точність прогнозування.

Для визначення результатів прогнозування була використана середня абсолютна процентна похибка (MAPE), також відома як середнє абсолютне процентне відхилення (MAPD) [4]. Варто відзначити, що більшість інфекційних захворювань не мають деякого періоду коливань, що відображається в передбаченнях нейронної мережі. Transformer Model дала кращі показники передбачень на часових рядах, які коливаються відносно стало (наприклад, кількість захворювань збільшується влітку і зменшується взимку). На рисунку 3 показано передбачення нейронних мереж на часових рядах, які коливаються відносно стало.

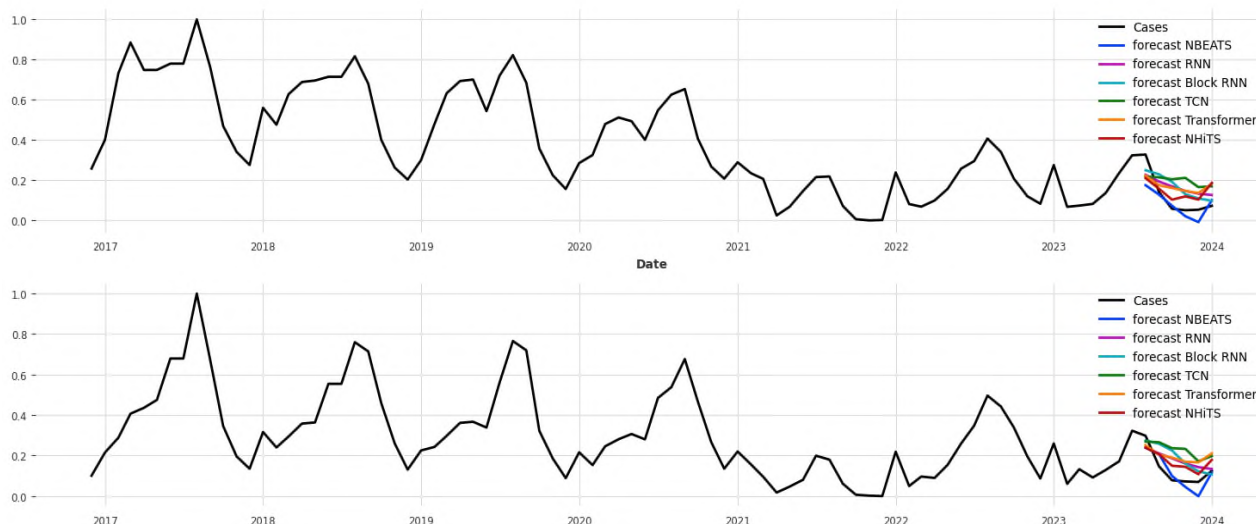


Рис. 3. Приклад прогнозування нейронних мереж.

Для проведення аналізу точності передбачення нейронних мереж, було використано інтервал впевненості в 95%. Таким чином, середні значення MAPE передбачення Transformer, яка навчалась з використанням всіх часових рядів, що охоплюють усі інфекційні захворювання в 95% випадків є 67.99%. В таблиці 1 представлено прогнози архітектури нейронної мережі, яка була навчена на всіх часових рядах.

Таблиця 1.

Прогнози нейронної мережі, яка були навчена на всіх часових рядах

Neural Network	Average MAPE, %
Transformer	67.99

Примітно, що нейронна мережа, навчена на кількох часових рядах, продемонструвала кращу продуктивність порівняно з тією, що було навчено на окремих часових рядах. Середнє значення MAPE для нейронної мережі Transformer Model, яка була навчена на окремих часових рядах склало 122.72%. Це покращення можна пояснити здатністю моделі використовувати спільну інформацію про різні захворювання, що збагатило процес навчання та сприяло більш надійним прогнозам. Результати підкреслюють потенціал підходів із багатовисхідними рядами для підвищення точності прогнозування інфекційних захворювань за допомогою нейронних мереж.

Це відкриття також свідчить про те, що інфекційні захворювання можуть впливати одна на одну та на їх поширення, що вказує на складну взаємодію між різними збудниками.

В результаті даної роботи проведено аналіз застосування нейронних мереж для вирішення задачі прогнозування часових рядів, зокрема, прогнозування розповсюдження інфекційних захворювань. Проведено аналіз та нормалізація датасету для тренування. Досліджено архітектуру нейронної мережі Transformer Model. Нейронна мережа з архітектурою Transformer в 95% випадків має 67.99% середню похибку. Відкрито, що інфекційні захворювання можуть впливати одна на одну та на їх поширення.

References

1. Goniewicz, K., Burkle, F.M., Home, S., Borowska-Stefańska, M., Wiśniewski, S., Khorram-Manesh, A. (2021) The Influence of War and Conflict on Infectious Disease: A Rapid Review of Historical Lessons We Have Yet to Learn. *Sustainability*, 13, 10783. Retrieved from: <https://doi.org/10.3390/su131910783>.
2. Public Health Center of the Ministry of Health of Ukraine (2024) Infectious disease in the population of Ukraine, Retrieved from: <https://phc.org.ua/kontrol-zakhvoryuvan/inshi-infekciyni-zakhvoryuvannya/infekciyna-zakhvoryuvannist-naseleennya-ukraini>.
3. Vaswani A, Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I., (2017) Attention Is All You Need, *Advances in Neural Information Processing Systems*, 2023, 3-4., Retrieved from: <https://arxiv.org/abs/1706.03762>
4. Myttenaere, Golden, B., Grand, B. Le., Rossi, F. (2016) Mean absolute percentage error for regression models, *Neurocomputing*. Retrieved from: <https://arxiv.org/abs/1605.02541>. DOI 10.34132/mspc2025.01.14.11

**Oleksandr Kovaliv,
Ievgen Sidenko,
Galyna Kondratenko**

This paper presents the results of time series forecasting using machine learning methods and mobile technologies. The main results of this work are a comparison of the effectiveness of different neural network architectures in predicting the spread of infectious diseases in Ukraine for the period from December 2016 to January 2024.

To achieve the goal, the following tasks were solved: the current state of the time series forecasting was analyzed; existing analogs of the systems were analyzed; the necessary neural network architectures were selected as one of the machine learning methods; a dataset for infectious diseases spread was analyzed and normalized; conducted testing. The developed system can be an effective tool for rational management decisions to ensure the epidemiological well-being and biosecurity of the population.

Keywords: *Forecasting, time series, machine learning methods, neural networks.*

Відомості про автора / Information about the Author

Олександр Ковалів, аспірант кафедри інтелектуальних інформаційних систем в Чорноморському національному університеті імені Петра Могили, м. Миколаїв, Україна. E-mail: kizutoamazed@gmail.com, orcid: <https://orcid.org/0009-0001-2047-6319>.

Oleksandr Kovaliv, PhD student of the Intelligent Information Systems Department at Petro Mohyla Black Sea National University (PMBSNU), Mykolaiv, Ukraine. E-mail: kizutoamazed@gmail.com, orcid: <https://orcid.org/0009-0001-2047-6319>.

Євген Сіденко, канд. техн. наук, доцент кафедри інтелектуальних інформаційних систем, Чорноморського національного університету імені Петра Могили, м. Миколаїв, Україна. E-mail: ievgen.sidenko@chmnu.edu.ua, orcid: <https://orcid.org/0000-0001-6496-2469>.

Ievgen Sidenko, PhD, Associate Professor of the Intelligent Information Systems Department at Petro Mohyla Black Sea National University, Mykolaiv, Ukraine. E-mail: ievgen.sidenko@chmnu.edu.ua, orcid: <https://orcid.org/0000-0001-6496-2469>.

Галина Кондратенко, канд. техн. наук, доцент кафедри інтелектуальних інформаційних систем, Чорноморського національного університету імені Петра Могили, м. Миколаїв, Україна. E-mail: halyna.kondratenko@chmnu.edu.ua, orcid: <https://orcid.org/0000-0002-8446-5096>.

Galyna Kondratenko, PhD, Associate Professor of the Intelligent Information Systems Department at Petro Mohyla Black Sea National University, Mykolaiv, Ukraine. E-mail: halyna.kondratenko@chmnu.edu.ua, orcid: <https://orcid.org/0000-0002-8446-5096>.

