

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет \_\_\_\_\_ Комп'ютерних наук  
(повна назва)  
Кафедра \_\_\_\_\_ Штучного інтелекту  
(повна назва)

**КВАЛІФІКАЦІЙНА РОБОТА**  
**Пояснювальна записка**

рівень вищої освіти \_\_\_\_\_ перший (бакалаврський)

\_\_\_\_\_ Стохастичне моделювання та глибоке навчання для  
\_\_\_\_\_ прогнозування динаміки епідемій  
(тема)

Виконав:  
здобувач \_\_\_\_\_ четвертого \_\_\_\_\_ року навчання,  
групи \_\_\_\_\_ ІТШ-22-1

\_\_\_\_\_ Владислав ЗЕЛІНСЬКИЙ  
(власне ім'я, прізвище)

Спеціальність 122 Комп'ютерні науки

(код і повна назва спеціальності)

Тип програми \_\_\_\_\_ освітньо-професійна

Освітня програма \_\_\_\_\_ Штучний інтелект

(повна назва освітньої програми)

Керівник \_\_\_\_\_ проф. Лариса ЧАЛА  
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ

\_\_\_\_\_  
(підпис)

\_\_\_\_\_ Лариса ЧАЛА  
(власне ім'я, прізвище)

2026 р.

Харківський національний університет радіоелектроніки

Факультет \_\_\_\_\_ Комп'ютерних наук \_\_\_\_\_

Кафедра \_\_\_\_\_ Штучного інтелекту \_\_\_\_\_

Рівень вищої освіти \_\_\_\_\_ перший (бакалаврський) \_\_\_\_\_

Спеціальність \_\_\_\_\_ 122 Комп'ютерні науки \_\_\_\_\_  
(код і повна назва)

Тип програми \_\_\_\_\_ освітньо-професійна \_\_\_\_\_

Освітня програма \_\_\_\_\_ Штучний інтелект \_\_\_\_\_  
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_

(підпис)

« \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ р.

**ЗАВДАННЯ**  
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві \_\_\_\_\_ Зелінському Владиславу Сергійовичу \_\_\_\_\_  
(прізвище, ім'я, по батькові)

1. Тема роботи Стохастичне моделювання та глибоке навчання для прогнозування динаміки епідемій

затверджена наказом університету від 5 червня 2026 р. № 488Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії 8 липня 2026 р.

3. Вихідні дані до роботи офіційна документація мови Python та її наукових бібліотек (NumPy, SciPy, statsmodels, PyTorch); відкритий набір даних Our World in Data щодо COVID-19 (локація – Україна); дані системи нагляду CDC ILINet (FluView) через програмний інтерфейс Delphi Epidata; наукові публікації з епідеміологічного моделювання, глибокого навчання та імовірного оцінювання прогнозів.

4. Перелік питань, що потрібно опрацювати в роботі \_\_\_\_\_

1) Аналіз предметної галузі та постановка задачі

2) Теоретичне дослідження

3) Проектування системи та її розробка

4) Експериментальні дослідження

## КАЛЕНДАРНИЙ ПЛАН

| №  | Назва етапів роботи                                                           | Строк / термін виконання етапів роботи | Примітка |
|----|-------------------------------------------------------------------------------|----------------------------------------|----------|
| 1  | Отримання завдання на кваліфікаційну роботу                                   | 08.06.2026                             | виконано |
| 2  | Аналіз предметної галузі та постановка задачі                                 | 09.06.2026                             | виконано |
| 3  | Теоретичне дослідження методів моделювання та глибокого навчання              | 13.06.2026                             | виконано |
| 4  | Підготовка даних, реалізація моделей і схеми валідації без витоку майбутнього | 14.06.2026                             | виконано |
| 5  | Реалізація та навчання гібридної моделі Neural-SEIR                           | 19.06.2026                             | виконано |
| 6  | Експериментальне дослідження та аналіз результатів і порівняння моделей       | 20.06.2026                             | виконано |
| 7  | Оформлення пояснювальної записки, висновків, переліку джерел та додатків      | 22.06.2026                             | виконано |
| 8  | Проходження нормконтролю                                                      | 25.06.2026                             | виконано |
| 9  | Рецензування                                                                  | 26.06.2026                             | виконано |
| 10 | Перевірка на академічний плагіат                                              | 28.06.2026                             | виконано |
| 11 | Підготовка презентації та доповіді                                            | 29.06.2026                             | виконано |
| 12 | Попередній захист                                                             | 30.06.2026                             | виконано |
| 13 | Захист роботи                                                                 | 08.07.2026                             |          |
|    |                                                                               |                                        |          |
|    |                                                                               |                                        |          |
|    |                                                                               |                                        |          |

Дата видачі завдання 8 червня 2026 р.

Здобувач \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_  
(підпис)

\_\_\_\_\_ проф. Лариса ЧАЛА  
(посада, власне ім'я, прізвище)

## РЕФЕРАТ

Пояснювальна записка: 86 с., 19 рис., 8 табл., 3 дод., 39 джерел.

ВАЛІДАЦІЯ, ГІБРИДНА МОДЕЛЬ, ГЛИБОКЕ НАВЧАННЯ, ЕПІДЕМІЯ, НЕВИЗНАЧЕНІСТЬ, ПРОГНОЗУВАННЯ, СТОХАСТИЧНЕ МОДЕЛЮВАННЯ, УНІВЕРСАЛЬНЕ ДИФЕРЕНЦІАЛЬНЕ РІВНЯННЯ, ЧАСОВІ РЯДИ, NEURAL-SEIR, SEIR.

Об'єкт дослідження – процес прогнозування динаміки епідемій за часовими рядами захворюваності.

Предмет дослідження – гібридна стохастично-нейромережева модель Neural-SEIR на основі універсального диференціального рівняння та методи оцінювання невизначеності прогнозу.

Мета роботи – розроблення, програмна реалізація та всебічна верифікація гібридної моделі Neural-SEIR для прогнозування динаміки епідемій і її об'єктивне порівняння з набором еталонних (бенчмарк) моделей.

Методи дослідження – компартментне та стохастичне моделювання; глибоке навчання з ансамблями глибоких мереж; універсальні диференціальні рівняння; валідація з наростаючим ковзним початком без витоку інформації з майбутнього; імовірнісні метрики оцінювання.

У роботі розглянуто актуальну проблему прогнозування динаміки епідемій за коротких, зашумлених і нерегулярно зведених рядів захворюваності. Запропоновано концепцію гібридної моделі Neural-SEIR, що поєднує епідеміологічну інтерпретовність SEIR з адаптивністю глибокого навчання та доповнюється стохастичним оцінюванням невизначеності прогнозу. Модель розроблено, програмно реалізовано й перевірено на даних щодо COVID-19 та сезонного грипу.

## **ABSTRACT**

Bachelor's thesis contains: 86 pp., 19 fig., 8 tabl., 3 ann., 39 references.

DEEP LEARNING, EPIDEMIC, FORECASTING, HYBRID MODEL, NEURAL-SEIR, SEIR, STOCHASTIC MODELLING, TIME SERIES, UNCERTAINTY, UNIVERSAL DIFFERENTIAL EQUATION, VALIDATION.

Object of research – the process of forecasting epidemic dynamics from incidence time series.

Subject of research – the hybrid stochastic–neural-network model Neural-SEIR based on a universal differential equation, together with methods for quantifying forecast uncertainty.

The aim of the work is to develop, implement in software and comprehensively verify the hybrid Neural-SEIR model for forecasting epidemic dynamics, and to provide its objective comparison against a set of benchmark models.

Research methods – compartmental and stochastic modelling; deep learning with deep ensembles; universal differential equations; rolling-origin validation with no leakage of future information; probabilistic evaluation metrics.

The work addresses the pressing problem of forecasting epidemic dynamics from short, noisy and irregularly aggregated incidence series. A concept of the hybrid Neural-SEIR model is proposed, combining the epidemiological interpretability of SEIR with the adaptivity of deep learning and complemented by stochastic estimation of forecast uncertainty. The model was developed, implemented in software and evaluated on COVID-19 and seasonal influenza data.

## ЗМІСТ

|                                                                |    |
|----------------------------------------------------------------|----|
| Перелік умовних позначень, символів, скорочень і термінів..... | 8  |
| Вступ.....                                                     | 10 |
| 1 Аналіз предметної галузі та постановка задачі.....           | 12 |
| 1.1 Опис предметної галузі .....                               | 12 |
| 1.2 Опис проблем та викликів.....                              | 14 |
| 1.3 Огляд існуючих підходів і рішень .....                     | 16 |
| 1.4 Інноваційність та переваги розроблюваної системи .....     | 20 |
| 1.5 Постановка задачі .....                                    | 22 |
| 2 Теоретичні дослідження .....                                 | 25 |
| 2.1 Огляд існуючих досліджень .....                            | 25 |
| 2.2 Компартментні моделі епідемій .....                        | 28 |
| 2.3 Стохастичне моделювання поширення інфекцій .....           | 31 |
| 2.4 Глибоке навчання для прогнозування часових рядів.....      | 34 |
| 2.5 Гібридні моделі та оцінювання невизначеності .....         | 37 |
| 2.6 Метрики оцінювання якості прогнозів .....                  | 41 |
| 3 Проектування системи та її розробка.....                     | 45 |
| 3.1 Загальна архітектура системи.....                          | 45 |
| 3.2 Структура програмного проєкту .....                        | 46 |
| 3.3 Підготовка та характеристика даних .....                   | 47 |
| 3.4 Схема валідації без витоку майбутнього.....                | 49 |
| 3.5 Реалізація базових і механістичних моделей.....            | 51 |
| 3.6 Реалізація моделей глибокого навчання.....                 | 53 |
| 3.7 Реалізація гібридної моделі Neural-SEIR .....              | 54 |
| 3.8 Відтворюваність і тестування.....                          | 58 |
| 4 Експериментальні дослідження.....                            | 61 |
| 4.1 Основне порівняння моделей .....                           | 61 |
| 4.2 Цінність нейромережевого коефіцієнта передачі.....         | 63 |
| 4.3 Ефективність за обсягом даних .....                        | 65 |

|                                                                   |    |
|-------------------------------------------------------------------|----|
| 4.4 Межі методу .....                                             | 67 |
| 4.5 Обговорення результатів.....                                  | 72 |
| Висновки .....                                                    | 74 |
| Перелік джерел посилання .....                                    | 76 |
| Додаток А Код гібридної моделі Neural-SEIR .....                  | 76 |
| Додаток Б Повні результати порівняння моделей за горизонтами..... | 84 |
| Додаток В Відомість кваліфікаційної роботи .....                  | 86 |

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ

ADR – Architecture Decision Record – запис архітектурного рішення;  
AIC – Akaike Information Criterion – інформаційний критерій Акаїке;  
API – Application Programming Interface – програмний інтерфейс застосунку;

CRPS – Continuous Ranked Probability Score – неперервна рангова ймовірнісна оцінка;

DL – Deep Learning – глибоке навчання;

EDA – Exploratory Data Analysis – розвідувальний аналіз даних;

GRU – Gated Recurrent Unit – вентиляльний рекурентний блок;

ILI / wILI – Influenza-Like Illness – грипоподібне захворювання та зважений відсоток;

LSTM – Long Short-Term Memory – мережа довгої короткочасної пам'яті;

MAE – Mean Absolute Error – середня абсолютна похибка;

MAPE – Mean Absolute Percentage Error – середня абсолютна відсоткова похибка;

MASE – Mean Absolute Scaled Error – масштабована абсолютна похибка;

MLP – Multilayer Perceptron – багатошаровий перцептрон;

ODE – Ordinary Differential Equation – звичайне диференціальне рівняння;

OWID – Our World in Data – джерело даних щодо COVID-19;

PINN – Physics-Informed Neural Network – фізично-інформована нейромережа;

RMSE – Root Mean Squared Error – корінь середньоквадратичної похибки;

RNN – Recurrent Neural Network – рекурентна нейронна мережа;

(S)ARIMA – (сезонна) інтегрована модель авторегресії – ковзного середнього;

SEIR – компартментна модель «сприйнятливі, експоновані, інфекційні, одужалі»;

UDE – Universal Differential Equation – універсальне диференціальне рівняння;

WIS – Weighted Interval Score – зважена інтервальна оцінка;

$\beta(t)$  – коефіцієнт передачі (нейромережевий у Neural-SEIR);

$\sigma, \gamma$  – швидкість інкубації ( $E \rightarrow I$ ) та одужання ( $I \rightarrow R$ ).

## ВСТУП

Прогнозування динаміки епідемій – одна з ключових задач сучасної охорони здоров'я. Точний і своєчасний прогноз захворюваності дає змогу заздалегідь планувати ресурси – ліжковий фонд, запаси ліків і медичний персонал. Він також допомагає вчасно вживати протиепідемічних заходів та оцінювати їхній результат. Пандемія COVID-19 і щорічні сезонні епідемії грипу показали, що практичну цінність має не лише сам прогноз, а й оцінка його невизначеності – саме на ній ґрунтується ухвалення рішень з урахуванням ризиків [1], [2].

До цієї задачі історично склалися два протилежні підходи, і кожен має свої межі. Перший – механістичний – спирається на компартментні моделі, передусім SEIR, які описують перебіг епідемії системою диференціальних рівнянь [3], [4], [5]. Їхні переваги – зрозуміла фізична інтерпретація та стійкість за невеликого обсягу даних. Головна вада – припущення про сталий коефіцієнт передачі, через яке одна модель не відтворює багатохвильову динаміку. Другий підхід – емпіричний – використовує моделі глибокого навчання, що виводять залежності прямо з даних [6], [7], [8]. Вони гнучкі, але сильно залежать від обсягу якісних даних і за їх нестачі схильні до перенавчання.

Найбільша потреба в прогнозі виникає на ранній стадії нового спалаху, коли спостережень ще мало, вони зашумлені, а рішення треба ухвалювати негайно. Саме тоді моделі глибокого навчання, яким потрібні великі обсяги якісних даних, працюють ненадійно, а суто механістичні моделі не вловлюють змін в інтенсивності поширення. Водночас провідні світові ініціативи – CDC FluSight для грипу та спільні платформи прогнозування COVID-19 – закріпили як стандарт імовірнісний прогноз з оцінкою невизначеності [2]. Попри це більшість прикладних рішень досі дає лише точкові оцінки. Усе це показує потребу в методі, який поєднав би

фізичну зрозумілість механістичних моделей з гнучкістю глибокого навчання й зберігав точність навіть за нестачі даних.

Такий метод належить до напряму наукового машинного навчання (scientific machine learning), що активно розвивається й прагне поєднати інтерпретованість механістичних моделей з гнучкістю нейромереж [9], [10]. У роботі ідею універсальних диференціальних рівнянь застосовано до моделі SEIR. Сталий коефіцієнт передачі замінено нейромережею, яка обчислює його за поточним станом системи. Цей гібрид – Neural-SEIR. Він відповідає загальній тенденції до поєднання фізичних і нейромережеских моделей у прогнозуванні часових рядів [9].

У роботі проаналізовано предметну галузь, реалізовано набір базових моделей за єдиним програмним інтерфейсом і на їхньому тлі перевірено модель Neural-SEIR. Окрему увагу приділено чесності порівняння. Дані (компактний набір OWID щодо COVID-19 та дані CDC FluView/ILINet через Delphi Epidata API) і методику валідації без витоку майбутнього дібрано так, щоб результати були відтворюваними, зокрема за нестачі даних.

Практичне значення роботи – створення коректного й відтворюваного інструмента прогнозування епідемій, придатного до роботи за обмежених даних. Запропонований метод і протокол порівняння можна застосувати в системах епідеміологічного нагляду, для планування ресурсів охорони здоров'я та для раннього прогнозування спалахів, коли даних замало для надійного застосування глибокого навчання. Матеріали роботи також придатні для використання в навчальному процесі.

## 1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ ТА ПОСТАНОВКА ЗАДАЧІ

### 1.1 Опис предметної галузі

Предметною галуззю роботи є прогнозування динаміки епідемій. Тобто передбачення майбутніх показників захворюваності на інфекційну хворобу за наявним рядом спостережень. Така задача належить до ширшої сфери епідеміологічного нагляду, у межах якої органи охорони здоров'я відстежують поширення хвороб, оцінюють поточну ситуацію та готують рішення на майбутнє. Прогноз потрібен не сам по собі, він є вхідними даними для цілком конкретних рішень – скільки лікарняних ліжок зарезервувати, який запас ліків і засобів захисту створити, коли й де посилити протиепідемічні заходи. Що раніше й точніше зроблено такий прогноз, то більше часу має система охорони здоров'я на підготовку.

Масштаб проблеми добре ілюструють дві хвороби, на яких зосереджено цю роботу. Пандемія COVID-19, що почалася на межі 2019 та 2020 років, за даними Всесвітньої організації охорони здоров'я призвела до понад 770 мільйонів підтверджених випадків і більш ніж 7 мільйонів зареєстрованих смертей у світі [11]; при цьому фактичні втрати, за оцінками надлишкової смертності, є помітно більшими. Сезонний грип, попри звичність, теж залишається вагомим тягарем. За модельною оцінкою глобальна смертність, пов'язана з респіраторними ускладненнями грипу, становить від 290 до 650 тисяч випадків щороку [12]. До прямих медичних втрат додаються економічні та соціальні: перевантаження лікарень, зупинка підприємств, обмеження пересування. Тому навіть невелике підвищення точності прогнозу має відчутну практичну вагу.

Вихідним матеріалом для прогнозу слугують ряди, які формують системи нагляду. Для грипу у США це мережа ILINet, що збирає частку звернень до лікарів із приводу грипоподібних захворювань і подає її як зважений відсоток wILI; для COVID-19 – офіційні реєстри нових випадків,

які ведуть національні органи охорони здоров'я. Такі ряди мають спільні особливості. Вони здебільшого тижневі, містять шум і затримки звітування, а їхні значення нерідко переглядають заднім числом, коли надходять уточнені дані. Усе це треба враховувати, інакше прогноз спиратиметься на інформацію, якої насправді не було в розпорядженні на момент його побудови.

Формально задачу прогнозування можна описати так. Нехай відомий ряд значень захворюваності до деякого моменту часу  $t$  (його називають *origin* – момент побудови прогнозу). Потрібно оцінити майбутні значення цього ряду на кілька кроків наперед, на так званому горизонті  $h$ . На практиці важливий не лише точковий прогноз, а й міра його невизначеності, тому сучасний стандарт вимагає видавати ймовірнісний прогноз – набір квантилів або інтервалів, що описують можливий розкид майбутнього значення. Робочими в епідеміології вважають горизонти від одного до чотирьох тижнів. Саме на такому проміжку прогноз і досяжний за точністю, і ще корисний для оперативних рішень.

Математичне моделювання поширення інфекцій має тривалу історію. Його відлік зазвичай ведуть від праці Кермака й Мак-Кендрика 1927 року [3], у якій епідемію вперше описано системою диференціальних рівнянь через поділ населення на групи за станом щодо хвороби. Цей підхід згодом розвинули в цілий клас компартментних моделей [4], [5]. Останнім часом галузь відчутно змінилася, завдяки відкритим платформам та регулярним змаганням із прогнозування, як-от спільне оцінювання прогнозів грипу в США [2], стандартом стало не разове моделювання, а систематичне порівняння багатьох методів на однакових даних із ймовірнісними метриками. Саме в цьому сучасному, орієнтованому на чесне порівняння руслі й виконано цю роботу.

Актуальність галузі підтвердила пандемія COVID-19, яка наочно показала ціну неточних і несвоєчасних прогнозів. Для України задача має додаткове практичне значення. Вітчизняні ряди захворюваності короткі,

неповні та ускладнені зовнішніми обставинами, тому методи, які зберігають точність за браку якісних даних, тут особливо потрібні.

## 1.2 Опис проблем та викликів

Під час розроблення системи прогнозування динаміки епідемій доводиться долати низку проблем, що впливають як із природи самих даних, так і з обмежень наявних методів. Нижче розглянуто основні з них, саме вони визначають вимоги до моделі та пояснюють, чому жоден окремий клас підходів не дає вичерпного розв'язку.

Найгостріша проблема – дефіцит чистих даних на ранній стадії спалаху. Саме тоді, коли прогноз найпотрібніший, спостережень ще мало, накопичилося кілька десятків точок, вони зашумлені, а рішення треба ухвалювати негайно. Моделі, що виводять залежності винятково з даних, у таких умовах працюють ненадійно, бо їм просто немає на чому вчитися. Цей дефіцит – не виняткова, а типова ситуація початку будь-якої нової епідемії, тому стійкість методу за малого обсягу даних є ключовою вимогою.

Друга група труднощів пов'язана з якістю та структурою рядів. Захворюваність звітують із затримкою, а вже опубліковані значення згодом переглядають у міру надходження уточнень – це явище провізорності (backfill). Якщо під час навчання використати остаточні, переглянуті значення, модель отримає інформацію, якої на момент прогнозу ще не існувало, і її оцінки виявляться надто оптимістичними. Додають складності тижневе агрегування, «вихідні» провали у звітності та нерегулярність подання даних. Огляди сучасних підходів окремо наголошують, що робота саме з такими «брудними» рядами, а не з ідеалізованими даними, є визначальною для практичної цінності прогнозу [13].

Третій виклик – нестаціонарність епідемічного процесу. Реальні епідемії проходять кількома хвилями, інтенсивність поширення змінюється

під впливом сезону, протиепідемічних заходів і появи нових штамів. Через це припущення про сталу інтенсивність передачі, на якому ґрунтуються найпростіші механістичні моделі, швидко втрачає чинність. Одна крива не лягає на кілька різних за формою хвиль. Водночас і суто статистичні методи погано вловлюють різкі зміни режиму, бо орієнтуються переважно на нещодавню поведінку ряду.

Окремо стоять структурні розриви й аномалії, не пов'язані з самою хворобою. Для українського ряду COVID-19 таким розривом стало повномасштабне вторгнення 2022 року, що зруйнувало звичну звітність; для американського ILINet – обвал показника під час пандемії 2020–2022 років, коли карантинні заходи різко змінили звертання до лікарів. Такі ділянки відображають форс-мажор, а не нормальну динаміку хвороби, тому їх не можна використовувати для оцінювання якості моделей нарівні зі звичайними періодами – інакше метрики характеризуватимуть реакцію на катастрофу, а не прогностичну спроможність.

Найпринциповіший методологічний виклик – суперечність між двома історично сформованими підходами. Механістичні компартментні моделі [3], [4], [5] мають зрозумілу фізичну інтерпретацію та зберігають притомність навіть за малого обсягу даних, проте їхня жорстка структура погано пристосовується до мінливої динаміки. Моделі глибокого навчання [6], [7], [8] навпаки – дуже гнучкі й точні, коли даних достатньо, але вкрай залежні від їхнього обсягу та схильні до перенавчання за дефіциту. Огляди застосувань машинного навчання в епідеміології підтверджують, що жоден із цих полюсів сам по собі не покриває всього діапазону умов [14]. Потрібен метод, який поєднав би сильні сторони обох.

Ще одна проблема – оцінювання невизначеності. Для рішень, що враховують ризик, важливо знати не лише найімовірніше значення, а й діапазон, у якому майбутній показник опиниться з заданою ймовірністю. Провідні міжнародні ініціативи з прогнозування епідемій закріпили ймовірнісний прогноз як сучасний стандарт [2], а для його об'єктивного

оцінювання запропонували спеціалізовані інтервальні метрики [1]. Попри це більшість прикладних досліджень і реальних інформаційних систем досі обмежується лише точковими прогнозами, залишаючи питання надійності та невизначеності результатів поза увагою.

Нарешті, серйозним викликом залишається неоднорідність задач прогнозування та потреба в коректному порівнянні різних моделей. Різні інфекційні захворювання описуються показниками різної природи й масштабу та характеризуються відмінною сезонністю, тривалістю часових рядів і динамікою поширення. У таких умовах порівняння результатів окремих досліджень є складним, оскільки вони часто використовують різні набори даних, способи їх розбиття та критерії оцінювання. Тому для отримання об'єктивних висновків усі моделі необхідно випробовувати за єдиною методикою, використовуючи однакові набори даних, ідентичні схеми навчання й тестування та спільний набір метрик. Лише такий підхід забезпечує чесне порівняння методів і дозволяє достовірно оцінити їхні переваги та недоліки. Без дотримання такого протоколу будь-які твердження про вищу точність окремого підходу залишаються недостатньо обґрунтованими.

### 1.3 Огляд існуючих підходів і рішень

Підходи до прогнозування динаміки епідемії можна умовно поділити на чотири групи: механістичні (компаратментні) моделі, статистичні методи аналізу часових рядів, моделі машинного й глибокого навчання та гібридні рішення. Окремо варто розглянути готові програмні платформи, що функціонують як прикладні аналоги розроблюваної системи. Кожен із цих підходів ґрунтується на різних принципах моделювання процесу поширення інфекції та має власні переваги, обмеження й сфери практичного застосування. Загальну класифікацію наведено на рисунку 1.1.

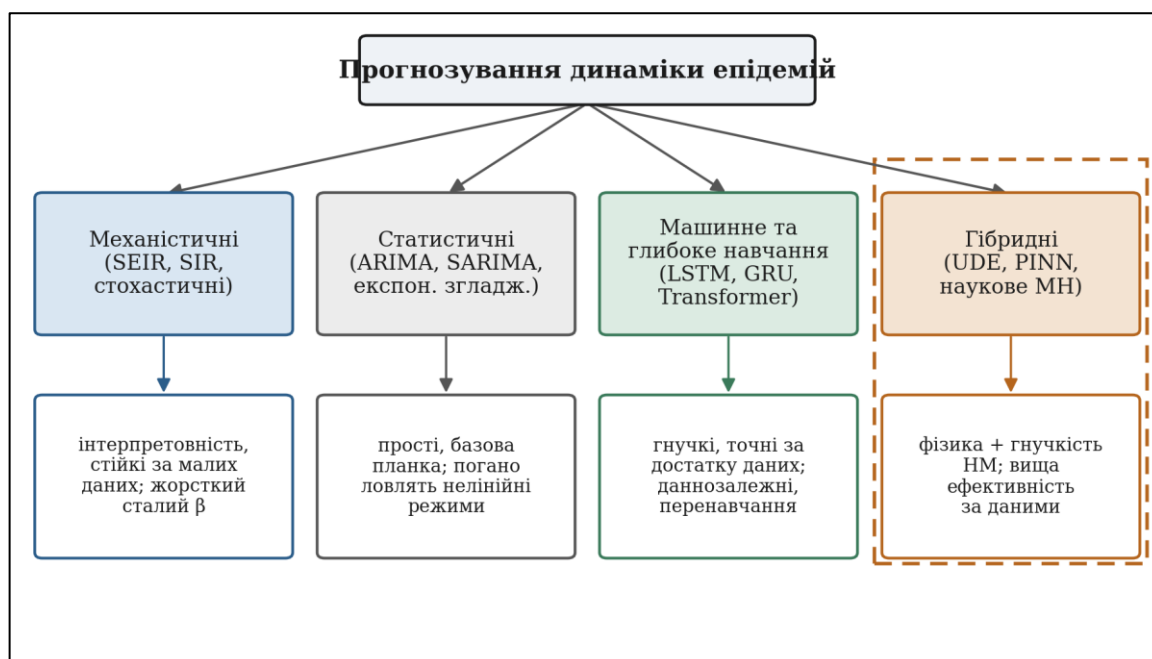


Рисунок 1.1 – Класифікація підходів до прогнозування динаміки епідемій

Механістичні моделі описують епідемію як перебіг хвороби в популяції, поділеній на компартменти за станом щодо інфекції. Найвідоміша з них – модель SEIR, у якій населення поділяють на сприйнятливих, експонованих (інфікованих у латентному періоді), інфекційних та одужалих, а переходи між групами задають коефіцієнтами передачі, інкубації й одужання (рисунок 1.2). Для підвищення реалістичності прогнозів базову модель SEIR часто розширюють додатковими компартментами або змінними параметрами, що дає змогу враховувати особливості перебігу конкретної епідемії. Такі моделі мають прозору інтерпретацію та стійкі за малого обсягу даних, адже спираються на фізику процесу, а не лише на спостереження [3], [4], [5]. Їхня головна вада вже згадувалася, припущення про сталий коефіцієнт передачі не дає однією моделлю відтворити багатохвильову динаміку. Часткове розв'язання дають стохастичні варіанти, які додають у модель випадковість переходів і так дозволяють оцінювати невизначеність, проте проблему жорсткої структури вони не знімають.

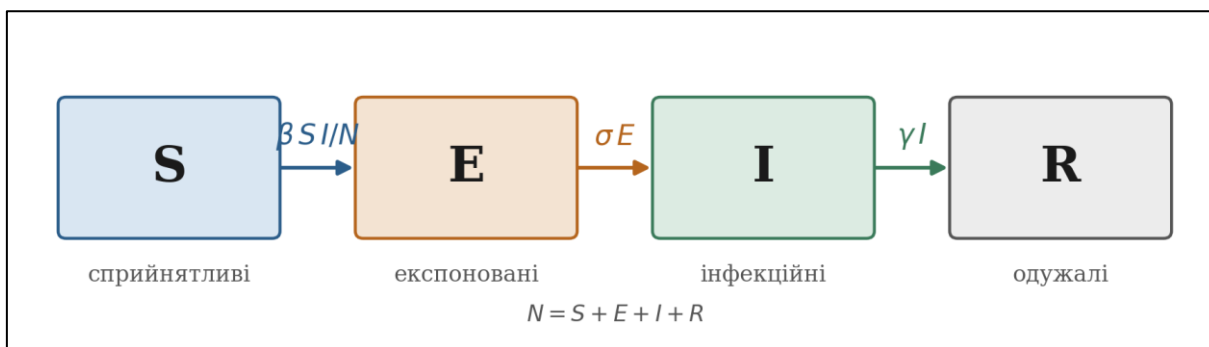


Рисунок 1.2 – Структурна схема компартментної моделі SEIR

Статистичні методи аналізу часових рядів підходять до задачі з іншого боку. Вони не моделюють механізм хвороби, а описують ряд формально. Найпоширеніші тут моделі сімейства (S)ARIMA та експоненційного згладжування, докладно викладені у класичних посібниках із прогнозування [15]. Їхні переваги – простота, мала кількість параметрів і добра робота як базової планки, особливо коли в ряді є виражена сезонність, яку вловлює сезонна складова SARIMA. Проте лінійна природа цих моделей заважає їм відтворювати різкі нелінійні підйоми та спади, характерні для епідемій, а ймовірнісні інтервали вони будують на досить сильних припущеннях про розподіл похибок.

Моделі машинного й особливо глибокого навчання останніми роками стали найдинамічнішим напрямом. Рекурентні мережі, передусім LSTM [6] та GRU [7], здатні вловлювати складні часові залежності, а архітектура Transformer із механізмом уваги [8] дала змогу ефективно опрацьовувати довгі послідовності. Огляди напряму [13], [14] показують, що такі моделі досягають високої точності, коли даних достатньо, і добре інтегрують додаткові джерела інформації. Водночас вони залишаються «чорними скриньками» з обмеженою інтерпретацією та, що головне для нашої задачі, потребують великих обсягів якісних даних; за їх дефіциту якість різко падає, а ризик перенавчання зростає.

Гібридні, або фізично-інформовані, підходи виникли як спроба поєднати переваги механістичних і нейромережових моделей. Ідея фізично інформованих нейромереж [10] полягає в тому, щоб уводити рівняння процесу безпосередньо у функцію втрат мережі. Більш загальну схему пропонують універсальні диференціальні рівняння (UDE) [9]. У відомому рівнянні процесу невідому або змінну складову замінюють невеликою нейромережею, зберігаючи решту фізичного кістяка. Цей напрям наукового машинного навчання активно розвивається й слугує теоретичною основою для моделі, запропонованої в цій роботі.

Окрему категорію становлять готові програмні платформи, які можна розглядати як прикладні аналоги. Найвідоміший приклад – COVID-19 Forecast Hub у США, що з квітня 2020 року щотижня збирав прогнози від десятків наукових груп у єдиному форматі [16], а згодом дав матеріал для ґрунтовного порівняння точності окремих моделей та їхнього ансамблю [17]. Аналогічну роль у Європі відіграв European COVID-19 Forecast Hub [18], а для грипу подібну функцію тривалий час виконує ініціатива CDC FluSight [2]. Спільний висновок цих платформ показовий. Найстабільніше точним виявляється не окрема модель, а ансамбль прогнозів, причому всі вони оцінюються ймовірносно, переважно за зваженою інтервальною оцінкою WIS [1]. Це задає орієнтир і для цієї роботи – ансамблевисть та ймовірісне оцінювання.

Щоб звести воедино сильні й слабкі сторони розглянутих груп методів, у таблиці 1.1 наведено їх порівняння за критеріями, найважливішими для поставленої задачі: інтерпретовністю, потребою в обсязі даних, робастністю за дефіциту даних та здатністю оцінювати невизначеність. Таке узагальнення дозволяє наочно оцінити переваги й обмеження кожного підходу. Воно також слугує основою для обґрунтування вибору методів, використаних у подальшому дослідженні.

Таблиця 1.1 – Порівняння підходів до прогнозування динаміки епідемій

| Підхід                   | Інтерпретовність | Потреба в даних | Робастність за дефіциту | Оцінка невизначеності |
|--------------------------|------------------|-----------------|-------------------------|-----------------------|
| SEIR                     | висока           | низька          | висока                  | часткова              |
| (S)ARIMA                 | середня          | середня         | середня                 | за припущеннями       |
| LSTM / GRU / Transformer | низька           | висока          | низька                  | через ансамблі        |
| Neural-SEIR              | висока           | низька          | висока                  | через ансамблі        |

Як видно з таблиці, жоден окремий підхід не є оптимальним за всіма критеріями водночас. Механістичні моделі виграють в інтерпретовності й стійкості за малих даних, але програють у гнучкості. Тоді як глибоке навчання, навпаки. Тому сучасні дослідження дедалі частіше орієнтуються на комбінування різних підходів для використання їхніх взаємних переваг. Такий підхід дозволяє досягти вищої точності прогнозування без втрати інтерпретованості та надійності результатів. Найбажанішим є метод, що поєднує робастність і прозорість фізичних моделей зі здатністю глибокого навчання адаптуватися до даних, а невизначеність оцінює ансамблево. Саме цю нішу й займає гібридна модель, запропонована в роботі.

#### 1.4 Інноваційність та переваги розроблюваної системи

Запропонована в роботі система побудована навколо гібридної стохастично-нейромережевої моделі Neural-SEIR. В її основі лежить ідея універсального диференціального рівняння [9]. Такий підхід поєднує знання про механізми поширення інфекцій із можливістю автоматичного навчання складних нелінійних залежностей на основі даних. Це дозволяє підвищити точність прогнозування без втрати фізичної інтерпретованості моделі. Жорсткий кістяк класичної моделі SEIR збережено повністю, але сталий коефіцієнт передачі замінено невеликою нейромережею, яка обчислює його значення за поточним станом системи. Концептуальну схему такого поєднання показано на рисунку 1.3.

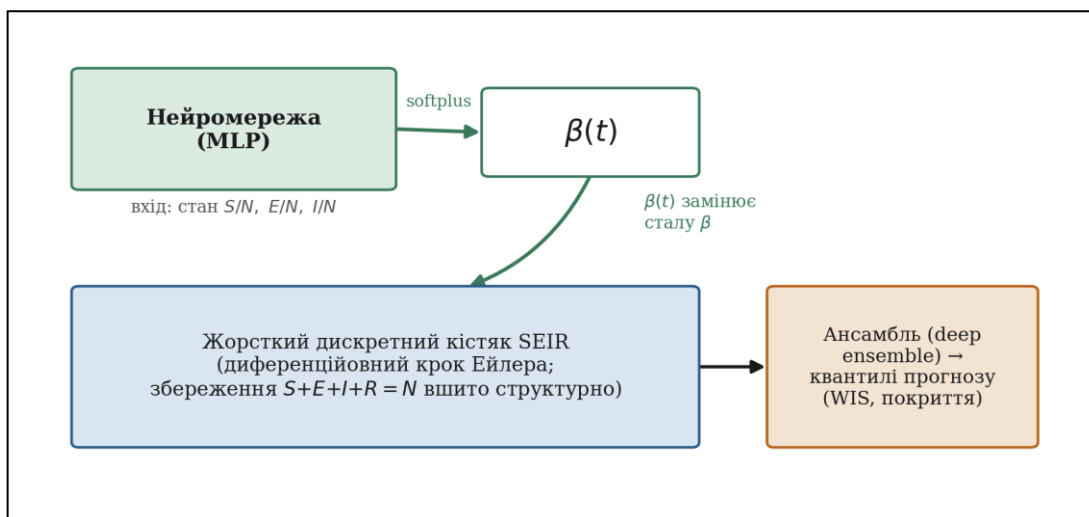


Рисунок 1.3 – Концептуальна схема гібридної моделі Neural-SEIR

Інноваційність рішення полягає в способі поєднання двох парадигм. На відміну від чистого глибокого навчання, тут фізичні обмеження не пропонуються моделі як побажання, а вшиті в саму структуру обчислень. Дискретні рівняння SEIR гарантують збереження чисельності населення й невід'ємність компартментів незалежно від того, що видасть нейромережа. На відміну від чистої механістики, коефіцієнт передачі більше не є сталим, а вільно змінюється в часі відповідно до стану епідемії, що дає змогу однією моделлю описати багатохвильову динаміку. При цьому нейромережевий внесок ізольовано саме в коефіцієнті передачі, тому порівняння з класичною SEIR лишається строго коректним. Єдина відмінність між ними – змінний коефіцієнт проти сталого.

Невизначеність прогнозу система оцінює за допомогою ансамблю глибоких мереж (deep ensemble) [19]. Кілька однакових моделей навчають із різною ініціалізацією, а розкид їхніх прогнозів дає епістемічну складову невизначеності. Підсумковий прогноз подається не точкою, а набором квантилів, який оцінюють зваженою інтервальною оцінкою WIS [1] – тією самою метрикою, що є стандартом провідних платформ. Так розроблювана система від початку відповідає сучасним вимогам до ймовірнісного прогнозування.

Сукупно це дає кілька практичних переваг. По-перше, висока ефективність за даними. Фізичний кістяк дозволяє моделі залишатися притомною там, де глибоке навчання ще не має на чому вчитися, тобто саме за дефіциту даних на початку спалаху. По-друге, інтерпретовність, відновлений нейромережею коефіцієнт передачі має зрозумілий епідеміологічний зміст і показує, як змінювалась інтенсивність поширення. По-третє, структурна коректність, закони збереження виконуються автоматично, без додаткових штрафів. По-четверте, чесність порівняння, усі моделі випробувано за єдиним протоколом без витоку інформації з майбутнього, тому будь-яка перевага є вимірюваною, а не задекларованою.

Додатковим результатом роботи стала точна характеристика самих даних. Під час аналізу ряду COVID-19 для України виявлено, що він є фактично тижневим по всьому робочому вікну, а не лише з 2022 року, як вважалося спершу. Це спостереження змінило спосіб формування цільового показника й усунуло викривлення метрик на найкоротшому горизонті. Чесна характеристика даних замість хибної – самостійний методологічний внесок, який також підвищує довіру до результатів.

Варто одразу окреслити й межі запропонованого рішення, щоб уникнути перебільшень. Перевага гібрида є відносною та нішевою. Він стабільно переважає чисту механістику й виграє в глибокого навчання за дефіциту даних, проте за достатку даних сильнішим лишається глибоке навчання, а на тижневому ряді COVID-19 на коротких горизонтах дуже сильною планкою є найпростіший наївний прогноз. Саме таке тверезе окреслення ніші, а не заяви про універсальну перевагу, робить результат науково вагомим.

## 1.5 Постановка задачі

На підставі проведеного аналізу галузі, її проблем та наявних рішень сформульовано задачу роботи. Об'єктом дослідження є процес

прогнозування динаміки епідемій за часовими рядами захворюваності. Предметом дослідження є гібридна стохастично-нейромережева модель Neural-SEIR на основі універсального диференціального рівняння та методи оцінювання невизначеності прогнозу. Метою роботи є розроблення, програмна реалізація та всебічна верифікація моделі Neural-SEIR для прогнозування динаміки епідемій і її об'єктивне порівняння з набором еталонних моделей.

Формально задачу прогнозування ставимо так. Нехай задано ряд значень захворюваності  $y$  до моменту  $t$ . Для кожного горизонту  $h$  потрібно за інформацією, доступною лише на момент  $t$ , оцінити набір квантилів прогнозного розподілу майбутнього значення  $y$  у момент  $t + h$ :

$$q^{(\tau)}_{t+h} = F^{-1}_{t+h}(\tau), \quad (1.1)$$

де  $q$  – квантиль рівня  $\tau$  прогнозного розподілу для горизонту  $h$ ;

$F$  – функція розподілу прогнозу на момент  $t + h$ .

Рівні квантилів беруть зі стандартного для платформ прогнозування набору  $\tau \in \{0,025; 0,1; 0,25; 0,5; 0,75; 0,9; 0,975\}$ , а горизонт – зі значень  $h = 1 \dots 4$ . Принципова вимога – відсутність витоку майбутнього. Усі ознаки й параметри на момент  $t$  обчислюють лише з даних, доступних до цього моменту включно.

Дослідження виконано на двох хворобах із контрастними режимами. Перший набір – тижневі нові випадки COVID-19 для України з відкритого джерела Our World in Data [20], робоче вікно 2020–2021 років. Другий – це тижневий показник wILI для США з мережі ILINet, отриманий через програмний інтерфейс Delphi Epidata [21], із тестовими до-ковідними сезонами 2014/15–2018/19. Для обох хвороб розглянуто горизонти від одного до чотирьох тижнів.

Для досягнення мети в роботі поставлено й розв'язано такі задачі:

- виконати аналіз предметної галузі, наявних підходів і готових рішень та сформулювати постановку задачі;
- опрацювати теоретичні засади компартментного й стохастичного моделювання, глибокого навчання та універсальних диференціальних рівнянь;
- підготувати дані двох хвороб і розробити схему валідації з наростаючим ковзним початком без витоку інформації з майбутнього;
- спроектувати та програмно реалізувати повний набір моделей за єдиним інтерфейсом: базові, механістичні, глибокого навчання й гібридну Neural-SEIR;
- реалізувати систему точкових та імовірнісних метрик і протокол чесного порівняння;
- провести експериментальне дослідження, зокрема стрес-тест дефіциту даних, і оцінити переваги та межі гібридної моделі;
- сформулювати висновки щодо досягнутих показників і напрямів подальшого вдосконалення.

Критеріями успішного розв'язання задачі вважаємо: стабільну перевагу гібридної моделі над чистою механістичною SEIR на обох хворобах за метрикою WIS; вищу ефективність за даними порівняно з глибоким навчанням в умовах дефіциту; а також повноту й коректність порівняння, проведеного на ідентичних розбиттях без витоку майбутнього. Усі результати мають бути відтворюваними за фіксованих початкових умов і збережених конфігурацій.

Програмну реалізацію виконано мовою Python з використанням бібліотек наукових обчислень і глибокого навчання; докладний опис інструментів і архітектури системи наведено в третьому розділі. Теоретичні засади використаних методів розглянуто в наступному, другому розділі.

## 2 ТЕОРЕТИЧНІ ДОСЛІДЖЕННЯ

### 2.1 Огляд існуючих досліджень

Дослідження з математичного моделювання епідемій мають майже столітню історію. Її прийнято відлічувати від праці Кермака й Мак-Кендрика [3], де вперше показано, що перебіг епідемії можна описати невеликою системою диференціальних рівнянь, а її результат визначається пороговою умовою. Згодом цей підхід оформився в окремий напрям, фундамент якого заклали монографії Андерсона й Мея [4] та Кілінга й Рохані [5]. У них компартментні моделі узагальнено на різні механізми передачі, вікову й просторову структуру населення, а також виведено ключові інтегральні характеристики, передусім базове репродуктивне число. Ці роботи закріпили за механістичними моделями репутацію інтерпретовного й теоретично прозорого інструмента, водночас не приховуючи їхнього головного обмеження – класична модель спирається на сталі в часі параметри, через що одна крива погано описує реальну багатохвилову динаміку.

Паралельно розвивався статистичний напрям, що підходить до прогнозування з боку аналізу часових рядів, не вдаючись до механізму хвороби. Його методологічну основу сформулювали Бокс і Дженкінс [22], запропонувавши системний підхід до побудови моделей авторегресії та інтегрованого ковзного середнього (ARIMA), а також їхніх сезонних варіантів. У задачах прогнозування захворюваності такі моделі десятиліттями слугують стандартною базовою планкою. Вони прості, потребують небагато даних і добре вловлюють регулярну сезонність, характерну, зокрема, для грипу. Їхня слабкість впливає з лінійної природи. Різкі нелінійні підйоми та спади епідемічних хвиль вони відтворюють лише наближено, а інтервали невизначеності будують на досить жорстких припущеннях про розподіл похибок.

Останнє десятиліття позначене стрімким поширенням методів машинного й глибокого навчання. Оглядові праці [13], [14] фіксують справжній вибух застосувань нейромереж до епідеміологічного прогнозування: рекурентних мереж LSTM [6] і GRU [7], а згодом і архітектур на основі механізму уваги [8]. Дослідники показали, що за наявності достатнього обсягу якісних даних і додаткових інформаційних потоків такі моделі досягають точності, недосяжної для класичних підходів. Водночас у тих самих оглядах наголошено на двох системних вадах: сильній залежності якості від обсягу навчальних даних і обмеженій інтерпретовності, через яку модель важко проаналізувати чи пояснити особі, що ухвалює рішення. Для задачі раннього прогнозування, де даних мало, перша вада є критичною.

Окремий і дуже впливовий напрям досліджень сформувався навколо колективних платформ прогнозування. Аналіз прогнозів, зібраних у COVID-19 Forecast Hub у США [16], [17] та в його європейському аналозі [18], а також багаторічний досвід ініціативи CDC FluSight для грипу [2] дали емпірично обґрунтований висновок. Найстабільнішу точність демонструє не окрема модель, а ансамбль прогнозів багатьох команд. Ці ж дослідження закріпили як галузевий стандарт ймовірнісну форму прогнозу та його оцінювання за зваженою інтервальною оцінкою [1]. Саме завдяки цим платформам прогнозування епідемій перетворилося з набору ізольованих моделей на дисципліну з чіткими правилами чесного порівняння.

На перетині механістичного й нейромережевого напрямів виник науковий машинний навчання (scientific machine learning). Його ідея полягає в тому, щоб не протиставляти фізику й нейромережі, а поєднувати їх. Робота про нейронні звичайні диференціальні рівняння [23] показала, що неперервну динаміку можна параметризувати нейромережею, а праці про універсальні диференціальні рівняння [9] та фізично-інформовані нейромережі [10] запропонували вбудовувати відомі рівняння процесу

безпосередньо в модель. Дослідження цього напрямку демонструють, що такий гібрид здатний зберігати інтерпретовність і стійкість фізичних моделей, додаючи гнучкість нейромереж. Проте застосувань цих ідей саме до прогнозування коротких епідемічних рядів поки небагато, і це залишає простір для внеску.

Нарешті, окрему лінію досліджень становить оцінювання невизначеності прогнозу. Якщо ранні прикладні системи здебільшого видавали точкову оцінку, то сучасні роботи розглядають невизначеність як невіддільну частину прогнозу. Серед практичних методів її оцінювання для нейромереж особливе місце посідають ансамблі глибоких мереж [19], привабливі простотою та масштабованістю. Дослідження в цій царині розрізняють епістемічну невизначеність, пов'язану з браком знань про модель, та алеаторну, зумовлену власним шумом процесу, і наголошують, що для відповідальних рішень потрібні саме калібровані, а не лише вузькі інтервали.

Зіставлення цих напрямів дозволяє зробити висновок, важливий для постановки роботи. Жодна з груп методів не є універсально найкращою. Механістичні моделі виграють в інтерпретовності й стійкості за дефіциту даних, але програють у гнучкості; глибоке навчання – навпаки. Найперспективнішим виглядає напрям, що структурно поєднує фізичні обмеження з адаптивністю нейромереж і доповнює прогноз ансамблевою оцінкою невизначеності. Саме він і покладений в основу запропонованої моделі Neural-SEIR. Такий вибір дозволяє поєднати переваги моделей різної природи та компенсувати їхні індивідуальні обмеження. Крім того, використання ансамблевого підходу забезпечує не лише отримання прогнозу, а й кількісну оцінку його надійності, що є важливим для підтримки прийняття рішень у сфері охорони здоров'я. Щоб обґрунтувати її будову, у наступних підрозділах докладно розглянуто теоретичний апарат кожної зі складових – від компартментних рівнянь до метрик оцінювання.

## 2.2 Компартментні моделі епідемій

Компартментні моделі описують епідемію як перебіг хвороби в популяції, яку поділяють на групи (компартменти) за станом щодо інфекції. Перехід між групами трактують як потік, інтенсивність якого задають кількома параметрами зі зрозумілим епідеміологічним змістом. У роботі використано модель SEIR – одну з найпоширеніших у цьому класі, бо вона враховує інкубаційний період, істотний для більшості респіраторних інфекцій. Населення чисельністю  $N$  поділяють на чотири групи: сприйнятливі  $S$ , які ще можуть заразитися; експоновані  $E$ , тобто інфіковані в латентному періоді, які ще не є заразними; інфекційні  $I$ , здатні передавати збудника; та одужалі або вилучені  $R$ , які вже не беруть участі в передачі [3], [4], [5]. Класичний виклад моделі та її властивостей подано в оглядовій праці Гезкоута [24], яку прийнято вважати довідковою в математичній епідеміології.

В основу моделі покладено два припущення. Перше – замкненість популяції. Упродовж розглядуваного періоду нехтують народжуваністю, природною смертністю та міграцією, тож сума всіх компартментів стала. Друге – однорідність змішування. Вважають, що кожен сприйнятливий контактує з кожним інфекційним з однаковою ймовірністю, без урахування просторової чи соціальної структури. Ці спрощення роблять модель аналітично прозорою і саме завдяки їм вона зберігає притомність навіть за невеликого обсягу спостережень, адже спирається на структуру процесу, а не лише на дані. У такій постановці модель дозволяє формально описати динаміку переходів між станами та отримати узгоджену еволюцію епідемічного процесу в часі. Кожне диференціальне рівняння описує швидкість зміни чисельності відповідного компартменту під впливом переходів між станами. Інтенсивність цих переходів визначається параметрами моделі, які характеризують процес інфікування, латентний

період та одужання. Неперервну форму моделі SEIR подають системою звичайних диференціальних рівнянь.

$$\begin{aligned}\frac{dS}{dt} &= -\beta \frac{SI}{N}, \\ \frac{dE}{dt} &= \beta \frac{SI}{N} - \sigma E, \\ \frac{dI}{dt} &= \sigma E - \gamma I, \\ \frac{dR}{dt} &= \gamma I,\end{aligned}\tag{2.1}$$

де  $\beta$  – коефіцієнт передачі, що відображає інтенсивність ефективних контактів за одиницю часу;

$N$  – чисельність популяції;

$\sigma$  – швидкість переходу з латентного стану в інфекційний, обернена до середньої тривалості інкубації;

$\gamma$  – швидкість одужання, обернена до середньої тривалості заразності.

Перший доданок  $\beta \cdot S \cdot I/N$  виражає закон дії мас: кількість нових заражень за одиницю часу пропорційна добутку часток сприйнятливих та інфекційних. Легко перевірити, що сума правих частин усіх чотирьох рівнянь дорівнює нулю, а отже величина  $S + E + I + R$  лишається сталою й дорівнює  $N$ .

Найважливішою інтегральною характеристикою моделі є базове репродуктивне число  $R_0$  — середня кількість вторинних заражень, які спричиняє один інфекційний індивід у повністю сприйнятливій популяції.

$$R_0 = \frac{\beta}{\gamma}.\tag{2.2}$$

Базове репродуктивне число задає пороговий критерій розвитку епідемії. Якщо  $R_0$  більше за одиницю, кількість інфекційних на початку

зростає експоненційно і виникає спалах; якщо менше – занесена інфекція згасає; значення  $R_0 = 1$  відповідає епідемічному порогу [24]. Оскільки з перебігом епідемії частка сприйнятливих зменшується, фактичну інтенсивність поширення в момент  $t$  характеризує ефективне репродуктивне число, яке враховує поточну частку сприйнятливих.

$$R_t = \frac{\beta}{\gamma} \cdot \frac{S(t)}{N}. \quad (2.3)$$

Коли ефективне репродуктивне число падає нижче одиниці, епідемічна крива проходить пік і починає спадати. Типову динаміку всіх чотирьох компартментів за сталих параметрів показано на рисунку 2.1: частка сприйнятливих монотонно зменшується, інфекційні дають характерний одногорбий спалах, а одужалі поступово накопичуються. Саме одногорбість кривої за сталого  $\beta$  є наочною ілюстрацією головного обмеження класичної моделі: одна така крива не здатна відтворити кілька послідовних хвиль, які спостерігаються в реальних епідеміях через сезонність, заходи й нові штами. Подолання цього обмеження – основний мотив гібридної моделі, у якій  $\beta$  робиться змінним у часі.

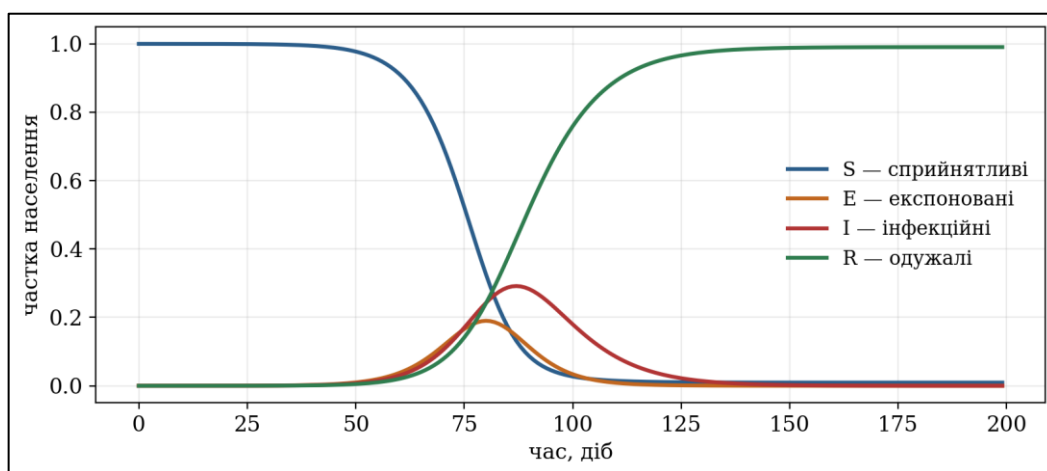


Рисунок 2.1 – Типова динаміка компартментів моделі SEIR за сталих параметрів

Для практичних обчислень неперервну систему дискретизують. У роботі застосовано явну схему Ейлера з кроком в одну добу, яка перетворює диференціальні рівняння на покрокові правила оновлення станів. Спершу обчислюють три потоки переходів – нові інфекції  $s2e$ , перехід латентних в інфекційні  $e2i$  та одужання  $i2r$ , після чого оновлюють значення компартментів.

$$\begin{aligned} s2e &= \min(\beta S_t I_t / N, S_t), & e2i &= \min(\sigma E_t, E_t), \\ i2r &= \min(\gamma I_t, I_t), \\ S_{t+1} &= S_t - s2e, & E_{t+1} &= E_t + s2e - e2i, \\ I_{t+1} &= I_t + e2i - i2r, & R_{t+1} &= R_t + i2r. \end{aligned} \quad (2.4)$$

Істотним елементом схеми є обмеження потоків знизу й зверху, записане через операцію мінімуму. Воно гарантує, що з компартмента не може вийти більше осіб, ніж у ньому є, а отже жодне значення не стає від'ємним. Разом із симетричною побудовою правил оновлення це забезпечує точне збереження суми  $S + E + I + R$  на кожному кроці. Така дискретна форма є базовою для всієї роботи. Її без змін успадковують і стохастичний варіант моделі, де детерміновані потоки замінюють випадковими, і гібридна модель Neural-SEIR, де сталий коефіцієнт  $\beta$  замінюють на нейромережвий. Завдяки цьому всі три моделі мають спільний кістяк і відрізняються лише способом задання потоків, що робить їх порівняння коректним.

### 2.3 Стохастичне моделювання поширення інфекцій

Детермінована модель SEIR описує лише середній, очікуваний перебіг епідемії й не відображає її випадкової природи. Тим часом реальне поширення інфекції є низкою дискретних випадкових подій. Конкретний контакт або призводить до зараження, або ні. Коли число інфікованих велике, ці випадковості усереднюються і детермінована крива добре

наближає дійсність. Проте саме на початку спалаху, коли інфікованих одиниці, і наприкінці, коли епідемія згасає, випадкові коливання стають визначальними. Занесена інфекція може як спричинити велику хвилю, так і згаснути, навіть якщо базове репродуктивне число більше за одиницю. Цю демографічну стохастичну детермінована модель принципово не вловлює, тому для коректного оцінювання невизначеності її розширюють стохастичним варіантом [25].

Стохастичну SEIR будують на тому самому дискретному кістяку, що й детерміновану, але потоки переходів роблять випадковими. Кількість подій кожного типу за крок моделюють пуассонівською випадковою величиною, середнє якої дорівнює відповідному детермінованому потоку. Такий підхід відомий як метод приблизного прискореного моделювання, або tau-leaping; його запропонував Гіллеспі для систем хімічних реакцій [26], а згодом перенесли на епідемічні процеси, де переходи між компартментами формально аналогічні реакціям. Правила одного кроку стохастичної моделі подано формулою.

$$\begin{aligned} s2e &\sim \min(\text{Pois}(\beta S_t I_t / N), S_t), \\ e2i &\sim \min(\text{Pois}(\sigma E_t), E_t), \\ i2r &\sim \min(\text{Pois}(\gamma I_t), I_t). \end{aligned} \tag{2.5}$$

Як і в детермінованому випадку, кожен потік обмежують зверху наявною чисельністю компартмента через операцію мінімуму, щоб уникнути від'ємних значень. Відмінність лише в тому, що замість точного добутку інтенсивності на чисельність береться випадкове число подій із пуассонівського розподілу з тим самим середнім. Через це кожен прогін моделі дає трохи іншу траєкторію, навіть за однакових початкових умов і параметрів. Щоб охопити весь діапазон можливих сценаріїв, модель проганяють багато разів – у роботі використано Монте-Карло-ансамбль із трьохсот незалежних траєкторій. Сукупність цих траєкторій і є наближенням розподілу майбутніх значень.

За ансамблем траєкторій будують ймовірнісний прогноз. Для кожного моменту часу обчислюють емпіричні квантілі по всіх прогонах і так отримують прогнозні інтервали різного рівня. Невизначеність, пов'язану з невідомими параметрами моделі, додатково враховують параметричним бутстрапом – повторним оцінюванням параметрів на варіаціях даних. Результат такого моделювання показано на рисунку 2.2. Тонкі лінії відповідають окремим випадковим траєкторіям, центральна крива – медіані ансамблю, а затінена смуга – дев'яностовідсотковому інтервалу.

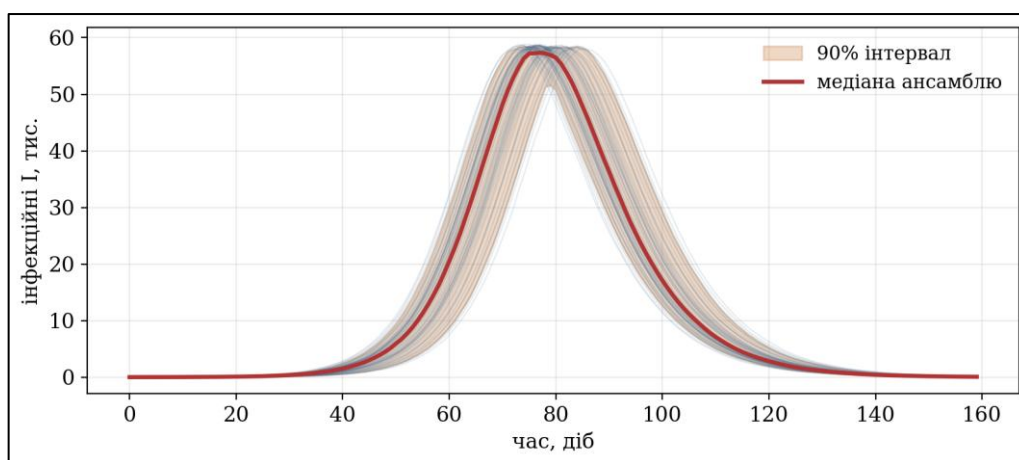


Рисунок 2.2 – Ансамбль стохастичних траєкторій моделі SEIR та побудований за ним інтервал

Розкид траєкторій наочно показує головну перевагу стохастичного підходу. Невизначеність виникає в моделі природно, як наслідок випадковості самого процесу, а не доклеюється штучно. Видно також, що розкид найбільший поблизу піку й на спаді, що узгоджується з інтуїцією про чутливість епідемії до випадковостей. Платою за цю перевагу є обчислювальна вартість – потреба в багатьох прогонах – і те, що стохастична модель успадковує від детермінованої припущення про сталий коефіцієнт передачі. Тому в роботі стохастична SEIR слугує одним із учасників бенчмарк-ростера, а закладений у ній принцип оцінювання

невизначеності через ансамбль траєкторій згодом використано і в гібридній моделі, де роль джерела розкиду відіграє вже ансамбль нейромереж.

## 2.4 Глибоке навчання для прогнозування часових рядів

На відміну від механістичних моделей, методи глибокого навчання не закладають у себе жодних припущень про механізм хвороби, а виводять залежності безпосередньо з даних. Загальні принципи цього напряму систематично викладені в монографії Гудфеллоу, Бенжіо та Курвілля [27]. Для роботи з часовими рядами задачу формують так, на вхід моделі подають вікно минулих спостережень фіксованої довжини, а на виході вона має передбачити значення на кілька кроків наперед. Ключове питання полягає в тому, як ефективно врахувати порядок і взаємозв'язок елементів послідовності, і відповіддю на нього стали спеціалізовані нейромережеві архітектури.

Першими таку здатність дали рекурентні нейронні мережі (RNN). Вони опрацьовують послідовність елемент за елементом, підтримуючи прихований стан, який відіграє роль пам'яті й оновлюється на кожному кроці за поточним входом і попереднім станом.

$$h_t = \tanh(W_x x_t + W_h h_{t-1} + b). \quad (2.6)$$

Прихований стан накопичує інформацію про вже побачену частину ряду, тож теоретично проста рекурентна мережа здатна вловлювати залежності будь-якої давності. На практиці, однак, її можливості обмежені відомою проблемою згасання й вибуху градієнтів, докладно проаналізованою Бенжіо зі співавторами [28]. Під час навчання градієнт, що поширюється на багато кроків назад, з кожним кроком множиться на близькі величини й тому або експоненційно зменшується до нуля, або необмежено зростає. Унаслідок цього проста RNN фактично не здатна

навчитися довготривалих залежностей, що для епідемічних рядів із їхньою сезонністю є суттєвою вадою.

Розв'язанням цієї проблеми стала мережа довгої короткочасної пам'яті (LSTM), запропонована Хохрайтером і Шмідхубером [6]. У ній, окрім прихованого стану, уведено окремий стан комірки, що проходить крізь усю послідовність майже без змін і слугує «магістраллю» для збереження інформації на довгих проміжках. Доступ до цього стану регулюють три навчувані вентиля: вентиль забування вирішує, яку частину пам'яті стерти, вхідний вентиль – яку нову інформацію записати, а вихідний – що передати далі як прихований стан. Рівняння комірки LSTM подано формулою 2.7, а її будову схематично показано на рисунку 2.3.

$$\begin{aligned}
 f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f), \\
 i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i), \\
 o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o), \\
 \tilde{c}_t &= \tanh(W_c[h_{t-1}, x_t] + b_c), \\
 c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \\
 h_t &= o_t \odot \tanh(c_t).
 \end{aligned}
 \tag{2.7}$$

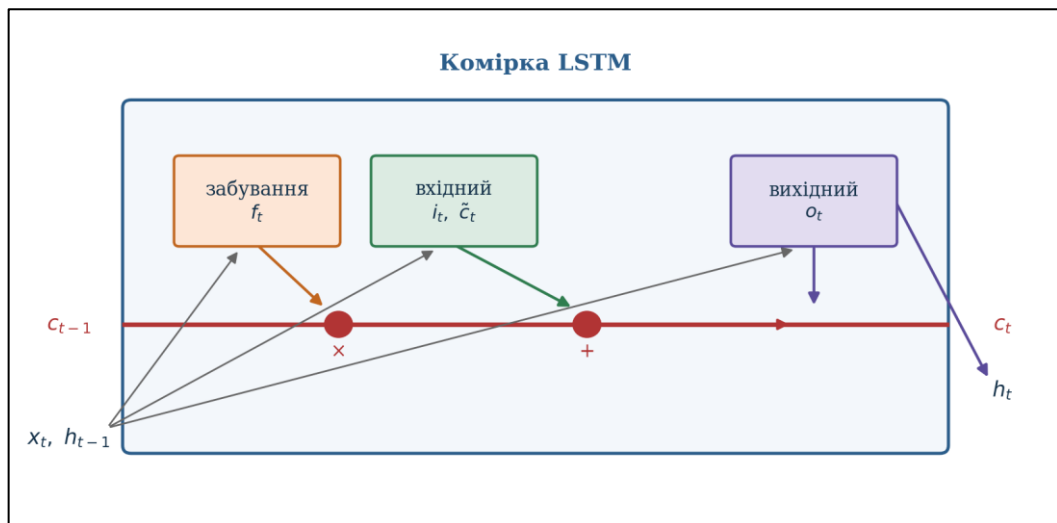


Рисунок 2.3 – Схема комірки LSTM з вентилями забування, входу та виходу

На схемі добре видно центральну ідею архітектури. Горизонтальна лінія стану комірки проходить крізь усю комірку, і інформація на ній змінюється лише через дві контрольовані операції – множення на вихід вентиля забування та додавання нового кандидата, відфільтрованого вхідним вентилям. Завдяки тому, що пам'ять переноситься переважно додаванням, а не повторним множенням, градієнти вже не згасають так швидко, і мережа здатна утримувати залежності на десятки кроків. Саме ця властивість зробила LSTM однією з основних архітектур для прогнозування часових рядів.

Близькою за духом, але простішою є мережа з вентиляльними рекурентними блоками (GRU), яку запропонували Чо зі співавторами [7]. У ній два вентиля замість трьох і немає окремого стану комірки, вентиль оновлення визначає, наскільки замінити старий прихований стан новим, а вентиль скидання – яку частину минулого врахувати під час обчислення кандидата.

$$\begin{aligned} z_t &= \sigma(W_z[h_{t-1}, x_t]), & r_t &= \sigma(W_r[h_{t-1}, x_t]), \\ \tilde{h}_t &= \tanh(W_h[r_t \odot h_{t-1}, x_t]), \\ h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t. \end{aligned} \quad (2.8)$$

За меншої кількості параметрів GRU нерідко показує точність, порівнянну з LSTM, а швидше навчається й менше схильна до перенавчання. Остання властивість особливо цінна за обмеженого обсягу даних, тому в роботі GRU розглянуто нарівні з LSTM як окрему модель ростера.

Принципово інший підхід реалізує архітектура Transformer, запропонована Васвані зі співавторами [8]. Вона повністю відмовляється від рекурентності й опрацьовує всю послідовність одночасно, а зв'язки між її елементами моделює механізмом самоуваги. Для кожного елемента механізм обчислює зважену суму всіх інших, де ваги визначаються

подібністю запитів і ключів, а нормування на корінь з розмірності утримує значення в придатному для навчання діапазоні.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \quad (2.9)$$

Оскільки самоувага не враховує порядку елементів, до входу додають позиційне кодування, що повертає інформацію про їхнє розташування. Перевага Transformer – у здатності напряду пов'язувати далекі в часі точки й у природній паралельності обчислень; його вада для нашої задачі це підвищена потреба в даних, адже без вбудованого припущення про послідовність він має вивести всі залежності сам. Усі три архітектури навчають мінімізацією середньоквадратичної похибки в масштабованому просторі ознак. Спільна для них риса – висока залежність якості від обсягу навчальної вибірки та схильність до перенавчання на коротких рядах – і є тим обмеженням, яке в роботі долають двома засобами: ансамблюванням, розглянутим далі, та поєднанням нейромережі з фізичним кістяком у гібридній моделі.

## 2.5 Гібридні моделі та оцінювання невизначеності

Розглянуті дотепер підходи утворюють два полюси: механістичні моделі повністю задають динаміку рівняннями, а нейромережі повністю виводять її з даних. Гібридні моделі займають проміжне положення й намагаються поєднати сильні сторони обох. Цей напрям, відомий як науковий машинний навчання, спирається на ідею, що частину відомої фізики процесу варто зберегти у структурі моделі, а нейромережею замінювати лише те, що насправді невідоме. Відправною точкою для нього стала концепція нейронних звичайних диференціальних рівнянь [23], у якій

праву частину рівняння динаміки задають неймережею, а її узагальненням – універсальні диференціальні рівняння (UDE) [9].

Ідея UDE полягає в тому, щоб у відомому рівнянні процесу залишити всю наявну фізику, а невідому або погано визначену складову замінити невеликою неймережею, яку навчають разом із рештою системи. Так зберігається інтерпретовний кістяк, але модель отримує гнучкість там, де знань бракує.

$$\frac{dx}{dt} = f(x, t) + \text{NN}_{\theta}(x). \quad (2.10)$$

Споріднений підхід реалізують фізично-інформовані неймережі (PINN), запропоновані Райссі зі співавторами [10]. У них рівняння процесу не вбудовують у структуру, а додають як штрафний доданок до функції втрат: повна втрата складається з похибки на даних і доданку, що карає за відхилення від рівнянь.

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda \mathcal{L}_{\text{phys}}. \quad (2.11)$$

Між цими двома підходами є принципова різниця. У PINN фізичні закони виконуються лише наближено, як компроміс під час оптимізації, і коефіцієнт  $\lambda$  доводиться ретельно підбирати; за невдалого вибору модель може або знехтувати фізикою, або недонавчитися на даних. У підході UDE фізичний кістяк вшито в саму структуру обчислень, тому закони збереження виконуються точно й безумовно, незалежно від того, що видасть неймережа. Через цю надійнішу гарантію, а також через відому нестабільність навчання PINN на коротких рядах, у роботі за основу гібридної моделі обрано саме UDE.

Застосування ідеї UDE до моделі SEIR і дає центральну модель роботи Neural-SEIR. У ній жорсткий дискретний кістяк SEIR зберігають повністю,

а сталий коефіцієнт передачі  $\beta$  замінюють невеликою нейромережею, яка обчислює його значення за поточним нормованим станом системи. На вхід мережі подають частки сприйнятливих, експонованих та інфекційних, а функція `softplus` на виході гарантує, що коефіцієнт передачі завжди невід'ємний.

$$\begin{aligned}\beta(t) &= \text{softplus}(W_2 \tanh(W_1 s + b_1) + b_2), \\ s &= [S/N, E/N, I/N]^T.\end{aligned}\tag{2.12}$$

Принципово, що нейромережевим зроблено лише коефіцієнт передачі, тоді як решта параметрів і структура збігаються з класичною SEIR. Завдяки цьому, по-перше,  $\beta$  перестає бути сталим і вільно змінюється відповідно до фази епідемії, що знімає обмеження на відтворення багатохвильової динаміки; по-друге, порівняння гібрида з чистою SEIR лишається строго коректним, бо моделі відрізняються рівно однією деталлю – змінним коефіцієнтом проти сталого.

Окремої уваги потребує оцінювання невизначеності прогнозу нейромережевих моделей. Для цього в роботі застосовано ансамблі глибоких мереж [19]: кілька однакових за архітектурою моделей навчають незалежно, з різною випадковою ініціалізацією, після чого їхні прогнози усереднюють. У теорії невизначеності прийнято розрізняти дві її складові [29]: епістемічну, що відображає брак знань про модель і зменшується з надходженням даних, та алеаторну, зумовлену власним шумом процесу й непереборну. Такий поділ дає змогу окремо оцінити внесок невизначеності моделі та випадковості самого процесу в загальну похибку прогнозу. Це особливо важливо для задач прогнозування епідемій, де обмежений обсяг даних і стохастичний характер поширення захворювання можуть суттєво впливати на надійність передбачень. Розкид прогнозів членів ансамблю природно оцінює епістемічну складову, а

алеаторну враховують окремо через залишкову дисперсію. Підсумкові квантилі будують, поєднавши обидві складові.

$$\begin{aligned}\mu &= \frac{1}{K} \sum_{k=1}^K \hat{y}_k, & \sigma_{\text{epi}}^2 &= \frac{1}{K} \sum_{k=1}^K (\hat{y}_k - \mu)^2, \\ \sigma &= \sqrt{\sigma_{\text{epi}}^2 + \sigma_{\text{ale}}^2}, & q_\tau &= \mu + \sigma \Phi^{-1}(\tau).\end{aligned}\quad (2.13)$$

Схему такого ансамблю показано на рисунку 2.4: один і той самий вхід подають на  $K$  незалежно навчених мереж, а їхні прогнози зводять у середнє значення та оцінку розкиду, з яких далі будують квантилі прогнозного розподілу.

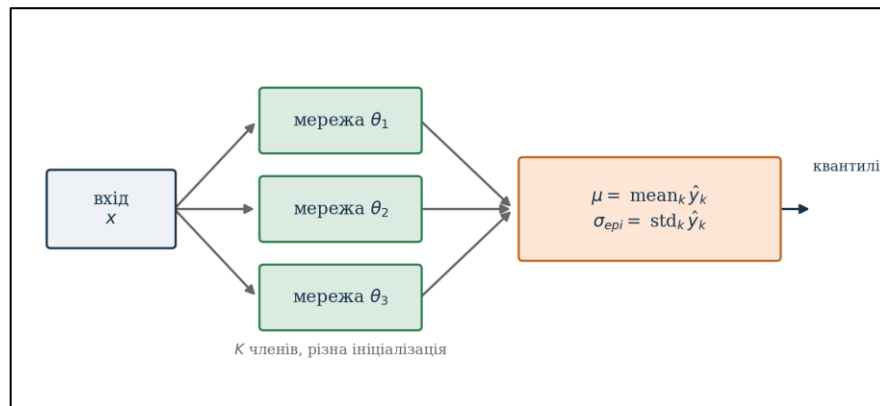


Рисунок 2.4 – Оцінювання невизначеності ансамблем глибоких мереж

Важливо, що той самий механізм ансамблювання застосовано однаково до всіх нейромережових моделей ростера – і до LSTM, GRU та Transformer, і до гібрида Neural-SEIR. Це зрівнює умови порівняння. Тобто, якщо гібридна модель показує кращі результати, це точно не через якийсь особливий спосіб оцінки невизначеності, адже він однаковий для всіх. Маючи від кожної моделі не точкову оцінку, а повний набір квантилів, можна перейти до питання про те, як саме вимірювати й порівнювати якість таких ймовірнісних прогнозів.

## 2.6 Метрики оцінювання якості прогнозів

Щоб порівнювати моделі об'єктивно, потрібен чіткий набір метрик. Їх поділяють на дві групи. Точкові метрики оцінюють, наскільки близькою до факту виявилася центральна оцінка прогнозу, а ймовірнісні – наскільки добре модель оцінила власну невизначеність, тобто чи відповідають заявлені інтервали реальному розкиду. Оскільки ряди різних хвороб мають різний масштаб, наприкінці вводять відносну метрику, що дає змогу зіставляти якість на різних рядах. Усі формули цього підрозділу наведено в тому вигляді, у якому їх реалізовано в системі оцінювання.

Найпростіша точкова метрика – середня абсолютна похибка (MAE), що дорівнює середньому модулю відхилення прогнозу від факту. Близька до неї за змістом середньоквадратична похибка (RMSE) усереднює квадрати відхилень і бере корінь.

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|, \quad (2.14)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}. \quad (2.15)$$

Різниця між ними тонка, але важлива: через піднесення до квадрата RMSE значно сильніше карає за великі поодинокі промахи, тоді як MAE оцінює типову похибку. Обидві метрики виражено в одиницях самого ряду, тому їхні значення не можна напрямую порівнювати між хворобами з різним масштабом захворюваності. Крім того, поширена відсоткова похибка MAPE для епідемічних рядів непридатна: біля нульових і близьких до нуля значень, типових для міжсезоння, вона необмежено зростає й дає оманливі результати.

Щоб усунути залежність від масштабу, застосовують масштабовану абсолютну похибку (MASE), запропоновану Гайндманом і Кьолером [30]. Її обчислюють як відношення MAE моделі до середньої абсолютної похибки наївного сезонного прогнозу на навчальній вибірці.

$$\text{MASE} = \frac{\frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|}{\frac{1}{n-m} \sum_{t=m+1}^n |y_t - y_{t-m}|}. \quad (2.16)$$

Перевага MASE в тому, що вона безрозмірна й має природний орієнтир: значення менше за одиницю означає, що модель точніша за наївну планку, а більше – що гірша. Завдяки цьому MASE можна усереднювати по різних рядах і горизонтах. Проте всі точкові метрики мають спільну ваду: вони оцінюють лише центральну оцінку й повністю ігнорують якість інтервалів невизначеності, яка для відповідальних рішень не менш важлива.

Якість ймовірнісного прогнозу оцінюють так званими строго належними оцінними правилами, теорію яких узагальнили Гнайтінг і Рафтері [31]. Належність означає, що правило заохочує модель повідомляти свій справжній розподіл: найкращий очікуваний бал досягається лише тоді, коли заявлений прогноз збігається з істинним. Такі правила водночас винагороджують і калібрування (відповідність інтервалів дійсності), і гостроту (вузькість інтервалів), не дозволяючи покращити оцінку за рахунок лише одного з цих складників. Базовим прикладом є втрата вільотом (pinball loss) для окремого квантиля, яка несиметрично карає за відхилення залежно від рівня квантиля.

$$L_{\tau}(y, q) = \max(\tau(y - q), (\tau - 1)(y - q)). \quad (2.17)$$

Для оцінювання цілого центрального інтервалу використовують інтервальну оцінку. Вона додає до ширини інтервалу штраф за кожне

спостереження, що випало за його межі, причому штраф тим більший, чим вужчий заявлений інтервал.

$$IS_{\alpha} = (u - l) + \frac{2}{\alpha} (l - y)_{+} + \left(\frac{2}{\alpha} (y - u)_{+}\right). \quad (2.18)$$

Головною ймовірнісною метрикою роботи обрано зважену інтервальну оцінку (WIS), яку для прогнозування епідемій запропонували Брахер зі співавторами [1]. Вона є зваженою сумою інтервальних оцінок для кількох рівнів довіри та абсолютної похибки медіани і слугує дискретним наближенням неперервної рангової ймовірнісної оцінки. Саме WIS прийнято як стандарт у провідних платформах прогнозування, тому її використано як основну метрику й тут.

$$WIS = \frac{1/2 |y - m| + \sum_{k=1}^K \alpha_k / 2 IS_{\alpha_k}}{K + 1/2}. \quad (2.19)$$

Принцип, за яким WIS оцінює прогноз, унаочнює рисунок 2.5: навколо медіанного прогнозу будують вкладені інтервали різного рівня, і метрика винагороджує модель за те, що факт потрапляє всередину вузьких інтервалів, та карає за промахи й надмірну ширину.

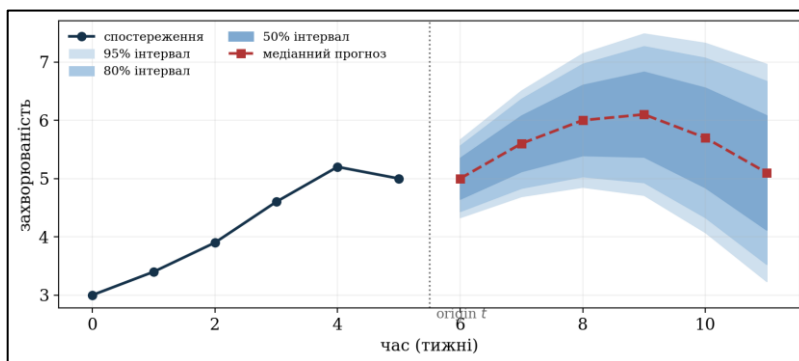


Рисунок 2.5 – Ймовірнісний прогноз із вкладеними інтервалами 50, 80 і 95 відсотків

Окремо контролюють калібрування інтервалів за допомогою покриття, частки спостережень, що фактично потрапили в інтервал заданого рівня. Для добре каліброваної моделі покриття має наближатися до номінального значення: приблизно половина спостережень мають потрапляти в п'ятдесятивідсотковий інтервал, і так далі. Систематичне відхилення покриття від номіналу свідчить про те, що модель або переоцінює, або недооцінює власну невизначеність.

$$\text{Cov} = \frac{1}{n} \sum_{t=1}^n \mathbf{1} [l_t \leq y_t \leq u_t]. \quad (2.20)$$

Нарешті, щоб зіставляти моделі на рядах різного масштабу й зводити результати в єдиний рейтинг, використовують відносний skill стосовно наївного прогнозу. Він показує, на скільки відсотків модель краща або гірша за просту персистенс-планку: додатне значення означає перевагу над наївним прогнозом, нульове – рівність, від'ємне – програш [2], [15].

$$\text{Skill} = 1 - \frac{\text{WIS}_{\text{model}}}{\text{WIS}_{\text{naive}}}. \quad (2.21)$$

Описані метрики разом зі схемою валідації без витоку майбутнього утворюють єдину систему оцінювання, на якій ґрунтується все експериментальне порівняння в четвертому розділі. На цьому теоретичний апарат роботи викладено повністю: розглянуто компартментне та стохастичне моделювання, основні архітектури глибокого навчання, концепцію гібридних моделей і методи оцінювання якості прогнозів. У наступному розділі ці теоретичні засади покладено в основу проектування й програмної реалізації системи моделей.

## 3 ПРОЕКТУВАННЯ СИСТЕМИ ТА ЇЇ РОЗРОБКА

### 3.1 Загальна архітектура системи

Систему спроектовано як конвеєр, що послідовно перетворює сирі дані на ймовірнісний прогноз і його оцінку. Конвеєр складається з чотирьох ланок. Підготовка даних, нарізання їх на вікна валідації без витоку майбутнього, прогнозування набором моделей і, нарешті, оцінювання прогнозів за єдиними метриками. Загальну архітектуру показано на рисунку 3.1.

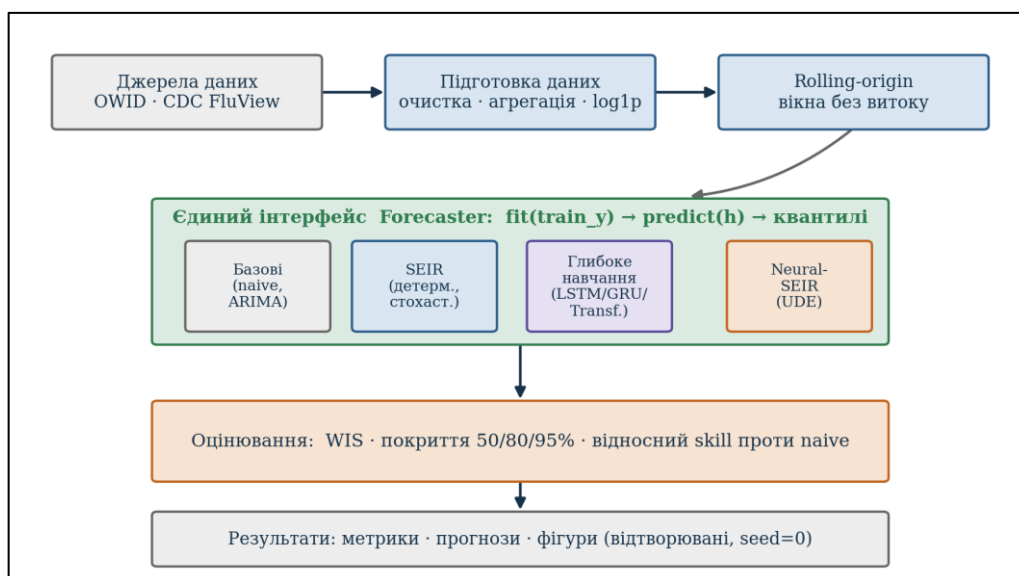


Рисунок 3.1 – Загальна архітектура системи прогнозування

Ключовим архітектурним рішенням є єдиний програмний інтерфейс для всіх моделей. Кожна модель, незалежно від її природи, реалізує той самий контракт: метод навчання приймає лише значення таргета до моменту прогнозу, а метод передбачення повертає прогноз на горизонти від першого до заданого. Прогноз завжди подається не точкою, а набором із семи квантилів рівнів 0,025, 0,1, 0,25, 0,5, 0,75, 0,9 та 0,975, що дає змогу одразу обчислювати і медіанний прогноз, і інтервали 50, 80 та 95 відсотків. Завдяки

такому уніфікованому контракту каркас оцінювання працює з усіма моделями однаково й порівнює їх на ідентичних розбиттях і горизонтах. Це не просто зручність реалізації, а методологічна вимога: будь-яка перевага складної моделі має бути виміряна в абсолютно однакових умовах.

Систему реалізовано мовою Python. Чисельне ядро спирається на бібліотеку NumPy для роботи з масивами [32] та SciPy для оптимізації й статистичних функцій [33]; підготовку й вирівнювання рядів виконано засобами pandas [34]; статистичні моделі сімейства (S)ARIMA взято з пакета statsmodels [35], а всі нейромережеві моделі, зокрема й гібрид, побудовано на PyTorch [36]. Такий вибір забезпечує, з одного боку, надійність перевірених бібліотек, а з іншого – повну прозорість і відтворюваність обчислень. Розмежування відповідальності між ланками конвеєра дає змогу змінювати або додавати моделі, не зачіпаючи логіки підготовки даних чи оцінювання.

### 3.2 Структура програмного проєкту

Програмний проєкт організовано за принципом розділення відповідальності: кожна логічна ланка конвеєра винесена в окремий пакет, що спрощує супровід і тестування. Код зосереджено в трьох основних пакетах. Пакет роботи з даними відповідає за завантаження сирих наборів, їх очищення та вирівнювання рядів, а також за нарізання даних на вікна валідації. Пакет моделей містить усі прогностні моделі за єдиним інтерфейсом: базовий контракт, базові моделі, механістичні SEIR, моделі глибокого навчання та гібрид. Пакет оцінювання реалізує метрики, прогін за схемою ковзного початку, зведення результатів у таблиці та окремі стрес-тести.

Окремо від коду зберігаються дані, результати й допоміжні матеріали. Сирі та оброблені набори лежать у каталозі даних разом із картками даних, що документують джерела й відомі особливості. Усі обчислені метрики,

прогнози, чекпойнти моделей та згенеровані фігури зберігаються в каталозі результатів, причому кожен експеримент супроводжується збереженням конфігураційним файлом із усіма параметрами запуску. Завдяки цьому будь-який результат можна відтворити, не перезапускаючи всього конвесра, а фінальні графіки рендеряться зі збережених таблиць, а не обчислюються наново.

Такий поділ забезпечує важливу властивість: логіку прогнозування й логіку оцінювання можна розвивати незалежно. Додавання нової моделі зводиться до реалізації того самого інтерфейсу в пакеті моделей і не потребує жодних змін у пакетах даних чи оцінювання. Це безпосередньо служить головній меті роботи – чесному порівнянню: усі моделі проходять через однаковий код підготовки даних і однаковий код метрик, тож відмінності в результатах відображають саме якість моделей, а не випадкові розбіжності в обробці.

### 3.3 Підготовка та характеристика даних

Експериментальну перевірку проведено на двох хворобах із контрастними режимами динаміки. Перша – COVID-19 в Україні за відкритим набором Our World in Data [20]: пандемічний спалах без вираженої річної сезонності, що дає локально релевантний і складний для прогнозу матеріал. Друга – сезонний грип у США за показником ILINet, отриманим через програмний інтерфейс Delphi Epidata [21]: регулярна щорічна повторюваність, на якій тривалий час відпрацьовують методи прогнозування. Оскільки структура зовнішніх джерел періодично змінюється, фактичний перелік колонок щоразу звіряли програмно перед обробкою, не покладаючись на очікувані назви полів.

Під час підготовки даних виявлено особливість, що вплинула на всю методику й сама є самостійним результатом роботи. Виявилося, що ряд нових випадків COVID-19 для України де-факто є тижневим по всьому

робочому вікну, а не лише з 2022 року, як здавалося спершу: значення звітуються тижневими «лампами», тобто одне ненульове число раз на сім днів з нулями між ними, а «згладжений» добовий ряд є просто цією лампою, поділеною на сім, тобто кусково-сталим семиденним плато. Хибне трактування ряду як добового призводило до вироджених метрик на горизонті в один крок, бо такий крок завжди потрапляв усередину плато. Щоб усунути це викривлення, COVID-ряд агреговано до справді тижневого: таргетом узято суму сирих нових випадків за епідеміологічний тиждень, що перевірено на узгодженість із приростом кумулятивного підсумку. Після переагрегації виродженість зникла, а характеристика даних стала чесною.

Відповідно остаточною цільовими змінними обрано тижневі нові випадки для COVID-19 та тижневий зважений відсоток грипоподібних захворювань для грипу. Прогноз будують на горизонти від одного до чотирьох тижнів. Робоче вікно COVID-ряду охоплює період з березня 2020 до грудня 2021 року – близько дев'яноста п'яти тижнів; для грипу тестовими взято п'ять до-пандемічних сезонів. Аномальні ділянки, не пов'язані з нормальною динамікою хвороби, – розрив звітності через повномасштабне вторгнення 2022 року для COVID-19 та обвал показника ILINet під час пандемії для грипу – виключено з тестування, щоб метрики відображали якість прогнозу, а не реакцію на форс-мажор.

Для навчання моделей таргет піддають логарифмічному перетворенню, яке стискає широкий діапазон значень і стабілізує дисперсію, водночас залишаючись визначеним при нульових значеннях, типових для міжсезоння.

$$z_t = \ln(1 + y_t). \quad (3.1)$$

Загальний вигляд обох рядів подано на рисунку 3.2. На верхній панелі – тижневі нові випадки COVID-19 для України з характерними хвилями та збільшеним фрагментом, що демонструє сходинкову, тижневу

природу ряду; на нижній – багаторічний ряд грипу wILI з виділеними тестовими сезонами.

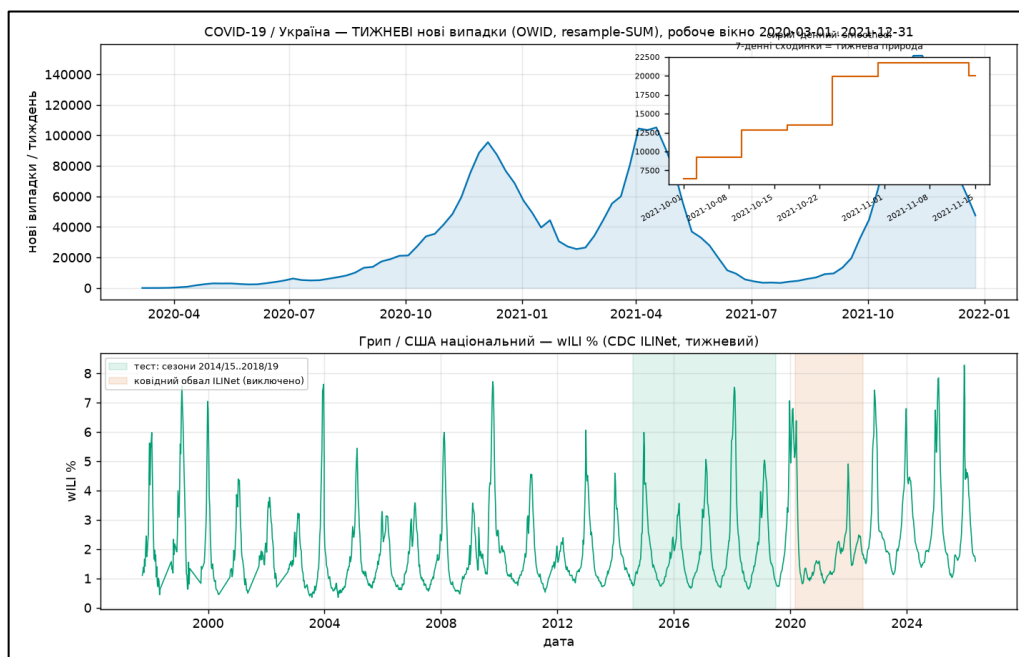


Рисунок 3.2 – Огляд даних: тижневі нові випадки COVID-19 (Україна) та показник wILI грипу (США)

Зіставлення двох панелей наочно показує, наскільки різними є режими хвороб: COVID-ряд багатохвильовий і не має регулярної сезонності, тоді як грип демонструє чітку щорічну повторюваність. Саме ця контрастність робить пару хвороб придатною для всебічної перевірки моделей: метод, що добре працює на обох, можна вважати робастним, а не підлаштованим під один тип динаміки.

### 3.4 Схема валідації без витоку майбутнього

Коректність порівняння моделей цілком залежить від схеми валідації. Головним її правилом є відсутність витоку майбутнього: для будь-якого моменту побудови прогнозу всі ознаки, параметри та коефіцієнти

масштабування мають обчислюватися виключно з даних, доступних до цього моменту включно. Порухення цього правила – наприклад, нормування ряду з використанням його майбутніх значень – робить прогноз нереалістично точним і знецінює всі висновки. Тому в системі це правило винесено в окреме архітектурне обмеження, що пронизує весь код підготовки даних.

Як схему перевірки обрано наростаюче ковзне вікно з рухомим початком (expanding rolling-origin). Її ідею показано на рисунку 3.3: початок прогнозу послідовно зсувається на один тиждень уперед, і для кожного його положення модель навчають на всій доступній історії до цього моменту, після чого вона прогнозує наступні від одного до чотирьох тижнів.

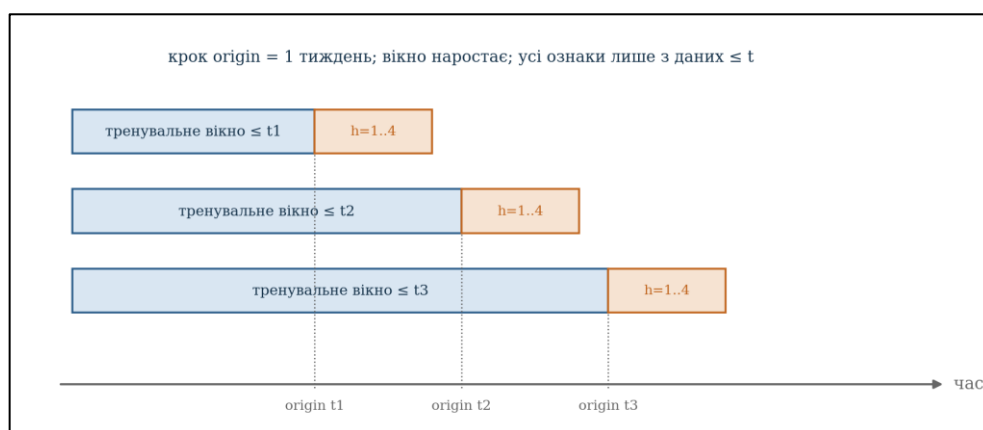


Рисунок 3.3 – Схема валідації з наростаючим ковзним початком

Така схема природна для часових рядів, бо зберігає їхній хронологічний порядок. Випадкове розбиття на навчальну й тестову частини, прийнятне в задачах без часу, тут неприпустиме: воно дало б моделі підглянути в майбутнє й оцінити її надто оптимістично. Наростання тренувального вікна, навпаки, відтворює реальну ситуацію, коли з часом нагромаджується дедалі більше спостережень. Усі моделі без винятку проходять перевірку на тих самих положеннях початку прогнозу й на тих самих горизонтах, тож їхні результати безпосередньо зіставні.

Технічно відсутність витоку забезпечено двома засобами. По-перше, коефіцієнти стандартизації для нейромережових моделей обчислюються лише з тренувального зрізу й застосовуються до решти даних, не зазираючи в майбутнє. По-друге, під час нарізання ряду на навчальні зразки кожен зразок використовує як вхід лише власне минуле, а як ціль – наступні за ним точки, не перетинаючись із майбутнім за межами свого горизонту. Разом із виключенням аномальних ділянок із тестування ця схема й утворює доказову базу роботи: оскільки порівняння проведено чесно й на ідентичних умовах, будь-яка виміряна перевага моделі є реальною, а не артефактом методики.

### 3.5 Реалізація базових і механістичних моделей

Базові моделі утворюють планку, яку кожна складніша модель зобов'язана перевершити; інакше її перевага лишається недоведеною, і це фіксують як результат. Усі базові моделі будують ймовірнісний прогноз однаково: навколо центральної оцінки на кожному горизонті припускають нормальний розподіл і обчислюють потрібні квантілі, де оберненою функцією нормального розподілу служить квантильна функція.

$$q_{h,\tau} = c_h + \sigma_h \Phi^{-1}(\tau). \quad (3.2)$$

Найпростіша базова модель – наївний прогноз (випадкове блукання): центром на всіх горизонтах є остання спостережена точка, а невизначеність зростає з горизонтом пропорційно кореню з кроку, де масштаб береться зі стандартного відхилення одно-крокових різниць ряду. Сезонний наївний прогноз повторює значення з минулого сезону: центром на горизонті є відповідна точка останнього повного сезону, а невизначеність росте з кількістю сезонів уперед.

$$c_h = y_n, \quad \sigma_h = \sigma_\Delta \sqrt{h}, \quad \sigma_\Delta = \text{std}(\Delta y), \quad (3.3)$$

$$c_h = p_{((h-1) \bmod m)+1}, \quad \sigma_h = \sigma_s \sqrt{\lceil h/m \rceil}. \quad (3.4)$$

Третя базова модель – (S)ARIMA, реалізована через клас SARIMAX бібліотеки statsmodels [35]. Її порядок підбирають один раз на старті за інформаційним критерієм Акаїке й далі фіксують для всіх положень початку прогнозу. Таке рішення є свідомим інженерним компромісом: повторний підбір порядку на кожному кроці зробив би модель неоднорідною й обчислювально дорогою, тоді як фіксований порядок забезпечує однакові умови та відтворюваність. Інтервали невизначеності тут беруть нативні, гаусові, які надає сам пакет.

Механістичну модель SEIR реалізовано на тому самому дискретному кістяку, що описано в другому розділі. Оскільки модель оперує чисельностями компартментів, а дані подано в інших одиницях (нові випадки для COVID-19, відсоток для грипу), спостережувану величину беруть як масштабовану захворюваність – добуток коефіцієнта масштабу на модельну incidence. Сам масштаб не підбирають разом з іншими параметрами, а профілюють аналітично: за фіксованої траєкторії оптимальний у сенсі найменших квадратів масштаб має замкнену форму, що зменшує розмірність задачі оптимізації.

$$k = \frac{\langle \text{inc}, \text{target} \rangle}{\langle \text{inc}, \text{inc} \rangle} = \frac{\sum_t \text{inc}_t \text{target}_t}{\sum_t \text{inc}_t^2}. \quad (3.5)$$

Решту параметрів – коефіцієнти передачі, інкубації та одужання, а також початкові чисельності експонованих та інфекційних – оцінюють методом найменших квадратів через функцію least\_squares бібліотеки SciPy [33], мінімізуючи відхилення масштабованої incidence від таргета в межах фізично осмислених обмежень.

$$\min_{\theta} \sum_t (k \cdot \text{inc}_t(\theta) - \text{target}_t)^2, \quad \theta = (\beta, \sigma, \gamma, E_0, I_0). \quad (3.6)$$

Калібрування виконують не на всій історії, а на недавньому ковзному вікні, що відповідає поточній хвилі. Це вимушене рішення: за сталого коефіцієнта передачі одна крива принципово не здатна лягти на кілька різних за формою хвиль, тож спроба підігнати модель під усю історію дала б розмиту, погану траєкторію. Стохастичну SEIR будують поверх відкаліброваної детермінованої: детерміновані потоки замінюють пуассонівськими, проганяють триста незалежних траєкторій методом Монте-Карло й за ними обчислюють емпіричні квантілі, додатково враховуючи невизначеність параметрів параметричним бутстрапом. Як буде показано в четвертому розділі, інтервали механістичних моделей виявляються дещо завузкими — це задокументоване обмеження калібрування, а не помилка реалізації.

### 3.6 Реалізація моделей глибокого навчання

Три архітектури глибокого навчання, розглянуті в другому розділі, реалізовано за тим самим єдиним інтерфейсом і об'єднано в один механізм ансамблювання. Перед навчанням ряд нарізають на навчальні зразки: входом кожного зразка є вікно минулих значень фіксованої довжини, а ціллю — наступні за ним точки на всю глибину горизонту. Щоб мережі навчалися стабільно, дані стандартизують — віднімають середнє й ділять на стандартне відхилення; важливо, що ці коефіцієнти обчислюють лише з тренувального зрізу, тож майбутнє не впливає на масштаб.

$$\tilde{y}_t = \frac{y_t - \mu_{\text{train}}}{\sigma_{\text{train}}}. \quad (3.7)$$

Навчання й прогнозування відбуваються в стандартизованому просторі, а готовий прогноз повертають у вихідні одиниці зворотним перетворенням. Усі три моделі побудовано на PyTorch [36] і навчають оптимізатором Adam [37] із мінімізацією середньоквадратичної похибки. Рекурентні моделі LSTM та GRU мають спільний кістяк: вхідну послідовність подають у рекурентний шар із прихованим станом розмірності тридцять два, з останнього прихованого стану лінійний шар одразу видає прогноз на всі горизонти. Модель Transformer спочатку проєктує одновимірний вхід у простір ознак тієї самої розмірності, додає навчані позиційні ембединги, проганяє послідовність через один шар енкодера з чотирма головами уваги, а з останнього токена лінійний шар формує прогноз.

Невизначеність прогнозу оцінюють ансамблем глибоких мереж [19]. Для кожної архітектури навчають п'ять однакових за будовою моделей, що відрізняються лише випадковою ініціалізацією; розкид їхніх прогнозів дає епістемічну складову невизначеності. Алеаторну складову оцінюють залишковою дисперсією прогнозу на кожен горизонт, після чого обидві складові поєднують і будують гаусові квантилі так само, як описано в підрозділі про оцінювання невизначеності. Принципово, що цей механізм ансамблювання застосовано однаково до всіх нейромережових моделей, тож жодна з них не має переваги в способі оцінювання невизначеності. Водночас, як показано далі, моделі глибокого навчання істотно залежать від обсягу навчальних даних, і саме цю їхню межу долає гібридна модель, до якої переходимо.

### 3.7 Реалізація гібридної моделі Neural-SEIR

Гібрид Neural-SEIR є центральною моделлю роботи, тому його реалізацію описано найдетальніше. В основу покладено ідею універсального диференціального рівняння [9]: жорсткий дискретний кістяк

SEIR зберігається повністю, а сталий коефіцієнт передачі замінює нейромережа, яка обчислює його за поточним станом системи. Принципово, що коефіцієнт передачі робиться функцією стану, а не часу. Це автономна постановка: прогноз отримують інтегруванням вивченого векторного поля вперед, від власного стану моделі, без екстраполяції в часі. Завдяки цьому уникають типового для рекурентних моделей зміщення, коли навчання й розгортання спираються на різні джерела вхідних даних. Загалом поєднання різномірних нейромережових компонентів у межах однієї прогнозовної моделі узгоджується із сучасними дослідженнями, зокрема з працями кафедри штучного інтелекту, де гібридне використання нейромережових архітектур застосовано для прогнозування часових характеристик [38].

Коефіцієнт передачі видає невеликий багат шаровий перцептрон, наведений у лістингу 3.1. На вхід подають нормований стан – частки сприйнятливих, експонованих та інфекційних, – мережа має один прихований шар із шістнадцяти нейронів з гіперболічним тангенсом, а функція `softplus` на виході гарантує, що коефіцієнт передачі завжди додатний. Повний текст модуля реалізації наведено в додатку А.

### Лістинг 3.1 – Архітектура нейромережі коефіцієнта передачі $\beta$

```
class _BetaNet(nn.Module):
    """Малий MLP: нормалізований стан (S/N, E/N, I/N) -> beta
    >= 0."""

    def __init__(self, hidden: int = 16):
        super().__init__()
        self.net = nn.Sequential(
            nn.Linear(3, hidden), nn.Tanh(),
            nn.Linear(hidden, 1), nn.Softplus(),
        )

    def forward(self, state_norm):
        # state_norm: (., 3) = (S/N, E/N, I/N); Softplus
        гарантує beta > 0
        return self.net(state_norm).squeeze(-1)
```

Динаміку інтегрують диференційовним кроком Ейлера, реалізованим засобами PyTorch. На кожному кроці мережа за поточним станом обчислює коефіцієнт передачі, після чого потоки переходів рахують і кламбують точно так само, як у чистій SEIR. Дискретні правила повністю дзеркалять механістичну модель, тож єдиною відмінністю гібрида від неї лишається змінний нейромережевий коефіцієнт замість сталого. Оскільки всі операції диференційовні, градієнт проходить наскрізь через усю траєкторію, що й уможливорює навчання мережі.

$$\begin{aligned}\beta_t &= \text{NN}_\theta(S_t/N, E_t/N, I_t/N), \\ s2e &= \min(\beta_t S_t I_t / N, S_t), \quad e2i = \min(\sigma E_t, E_t), \\ i2r &= \min(\gamma I_t, I_t).\end{aligned}\tag{3.8}$$

Навчання влаштовано так, щоб ізолювати нейромережевий внесок саме в коефіцієнті передачі. Спершу механістичні параметри інкубації та одужання й початковий стан вікна отримують теплим стартом – короткою калібрацією звичайної SEIR – і далі фіксують. Після цього навчають лише  $\beta$ -мережу. Функція втрат є траєкторною, а не поточною: вікно інтегрують від початкового стану з поточною  $\beta$ -мережею, дістають модельну траєкторію захворюваності, аналітично профілюють масштаб і мінімізують середньоквадратичне відхилення масштабованої траєкторії від таргета. Оптимізацію виконують методом Adam лише за параметрами  $\beta$  мережі впродовж фіксованого числа епох.

$$\mathcal{L}(\theta) = \frac{1}{n} \sum_{t=1}^n (k \cdot \text{inc}_t(\theta) - \text{target}_t)^2.\tag{3.9}$$

Невизначеність прогнозу оцінюють точно так само, як для решти нейромережевих моделей, що зрівнює умови порівняння. Навчають ансамбль із п'яти  $\beta$ -мереж з різною ініціалізацією; розкид їхніх прогнозів дає

епістемічну складову, а залишкова дисперсія на горизонт – алеаторну. Обидві складові поєднують і будують гаусові квантилі, які наприкінці обмежують знизу нулем, бо таргет невід'ємний.

$$\mu = \text{mean}_k \hat{y}_k, \quad \sigma = \sqrt{\sigma_{\text{epi}}^2 + \sigma_{\text{ale}}^2},$$

$$q = \text{normal\_quantiles}(\mu, \sigma). \quad (3.10)$$

Повний конвеєр навчання та прогнозування гібрида показано на рисунку 3.4: від вікна даних і теплового старту механістики через  $\beta$ -мережу й диференційовний Ейлер до траєкторного лосу, ансамблю та підсумкових квантилів.

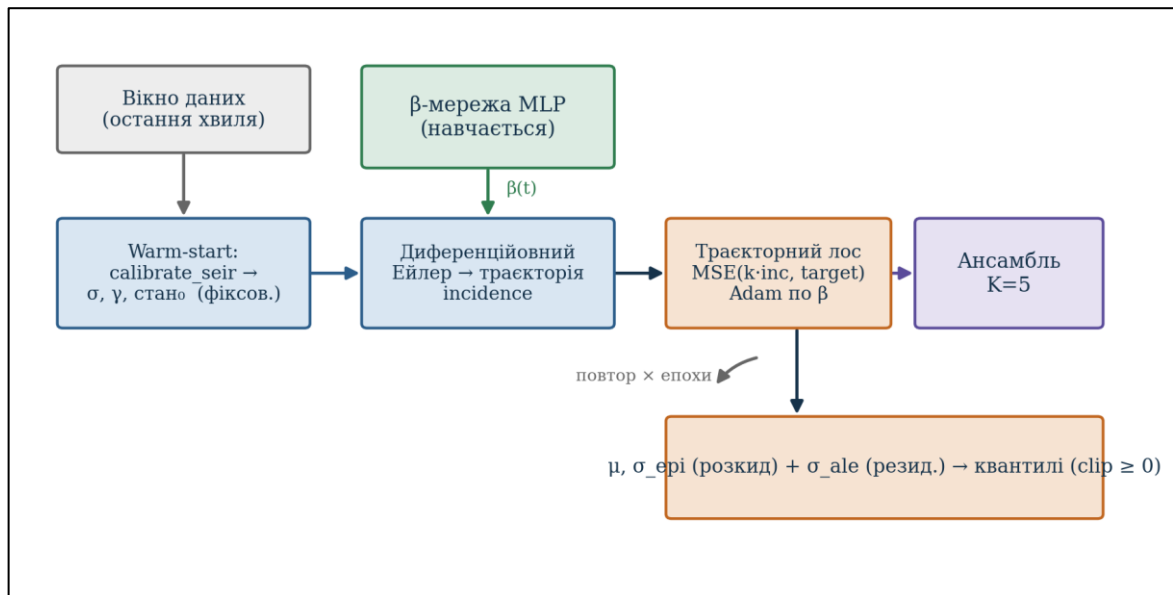


Рисунок 3.4 – Конвеєр навчання та прогнозування гібридної моделі Neural-SEIR

Запропонована архітектура гібридної моделі має чітку методологічну цінність. Фізичні обмеження – збереження населення й невід'ємність компартментів – вшиті в структуру обчислень, а не додано як штраф, тому вони виконуються безумовно. Нейромережевий внесок цілком ізольовано в

коефіцієнті передачі, тож порівняння з чистою SEIR вимірює рівно одну річ: що саме додає змінний коефіцієнт замість сталого. А оскільки невизначеність оцінюють спільним для всіх моделей механізмом, перевага гібрида не може пояснюватися кращим способом її оцінювання. Результати цього порівняння наведено в наступному розділі.

### 3.8 Відтворюваність і тестування

Відтворюваність результатів закладено в архітектуру системи як окрему імперативну вимогу, що гарантує валідність наукових висновків. Усі потенційні джерела випадковості й нелінійності обчислень – від ініціалізації вагових коефіцієнтів нейромереж та випадкового перемішування зразків (shuffling) у процесі батчезації до генерації стохастичних траєкторій епідемічного процесу – жорстко керуються фіксованим початковим значенням генератора псевдовипадкових чисел (seed numbers). Це забезпечує повний детермінізм обчислювального графа на рівні використовуваних бібліотек глибокого навчання та системних потоків, тож кожен повторний запуск конвеєра на ідентичних даних дає абсолютно однаковий результат.

Повна конфігурація кожного експерименту, включаючи гіперпараметри моделей, архітектурні складові та налаштування оптимізаторів, серіалізується й зберігається в окремому декларативному файлі (наприклад, у форматі YAML або JSON) безпосередньо поряд із отриманими результатами. Самі ж результати – розраховані інтегральні метрики, прогнозні часові ряди, матриці помилок та збережені стани ваг (чекпойнти моделей на різних епохах навчання) – автоматично структурують і каталогізують у централізованому сховищі артефактів. Завдяки такій ізоляції даних будь-яку таблицю, графік чи аналітичну діаграму, представлені в роботі, можна миттєво відтворити або верифікувати без повторного ресурсомісткого прогону всього

обчислювального конвеєра. Фінальні графічні фігури рендеряться безпосередньо зі збережених сирих даних, а не перераховуються наново, що повністю виключає ризик виникнення розбіжностей між емпіричним матеріалом у тексті дисертації та поточним станом програмного коду.

Коректність програмної реалізації та стійкість обчислювальних алгоритмів підтверджено розгалуженим набором автоматичних тестів (unit та integration tests), які охоплюють як ізольовані математичні функції, так і наскрізну взаємодію макромодулів системи. Тестове покриття фокусується на автоматичній верифікації ключових архітектурних інваріантів та бізнес-логіки. Зокрема, тести контролюють суворе дотримання контракту єдиного програмного інтерфейсу (API) всіма без винятку моделями ростера, що є критичним для їхньої безшовної взаємозамінності. На рівні епідеміологічного ядра системи виконується ретельна перевірка законів збереження: точного дотримання чисельності популяції та безумовної невід'ємності компартментів у дискретній різницевій схемі моделі SEIR за будь-яких параметрів інтенсивності.

Окрему увагу приділено тестуванню часової чистоти експерименту: архітектура тестів гарантує повну відсутність явища «витоку майбутнього» (data leakage) – розрахунок коефіцієнтів масштабування, нормалізація ознак та формування навчальних підмножин здійснюються виключно на основі історичних даних, доступних до фіксованого моменту генерації прогнозу. Додатково розроблено стрес-тести для перевірки поведінки системи в крайових умовах та на аномальних вхідних даних (Edge Cases) – наприклад, при подачі надто коротких часових рядів, за наявності пропусків або нульової дисперсії. У таких сценаріях модулі мають коректно й контрольовано відмовлятися від ініціалізації чи навчання із генерацією відповідних логів, замість видачі помилкових або викривлених результатів. На момент фінального збирання системи увесь набір тестів проходить успішно, демонструючи стабільність кодової бази. Методичні засади автоматизованого тестування програмних модулів складних інформаційних

систем, на які спирається застосована у роботі практика верифікації, детально розвинуто й обґрунтовано в фундаментальних дослідженнях кафедри [39].

Було комплексно спроектовано, математично обґрунтовано та програмно реалізовано цілісну інформаційну систему прогнозного моделювання. Усі гетерогенні базові моделі, включно з класичними рекурентними архітектурами, трансформерами та розробленим центральним гібридом Neural-SEIR, успішно зведено під єдиний уніфікований інтерфейс абстракції. Побудовано та валідовано гнучку схему ковзного контролю, яка унеможливує зміщення оцінок через часовий витік, а також розгорнуто надійні інструменти забезпечення відтворюваності, логування артефактів та автоматичного тестування коду. Архітектуру системи побудовано за модульним принципом, що забезпечує просте розширення її функціональних можливостей. Це дозволяє інтегрувати нові моделі прогнозування та методи обробки даних без істотних змін у програмному коді. Єдиний програмний інтерфейс гарантує однакову процедуру навчання, прогнозування та оцінювання для всіх реалізованих моделей. Автоматизація експериментального конвеєра забезпечує послідовність виконання обчислень і спрощує проведення серій порівняльних досліджень. У сукупності це створює надійну основу для об'єктивного аналізу ефективності різних підходів до прогнозування епідемічної динаміки. Створений програмно-методологічний базис дозволяє перейти від етапу конструювання до практичного застосування системи. Саме на цій основі у наступному розділі роботи проведено масштабне експериментальне дослідження, виконано порівняльний аналіз моделей на реальних епідемічних даних та інтерпретовано отримані результати.

## 4 ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

### 4.1 Основне порівняння моделей

Спершу розглянемо повну картину – як усі моделі ростера співвідносяться з наївною планкою на заголовному горизонті. Результати для COVID-19 наведено в таблиці 4.1, для грипу – у таблиці 4.2; додатне значення відносного skill означає перевагу над наївним прогнозом. Повні результати по всіх горизонтах для обох хвороб наведено в додатку Б.

Таблиця 4.1 – Метрики моделей для COVID-19 на горизонті 4 тижні

| Модель             | WIS     | Відносний skill |
|--------------------|---------|-----------------|
| transformer        | 17 526  | +0,037          |
| naive              | 18 208  | 0,000           |
| gru                | 23 554  | −0,294          |
| arima              | 24 497  | −0,345          |
| lstm               | 28 439  | −0,562          |
| neural_seir        | 32 149  | −0,766          |
| seir               | 38 499  | −1,114          |
| stochastic_seir    | 45 992  | −1,526          |
| seasonal_naive_m52 | 170 270 | −8,352          |

Таблиця 4.2 – Метрики моделей для грипу на горизонті 4 тижні

| Модель             | WIS   | Відносний skill |
|--------------------|-------|-----------------|
| lstm               | 0,263 | +0,368          |
| gru                | 0,283 | +0,321          |
| transformer        | 0,295 | +0,292          |
| sarima52           | 0,347 | +0,167          |
| naive              | 0,417 | 0,000           |
| seasonal_naive_m52 | 0,466 | −0,117          |
| neural_seir        | 0,518 | −0,243          |
| seir               | 0,624 | −0,497          |
| stochastic_seir    | 1,476 | −2,542          |

Картина для двох хвороб виявилася принципово різною. На грипі моделі глибокого навчання впевнено домінують: найкраща з них, LSTM, перевершує наївну планку на тридцять сім відсотків, а виражену річну сезонність добре вловлює і сезонна SARIMA. Це очікувано: грип має

регулярну повторювану структуру, з якої моделі є чого навчитися. На тижневому ряді COVID-19 на горизонтах від одного до чотирьох тижнів ситуація протилежна: наївний персистенс виявляється дуже сильною планкою, яку ледь перевершує лише Transformer, тоді як решта складних моделей, зокрема механістичні й гібрид, їй програють. Загальну картину унаочнює рисунок 4.1.

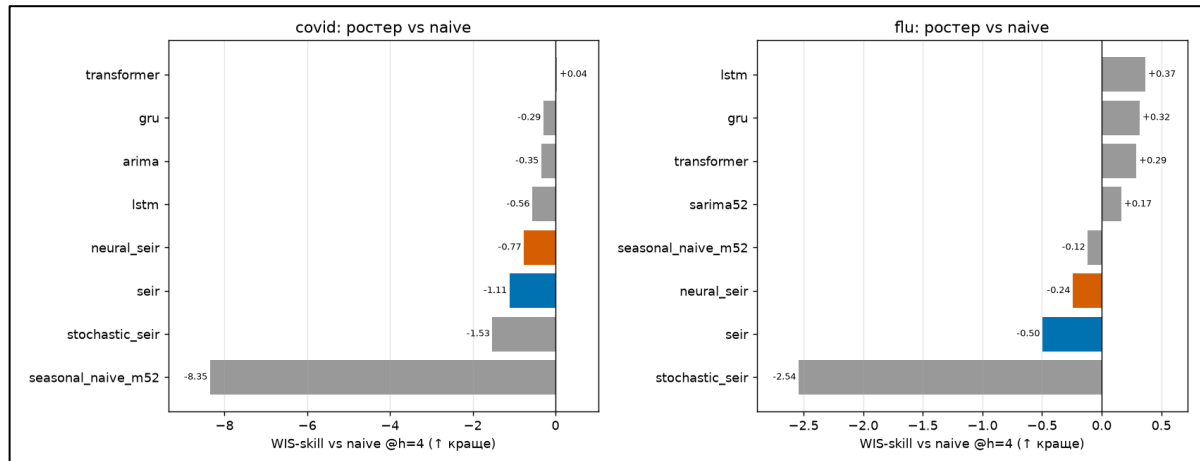


Рисунок 4.1 – Порівняння моделей ростера з наївною планкою на обох хворобах

Цей результат не приховано, а зафіксовано як чесний висновок. Його пояснення цілком прозоре: тижневий епідемічний ряд на короткому горизонті близький до випадкового блукання, для якого персистенс є майже оптимальним прогнозом – це відома властивість таких рядів, а не провал складних моделей. Важливо, що сильна наївна планка на COVID-ряді стосується саме абсолютного порівняння на заголовному горизонті. Справжню цінність гібридної моделі та її нішу розкривають наступні підрозділи, де порівняння ведуть не з наївним прогнозом узагалі, а з прямими конкурентами – чистою механістикою та глибоким навчанням за різного обсягу даних.

## 4.2 Цінність нейромережевого коефіцієнта передачі

Найважливіше для тези роботи порівняння – це зіставлення гібрида Neural-SEIR із чистою механістичною SEIR. Воно є зразковим прикладом контрольованого експерименту: обидві моделі мають однаковий кістяк, однакові параметри інкубації та одужання, однаковий спосіб калібрування й однакове оцінювання невизначеності, а відрізняються рівно однією деталлю – у чистій SEIR коефіцієнт передачі сталий, а в гібриді його обчислює нейромережа за поточним станом. Тому будь-яка різниця в точності є мірою саме цінності нейромережевого коефіцієнта. Усереднену по горизонтах зважену інтервальну оцінку для обох моделей наведено в таблиці 4.3.

Таблиця 4.3 – Гібрид Neural-SEIR проти чистої SEIR (середній WIS по горизонтах)

| Хвороба  | SEIR   | Neural-SEIR | Покращення |
|----------|--------|-------------|------------|
| COVID-19 | 22 443 | 19 647      | +12,5 %    |
| Грип     | 0,4414 | 0,3692      | +16,4 %    |

Результат однозначний: гібрид стабільно перевершує чисту механістику на обох хворобах – на дванадцять із половиною відсотків на COVID-19 і на шістнадцять із половиною відсотків на грипі. Це не випадковий виграш на одному наборі, а узгоджена перевага на двох контрастних режимах динаміки, що робить висновок надійним. При цьому перевага моделі зберігається незалежно від особливостей часових рядів і характеру поширення захворювання. Отримані результати свідчать, що поєднання механістичних знань із можливостями глибокого навчання забезпечує вищу точність прогнозування, ніж використання лише класичної епідеміологічної моделі. Наочно різницю показано на рисунку 4.2.

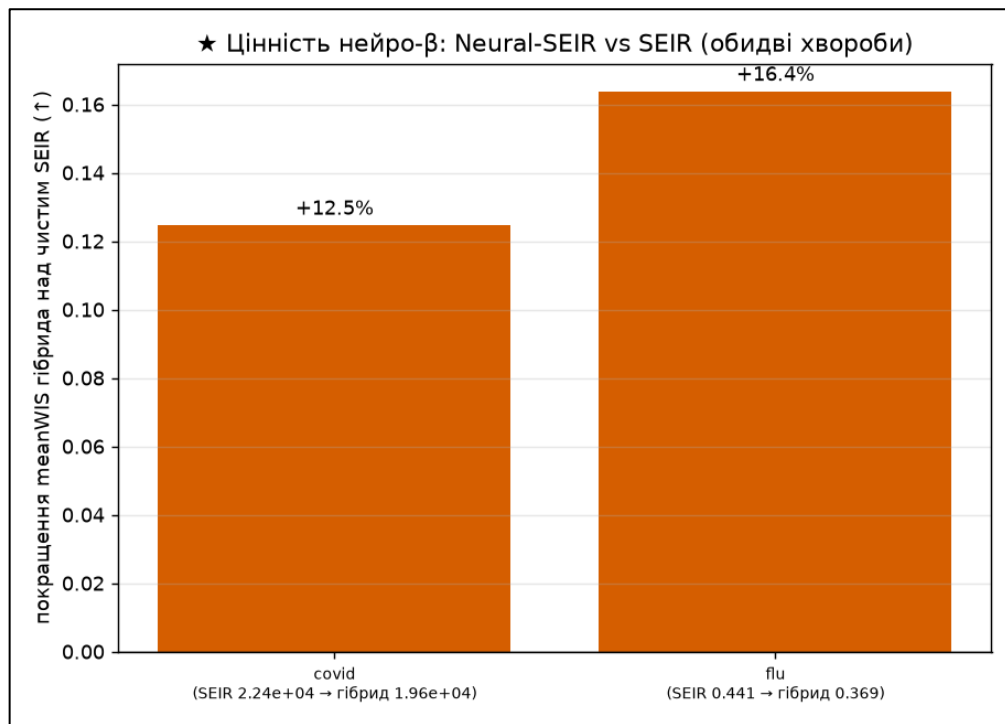


Рисунок 4.2 – Гібрид Neural-SEIR проти чистої механістичної SEIR за середнім WIS

Природа цього покращення цілком зрозуміла. Сталий коефіцієнт передачі змушує одну криву описувати всю історію, тож на багатохвильовому ряді вона неминуче виходить компромісною й розмитою. Нейромережевий коефіцієнт, навпаки, вільно змінюється відповідно до фази епідемії: він зростає на підйомі хвилі й спадає на спаді, тож модель здатна відтворити кілька різних за формою хвиль однією структурою. Оскільки нейромережевий внесок ізольовано саме в коефіцієнті передачі, а невизначеність оцінюють спільним для обох моделей механізмом, виміряну перевагу не можна пояснити нічим іншим, окрім переходу від сталого коефіцієнта до змінного. Це й становить головний, доведений та вимірюваний внесок роботи: запропонований гібрид додає реальну прогностичну цінність порівняно з класичною механістикою. Здобуті результати узгоджуються із загальним напрямом кафедри щодо

застосування гібридних нейромережових моделей у задачах прогнозування [38].

### 4.3 Ефективність за обсягом даних

Головний експеримент роботи перевіряє центральну гіпотезу про вищу ефективність гібрида за даними. Для цього проведено абляцію обсягу навчальної вибірки: ту саму задачу розв'язують за різної кількості доступних тижнів історії, від дуже малої до повної, зберігаючи однакові положення початку прогнозу. Це моделює реальну ситуацію раннього спалаху, коли спостережень обмаль. Усереднену по горизонтах зважену інтервальну оцінку залежно від обсягу вікна наведено в таблиці 4.4, де прочерк означає, що моделі бракує даних для навчання.

Таблиця 4.4 – Середній WIS залежно від обсягу навчальної вибірки W

| W, тижнів | naive  | seir   | neural_seir | lstm   | gru    |
|-----------|--------|--------|-------------|--------|--------|
| 6         | 15 002 | 17 164 | 20 372      | –      | –      |
| 9         | 14 216 | 20 247 | 17 469      | –      | –      |
| 12        | 13 691 | 22 443 | 19 647      | 32 427 | 34 302 |
| 18        | 11 808 | 22 443 | 19 647      | 23 543 | 19 748 |
| 24        | 11 830 | 22 443 | 19 647      | 29 852 | 26 369 |
| 48        | 12 231 | 22 443 | 19 647      | 14 568 | 13 571 |
| ≈ 90      | 11 674 | 22 443 | 19 647      | 17 658 | 14 935 |

Таблиця виявляє три чіткі закономірності. По-перше, моделі глибокого навчання катастрофічно залежні від обсягу даних: за дванадцяти тижнів історії LSTM і GRU дають оцінку близько тридцяти двох – тридцяти чотирьох тисяч, що майже втричі гірше за фізичні моделі, а за меншого обсягу вони взагалі не навчаються. По-друге, фізична й гібридна моделі робастні: вони тримають точність навіть на найменших вікнах. По-третє, з нарощуванням даних глибоке навчання різко покращується й зрештою обганяє фізику – відбувається кросовер десь у діапазоні від вісімнадцяти до

сорока восьми тижнів. Цей перехід наочно показано на рисунку 4.3, який є головним результатом роботи.

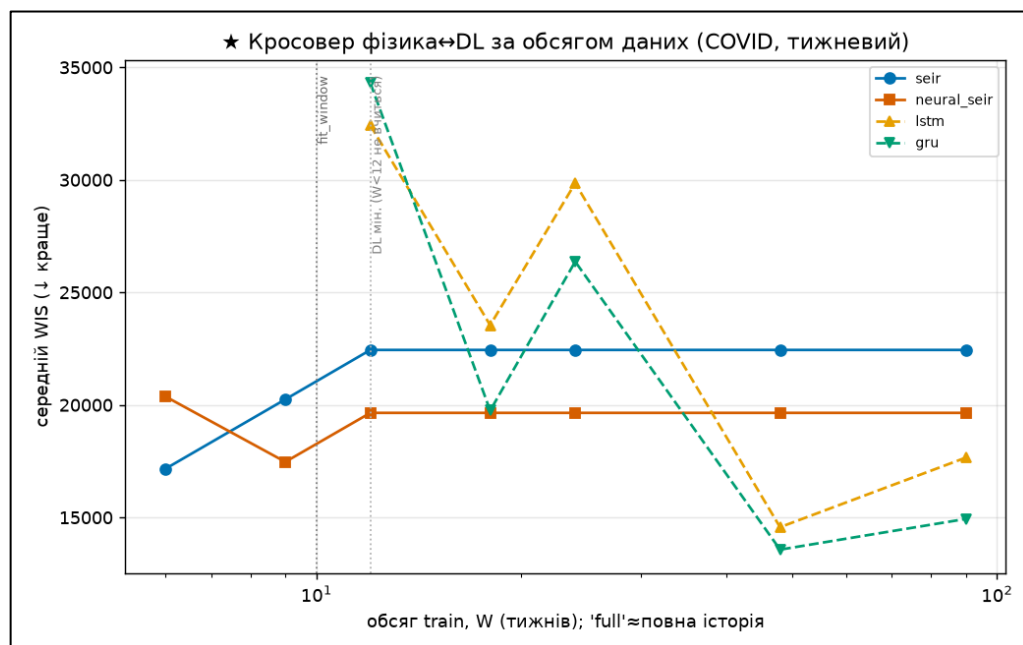


Рисунок 4.3 – Кросовер фізичних і нейромережевих моделей за обсягом даних

Тут потрібне важливе чесне застереження. На тижневому ряді COVID-19 наївний персистенс лишається сильнішим за всіх на кожному обсязі вікна, тож перевага фізики й гібрида над глибоким навчанням за дефіциту даних є відносною – це вища ефективність за даними, а не абсолютна перемога над наївною планкою. Слід також пояснити, чому крива фізичних моделей на таблиці пласка: за обсягу, що перевищує розмір вікна калібрування, ці моделі просто ігнорують надлишок історії й дивляться на фіксоване недавнє вікно, тож їхня незмінність – не магія, а наслідок конструкції. Саме тому сітку обсягів навмисно опущено нижче розміру вікна калібрування, де й видно справжню ефективність за даними.

З цього експерименту випливає чітко окреслена ніша гібрида. Він не претендує бути найкращим скрізь, але має конкретну й цінну сферу переваги: ранній спалах з малою кількістю чистих даних, коли глибоке

навчання ще безпорадне, а фізичний кістяк дозволяє гібриду залишатися притомним і точнішим за чисту механістику. Це і є практична відповідь на запит, сформульований у вступі, – метод, що зберігає якість саме там, де її найважче забезпечити.

#### 4.4 Межі методу

Наукова цінність роботи полягає не лише в перевагах гібрида, а й у чесному окресленні його меж. Нижче розглянуто три аспекти, у яких метод поступається або поводить себе неідеально: стійкість до шуму, поведінку на холодному старті крізь розворот хвилі та калібрування інтервалів. Кожен із них подано не як ваду, що знецінює результат, а як конкретний напрям подальшого вдосконалення.

Перший аспект – стійкість до шуму. Гіпотезу про те, що фізичний кістяк робить гібрид стійкішим до зашумлення даних, експеримент не підтвердив. У таблиці 4.5 наведено середню зважену інтервальну оцінку за різного рівня лог-нормального шуму в навчальних даних.

Таблиця 4.5 – Середній WIS за різного рівня шуму в навчальних даних

| Шум $\sigma$ | naive  | seir   | neural_seir | lstm   | gru    |
|--------------|--------|--------|-------------|--------|--------|
| 0,00         | 11 674 | 22 443 | 19 647      | 17 658 | 14 935 |
| 0,20         | 13 662 | 27 992 | 25 820      | 13 012 | 11 705 |
| 0,40         | 20 911 | 70 697 | 80 473      | 17 027 | 18 225 |

Як видно, найробастнішими виявилися саме моделі глибокого навчання: ансамбль усереднює шум по багатьох зразках, тому за сильного зашумлення LSTM і GRU деградують найменше. Фізична й гібридна моделі, навпаки, за великого шуму погіршуються найсильніше, бо калібрування на невеликому вікні чутливе до спотворень. Це чесно зафіксована межа, унаочнена на рисунку 4.4.

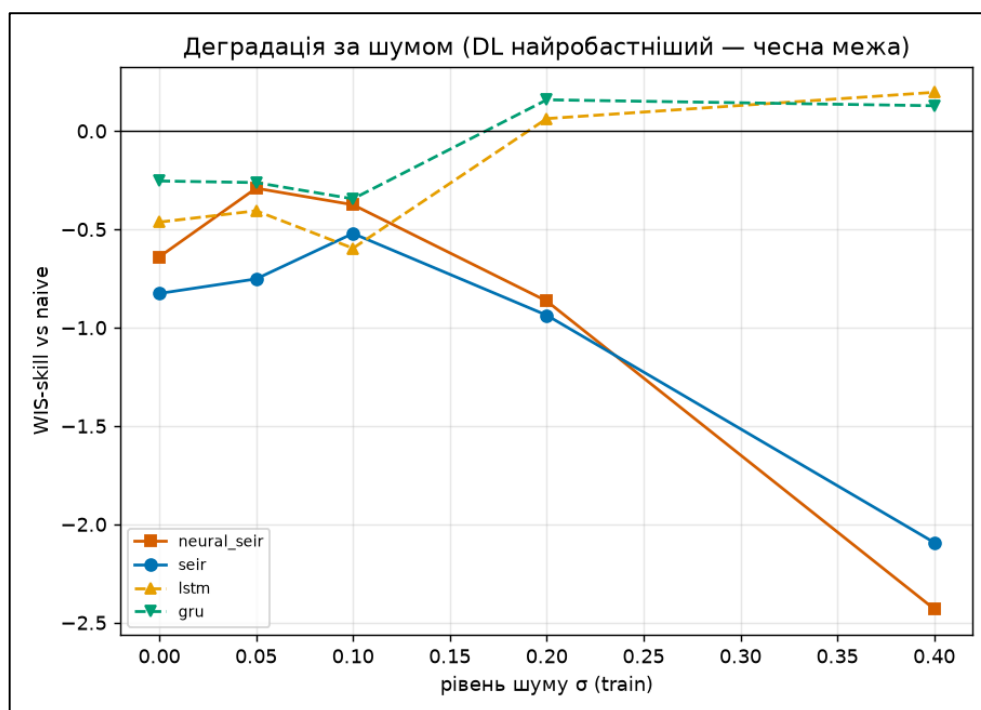


Рисунок 4.4 – Стійкість моделей до шуму в навчальних даних

Другий аспект – холодний старт, тобто прогнозування крізь розворот хвилі з малої історії, коли модель ще не бачила перелому тренду. Середню оцінку для першої хвилі 2020 року та хвилі «дельта» 2021 року наведено в таблиці 4.6.

Таблиця 4.6 – Середній WIS на холодному старті крізь розворот хвилі

| Хвиля         | naive  | seir   | gru    | neural_seir |
|---------------|--------|--------|--------|-------------|
| Перша (2020)  | 10 190 | 37 704 | 39 246 | 42 973      |
| Дельта (2021) | 23 702 | 77 390 | 61 379 | 49 427      |

На розвороті хвилі домінує наївний прогноз, і це закономірно: проста персистенс-планка не схильна до різких екстраполяцій. Чиста механістична SEIR тут поводить себе найгірше – вона катастрофічно перестрілює, екстраполуючи експоненційне зростання вгору вже після перелому. Проте саме тут нейромережевий коефіцієнт виявляє несподівану користь: він гасить цей переліт. На хвилі «дельта» гібрид різко кращий за чисту SEIR, близько сорока дев'яти тисяч проти сімдесяти семи. На першій хвилі 2020

року гібрид приблизно дорівнює SEIR, бо моделі ще немає на чому навчити коефіцієнт передачі. Поведінку обох хвиль показано на рисунках 4.5 та 4.6.

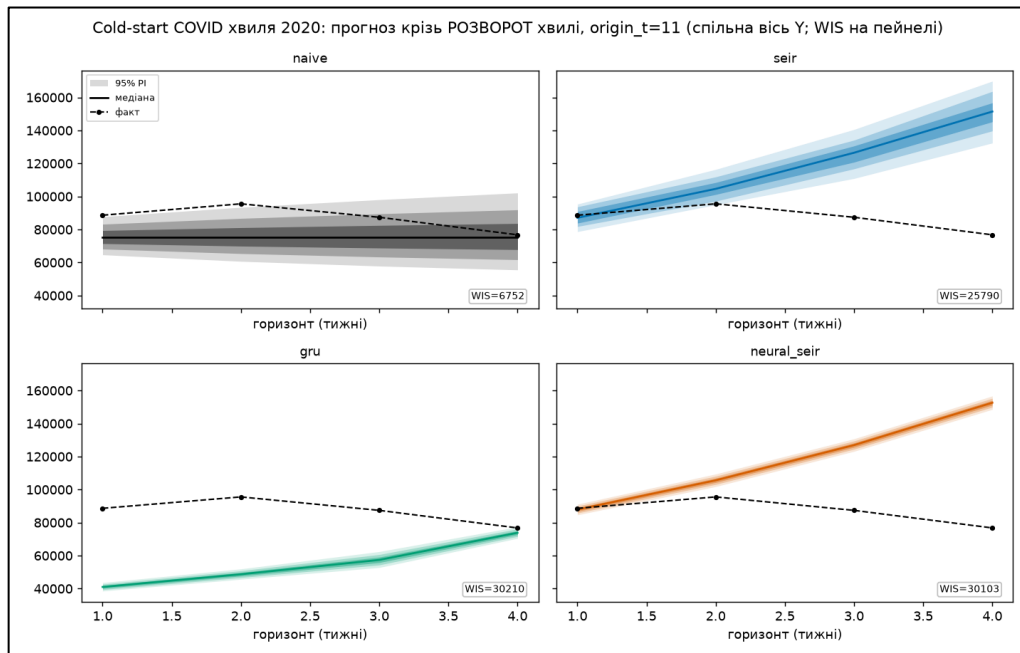


Рисунок 4.5 – Холодний старт: перша хвиля 2020 року

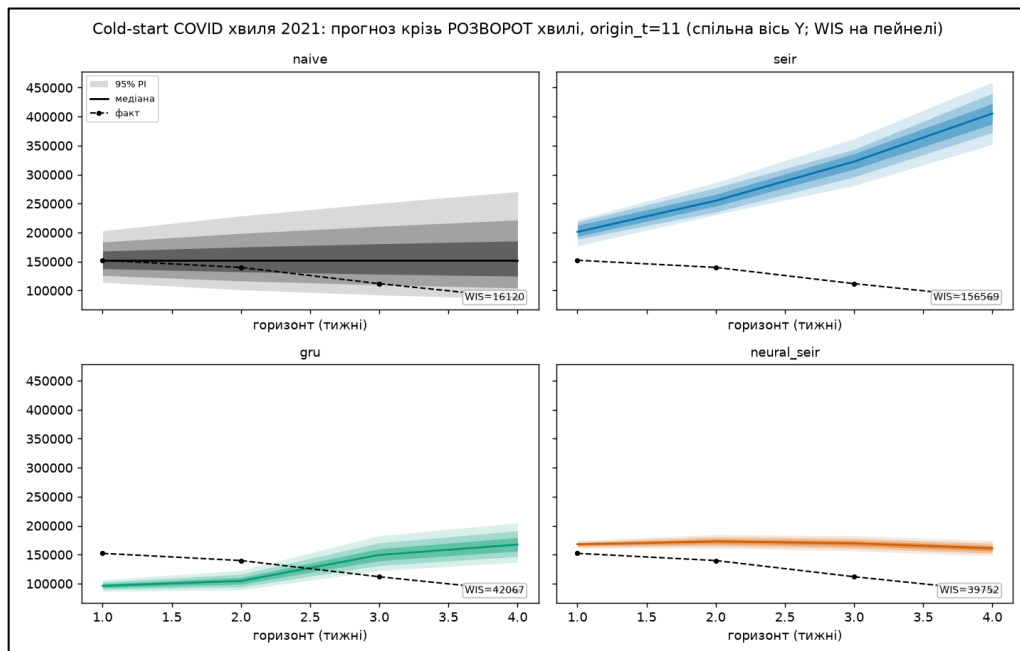


Рисунок 4.6 – Холодний старт: хвиля «дельта» 2021 року

Отже, навіть у несприятливому для механістики сценарії нейромережевий коефіцієнт радше допомагає, ніж шкодить, пом'якшуючи головну ваду чистої SEIR. Це додатковий, хоч і нішевий, аргумент на користь гібридного підходу.

Третій аспект – калібрування інтервалів. Емпіричне покриття для COVID-19, тобто фактична частка спостережень, що потрапили в інтервали номінальних рівнів п'ятдесят, вісімдесят і дев'яносто п'ять відсотків, наведено в таблиці 4.7.

Таблиця 4.7 – Емпіричне покриття інтервалів для COVID-19 проти номінального

| Модель      | 50 % | 80 % | 95 % |
|-------------|------|------|------|
| naive       | 0,44 | 0,86 | 1,00 |
| seir        | 0,18 | 0,38 | 0,41 |
| neural_seir | 0,18 | 0,31 | 0,37 |
| gru         | 0,11 | 0,22 | 0,32 |

Інтервали наївного прогнозу виявилися надмірно консервативними, тобто систематично ширшими, ніж це впливає з реальної варіативності процесу. Натомість інтервали фізичної та гібридної моделей демонструють протилежну поведінку – вони є помітно звуженими, що проявляється у стабільно заниженому фактичному покритті відносно номінального рівня довіри. Така асиметрія свідчить про різну природу похибок у підходах. Якщо наївний прогноз «перестраховується», то модельно-обґрунтовані підходи недообліковують розкид.

Ключова причина полягає в тому, що залишки, оцінені всередині навчального вікна тієї самої моделі, не відтворюють повної структури невизначеності, яка виникає під час екстраполяції за межі історичного інтервалу. Вони відображають лише внутрішню варіацію в межах вже спостережених даних і тому систематично занижують реальну похибку прогнозу на кінці траєкторії. Це є добре відоме обмеження стандартної

процедури калібрування, яке емпірично підтверджується результатами, наведеними на рисунку 4.7.

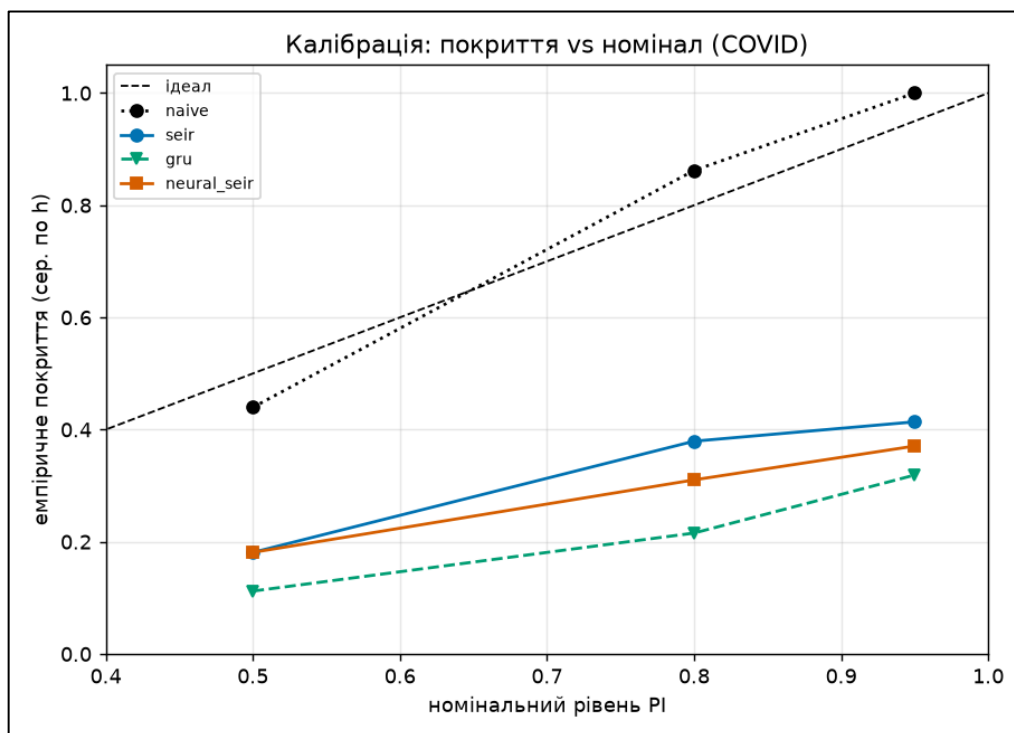


Рисунок 4.7 – Калібрування інтервалів: емпіричне покриття проти номінального

Усі три обмеження фактично задають природні напрями подальшого вдосконалення. Проблему калібрування покриття можна усунути через конформні підходи, окрему калібрувальну декомпозицію або розширення ансамблю моделей; чутливість до шуму – за рахунок попереднього згладжування даних або використання більш робастної функції втрат; а особливості поведінки на етапі холодного старту – через введення явних обмежень на динаміку зміни коефіцієнта передачі. Водночас жодне з цих обмежень не заперечує основного результату, а радше визначає потенційні напрями його подальшого розвитку.

## 4.5 Обговорення результатів

Зведемо отримані результати, відповівши на три ключові питання, які природно постають перед роботою. Така форма обговорення дозволяє оцінити і повноту виконання теми, і реальну вагу запропонованого методу, не уникаючи незручних місць.

Перше питання: чи виконано заявлену тему – стохастичне моделювання та глибоке навчання для прогнозування епідемій? Так, у повному обсязі. Реалізовано і стохастичне моделювання у вигляді стохастичної SEIR з пуассонівськими переходами, Монте-Карло-ансамблем і параметричним бутстрапом, і глибоке навчання у вигляді трьох архітектур з ансамблюванням, і – головне – їхнє поєднання в гібриді Neural-SEIR. Усе це перевірено за суворою методикою без витоку майбутнього, з імовірнісними метриками та стрес-тестами на двох контрастних хворобах. Тема розкрита цілком.

Друге питання: чи має сенс розроблений алгоритм? Так, він має чіткий і доведений сенс та власну нішу. Гібрид стабільно перевершує чисту механістику на обох хворобах, додаючи вимірювану цінність нейромережовим коефіцієнтом передачі. Він суттєво ефективніший за глибоке навчання за дефіциту чистих даних, тобто саме в тій ситуації раннього спалаху, де прогноз найпотрібніший. Нарешті, нейромережовий коефіцієнт пом'якшує катастрофічний переліт чистої SEIR на розвороті хвилі. При цьому метод методологічно коректний: фізичні обмеження вшиті структурно, а невизначеність оцінено ансамблево. Це не іграшка, а обґрунтований гібрид із доведеною перевагою.

Третє питання: чи кращий він за готові рішення? Відповідь чесна й виважена: частково й нішево. Гібрид завжди кращий за готові механістичні рішення – чисту та стохастичну SEIR. Він кращий за глибоке навчання, але лише в ніші дефіциту чистих даних, де готові нейромережові моделі катастрофічно програють. Водночас він не кращий універсально: за

достатку даних виграють готові моделі глибокого навчання, особливо на сезонному грипі, а на тижневому COVID-ряді на коротких горизонтах найпростіша готова планка – наївний персистенс – лишається сильнішою за всі складні моделі. Внесок гібрида є відносним і нішевим – це вища ефективність за даними та стабільна перевага над механістикою, доведені чесним порівнянням.

Саме чесність окреслення меж робить результат науково вагомим. Робота не оголошує гібрид універсальним розв'язком, а точно вказує, де він кращий, де гірший і чому. До цього додається самостійний методологічний результат – чесна характеристика тижневої природи ряду COVID-19 для України, яка усунула викривлення метрик і підвищила довіру до всіх подальших висновків. Отримані результати підтверджують, що поєднання механістичних моделей із методами глибокого навчання є перспективним напрямом розвитку систем прогнозування епідемій. Запропонований підхід не лише забезпечує підвищення точності прогнозів у складних умовах обмежених даних, а й зберігає інтерпретованість, притаманну класичним епідеміологічним моделям. Розроблена інформаційна система може слугувати основою для подальших досліджень, спрямованих на інтеграцію додаткових джерел даних, удосконалення механізмів оцінювання невизначеності та розширення спектра підтримуваних моделей. Практична цінність роботи полягає у можливості використання запропонованого підходу для підтримки прийняття рішень під час моніторингу та прогнозування розвитку епідемічних процесів. У сукупності це свідчить про досягнення поставленої мети та виконання всіх завдань кваліфікаційної роботи.

## ВИСНОВКИ

У кваліфікаційній роботі розв'язано задачу прогнозування динаміки епідемій за умов коротких та нерегулярно спостережуваних рядів захворюваності. Розроблено, програмно реалізовано та всебічно верифіковано гібридну стохастично-нейромережеву модель Neural-SEIR, а також виконано її коректне порівняння з повним набором еталонних підходів.

У межах роботи реалізовано як стохастичне епідеміологічне моделювання на основі SEIR із пуассонівськими переходами та Монте-Карло ансамблюванням, так і підходи глибокого навчання, представлені архітектурами LSTM, GRU та Transformer з ансамблюванням прогнозів. Окремо розроблено їх поєднання у вигляді гібридної моделі Neural-SEIR, побудованої на основі універсального диференціального рівняння. Усі моделі приведено до єдиного програмного інтерфейсу та оцінено за єдиною методологією з наростаючим ковзним стартом без витоку інформації з майбутнього, на ідентичних розбиттях та з використанням імовірнісних метрик, зокрема зваженої інтервальної оцінки.

Основний кількісний результат полягає у стабільному перевищенні гібридною моделлю Neural-SEIR класичної механістичної SEIR-моделі на обох досліджених захворюваннях. Середнє покращення зваженої інтервальної оцінки становить 12.5% для COVID-19 та 16.5% для грипу. Оскільки відмінність гібридної моделі від базової SEIR полягає лише у заміні сталого коефіцієнта передачі на нейромережевий, отриманий ефект безпосередньо відображає внесок запропонованого нейромережевого розширення.

Додатково встановлено, що гібридна модель має чітко визначену область ефективності. Результати абляційного експерименту показали високу залежність чистих моделей глибокого навчання від обсягу даних, тоді як за умов дефіциту спостережень вони суттєво поступаються фізичним

підходам. Натомість гібрид Neural-SEIR завдяки наявності механістичного каркаса зберігає стабільність і демонструє значно кращу придатність саме в режимі раннього спалаху, коли якість прогнозу є критично важливою. Також підтверджено, що нейромережеве узагальнення коефіцієнта передачі дозволяє зменшити ефект надмірного перелаштування SEIR-моделі на етапах зміни епідемічної динаміки.

Порівняння з існуючими підходами виконано коректно та без завищених узагальнень. Гібридна модель послідовно перевершує класичні механістичні рішення, є ефективнішою за глибокі моделі в умовах обмежених даних, однак не є універсально домінуючою. За достатнього обсягу даних перевагу можуть отримувати моделі глибокого навчання, а на коротких горизонтах для деяких рядів конкурентною залишається наївна екстраполяція. Така збалансована оцінка визначає реальні межі застосовності підходу.

Самостійним методологічним результатом роботи є уточнення властивостей даних. Встановлено фактичну тижневу структуру ряду нових випадків COVID-19 в Україні, що дозволило зменшити викривлення метрик на коротких горизонтах і підвищити коректність подальших висновків.

Виявлені обмеження методу визначають напрями подальшого розвитку, серед яких підвищення якості калібрування інтервалів, підвищення стійкості до шуму та покращення поведінки на етапі холодного старту, а також розширення моделі на просторово-структуровані компартментні системи й інтеграцію додаткових джерел даних. Практична цінність роботи полягає у створенні відтворюваного інструмента прогнозування епідемій, придатного для застосування в умовах обмеженої інформації, а також у формалізації протоколу чесного порівняння моделей, який може бути використаний у системах епідеміологічного нагляду та планування ресурсів охорони здоров'я.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Evaluating epidemic forecasts in an interval format / J. Bracher et al. *PLOS computational biology*. 2021. Vol. 17, no. 2. P. e1008618. URL: <https://doi.org/10.1371/journal.pcbi.1008618> (date of access: 23.06.2026).
2. A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States / N. G. Reich et al. *Proceedings of the national academy of sciences*. 2019. Vol. 116, no. 8. P. 3146–3154. URL: <https://doi.org/10.1073/pnas.1812594116> (date of access: 28.06.2026).
3. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. series A, containing papers of a mathematical and physical character*. 1927. Vol. 115, no. 772. P. 700–721. URL: <https://doi.org/10.1098/rspa.1927.0118> (date of access: 28.06.2026).
4. Anderson R. M., May R. M. Infectious diseases of humans: dynamics and control. Oxford : Oxford University Press, 1991. 757 p.
5. Keeling M. J., Rohani P. Modeling infectious diseases in humans and animals. Princeton : Princeton University Press, 2008. 408 p.
6. Hochreiter S., Schmidhuber J. Long short-term memory. *Neural computation*. 1997. Vol. 9, no. 8. P. 1735–1780. URL: <https://doi.org/10.1162/neco.1997.9.8.1735> (date of access: 28.06.2026).
7. Learning phrase representations using RNN encoder–decoder for statistical machine translation / K. Cho et al. *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, Doha, Qatar. Stroudsburg, PA, USA, 2014. URL: <https://doi.org/10.3115/v1/d14-1179> (date of access: 28.06.2026).
8. Attention is all you need / A. Vaswani et al. *Advances in neural information processing systems (NeurIPS)*. 2017. Vol. 30. P. 5998–6008. URL: <https://doi.org/10.48550/arXiv.1706.03762> (date of access: 27.06.2026).

9. Universal differential equations for scientific machine learning / C. Rackauckas et al. *arXiv*. 2020. URL: <https://doi.org/10.48550/arXiv.2001.04385> (date of access: 26.06.2026).
10. Raissi M., Perdikaris P., Karniadakis G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of computational physics*. 2019. Vol. 378. P. 686–707. URL: <https://doi.org/10.1016/j.jcp.2018.10.045> (date of access: 25.06.2026).
11. WHO COVID-19 dashboard: cases and deaths. *World Health Organization*, 2024. URL: <https://data.who.int/dashboards/covid19/> (date of access: 12.06.2026).
12. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study / A. D. Iuliano et al. *The lancet*. 2018. Vol. 391, no. 10127. P. 1285–1300. URL: [https://doi.org/10.1016/S0140-6736\(17\)33293-2](https://doi.org/10.1016/S0140-6736(17)33293-2) (date of access: 24.06.2026).
13. Data-centric epidemic forecasting: a survey / A. Rodríguez et al. *arXiv*. 2022. URL: <https://doi.org/10.48550/arXiv.2207.09370> (date of access: 23.06.2026).
14. Machine learning for data-centric epidemic forecasting / A. Rodríguez et al. *Nature machine intelligence*. 2024. Vol. 6, no. 10. P. 1122–1131. URL: <https://doi.org/10.1038/s42256-024-00895-7> (date of access: 22.06.2026).
15. Hyndman R. J., Athanasopoulos G. Forecasting: principles and practice. 3rd ed. Melbourne : OTexts, 2021. 442 p.
16. The United States COVID-19 forecast hub dataset / E. Y. Cramer et al. *Scientific data*. 2022. Vol. 9. Article 462. URL: <https://doi.org/10.1038/s41597-022-01517-w> (date of access: 21.06.2026).
17. Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States / E. Y. Cramer et al. *Proceedings of the national academy of sciences*. 2022. Vol. 119, no. 15. P. e2113561119. URL: <https://doi.org/10.1073/pnas.2113561119> (date of access: 20.06.2026).

18. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations / K. Sherratt et al. *eLife*. 2023. Vol. 12. P. e81916. URL: <https://doi.org/10.7554/eLife.81916> (date of access: 19.06.2026).
19. Lakshminarayanan B., Pritzel A., Blundell C. Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems (NeurIPS)*. 2017. Vol. 30. P. 6402–6413. URL: <https://doi.org/10.48550/arXiv.1612.01474> (date of access: 18.06.2026).
20. Our World in Data. *COVID-19 dataset* (Creative Commons BY 4.0). URL: <https://catalog.ourworldindata.org/garden/covid/latest/compact/compact.csv> (date of access: 09.06.2026).
21. Centers for Disease Control and Prevention. *FluView / ILINet*. Delphi Epidata API, Carnegie Mellon University. URL: <https://api.delphi.cmu.edu/epidata/fluview/> (date of access: 09.06.2026).
22. Box G. E. P., Jenkins G. M., Reinsel G. C. Time series analysis: forecasting and control. 4th ed. Hoboken : John Wiley & Sons, 2008. 746 p.
23. Neural ordinary differential equations / R. T. Q. Chen et al. *Advances in neural information processing systems (NeurIPS)*. 2018. Vol. 31. P. 6571–6583. URL: <https://doi.org/10.48550/arXiv.1806.07366> (date of access: 17.06.2026).
24. Hethcote H. W. The mathematics of infectious diseases. *SIAM review*. 2000. Vol. 42, no. 4. P. 599–653. URL: <https://doi.org/10.1137/S0036144500371907> (date of access: 16.06.2026).
25. Allen L. J. S. An introduction to stochastic epidemic models. *Mathematical epidemiology*. Berlin : Springer, 2008. P. 81–130.
26. Gillespie D. T. Approximate accelerated stochastic simulation of chemically reacting systems. *The journal of chemical physics*. 2001. Vol. 115, no. 4. P. 1716–1733. URL: <https://doi.org/10.1063/1.1378322> (date of access: 15.06.2026).
27. Goodfellow I., Bengio Y., Courville A. Deep learning. Cambridge : MIT Press, 2016. 800 p.

28. Bengio Y., Simard P., Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*. 1994. Vol. 5, no. 2. P. 157–166. URL: <https://doi.org/10.1109/72.279181> (date of access: 14.06.2026).
29. Kendall A., Gal Y. What uncertainties do we need in Bayesian deep learning for computer vision? *Advances in neural information processing systems (NeurIPS)*. 2017. Vol. 30. P. 5574–5584. URL: <https://doi.org/10.48550/arXiv.1703.04977> (date of access: 13.06.2026).
30. Hyndman R. J., Koehler A. B. Another look at measures of forecast accuracy. *International journal of forecasting*. 2006. Vol. 22, no. 4. P. 679–688. URL: <https://doi.org/10.1016/j.ijforecast.2006.03.001> (date of access: 11.06.2026).
31. Gneiting T., Raftery A. E. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical association*. 2007. Vol. 102, no. 477. P. 359–378. URL: <https://doi.org/10.1198/016214506000001437> (date of access: 10.06.2026).
32. Array programming with NumPy / C. R. Harris et al. *Nature*. 2020. Vol. 585. P. 357–362. URL: <https://doi.org/10.1038/s41586-020-2649-2> (date of access: 08.06.2026).
33. SciPy 1.0: fundamental algorithms for scientific computing in Python / P. Virtanen et al. *Nature methods*. 2020. Vol. 17. P. 261–272. URL: <https://doi.org/10.1038/s41592-019-0686-2> (date of access: 08.06.2026).
34. McKinney W. Data structures for statistical computing in Python. *Proceedings of the 9th Python in science conference (SciPy)*. 2010. P. 56–61. URL: <https://doi.org/10.25080/Majora-92bf1922-00a> (date of access: 10.06.2026).
35. Seabold S., Perktold J. Statsmodels: econometric and statistical modeling with Python. *Proceedings of the 9th Python in science conference (SciPy)*. 2010. P. 92–96. URL: <https://doi.org/10.25080/Majora-92bf1922-011> (date of access: 10.06.2026).

36. PyTorch: an imperative style, high-performance deep learning library / A. Paszke et al. *Advances in neural information processing systems (NeurIPS)*. 2019. Vol. 32. P. 8024–8035. URL: <https://doi.org/10.48550/arXiv.1912.01703> (date of access: 26.06.2026).
37. Kingma D. P., Ba J. Adam: a method for stochastic optimization. *Proceedings of the international conference on learning representations (ICLR)*. 2015. URL: <https://doi.org/10.48550/arXiv.1412.6980> (date of access: 27.06.2026).
38. Neural network technologies for monitoring and analysis of destructive damage to agricultural plots / K. V. Silvanovych et al. *Bionics of intelligence*. 2025. Vol. 2, no. 103. P. 22–33. URL: [https://doi.org/10.30837/bi.2025.2\(103\).04](https://doi.org/10.30837/bi.2025.2(103).04) (date of access: 28.06.2026).
39. Automated software testing of information system modules / S. H. Udovenko et al. *Bionics of intelligence*. 2025. Vol. 1, no. 102. P. 52–63.