

РЕЦЕНЗІЯ
на дисертаційну роботу
Буткевича Миколи Віталійовича
на тему «Моделі, методи та інформаційна технологія прогнозування
часових рядів епідемічних процесів»,
представлену на здобуття ступеня доктора філософії
в галузі знань 12 Інформаційні технології за спеціальністю
122 Комп'ютерні науки

Актуальність теми дисертації.

Прогнозування захворюваності на інфекційні хвороби належить до ключових завдань системи громадського здоров'я, оскільки безпосередньо впливає на ефективність планування ресурсів, закупівель вакцин та антивірусних препаратів, а також своєчасність реагування на епідемічні загрози. Традиційні підходи до епідемічного прогнозування на основі детерміністичних та стохастичних компартментних моделей (SIR, SEIR, SEIRS) вимагають точного знання біологічних параметрів передачі інфекції, які у реальних умовах епідемічного нагляду є невідомими або суттєво змінюваними; статистичні моделі часових рядів типу ARIMA, SARIMA та SARIMAX обмежені припущенням лінійності, стаціонарності або сезонної стаціонарності і втрачають точність за наявності складних нелінійних залежностей та структурних зламів. Методи глибокого навчання, передусім рекурентні нейронні мережі з довгою короткочасною пам'яттю (LSTM), демонструють здатність автоматично виявляти складні нелінійні патерни у послідовних даних без явної специфікації функціональної форми залежностей, що робить їх перспективним інструментом моделювання епідемічних часових рядів із вираженою сезонністю та непередбачуваними структурними зламами.

В умовах сучасних викликів, з якими стикається національна система громадського здоров'я України, – порушення безперервності епідемічного спостереження у прифронтових і деокупованих регіонах, масові міграційні процеси, зниження охоплення планової імунізації – особливої ваги набувають методи прогнозування, що зберігають працездатність на коротких та неповних вибірках і не потребують значних обчислювальних ресурсів. У зв'язку з цим виникає актуальне наукове завдання розроблення інформаційної технології прогнозування захворюваності на інфекційні хвороби на основі інтеграції статистичних моделей часових рядів і рекурентних нейронних мереж з довгою короткочасною пам'яттю в умовах структурних зламів системи епідемічного нагляду, що й обумовлює актуальність теми дисертаційного дослідження.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.

Матеріали, представлені у дисертаційній роботі Буткевича Миколи Віталійовича, свідчать про наукову обґрунтованість і достовірність отриманих результатів, що забезпечується системним застосуванням сучасних методів теорії часових рядів, статистичного моделювання та глибокого навчання. Сформульовані автором наукові положення спираються на коректно проведений аналіз сучасного стану проблематики, чітко визначений математичний апарат моделей ARIMA, SARIMA та LSTM, а також на повноцінну експериментальну валідацію на реальних даних епідемічного нагляду України та країн Балтії з використанням стратегії walk-forward валідації, що забезпечує строгу темпоральну каузальність оцінок. У роботі сформульовано та вирішено актуальне наукове завдання – розроблення інформаційної технології прогнозування захворюваності на інфекційні хвороби на основі інтеграції статистичних моделей часових рядів і рекурентних нейронних мереж з довгою короткочасною пам'яттю в умовах структурних зламів системи епідемічного нагляду.

Розроблені моделі, методи та програмне забезпечення впроваджено у практичну діяльність Державної установи «Харківський обласний центр контролю та профілактики хвороб Міністерства охорони здоров'я України» (акт впровадження від 26 листопада 2025 р.); у навчальний процес кафедри епідеміології Харківського національного медичного університету (акт впровадження від 10 січня 2026 р.); у науково-дослідну роботу кафедри епідеміології Харківського національного медичного університету в рамках міжнародного проєкту STCU № 7136 (акт впровадження від 5 січня 2026 р.); у практичну діяльність Комунального некомерційного підприємства «Міська клінічна лікарня № 13» Харківської міської ради (акти впровадження від 15 листопада 2025 р. та від 17 листопада 2025 р.).

Наукова новизна результатів дисертаційного дослідження полягає в такому:

1. Уперше запропоновано метод кумулятивного представлення епідемічних часових рядів для прогнозування мережами довгої короткочасної пам'яті, заснований на принципі заміни рівневих значень кількості випадків захворювання наростаючим підсумком від початку періоду спостереження з подальшою min-max нормалізацією, параметри якої обчислюються виключно на навчальній вибірці, який відрізняється від існуючих підходів до прогнозування захворюваності використанням монотонно зростаючої траєкторії, що згладжує короткострокові флуктуації та знижує відносну дисперсію, що дозволяє знизити середню абсолютну відсоткову похибку прогнозування у 16–94 рази залежно від захворювання (до 0,28% для гепатиту С, 0,82–0,85% для гепатиту В).

2. Уперше запропоновано метод циклічного кодування сезонних ознак для LSTM моделей епідемічного прогнозування, заснований на тригонометричному відображенні дискретного часового індексу на одиничне коло у двовимірному просторі за допомогою синусоїдальної та косинусоїдальної функцій від номера місяця, який відрізняється від існуючих підходів до моделювання сезонності в

епідемічних нейромережевих моделях забезпеченням безперервності та рівновіддаленості сусідніх місяців, включаючи перехід грудень–січень, що дозволяє покращити точність прогнозування на 46,6% порівняно з моделями без явних сезонних ознак.

3. Удосконалено метод валідації прогностичних моделей епідемічних часових рядів для коротких вибірок від 72 місяців в умовах структурних зламів, заснований на стратегії walk-forward з послідовним розширенням навчального вікна та принципі строгої темпоральної каузальності, який відрізняється від існуючих обов'язковим обчисленням параметрів усіх трансформацій, включаючи нормалізацію та формування ознак, виключно на навчальній частині вибірки на кожному кроці, що дозволяє отримати реалістичну оцінку продуктивності моделі в режимі послідовного прогнозування та запобігає витоку інформації з тестового періоду.

4. Набула подальшого розвитку архітектура LSTM мереж для епідемічного прогнозування, заснована на принципі ієрархічного представлення часових залежностей із адаптацією складності моделі до обсягу та характеру доступних даних, яка відрізняється від існуючих систематичним емпіричним обґрунтуванням достатності компактних двошарових конфігурацій з 64 одиницями пам'яті на шар для коротких рядів від 72 місяців та необхідності глибоких тришарових пірамідальних архітектур з 11,89 мільйонами параметрів для рядів з множинними сезонними компонентами та структурними зламами, що дозволяє забезпечити навчання моделей за 40 секунд на стандартному CPU без GPU-прискорювачів та обирати архітектуру з мінімальними обчислювальними витратами без втрати точності прогнозування.

5. Набула подальшого розвитку архітектура інформаційної технології прогнозування інфекційної захворюваності, заснована на розділенні тренувального та інференсного шарів з версіонованим сховищем артефактів моделей, яка відрізняється від існуючих інтеграцією статистичних моделей

ARIMA та неймережевих моделей LSTM у єдиний конвеєр з автоматизованим контролем якості через маніфести версій та метрики валідації, що дозволяє забезпечити відтворюваність прогнозів та оперативне перенавчання при надходженні нових епідемічних даних.

Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності, мова та стиль викладення результатів.

За своїм змістом дисертаційна робота Буткевича Миколи Віталійовича повністю відповідає Стандарту вищої освіти за спеціальністю 122 Комп'ютерні науки. Дисертаційна робота є завершеною науковою працею, а її результати свідчать про вагомий особистий внесок здобувача у науковий напрям комп'ютерного моделювання та прогнозування епідемічних процесів засобами штучного інтелекту. Дисертація складається з анотації, змісту, переліку умовних скорочень, вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Повний обсяг роботи становить 211 сторінок друкованого тексту, з них основний текст – 164 стор., список зі 143 використаних джерел – 18 стор., додатки – 12 стор. Дисертація містить 28 рисунків та 20 таблиць. Оформлення дисертаційної роботи відповідає вимогам наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

За результатами ознайомлення зі звітом подібності дисертаційну роботу Буткевича Миколи Віталійовича слід кваліфікувати як результат самостійних досліджень здобувача, в якій відсутні ознаки фальсифікації, компіляції, фабрикації та плагіату. Цитування праць інших авторів виконано коректно і супроводжується відповідними посиланнями; виявлений відсоток текстових співпадінь зумовлений посиланнями на власні попередні публікації здобувача. Дисертаційну роботу написано з дотриманням принципів академічної доброчесності.

Дисертаційна робота написана сучасною українською науковою мовою з належною термінологічною точністю у сферах інформаційних технологій,

математичного моделювання та аналізу часових рядів. Виклад відзначається логічною стрункістю та обґрунтованістю авторських тверджень, які послідовно ілюструються числовими прикладами, таблицями та графіками. Оформлення дисертаційної роботи відповідає вимогам наказу МОН України від 12 січня 2017 р. № 40.

Зміст дисертації.

У вступі обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету, об'єкт, предмет і завдання дослідження, представлено методи досліджень, наукову новизну отриманих результатів, повноту їх опублікування й апробацію, а також наведено зв'язок дисертаційної роботи з науковими програмами та темами.

Перший розділ «Аналіз проблеми прогнозування епідемічних процесів» присвячено систематичному огляду сучасних підходів до моделювання та прогнозування епідемічних процесів. Систематизовано методологічні підходи: детерміністичні та стохастичні компартментні моделі (SIR, SEIR, SEIRS), моделі часових рядів (ARIMA, SARIMA, SARIMAX), методи машинного навчання (багатошарові нейронні мережі, мережі довгої короткочасної пам'яті LSTM) та гібридні підходи. Окремо проаналізовано стан досліджень прогнозування гострого та хронічного гепатиту В із застосуванням рекурентних нейронних мереж, моделювання динаміки грипу на основі ARIMA та прогнозування захворюваності на лептоспіроз. За результатами проведеного аналізу виявлено прогалини у дослідженнях, зосереджених на середовищах з порушеними даними та обмеженими обчислювальними ресурсами, характерних для країн у стані конфлікту, та сформульовано задачі дисертаційного дослідження.

Другий розділ «Розроблення методів та теоретичних основ прогнозування епідемічних процесів» присвячено формалізації математичного апарату моделей ARIMA, SARIMA та LSTM. Розкрито теорію функціонування рекурентних

мереж з системою вентилів (вентиль забування, вентиль входу, вентиль виходу) та механізм збереження довгострокових залежностей через адитивне оновлення стану комірки. Запропоновано процедуру ідентифікації порядків ARIMA-моделей на основі аналізу автокореляційних та частково автокореляційних функцій, тестів стаціонарності Дікі–Фуллера та Квятковського–Філіпса–Шмідта–Шіна, відбору оптимальної специфікації за інформаційними критеріями Акаїке (AIC) та Байєса (BIC), а також метод діагностики залишків з використанням критерію Льюнга–Бокса. Сформульовано функції втрат, метрики оцінювання (MAE, MSE, RMSE, MAPE) та валідаційні стратегії для епідемічних часових рядів. Описано архітектурні рішення першої ітерації LSTM-моделей із методами попередньої обробки, навчання, оптимізації та регуляризації.

Третій розділ «Покращення методів та прикладне застосування методів прогнозування епідемічних процесів» присвячено вдосконаленням третьої ітерації LSTM-архітектури. Запропоновано метод кумулятивного представлення часових рядів та метод циклічного кодування сезонних ознак через синусоїдальне та косинусоїдальне відображення номера місяця, що забезпечує безперервність переходу грудень–січень. Описано розширені методи попередньої обробки даних та представлено архітектуру вебзастосунку для онлайн-трекінгу епідемічних прогнозів: серверний модуль збереження артефактів моделей, інференсний шар з версіонованими маніфестами, методологію оцінювання якості і валідації, а також графічний інтерфейс системи моніторингу. Розглянутий програмно-апаратний комплекс становить ядро запропонованої інформаційної технології прогнозування інфекційної захворюваності.

Четвертий розділ «Результати експериментальних досліджень» містить триітераційну експериментальну валідацію розроблених методів на реальних даних епідемічного нагляду України та країн Балтії за 2013–2024 роки. У

першій ітерації досліджено базові LSTM-моделі для гострого гепатиту А в Україні, грипу в Харківській області та сальмонельозу (тришарова пірамідальна архітектура $3 \times \text{LSTM}(256/128/64)$ з MAPE на рівні 6,98% у фізичних одиницях для сальмонельозу та 5,40% для гострого гепатиту А з компактною двошаровою архітектурою $2 \times \text{LSTM}(64)$). Друга ітерація присвячена побудові ARIMA-моделей для лептоспірозу в країнах Балтії (Естонія, Литва, Латвія) та грипу в Харківській області. Третя ітерація реалізує методологічні вдосконалення: для гепатиту В (гострий і хронічний, 2018–2023) MAPE кумулятивного прогнозу становить 0,82–0,85% (покращення у 16–21 раз); для гепатиту С – MAPE = 0,28% (покращення у 94 рази); для ГРВІ глибока архітектура $3 \times \text{LSTM}(300/1200/600)$ з циклічними сезонними ознаками суттєво знизила похибку при включенні даних воєнного періоду. Окремо подано контрфактуальний аналіз впливу повномасштабного вторгнення на динаміку COVID-19 у Польщі. Розділ завершується порівняльним аналізом LSTM- та ARIMA-підходів за критеріями точності, обчислювальної ефективності й практичної придатності.

У висновках узагальнено отримані в дисертаційній роботі результати, їх наукову новизну та практичне значення, окреслено напрями подальших досліджень.

Список використаних джерел містить достатню кількість актуальних публікацій провідних вітчизняних і міжнародних дослідників, які безпосередньо пов'язані з темою дисертаційного дослідження. Додатки містять список публікацій здобувача за темою дисертації, відомості про апробацію результатів дослідження та акти впровадження.

Повнота викладення результатів досліджень в опублікованих працях.

Дослідження, результати яких викладено в дисертації, виконано на кафедрі математичного моделювання та штучного інтелекту Національного аерокосмічного університету «Харківський авіаційний інститут» в рамках

виконання науково-дослідних робіт за проектами «Мультидисциплінарне дослідження впливу надзвичайних ситуацій на поширення інфекційних захворювань для підтримки прийняття управлінських рішень у сфері біобезпеки населення» (Національний фонд досліджень України, № 2023.03/0197) та «Моделювання та прогнозування поширення інфекцій у воєнних та повоєнних умовах з використанням даних епідеміологічного, поведінкового та геномного спостереження» (фінансування Національною академією наук США (US NAS) через Науково-технологічний центр в Україні (STCU), № 7136).

Основні положення, ідеї та висновки дисертаційної роботи висвітлено у 23 наукових публікаціях здобувача, у тому числі: 5 статей у періодичних наукових виданнях, проіндексованих у наукометричній базі Scopus (дві статті у закордонних виданнях – одна Q1, одна Q2; три статті у наукових фахових виданнях України категорії А – Q3); 8 публікацій у матеріалах міжнародних науково-практичних конференцій, індексованих у Scopus; 2 публікації у матеріалах міжнародних науково-практичних конференцій; 8 свідоцтв про реєстрацію авторського права на комп'ютерні програми. Результати дисертації доповідалися та обговорювалися на IV та V Міжнародних науково-практичних конференціях «ProfIT Conference» (м. Харків, Україна, 2021, 2023); 4th та 5th International Workshop of IT-professionals on Artificial Intelligence ProfIT AI (м. Кембрідж, США, 2024; м. Ліверпуль, Велика Британія, 2025); 7th International Conference on Informatics & Data-Driven Medicine IDDM 2024 (м. Бірмінгем, Велика Британія, 2024); 2024 14th International Conference on Dependable Systems, Services and Technologies DESSERT 2024 (м. Афіни, Греція, 2024); 2024 IEEE 17th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering TCSET 2024 (м. Львів – Славське, Україна, 2024); International Conference «Integrated Computer Technologies in Mechanical Engineering – 2024» (м. Харків, Україна, 2024); 2025 15th International Conference on Advanced Computer Information Technologies

АСІТ 2025 (м. Сібернік, Хорватія, 2025); International Conferences on Applied Computing and WWW/Internet 2025 (м. Порту, Португалія, 2025). Опубліковані матеріали повністю відображають зміст дисертації та відповідають вимогам пункту 8 «Порядку присудження ступеня доктора філософії...», затвердженого Постановою КМУ від 12.01.2022 р. № 44.

Наукові публікації здобувача містять опис проведених досліджень, аналіз сутності проблеми, методів і результатів, а також обґрунтовані висновки. Поршень принципів академічної доброчесності у публікаціях не виявлено; висновки є оригінальними. Таким чином, наукові результати дисертаційної роботи повністю висвітлено у опублікованих працях здобувача.

Недоліки та зауваження до дисертаційної роботи.

1. Опис процедури walk-forward валідації та формування навчальних послідовностей методом ковзного вікна у підрозділах 2.4 та 3.2 наведено переважно у словесній формі; виокремлений псевдокод процедури з покроковим розширенням навчального вікна не наведено, що ускладнює точне відтворення обчислювальних експериментів сторонніми дослідниками.

2. У підрозділі 3.5 розроблено архітектуру вебзастосунку для онлайн-трекінгу прогнозів з версіонованим сховищем артефактів моделей, проте у роботі не наведено кількісних характеристик продуктивності системи (час відповіді на запит, обсяг збереженого артефакту моделі, типове споживання оперативної пам'яті), які підтверджували б практичну придатність розгортання у режимі реального часу.

3. У підрозділі 4.3 наведено результати прогнозування на горизонті у межах річного циклу; водночас оцінок поведінки запропонованих моделей на довших горизонтах (24, 36 місяців) – з характеристикою темпу накопичення похибки прогнозу з ростом горизонту – у дисертаційній роботі не представлено, що ускладнює формулювання рекомендацій щодо застосування моделей для середньострокового планування ресурсів системи охорони здоров'я.

4. У підрозділі 3.3 додавання циклічних ознак \sin/\cos описано на рівні методологічного рішення; у тексті не наведено прямого числового порівняння точності прогнозу з циклічним кодуванням проти простіших альтернатив (one-hot кодування номера місяця, лагова ознака за 12 місяців), що ускладнює оцінку того, наскільки приріст точності зумовлено саме циклічним характером кодування.

5. У постановці задачі підкреслено, що епідемічні часові ряди містять прогалини та запізнення у поданні даних; однак у дисертаційній роботі не наведено окремого числового експерименту з оцінювання чутливості якості прогнозу до штучних пропусків та запізнень останніх місяців тестової вибірки, який підтвердив би стійкість запропонованих моделей до характерних для епідеміологічного нагляду викривлень первинних даних.

Вважаю, що висловлені зауваження не є визначальними і не зменшують загальну наукову новизну та практичну значимість результатів дисертаційної роботи та не впливають на її позитивну оцінку.

Висновок про дисертаційну роботу.

Враховуючи теоретичну ґрунтовність, повноту експериментального обґрунтування та практичну значущість запропонованих моделей і методів, вважаю, що дисертаційна робота здобувача ступеня доктора філософії Буткевича Миколи Віталійовича на тему «Моделі, методи та інформаційна технологія прогнозування часових рядів епідемічних процесів» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних і практичних результатів якого розв'язує наукове завдання, що має істотне значення для галузі знань 12 Інформаційні технології.

Дисертаційна робота Буткевича Миколи Віталійовича на тему «Моделі, методи та інформаційна технологія прогнозування часових рядів епідемічних процесів» за своїм змістом, структурою, обсягом та оформленням відповідає

вимогам Наказу МОН України № 40 від 12.01.2017 «Про затвердження вимог щодо оформлення дисертації (зі змінами)» та «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12.01.2022 р. № 44.

Таким чином, здобувач Буткевич Микола Віталійович заслуговує на присудження наукового ступеня доктора філософії у галузі знань 12 Інформаційні технології за спеціальністю 122 Комп'ютерні науки.

Рецензент:

доктор технічних наук, професор,
професор кафедри кібербезпеки та
інтелектуальних інформаційних технологій
Національного аерокосмічного університету
«Харківський авіаційний інститут»

Ольга МОРОЗОВА