

# **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.**

Сучасний стан епідеміологічної ситуації в Україні визначається сукупністю чинників - наслідками пандемії COVID-19, бойовими діями, міграційними процесами та обмеженням доступу до медичної допомоги у низці регіонів, - що в комплексі формує запит на оперативні інструменти прогнозування інфекційної захворюваності, орієнтовані на роботу з реальними, часто неповними та зашумленими даними щомісячного епідемічного нагляду. У світовій практиці методи глибокого навчання, зокрема рекурентні нейронні мережі з довгою короткочасною пам'яттю (LSTM), вже зарекомендували себе як ефективний інструмент моделювання нелінійних часових рядів, проте їх застосування для епідемічних даних з коротким горизонтом спостереження (від 72 місяців) у поєднанні з класичними статистичними моделями ARIMA залишається недостатньо опрацьованим.

Таким чином, актуальним є наукове завдання створення інформаційної технології, яка б поєднувала статистичні та нейромережеві моделі прогнозування у єдиному відтворюваному конвеєрі з версіонуванням артефактів, дозволяла б оперативне перенавчання при надходженні нових даних і функціонувала на стандартному апаратному забезпеченні без GPU-

прискорювачів. Саме на розв'язання цього завдання спрямовано рецензовану дисертаційну роботу, що й обумовлює її актуальність.

### **Обґрунтованість, достовірність та новизна наукових результатів.**

Наукові результати, отримані Буткевича Миколи Віталійовича, є обґрунтованими та достовірними. Достовірність забезпечується використанням актуальних реальних даних епідемічного нагляду України та країн Балтії, коректним застосуванням сучасних інструментів машинного навчання (scikit-learn, TensorFlow/Keras, statsmodels), а також послідовним проведенням експерименту у три ітерації, що дозволяє чітко атрибутувати отриманий приріст точності саме запропонованим методичним рішенням. Здобувач коректно сформулював актуальне прикладне наукове завдання - побудови інформаційної технології прогнозування інфекційної захворюваності на основі поєднання моделей ARIMA та LSTM - і повноцінно його вирішив у межах дослідження.

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 нейромережових 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 нейромережевих 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.

**Завершеність дисертаційної роботи, дотримання академічної доброчесності, мова та стиль викладення.**

За змістом і структурою дисертаційна робота Буткевича Миколи Віталійовича відповідає Стандарту вищої освіти за спеціальністю 122 Комп'ютерні науки. Це завершена наукова праця з вираженою прикладною спрямованістю; отримані результати свідчать про самостійний і вагомий внесок здобувача у розвиток методів прогнозування часових рядів епідемічних процесів. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Повний обсяг роботи - 211 сторінок друкованого тексту, зі 143 використаних джерел; дисертація містить 28

рисунків та 20 таблиць. Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40.

Аналіз звіту перевірки на текстові співпадіння свідчить про те, що подана робота є власним самостійним дослідженням Буткевича Миколи Віталійовича і не містить плагиату, фальсифікації, компіляції чи фабрикації. Запозичені фрагменти супроводжуються коректними посиланнями на першоджерела; наявні текстові співпадіння є відсилками до раніше опублікованих праць самого здобувача. Принципи академічної доброчесності у роботі дотримано в повному обсязі.

Робота написана українською мовою з використанням загальноприйнятої термінології сфери машинного навчання, прикладної статистики та епідеміологічного моделювання. Тексту властиві ясність формулювань, лаконічність та логічна послідовність викладу; теоретичні положення підкріплено графіками, таблицями метрик і фрагментами програмного коду. Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40.

### **Структура та зміст дисертації.**

Дисертаційну роботу побудовано за класичною структурою. У вступі обґрунтовано актуальність теми, сформульовано мету, об'єкт, предмет і завдання дослідження, охарактеризовано наукову новизну та практичне значення отриманих результатів.

У першому розділі здобувач системно опрацьовує доступні підходи до прогнозування епідемічних процесів - від детерміністичних компартментних моделей до сучасних нейромережевих архітектур - і обґрунтовано формулює прикладну прогалину: брак методик, придатних для коротких рядів і обмежених обчислювальних ресурсів. Це задає чітку прикладну рамку для наступних розділів.

Другий розділ містить компактний, але достатній виклад математичного апарату ARIMA та LSTM у формі, придатній для безпосередньої програмної реалізації: описано процедуру вибору порядків моделей за тестами стаціонарності, інформаційними критеріями та аналізом залишків, а також механізм роботи вентилів LSTM. Особливу цінність має покрокове формулювання функцій втрат, метрик оцінювання та стратегії walk-forward валідації, орієнтованих саме на епідемічні дані.

Третій розділ - методично найважливіший з прикладної точки зору. Тут запропоновано метод кумулятивного представлення епідемічних рядів і метод циклічного (sin/cos) кодування сезонних ознак, а також спроектовано вебзастосунок з версіонованим сховищем артефактів моделей. Така архітектура з розділенням тренувального та інференсного шарів є типовим інженерним рішенням сучасних ML-систем і робить роботу не лише науковою, а й готовою до практичного впровадження.

У четвертому розділі представлено триітераційну експериментальну валідацію на реальних даних України та країн Балтії за 2013–2024 роки. Послідовно продемонстровано, що запропоновані методичні вдосконалення дають значне зниження MAPE для гепатитів В та С (до 0,28%), а також стійке покращення прогнозів ГРВІ при включенні даних воєнного періоду; окремо показано можливості контрфактуального аналізу на прикладі COVID-19 у Польщі.

Висновки лаконічно узагальнюють отримані результати; список джерел відображає сучасний стан проблематики; додатки містять список публікацій здобувача та акти впровадження.

### **Повнота відображення результатів у наукових публікаціях здобувача.**

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 ACIT 2025 (м.

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

Оpubліковані наукові праці здобувача повно відображають як методологічну, так і експериментальну складові дисертації. Зокрема, статті у виданнях Q1/Q2 Scopus та доповіді на профільних міжнародних конференціях ProfIT AI, IDDM, DESSERT, ACIT, TCSET і ICTM формують достатню апробаційну базу для висновку про повну відображеність результатів дисертації в наукових публікаціях здобувача.

### **Зауваження до дисертаційної роботи.**

1. У дисертаційній роботі здійснено детальне порівняння LSTM-моделей зі статистичними моделями ARIMA та SARIMA; водночас бракує бодай орієнтовного порівняння з простими, але часто ефективними прикладними baseline-моделями для часових рядів — методом сезонного наївного прогнозу (seasonal naïve) та методом експоненційного згладжування (Holt–Winters), — що могло б додатково посилити позиціювання запропонованого підходу.

2. Експериментальна валідація третьої ітерації проведена на даних 2018–2023 років (гепатити В/С) та різних періодів для інших нозологій; у роботі не обговорюється, чи впливає рік початку навчальної вибірки (наприклад, 2013 проти 2016 або 2018 року) на якість прогнозу: відсутність такого аналізу ускладнює формулювання рекомендації щодо мінімальної довжини історії спостереження, достатньої для досягнення задовільної точності прогнозу з практичної точки зору.

3. У підрозділі 4.3 емпірично показано стабільну роботу запропонованих моделей на рядах довжиною від 72 місяців, проте систематичного дослідження чутливості якості прогнозу до обсягу навчальної вибірки (наприклад, при її



скороченні до 60, 48 чи 36 місяців) у роботі не наведено; це ускладнює формулювання практичних рекомендацій щодо мінімально достатньої довжини ряду для надійного застосування методів.

4. Для оцінювання якості прогнозів LSTM-моделей у дисертаційній роботі використано точкові метрики (MAE, MAPE, RMSE); попри те, що для ARIMA-моделей в підрозділі 4.2 побудовано довірчі інтервали прогнозу, для LSTM-моделей аналогічних інтервальних характеристик не наведено; це дещо обмежує прикладне використання результатів LSTM-прогнозування у системі планування закупівель вакцин та антивірусних препаратів, де потрібна кількісна оцінка ризику недо- або переоцінки прогнозованого попиту.

5. Підрозділ 4.3.4 містить контрфактуальний аналіз впливу повномасштабного вторгнення на динаміку COVID-19 у Польщі, що є цікавим прикладним результатом; водночас аналогічного контрфактуального експерименту для української вибірки гепатитів В/С або ГРВІ — з оцінкою того, як виглядав би прогноз без структурного злому 2022 року — у роботі не наведено, що дещо обмежує обсяг інтерпретації виявленого впливу воєнного фактору на епідемічну динаміку в Україні.

Наведені зауваження мають характер побажань щодо подальшого розширення дослідження і не впливають на загальну позитивну оцінку дисертаційної роботи.

#### **Загальний висновок щодо дисертаційної роботи.**

З урахуванням викладеного дисертаційна робота Буткевича Миколи Віталійовича на тему «Моделі, методи та інформаційна технологія прогнозування часових рядів епідемічних процесів» є завершеним самостійним науковим дослідженням, у якому отримано сукупність нових теоретичних та прикладних результатів, що в комплексі розв'язує актуальне наукове завдання, яке 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:**

кандидат технічних наук,  
доцент кафедри  
інформаційно-комунікаційних  
технологій  
ім. О. О. Зеленського  
Національного аерокосмічного  
університету  
«Харківський авіаційний інститут»

**Олексій РУБЕЛЬ**