

## ВІДГУК

офіційного опонента – доктора фізико-математичних наук,  
старшого наукового співробітника, провідного наукового співробітника  
відділу математичних методів дослідження операцій № 130 інституту  
кібернетики ім. В.М. Глушкова НАН України Норкіна Володимира Івановича  
на дисертаційне дослідження Токаревої Катерини Анатоліївни на тему  
«Гібридні та ансамблеві методи та моделі машинного навчання  
прогнозування фінансових часових»,  
поданої на здобуття наукового ступеня доктора філософії  
за спеціальністю 113 – «Прикладна математика»  
галузі знань – 11 «Математика та статистика»

**Актуальність теми виконаної роботи.** Дослідження фінансових ринків на основі моделей часових рядів являється одним із основних підходів як в теоретичному аспекті, так і з точки зору практичного використання. Використання часових рядів є зрозумілим, оскільки дозволяє достатньо точно оцінювати прогноз реального фінансового активу, базуючись на припущеннях щодо основних характеристик часового ряду:

- сумісному розподілу між елементами часового ряду в різні моменти часу, тобто сумісного розподілу  $x_{t_1}, \dots, x_{t_n}$ ;
- властивостям залишків часового ряду  $\varepsilon_t$ , які описують наявність випадкових чинників в математичній моделі.

Основна проблема у використанні теоретичних моделей часових рядів полягає у невиконанні припущень на сумісні розподіли елементів часового ряду та властивості залишків (гомоскедастичність, незалежність тощо). В цьому випадку для оцінки прогнозів вибирається найкраща модель із наявних на основі деякої метрики (інформаційних критеріїв, RMSE, MAE тощо). Іншим підходом до розв'язання проблеми із невиконанням основних припущень щодо елементів часового ряду є використання гібридних моделей, а саме моделей машинного навчання та глибинного навчання. Основна ідея гібридних моделей полягає в тому, що деякі припущення на елементи часового ряду замінюються

більш слабшими обмеженнями, які відображаються у моделях машинного навчання.

У дисертаційній роботі реалізовано один із підходів до конструювання гібридних систем, згідно якого зроблено припущення про те, що динаміка залишків описується за допомогою штучної нейронної мережі. Це в свою чергу дало змогу пом'якшити вимоги до класичних моделей (AR, ARMA, ARIMA) та збільшити прогностичну точність класичних моделей.

Іншим підходом дисертаційної роботи являється узагальнення класичної LSTM нейронної мережі із врахуванням довільної скінченної передісторії. Даний підхід дозволить відслідковувати сезонні компоненти, які не можуть бути відслідковані на основі класичної LSTM мережі.

Таким чином, основна увага дисертаційного дослідження зосереджена на актуальному питанні розробки гібридних алгоритмів, в яких математичною моделлю досліджуваного часового ряду є класичні моделі часових рядів, а залишки описуються на основі моделей машинного навчання. Отже, дисертаційне дослідження Токаревої Катерини Анатоліївни на тему «Гібридні та ансамблеві методи та моделі машинного навчання прогнозування фінансових часових» є своєчасним та актуальним.

**Структура та зміст дисертаційної роботи.** Дисертаційна робота має наступну структуру: вступ, чотири розділи, висновки до кожного розділу, загальні висновки, список літератури, що містить 158 джерел, додаток зі списком публікацій, що стосуються теми дисертації.

У першому розділі представлений всебічний огляд поточних наукових досліджень в області аналізу та прогнозування фінансових часових рядів, розглянуто історію розвитку часових рядів та їх використання для аналізу фінансових ринків. Основну увагу даного розділу зосереджено на переваги та недоліки моделей часових рядів та їх застосуванні в реальному житті. Також, даний розділ концентрує увагу на деяких моделях машинного навчання та їх

використанні у фінансовій математиці. В кінці даного розділу визначено основні моделі машинного навчання, які надалі будуть використані для побудови гібридних моделей.

У другому розділі дисертації розглядаються проблеми побудови гібридних моделей та основні підходи до конструювання останніх. Наведено ряд класичних гібридних моделей, побудованих на основі часових рядів та моделей машинного навчання, наведено загальний підхід до побудови нейронних мереж для аналізу часових рядів на основі моделей машинного навчання. Проаналізовано ряд фінансових показників та виявлено, що використання гібридних моделей на основі моделей машинного навчання забезпечує кращу точність, а також дає ряд інших переваг.

Третій розділ присвячений виявленню та аналізу хаотичних складових у часових рядах, у тому числі визначенні та аналізу хаотичної складової. Для оцінки хаотичності часового ряду використано чотири основні підходи (на основі перетворення Фур'є, на основі показника Ляпунова, на основі показника Герста та на основі фрактальної розмірності часового ряду). На реальних прикладах показано, що визначення хаотичної складової дозволяє значно покращити точність прогностичної моделі.

У четвертому розділі дисертаційного дослідження розглянуто два розширення класичних моделей. А саме розглянуто розширення лінійного часового ряду до стохастичного диференціального рівняння за припущення про нерівномірні вимірювання у часовому ряді. Відносно розширеної моделі розглянуто оптимізаційну задачу для оцінки параметрів моделі. Іншим узагальненням даного розділу являється LSTM нейронна мережа. В узагальненій LSTM мережі зроблено припущення про наявність довгострокових детермінованих залежностей, які не можуть бути відслідковані на основі класичної моделі.

**Обґрунтованість і достовірність** наукових положень, висновків і рекомендацій забезпечена публікаціями відповідних результатів у рейтингових рецензованих міжнародних виданнях. П'ять статей (дві статті у виданнях, що індексується у наукометричній базі даних SCOPUS, три статті у наукових фахових виданнях України) за темою дисертації повністю охоплюють проблематику та основні положення наукової роботи. Апробація основних результатів дисертації відбулася під час представлення доповідей на 4 Всеукраїнських та Міжнародних науково-практичних конференціях. Розгорнутий та повний огляд наукової літератури з тематики представленого дослідження, аналіз та застосування сучасних підходів до розв'язання задач дослідження свідчать про те, що дисертантка повністю володіє необхідною методологією наукового дослідження.

**Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.** Дисертаційне дослідження містить ряд нових та значущих результатів, отриманих дисертанткою, а саме:

1. Показано, що часові ряди фондових індексів містять як лінійну, так і не лінійну складову, а отже, окремо лінійна модель ARIMA та нелінійна модель ANN не можуть дати точну оцінку таких часових рядів.

2. Розроблено гібридну модель моделі авторегресії-ковзної середньої та штучної нейронної мережі прогнозування часових рядів фондових індексів. Побудована модель показала більшу точність прогнозів, отриманих за допомогою алгоритму гібридної інтегрованої моделі авторегресії-ковзної середньої та штучної нейронної мережі.

3. Використовуючи підхід оцінки наявності хаосу в динамічній системі, розроблено ряд гібридних моделей на основі класичної моделі нейронної мережі прямого поширення та рекурентній нейронної мережі Елмана, які дозволили більш точно оцінювати прогноз для моделі Маккея-Глесса, логістичного рівняння та моделі Ено.

4. Показано, що використання гібридних алгоритмів для підбору стартової точки оцінки параметрів моделі, може значно зменшити як обчислювальну складність моделі, так і підвищити точність прогнозу.

5. Отримано явне представлення розв'язку одного стохастичного диференційного рівняння з запізненням аргументу (теорема 4.1).

6. Класичні моделі LSTM та багатовимірні моделі MV-LSTM узагальнено для явного врахування передісторії часового ряду за рахунок додаткового параметру глибини передісторії.

Результати, що отримані у дисертаційному дослідженні, роблять значний внесок у розвиток теорії та практики теорії часових рядів та побудові прогностичних систем на основі гібридних алгоритмів. Запропоновані алгоритми та методи можуть бути використані для вирішення задач оптимізації в різних галузях науки та техніки, в тому числі і в задачі оптимізації портфелю цінних паперів.

**Зв'язок роботи з науковими програмами, планами, темами.** Обраний напрям дослідження відповідає науковому напрямку кафедри прикладної математики та інформаційних технологій Чернівецького національного університету імені Юрія Федьковича. Дослідження дисертаційної роботи було проведено в рамках науково-дослідної теми кафедри «Математичне моделювання і числово-аналітичні методи дослідження динамічних та інформаційних процесів» (номер державної реєстрації 0102U006591)

**Зауваження та побажання.** Загалом позитивно оцінюючи отримані результати дослідження, їх наукову новизну та практичну значущість, водночас, вважаю за доцільне звернути увагу на деякі недоліки:

1. Загальне зауваження. В роботі використовуються терміни «методи штучного інтелекту, глибинного інтелекту, машинного навчання». Треба було б дати якісь визначення для них, або посилання на авторитетів. Чим «методи штучного інтелекту» відрізняються від «методів глибинного інтелекту»?

2. Зауваження до оформлення. В роботі використовується дуже багато аббревіатур, але списку аббревіатур з поясненням їх значення нема. Це суттєво ускладнює читання та розуміння тексту. Недаремно державний стандарт вимагає створення окремого списку скорочень.

3. Зауваження до Розділу 1. Багато разів зустрічається комбінація слів «алгоритми штучного інтелекту», але не дано визначення (у розумінні автора),

що є «штучний інтелект». Чому не менш пафосне «алгоритми машинного навчання»?

4. Зауваження до викладання матеріалу. В роботі багато числових результатів у вигляді таблиць та графіків, але майже відсутні безпосередні коментарі до них.

5. В Розділі 3 активно використовується мережа Елмана (рекурентна мережа зі зворотним зв'язком). Чому немає посилання на роботу Елмана? Elman, J. L. (1990). Finding structure in time. Cognitive science, 14(2), 179-219.

6. Посилання [154 - 155] зайві.

7. Застереження до Розділу 4, п.4.2.3. В цьому розділі присутні розривні функції перетворення вихідного сигналу, це може привести до розривних функцій похибки, для яких градієнтні методи мінімізації можуть бути незастосовні.

**Відсутність порушень академічної доброчесності.** Кваліфікаційне дослідження є самостійною науковою працею авторки. Висновки, рекомендації та пропозиції, що характеризують наукову новизну кваліфікаційного дослідження, одержані авторкою особисто. При використанні праць інших вчених для аргументації актуальних положень дослідження обов'язково вказано посилання на відповідні праці.

**Висновок про дисертаційну роботу.** Вважаю, що дисертаційна робота на здобуття наукового ступеня доктора філософії Токаревої Катерини Анатоліївни на тему «Гібридні та ансамблеві методи та моделі машинного навчання прогнозування фінансових часових» виконана на високому науковому рівні, не порушує принципів доброчесності та є завершеним науковим дослідженням, що має важливе значення для розвитку теорії часових рядів та гібридних моделей.

Дисертаційне дослідження за актуальністю, науково-технічним рівнем, новизною постановки та розв'язання проблем, практичним значенням отриманих результатів відповідає вимогам пунктів 6, 7, 8, 9 «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 січня 2022 р. №44 (із змінами, внесеними згідно з Постановою Кабінету Міністрів України №341 від 21.03.2022 р. та №502 від 19.05.2023), а її авторка Токарева К.А. заслуговує на присудження їй ступеня доктора філософії у галузі знань 11 – Математика та статистика за спеціальністю 113 – Прикладна математика.

Офіційний опонент:  
доктор фізико-математичних наук,  
старший науковий співробітник,  
провідний науковий співробітник  
відділу математичних методів  
дослідження операцій  
Інституту кібернетики ім. В.М. Глушкова  
НАН України



Володимир НОРКІН

Підпис Норкіна В.І. засвідчую,  
Вчений секретар Інституту кібернетики  
ім. В.М.Глушкова НАН України, д.ф.-м.н.

Сергій ЄРШОВ