

ВІДГУК

офіційного опонента – доктора технічних наук, професора,
завідувача кафедри Інформаційних систем Інституту комп'ютерних систем

Національного університету «Одеська політехніка»

Арсірій Олени Олександрівни

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. Аналіз фінансових стратегій являється одним із секторів із найбільшою прибутковістю. Це можна прослідкувати за багатьма сучасними тенденціями, включаючи швидке зростання/падіння вартостей акцій, цінних металів та криптовалют. Проте поряд із високою прибутковістю наявний великий ризик швидкого знецінення фінансового портфелю, який складеться як із основних фінансових активів, так і похідних – дериватив.

Аналіз роботи багатьох хедж-фондів вказує на необхідність чіткості у математичних моделях, які використовуються для аналізу фінансових часових рядів та їх прогнозування. Крім того було помічено, що реальні часові ряди (із врахуванням різного роду корекцій) не підпорядковуються класичним моделям часових рядів. Тому перед дослідниками та теоретиками постає питання щодо пом'якшення основних припущень на класичні моделі поряд із покращенням прогностичної точності нових моделей. Таким чином, основна увага дослідників часових рядів зосереджена на розробці нових моделей, які по-перше покращують точність моделі, по-друге, пом'якшують припущення щодо часового ряду.

Одним із найбільш оптимістичних підходів в даному напрямку є створення гібридних моделей на основі синтезу класичних моделей та новітніх моделей штучного інтелекту, а саме моделей машинного навчання та глибинного навчання. Використання таких моделей розширило область використання теорії часових рядів і дало змогу усунути деякі недоліки класичних часових рядів, таких як гомоскедастичність даних, однорідність часового ряду тощо.

Отже, основна увага в дисертаційному дослідженні звернена на побудову гібридних алгоритмів як симбіозу алгоритмів часових рядів та методів машинного навчання. Даний підхід дозволив визначити ряд більш точних алгоритмів, що було перевірено в дисертаційному дослідженні на ряді реальних фінансових часових рядів. Виходячи із вищесказаного, вважаю, що дисертаційне дослідження Токаревої Катерини Анатоліївни на тему «Гібридні та ансамблеві методи та моделі машинного навчання прогнозування фінансових часових» актуальним та має прикладний характер.

Структура та зміст дисертаційної роботи. Дисертаційна робота Токаревої Катерини Анатоліївни на тему «Гібридні та ансамблеві методи та моделі машинного навчання прогнозування фінансових часових» складається зі вступу, чотирьох розділів, висновків до кожного розділу, загальних висновків, списку використаної літератури (158 джерел), додатку зі списком публікацій авторки дисертаційного дослідження.

Основним завданням першого розділу являється розкриття проблематики прогнозування фінансових часових рядів та проблем які виникають при підборі моделі. Тут описано основні підходи для аналізу реальних процесів на основі класичних моделей часових рядів, моделей машинного навчання та гібридних моделей, як симбіозу моделей часових рядів та моделей машинного навчання. Слід зауважити що для аналізу вибрано простіші моделі нейронних мереж для кращої візуалізації побудови гібридних алгоритмів.

Більше уваги теорії гібридних алгоритмів присвячено в розділі 2, в якому проаналізовано основні принципи побудови гібридних алгоритмів. В даному розділі запропоновано гібридну модель, яка включає використання нейронної мережі для аналізу помилок основного часового ряду. Таким чином, основний фінансовий актив описується на основі часового ряду, в той же час помилка часового ряду описується на основі нейронної мережі.

Ще однією характеристикою часових рядів, яка може вносити особливості при аналізі, є наявність хаотичних складових (детермінований хаос), які попри свою дискретну природу можуть виявлятися як псевдо-стохастичні складові. Для аналізу наявності хаотичних складових перетворення Фур'є, показник Ляпунова, показник Герста та фрактальну розмірність випадкового процесу. Використовуючи реальні приклади динамічних систем було встановлено, що виділення хаотичної складової значно підвищує точність моделі.

Четвертий розділ присвячено аналізу узагальненої LSTM мережі та розширенню часового ряду до неперервного випадку (стохастичне диференціальне рівняння Іто). Було встановлено, що розроблені моделі володіють більш точним прогнозом для реальних даних.

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 Токаревої Катерини Анатоліївни містить наступні нові або узагальнює відомі результати, а саме у роботі:

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.
3. Доведено, якщо часовий ряд є строго стаціонарним, то ARIMA модель є більш адекватною для гауссівських рядів.
4. Доведено практичність використання теорії хаотичних динамічних систем із використанням різних показників хаотичності для аналізу як реальних часових рядів, так і класичних теоретичних моделей. Використовуючи підхід оцінки наявності хаосу в динамічній системі, розроблено ряд гібридних моделей на основі класичної моделі нейронної мережі прямого поширення та нейронної мережі Елмана, які дозволили більш точно оцінювати прогноз для моделі Маккея-Глесса, логістичного рівняння та моделі Ено.
5. Проілюстровано, що використання гібридних алгоритмів для підбору стартової точки оцінки параметрів моделі, може значно зменшити як обчислювальну складність моделі, так і підвищити точність прогнозу.
6. Розроблено загальний підхід побудови гібридних моделей на основі узагальненої моделі LSTM із врахуванням періодичних складових

7. Отримані теоретичні результати протестовано на реальних даних із використанням динаміки трьох реальних фінансових процесів.

8. Розроблено підхід узагальнення LSTM мережі на випадок використання довільного часового ряду $AR(p)$ замість $AR(1)$ в класичній моделі.

Результати дисертаційного дослідження значно розширюють теорію часових рядів та деякі розділи машинного навчання.

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. До зауважень та побажань слід віднести наступне:

1. Назва роботи дуже загальна. Варто було б звужити назву до більш конкретних задач.

2. Вступ та опис літератури являється дуже детальним. Варто акцентувати більше уваги на результатах дисертаційного дослідження та використовувати більшість джерел в якості порівняння.

3. Автор при перевірці точності прогнозування часових рядів на запропонованих гібридних моделях в другому розділі використовує такі метрики як MAE MSE, MAPE, RMSE, а для моделей із налаштуванням параметрів за допомогою ГА в третьому розділі – метрики RMSE та NRMSE. При цьому автор не обґрунтовує вибір метрик та детально не аналізує їх значення. Думаю, що

доцільно було б також використати коефіцієнт детермінації R^2 значення якого є показовим при прогнозуванні часових рядів.

4. В роботі розглянута декілька основних задач та методів їх вирішення, проте не наведено загального алгоритму використання цих всіх методів разом, тобто не наведено прикладів ансамблевого підходу до прогнозування.

5. Робота містить цілий ряд описок та неточностей.

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.

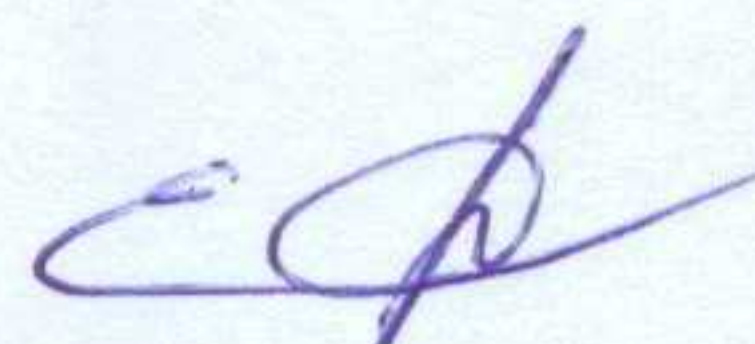
Офіційний опонент:

доктор технічних наук, професор,

завідувач кафедри Інформаційних систем

Інституту комп'ютерних систем

Національного університету «Одеська політехніка»

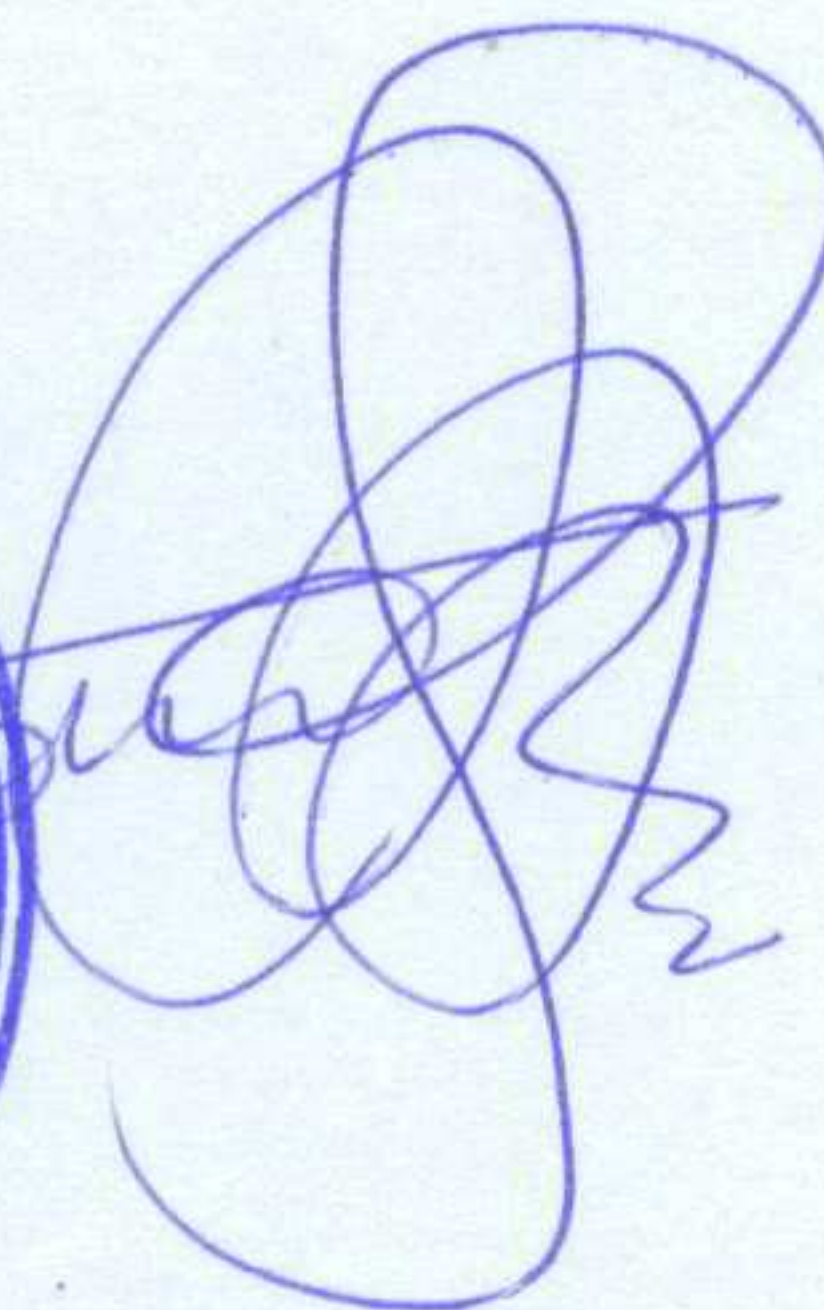


Олена АРСІРІЙ

Вчений секретар Вченої ради

Національного університету

«Одеська політехніка»



Лада ПРОКОПОВИЧ