

РЕЦЕНЗІЯ

доктора філософії, доцента

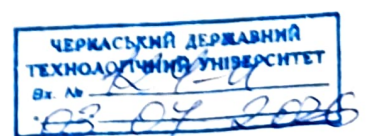
Олексюка Вадима Володимировича

на дисертаційну роботу ОСТАПЮКА Володимира Вікторовича
«Агентні моделі в інформаційній технології інтелектуального моніторингу»,
подану на здобуття наукового ступеня доктора філософії
з галузі знань 12 – «Інформаційні технології» за спеціальністю
121 – «Інженерія програмного забезпечення».

Актуальність теми дослідження

Сучасний етап розвитку розподілених інформаційних систем, мікросервісних архітектур та систем критичної інфраструктури характеризується стрімким зростанням їхньої складності та динамічності середовищ функціонування. У процесі експлуатації таких систем неминуче відбувається зміна статистичних властивостей даних (явище «дрейфу концепцій») та режимів роботи. За таких умов традиційні статичні засоби моніторингу втрачають свою адекватність, що може призводити до пропуску критичних збоїв. Виникає гостра суперечність між динамічною природою об'єктів спостереження та статичністю класичних інструментів моніторингу.

Проблемою, що потребує вирішення, є розробка методів і засобів для створення моніторингових програмних агентів, здатних до автономного синтезу та адаптації своїх прогностичних моделей без втручання людини. Це підтверджується актуальними викликами у галузі MLOps (Machine Learning Operations), зокрема проблемою автоматизації багатоетапних процесів конструювання складних ансамблів та забезпечення машинної відтворюваності експериментів при перенесенні моделей у середовище експлуатації.



Таким чином, дисертаційна робота В.В. Остапюка, спрямована на вдосконалення методів багат шарового синтезу моделей та розроблення інструментів їх відтворюваного виконання, зокрема на основі декларативного підходу та подання у вигляді спрямованого ациклічного графа DAG, є надзвичайно своєчасною та актуальною. Вирішення цієї науково-прикладної проблеми сприятиме підвищенню надійності та ефективності систем інтелектуального моніторингу.

Наукова новизна одержаних результатів

У дисертаційній роботі В.В. Остапюка розв'язано актуальне науково-прикладне завдання підвищення ефективності функціонування моніторингового програмного агента шляхом розробки методів побудови та засобів програмної реалізації багат шарових агентних моделей.

Зокрема, автором отримано такі основні наукові результати:

1. **Вперше запропоновано** метод багат шарового синтезу агентних моделей, який, на відміну від існуючих, використовує попередню кластеризацію вхідних даних і побудову спеціалізованих моделей для кожного кластеру з подальшим додаванням їх прогнозів до масиву вхідних даних для повторного синтезу. Це дозволяє суттєво підвищити однорідність точок спостереження у кластерах та, як наслідок, точність моделювання.

2. **Удосконалено** метод синтезу багат шарових моделей та їх програмної реалізації шляхом залучення гетерогенних алгоритмів машинного навчання для формування структурних шарів. Це дозволяє компенсувати алгоритмічні обмеження окремих класів машинного навчання, підвищити повноту опису об'єкта та ефективніше виявляти приховані закономірності.

3. **Набув подальшого розвитку** метод багат шарового синтезу агентних моделей за рахунок формалізації процесів конструювання алгоритмів на основі спрямованого ациклічного графу (DAG). Такий

декларативний підхід забезпечує багатоетапність та машинний детермінізм обчислень при проектуванні програмного забезпечення агента.

Практична цінність та значимість результатів роботи

Практичне значення отриманих результатів полягає у створенні цілісної інформаційної технології проектування програмного забезпечення моніторингових агентів.

Розроблений метод підвищення однорідності даних та вдосконалений метод гетерогенної рециркуляції дозволяють моніторинговому агенту автономно розпізнавати зміни режимів функціонування інформаційної системи та ефективно адаптувати до них свою прогностичну логіку.

Важливим практичним здобутком є застосування декларативного підходу та архітектури на основі DAG, що вирішує критичну прикладну інженерну проблему MLOps – надійне перенесення навчених складних ансамблів у середовище експлуатації із гарантуванням повної машинної відтворюваності обчислень.

Результати роботи мають високий потенціал для впровадження в системах технічної діагностики, моніторингу мікросервісних архітектур та автоматизованого контролю, про що свідчать акти впровадження (Додаток А). Окрім того, результати дослідження використані у виконанні науково-дослідних робіт кафедри ПЗАС ЧДТУ («Програмні агенти в інформаційній технології кризового інтелектуального моніторингу» та «Інтелектуальний моніторинг кібербезпеки корпоративної мережі програмними агентами»).

Достовірність отриманих результатів

Наукові положення, висновки та рекомендації дисертаційної роботи В.В. Осталюка є обґрунтованими. Достовірність підтверджується коректним застосуванням методів системного аналізу, машинного навчання,

кластеризації (K-Means, GMM, ієрархічна), методів оптимізації (алгоритм TPE) та граф-орієнтованого моделювання (DAG).

Теоретичні висновки надійно підкріплені масштабними експериментальними дослідженнями, проведеними на трьох структурно відмінних відкритих наборах табличних даних. Використання загальноприйнятих метрик оцінки якості регресійних моделей (RMSE, MAE, R^2) та методології перехресної перевірки доводить адекватність запропонованих моделей та ефективність алгоритмів. Збіг результатів повторних запусків складної архітектури підтверджує машинний детермінізм запропонованого підходу.

Структура та зміст дисертаційного дослідження

Представлена дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел (100 найменувань) і 4 додатків. Загальний обсяг становить 149 сторінок, з яких 119 – основний текст, що містить 8 таблиць та 17 рисунків.

У **Вступі** обґрунтовано актуальність теми дослідження, яка визначається потребою в автоматизації процесів моніторингу складних розподілених систем в умовах дрейфу концепцій. Сформульовано мету роботи, об'єкт, предмет та методи дослідження. Чітко визначено завдання, які необхідно вирішити для досягнення поставленої мети. Наведено положення, що визначають наукову новизну та практичну цінність отриманих результатів, а також відомості щодо апробації результатів дослідження та структури роботи.

Розділ 1 «Аналіз стану проблеми інтелектуального моніторингу та інженерії програмного забезпечення агентних систем» присвячений системному аналізу сучасного стану та тенденцій розвитку інформаційних технологій моніторингу динамічних середовищ. Автор детально досліджує феномен зміни статистичних властивостей вхідних потоків даних, так званий

«дрейф концепцій» та обґрунтовує, чому класичні статичні інструменти спостереження втрачають ефективність. Значну увагу приділено аналізу сучасних інженерних концепцій у межах MLOps (Machine Learning Operations). Дисертант критично аналізує існуючі інструменти та платформи управління життєвим циклом моделей (такі як MLflow, Kubeflow тощо), виявляючи їхні обмеження у контексті забезпечення повної машинної відтворюваності складних ансамблів та автоматизації багатоетапних обчислень при перенесенні моделей у промислове середовище експлуатації. На основі проведеного аналізу автор обґрунтовує доцільність застосування парадигми автономних програмних агентів, здатних самостійно адаптувати власну прогностичну логіку, та формулює чіткі завдання дисертаційного дослідження.

Розділ 2 «Методи багат шарового синтезу та адаптації агентних моделей» містить теоретико-методологічний базис роботи. У ньому автор пропонує та формалізує комплекс підходів до автоматизованого конструювання моделей програмного агента.

По-перше, детально описано метод підвищення однорідності вхідних даних. Його суть полягає у попередній сегментації простору ознак за допомогою алгоритмів кластеризації, зокрема K-Means, GMM та ієрархічної кластеризації, з наступним навчанням локальних моделей під кожен виділений кластер та інтеграцією класифікатора для маршрутизації нових точок спостереження.

По-друге, викладено вдосконалений метод багат шарового синтезу моделей. Особливістю методу є використання гетерогенних пулів алгоритмів машинного навчання на кожному структурному шарі, де прогнози моделей попереднього кроку виступають як додаткові ознаки для наступного етапу моделювання. Автор подає концептуальні схеми та детальні діаграми

процесів, що описують логіку взаємодії елементів у процесі автономного донавчання агента.

Розділ 3 «Експериментальні дослідження методів синтезу та адаптації моделей» присвячений практичній перевірці та оцінці ефективності запропонованих теоретичних рішень. Експерименти виконано на трьох структурно відмінних відкритих наборах табличних даних, що моделюють реальні процеси, включаючи прогнозування цільових метрик, пов'язаних із транспортними та інфраструктурними параметрами, зокрема Trip Price. Для оптимізації гіперпараметрів моделей на кожному шарі рециркуляції використано сучасний алгоритм Tree-structured Parzen Estimators (TPE). Оцінювання якості моделей здійснювалося за допомогою загальноприйнятих метрик регресійного аналізу: середньоквадратичної похибки (RMSE), середньої абсолютної похибки (MAE) та коефіцієнта детермінації (R^2). Автор наводить порівняльні таблиці та графіки, які наочно демонструють, що комплексне застосування методів кластеризації та гетерогенної рециркуляції дозволяє суттєво знизити похибку прогнозування порівняно як з базовими ізолюваними алгоритмами, так і з класичними лінійними ансамблями.

Розділ 4 «Проектування та реалізація інформаційної технології інтелектуального моніторингу» відображає інженерно-програмну складову дослідження у межах спеціальності 121. У цьому розділі охарактеризовано архітектурні шаблони побудови програмного забезпечення інтелектуального агента. Автор пропонує компонентну об'єктно-орієнтовану архітектуру системи та декларативний опис конфігурації, реалізований у форматі структурованої JSON-схеми. Головним інженерним здобутком розділу є розробка логіки подання багатоетапних процесів конструювання алгоритмів синтезу моделей у вигляді спрямованого ациклічного графа (DAG). Описано механізм функціонування цього обчислювального графа та процес

формування відтворюваного пакета результатів (bundle), що містить повну інформацію для аудиту, відтворення та експлуатації навченої модельної структури. Таке рішення гарантує детермінізм обчислень та машинну відтворюваність результатів під час перенесення складних ансамблів у середовище експлуатації.

У **Загальних висновках** підсумовано головні наукові та практичні результати дисертаційної роботи, констатовано виконання всіх поставлених завдань і підтверджено досягнення мети дослідження.

Додатки містять допоміжні матеріали, зокрема акти про впровадження результатів дисертаційного дослідження у виробничу діяльність підприємств та в освітній процес, що підтверджує практичну цінність роботи, а також лістинги коду та фрагменти розроблених декларативних конфігураційних файлів (DSL JSON).

Повнота викладення наукових положень, висновків і рекомендацій, сформульованих у дисертації та опублікованих працях

Основні наукові результати, положення та висновки дисертації повною мірою висвітлені у 14 наукових працях автора. Серед них 4 статті у фахових наукових виданнях України, 1 стаття у виданні, що індексується базою Scopus, та 9 тез доповідей на міжнародних і всеукраїнських науково-практичних конференціях. Опубліковані праці адекватно і повно відображають зміст дисертації.

Відсутність порушення академічної доброчесності

Дисертаційна робота виконана самостійно. У тексті дисертації відсутні ознаки академічного плагіату, фабрикації чи фальсифікації результатів. Використані ідеї, твердження та результати інших авторів супроводжуються

належними посиланнями на першоджерела. Внесок дисертанта у працях, опублікованих у співавторстві, чітко розмежований і коректно вказаний у тексті дисертації. Робота відповідає принципам академічної доброчесності.

Зауваження та недоліки дисертації щодо її оформлення і змісту

Незважаючи на високий науковий рівень та практичну значущість роботи, можна відзначити окремі зауваження:

1. У дисертації слушно наголошується на проблемі непередбачуваної зміни властивостей даних з часом, що є критичним для систем моніторингу. Запропонована архітектура дозволяє автоматизувати процес перенавчання моделей. Проте в роботі недостатньо розкрито механізм запуску цього процесу: дослідження виграло б від опису конкретних умов або критеріїв, які б автоматично сигналізували програмному агенту про необхідність оновлення моделей на нових даних.

2. Автор наводить ґрунтовну критику існуючих систем автоматизованого машинного навчання (AutoML), вказуючи на складність управління їхньою архітектурою. В експериментах розроблений метод демонструє вищу точність порівняно з базовими моделями та класичними ансамблями. Проте для повнішого розкриття переваг запропонованого підходу доцільно було б навести порівняння із сучасними платформами автоматизованого машинного навчання не лише за точністю, а й за показниками гнучкості налаштувань.

3. Удосконалений метод рециркуляції базується на використанні різнотипних алгоритмів на кожному структурному шарі. Хоча автор наводить результати ефективності таких рішень, принципи спільного використання конкретних класів моделей висвітлено недостатньо. Робота набула б більшої завершеності за умови формування чітких рекомендацій: які саме комбінації

методів (наприклад, дерева рішень і лінійні моделі) є найбільш сумісними при передачі даних між шарами.

4. У першому розділі наведено розлогий огляд сучасних методів прогнозування часових рядів. Проте в експериментальній частині задачу зведено до табличної регресії за окремими зрізами стану системи, що певною мірою нівелює часову залежність даних. Враховуючи, що моніторингова інформація за своєю природою є послідовністю подій у часі, доцільно було б пояснити механізми подальшої адаптації запропонованого підходу для безпосередньої роботи з масивами історичних даних за попередні періоди.

Зазначені недоліки не мають принципового характеру, не применшують наукової новизни та практичної цінності отриманих результатів і не впливають на загальну позитивну оцінку дисертаційної роботи.

Загальні висновки за дисертаційним дослідженням

Дисертаційна робота Остапюка Володимира Вікторовича на тему «Агентні моделі в інформаційній технології інтелектуального моніторингу» є завершеною, самостійно виконаною науково-дослідною працею, у якій розв'язано актуальне науково-прикладне завдання підвищення ефективності моніторингових програмних агентів шляхом удосконалення методів синтезу та реалізації багат шарових адаптивних моделей.

Вважаю, що представлена до захисту дисертаційна робота за своїм змістом, ступенем актуальності, рівнем обґрунтованості наукових положень, новизною, повнотою викладу в наукових публікаціях та практичною значимістю цілком відповідає вимогам «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 (зі змінами), а її автор ОСТАПЮК Володимир Вікторович заслуговує на присудження йому наукового ступеня доктора філософії з галузі знань 12 «Інформаційні технології» за спеціальністю 121 «Інженерія програмного забезпечення».

Рецензент:

доцент кафедри програмного
забезпечення автоматизованих
систем, Черкаського державного
технологічного університету, доктор
філософії, доцент



Вадим ОЛЕКСЮК

Підпис доктора філософії, доцента
Вадима ОЛЕКСЮКА засвідчую:

Учений секретар Черкаського
державного
технологічного університету
к.т.н., доцент



Ірина МИРОНЕЦЬ