

РЕЦЕНЗІЯ

на дисертаційну роботу

Яременка Вадима Сергійовича

на тему «Модель мультиагентної системи для автоматизованої побудови

словника предметної області при обробці поточкових даних»,

представлену на здобуття ступеня доктора філософії

в галузі знань 12 «Інформаційні технології»

за спеціальністю 122 «Комп'ютерні науки»

Актуальність теми дисертації.

Актуальність теми дисертації зумовлена необхідністю створення ефективних інтелектуальних систем для обробки великих потоків текстових даних, що неперервно надходять із численних джерел. Одним із ключових викликів є зменшення обсягів інформації, яка передається на подальшу обробку, при збереженні її релевантності й аналітичної цінності. Важливу роль у цьому відіграють автоматизовані доменні словники, які є основою для тематичної фільтрації та класифікації.

У сучасних умовах особливого значення набуває використання моделей мультиагентних систем, які дозволяють реалізовувати розподілену, масштабовану та адаптивну обробку текстової інформації. Такі моделі забезпечують паралельну взаємодію незалежних агентів, кожен з яких виконує локальний аналіз даних, бере участь у колективному прийнятті рішень, а також у формуванні та динамічному оновленні словникової бази. Це дозволяє досягти високої стійкості системи до збоїв, знизити обчислювальні витрати та підвищити ефективність у розподіленому середовищі.

Таким чином, тема дисертації є актуальною як у теоретичному аспекті — через розвиток моделей мультиагентної взаємодії в задачах текстової аналітики, так і з прикладної точки зору — у контексті побудови практично орієнтованих систем для обробки інформаційних потоків із мінімальними ресурсними витратами.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.

Наукова новизна результатів дисертаційного дослідження полягає в наступному:

Вперше запропоновано модель мультиагентної системи, яка поєднує модифікований фільтр Блума, нейронну мережу для класифікації текстів, мультиагентний підхід для побудови та оновлення словників і механізм голосування методом Шульце з використанням методу TF-IDF, що дозволяє автоматизувати процес створення словників предметної області в умовах потокової обробки текстових даних.

Вперше запропоновано модифікацію класичного фільтра Блума, який відрізняється тим, що він забезпечує швидке виявлення релевантних текстів і виконання їх попередньої класифікації, що забезпечує значне зменшення обсягу необроблених даних на наступних етапах системи та підвищує ефективність роботи в умовах обробки поточкових даних.

Вперше запропоновано модифікацію методу TF-IDF в розподіленому середовищі для вирішення задачі побудови словника предметної області, яка відрізняється застосуванням адаптованого методу Шульце для використання у мультиагентних системах при голосуванні між агентами щодо оновлення доменних словників, що забезпечує ухвалення рішень на основі колективного аналізу текстових даних.

Вперше запропоновано метод оптимізації параметрів налаштування фільтра Блума з використанням генетичного алгоритму для застосування у задачі багатокласової фільтрації поточкових текстових даних для підвищення точності їх попередньої класифікації.

Розроблено модель нейронної мережі, адаптовану для класифікації текстових даних за кількома предметними областями для заданого набору даних.

Розроблено механізм динамічної адаптації системи зворотного зв'язку, який передає оновлений словник назад у фільтр Блума. Це дозволяє динамічно адаптувати систему до нових даних і підвищити точність фільтрації текстів. Такий підхід забезпечує постійну актуальність системи в умовах змінного середовища даних.

Достовірність отриманих результатів забезпечується використанням достатньо повних і характерних для задачі наборів даних, а також коректною побудовою моделей і чітким дотриманням експериментальної методології. Узагальнення та висновки підтверджуються внутрішньою логічною відповідністю між теоретичними положеннями та практичними результатами дослідження.

Наукові дослідження були виконані здобувачем на кафедрі системного проектування КПІ ім. Ігоря Сікорського в рамках ініціативної теми під керівництвом доктора технічних наук, професора Рогози Валерія Станіславовича.

Отже, в дисертаційній роботі поставлене наукове завдання розробки моделі мультиагентної системи для автоматизованої побудови словника предметної області при обробці потокових даних виконано повністю, здобувач повною мірою оволодів методологією наукової діяльності.

Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності.

За своїм змістом дисертаційна робота здобувача Яременка Вадима Сергійовича повністю відповідає Стандарту вищої освіти зі спеціальності 122 «Комп'ютерні науки» та напрямкам досліджень відповідно до освітньої програми «Комп'ютерні науки».

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача у науковий напрям інформаційних та комунікаційних технологій.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадіння, можна зробити висновок, що дисертаційна робота Яременка Вадима Сергійовича є результатом самостійних досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело.

Мова та стиль викладення результатів.

Дисертаційна робота написана українською мовою. Текст викладено послідовно та логічно, з дотриманням наукового стилю. Використано загальноприйнятну термінологію. Матеріал подано доступно та аргументовано.

Дисертація складається зі вступу, 4 розділів, висновків, списку літератури та додатків. Загальний обсяг дисертації 177 сторінок.

У вступі обґрунтовано необхідність автоматизації процесу побудови словників предметної області для ефективної обробки текстових потоків. Акцентовано увагу на доцільності використання мультиагентного підходу, який забезпечує масштабовану, стійку до збоїв і розподілену обробку даних. Словники розглядаються як ключовий компонент для фільтрації та класифікації текстів, а також подано загальний опис системи, що реалізує послідовність таких етапів, як фільтрація, класифікація, виявлення термінів і їхнє узгодження за допомогою агентного голосування.

Перший розділ присвячено аналізу сучасних досліджень у сферах мультиагентних систем, обробки потокових текстових даних і машинного навчання. На підставі аналізу окреслено наукову проблему та сформульовано задачу створення автоматизованого механізму побудови предметних словників із використанням мультиагентної парадигми.

У другому розділі розроблено концепцію побудови мультиагентної системи для задач обробки текстів та словникового моделювання. Надано опис архітектури взаємодії агентів, обґрунтовано вибір технологічної основи реалізації та запропоновано механізм колективного прийняття рішень як складову системної координації.

У третьому розділі розглянуто методи класифікації текстових даних, зокрема вдосконалений варіант фільтра Блума, адаптований до багатокласової класифікації, параметри якого оптимізуються за допомогою генетичного алгоритму. Оцінено можливості нейронних мереж різних типів та методів формування словникових структур, що дозволило реалізувати ефективну агентну модель для обробки текстових потоків.

Четвертий розділ присвячено практичній реалізації системи, яка об'єднує модифікований фільтр Блума, LSTM-мережу та інструменти динамічного оновлення словників. Наведено результати експериментів, що демонструють високу точність класифікації, фільтрації й ефективну роботу системи в умовах безперервного надходження текстових даних.

Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

Оприлюднення результатів дисертаційної роботи.

Наукові результати дисертації висвітлені у 8 наукових публікаціях здобувача, серед яких: 7 статей у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України; 1 стаття у періодичних наукових виданнях, проіндексованих у базі даних Scopus.

У співавторських публікаціях чітко простежується особистий внесок Яременка В. С., що полягає у створенні архітектурних рішень мультиагентних систем, розробці модифікованих підходів до багатокласової фільтрації з використанням фільтра Блума, розробленні технічних засад побудови предметних словників, а також у проведенні порівняльного аналізу алгоритмів класифікації текстових даних. Усі наукові праці безпосередньо пов'язані з тематикою дисертації та відображають її ключові положення, що підтверджує як практичну значущість дослідження, так і його наукову новизну.

У публікації "Application of neural network algorithms and naïve bayes for text classification" висвітлено результати порівняння ефективності класифікаторів на основі нейронних мереж і методу наївного Байєса. Яременком В. С. було визначено параметри моделей, обрано архітектури нейронних мереж, підібрано тестові дані та проведено аналіз результатів експериментів.

У статті "Огляд наявних мультиагентних систем для задач інтелектуального аналізу даних" систематизовано існуючі підходи до побудови мультиагентних архітектур. Особистий внесок здобувача полягає в проведенні

огляду джерел, виокремленні ключових компонентів систем і формуванні узагальненої моделі.

У роботі "Підхід до використання фільтра Блума для багатокласової класифікації текстових даних в режимі реального часу" представлено вдосконалення класичного фільтра Блума для підтримки багатокласової класифікації в потоковому середовищі. Яременком В. С. було запропоновано модифікацію структури фільтра, узгоджену з вимогами до багатокласової обробки, та обґрунтовано доцільність такого підходу.

У публікації "Модель мультиагентної системи для семантичного аналізу текстів" подано архітектурну модель системи для розподіленої обробки текстів. Внесок здобувача включає постановку задачі, побудову структурної схеми системи та реалізацію її базових компонентів.

У статті "Development of a multi-agent system for solving domain dictionary construction problem" описано підхід до автоматизованої побудови словників предметної області на основі агентної взаємодії. Здобувачем запропоновано загальну структуру системи, визначено методологічну основу побудови словника, запропоновані підходи до голосування між агентами в розподіленому середовищі та сформульовано технічне завдання для практичного експерименту.

Публікація "A comparative analysis of text data classification accuracy and speed using neural networks, Bloom filter and Naive Bayes" присвячена порівнянню популярних підходів до класифікації текстів за критеріями точності та швидкодії. Яременком В. С. було ініційовано експериментальне дослідження, визначено конфігурації моделей та проведено обробку отриманих результатів.

У статті "Порівняльний аналіз програмних бібліотек для класифікації текстових даних із використанням штучних нейронних мереж" подано огляд сучасних бібліотек, що використовуються для побудови систем класифікації. Особистий внесок здобувача полягає в оцінці функціональних можливостей бібліотек, формулюванні критеріїв порівняння та обґрунтуванні вибору технологічного інструментарію для експериментів.

У роботі "Neural networks and Monte-Carlo method usage in multi-agent systems for sudoku problem solving" розглянуто поєднання нейронних мереж і методу Монте-Карло в рамках мультиагентного підходу. Здобувачем було розроблено модель взаємодії агентів і реалізовано принципи розв'язання задач підвищеної обчислювальної складності.

Наукові публікації здобувача повною мірою відображають основні положення, результати та висновки дисертаційного дослідження, виконані на високому фаховому рівні та не містять ознак порушення академічної доброчесності.

Таким чином, наукові результати описані в дисертаційній роботі повністю висвітлені у наукових публікаціях здобувача.

Недоліки та зауваження до дисертаційної роботи.

1. Для додаткового підкріплення отриманих результатів було б доцільно провести окремий аналіз чутливості результатів до змін параметрів мультиагентної системи, зокрема — кількості агентів, розміру словника та порогів фільтрації.
2. У роботі недостатньо розкрито критерії зупинки процесу оновлення словника або умови, за яких нові терміни вважаються достатньо релевантними для включення, що важливо для підтримки якості класифікації впродовж тривалого часу.
3. Інформація, наведена на рис. 1.4 (стор. 30), ілюструє динаміку зростання обсягів даних лише до 2015 року, що з урахуванням актуальності дослідження на 2025 рік виглядає дещо застарілою; доцільним було б доповнити її більш сучасними статистичними даними або зазначити це як умовне узагальнення.
4. Опис математичних виразів, наведених на сторінці 86, доцільно винести з основного тексту абзацу та подати у вигляді окремих нумерованих формул, що сприятиме кращому сприйняттю матеріалу та забезпечить структурованість викладення.
5. Для наглядності доцільно було б представити опис алгоритму роботи модифікованого фільтру Блума на (ст. 108) та кроки генетичного алгоритму (ст. 110) у вигляді блок-схеми, слідуючи стандарту ДСТУ ISO5807:2016.
6. Була помічна незначна кількість граматичних помилок.

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

Висновок про дисертаційну роботу.

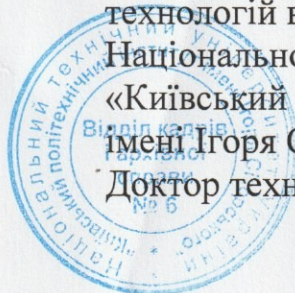
Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Яременка Вадима Сергійовича на тему «Модель мультиагентної системи для автоматизованої побудови словника предметної області при обробці потокових даних» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує наукове завдання, що має істотне значення для галузі інформаційних технологій. Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в п.6 – 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.

Здобувач Яременко Вадим Сергійович заслуговує на присудження ступеня доктора філософії в галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Рецензент:

Завідувач кафедри цифрових
технологій в енергетиці
Національного технічного університету України
«Київський політехнічний інститут
імені Ігоря Сікорського»
Доктор технічних наук, професор



9 червня 2025 р.

| | |
|---------------|----------------|
| Підпис гр. | Натадія АУШЕВА |
| ЗАСВІДЧУЮ | |
| Відділ кадрів | |
| підпис | пр-ще |