

РЕЦЕНЗІЯ

на дисертаційну роботу

Яременка Вадима Сергійовича

на тему «Модель мультиагентної системи для автоматизованої побудови

словника предметної області при обробці потокових даних»,

представлену на здобуття ступеня доктора філософії

в галузі знань 12 «Інформаційні технології»

за спеціальністю 122 «Комп'ютерні науки»

Актуальність теми дисертації.

Актуальність теми дослідження полягає в необхідності підвищення ефективності обробки потокових текстових даних у розподілених системах шляхом зменшення обсягів інформації, що підлягає зберіганню, передаванню та аналізу. Для цього важливим є використання оптимально налаштованих структур даних, зокрема модифікованого фільтра Блума, параметри якого обираються з урахуванням балансу між ймовірністю хибних спрацьовувань і використанням пам'яті.

Запропонована мультиагентна архітектура дає змогу реалізувати локальну попередню обробку даних із мінімальними комунікаційними витратами, що підвищує масштабованість і дозволяє ефективно використовувати ресурси обчислювальних вузлів. Особливу роль у системі відіграють автоматизовано побудовані доменні словники, які слугують основою для класифікації текстів і дозволяють адаптувати систему до нових тематик без повного перенавчання моделей.

Таким чином, дослідження є актуальним з погляду побудови обчислювально ефективної архітектури, що поєднує механізми локальної оптимізації обробки інформації, адаптивність словникового представлення та стійкість у розподіленому середовищі.

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,

мультіагентний 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 мультіагентних 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 фільтру 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 підтверджується застосуванням репрезентативних даних, формальною коректністю побудованих моделей і послідовністю експериментальної перевірки. Висновки роботи ґрунтуються на логічному узгодженні теоретичних положень з отриманими практичними результатами.

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 мультіагентної системи для автоматизованої побудови словника предметної області при обробці поточкових 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 з вступу, 4 розділів, висновків, списку літератури та додатків. Загальний обсяг дисертації 177 сторінок.

У вступі викладено аргументацію щодо доцільності автоматизованої побудови словників предметної області для задач обробки поточкових текстів. Наголошено на перевагах мультиагентного підходу, зокрема його здатності забезпечувати масштабованість, розподілену організацію обчислень та стійкість до відмов. Окреслено роль словників як базового механізму для здійснення фільтрації та класифікації текстової інформації, а також подано загальну схему функціонування системи, яка включає етапи фільтрації, класифікації, виділення термінів і колективного формування словника за допомогою голосування агентів.

Перший розділ містить огляд сучасного стану досліджень у галузі мультиагентних систем, машинного навчання та методів обробки великих обсягів текстових потоків. На основі цього аналізу обґрунтовано вибір напрямку дослідження та сформульовано задачу побудови автоматизованого механізму створення предметних словників із використанням мультиагентної моделі.

У другому розділі розроблено підхід до проектування мультиагентної системи, орієнтованої на обробку потоків текстової інформації та формування словникової бази. Детально описано архітектуру взаємодії агентів,

обґрунтовано вибір технологічного інструментарію та запропоновано модель агентної комунікації з використанням процедури колективного прийняття рішень.

Третій розділ містить аналіз методи класифікації текстів. Наведено модифікацію фільтра Блума, адаптовану для багатокласової класифікації з оптимізацією параметрів на основі генетичного алгоритму. Розглянуто архітектури нейронних мереж та підходи до побудови словників, на основі чого реалізовано ефективну модель мультиагентної системи для автоматизованої обробки текстових потоків.

У четвертому розділі подано реалізацію розробленої мультиагентної системи, яка поєднує вдосконалений фільтр Блума, нейронну мережу типу LSTM та алгоритмічні засоби для оновлення словників. Показано, що така комбінація забезпечує високу точність класифікації та стабільну продуктивність системи в умовах обробки поточкових 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.

Усі публікації мають безпосередній зв'язок із тематикою дисертації та підтверджують наукову і практичну цінність проведеного дослідження.

Публікації здобувача відзначаються високим науковим рівнем і містять глибокий аналіз основних результатів проведеного дослідження. Вагомий особистий вклад здобувача в співавторські роботи. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело.

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. У дисертації відсутні публікації, що висвітлюють результати дослідження за останні три роки, що дещо ускладнює оцінку поточної динаміки розвитку тематики у порівнянні з результатами, отриманими в дисертації.

2. Опис застосованого генетичного алгоритму подано на концептуальному рівні, без математичного опису та формалізованого викладення всіх етапів його роботи, що обмежує можливість повного відтворення та порівняння з іншими оптимізаційними підходами.
3. У формулі 3.1 на сторінці 104 відсутній опис параметру k , що відображає кількість обраних хеш-функцій для встановлення параметрів фільтрації.
4. В роботі присутні граматичні та стилістичні помилки в незначній кількості.

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».

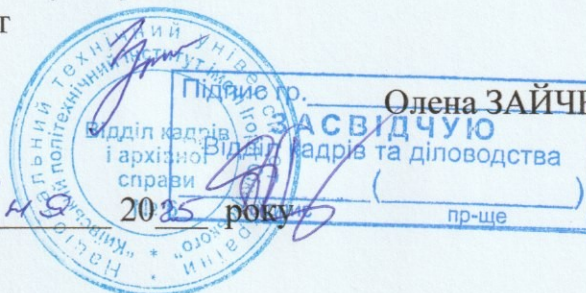
Рецензент:

Професор кафедри математичних методів
системного аналізу
Національного технічного університету України
«Київський політехнічний інститут
імені Ігоря Сікорського»
Доктор технічних наук, доцент

М.П.

« 9 » червня

2025 року



Олена ЗАЙЧЕНКО