

ВІДГУК офіційного опонента

на дисертаційну роботу Хаустової Яни Володимирівни
«Методи нечіткої кластеризації на основі ядерних функцій в задачах
інтелектуального аналізу даних», подану до захисту
на здобуття наукового ступеня кандидата технічних наук
за спеціальністю 05.13.23 – системи та засоби штучного інтелекту

Актуальність теми дисертаційної роботи. Питаннями динамічного інтелектуального аналізу потоків даних з особливими характеристиками в умовах априорної та поточної невизначеності займається багато дослідників і наукових організацій в усьому світі. Розвиток цього напрямку досліджень відбувається в межах підходів обчислювального інтелекту і характеризується створенням нових методів та засобів для послідовної обробки інформації, заснованих, зокрема, на гібридних ядерних нейронних мережах та нейро-фаззі-системах, що поєднують переваги теорії штучних нейронних мереж, фаззі-систем, методів ядерної апроксимації, тощо. Це свідчить про надзвичайну актуальність цього напрямку досліджень, що підтверджується його широким представленням на міжнародних наукових конференціях, конгресах та симпозіумах в Україні та за кордоном. З огляду на це тема дисертації Я.В. Хаустової, що присвячена розробленню методів нечіткої кластеризації на основі ядерних нейронних мереж і нейро-фаззі систем, які налаштовують свою архітектуру в процесі навчання-самонавчання за умов, коли кластери можуть перетинатися і мати довільну форму, є актуальнюю як в теоретичному, так і прикладному аспекті, оскільки такого типу задачі, пов'язані з реальними процесами різної природи, є дуже поширеними на практиці.

Ступінь обґрунтованості наукових положень, висновків та рекомендацій. Авторкою виконано ґрунтовний аналіз наявних методів інтелектуальної обробки потоків даних на основі різних підходів. За результатами цього аналізу виділено переваги та недоліки відомих методів, обґрунтовано висновок про доцільність їх удосконалення та розроблення нових методів динамічного інтелектуального аналізу потоків даних, здатних функціонувати в умовах істотної априорної та поточної невизначеності. Цей аналіз дозволив авторці обґрунтовано сформулювати наукову-практичну задачу, мету роботи, виділити об'єкт та предмет дослідження.

Основні припущення, покладені в основу теоретичних досліджень, є коректними, а отримані результати не суперечать відомим уявленням теорії інтелектуальної обробки даних. Подальша перевірка теоретичних результатів за допомогою імітаційного моделювання розроблених ядерних методів кластеризації продемонструвала високий ступінь їх адекватності в обраному

класі задач. Застосування методів математичного та імітаційного моделювання, а також обчислювальних експериментів, що використовуються в роботі, дозволяє отримати цілком обґрунтовані оцінки ефективності розроблених авторкою оригінальних методів та моделей.

Достовірність отриманих результатів. Достовірність викладених в дисертації основних наукових положень, висновків і результатів, отриманих здобувачем, забезпечується коректними постановками розв'язуваних у роботі задач та подальшим їх теоретичним аналізом, висновки якого узгоджуються з одержаними практичними результатами.

Достовірність отриманих результатів підтверджується також узгодженістю теоретичних положень з даними, отриманими при імітаційному моделюванні, належною апробацією на міжнародних конференціях і семінарах, а також впровадженням результатів дисертаційної роботи.

Застосування всіх наукових положень і результатів роботи в реальних задачах ґрунтується на детальному аналізі суті кожного об'єкта дослідження, що забезпечує коректність висновків про практичну ефективність розроблених методів та алгоритмів.

Наукова новизна результатів дисертації. Аналіз дисертаційної роботи дозволяє зробити висновок, що здобувачем у процесі досліджень отримані такі істотно нові наукові результатати:

А) вперше:

1. запропоновано ядерні кластерувальні нейронні мережі, які засновані на радіально-базисній нейронній мережі та узагальненій регресійній мережі, що дозволяють обробляти потоки даних різної фізичної природи в послідовному режимі;

2. запропоновано багатошарову гіbridну нейро-фаззі систему обчислювального інтелекту на основі системи Ванга-Менделя і нечіткої кластерувальної самоорганізовної мережі, що дозволяє в процесі самонавчання налаштовувати не тільки свої параметри, але і архітектуру в on-line режимі і вирішувати задачі кластеризації потоку даних за умов априорно невідомої форми кластерів і рівнів їх перетинання;

Б) уdosконалено:

3. метод кластеризації ЕМ шляхом використання ядерних функцій спеціального виду, що дозволяє на відміну від стандартного підходу вирішувати задачу кластеризації в умовах перетинних кластерів з розрахунком оцінки належності кожного спостереження до кожного кластеру.

4. штучну нейронну мережу для аналізу головних компонент шляхом

введення додаткових шарів ядерних функцій для підвищення розмірності вхідного простору, що дозволило обробляти інформацію, яка міститься в класах довільної форми

Зв'язок з науковими програмами, планами, темами.

Дисертаційна робота виконана в рамках держбюджетних науково-дослідних робіт: «Нейро-фаззі системи для поточної кластеризації та класифікації послідовностей даних за умов їх викривленості відсутніми та аномальними спостереженнями» (№ДР 0113U000361) та «Динамічний інтелектуальний аналіз послідовностей нечіткої інформації в умовах істотної невизначеності на основі гібридних систем обчислювального інтелекту» (№ДР 0116U002539). Здобувачкою як виконавцем розроблено методи нечіткої ядерної кластеризації, які призначені для обробки даних в on-line режимі, а кластери можуть перетинатися.

Практичне значення результатів дисертаційної роботи.

Розроблені в дисертаційній роботі нові методи нечіткої кластеризації на основі ядерних функцій і нейро-фаззі системи обчислювального інтелекту для on-line кластеризації дозволяють підвищити ефективність вирішення проблеми інтелектуального аналізу потоку інформації за умов невизначеності як про форму класів, так і про рівень їх перетинання. Використання розроблених методів кластеризації дозволило підвищити ефективність вирішення задачі автоматичної категоризації і аналізу документації програмного забезпечення, що розроблюється у ТОВ «Академія СМАРТ» (акт впровадження від 10.10.2016 р.).

Також результати дисертаційної роботи, що пов'язані з розробкою нових методів ядерної кластеризації потоків даних за наявності часткової априорної інформації про властивості даних, впроваджені в навчальний процес на кафедрі штучного інтелекту в курсах «Штучні нейронні мережі: архітектури, навчання та застосування» та «Нейромережеві методи обчислювального інтелекту» (акт впровадження від 20.10.2016 р.).

Повнота викладу основних результатів у наукових виданнях та апробація. Основні положення дисертаційної роботи опубліковані в 13 наукових працях: у тому числі у 5 статтях (з них 4 статті у фахових виданнях з технічних наук, що включені до переліку МОН України та 1 стаття за кордоном) серед яких 3 статті, що входять до міжнародних наукометричних баз Web of Science, Index Copernicus, BASE, 8 публікацій в матеріалах міжнародних наукових конференцій, з них 2 публікації входять до міжнародних наукометричних баз Scopus, Web of Science.

Результати дисертаційної роботи пройшли необхідну апробацію на багатьох міжнародних наукових конференціях, присвячених питанням обчислювального інтелекту та інтелектуального аналізу даних.

Відповідність змісту автореферату основним положенням дисертації. Оформлення автореферату за своїм обсягом, структурою та змістом відповідає чинним вимогам. Зміст автореферату ідентичний змісту основних положень дисертації, автореферат адекватно відображає результати дисертації.

Відповідність дисертації встановленим вимогам. Рецензована дисертаційна робота є завершеним і цілісним дослідженням, матеріал її добре структуровано і характеризується логічним викладом, що узагальнює дослідження авторки. Роботу написано коректною мовою з використанням сучасної науково-технічної термінології.

Оформлення дисертації відповідає вимогам п. 9, пп. 11-14 «Порядку присудження наукових ступенів». Стиль викладу матеріалів досліджень, наукових положень і рекомендацій забезпечує їх адекватне і належне сприйняття.

Зауваження по дисертаційній роботі.

Серед недоліків дисертації слід зазначити такі:

1. Формули (2.17), (2.18) є коректними лише для стаціонарних потоків даних. Замість цих виразів коректніше було б використати ідею «ковзного вікна».

2. Слід було б обґрунтувати, в яких ситуаціях є доцільним використовувати міру «схожості», а в яких – міру «несхожості».

3. Слід було б прокоментувати з яких міркувань обирається розмірність h простору підвищеної розмірності та чи можна її переналаштовувати в процесі навчання.

4. Не сформульовано загальну методику проведення експериментів із чисельного моделювання та рекомендації щодо її застосування, зокрема, відсутні належні коментарі щодо практичного вибору архітектури мереж, структури вузлів, кількості налаштовних параметрів, вибору їх початкових значень, обсягу навчальної та контрольної вибірок тощо – все, як правило, дається без пояснень в остаточному вигляді.

5. Серед результатів чисельних експериментів авторка подає переважно порівняльні графічні та табличні характеристики точності побудованих нейромереж різного типу, проте відсутні дані про показники швидкодії, що є визначальними з точки зору прикладної ефективності запропонованих розробок з огляду на складність відповідних гіbridних архітектур та алгоритмів їх навчання.

6. У роботі бракує повного опису (чи протоколу) хоча б кількох експериментів на штучних чи реальних прикладах, які б демонстрували роботу розробленого інструментарію в реальному часі.

7. Треба було більше обґрунтувати переваги ядерної функції Спанечнікова перед іншими ядерними функціями.

Вказані недоліки децо знижують загалом позитивне враження від цієї дисертації, але не впливають на високу оцінку виконаних авторкою досліджень.

Висновок по дисертаційній роботі. Дисертаційна робота Хаустової Я. В. є завершеною науковою працею, у якій вирішено актуальну науково-практичну задачу: розробку методів нечіткої кластеризації на основі ядерних нейронних мереж і нейро-фаззі систем, які налаштовують свою архітектуру в процесі навчання-самонавчання в умовах, коли кластери можуть перетинатися і мати довільну форму. Виконані в дисертаційній роботі дослідження та отримані наукові результати відповідають паспорту спеціальності 05.13.23 – системи та засоби штучного інтелекту: 11 - "Створення математичних моделей на принципах нечіткої логіки для застосування в системах штучного інтелекту", 14 - "Аналіз, синтез та моделювання нейронних мереж, розроблення методів їх проектування, оптимізації та навчання". Автореферат повністю відповідає змісту дисертації й описує суть одержаних результатів та висновків у дисертаційній роботі і оформленний згідно з чинними вимогами, що висуваються до кандидатських дисертацій.

Дисертація відповідає вимогам п. 9, 11, 12, 13, 14 «Порядку присудження наукових ступенів» (Постанова КМУ № 567, від 24 липня 2013 р.) щодо кандидатських дисертацій, а її авторка Хаустова Яна Володимирівна заслуговує на присудження її наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту.

Офіційний опонент

Профессор кафедри інформаційних
технологій видавничої справи Національного
університету «Львівська політехніка»,
доктор технічних наук, професор

Д.Д. Пелешко

Підпись Д.Д. Пелешка засвідчує
Вчений секретар Національного університету «Львівська політехніка»
к.т.н., доц. Брилинський Роман Григорович

