

ВІДГУК

офіційного опонента доктора технічних наук, професора

Зайченка Юрія Петровича на дисертацію

Мельникової Наталії Іванівни

«Моделі та методи підтримки персоналізованих рішень у медичних системах»,

яка подана на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.23. - системи та засоби штучного інтелекту

Актуальність теми дисертації.

Дослідження у напрямку персоналізованої медицини наразі є новою та потенційною темою в галузі медицини та охорони здоров'я. Інтелектуальні системи для підтримки персоналізованих рішень лікаря базуються на даних орієнтованих на пацієнта та аналізі застосування медичних протокольних рішень. Специфіка персоналізованих медичних даних полягає у низці характеристик, а саме: великій кількості якісних ознак, значній варіабельності медичних даних, малих вибірках з невідомими законами розподілу, нечіткості, наявності викидів, шумів, надлишковості та пропусків даних, що є елементами набору даних, подій або спостережень. Враховуючи особливості цього типу даних, процеси їхнього аналізу та опрацювання, а також сторонні фактори впливу, виникає необхідність пошуку комплексного рішення у задачах прогнозування станів пацієнта чи приналежності його до певного класу.

Дисертаційна робота Мельникової Н.І. присвячена науковій проблемі розроблення та удосконалення моделей, методів і засобів інтелектуального аналізу в задачах класифікації, кластеризації, прогнозування та візуалізації результатів опрацювання персональних даних для адаптації медичних рішень до пацієнта. У роботі ставляться і розв'язуються задачі аналізу проблеми опрацювання мультимодальних персоналізованих медичних даних, розроблення інформаційної моделі стану пацієнта у багатовимірному просторі умов, розробки методу двоетапної обробки малих наборів даних, методу групування моделей інтелектуального аналізу персоналізованих медичних даних, методу зменшення розмірності вхідних даних для підвищення точності вибору пріоритетних ознак на великих наборах даних,

методу заповнення пропусків даних для підвищення стійкості моделі до помилок даних та паралелізації обчислень, методу пошуку зміни стану пацієнта з урахуванням простору умов та аналізу характеристик його стану, методу консолідації мультимодальних даних для забезпечення агрегації даних з різною структурою.

Актуальність обраної теми підтверджується зв'язком дисертаційної роботи з важливими програмами та темами, зокрема, «Методи та засоби обробки, консолідації та аналізу персоналізованої медичної інформації» за темами: ДБ «Інформаційна технологія опрацювання персоналізованої медичної інформації» №0119U002257. 2019-2020 (авторка розробила методи та засоби підтримки прийняття лікарських рішень для покращення якості надання медичної допомоги); Госпдоговір ТОВ «БІОФАРМА ПЛАЗМА», (авторкою розроблена інформаційна система обліку та аналізу пацієнтів з орфаними хворобами «Реєстр первинних імунодефіцитів»); CEI Extraordinary (STOP COVID 19). 2020. (авторкою проведені дослідження щодо вирішення задачі пошуку залежностей між цільовими ознаками хворих на Covid-19, що дозволило довести закономірності залежності від поширеності COVID-19 на території різних країн та знайти рішення щодо зниження поширеності COVID-19 серед людей).

Роботи здобувачем виконані як відповідальним виконавцем.

Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації, їхня достовірність.

Обґрунтованість і достовірність отриманих наукових положень, висновків та рекомендацій забезпечується використанням перевірених підходів, підтверджених теорією та практикою за рахунок використання методів прийняття рішень – для забезпечення ціле-направленого пошуку рішень стосовно вибору цільових схем лікування; теорії реляційних баз даних – для відображення простору умов щодо ідентифікації стану хворого та урахування додаткових вимірів та параметрів; системи операцій реляційної алгебри – для формалізації залежностей між множинами параметрів пацієнта та показників оцінки його стану; методів штучного інтелекту – для виявлення закономірностей та залежностей між цільовими змінними, кластеризації груп пацієнтів за відповідними ознаками, для усунення невизначеностей у просторі даних, для прогнозування цільових змінних.

Оцінюючи дисертаційну роботу в цілому, слід відзначити новизну запропонованої і реалізованої автором моделі відображення стану пацієнта в багатовимірному просторі умов та наукову цінність введеної теорії підтримки прийняття персоналізованих рішень. Характерними особливостями запропонованої теорії є розроблення методів двоетапної обробки малих наборів даних, групування моделей машинного навчання, зменшення розмірності вхідних даних, заповнення пропусків даних, пошуку зміни стану пацієнта та консолідації мультимодальних даних, що забезпечило адаптацію прийняття персоналізованих медичних рішень.

Достовірність результатів дисертаційної роботи підтверджується експериментальними дослідженнями в галузі медицини, а саме:

- методи та ансамблі моделей прогнозування даних для оцінки стану хворого доведено до практичної реалізації інформаційної системи підтримки прийняття рішень для лікування хворих з орфанними хворобами використано ТОВ «БІОФАРМА ПЛАЗМА»;
- програмний модуль щодо прогнозування динаміки поширення COVID-19 в межах України, Білорусії, Німеччини використано в дослідженнях Львівського Медичного Університету ім. Д.Галицького;
- програмний модуль щодо аналізу наявності ішемічного інсульту та визначає час від його появи, програмний модуль для розв'язання задачі класифікації щодо виявлення наявності COVID-19 у пацієнта використано під час аналізу стану хворого під час лікування та корекції призначення схеми лікування в Клінічній лікарні швидкої допомоги, Першому медичному об'єднанні міста Львова.

Наукова новизна дисертаційної роботи.

Ознайомлення із змістом та результатами дисертаційної роботи дає змогу зробити висновок, що дисертанткою запропоновано модель відображення стану пацієнта в багатовимірному просторі умов та наявні методи опрацювання, аналізу персоналізованих медичних даних. Сформульовано і вирішено важливу науково-прикладну проблему розроблення та удосконалення моделей, методів і засобів інтелектуального аналізу персоналізованих медичних даних в задачах класифікації, кластеризації, прогнозування та візуалізації результатів опрацювання персональних даних для адаптації медичних рішень до пацієнта. Отримано такі нові наукові результати:

вперше

- розроблено модель відображення стану пацієнта в багатовимірному просторі умов, що за рахунок додаткових вимірів та параметрів забезпечила підвищення точності прогнозування;
- розроблено метод двоетапної обробки малих наборів даних на основі ієрархічного предиктора, який за рахунок кластеризації та прогнозування забезпечує підвищення точності опрацювання нових вхідних наборів даних;
- розроблено спосіб групування моделей інтелектуального аналізу, який, враховуючи основні ознаки та повторне навчання, забезпечує підвищення точності прогнозування даних;
- розроблено метод зменшення розмірності вхідних даних, який завдяки класифікаторам, асоціативним правилам та узагальненому рангу ознак на основі індексу Жакара, забезпечує підвищення точності вибору пріоритетних ознак на великих наборах даних;

удосконалено

- метод заповнення пропусків даних, який, використовуючи асоціативні та продукційні правила, забезпечує підвищення стійкості моделі до помилок даних;
- метод пошуку зміни стану пацієнта, який, використовуючи простір умов та аналіз характеристик стану, забезпечує підвищення точності підбору схеми лікування;
- метод консолідації мультимодальних даних, який за рахунок попереднього визначення структури та семантики даних забезпечує агрегування даних з різною структурою;

одержала подальший розвиток

- теорія підтримки прийняття персоналізованих рішень, яка ґрунтується на розробленні моделі відображення стану пацієнта в багатовимірному просторі умов та розроблених методах двоетапної обробки малих наборів даних, зменшенні розмірності вхідних даних, заповненні пропусків даних, пошуку зміни стану пацієнта та консолідації мультимодальних даних, що забезпечило адаптацію прийняття персоналізованих медичних рішень.

Значення одержаних результатів для науки і практики.

Практичну цінність роботи підтверджує впровадження її результатів у дослідженнях проведених в межах проектів, а також використання її результатів при виконанні держбюджетних тематик Національного університету «Львівська політехніка».

Використання розроблених в дисертації методів та засобів та інших результатів дозволило:

- підвищити точність прогнозування цільових показників на 5% в підмножині простору умов, що дало змогу забезпечити індивідуальний підхід до аналізу стану пацієнта;
- підвищити точність класифікації для набору даних по Covid-19 за рахунок ієрархічної класифікації та поєднання різних моделей ІАД;
- підвищити точність прогнозування даних на 7-9% та забезпечити паралелізацію процесу обробки даних як малої, так і великої розмірності за рахунок розробленої стекінгової моделі на основі випадкового лісу;
- зменшити ймовірність появи похибки при виборі схеми лікування за рахунок персоналізації стандартних схем лікування аналізу динаміки приросту значень часово-залежних даних;
- розробити інформаційну систему підтримки прийняття рішень у лікуванні хворих з первинним імунодефіцитом;
- розробити програмний модуль для підтримки прийняття рішень у визначенні схеми лікування пацієнтів з постковідним синдромом.

Рекомендації щодо використання результатів дисертації

Використання запропонованих підходів щодо опрацювання даних малої та великої розмірності дасть змогу одержати результати з вищою точністю в задачах пошуку персоналізованих рішень за рахунок прогнозування станів досліджуваного об'єкта, кластеризації нових вхідних наборів даних, зменшення розмірності вхідних даних, заповнення пропусків даних у таких предметних областях, що мають потребу у проведенні такого типу обчислень, а саме: медицина – для підтримки прийняття рішень, як навчальна база для лікарів-початківців та студентів-медиків, соціальні служби для покращення якості життя громадян за рахунок надання якісного обслуговування самотніх та/чи хворих людей в умовах не стаціонару, у

військовій сфері надання медичних послуг в умовах обмеження часу на прийняття рішень.

Конкретні шляхи використання запропонованих в дисертації Мельникової Н.І. моделей, методів, алгоритмів та програмних засобів полягають у підвищенні точності пошуку рішень на базі опрацювання та аналізу персоналізованих медичних даних під час отримання нових знань та розроблення якісних інтелектуальних медичних систем.

Повнота викладу в опублікованих працях.

Результати дисертаційного дослідження опубліковано у 56 наукових публікаціях, із них 21 стаття, 35 публікацій у матеріалах конференцій. Загалом опубліковано 5 монографічних робіт, дві з яких – одноосібні, а також 3 навчальні посібники, 4 авторських свідоцтва на твір.

Оцінка змісту дисертаційної роботи, її завершеність.

На основі аналізу відомих підходів до опрацювання персоналізованих медичних даних визначено об'єкт та предмет досліджень, постановку проблеми дослідження, розроблено модель відображення стану пацієнта в багатовимірному просторі умов та методи двоетапної обробки малих наборів даних, групування моделей машинного навчання, зменшення розмірності вхідних даних, заповнення пропусків даних, пошуку зміни стану пацієнта та консолідації мультимодальних даних, що забезпечило адаптацію прийняття персоналізованих медичних рішень. Розроблено відповідні алгоритми та наведено результати розробки інструментальних алгоритмічних та програмних засобів та їх використання в умовах реальних досліджень.

У вступі виконано аналіз наукової проблеми, обґрунтовано актуальність досліджень, наукову новизну та практичну значимість досліджень, розглянуто загальну характеристику роботи, особистий внесок здобувача в публікації, виконані у співавторстві, апробацію роботи, кількість та якість публікацій.

У першому розділі проаналізовано існуючі практичні рішення, що відображають основні підходи до використання медичних даних та урахування персоналізації. Проведено порівняльний аналіз класичних методів та визначені обмеження, які виникають при опрацюванні медичних персоналізованих даних. За результатами досліджень сформульовані основні етапи досліджень. Визначені обмеження під час опрацювання універсальними методами великих та малих мультимодальних медичних

наборів даних, що сформувало актуальну науково-прикладну проблему розроблення нових чи удосконалення існуючих методів штучного інтелекту щодо підвищення точності процесу обробки медичних даних для пошуку персоналізованих рішень.

У другому розділі введено модель відображення стану пацієнта, яка відображає його стани; формалізовано відображення фізичного стану пацієнта з урахуванням часовозалежних та часовонезалежних даних пацієнта, що дає змогу оцінити стан пацієнта в певний момент часу; введена інформаційна модель простору станів пацієнта представлено у вигляді гіперкуба як відображення функціонального відношення загального стану пацієнта; розроблено метод пошуку послідовних шаблонів та визначено етапи пошуку, що базуються на модифікації методу асоціативних правил.

У третьому розділі розроблено метод двоетапної обробки малих наборів даних, що базується на основі ієрархічного предиктора для підвищення точності процесу прогнозування та узагальнення результатів; розроблено новий метод групування моделей інтелектуального аналізу на основі модифікації стекінгу для збільшення повторюваності результатів на малих вибірках даних.

У четвертому розділі проаналізовано існуючі методи відбору важливих ознак: фільтри, вбудовані алгоритми. Розроблено гібридний ансамбль вибору ознак на основі комбінації селекторів та агрегування результатів, який використовується на етапі попередньої обробки на великих наборах даних.

У п'ятому розділі визначені типи аномалій та шляхи їхнього пошуку, що дає можливість диференціації підходів щодо аналізу персоналізованих медичних даних та визначені особливості їхнього застосування. Розроблено метод заповнення пропусків даних на основі ймовірнісних продукційних залежностей, що збільшує стійкість моделі до помилок у даних та забезпечення аналізу мультимодальних даних при паралельній реалізації в розподілених базах даних.

У шостому розділі розроблено архітектуру системи підтримки прийняття медичних рішень щодо прогнозування станів пацієнта на підставі опрацювання та аналізу персоналізованих медичних даних; подано опис імплементації рішень у системі підтримки прийняття рішень для лікування хворих на первинні імунodefіцити з порушеннями антитілоутвореннями; подано результати розробки програмного модулю щодо прогнозування

динаміки поширення COVID - 19; подано результати імплементації системи для виявлення помилок.

У висновках здобувач навела одержані в роботі результати, відзначила новизну та практичну значимість дисертаційних досліджень.

У додатках подано результати впроваджень дисертаційних досліджень.

Загалом дисертаційна робота написана грамотно, технічною мовою. Матеріали досліджень подано логічно, послідовно та доказово. Текст автореферату відповідає змісту дисертаційної роботи.

Зауваження до дисертації:

1. Багато показників стану здоров'я пацієнтів мають якісний характер, деякі є нечіткими, тому для їх опису доцільно було б використати лінгвістичні змінні, а для побудови логічних висновків – використати системи нечіткої логіки та нечіткі нейронні мережі, що дало б змогу проводити більш змістовний аналіз стану пацієнтів та більш якісні висновки. Про те на жаль, автор не використовує в своїй роботі цей потужний сучасний апарат штучного інтелекту.

2. Не зовсім вдало розкрито практичне значення результатів дисертації. Автор вважає, що « науково- практичне значення результатів роботи полягає в тому, що вони є структурною складовою процесу підтримки прийняття рішень з урахуванням персоналізованих медичних даних...» .

А на справді практичне значення полягає в тому, що розроблені методи, алгоритми та запропонована модель стану пацієнтів в просторі ознак дозволяють підвищити якість оцінки стану пацієнтів та точність прогнозування розвитку хвороби і тривалість лікування в порівнянні з відомими методами.

3. Сумнівне твердження, що «...логістична регресія є найкращим класифікатором для набору даних по Covid-19», стор. 237. Це ніким не доведено!

4. Є зауваження , щодо формулювання мети дисертації .

Автор зазначає, що «метою дисертаційного дослідження є розробка моделей та методів машинного навчання для підвищення прогнозованої точності та візуалізації результатів доопрацювання персоналізованих даних щодо стану хворого ».

Проте насправді, розроблені автором методи не є методами машинного навчання, а є методами інтелектуального аналізу медичних даних для вирішення конкретних задач, які виникають в процесі обробки персоналізованої інформації пацієнтів, зокрема, вибору методів та протоколів лікування та прогнозування результатів їх застосування, прогнозування тривалості лікування тощо.

5. Висновок 6 до розділу 3 дисертації не відредагований і тому не зовсім зрозумілий.

6. Погано сформульований висновок 4 дисертації, а саме:

« Вперше розроблено метод групування моделей машинного навчання як стекінгову модель алгоритмів машинного навчання,,». Метод не може бути моделлю, навіть стекінговою. Крім того, на мій погляд це не метод, а спосіб групування алгоритмів і не машинного навчання, а класифікації (Random Forest, SVM і т.д.).

7. Є зауваження редакційного плану:

- Рис. 6.52- це не матриця помилок, а матриця якості класифікатора, оскільки в ній по діагоналі вказані правильні класифікації;

- Деякі таблиці автор чомусь називає рисунками, зокрема, рис. 6.53, рис 6.54 та рис. 6.56 та деякі інші;

- Аналіз зміни приросту значень часово-залежних даних- краще «аналіз зміни значень даних», оскільки зміни та прирост це одне й те ж саме.

8. Є деякі помилки та описки в формулах 2.43 та 2.44 на стор. 121, а саме:

- В формулі (2.43) не повинен бути індекс і зліва, оскільки справа є сума по цьому індексу;

- В (2.44) якась плутанина-зліва та справа одна і та ж змінна Q_{il} .

- Крім того, не пояснена змінна Q_{const} .

Вказані зауваження дещо знижують враження від рецензованої дисертації, проте те впливають на основні наукові результати та висновки дисертації.

Висновок про відповідність дисертації встановленим вимогам

- Дисертація Мельникової Н.І. «Моделі та методи підтримки персоналізованих рішень у медичних системах» є завершеною працею, в якій отримано нові науково обґрунтовані результати, що в сукупності вирішують проблему розроблення та удосконалення моделей, методів і засобів класифікації, кластеризації, прогнозування та візуалізації результатів опрацювання персональних даних для адаптації медичних рішень до пацієнта.
- Дисертаційна робота Мельникової Н.І. відповідає спеціальності 05.13.23 – системи та засоби штучного інтелекту, зокрема, результати роботи відповідають вимогам пунктів: розроблення теоретичних засад створення та застосування систем штучного інтелекту різноманітного призначення; моделювання інтелектуальної діяльності людини та його застосування в системах штучного інтелекту, розроблення принципів, методів й архітектури побудови баз знань і технологія їх експертизи (експертні та багатоагентні системи).
- Основні результати роботи достатньо висвітлені у 56 наукових працях, серед яких 21 стаття, з них: 2 статті опубліковано в журналах з Q1, 3 статті опубліковано в журналах з Q2, 7 статей у інших фахових іноземних виданнях, 7 статей у наукових фахових виданнях України, 2 монографії.
- На підставі вищевказаного вважаю, що дисертаційна робота відповідає вимогам п. 7 та 9 Порядку присудження та позбавлення наукового ступеня доктора наук, затвердженого постановою Кабінету Міністрів України, а її автор Мельникова Наталія Іванівна заслуговує присвоєння наукового ступеня доктора технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту.

Офіційний опонент

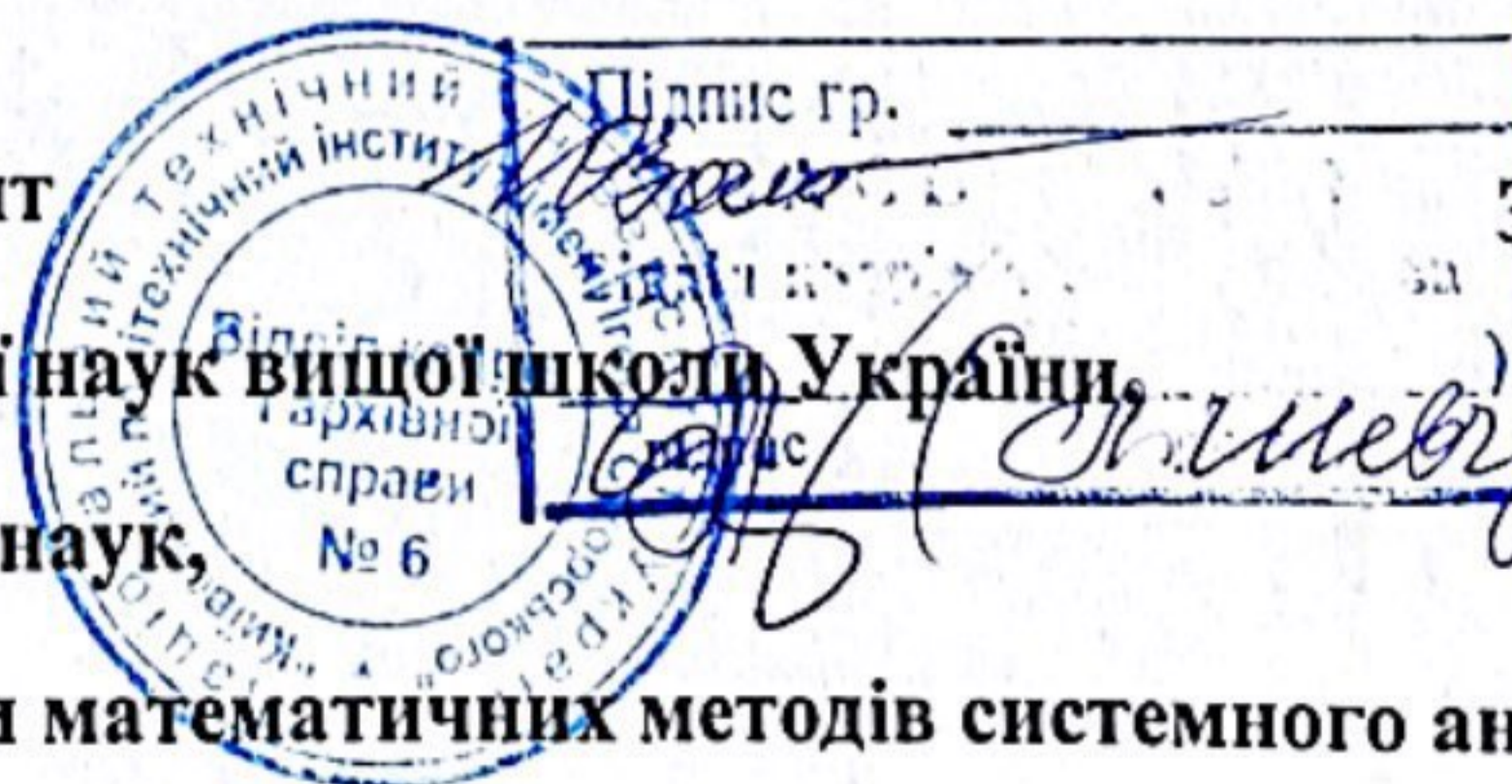
академік Академії наук вищої школи України,

доктор технічних наук,

професор кафедри математичних методів системного аналізу,

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»



Зайченко Ю.П.