

А. І. Поворознюк, О. А. Поворознюк, Х. Шехна

Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

СИНТЕЗ КОМБІНОВАНОГО ДІАГНОСТИЧНОГО ВИРІШАЛЬНОГО ПРАВИЛА В МЕДИЧНИХ СИСТЕМАХ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

Анотація. Робота присвячена вирішенню актуальної науково-технічної задачі побудови системи підтримки прийняття рішень на основі реалізації розробленої моделі діагностичного вирішального правила засобами сучасних інформаційних технологій, використання яких дозволило забезпечити працездатність розробленої системи на різних апаратних платформах під управлінням різних операційних систем. На основі аналізу методів, які використовуються для побудови вирішальних правил в системах підтримки прийняття рішень, запропоновані складові комбінованого вирішального правила, що виражають два підходи до формування діагностичного висновку: об'єктивна, яка заснована на аналізі навчальної вибірки, і суб'єктивна, яка заснована на експертній інформації про структуру симптомокомплексів. **Мета дослідження** – синтез комбінованого вирішального правила на основі методу порівняння з прототипом, яке б дозволило врахувати як об'єктивну, так і суб'єктивну складову процесу постановки діагнозу. **Результати.** В роботі розроблена математична модель комбінованого діагностичного вирішального правила і обґрунтовано вибір його складових. В якості об'єктивної складової вибрано метод порівняння з прототипом, в якому діагностуємі стани представляються їхніми прототипами в просторі ознак. Формалізована експертна інформація про структуру симптомокомплексів шляхом представлення симптомом комплексів захворювань числовими інтервалами лінгвістичних змінних. Розглянуті варіанти врахування експертних оцінок щодо структури симптомокомплексів при обчисленні координат прототипів класів. Сформульовані вимоги до функціональних можливостей системи, визначено засоби проектування, основну платформу розробки (Java), систему управління базою даних (MySQL). Виконано проектування системи підтримки прийняття рішень та комплексна перевірка розробленої системи на реальних медичних даних, яка підтвердила ефективність роботи системи.

Ключові слова: медична діагностика, вирішальне правило, метод порівняння з прототипом, симптомокомплекс, експертна інформація, система підтримки прийняття рішень.

Вступ

Активний розвиток високотехнологічного суспільства сприяє впровадженню науково-технічних розробок в таку важливу, проте мало формалізовану сферу діяльності, як медицина, завдяки чому бурхливого розвитку набуває напрям біомедичної інженерії, який охоплює різноманітні галузі: біомеханіка; розробка нових біоматеріалів, біосенсорів, інструментів та приладів; реабілітаційна інженерія; медичний та біологічний аналіз та інші.

На сьогодні нараховується сім рівнів медичних інформаційних систем (МІС) [1-3]: від електронних амбулаторних карт до інтелектуальних систем підтримки прийняття рішень з синергетичними базами даних, методами штучного інтелекту та можливостями телемедицини [4]. Переважна більшість цих систем спрямована на автоматизацію проведення первинних обстежень в різних предметних галузях медицини, що пов'язані з виявленням та фіксацією діагностичних ознак, обробкою біомедичних сигналів та зображень. Не зважаючи на широке розмаїття існуючих МІС та складність задач, що ними вирішуються, однією з актуальних невіршених оптимізаційних задач залишається синтез діагностичного вирішального правила, яке б забезпечило достовірність та обґрунтованість поставленого діагнозу.

Постановка проблеми в загальному вигляді. Традиційна первинна діагностика заснована на систематичному огляді пацієнта, аналізі анамнезу, скарг та об'єктивних ознак захворювання, що були виявлені при фізичному дослідженні – огляді, перкусії та ін. а також за допомогою лабораторно-інструментальних досліджень. На інтуїтивному рівні практикуючий лікар використовує наступну логіку постановки діагнозу, яка відповідає принципам доказової медицини.

Виявленні ознаки (симптоми) захворювання лікар об'єднує в синдроми (сукупність симптомів, що мають спільний патогенез), на основі яких робиться остаточний висновок про можливе захворювання.

Діагноз (грец. $\Delta\iota\alpha\gamma\nu\sigma\tau\acute{\iota}\varsigma$ - розпізнавання, визначення) – це медичний висновок про функціонально-морфологічний стан здоров'я людини, про наявні хвороби або травми, про зміни нормального функціонування органів або систем, а потім і організму взагалі, або ж про причини, які привели до смерті, тобто висновок, виражений в термінах, передбачених прийнятими класифікаціями та номенклатурою захворювань. При постановці діагнозу лікарем, умовно виділяються типові етапи процесу формування діагнозу [5]: 1) виявлення відхилення від норми в показниках стану організму; 2) тлумачення відхилення і їх значимість (фізіологічний або патологічний характер, можливий зв'язок з певною нозологічною формою); 3) формування попереднього діагностичного висновку; 4) підтвердження діагнозу, тобто встановлення остаточного діагнозу (в іншому випадку процес діагностики поновлюється, починаючи з першого, другого або третього етапу). При цьому важливу роль відіграє кваліфікація та досвід лікаря.

В МІС різних рівнів постановка комп'ютерного діагнозу в формалізованому виді представляється в класичній постановці задачею класифікації, в якій моделлю об'єкта діагностики (ОД) є "чорна шухляда", і шукається залежність між формалізованими станами ОД Y і вектором вхідних ознак X , тобто $Y = f(X)$ [1]. Вказана залежність визначається на етапі навчання при аналізі навчальної вибірки, елементами якої виступають пацієнти з підтвердженим діагнозом. При цьому репрезентативність вибірки не завжди можна досягти в існуючих медичних базах даних, особливо при діагностиці рідкісних захворювань. Крім того, практикуючі лікарі з недовірою

відносяться до результатів комп'ютерного діагнозу та зазнають труднощів при інтерпретації результатів діагностики. Тому перспективним напрямком дослідження є розробка методу постановки діагнозу – синтез діагностичного вирішального правила (ВП), який об'єднує вказані підходи.

Аналіз останніх досліджень і публікацій. В даний час є широкий спектр комп'ютерних систем підтримки прийняття рішень (СППР) в різних предметних областях медицини [2, 3], в яких використовуються такі математичні методи підтримки прийняття рішень: детерміністичні методи [6]; імовірнісні методи [7]; методи, засновані на теорії розпізнавання образів [8]; методи, засновані на нечіткій логіці [9]; логіко-лінгвістичні методи [8]; методи на основі штучних нейронних мереж [10] та ін. Розглянемо більш детально найбільш вживані методи.

Детерміністичні методи. Застосовуються у випадках наявності детермінованих зв'язків між ознаками і формалізованими станами об'єктів, як правило на етапі попередньої класифікації. Детерміністичні методи і методи нечіткої логіки застосовуються для формалізації експертних оцінок щодо структури симптомокомплексів.

Імовірнісні (статистичні) методи. Ці методи засновані на використанні апарата математичної статистики [11]. Вони найчастіше застосовуються у випадках, коли відомі ймовірнісні характеристики класів або коли вони можуть бути визначені за наявною навчальною вибіркою, що звужує область їхнього застосування.

Методи, засновані на теорії розпізнавання образів. У даній групі методів результати вимірювання характеристик об'єктів представляються точками в просторі діагностичних ознак. При цьому різні класи повинні утворювати компактні множини в просторі ознак. Діагностика нового об'єкта зводиться до обчислення міри близькості до кожного класу. До даної групи методів належить *метод порівняння з прототипом (еталоном)*, який найчастіше використовується при аналізі числових ознак, в випадку, коли класи Ω_m ($m = \overline{1, M}$) утворюють компактні множини об'єктів, що мають сферичну форму в просторі ознак. В роботі використовується цей метод в якості об'єктивної складової ВП.

Мета і задачі дослідження. Метою дослідження є синтез вирішального правила на основі методу порівняння з прототипом, яке б дозволило врахувати як об'єктивну, так і суб'єктивну складову процесу постановки діагнозу.

Результати досліджень

Розглянемо формальну модель комбінованого вирішального правила. Проаналізуємо особливості використання методу порівняння з прототипом. Як було відмічено раніше, в методі передбачається, що діагностуємі класи утворюють компактні множини об'єктів, що мають сферичну форму в просторі ознак. У цьому випадку кожний із класів Ω_m ($m = \overline{1, M}$) описується прототипом або еталоном $\omega^{m\bar{e}}$, у якості якого вибирається геометричний центр угруповання класу. При реалізації методу розрізня-

ють два етапи: етап навчання, та етап класифікації. На етапі навчання обчислюються координати прототипу кожного класу по формулі

$$x_i^{m\bar{a}} = \frac{1}{n_m} \sum_{j=1}^{n_m} x_i^j, \quad i = \overline{1, p}, \quad m = \overline{1, M}, \quad (1)$$

де $x_i^{m\bar{a}}$ – i -та координата етального об'єкта класу m ; n_m – кількість об'єктів класу m в навчальній вибірці; x_i^j – i -та координата j -го об'єкта класу m в навчальній вибірці; p – розмір координатного простору (кількість діагностичних ознак); M – кількість класів, на які виконується класифікація невідомого об'єкта (кількість діагностуємих станів в даній предметній області медицини);

На етапі класифікації (постановка діагнозу) обчислюються відстані $R(\omega, \omega^{m\bar{y}})$ від точки класифікуємого об'єкта ω в просторі ознак до кожного еталоноу $\omega^{m\bar{y}}$ класу m по формулі (в даному випадку наведена евклідова відстань, яка зазвичай застосовується до чисельних даних)

$$R(\omega, \omega^{m\bar{a}}) = \sqrt{\sum_{i=1}^p \left(x_i^\omega - x_i^{m\bar{a}} \right)^2}, \quad m = \overline{1, M}, \quad (2)$$

де $x_i^{m\bar{a}}$ – i -та координата етального об'єкта класу m ; x_i^ω – i -та координата класифікуємого об'єкта; p – розмір координатного простору (кількість діагностичних ознак); M – кількість класів, на які виконується класифікація. Невідомий об'єкт ω належить до класу $\omega^{\bar{y}}$, відстань до прототипу якого $R(\omega, \omega^{\bar{y}})$ буде мінімальною:

$$R(\omega, \omega^{\bar{y}}) = \min_{m=\overline{1, M}} R(\omega, \omega^{m\bar{y}}), \quad (3)$$

де $R(\omega, \omega^{m\bar{y}})$ – відстань між об'єктом ω та еталоном $\omega^{m\bar{y}}$ класу ω^m , яка обчислюється по формулі (2).

Застосування методу порівняння з прототипом для задач медичної діагностики має ряд обмежень:

- в медичній діагностиці використовується система різномірних ознак (числові, рангові та бінарні), а реалізація методу вимагає використання числових ознак, тому систему різномірних ознак необхідно привести до єдиної шкали;

- метод ефективно працює, коли класи об'єктів утворюють компактні множини об'єктів сферичної форми в просторі ознак, що не завжди має місце в медичних БД;

- точність обчислення координат етального об'єкта кожного класу залежить не тільки від точності вимірювання ознак, але і від кількості об'єктів навчальної вибірки в кожному класі (чим менше об'єктів, тим вище дисперсія оцінки математичного очікування), тому невизначеність координат етального об'єкта може привести до неправильної класифікації.

- даний метод не враховує структуру симптомокомплексів, що визиває недовіру до його результатів з боку лікарів-практиків.

Тому в роботі при розробці комбінованого ВП виконується модифікація методу для подолання наведених обмежень.

Врахування експертної інформації щодо структури симптомокомплексу. Інформацію про симптоми захворювань у неформалізованому виді можна знайти в різних медичних довідниках [12], відомості яких можна вважати експертною оцінкою захворювання, виробленої багатьма поколіннями лікарів. Розрізняють наступні типи симптомів, що формують симптомокомплекс: *патогномонічні* (однозначно свідчать про наявність заданого захворювання), *специфічні* (присутні при заданому захворюванні, але не вказують однозначно на його наявність), *неспецифічні* (можуть бути присутнім при заданому захворюванні).

Структура симптомокомплексу є вираженням поглядів експертів щодо конкретного діагнозу, причому вплив кожної ознаки розглядається в термінах деякої лінгвістичної змінної, що описує розглянуту ознаку (наприклад, “висока температура” або “підвищений артеріальний тиск”).

Крім того, кожній з ознак x_j виставляється експертна оцінка, що виражає її вагу симптомокомплексу. Ці оцінки приймають значення e_0, e_1, e_2, e_3 , де e_0 – вага патогномонічних симптомів; e_1 – вага специфічних симптомів; e_2 – вага неспецифічних симптомів; e_3 – вага показників, що не входять у симптомокомплекс даного захворювання. При цьому:

$$e_0 \geq e_1 \geq e_2 \geq e_3; \quad \sum_{i=1}^3 e_i = 1. \quad (4)$$

Розглянемо врахування експертних оцінок щодо структури симптомокомплексу при обчисленні координат еталонів класів. Лікарі-фахівці частіше за все оперують поняттями норми того чи іншого показника, що виражається у розбитті динамічного діапазону ознаки на три інтервали: “нижче норми”, “норма”, “вище норми”. Таким чином, симптомокомплекс кожного захворювання (еталон кожного класу) описується множиною діагностичних ознак x_j , кожна з яких приймає одне з наведених вище значень лінгвістичної змінної. Аналогічним чином задаються бінарні ознаки, для яких лінгвістичними змінами є два терми: «ознака присутня» чи «ознака відсутня».

Представлення симптомокомплексу прототипу лінгвістичними змінами є експертною оцінкою прототипу. Якщо відомі значення динамічних діапазонів всіх ознак, а також граничні значення «норми», то центри цих діапазонів і є числовим вираженням еталона класу ω^{ml} , який одержано не по навчальній вибірці, а по експертній оцінці симптомокомплексу.

В остаточному підсумку, різні підходи (підхід на основі навчальної вибірки, і підхід на основі експертної оцінки симптомокомплексу) призначені для рішення однієї й тієї ж задачі класифікації. Будучи отриманими з різних передумов (статистика й думка експерта), вони описують загальну проблему з різних позицій. Природним є припущення, що їхнє спільне використання перспективно, тому можуть бути запропоновані такі варіанти спільного використання:

1. Колектив ВП. Метод проілюстрований рис. 1.

Симптомокомплекс X аналізується за допомогою кожного із ВП, яким заданий ступінь довіри ω . Блок логічного виводу формулює остаточний діагноз D_k на підставі зважених результатів роботи вирішальних правил. Варіанти реалізації блоку логічного виводу:

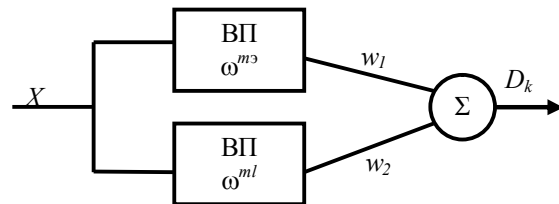


Рис. 1. Структурна схема колективу вирішальних правил

$$\min(R(\omega, \omega^{mā}), (R(\omega, \omega^{ml}))) \rightarrow D_k, \quad (5)$$

$$\min(w_1 R(\omega, \omega^{mā}) + w_2 (R(\omega, \omega^{ml}))) \rightarrow D_k. \quad (6)$$

2. Підсумовування оцінок. Остаточний варіант комбінованого вирішального правила, що пропонується в даній роботі полягає у використанні суми оцінок складових кожної координати еталона у методи порівняння з еталоном. При цьому координати еталона на етапі навчання обчислюються за формулою

$$x_i^m = k_1 x_i^{mā} + k_2 x_i^{ml}, \quad (7)$$

де $k_i > 0, \sum k_i = 1, i = \overline{1, 2}$.

Вагові коефіцієнти k_i визначаються на етапі навчання системи і відповідають ступеню довіри до кожної складової. Кінцевий користувач-експерт повинен мати можливість корегувати ці значення на власний розсуд під особисту відповідальність.

В подальшому класифікація нових об'єктів виконується за стандартним алгоритмом методу порівняння з еталоном. У будь-якому разі, діагноз визначений за допомогою запропонованого ВП носить дорадчий характер. Рішення про постановку остаточного діагнозу приймає лікар.

Реалізація СППР та тестова перевірка. Виконано обґрунтування реалізації системи підтримки прийняття рішень. Найбільш привабливими засобами розробки програмного забезпечення на сьогодні є ті, що дозволяють створювати додатки, які є кросплатформним на рівні запуску. Лідерами в цій галузі є платформи. NET (Microsoft) і Java (Oracle Corporation (раніше Sun Microsystems)). Операційна система (ОС) Windows займає майже 90% ринку ОС для персональних комп'ютерів, таким чином, була обрана ОС Windows. Також її API дозволяють побудувати необхідний інтерфейс програм. Враховуючи економічну складову і наявність останніх версій середовищ виконання для більшості сучасних платформ (Windows, UNIX, Linux) в якості платформи розробки обрано Java. Дана мова дозволяє використовувати на пряму API системи Windows, має гнучкі можливості для роботи з пам'яттю, має багато бібліотек для роботи з інтерфейсом. В якості засобів проектування баз даних обрано MySQL. Сформульовані вимоги до функціональних можливостей системи. Розроблена архітектура програмного забезпечення системи, яка включає складові: ядро системи (забезпечує роботу з даними, формування знань, здійснює діагностику), базу даних та інтерфейс користувача (відповідність шаблону MVC в архітектурному плані). Виконано тестування системи на реальній медичній базі даних навчальної вибірки, яка містить 253 записи навчальної вибірки, що включає 9 ознак клінічного аналізу крові та 4 діагнози. Її структура наведена у табл. 1.

Результати тестових випробувань показують, що спільне використання складових вирішального правила дозволяє підвищити ефективність постановки діагнозу на 13.4% відносно самостійного використання методу порівняння з прототипом (ефективність самостійного використання методу порівняння з прототипом 76.73%, самостійної діагностики, заснованої на описі симптомокомплексів – 37.56%, спільне використання – 90.1%).

Таблиця 1 – Структура навчальної вибірки

МКХ-10	Назва	Кількість осіб
D50	Залізодефіцитна анемія	60
D53.0	Анемія внаслідок недостатності білків	40
D55	Анемія внаслідок ферментних порушень	53
-	Умовно здоровий	100

Висновки

У статті обґрунтовано необхідність синтезу комбінованого вирішального правила, яке поєднує

два підходи до формулювання діагностичного висновку: об'єктивна, що заснована на аналізі навчальної вибірки, і суб'єктивна, яка заснована на експертній інформації про структуру симптомокомплексів. Розроблено математичну модель вирішального правила на основі модифікації методу порівняння з прототипом шляхом внесення експертної інформації щодо структури симптомокомплексів при обчисленні координат прототипів класів. Запропоновано варіанти спільного використання його складових (колектив вирішальних правил, зважування та підсумовування оцінок). Розроблено програмне забезпечення системи підтримки прийняття рішень та виконано тестову перевірку на реальних медичних даних. Результати тестових випробувань підтвердили працездатність системи та показали ефективність постановки діагнозу.

Подальші дослідження направлені на застосування розробленого вирішального правила при аналізі не тільки числових ознак, але і різномірної діагностичної інформації, включаючи біомедичні сигнали та зображення.

СПИСОК ЛІТЕРАТУРИ

1. Поворозник А. И. Системы поддержки принятия решений в медицинской диагностике. Синтез структурированных моделей и решающих правил. Saarbrücken Germany: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2011. 314 с.
2. Аврунін О.Г. Сучасні інтелектуальні технології функціональної медичної діагностики: монографія / О.Г. Аврунін, Є.В. Бодяньський, М.В. Калашник, В.В. Семенець, В.О. Філатов. – Харків: ХНУРЕ, 2018. – 236 с.
3. Тимчик С.В., Класифікація медичних інформаційних систем і технологій за інтегральним сукупним критерієм / С.В. Тимчик, С.М. Злепко, С.В. Костішин // Системи обробки інформації – 2016 – 3 (140) – С. 194-198.
4. Yang Y. T., Iqbal U., Horn-Yu Ching J. et. "Trends in the growth of literature of telemedicine: A bibliometric analysis". *Computer Methods and Programs in Biomedicine*. 2015; Vol. 122 (3): 471–479. DOI: 10.1016/j.cmpb.2015.09.008
5. Весненко А.И., Попов А.А., Проненко М.И. Топо-типология структуры развернутого клинического диагноза в современных медицинских информационных системах и технологиях. *Кибернетика и системный анализ*. 2002. № 6. С. 143-154.
6. Кобринский Б.А. Системы поддержки принятия решений в здравоохранении и обучении / Б.А. Кобринский // *Врач и информационные технологии*, 2010, 2: 39–45.
7. Zhukovskaya O.A., Fainzilberg L.S. "Bayesian Strategy for Group Decision Making and its Interval Generalization". *Journal of Automation and Information Sciences*. 2019; Vol. 51, Issue 1: 1-14. DOI: 10.1615/JAutomatInfScien.v51.i1.10
8. Дюк В.А. Компьютерная психодиагностика / В.А. Дюк – СПб.: Братство, 1994. – 364 с.
9. Innocent P.R., John R.I., Garibaldi J.M. "Fuzzy Methods for Medical Diagnosis". *Applied Artificial Intelligence*. 2004; Vol. 19, Issue 1: 69-98. DOI: 10.1080/08839510590887414
10. Das S., Sanyal M.K. "Machine intelligent diagnostic system (MIDS): an instance of medical diagnosis of tuberculosis". *Neural Comput & Applic.* 2020; Vol. 32: 15585–15595. <https://doi.org/10.1007/s00521-020-04894-8>
11. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. К.: Наук. думка, 2004. 545 с.
12. Бородулин В.И. Тополянский А.В. Справочник практического врача в 2-х книгах. Книга 1 М.: Оникс; Мир и Образование, 2007. — 752 с.

Received (Надійшла) 19.11.2020

Accepted for publication (Прийнята до друку) 03.02.2021

Synthesis of a combined diagnostic decision rule in a medical decision support system

Anatoly Povoroznyuk, Oksana Povoroznyuk, Khaled Shekhna

Abstract. The work is devoted to solving the urgent scientific and technical problem of building a decision support system based on the implementation of the developed model of the diagnostic decision rule by means of modern information technologies, the use of which made it possible to ensure the operability of the system on various hardware platforms under the control of various operating systems. Based on the analysis of the methods which are used to construct decision rules in decision support systems, the components of the combined decision rule are proposed, expressing two approaches to the formulation of a diagnostic conclusion: objective, based on the analysis of the training sample, and subjective, based on expert information about the structure of symptom complexes. **The aim** of the study is to synthesize a combined decision rule based on the method of comparison with a prototype, which would take into account both the objective and subjective components of the diagnosis process. **Results.** It was developed a mathematical model of the combined diagnostic decision rule and was substantiated the choice of its components in the work. The method of comparison with the prototype, in which the diagnosed states are represented by their prototypes in the feature space, was chosen as an objective component. The expert information on the structure of symptom complexes is formalized by presenting the symptom complex of diseases with numerical intervals of linguistic variables. Variants of accounting of expert assessments on the structure of symptom complexes when calculating the coordinates of class prototypes are considered. Requirements for the functionality of the system are formulated, design tools, the main development platform (Java), and the database management system (MySQL) are defined. The design of a decision support system and a comprehensive check of the developed system on real medical data were carried out, which confirmed the efficiency of the system.

Keywords: medical diagnostics, decision rule, method of comparison with the prototype, symptom complex, expert information, decision support system.