

А.В. ТЕЛІШЕВСЬКА, асистент, ЧФ НТУ "ХПІ", Чернівці,
А.І. ПОВОРОЗНЮК, к.т.н., доц. НТУ "ХПІ", Харків

ФОРМАЛІЗАЦІЯ ВХІДНОЇ ІНФОРМАЦІЇ ДЛЯ ДІАГНОСТИКИ НЕВРОЛОГІЧНИХ ЗАХВОРЮВАНЬ

В роботі розглядається проблема постановки діагнозу неврологічного захворювання. В якості вхідних ознак розглянуті лабораторні дані і нейровізуальні дослідження, а також неврологічний статус. В результаті побудована множина, що містить підмножини ознак по лабораторним та клінічним дослідженням, які проведені в Кам'янець-Подільській міській лікарні № 1. Табл.: 1. Бібліогр.: 8 назв.

Ключові слова: діагностика, ознака, формалізація, неврологічні захворювання, клінічні дослідження.

Постановка проблеми та аналіз літератури. За останні десять років відмічається значне зростання неврологічних захворювань (НЗ) в Україні: захворюваність на нервові хвороби, а також поширеність їх збільшились майже вдвічі [1]. Інсульт, або, як його називали раніше, мозковий удар, обумовлений раптовим припиненням кровопостачання частини головного мозку або крововиливом у порожнину черепу та стисненням тканини мозку кров'ю, що вилилась. Обидві ці причини призводять до припинення функціонування та до загибелі клітин мозку в ушкодженій ділянці, що призводить до порушення або втрати функції тих частин тіла, якими вони керують. В Україні захворюваність на інсульт складає приблизно 120 тис. випадків на рік. Серед наших співвітчизників, які перенесли інсульт, відносна кількість загиблих або таких що залишилися інвалідами, суттєво вища, ніж в цивілізованих країнах. На відміну від розвинених країн, в Україні інсульт як причина смерті посідає не третє, а друге місце, поступаючись лише захворюванням серця. Витрати на лікування хворих на інсульт або його наслідки в Україні набагато нижчі, ніж в розвинених країнах, та й навіть в багатьох країнах, що розвиваються. На сьогоднішній день є чітке уявлення про НЗ [2, 3]. Вони зумовлені навколишніми чинниками, такими, як забрудненість навколишнього середовища та шкідливі звички, та внутрішніми чинниками: високим кров'яним тиском, холестерином, тромбозом та ін. Тому для діагностики необхідні дослідження, що дозволяють оцінити збереженість чи порушення кровоплину та визначити місця крововиливів (ультразвукове дослідження сонних артерій, ангіографія – рентгенологічне обстеження судин). Дослідження крові для визначення порушень здатності крові до згортання, що сприяють чи утворенню тромбів, чи кровоточивості. Електрокардіографія (ЕКГ) або ультразвукове дослідження серця (ехокардіографія) для

виявлення серцевих джерел тромбів, які можуть, відірвавшись, пересуватися в судини головного мозку [4]. Таким чином, для діагностики НЗ необхідно виконати аналіз різнорідних діагностичних ознак (дихотомічних, числових та візуальних), тому необхідно формалізувати множину діагностичних ознак. При цьому необхідно вирішити задачі формального опису ознак [5, 6] та відновлення пропущених даних [7, 8].

Ціллю статті є аналіз та формалізація вхідних даних для діагностування неврологічних захворювань.

Формалізація вхідних даних. Вхідні дані для діагностування НЗ представлені Кам'янець-Подільською міською лікарнею № 1. Для аналізу були відібрані 200 пацієнтів з підозрою на НЗ. В ході комплексного обстеження встановлені неврологічні порушення, а саме геморагічний інсульт, ішемічний інсульт, епілепсія, остеохондроз, енцефаліт та мігрень. Особливістю вхідних даних являється наявність великої кількості різнорідної інформації. При цьому багато ознак мають описовий характер.

Нехай кожен пацієнт представляє собою об'єкт ω_i ($i = \overline{1, N}$, N – кількість пацієнтів) в багатовимірному просторі ознак. Простір ознак являє собою множину X . Із елементів простору формується вектор ознак $\vec{x} = (x_1, x_2, \dots, x_m)$. Отже, кожен об'єкт ω_i в просторі ознак описується вектором $\vec{x}_i^0 = (x_{i1}, x_{i2}, \dots, x_{im})$. Задачею діагностики є віднесення пацієнта ω_i до одного з формалізованих станів НЗ D_i^j ($i = \overline{1, 6}$), де D_1^1 – ішемічний інсульт, D_1^2 – геморагічний інсульт, D_1^3 – епілепсія, D_1^4 – остеохондроз, D_1^5 – енцефаліт, D_1^6 – мігрень.

Далі необхідно розбити вхідну множину ознак X на підмножини X_k , які не перетинаються, так, щоб $\bigcup_{i=1}^k X_i = X$, $X_j \cap X_i = \emptyset$, $j, i = \overline{1, k}$, $j \neq i$. В результаті були виділені наступні підмножини: X_1 – неврологічний статус, X_2 – лабораторні дослідження та X_3 – нейровізуальні дослідження. Кожна підмножина, в свою чергу, розбивається на підмножини X_k^l . Підмножина X_1 розбита на підмножини: X_1^1 – черепномозкові нерви (12 пар), X_1^2 – рухова сфера (5 ознак), X_1^3 – менінгальні контрактури (2 ознаки), X_1^4 – чутливість

(7 ознак), X_1^5 – сухожилкові та періостальні рефлекси (5 ознак), X_1^6 – патологічні ознаки (6 ознак), X_1^7 – координація рухів (8 ознак), X_1^8 – вегетативна нервова система (6 ознак). Підмножина X_2 розбита на підмножини: X_2^1 – аналіз крові на глюкозу, X_2^2 – загальний аналіз сечі, X_2^3 – загальний аналіз крові, X_2^4 – печінкові проби, X_2^5 – копрограма, X_2^6 – аналіз сечі на амілазу, X_2^7 – аналіз сечі на діастазу, X_2^8 – аналіз сечі на глюкозу. Підмножина X_3 , в свою чергу, розбита на підмножини: X_3^1 – електрокардіограма, X_3^2 – ультразвукова діагностика, X_3^3 – ехоенцефалоскопія, X_3^4 – магніто-резонансна томографія головного мозку, X_3^5 – комп’ютерна томографія.

Ознаки, які входять в підмножину X_1^i ($i = \overline{1,8}$), вимірюються в дихотомічній шкалі, тому 1 – присутність ознаки, 0 – відсутність. Значення показників підмножини X_2^i ($i = \overline{1,8}$) вимірюються в кількісній шкалі. Значення показників підмножини X_3^i ($i = \overline{1,5}$) мають візуальне представлення. Таким чином, в якості експериментальних даних для діагностики НЗ необхідно сформувати таблицю наступної структури:

Таблиця. Експериментальні дані

X	X_1						X_2			X_3			
	X_1^{1j}			...	X_1^{8j}			X_2^{1j}	...	X_2^{8j}	X_3^{1j}	...	X_3^{5j}
	$x_1^{1j,1}$...	$x_1^{1j,12}$...	$x_1^{8j,1}$...	$x_1^{8j,6}$						
1	$x_1^{11,1}$...	$x_1^{11,12}$...	$x_1^{11,1}$...	$x_1^{11,1}$	x_2^{11}	...	x_2^{81}	x_3^{11}	...	x_3^{51}
2	$x_1^{12,1}$...	$x_1^{12,12}$...	$x_1^{12,1}$...	$x_1^{12,1}$	x_2^{12}	...	x_2^{82}	x_3^{12}	...	x_3^{52}
3	$x_1^{13,1}$...	$x_1^{13,12}$...	$x_1^{13,1}$...	$x_1^{13,1}$	x_2^{13}	...	x_2^{83}	x_3^{13}	...	x_3^{53}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
j	$x_1^{1j,1}$...	$x_1^{1j,12}$...	$x_1^{8j,1}$...	$x_1^{8j,6}$	x_2^{1j}	...	x_2^{8j}	x_3^{1j}	...	x_3^{5j}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
N	$x_1^{1n,1}$...	$x_1^{1n,12}$...	$x_1^{8n,1}$...	$x_1^{8n,6}$	x_2^{1n}	...	x_2^{8n}	x_3^{1n}	...	x_3^{5n}

Для комп'ютерної обробки експериментальних даних необхідно, щоб всі ознаки x_k^{lj} були виражені в числовому представленні. Отже множина X_1 представлена в дихотомічній шкалі, то x_1^{lj} будуть набувати значень 1 або 0, тобто присутність ознаки або відсутність. Множина X_2 представлена в номінальній шкалі, тобто ознаки x_2^{lj} мають цифрове представлення, яке відповідає нормам аналізів пацієнтів. Оскільки множина X_2 має числові дані, то необхідно виконати центрування і нормування. Першим етапом, як правило є центрування – знаходження точки середнього значення всіх ознак – геометричного центра багатовимірної множини точок даних. Зазвичай зручно зсунути всі точки даних на один і той же вектор таким чином, щоб центр множини опинився на початку координат. Наступний етап – це нормування – тобто ділення всіх значень ознак на певне число таким чином, щоб значення ознак потрапляли в подібні по величині інтервали. В якості такого числа, зазвичай вибирають одну із характерних відстаней. В багатовимірній множині існує декілька відстаней. Перша – це середньоквадратичне відхилення:

$$\sigma = \sqrt{\frac{1}{N} \sum_{l=1}^N (X_2^l - \bar{X})^2},$$

де $\bar{X} = \frac{1}{N} \sum_{l=1}^N X_2^l$ (X_2^l – вектор даних, x_2^{lj} – j -та координата l -го вектора).

У випадку, якщо вибірка може вважатися отриманою із нормального розподілу, то в колі з центром в \bar{x} та радіусом σ знаходиться близько двох третин від числа точок даних. Існує відстань, яка характеризує максимальне розсіювання в множині даних

$$R = \max_{l=1, N} \|X_2^l - \bar{X}\|.$$

Нормування всіх ознак на R призводить до того, що вся множина даних поміщена в коло одиничного радіусу.

Якщо в якості відстані вибрані σ або R , то відповідні формули обробки (нормування на "одиничну дисперсію" і "на одиничне коло") мають вигляд:

$$\tilde{X}_2^l = \frac{X_2^l - \bar{X}}{\sigma}, \quad \tilde{X}_2^l = \frac{X_2^l - \bar{X}}{R},$$

де \tilde{X}_2^l – новий вектор ознак, X_2^l – старий вектор ознак.

Крім того, якщо діапазони значень для різних ознак сильно відрізняються один від одного, то краще для кожної з ознак застосувати власний масштаб. Тобто, для кожної з ознак можна ввести своє середньоквадратичне відхилення та розсіювання:

$$\sigma_j = \sqrt{\frac{1}{N} \sum_{l=1}^N (x_2^{lj} - \bar{x}_j)^2}, \quad R_j = \max_{l=1, N} \|x_2^{lj} - \bar{x}_j\|,$$

де $\bar{x}_j = \frac{1}{N} \sum_{l=1}^N x_2^{lj}$ (x_2^{lj} – значення j -ої ознаки на l -му векторі).

Як результат отримуємо формули для нормування на "одичинчу дисперсію для кожної ознаки" і "на одичинчний куб":

$$\tilde{x}_2^{lj} = \frac{x_2^{lj} - \bar{x}_j}{\sigma_j}, \quad \tilde{x}_2^{lj} = \frac{x_2^{lj} - \bar{x}_j}{R_j}.$$

Оскільки, множина X_3 має візуальне представлення, тому для подальшої роботи ознаки x_3^{lj} необхідно перевести у числове представлення.

Висновки. В даній роботі виконані етапи збору, формалізації і попереднього аналізу вхідних даних та вихідних діагностуємих станів. Побудовані множини, що містять підмножини ознак неврологічного статусу та клінічних досліджень з метою подальшої розробки інформаційної структури медичної бази даних.

Список літератури. 1. *Волошин П.В.* Епідеміологія мозкового інсульту в Україні / *П.В. Волошин, Т.С. Міценко, І.В. Здесенко* // Матеріали науково-практичної конференції з міжнародною участю "Діагностика, лікування, профілактика гострих та хронічних порушень мозкового кровообігу". – Харків. – 2005. – С. 74. **2.** *Міценко Т.С.* Стан та перспективи розвитку неврологічної служби в Україні / *Т.С. Міценко* // Матеріали науково-практичної конференції "Фармакотерапія захворювань нервової системи". – Харків. – 2005. – С. 167. **3.** *Віничук С.М.* Нервові хвороби. / *С.М. Віничук, Є.Г. Дубенко, Є.Л. Мачерет.* За ред. С.М. Віничука, Є.Г. Дубенка. – К.: Здоров'я, 2001. – 696 с. **4.** *Кареліна Т.І.* Медсестринство в неврології. Підручник / *Т.І. Кареліна, Н.М. Касевич.* — К.: ВСВ "Медицина", 2010. – 296 с. **5.** *Орлов А.И.* Прикладная статистика. Учебник / *А.И. Орлов.* – М.: Издательство "Экзамен", 2004. – 656 с. **6.** *Джоррратано Джозеф.* Экспертные системы: принципы разработки и программирование / *Джозеф Джоррратано, Гари Райли.* – М.: ООО "И. Д. Вильямс", 2007. – 1152 с. **7.** *Айвазян С.А.* Классификация многомерных наблюдений / *С.А. Айвазян, З.И. Бежаева, О.В. Староверов.* – М.: Статистика, 1974. – 240 с. **8.** *Дейвисон М.* Многомерное шкалирование: Методы наглядного представления данных / *М. Дейвисон.* – М.: Финансы и статистика, 1988.

Статья представлена д.т.н., с.н.с. ДУЭП Тараненко Ю.К.

УДК 681.513:620.1

Формализация входной информации для диагностики неврологических заболеваний / Телишевская А.В., Поворознюк А.И. // Вестник НТУ "ХПИ". Тематический выпуск: Информатика и моделирование. — Харьков: НТУ "ХПИ". — 2011. — № 17. — С. 162 – 167.

В работе рассматривается проблема постановки диагноза неврологического заболевания. В качестве входных признаков рассмотрены данные лабораторных и нейровизуальных исследований, а также неврологический статус. В результате построено множество, которое содержит подмножества признаков по лабораторным и клиническим исследованиям, которые проведены в Каменец-Подольской городской больнице № 1. Табл.: 1. Библиогр.: 8 назв.

Ключевые слова: диагностика, признак, формализация, неврологические заболевания, клинические исследования.

UDC 681.513:620.1

Formalization of entrance information is for diagnostics of neurological diseases / Telishevskaya A.V., Povoroznyuk A.I. // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkov: NTU "KhPI". – 2011. – №. 17. – P. 162 – 167.

The problem of raising of diagnosis of neurological disease is considered in work. As source signs were considered the laboratory and neurovisual research, and also neurological status. As a result built plural, which contains subsets of signs for by laboratory and clinical research, what are conducted in Kamenech-Podol'skoy a city hospital № 1. Tabl.: 1. Refs.: 8 titles.

Key words: diagnostics, sign, formalization, neurological diseases, clinical researches.

Поступила в редакцию 29.03.2011