УДК 519.226:519.816

О.В. ЗАКОРЧЕНИЙ, М.М. БУДНИК

ЗАГАЛЬНИЙ РОЗВ'ЯЗОК ЗАДАЧІ КЛАСИФІКАЦІЇ У ПРОСТОРІ ЙМОВІРНОСТЕЙ

O.V. Zakorchenyi, M.M. Budnyk

GENERAL SOLUTION OF CLASSIFICATION PROBLEM AT PROBABILITY SPACE

The probability approach for set classifying is developed. Threshold and triple decision rules are determined in probability space based on link between set theory and Boolean algebra.
Key words: classification, decision rule, probability.

Развит вероятностный подход к классификации множеств. Получены пороговое и троичное решающие правила в пространстве вероятностей на основе связи между теорией множеств и булевой алгеброй.
Ключеые слова: классификация, решающее правило, вероятность.

Розвинуто ймовірнісний підхід до класифікації множин. Отримані порогове і трійкове вирішувані правила у просторі ймовірностей на основі зв'язку між теорією множин і булевою алгеброю.
Ключові слова: класифікація, вирішуване правило, ймовірність.

Вступ. Автори дотримуються угоди згідно якої класи – це множини, що не вміщують підмножин і тому не можуть перекриватися з іншими множинами, тобто поняття «класи» і «класи, що не перекриваються» тотожні [1]. У медичній інформатиці проблема класифікації чи дискримінації полягає у розбитті множини (групи осіб) на декілька класів (патологій, діагнозів). При цьому, якщо кількість класів апріорі відома, то задача розбиття зводиться до задачі дискримінації, якщо невідома – до задачі класифікації [2]. Далі, якщо це не буде оговорено окремо, під класифікацією розуміємо задачу дискримінації. Обидві задачі по суті є оберненими і вимагають синтезу простору ознак та прикладного алгоритму розбиття – вирішуваного правила (ВП).

Відмінність методів дискримінації полягає у тому, що вони використовують апріорну інформацію про класи, а тому за допомогою процедури навчання їх можна оптимізувати, водночас як методи класифікації ґрунтуються на самонавчанні (кластерний аналіз) [1]. Оптимізація здійснюється на основі функціоналів (критеріїв) якості, які в найбільш простому двокритеріальному випадку мають сенс максимум ефекту при однакових затратах, мінімум затрат при однаковому ефекті або узагальнених критеріїв типу "ефект/затрати" [3].

Найпростіше порогове правило дає змогу розбити групу на два класи. При цьому мають місце помилки 1-го (пропуску цілі) та 2-го (хибної тривоги) роду. Цю задачу розглядає статистична теорія рішень (СТР), де оптимальне (найбільш потужне) правило згідно критерію Неймана – Пірсона [4] при заданому рівню помилок 1-го роду мінімізує помилки 2-го роду. В теоретичному плані зазначений критерій дає можливість розрахувати поріг (критичне значення параметра, що розділяє класи) у наближенні нормального розподілу. Однак цей підхід мало корисний на практиці, бо величини помилок і поріг залежать від емпіричних диференційних функцій розподілу (ФР), які, як правило, є негаусовими.

Іншим широко розповсюдженим методом вирішення задачі дискримінації є баєсівський підхід, який дозволяє визначити поріг та мінімізувати втрати від помилкової класифікації [5]. Проте його недолік полягає у необхідності знання повної апріорної інформації – ФР, апріорних та умовних ймовірностей, що вимагає досить великих затрат на отримання вхідної інформації [6].

Іншим недоліком СТР на основі порогового ВП є відсутність зони сумніву (ЗС), яка існує у загальному підході [2] та еквівалентна переходу від бінарної до трійкової логіки. Наприклад, людина – дуже складна біологічна система, тому застосування порогового ВП є великою примітивізацією і не може бути прийняте як адекватна модель організму людини. Причина в наявності «фізіологічного коридору», який приводить до градацій типу «дефіцит – норма – надлишок» (тиск крові) або «норма – слабкий ризик – суттєвий ризик» (холестерин) [7].

Постановка задачі. Отже, актуальним є розробка ВП на основі ЗС та багатозначних логік, що надасть змогу сформулювати детальні висновки про стан людини. При цьому, на думку авторів, при збільшенні “*N*-значності” логіки вона своєю континуальною границею має нечітку логіку (НЛ). Отже шлях вдосконалення ВП полягає в імплементацію у СТР методів НЛ (функції належності (ФН) замість ФР), особливо для малих вибірок [8]. Надалі класифікацію на основі порогового ВП назвемо чітким діагнозом, на основі більш складного трійкового ВП – напівчітким діагнозом, а на основі НЛ – нечітким діагнозом.

Теоретичним підґрунтям до введення ЗС автори вважають класичну теорію ймовірностей (ТЙ) [9]. Шлях до вдосконалення ВП полягає у строгому визначенні класів, які мають складатися з подій, що взаємно виключають одна одну – тобто кожний елемент множини може належати до негативного, позитивного чи проміжного класу. Згідно геометричної інтерпретації ТЙ класам відповідають несумісні події.

Отже завдання полягає у тому, щоб знайти несумісні події, які в принципі мають місце при розділенні множини на класи. У разі якщо ця множина складається з двох підмножин A та B, такий підхід має дати значення класів, на які можна розділити множину, а також ймовірності цих класів для оцінки конкретних алгоритмів класифікації.

З іншого боку, проблема класифікації є частковим випадком розпізнавання образів (РО). Відомо, що ефективність алгоритмів РО в першу чергу залежить від структури простору, в якому воно здійснюється [10]. Тому конструктивний підхід до розв’язання проблеми класифікації автори вбачають у першу чергу не на шляху розробки найбільш потужних ВП, а на шляху синтезу відповідного вибіркового (параметричного) простору. Проте, спершу потрібно побудувати загальний розв'язок задачі класифікації в абстрактному просторі ймовірностей, іншими словами – надати теоретико-множинну інтерпретацію порогового ВП в рамках класичної ТЙ та булевої алгебри, що на сьогодні невідомо з літератури.

Зв'язок між двома сумісними подіями. З точки зору ТЙ випадкова подія полягає у віднесенні елемента множини до певного класу – A чи B. Відомо, що ймовірності суми $P(A+B)$ сумісних подій $P(A)$ та $P(B)$ зв'язані згідно [9]

$$P(A+B) = P(A) + P(B) - P(AB). \tag{1}$$

Водночас, геометрична інтерпретація добутку подій $P(AB)$ діаграмами Ейлера (рис. 1) дає таку залежність від площ областей подій S_A, S_B та їх перетину S_0:

$$P(AB) = \frac{S_0}{S_A + S_B}. \tag{2}$$

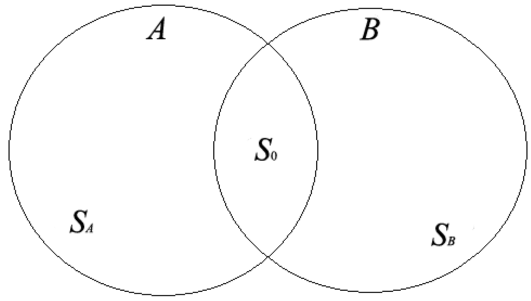

РИС. 1. Діаграма Ейлера для двох сумісних подій

Легко встановити, що мінімум S_0 залежить від співвідношення між площами областей подій A та B і має вигляд

$$S_0 = \min\{S_A, S_B\}. \tag{3}$$

Порівнюючи (2, 3) встановимо, що ймовірність добутку подій обмежена зверху

$$P(AB) \le 1/2. \tag{4}$$

З іншого боку відомі рівняння

$$P(AB) = P(A/B) * P(B),\ P(BA) = P(B/A) * P(A), \tag{5}$$

з яких не слідує ніяких обмежень на $P(AB)$, бо формально $P(AB) \le 1$; $P(BA) \le 1$; $P(A) \le 1$; $P(B) \le 1$.

Виявлене протиріччя дозволяє припустити, що існують невідомі зв'язки між $P(AB)$, $P(A/B)$, $P(B/A)$, $P(A)$ та $P(B)$, які забезпечують виконання умови (4). Для їх встановлення відмітимо, що при перекритті областей подій A та B, як частина площі області AS_{0A} перекривається з областю B, так і частина області BS_{0B} перекривається з областю A. $S_{0A} = S_{0B} = S_0$, що і забезпечує рівність

$P(AB) = P(BA)$. Однак перекриваються дві області, отже площа перекриття вдвічі більша за S_0

$$S_\Sigma = S_{0A} + S_{0B} = 2S_0. \tag{6}$$

Отже, сумарна ймовірність одночасного настання подій A та B має вигляд

$$P_\Sigma(AB) = P(AB) + P(BA) = 2P(AB) = 2P(BA). \tag{7}$$

Водночас в (1) входить тільки половина зазначеної ймовірності, причому немає значення – це $P(AB)$ чи $P(BA)$. Тому (1) не рівна ймовірності «настання події A чи B, виключаючи можливість їх одночасного настання».

Іншими словами, виразу (1) фактично відповідає диз'юнкція (функція АБО) $P(A \vee B)$ (при цьому згідно Уїттлу [9] позначення $P(A+B)$ залишимо тільки для несумісних подій). Водночас згідно відомого підходу до аксіоматики ТЙ на основі нормованої булевої алгебри [11] диз'юнкція має виражатися через функцію «Виключне АБО» (симетрична різниця). Останню для забезпечення відмінності від $P(A \vee B)$ позначимо, як

$$P(A \oplus B) = P(A \vee B) - P(BA), \tag{8}$$

де $\oplus$ – взятий по аналогії знак операції «Сума по модулю 2». Комбінуючи (1) та (7) отримаємо кінцевий вираз для ймовірності $\oplus$ двох сумісних подій:

$$P(A \oplus B) = P(A) + P(B) - P_\Sigma(AB). \tag{9}$$

Продовжуючи далі аналогію з булевою алгеброю ймовірність (7) необхідно розуміти як ймовірність логічного добутку подій A та B, тобто кон'юнкції $P_\Sigma = P(A \wedge B)$. Тоді вираз (9) можна подати у вигляді

$$P(A) + P(B) = P(A \oplus B) + P(A \wedge B), \tag{10}$$

який формально відповідає булевому виразу

$$A \vee B = (A \oplus B) \vee (A \wedge B). \tag{11}$$

У теоретичному плані вирази (10), (11) приводять у відповідність значення функцій булевої алгебри та їх розуміння в ТЙ на основі геометричної (теоретико-множинної) інтерпретації. Принципова відмінність (9), (10), (11) від (1) полягає у тому, що здійснено перехід від двох сумісних подій A та B до двох несумісних подій $A \oplus B$ та $A \wedge B$, еквівалентних згідно сумарної ймовірності. Прикладне значення (10), (11) полягає у тому, що несумісним подіям в абстрактному просторі відповідають області (класи) у вибірковому просторі, які не перекриваються.

Задача дискримінації полягає у розбитті множини подій $A \vee B$ на два класи, тобто на множини, що не перетинаються. З точки зору (10) множини $A \oplus B$ та $A \wedge B$ і тільки вони є такими класами, а отже моделюють результат задачі дискримінації в теоретико-множинному просторі. Для теорії дискримінації вираз (10) дає вирішення задачі класифікації у просторі ймовірностей, що вказує на:

1) які класи необхідно розбивати множину $A \vee B$,

2) метод оцінки алгоритму дискримінації, а саме – чим ближче ймовірність належності до класів наближається до величин апріорних ймовірностей

$P(A \oplus B)$ та $P(A \wedge B)$ – тим ближче даний алгоритм до ідеального.

Звичайно, (10) не може дати прикладного методу для синтезу «ідеального» алгоритму, – він, як і медичне значення введених класів, буде поданий далі.

Проте, в прикладному аспекті не менш важливим є також встановлення ймовірності різниці двох подій, що відповідають різницям площ ($A \setminus B \leftrightarrow \mathrm{S}_A - \mathrm{S}_0$ та $B \setminus A \leftrightarrow \mathrm{S}_B - \mathrm{S}_0$). Враховуючи співвідношення

$$P(A \oplus B) = P(A/B) + P(B/A), \tag{12}$$

а також (1), (7), (8), отримаємо вирази для різниці двох сумісних подій:

$$P(A \setminus B) = P(A) - P(AB); \;\; P(B \setminus A) = P(B) - P(BA). \tag{13}$$

Враховуючи також, що $A \setminus B = A \setminus AB$, та $B \setminus A = B \setminus BA$, з (13) отримаємо вирази

$$P(A \setminus AB) = P(A) - P(AB); \;\; P(B \setminus BA) = P(B) - P(BA), \tag{14}$$

які показують, що $P(AB)$ та $P(BA)$ не є тотожними (операція добутку некомутативна) і замінювати їх в (14) одна на одну принципово не можна. Це видно також з умови, що ймовірності не можуть бути від'ємні $P(A \setminus B) > 0$, $P(B \setminus A) > 0$, звідки отримаємо крім (4) ще обмеження зверху (14):

$$P(AB) < P(A) \text{ (бо } \mathrm{S}_{A0} < \mathrm{S}_A\text{)}; \; P(BA) < P(B) \text{ (бо } \mathrm{S}_{0B} < \mathrm{S}_B\text{)}. \tag{15}$$

Вирази (14) є важливими для теорії класифікації, де актуальним є встановлення ймовірності «настання однієї події, виключаючи можливість одночасного настання двох подій», або, що еквівалентно, «настання однієї події, виключаючи іншу подію». Таким чином вирази (13) дають «чисті» ймовірності однієї події без «домішок» іншої.

Розглянемо ще «сумарну» ймовірність, що описується виразом

$$P_+ = P(A) + P(B) + P(A \wedge B). \tag{16}$$

Враховуючи, що (1) фактично описує ймовірність диз'юнкції подій, тоді як (16) – «арифметичної» суми P_+ подій $A + B$, яку можна формально записати як

$$P_+(A+B) = P(1) + P(2), \;\; P(1) = P_{\text{П}}(A+B) = P(A) + P(AB),$$

$$P(2) = P_{\text{Л}}(A+B) = P(B) + P(BA), \tag{17}$$

де $P1(2)$ – «права» («ліва») сума подій.

Визначимо операцію, яку можна назвати «симетричною сумою» як

$$P_C(A+B) = \frac{P_+}{2} = \frac{P(A)+P(B)}{2} + \frac{P(A \wedge B)}{2}. \tag{18}$$

Враховуючи (8), вираз (18) можна подати також у вигляді

$$P_C(A+B) = P(A \wedge B) + (1/2)\left[P(A \setminus B) + P(B \setminus A)\right]. \tag{19}$$

Зазначимо, що аналогічно можна ввести «симетричну різницю» згідно

$$P_C(A-B) = \frac{P(A)+P(B)}{2} - \frac{P(A \wedge B)}{2}. \tag{20}$$

Тоді обидві симетричні операції алгебри подій можна записати у вигляді

$$P_C(A \pm B) = 1/2\left[P(A) + P(B)\right] \pm P(AB) \qquad (21)$$

звідки

$$P_C(A - B) = P(A \oplus B)/2; \qquad P_C(A + B) = P_C(A - B) + P(A \wedge B). \qquad (22)$$

Значення назв симетричних операцій видно з рис. 3, бо при $P(A) = P(B) = 1/2$ ймовірність «симетричної суми» рівна сумі відповідних ймовірностей, а ймовірність «симетричної різниці» – їх різниці. При несиметрії апріорних ймовірностей обидві симетричні операції алгебри подій збігаються до півсуми ймовірностей обох подій $\left[P(A) + P(B)\right]/2.$ Медична інтерпретація цих ймовірностей буде приведено в останньому пункті.

Наочне зображення операцій алгебри сумісних подій. Згідно геометричної інтерпретації на рис. 1, для апріорних та умовних ймовірностей маємо

$$P(A, B) = \frac{S_{A,B}}{S_A + S_B}; \quad P(A/B) = S_0 / S_B; \quad P(B/A) = S_0 / S_A \qquad (23)$$

звідки з урахуванням (3)

$$P(AB) = \min\left\{P(A), P(B)\right\}. \qquad (24)$$

Тоді маємо

$$P(A \vee B) = \max\left\{P(A), P(B)\right\}, \quad P(A) + P(B) = 1 \qquad (25)$$

Залежності ймовірностей різних складних подій (24 – 27) від апріорної ймовірності настання події $P(B)$ в нормованому просторі подій наглядно показано на рис. 2, а (28 – 30) – на рис. 3 (для повного простору нормування непотрібне).

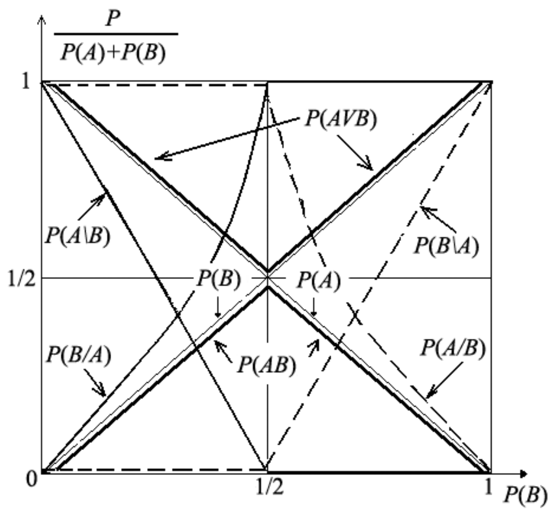

РИС. 2. Залежність добутку, диз'юнкції та різниці двох випадкових подій, а також умовних ймовірностей від апріорної ймовірності події B

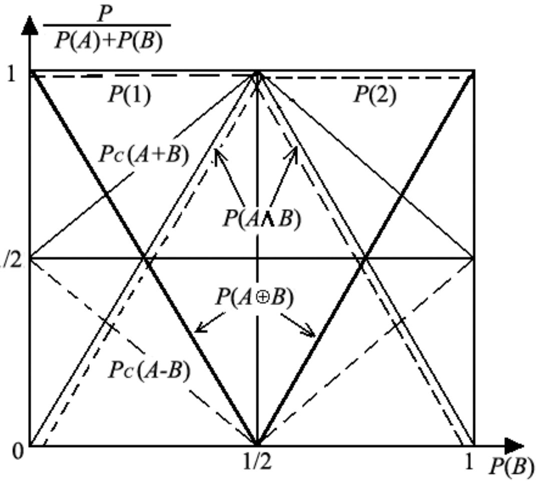

РИС. 3. Залежність правої та лівої сум, симетричної суми та різниці, кон’юнкції та суми по mod 2 двох випадкових подій від ймовірності події B

$$P(A\backslash B)=\begin{cases}0, & P(A)<P(B),\\ P(A)-P(B), & P(A)>P(B),\end{cases}\quad P(B\backslash A)=\begin{cases}P(B)-P(A); & P(A)<P(B),\\ 0, & P(A)>P(B),\end{cases}\qquad(26)$$

$$P(B/A)=\begin{cases}P(B)/P(A), & P(A)<P(B),\\ 1, & P(A)>P(B),\end{cases}\quad P(A/B)=\begin{cases}1, & P(A)<P(B),\\ P(A)/P(B), & P(A)>P(B),\end{cases}\qquad(27)$$

$$P(A\wedge B)=\begin{cases}2P(B), & P(B)<P(A),\\ 2P(A), & P(B)>P(A),\end{cases}\quad P(A\oplus B)=\begin{cases}P(A)-P(B); & P(B)<P(A),\\ P(B)-P(A); & P(B)>P(A).\end{cases}\qquad(28)$$

$$P(1)=\begin{cases}P(A)+P(B), & P(B)<P(A),\\ 2P(A), & P(B)>P(A),\end{cases}\quad P(2)=\begin{cases}2P(B), & P(B)<P(A),\\ P(A)+P(B), & P(B)>P(A),\end{cases}\qquad(29)$$

$$P_C(A+B)=\begin{cases}\left[3P(A)+P(B)\right]/2, & P(B)<P(A),\\ \left[P(A)+3P(B)\right]/2, & P(B)>P(A),\end{cases}\quad P_C(A-B)=\begin{cases}\left[P(A)-P(B)\right]/2, & P(B)<P(A),\\ \left[P(B)-P(A)\right]/2, & P(B)>P(A).\end{cases}\qquad(30)$$

Медична інтерпретація на прикладі двох груп пацієнтів. Нехай $P(B)$ – це апріорна ймовірність класифікації особи до групи хворих, а $P(A)$ – до групи здорових. Тоді $P(B)=\Pi$, де Π – преваленс хворих у групі, а відповідні ймовірності мають наступне значення:

$$P_H(B)=P(B\backslash A);\quad P_H(A)=P(A\backslash B);\quad P_\gamma=P_C(A-B);\qquad(31)$$

$$P_V=P_C(A+B);\qquad P_{\text{СУМНІВ}}=P(A\wedge B)\qquad(32)$$

та подані на рис. 4, де враховано, що простір подій повний, отже

$$P(A)+P(B)=1; \qquad P(A)=1-\Pi. \tag{33}$$

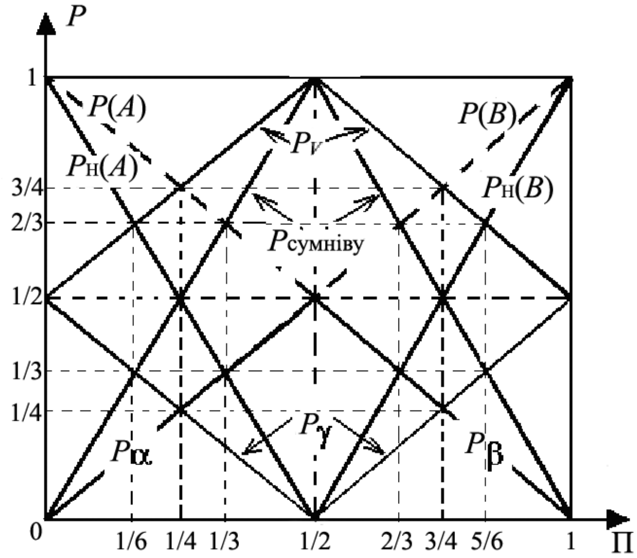

РИС. 4. Залежність операцій алгебри подій від превалeнсу груп

З рис. 4 видно, що $P_H(A)$ – це ймовірність належності тільки до групи хворих, а $P_H(B)$ – до групи здорових. $P_{\text{СУМНІВ}}$ означає ймовірність того, що особа належить одночасно обом групам, тобто це ймовірність належності до групи (зони) сумніву. Для узагальнення введено ймовірності похибок 1-го (P_α) та 2-го роду (P_β) як (таке визначення узгоджується із загальним визначенням (24))

$$P(AB)=\begin{cases} P_\alpha = P(B), & \text{якщо } P(B)<P(A); \\ P_\beta = P(A), & \text{якщо } P(B)>P(A). \end{cases} \tag{34}$$

Тоді ймовірності належності «чистим» групам (тобто класам) рівні

$$P_H(A)=P(A)-P_\alpha; \qquad P_H(B)=P(B)-P_\beta. \tag{35}$$

Таким чином дві групи, що перекриваються, еквівалентні 3-м класам:

1) $A \setminus B$ – клас негативних, тобто «чистих» здорових – здорових за виключенням хворих;

2) $B \setminus A$ – клас позитивних, тобто «чистих» хворих – хворих за виключенням здорових;

3) проміжний клас, тобто осіб із сумнівним станом $A \wedge B$, що включає осіб, які належать до обох груп одночасно.

Величини $P_H(A)$, $P_H(B)$ та $P_{\text{СУМНІВ}}$ належать проміжку $[0,1]$. Отже їх можна вважати функціями належності (ФН). Тоді, використовуючи прийняте у НЛ правило, згідно якого елемент належить до того нечіткого поняття, ФН якого при даному значенні нечіткого параметра (тут параметром є превалєнс Π) максимальна, можна вказати таке ВП для розділення груп

$$\text{ТРИЗНАЧНЕ ПРАВИЛО } (\Pi) = \begin{cases} \text{здоровий, якщо } 0 < \Pi < 1/4, \\ \text{проміжний, якщо } 1/4 < \Pi < 3/4, \\ \text{хворий, якщо } 3/4 < \Pi < 1. \end{cases} \quad (36)$$

Введення зони сумніву (ЗС) $\Pi \in [1/4, 3/4]$ зменшує ймовірності похибок 1-го та 2-го роду, теоретично у 2 рази, за рахунок того, що частина осіб належить до ЗС. Також з рис. 4 видно, що при $\Pi = 1/2$ ймовірності похибок обох родів рівні $P_\alpha = P_\beta$. Цьому оптимальному значенню превалєнсу відповідає найкраща якість розділення груп. Якість класифікації описується величиною P_V, яка лежить в діапазоні $50\,\% < P_V < 100\,\%$ і є загальним показником достовірності діагностичного тесту. Так для порогового ВП, що зазвичай застосовується в СТР

$$\text{ПОРОГОВЕ ПРАВИЛО } (\Pi) = \begin{cases} \text{здоровий, якщо } \Pi < 1/2, \\ \text{хворий, якщо } \Pi > 1/2 \end{cases} \quad (37)$$

показник P_V досягає максимуму. Це значить, що при відсутності інформації про апріорні ймовірності $P(A)$ і $P(B)$ у даній групі (сліпий тест) найбільш ефективне розділення буде досягнуто, якщо прийняти $P(B) = P(A) = 0{,}5$. Тобто ми маємо половину осіб віднести до здорових, а половину – до хворих.

Якщо ж $\Pi = 1(0)$, тобто всі особи в групі хворі (здорові), то зазначене правило приведе до того, що кожна друга особа буде класифікована невірно. Тому при $\Pi \neq 1/2$ показник достовірності тесту спадає і в крайніх точках $\Pi = 0$ та $\Pi = 1$ становить тільки $50\,\%$. Звідси випливає, що довільний алгоритм класифікації (вирішуване правило) математично коректне, якщо забезпечується достовірність тесту (наприклад, специфічність чи чутливість) не менша, ніж $50\,\%$.

Також зазначимо, що в оптимальній точці $\Pi_{\text{ОПТ}} = 1/2$, ймовірності похибок обох родів рівні $P_\alpha = P_\beta$. Це значить, що в ТЙ немає ніяких спеціальних підстав для того, щоб вони не дорівнювали один одному при класифікації у вибірковому просторі. Іншими словами, якщо $P_\alpha \neq P_\beta$, то це означає, що критичне значення $\Pi_{\text{КР}}$, застосоване у правилі (35) неоптимальне (у даному випадку $\Pi_{\text{КР}} \neq 1/2$). З огляду на це постає завдання розробити спосіб класифікації, що забезпечував би відшукання оптимального значення, яке задовольняло умову $P_\alpha = P_\beta$.

Висновок. У стандартній процедурі класифікації поріг обчислюють за допомогою диференціальної ФР. Це призводить до:

а) збільшення кількості показників тесту (специфічність (С), чутливість (Ч), негативна (НП) і позитивна (ПП) прогностичність);

б) нерівності помилок 1-го і 2-го роду;

в) погіршення достовірності.

При цьому два показники достовірності залежать від асиметрії помилок $\Delta\gamma = \alpha - \beta$; $С = С(\Delta\gamma)$; $Ч = Ч(\Delta\gamma)$, а прогностичності – додатково ще й від асиметрії кількості осіб у групах (або що те саме – превеленсу). Це утруднює як обробку, так і, що більш важливо, – інтерпретацію даних. Основний практичний результат роботи полягає у встановленні того факту, що залежність від зазначених факторів є паразитною, вона не випливає з ТЙ, а є недоліком саме способу обробки даних на основі традиційних методів класифікації СТР. Алгоритм, вільний від зазначених недоліків, раніше розроблено і описано в [12].

1. *Глушков В.М.* Основы безбумажной информатики. Изд. 2-е, испр. – М.: Наука, 1987. – 552 с.
2. *Математическая* энциклопедия. Гл. ред. И.М. Виноградов, **2**. – М: Советская энциклопедия, 1979. – 1104 с.
3. *Кветный Р.Н., Маликов В.Т.* Информационная теория измерений: от модели к изделию. – М.: Знание, 1998. – 32 с.
4. *Чистяков В.П.* Курс теории вероятностей. – М.: Наука, 1982. – 256 с.
5. *Файнзильберг Л.С.* Байесова схема принятия коллективных решений в условиях противоречий // Проблемы управления и информатики. – 2002. – № 3. – С. 112–122.
6. *Продеус А.Н., Захрабова Е.Н.* Экспертные системы в медицине. – Киев: ТОВ «ВЕК+», 1998. – 320 с.
7. *Эпидемиология* и факторы риска ишемической болезни сердца / Под ред. А.Н. Климова. – Л.: Медицина, 1989. – 176 с.
8. *Елисеев П.И., Гусаченко Р.П.* Алгоритм обработки результатов анализа данных по малым выборкам // УСиМ. – 2002. – № 6. – С. 18–20.
9. *Уиттл П.* Вероятность. – М.: Мир, 1984. – 640 с.
10. *Джордж Ф.* Основы кибернетики: Пер. с англ. / Под ред. А.Л. Горелика. – М.: Радио и связь, 1984. – 272 с.
11. *Математическая* энциклопедия. Гл. ред. И.М. Виноградов, 5. – М.: Советская энциклопедия, 1984. – 1248 с.
12. *Будник М.М., Закорчений О.В.* Спосіб класифікації групи пацієнтів. Патент UA 84884, заявл. 14.04.2006, опубл. 10.12.2008, Бюлетень «Промислова власність» № 23, 2008.

Отримано 20.10.2010