УДК 681.3.06.14

ПРАВДОПОДОБНЫЕ, НО НЕВЕРНЫЕ РЕШЕНИЯ ПРИ ПОСТРОЕНИИ ДИАГНОСТИЧЕСКИХ ПРАВИЛ

Л.С. Файнзильберг
МНУЦ ИТИС
e-mail: fainzilberg@voliacable.com

При построении систем медицинской диагностики часто приходится сталкиваться с ситуацией, когда существующие знания недостаточны для синтеза решающих правил формальными методами. В таких случаях не остается ничего иного, как полагаться на интуицию и предыдущий опыт экспертов-медиков. Другими словами, конструировать систему на основе «здравого смысла», а затем уже проверять свой замысел на имеющемся клиническом материале, как того требует современные представления о доказательной медицине [1]. Однако не следует забывать крылатую фразу о том, что «наука начинается там, где заканчивается здравый смысл…».

Цель настоящего доклада – привести несколько характерных примеров рассуждений, которые на первый взгляд кажутся правдоподобными, но оказываются необоснованными с научной точки зрения.

1. Построение многокритериальных диагностических правил. Пусть для диагностики некоторого заболевания предполагается использовать результаты измерения показателей (признаков, симптомов, критериев) $x_1,...,x_N$. При этом считается, что на основе предварительного обследования репрезентативной группы практически здоровых людей известны границы медицинской нормы этих признаков [2].

Тогда в простейшем случае диагностическое решение, основанное на измерении *отдельного* признака x_i, $i = 1,...,N$, может быть сведено к пороговому решающему правилу, например, правилу вида[1]

$$\begin{aligned} &\textbf{БОЛЕН, если } x_i \le x_i^0; \\ &\textbf{ЗДОРОВ, если } x_i > x_i^0, \end{aligned} \tag{1}$$

где x_i^0 – граница физиологической нормы признака x_i.

Для повышения достоверности принимаемых решений, по аналогии с (1), часто предлагается строить пороговое решающее правило, основанное на аддитивной линейной свертке признаков:

$$\begin{aligned} &\textbf{БОЛЕН, если } \sum_{i=1}^{N} k_i x_i \le x^0; \\ &\textbf{ЗДОРОВ, если } \sum_{i=1}^{N} k_i x_i > x^0, \end{aligned} \tag{2}$$

в котором коэффициенты k_i определяют «вес» признака в многокритериальном правиле, а x^0 – пороговое значение, найденное на основе предварительно проведенных экспериментов по выборке наблюдений с известными диагнозами.

Однако еще в работе [3] строго доказано, что аддитивная линейная свертка показателей $x_1,...,x_N$ не всегда допустима. Правило (2) можно применять лишь в тех

[1] Разумеется, диагнозы «БОЛЕН» и «ЗДОРОВ» означают всего лишь то, что у обследуемого выявлено или не выявлено рассматриваемое заболевания с определенными значениями чувствительности и специфичности [1].

случае, когда уменьшение значения одного показателя может быть компенсировано увеличением значения другого. В противном случае решения будут неверными.

Этот факт может быть продемонстрирован на следующем шуточном примере. Пусть для диагностики качества телевизора используются два показателя: качество звука и качество изображения, которые тем или иным способом оцениваются интервальными величинами $x_1 \in [0,1]$ и $x_2 \in [0,1]$ соответственно.

Понятно, что можно отдельно диагностировать качество («болезнь») звука и изображения пороговыми правилами, например, так

$$\textbf{ЗВУК} = \begin{cases} \textbf{"БОЛЕН",} & \textbf{если } x_1 \le 0{,}5, \\ \textbf{"ЗДОРОВ",} & \textbf{если } x_1 > 0{,}5, \end{cases} \tag{3}$$

$$\textbf{ИЗОБРАЖЕНИЕ} = \begin{cases} \textbf{"БОЛЕН",} & \textbf{если } x_2 \le 0{,}5, \\ \textbf{"ЗДОРОВ",} & \textbf{если } x_2 > 0{,}5. \end{cases} \tag{4}$$

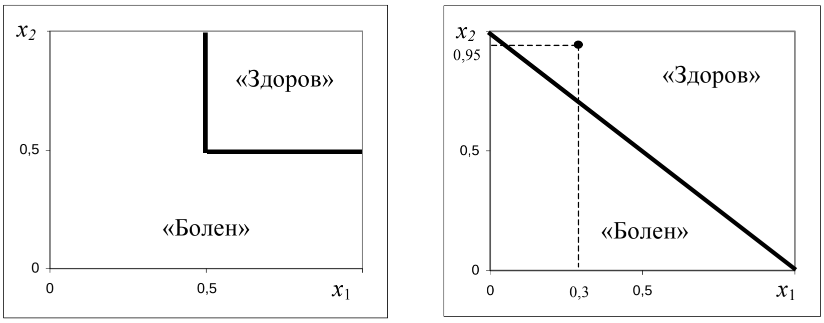
Однако совершенно очевидно, что невозможно компенсировать плохой звук хорошим качеством изображения и наоборот, а значит, область «больного» телевизора не является выпуклой (рис. 1, а). Поэтому использование двухкритериального правила

$$\textbf{ТЕЛЕВИЗОР} = \begin{cases} \textbf{"БОЛЕН",} & \textbf{если } \kappa_1 x_1 + \kappa_2 x_2 \le x^0, \\ \textbf{"ЗДОРОВ",} & \textbf{если } \kappa_1 x_1 + \kappa_2 x_2 > x^0 \end{cases} \tag{5}$$

при любых весах κ_1, k_2 и пороге x^0 лишено всякого смысла.

Например, положив в (5) $k_1 = k_2 = 0{,}5$ и $x^0 = 0{,}5$ приходим к тому, что телевизор с отвратительным звуком ($x_1 = 0{,}3$), но исключительно хорошим изображением ($x_2 = 0{,}95$) будет признан хорошим (рис. 1, б) только лишь потому, что выполняется условие

$$\kappa_1 x_1 + \kappa_2 x_2 = 0{,}625 > 0{,}5\,. \tag{6}$$



а б

Рис. 1. Диагностика качества телевизора по двум признакам

Несмотря на то, что приведенный пример наглядно иллюстрирует необходимость «осторожного» подхода к использованию диагностического правила (2), многие авторы, без должного на то обоснования, предлагают использовать подобные правила. Так в работе [4] для оценки функционального состояния сердца фактически предлагается аддитивная линейная свертка нормированные значений отдельных параметров электрокардиограммы. Однако весьма сомнительна обоснованность такого правила: вряд ли можно «компенсировать» высокую амплитуду зубца T малой продолжительностью интервала QT и наоборот. Список публикаций, в которых необоснованно предлагаются подобные диагностические правила, можно было бы продолжить.

2. Расширение пространства признаков. Рассмотрим теперь случай, когда непересекающиеся области возможных значений признаков в классах V_1 (больные) и V_2 (здоровые) выпуклы, а значит, для диагностики правомерно использовать линейную дискриминантную функцию

$$D(x_1,...,x_N) = k_0 + k_1 x_1 + ... + k_N x_N , \quad (7)$$

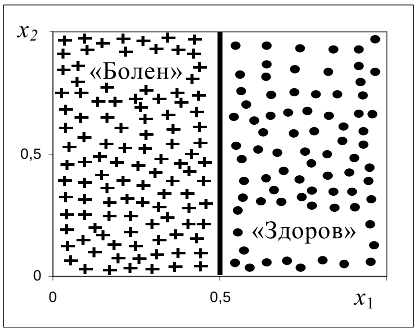
параметры которой k_0, $k_1,...,k_N$ находят на этапе обучения по выборке наблюдений векторов $(x_1,...,x_N)$ с известными принадлежностями к классам V_1 и V_2.

Существует распространенное мнение, что «вредных» признаков не бывает, а могут быть лишь бесполезные признаки. Опираясь на такую аргументацию, многие авторы предлагают [5, с. 62] включить в (7) все доступные признаки, а затем уже на этапе обучения отбросить «лишние» (бесполезные) признаки с малыми значениями весов $|k_i|$, $1 \le i \le N$.

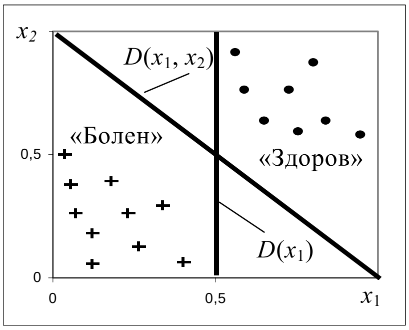
В то же время, если проводить обучение функции (7) по конечной выборке наблюдений, то «лишние» признаки могут ухудшить диагностическое правило [6]. Для иллюстрации этого факта рассмотрим еще один пример, но уже с радиоприемником.

Будем строить решающее правило для диагностики работоспособности радиоприемника по двум показателям: качество звука $x_1 \in [0,1]$ (полезный признак) и цвет его корпуса $x_2 \in [0,1]$ (бесполезный признак).

На рис. 2, а показана граница областей оптимальных решений, которая безошибочно разделяет всех представителей классов V_1 и V_2 генеральной совокупности, а на рис. 2, б – дискриминантная функция $D(x_1, x_2)$, построенная по ограниченной обучающей выборке.



а



б

Рис.2. Одномерные и двумерные дискриминантные функции

Легко видно, что в пространстве двух переменных x_1 и x_2 обучающая выборка оказалась нерепрезентативной: так случилось, что в выборку попали радиоприемники с хорошим звуком, которые имели темный цвет корпуса, а светлые имели плохой звук. Вследствие этого случайного факта построена неудачная дискриминантная функция

$$D(x_1, x_2) = 0{,}5x_1 + 0{,}5x_2 - 0{,}5 , \quad (8)$$

которая хорошо разделяет лишь обучающую выборку, но не разделяет генеральную совокупность представителей линейно разделимых классов V_1 и V_2.

Если бы «лишний» признак x_2 был заранее исключен из обучения, то по этой же выборке было настроено эффективное решающее правило (1), основанное на сравнении значений признака x_1 с порогом $x_1^0 = 0{,}5$, которое позволило бы безошибочно разделять не только ограниченную выборку, но и всю генеральную совокупность наблюдений.

3. Полезная совокупность «неинформативных» признаков. Получила распространение еще одна точке зрения: при построении диагностических правил целесообразно использовать лишь статистически независимые признаки. В качестве аргументации такой точки зрения считается, что зависимый признак мало информативен, так как он не несет дополнительной информации.

Ввиду ограниченного объема доклада не будем углубляться в математические тонкости понятий информативности статистически зависимых признаков, а лишь приведем один любопытный пример, иллюстрирующий несостоятельность таких аргументов.

На рис. 3 в пространстве признаков x_1 и x_2 показаны собственные области двух линейно разделимых классов V_1 и V_2. Каждый в отдельности такой признак неинформативен, поскольку совпадают их условные распределения в классах:

$$p(x_i \mid V_1) \equiv p(x_i \mid V_2),\ i = 1,2. \qquad (9)$$

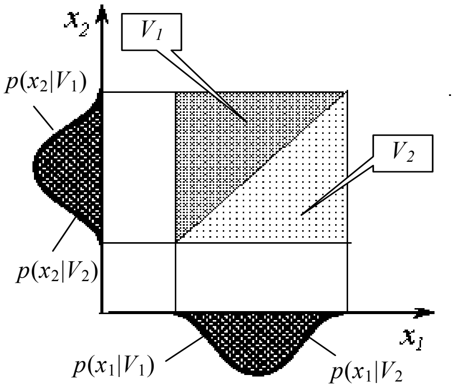


Рис. 3. Диагностика по двум «неинформативным» признакам

Однако легко видно, что в совокупности такие признаки не только полезны, но даже обеспечивают безошибочную диагностику классов V_1 и V_2. В работе [7] доказаны утверждения, из которых следует, что при выполнении (9) такое возможно лишь при наличии статистической связи между признаками x_1 и x_2 в обоих классах, т.е. при выполнении условия

$$p(x_1, x_2 \mid V_k) \not\equiv p(x_1 \mid V_k) p(x_2 \mid V_k), \quad k = 1,2. \qquad (10)$$

Литература

1. Власов В.В. Введение в доказательную медицину. – М.: Медиа Сфера, 2001. – 392 с.
2. Норма в медицинской практике: Справочное пособие. – М.: Медпресс, 2001. – 144 с.
3. Кини Р.Л., Райфа Х. Принятие решений при многих критериях: предпочтения и замещения. – М.: Радио и связь, 1981. – 560 с.
4. Чайковський І.А., Буднік М.М. Спосіб оцінки функціонального стану серця на основі аналізу форми електрокардіограми та варіабельності ритму серця / Патент на корисну модель № 61285 (Україна), опубл. 11.07.2011, Бюл. № 13, 2011 р.
5. Головкин Б.А. Машинное распознавание и линейное программирование. – М.: Сов. радио, 1973. – 100 с.
6. Загоруйко Н.В. Методы распознавания и их применение. – М.: Советское радио. – 1972. – 208 с.
7. Файнзильберг Л.С. Математические методы оценки полезности диагностических признаков. – Киев: Освита Украины, 2010. – 152 с.