

29. Bumping Decision Tree Regression Module. Azure. Machine learning. URL: <https://docs.microsoft.com/ru-ru/azure/machine-learning/algorithm-module-reference/boosted-decision-tree-regression> (accessed: 01.11.2021). (In Russ.)
30. Linear regression module. Azure. Machine learning. URL: <https://docs.microsoft.com/ru-ru/azure/machine-learning/algorithm-module-reference/linear-regression> (accessed: 01.11.2021). (In Russ.)
31. Decision forest regression module. Azure. Machine learning. URL: <https://docs.microsoft.com/ru-ru/azure/machine-learning/algorithm-module-reference/decision-forest-regression> (accessed: 01.11.2021). (In Russ.)

Received 08.11.2021

УДК 681.3.06:681

ВЕРОЯТНОСТНЫЕ КРИТЕРИИ ТЕСТОВ ИНФОРМАЦИОННО- ДИАГНОСТИЧЕСКИХ СИСТЕМ

Лихтциндер Б.Я.¹, Ларина В.А.¹, Муравец А.В.²

¹ Поволжский государственный университет телекоммуникаций и информатики, Самара, РФ

² Самарский областной центр медицинской профилактики

«Центр общественного здоровья», Самара, РФ

E-mail: lixt@psuti.ru

Данная статья посвящена вероятностным критериям тестов информационно-диагностической системы. Рассматриваются вероятностные критерии тестов информационно-диагностической системы, направленной на изучение состояния пациента. Система содержит объект диагностирования (пациент) и средства диагностирования. Система диагностирования реализует некоторый алгоритм диагностирования, задающий последовательность и способ анализа результатов обследования. Целью анализа результатов элементарных проверок является определение диагностируемого состояния, в котором действительно находится пациент. В данной статье рассматриваются такие разделы, как вероятностное диагностирование состояния пациента; матрица условных вероятностей; состояния; шкала состояний; простейшие проверки; условные вероятности. Все данные по каждому разделу сводятся в таблицу для визуализации. Вывод делается на основе анализа, что позволяет оценить эффективность применения данной системы диагностирования.

Ключевые слова: вероятностные критерии, информационно-диагностические системы, диагностирование состояния пациента, обследование

Введение

Автоматизированные медико-диагностические системы находят все большее применение при профилактике и диагностировании заболеваний [1; 2; 4; 8–11].

Система диагностирования состояния пациента содержит объект диагностирования ОД (пациент) и средства диагностирования (СД) [3].

СД реализуют некоторый алгоритм диагностирования, задающий последовательность и способ анализа результатов обследования на основе вероятностных критериев проверок [5–7; 12]. Под элементарной проверкой π_k ($k=1, \dots, n$) будем понимать некоторое тестирующее воздействие, заключающееся в постановке вопроса и анализе ответа на этот вопрос. Под k понимаем последовательное число проверок, прошедших с начала тестирования. (Проверки нумеруются последовательно, по мере их проведения.)

Вероятностное диагностирование состояния пациента

ОД может находиться в одном из диагностируемых состояний e_i на множестве $E = \{e_1, \dots, e_i, \dots, e_S\}$.

Элементами множества E являются состояния e_i ($i=1, 2, \dots, S$). Имеется набор элементарных проверок $\Pi = \{\pi_1, \dots, \pi_2, \dots, \pi_n\}$.

Целью анализа результатов элементарных проверок является определение диагностируемого состояния, в котором действительно находится ОД.

Каждой проверке π_k соответствует f_k исходов этой проверки, образующих множество $Q_k = \{q_1, \dots, q_{j_k}, \dots, q_{f_k}\}$. Исход q_{j_k} может быть равен единице, если проверка показала результат, соответствующий данному исходу, или равен нулю в противном случае. При каждой проверке

Таблица 1. Матрица состояний

Исходы проверки π_k	Состояния						
	e_1	...	e_r	...	e_i	...	e_s
q_{1_k}	$P_k(q_{1_k}/e_1)$...	$P_k(q_{1_k}/e_r)$...	$P_k(q_{1_k}/e_i)$...	$P_k(q_{1_k}/e_s)$
...
q_{j_k}	$P_k(q_{j_k}/e_1)$...	$P_k(q_{j_k}/e_r)$...	$P_k(q_{j_k}/e_i)$...	$P_k(q_{j_k}/e_s)$
...
q_{f_k}	$P_k(q_{0_k}/e_1)$...	$P_k(q_{0_k}/e_r)$...	$P_k(q_{0_k}/e_i)$...	$P_k(q_{0_k}/e_s)$

единичным может оказаться только один из всего множества исходов. Остальные исходы имеют нулевые значения.

Например, если на вопрос проверки π_k может быть дано лишь два ответа: «да» или «нет», то $f_k = 2$. Если при измерении температуры возможны ответы «пониженная», «нормальная», «повышенная», «высокая», значение $f_k = 4$.

Для каждого ОД всегда возможно определить априорные вероятности $P_0(e_i)$ нахождения его в e_i -м состоянии до первой проверки. Очевидно, что система вероятностей должна быть полной, поскольку считается, что ОД непременно находится в одном из перечисленных выше состояний $\sum_{i=1}^S P_0(e_i) = 1$.

Вероятности $P_0(e_i)$ являются априорными и характеризуют степень неопределенности множества состояний ОД до начала диагностирования.

После завершения некоторой проверки π_k возможно получение одного из f_k результатов. Такой результат мы назвали исходом q_{j_k} проверки π_k ($q_{j_k} = 1$).

Если проверка π_k информативна по отношению к множеству состояний ОД, то в результате проверки наши знания о его реальном состоянии возрастут.

Введем следующие обозначения: $P_k(e_i/q_{j_k})$ – вероятность того, что после проверки π_k , с исходом $q_{j_k} = 1$, ОД находится в состоянии e_i ; $P_k(q_{j_k}/e_i)$ – вероятность того, что в результате проверки π_k будет получен исход q_{j_k} при условии, что ОД находится в состоянии e_i ; $P_{k-1}(e_i)$ – вероятность того, что перед проверкой π_k ОД находится в состоянии e_i ; $P_k(e_i/q_{j_k})$ – вероятность того, что после проверки π_k с исходом q_{j_k} ОД находится в состоянии e_i (апостериорная вероятность).

Матрица условных вероятностей

Условные вероятности объединяются в матрицу (таблица 1).

Эти вероятности связаны с вероятностями состояний формулой Бейеса (1) для конкурирующих гипотез [3].

Очевидно, что сумма вероятностей в каждом из столбцов равна единице.

$$P_k(e_r/q_{j_k}) = \frac{P_{k-1}(e_r)P_k(q_{j_k}/e_r)}{\sum_{i=1}^S P_{k-1}(e_i)P_k(q_{j_k}/e_i)}. \quad (1)$$

Значения вероятностей $P_k(e_r/q_{j_k})$ существенно зависят от вида и очередности всех предыдущих проверок.

Значения вероятностей $P_k(q_{j_k}/e_i)$, напротив, не зависят от последовательности проведения проверок, а целиком определяются видом проверки π_k . Действительно, в процессе диагностирования ОД не изменяет своего состояния, результаты не зависят от момента времени и очередности, в которой производится проверка π_k .

Каждая проверка π_k характеризуется матрицей условных вероятностей $[P_k(q/e)]$.

$$P_{j_k i} = [P_k(q_{j_k}/e_i)].$$

Сумма вероятностей в каждом столбце матрицы $[P_k(q/e)]$ равна единице $\sum_{j_k=1}^{f_k} P_{j_k i} = 1$,

($i = 0, 1, \dots, s$), так как для любого состояния e_i проверка π_k непременно должна завершиться одним из исходов q_{j_k} .

Состояния

В качестве примера рассмотрим разделение состояний пациента по допустимым интервалам времени, в течение которых ему должна быть оказана медицинская помощь (таблица 2).

Таблица 2. Состояния пациента

Состояние e_i	Наименование состояния	Вероятность состояния $P_k(e_i)$
e_1	Экстренная помощь (до одного часа)	$P_k(e_1)$
e_2	Неотложная помощь (до одних суток)	$P_k(e_2)$
e_3	Плановое обследование (до одного месяца)	$P_k(e_3)$
e_4	Плановое обследование (до одного года)	$P_k(e_4)$
e_5	Необходимость обследования (свыше года)	$P_k(e_5)$

Таблица 3. Проверки

№	Проверки π_k	e_1	e_2	e_3	e_4	e_5
1	Сильная, сдавливающая боль за грудиной	*				
2	Нарастающее чувство беспокойства, страха			*		
3	Головокружение			*		
4	Резкое нарушение ритма сердца	*				
5	Затрудненное дыхание			*		
6	Повышенная потливость			*		
7	Слабость, тошнота и рвота	*				
8	Потеря сознания	*				
9	Ощущение холода	*				
10	Головная боль			*		
11	Сахарный диабет			*		
12	Гипертоническая болезнь			*		
13	Ишемическая болезнь сердца		*			
14	Приступы стенокардии		*			
15	Повышенное содержание в крови холестерина			*		
16	Чрезмерная масса тела				*	
17	Курение				*	
18	Злоупотребление алкоголем				*	
19	Большое потребление соли, животных жиров				*	
20	Сердечная аритмия			*		
21	Длительные стрессовые ситуации			*		
22	Давящая боль в области сердца		*			
23	Дискомфорт в области сердца			*		
24	Покалывание или дискомфорт в руках		*			
25	Покалывание или дискомфорт в спине		*			
26	Покалывание или дискомфорт в челюстях		*			
27	Покалывание или дискомфорт в желудке		*			
28	Одышка (затруднение дыхания)			*		
29	Тошнота, рвота, отрыжка или изжога		*			
30	Липкий холодный пот		*			
31	Учащение или замедление сердцебиения		*			
32	Наследственная предрасположенность к заболеваниям сердца				*	
33	Ожирение			*		
34	Отсутствие регулярных физических нагрузок				*	
35	Возраст старше 50 лет				*	
36	Мужской пол				*	

Выбор одного из состояний пациента происходит на основании апостериорных вероятностей $P_k(e_i)$, получаемых в результате ответа на ряд

последовательно задаваемых вопросов, называемых проверками π_k . Так, например, для определения действия в случае подозрения на инфаркт

Таблица 4. Значения интервалов

Состояния		T_{\min}	T_{\max}	\bar{T}	θ_{\min}	θ_{\max}	$\bar{\theta}_i$	$P_{\kappa}(e_i / q_{j_{\kappa}})$
e_1	Экстренная помощь (до одного часа)	1	6	3,5	0	0,78	0,47	$P_{\kappa}(e_1 / q_{j_{\kappa}})$
e_2	Неотложная помощь (до одних суток)	6	144	75	0,78	2,15	1,87	$P_{\kappa}(e_2 / q_{j_{\kappa}})$
e_3	Плановое обследование (до одного месяца)	144	4320	2232	2,15	3,63	3,35	$P_{\kappa}(e_3 / q_{j_{\kappa}})$
e_4	Плановое обследование (до одного года)	4320	51840	28080	3,63	4,71	4,48	$P_{\kappa}(e_4 / q_{j_{\kappa}})$
e_5	Необходимость обследования (свыше года)	28080	∞	–	4,71	∞	–	$P_{\kappa}(e_5 / q_{j_{\kappa}})$

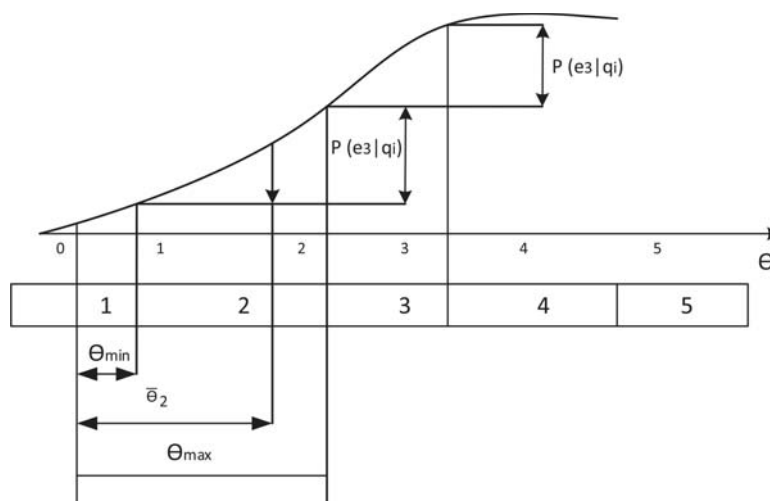


Рисунок. Шкала изменения времени

миокарда нами предлагается следующий ряд проверок (таблица 3). В таблице «звездочками» отмечены проверки, которые, с точки зрения экспертов, наиболее информативны для данного состояния. Такое состояние выбирается в качестве главного состояния e_r данной проверки (таблица 1).

Шкала состояний

Измерение длительности интервалов будем осуществлять в относительных единицах $T = t / t_0$. В качестве единицы измерения времени примем $t_0 = 10$ минут. T_{\min} и T_{\max} – начальное и конечное значения интервалов в относительных единицах, соответственно (таблица 4). Учитывая существенное различие длительностей интервалов, целесообразно принять логарифмический масштаб измерения времени $\theta = \log t / t_0$ единиц (десятичные логарифмы). θ_{\min} и θ_{\max} – начальное и конечное значения интервалов в логарифмических единицах соответственно. Нулевое значение логарифмической шкалы времени соответствует 10 минутам. Плотность распределения вероятностей принадлежности к данным состояниям предлагается представить в виде кривой нормального распределения в логарифмическом

масштабе времени θ . Вероятности $P_{\kappa}(e_i / q_{j_{\kappa}})$ принадлежности пациента к состоянию e_i , при условии ответа $q_{j_{\kappa}}$ характеризуются интегралом от плотности распределения вероятностей в пределах интервала времени, соответствующего данному состоянию.

$$P_{\kappa}(e_i / q_{j_{\kappa}}) = P(\theta_{i\max}, \theta_{i\min}) = \frac{1}{\sigma_{\theta} \sqrt{2\pi}} \int_{\theta_{\min}}^{\theta_{\max}} e^{-\frac{(\theta - \bar{\theta})^2}{2\sigma_{\theta}^2}} d\theta. \quad (2)$$

Здесь $\bar{\theta}$ – математическое ожидание и σ_{θ} – среднеквадратичное отклонение в логарифмическом масштабе.

На рисунке показана в логарифмическом масштабе θ шкала изменения времени, в течение которого следует оказать медицинскую помощь. Шкала разбита на 5 интервалов, соответствующих таблице 4. Для каждой из проверок выбирается главное состояние, на которое нацелена данная проверка (таблица 3). На рисунке в качестве главного выбрано состояние e_2 , которому соответствует второй интервал.

Показаны времена θ_{\min} , θ_{\max} начала и конца интервала соответственно, а также значение $\bar{\theta}$

Таблица 5. Матрица вероятностей проверки

Исходы проверки π_k	Состояния						
	e_1	...	e_r	...	e_i	...	e_s
q_{1_k}	$P_k(q_{1_k}/e_1)$...	$P_k(q_{1_k}/e_r)$...	$P_k(q_{1_k}/e_i)$...	$P_k(q_{1_k}/e_s)$
q_{0_k}	$P_k(q_{0_k}/e_1)$...	$P_k(q_{0_k}/e_r)$...	$P_k(q_{0_k}/e_i)$...	$P_k(q_{0_k}/e_s)$

Таблица 6. Значения начальных вероятностей

Исходы проверки π_k	e_1	e_2	e_3	e_4	e_s
$P_k(q_{1_k}/e_i)$	0,65	0,31	0,038	0,010	0,0
$P_k(q_{0_k}/e_i)$	0,35	0,69	0,962	0,990	1,0

математического ожидания данного интервала, соответствующие таблице 4. (В логарифмическом масштабе математическое ожидание не совпадает с серединой интервала).

Относительно данного математического ожидания построена функция распределения вероятностей $F(\theta, \sigma_\theta)$. Разность между значениями функции распределения вероятностей, соответствующими θ_{\max} и θ_{\min} , определяет базовое значение условной вероятности $P(e_2/q_j)$ для второго интервала. Аналогично определяется значение вероятности $P(e_3/q_j)$ для третьего интервала, показанное на рисунке, и значения вероятностей для всех остальных интервалов.

Дисперсии определяют скорость достижения порога при положительных ответах.

Простейшие проверки

Простейшей будем считать такую проверку π_k , которая имеет всего два исхода: положительный q_{1_k} и отрицательный q_{0_k} . Матрица вероятностей такой проверки имеет всего две строчки (таблица 5):

Используя рассмотренное выше распределение вероятностей при заданных значениях дисперсии, определим для значения начальных вероятностей при положительном исходе рассматриваемой проверки (таблица 6).

Очевидно, что $P_k(q_{0_k}/e_i) = 1 - P_k(q_{1_k}/e_i)$.

Заменим вероятности j -й строки таблицы 1 на другие вероятности, которые получаются, и умножим значения всех элементов j -й строки k -й проверки на постоянную величину x_{j_k}

$$\tilde{P}_k(q_{1_k}/e_i) = P_k(q_{1_k}/e_i)x_{1_k}.$$

Очевидно, что в соответствии с (1) значение

$$\begin{aligned} \tilde{P}_k(e_r/q_{1_k}) &= \\ &= \frac{P_{k-1}(e_r)\tilde{P}_k(q_{1_k}/e_r)}{\sum_{i=1}^S P_{k-1}(e_i)\tilde{P}_k(q_{1_k}/e_i)} = P_k(e_r/q_{1_k}). \end{aligned}$$

Таким образом, пропорциональное изменение вероятностей j -й строки k -й проверки не изменяет вероятностей результата исхода q_{j_k} . Однако пропорциональное изменение элементов первой строки таблицы 1 приводит к непропорциональному изменению элементов второй строки этой таблицы (сумма вероятностей по столбцам равна единице).

$$\tilde{P}_k(q_{0_k}/e_i) = 1 - \tilde{P}_k(q_{1_k}/e_i).$$

Следовательно,

$$\tilde{P}_k(e_r/q_{0_k}) = \frac{P_{k-1}(e_r)[1 - \tilde{P}_k(q_{1_k}/e_r)]}{\sum_{i=1}^S P_{k-1}(e_i)[1 - \tilde{P}_k(q_{1_k}/e_i)]},$$

$$\begin{aligned} \tilde{P}_k(e_r/q_{0_k}) \sum_{i=1}^S P_{k-1}(e_i)[1 - \tilde{P}_k(q_{1_k}/e_i)] &= \\ &= P_{k-1}(e_r)[1 - \tilde{P}_k(q_{1_k}/e_r)], \end{aligned}$$

$$\tilde{P}_k(e_r/q_{0_k}) \sum_{i=1}^S P_{k-1}(e_i) - \tilde{P}_k(e_r/q_{0_k}) \times$$

$$\times \sum_{i=1}^S P_{k-1}(e_i)\tilde{P}_k(e_i/q_{1_k}) =$$

$$= P_{k-1}(e_r) - P_{k-1}(e_r)\tilde{P}_k(q_{1_k}/e_r)].$$

Поскольку $\sum_{i=1}^S P_{k-1}(e_i) = 1$, получим:

$$\begin{aligned} \tilde{P}_k(e_r/q_{0_k}) - \tilde{P}_k(e_r/q_{0_k}) \sum_{i=1}^S P_{k-1}(e_i)\tilde{P}_k(q_{1_k}/e_i) &= \\ &= P_{k-1}(e_r) - P_{k-1}(e_r)\tilde{P}_k(q_{1_k}/e_r)], \end{aligned}$$

Таблица 7. Матрица вероятностей

Исходы проверки π_k	e_1	e_2	e_3	e_4	e_5
$\tilde{P}_\kappa(q_{1_k}/e_i)$	0,260	0,124	0,015	0,004	0,0
$\tilde{P}_\kappa(q_{0_k}/e_i)$	0,740	0,876	0,985	0,996	1,0

$$P_{\kappa-1}(e_r)\tilde{P}_\kappa(q_{1_k}/e_r)] - \tilde{P}_\kappa(e_r/q_{0_k}) \times \sum_{i=1}^s P_{\kappa-1}(e_i)\tilde{P}_\kappa(q_{1_k}/e_i) = P_{\kappa-1}(e_r) - \tilde{P}_\kappa(e_r/q_{0_k}).$$

Учитывая, что

$$\sum_{i=1}^s P_{\kappa-1}(e_i)\tilde{P}_\kappa(q_{1_k}/e_i) = \frac{P_{\kappa-1}(e_r)\tilde{P}_\kappa(q_{1_k}/e_r)}{\tilde{P}_\kappa(e_r/q_{1_k})},$$

получим:

$$P_{\kappa-1}(e_r)\tilde{P}_\kappa(q_{1_k}/e_r)] - \tilde{P}_\kappa(e_r/q_{0_k}) \times \frac{P_{\kappa-1}(e_r)\tilde{P}_\kappa(q_{1_k}/e_r)}{\tilde{P}_\kappa(e_r/q_{1_k})} = P_{\kappa-1}(e_r) - \tilde{P}_\kappa(e_r/q_{0_k}).$$

$$\tilde{P}_\kappa(q_{1_k}/e_r) = \frac{\tilde{P}_\kappa(e_r/q_{1_k})}{P_{\kappa-1}(e_r)} \times \left[\frac{P_{\kappa-1}(e_r) - \tilde{P}_\kappa(e_r/q_{0_k})}{\tilde{P}_\kappa(e_r/q_{1_k}) - \tilde{P}_\kappa(e_r/q_{0_k})} \right] = \frac{1 - \frac{\tilde{P}_\kappa(e_r/q_{0_k})}{P_{\kappa-1}(e_r)}}{1 - \frac{\tilde{P}_\kappa(e_r/q_{0_k})}{\tilde{P}_\kappa(e_r/q_{1_k})}} \quad (3)$$

Данное соотношение показывает, что требуемое значение главной условной вероятности $\tilde{P}_\kappa(q_{1_k}/e_r)$ проверки π_k определяется значениями трех вероятностей, которые мы предполагаем получить в результате проведения указанной проверки: $P_{\kappa-1}(e_r)$ – вероятность главного состояния перед началом указанной проверки; $\tilde{P}_\kappa(e_r/q_{1_k})$ – вероятность главного состояния в результате получения положительного ответа после указанной проверки; $\tilde{P}_\kappa(e_r/q_{0_k})$ – вероятность главного состояния в результате получения отрицательного ответа после указанной проверки.

Сделаем следующие допущения.

1. Обозначим через $P_{\max}(e_r)$ предельное значение вероятности состояний, при котором тестирование заканчивается. Тогда, при завершении тестирования положительным ответом при проверке π_k $\tilde{P}_\kappa(e_r/q_{1_k}) \geq P_{\max}(e_r)$.

2. Каждому состоянию e_r соответствует ряд проверок, для которых указанное состояние является главным (отмечены звездочками в таблице 3). Обозначим число таких проверок для состояния e_r через m_r и будем считать все эти

проверки равнозначными. Тогда при отрицательном ответе на одну из таких проверок апостериорная вероятность состояния $\tilde{P}_\kappa(e_r/q_{0_k})$ по сравнению с априорной вероятностью $P_{\kappa-1}(e_r)$ должна уменьшиться:

$$\tilde{P}_\kappa(e_r/q_{0_k}) = (1 - \frac{1}{m_r})P_{\kappa-1}(e_r).$$

3. Предположим, что перед началом проверки все состояния были равновероятны $P_{\kappa-1}(e_r) = 1/s$.

В результате указанных допущений получим:

$$\tilde{P}_\kappa(q_{1_k}/e_r) = \frac{sP_{\max}(e_r)}{1 + m_r[sP_{\max}(e_r) - 1]} \quad (4)$$

Так, в частности, в соответствии с таблицей 4 $s = 5$, для состояния e_1

$$(r = 1), \quad m_r = 5,$$

Если принять пороговое значение вероятности $P_{\max}(e_1) = 0,7$, получим $\tilde{P}_\kappa(q_{1_k}/e_1) = 0,26$.

Учитывая ранее принятое обозначение, определим значение

$$x_{1_k} = \frac{\tilde{P}_\kappa(q_{1_k}/e_1)}{P_\kappa(q_{1_k}/e_1)} = \frac{0,26}{0,65} = 0,4,$$

на которое необходимо умножить все вероятности первой строки данной проверки, полученные в соответствии с таблицей 6: $\tilde{P}_\kappa(q_{1_k}/e_i) = P_\kappa(q_{1_k}/e_i)x_{1_k}$. Указанные вероятности составляют первую строку матрицы рассматриваемой проверки π_k , соответствующую положительному ответу ($q_1 = 1$) и приведенную в таблице 7.

Вторая строка получается вычитанием из единицы значений элементов первой строки.

$$P_\kappa(e_r/q_{j_k}) = \frac{P_{\kappa-1}(e_r)}{P_{\kappa-1}(e_r) + \alpha_{j_k} \sum_{\substack{i=1 \\ i \neq r}}^s P_{\kappa-1}(e_i)} = \frac{P_{\kappa-1}(e_r)}{P_{\kappa-1}(e_r)(1 - \alpha_{j_k}) + \alpha_{j_k}} = \frac{1}{1 + \alpha_{j_k} \frac{1 - P_{\kappa-1}(e_r)}{P_{\kappa-1}(e_r)}}.$$

Нетрудно показать, что в случае равенства между собой всех элементов одной из строк таблицы 1 апостериорная вероятность равна апри-

орной вероятности $P_{\kappa}(e_r/q_{j_k}) = P_{\kappa-1}(e_r)$, то есть такая проверка является неинформативной.

При положительном ответе увеличению апостериорной вероятности $P_{\kappa}(e_r/q_{1_k})$, по сравнению с априорной, соответствует увеличение значения условной вероятности $P_{\kappa}(q_{1_k}/e_r)$ в верхней строке таблицы 1 по сравнению с остальными элементами указанной строки. Это означает, что коэффициенты α_{1_k} при $i \neq r$ должны быть существенно меньше единицы. Очевидно, что $\alpha_{1_r} = 1$.

Допустим, например, что априорная вероятность $P_{\kappa-1}(e_r) = 0,2$, условная вероятность $P_{\kappa}(q_{1_k}/e_r) = 0,2$ а условные вероятности $P_{\kappa}(q_{1_k}/e_i) = 0,02$, при $i \neq r$. Тогда $\alpha_{1_k} = 0,1$. При положительном ответе расчетное значение $P_{\kappa}(e_r/q_{1_k}) = 0,714$ превышает априорное значение более чем в три раза.

При отрицательном ответе

$$\alpha_{0_k} = \frac{P_{\kappa}(q_{0_k}/e_i)}{P_{\kappa}(q_{0_k}/e_r)} = \frac{1 - P_{\kappa}(q_{1_k}/e_i)}{1 - P_{\kappa}(q_{1_k}/e_r)} = \frac{0,98}{0,8} = 1,225,$$

а расчетное апостериорное значение $P_{\kappa}(e_r/q_{0_k}) = 0,169$, что мало отличается от априорного, которое было равно 0,2.

– Чем ближе значения α_{1_k} к нулю, тем ближе значения апостериорной вероятности к единице.

– Чем ближе значение α_{0_k} к единице, тем меньше апостериорная вероятность отличается от априорной.

Коэффициенты α_{j_k} могут быть выражены через значения априорных и апостериорных вероятностей:

$$\alpha_{j_k} = \frac{P_{\kappa-1}(e_r) \cdot [1 - P_{\kappa}(e_r/q_{j_k})]}{1 - P_{\kappa-1}(e_r) \cdot P_{\kappa}(e_r/q_{j_k})}.$$

Таким образом, если известны априорные вероятности $P_{\kappa-1}(e_r)$ нахождения пациента в состоянии e_r и эксперт считает, что в результате проверки π_k с исходом q_{j_k} должны получиться апостериорные значения вероятностей $P_{\kappa}(e_r/q_{j_k})$, то может быть определен коэффициент α_{j_k} , характеризующий отношение коэффициентов в строке исхода q_{j_k} указанной проверки.

И наоборот, апостериорная вероятность проверки может быть выражена через априорную вероятность и коэффициент α_{j_k} :

$$P_{\kappa}(e_r/q_{j_k}) = \frac{P_{\kappa-1}(e_r)}{P_{\kappa-1}(e_r) + \alpha_{j_k} [1 - P_{\kappa-1}(e_r)]}.$$

Как и было отмечено выше, при малых значениях коэффициента α_{j_k} , апостериорная вероятность $P_{\kappa}(e_r/q_{j_k})$ стремится к единице. Если же значения коэффициента α_{j_k} стремятся к единице, то апостериорная вероятность равна априорной.

Наконец, условные вероятности $P_{\kappa}(q_{1_k}/e_r)$ и $P_{\kappa}(q_{0_k}/e_r)$ легко определяются значениями α_{1_k} и α_{0_k} :

$$P_{\kappa}(q_{1_k}/e_r) = \frac{\alpha_{0_k} - 1}{\alpha_{0_k} - \alpha_{1_k}} = 0,2,$$

$$P_{\kappa}(q_{0_k}/e_r) = \frac{1 - \alpha_{1_k}}{\alpha_{0_k} - \alpha_{1_k}} = 0,8,$$

что и было принято раньше.

В дальнейшем при проведении проверок полученное в результате k -й проверки расчетное значение априорной вероятности $P_{\kappa}(e_r/q_{j_k})$ может быть принято за исходное априорное значение для следующей выбранной проверки $P_{\kappa}(e_r) = P_{\kappa}(e_r/q_{j_k})$.

Дальнейшие проверки выбираются по критерию максимальной информативности [...].

Условные вероятности

Ключевым вопросом системы является создание базы условных вероятностей (таблица 1).

При создании базы необходимо учитывать полученные выше рекомендации, а также ряд свойств проверок по отношению к диагностируемому состоянию. Так, например, при положительном исходе проверки № 8 «Потеря сознания» несомненно, необходимо обращение к скорой помощи (состояние $e_r = e_1$). В то же время отрицательный исход этой проверки совершенно не исключает необходимость указанного обращения (апостериорная вероятность после такой проверки с отрицательным результатом не изменяется по сравнению с априорной вероятностью (проверка при данном исходе малоинформативна).

Выполнение этих условий обеспечивается соответствующим выбором коэффициентов α_{j_k} , которые определяют значения условных вероятностей.

Для проверки «Потеря сознания» нами предлагаются следующие значения в таблице условных вероятностей (таблица 8).

Для указанных значений условных вероятностей получаем значения коэффициентов $\alpha_{1_k} = 0,033$; $\alpha_{0_k} = 1,41$. Если до начала тестирования о состоянии пациента ничего не известно, то естественно принять все пять состояний равновероятными и все начальные априорные веро-

Таблица 8. Проверка «Потеря сознания»

Исходы проверки π_k	Состояния				
	e_1 Вызвать скорую помощь	e_2 Немедленно обратиться к врачу	e_3 Пройти обследование	e_4 Учесть склонность к данному заболеванию	e_5 Отсутствие показаний к данному заболеванию
«Да»	0,3	0,01	0,01	0,01	0,01
«Нет»	0,7	0,99	0,99	0,99	0,99

Таблица 9. Проверка «Головная боль»

Исходы проверки π_k	Состояния				
	e_1 Вызвать скорую помощь	e_2 Немедленно обратиться к врачу	e_3 Пройти обследование	e_4 Учесть склонность к данному заболеванию	e_5 Отсутствие показаний к данному заболеванию
«Да»	0,09	0,09	0,1	0,09	0,09
«Нет»	0,91	0,91	0,9	0,91	0,91

ятности $P_0(e_i) = 0,2$. После завершения проверки положительным исходом «Да» апостериорная вероятность

$$P_1(e_1 / q_1) = \frac{P_0(e_1)}{P_0(e_1) + \alpha_1 [1 - P_0(e_1)]} = \frac{0,2}{0,2 + 0,033[1 - 0,2]} = 0,88.$$

Это весьма высокая вероятность, которая указывает, что при потере сознания, несомненно, необходимо обращение к скорой помощи. При отрицательном ответе (потеря сознания отсутствует), получаем:

$$P_1(e_1 / q_0) = \frac{P_0(e_1)}{P_0(e_1) + \alpha_0 [1 - P_0(e_1)]} = \frac{0,2}{0,2 + 1,414[1 - 0,2]} = 0,15,$$

в то время как перед проверкой вероятность $P_0(e_i) = 0,2$. Апостериорная вероятность после проверки с отрицательным ответом мало изменилась по сравнению с априорной.

Другой характер должна иметь проверка «Головная боль». При получении положительного ответа на нее вряд ли следует давать совет о немедленном обращении к врачу, и тем более – о немедленном вызове скорой помощи. Как следует из таблицы 3, проверка ориентирована на состояние e_3 , при котором следует «Пройти обследование». Однако получение положительного ответа на эту проверку еще не означает, что однозначно пациент должен быть направлен на прохождения обследования. Апостериорная вероятность указанного состояния e_3 после получения положительного ответа по такой проверке не должна резко возрастать так же, как и вероятность этого

состояния при отрицательном ответе не должна резко убывать. Такие характеристики проверки достигаются соответствующим выбором значений коэффициентов α_{1k} и α_{0k} . Для рассматриваемой проверки предлагается следующая таблица условных вероятностей (таблица 9).

Для указанных условных вероятностей получаем значения коэффициентов $\alpha_{1k} = 0,9$ и $\alpha_{0k} = 1,011$. При начальных условиях, аналогичных рассмотренной выше проверке,

$$P_1(e_3 / q_1) = \frac{P_0(e_3)}{P_0(e_3) + \alpha_1 [1 - P_0(e_3)]} = \frac{0,2}{0,2 + 0,9[1 - 0,2]} = 0,217,$$

$$P_1(e_3 / q_0) = \frac{P_0(e_3)}{P_0(e_3) + \alpha_0 [1 - P_0(e_3)]} = \frac{0,2}{0,2 + 1,011[1 - 0,2]} = 0,198.$$

Заключение

Апостериорные вероятности весьма незначительно отличаются от априорных, которые принимались нами равными 0,2. Это означает, что при данных условиях рассматриваемая проверка малоинформативна. Поскольку в системе выбор очередности проверок производится, исходя из их информативности, эта проверка должна проводиться в конце процесса диагностирования.

По мере ответов на задаваемые вопросы проверок изменяются значения апостериорных вероятностей $P_k(e_r / q_{jk})$, которые после проведения очередной проверки сравниваются с граничными значениями $P_{\max}(e_i)$. Если полученное значение превышает максимальное, то дальнейшая проверка прекращается и значения апостериорных

вероятностей выводятся пользователю. На основе вышесказанного можно сделать вывод, что данная система диагностирования является достаточно эффективной.

Литература

1. Акимова Э.К. Экспертная система-консультант для врача-эпидемиолога // Медицинская техника. 1989. № 3. С. 24–28.
2. Ахутин В.М., Шаповалов В.В., Мансур Д. Автоматизированные системы профилактических осмотров детей (АСПОН-Д) – состояние и перспективы // Биотехнические и медицинские системы: сб. науч. тр. 1990. С. 3–6.
3. Борецкий А.Б., Маслов В.Г., Хавронина М.А. Идентификация экспертных знаний на основе компетентных решающих правил в медицинских экспертных системах // Информатика в здравоохранении: материалы Всесоюз. научн. конф. 1990. С. 19.
4. Долина О.Н., Шварц Ю.Г. Построение и отладка медицинских экспертных систем // Математические методы в технике и технологиях – ММТТ-2000: сб. тр. 13 Международ. науч. конф. 2000. С. 37–39.
5. Использование вероятностных методов оценки знаний при разработке тестирующих модулей распределенных тренинг-систем / Б.Я. Лихтциндер [и др.] // Инфокоммуникационные технологии. 2003. Т. 1, № 3. С. 40–45.
6. Лихтциндер Б.Я., Иванова Л.Б. Вероятностные методы контроля и диагностирования многопараметрических объектов. М.: Информприбор, 1987. № 3682.
7. Лихтциндер Б.Я., Погодина Л.И., Пугин В.В. Автоматизированная справочная система терапевта. Самара: АТИ, 2004. 16 с.
8. Поляков М.И., Баллюзек Ф.В., Добрынин Е.В. Особенности представлений и использование знаний в медицинских экспертных системах // Информатика в здравоохранении: материалы Всесоюз. научной конф. 1990. С. 67.
9. Устинов А.Г., Ситарчук В.А., Кореневский Н.А. Автоматизированные медико-технологические системы: монография в 3-х ч. / под ред. А.Г. Устинова. Курск: Гос. техн. ун-т, 1995. 390 с.
10. Пугин В.В. Экспертные системы в медицине // Тезисы доклада VIII Российской научной конференции ПГАТИ. 2001. С. 96–97.
11. Пугин В.В. Диагностические системы в медицине // Тезисы доклада X Российской научной конференции ПГАТИ. 2003. С. 38.
12. Пугин В.В., Иванова Л.Б. Эффективность применений вероятностных методов диагностики в многопользовательской сетевой среде // Проблемы техники и технологии телекоммуникаций: материалы VII Международной научно-технической конференции. 2006. С. 150–152.

Получено 18.02.2021

Лихтциндер Борис Яковлевич, д.т.н., профессор кафедры сетей и систем связей Поволжского государственного университета телекоммуникаций и информатики (ПГУТИ). 443010, Российская Федерация, г. Самара, ул. Л. Толстого, 23. Тел. +7 846 333-47-69. E-mail: lixt@psuti.ru

Ларина Валерия Анатольевна, студентка второго курса кафедры программного обеспечения и управления в технических системах ПГУТИ. 443010, Российская Федерация, г. Самара, ул. Л. Толстого, 23. Тел. +7 987 957-94-46. E-mail: valeriya-larina-2000@mail.ru

Муравец Александр Владимирович, главный врач ГБУЗ «Самарский областной центр медицинской профилактики «Центр общественного здоровья». 443020, Российская Федерация, г. Самара, ул. Самарская, 93. Тел. +7 846 333-44-52. E-mail: sam_ompr@mail.ru

DEVELOPMENT OF THE AUTOMATED SYSTEM OF THE CONVEYOR-TELETRAFIC ANALYSIS (ASKAT)

Likhhtsinder B. Ya.¹, Larina V. A.¹, Moravec A. V.²

¹ *Povolzhskiy State University of Telecommunications and Informatics, Samara, Russian Federation*

² *Samara Regional Center for Medical Prevention «Public Health Center»,*

Samara, Russian Federation

E-mail: lixt@psuti.ru

This article deals with the probabilistic test criteria of an information-diagnostic system. Probabilistic test criteria for an information and diagnostic system aimed at studying a patient's condition are considered. The system contains a diagnostic object (patient) and diagnostic tools. The diagnostic system implements some diagnostic algorithm that specifies the sequence and method of analysis of examination results. The purpose of analyzing the results of elementary tests is to determine the diagnosable condition in which the patient really is. This paper covers such sections as probabilistic diagnosis of a patient's condition; conditional probability matrix; states; scale of states; elementary checks; conditional probabilities. All data for each section is summarized in a table for visualization. The conclusion is made on the basis of the analysis, which allows evaluating the effectiveness of the application of this diagnostic system.

Keywords: *probabilistic criteria, information and diagnostic systems, patient diagnosis, examination*

DOI: 10.18469/ikt.2021.19.4.09

Likhhtsinder Boris Yakovlevich, Povolzhskiy State University of Telecommunications and Informatics, 23, L. Tolstoy Street, Samara, 443010, Russian Federation; Doctor of Technical Science, Professor of Network and Communication Systems Department. Tel. +7 846 333-47-69. E-mail: lixt@psuti.ru

Larina Valeriya Anatolievna, Povolzhskiy State University of Telecommunications and Informatics, 23, L. Tolstoy Street, Samara, 443010, Russian Federation; Student. Tel. +7 987 957-94-46. E-mail: valeriya-larina-2000@mail.ru

Muravets Alexander Vladimirovich, Samara Regional Center for Medical Prevention «Public Health Center», 93, Samarskaya Street, Samara, 443020, Russian Federation; Chief Physician. Tel. +7 846 333-44-52. E-mail: sam_ocmp@mail.ru

References

1. Akimova E.K. Expert system-consultant for epidemiologist. *Meditinskaja tehnika*, 1989, no. 3, pp. 24–28. (In Russ.)
2. Ahutin V.M., Shapovalov V.V., Mansur D. Automated systems for preventive examinations of children (ASPO-N-D) - state and prospects. *Biotehnicheskie i meditsinskie sistemy: sb. nauch. tr.*, 1990, pp. 3–6. (In Russ.)
3. Boretskij A.B., Maslov V.G., Havronina M.A. Expert knowledge identification based on competent decision rules in medical expert systems. *Informatika v zdravoohranenii: materialy Vsesojuz. nauchn. konf.*, 1990, p. 19. (In Russ.)
4. Dolina O.N., Shvarts Yu.G. Building and debugging medical expert systems. *Matematicheskie metody v tehnike i tehnologijah – MMTT-2000: sb. tr. 13 Mezhdunar. nauch. konf.*, 2000, pp. 37–39. (In Russ.)
5. Likhhtsinder B. Ya. et al. The use of probabilistic methods for assessing knowledge in the development of testing modules for distributed training systems. *Infokommunikacionnye tehnologii*, 2003, vol. 1, no. 3, pp. 40–45. (In Russ.)
6. Likhhtsinder B. Ya., Ivanova L.B. *Probabilistic Methods for Monitoring and Diagnosing Multi-Parameter Objects*. Moscow: Informpribor, 1987, no. 3682. (In Russ.)
7. Likhhtsinder B. Ya., Pogodina L.I., Pugin V.V. *Automated Therapist Help System*. Samara: ATI, 2004, 16 p. (In Russ.)
8. Poljakov M.I., Balljuzek F.V., Dobrynin E.V. Features of representations and use of knowledge in medical expert systems. *Informatika v zdravoohranenii: materialy Vsesojuz. nauchnoj konf.*, 1990, p. 67. (In Russ.)
9. Ustinov A.G., Sitarchuk V.A., Korenevsky N.A. *Automated Medical And Technological Systems: monograph in 3 parts*. Ed. by A.G. Ustinov. Kursk: Gos. tehn. un-t, 1995, 390 p. (In Russ.)

10. Pugin V.V. Expert systems in medicine. *Tezisy doklada VIII Rossijskoj nauchnoj konferentsii PGATI*, 2001, pp. 96–97. (In Russ.)
11. Pugin V.V. Diagnostic systems in medicine. *Tezisy doklada X Rossijskoj nauchnoj konferentsii PGATI*, 2003, p. 38. (In Russ.)
12. Pugin V.V., Ivanova L.B. Efficiency of applications of probabilistic diagnostic methods in a multi-user network environment. *Problemy tehniki i tehnologii telekommunikatsij: materialy VII Mezh-dunarodnoj nauchno-tehnicheskoy konferentsii*, 2006, pp. 150–152. (In Russ.)

Received 18.02.2021

УДК 338.2:004

ОЦЕНКА ЭФФЕКТИВНОСТИ ДЕЯТЕЛЬНОСТИ МЕДИЦИНСКИХ УЧРЕЖДЕНИЙ НА БАЗЕ СБАЛАНСИРОВАННОЙ СИСТЕМЫ ПОКАЗАТЕЛЕЙ

Диязитдинова А.Р.¹, Сударушкина Е.В.², Федорова М.В.²

¹ *Поволжский государственный университет телекоммуникаций и информатики, Самара, РФ*

² *Самарский областной медицинский информационно-аналитический центр (МИАЦ), Самара, РФ*
E-mail: dijazitdinova@mail.ru

В статье рассмотрена актуальность проведения исследований по оценке эффективности деятельности медицинских учреждений в условиях цифровизации. В настоящее время все большее число медицинских учреждений в условиях фиксированного бюджета заинтересовано вопросами внедрения в практику действенных методов управления для достижения своих целей. Одной из таких методик управления, неплохо зарекомендовавших себя и доказавших на практике собственную эффективность, является система сбалансированных показателей (ССП, Balanced ScoreCard, BSC). Целью работы является исследование вопросов применения СПП в управлении лечебно-профилактическими учреждениями (ЛПУ) и выделение наиболее индикативных показателей, оценивающих достижение целей. В ходе выполнения исследования были изучены методики оценки эффективности деятельности медицинских организаций, установленные в разных субъектах РФ, а также зарубежный опыт. Существующая отчетность государственных медучреждений различается, так как в каждом субъекте РФ исполнительный орган в сфере здравоохранения с учетом методических рекомендаций Минздрава РФ устанавливает показатели для подведомственных организаций. Это затрудняет процесс комплексного определения результатов работы медорганизации за отчетный период. В работе предлагается набор универсальных показателей, основанных на СПП, позволяющих связать медицинскую результативность деятельности с достижением экономических показателей работы.

Ключевые слова: *цифровая медицина, оценка эффективности, показатели эффективности, лечебно-профилактическое учреждение (ЛПУ), сбалансированная система показателей (ССП), индикаторы результативности оказания медицинской помощи*

Введение

Одним из приоритетных направлений реализации нацпроекта «Здравоохранение» является цифровизация здравоохранения. Цель создания единого цифрового контура заключается в повышении эффективности управления, точности предоставляемых статистических данных и качества оказываемой медпомощи. Задачей цифровой трансформации здравоохранения выступает повышение качества, доступности и эффективности оказания медицинских услуг за счет использования цифровых медицинских сервисов врачами, пациентами, администраторами в сфере медицины для достижения стратегической цели – увеличения средней продолжительности

жизни населения РФ и роста активного трудоспособного возраста [1].

Законодательно государственная политика цифровой трансформации регулируется следующими ключевыми документами:

- Указом Президента РФ от 07.05.2018 г. № 204 «О национальных целях и стратегических задачах развития РФ на период до 2024 года»;
- Указом Президента РФ от 21.07.2020 г. № 474 «О национальных целях развития Российской Федерации на период до 2030 года»;
- Национальной программой «Цифровая экономика РФ»;
- Национальными проектами «Здравоохранение» и «Демография»;