

Интеллектуальный анализ данных и знаний по стентированию коронарных артерий*

Янковская А. Е.^{1,2}, Китлер С. В.³

ayuankov@gmail.com, svkitler@gmail.com

¹Томский государственный университет; ²Томский государственный архитектурно-строительный университет; ³Томский государственный университет систем управления и радиоэлектроники

Статья посвящена интеллектуальному анализу данных и знаний по стентированию коронарных артерий. Излагаются основные этапы интеллектуального анализа данных и знаний по стентированию коронарных артерий. Описывается подход к анализу данных и знаний по стентированию коронарных артерий, реализованный в интеллектуальной системе. Интеллектуальная система разработана в виде динамически подключаемых модулей к интеллектуальному инструментальному средству ИМСЛОГ, на базе которого конструируются прикладные интеллектуальные системы. Приводятся результаты исследования интеллектуальной системы.

Ключевые слова: интеллектуальный анализ данных и знаний, интеллектуальная система, матричная модель представления данных и знаний, выявление различного рода закономерностей, стентирование коронарных артерий.

Intelligent analysis of data and knowledge for stenting of coronary artery*

Yankovskaya A. E.^{1,2}, Kitler S. V.³

¹Tomsk State University; ²Tomsk State University of Architecture and Building; ³Tomsk State University of Control Systems and Radioelectronics

The paper is devoted to intelligent analysis of data and knowledge for stenting of coronary artery. The basic stages of the intelligent analysis of data and knowledge for stenting of coronary artery are described. An approach for analysis of data and knowledge for stenting of coronary artery is presented. Software realization of the approach is implemented in the intelligent system. The intelligent system is developed as dynamical plug-ins in the intelligent instrumental software (IIS) IMSLOG. The applied intelligent systems are constructed on the base of the IIS IMSLOG. Aprobatation results of the intelligent system is given.

Keywords: intelligent analysis of data and knowledge, intelligent system, matrix model of data and knowledge representation, revealing of various kinds regularities, stenting of coronary artery.

Введение

В настоящее время весьма актуально развитие технологии интеллектуального анализа данных и знаний [5–4]. Необходимость интеллектуальной обработки и управления знаниями в слабоструктурированных областях, таких как медицина, психология, социология, геология, биомедицина, обусловлена увеличением объема обрабатываемых данных и знаний, используемых при решении диагностических задач [5].

Работа частично выполнена при финансовой поддержке РФФИ, проекты № 13-07-00373, № 13-07-98037-р_сибирь_а, № 12-07-31109_мол-а и РГНФ № 13-06-00709.

Одной из наиболее значимых медицинских областей, несомненно, является кардиология, включающая такие важные задачи, как диагностика ишемической болезни сердца, диагностика окклюзий и рестенозов. Для их решения разработаны и внедрены различные системы [8–6]. Системы «Fuzzy expert system approach for coronary artery disease screening using clinical parameters» [8] и «A multilayer perceptron-based medical decision support system for heart disease diagnosis» [6] основаны на мягких вычислениях и предназначены для обнаружения и прогнозирования ишемической болезни сердца на ранней стадии и для диагностики сердечно-сосудистых заболеваний соответственно, система «Knowledge Management System for Clinical Decision Support — Application in Cardiology» [7] основана на статистических методах и используется для поддержки принятия решения по лечению больных с острым коронарным синдромом.

В отличие от вышеперечисленных систем в настоящем исследовании предлагается интеллектуальная система, основанная на логико-комбинаторных методах распознавания образов и предназначенная для диагностики окклюзий и рестенозов коронарных артерий.

Интеллектуальный анализ данных и знаний по стентированию коронарных артерий включает в себя формирование базы данных и знаний по исследуемой проблемной области, подход к анализу данных и знаний и его реализацию в интеллектуальной системе.

Следует отметить, что, заменив данные и знания из рассматриваемой области на данные и знания из другой проблемной области, предлагаемая интеллектуальная система будет ориентирована на анализ данных и знаний из вновь заполненной проблемной области, т.е. система не является узконаправленной.

В статье приводятся основные понятия и определения, излагается подход к выявлению различного рода закономерностей в данных и знаниях, описывается интеллектуальная система анализа данных и знаний по стентированию коронарных артерий, приводятся результаты ее исследования.

Основные понятия и определения

Для дальнейшего изложения воспользуемся понятиями и определениями, введенными в публикациях [5, 9].

Формирование базы данных и знаний осуществляется на основе матричной модели представления данных и знаний [9], включающей целочисленную матрицу описаний (\mathbf{Q}), задающую описание объектов в пространстве k -значных характеристических признаков z_1, \dots, z_m и целочисленную матрицу различий (\mathbf{R}), задающую разбиение объектов на классы эквивалентности по каждому механизму классификации. Если значение признака несущественно для объекта, то данный факт отмечается прочерком («-») в соответствующем элементе матрицы \mathbf{Q} . Для каждого признака z_j ($j \in \{1, 2, \dots, m\}$) задается интервал изменения его значений.

Множество всех неповторяющихся строк матрицы \mathbf{R} сопоставлено множеству выделенных образов, представленных однострочковой матрицей \mathbf{R}' , элементами которой являются номера образов.

На рис. 1 приведена матричная модель представления данных и знаний.

Данная модель позволяет представлять не только данные, но и знания экспертов, поскольку одной строкой матрицы \mathbf{Q} можно задавать в интервальной форме подмножество объектов, для которых характерны одни и те же итоговые решения, задаваемые соответствующими строками матрицы \mathbf{R} .

Диагностическим тестом (ДТ) называется совокупность признаков, различающих любые пары объектов, принадлежащих разным образам.

$$\mathbf{Q} = \begin{matrix} & \begin{matrix} z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 & z_8 & z_9 & z_{10} & z_{11} \end{matrix} \\ \begin{matrix} 1 \\ 4 \\ 3 \\ 3 \\ 2 \\ 2 \\ 1 \\ 3 \\ 5 \\ 4 \end{matrix} & \begin{bmatrix} 4 & 6 & 3 & 2 & 2 & 1 & 2 & 3 & 4 & 1 \\ 4 & 5 & 2 & 3 & 2 & 7 & 8 & 3 & 4 & 1 \\ 4 & 5 & 3 & 3 & 2 & 4 & 5 & 3 & 4 & 1 \\ 4 & 4 & 1 & 4 & 4 & 2 & 3 & 1 & 5 & 1 \\ 4 & 2 & 1 & 6 & 3 & 4 & 5 & 2 & 3 & 1 \\ 4 & 5 & 1 & 3 & - & 4 & 3 & 1 & 2 & 1 \\ 4 & 3 & 2 & 5 & 2 & 1 & 2 & 3 & 4 & 1 \\ 4 & 2 & 2 & 6 & 2 & 2 & 3 & 3 & 2 & 1 \\ 4 & 2 & 2 & 6 & 3 & 5 & 6 & 2 & 4 & 1 \\ 4 & 6 & 1 & 2 & 5 & 5 & 6 & 1 & 4 & 2 \end{bmatrix} & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{matrix} \end{matrix}$$

$$\mathbf{R} = \begin{matrix} & \begin{matrix} k_1 & k_2 \end{matrix} \\ \begin{matrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \\ 1 \\ 1 \\ 3 \\ 3 \end{matrix} & \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \\ 2 & 1 \\ 2 & 1 \\ 2 & 1 \\ 1 & 3 \\ 1 & 3 \\ 3 & 2 \\ 3 & 2 \end{bmatrix}$$

$$\mathbf{R}' = \begin{matrix} & \begin{matrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \\ 3 \\ 3 \\ 4 \\ 4 \end{matrix} \\ \begin{matrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \\ 3 \\ 3 \\ 4 \\ 4 \end{matrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \\ 3 \\ 3 \\ 4 \\ 4 \end{bmatrix}$$

Рис. 1: Матричная модель представления данных и знаний

Диагностический тест называется безызбыточным (тупиковым [5]), если содержит безызбыточное количество признаков.

Безызбыточный безусловный диагностический тест (ББДТ) характеризуется одновременным предъявлением всех входящих в него признаков исследуемого объекта при принятии решений.

Для представления диагностических тестов будем использовать двоичную матрицу тестов (\mathbf{T}) [10], столбцы которой сопоставлены столбцам матрицы \mathbf{Q} , строки — диагностическим тестам. Единичными значениями в каждой строке матрицы \mathbf{T} отмечены признаки, вошедшие в соответствующий данной строке диагностический тест.

На рис. 2 приведен пример построенной двоичной матрицы отказоустойчивых диагностических тестов \mathbf{T} , т. е. устойчивых к заданному числу t ошибок измерения (занесения) значений характеристических признаков исследуемых объектов [11].

$$\mathbf{T} = \begin{matrix} & \begin{matrix} z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 & z_8 & z_9 & z_{10} & z_{11} \end{matrix} \\ \begin{matrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \end{matrix} \end{matrix}$$

Рис. 2: Двоичная матрица отказоустойчивых диагностических тестов

Закономерности в данных и знаниях

Понятие закономерностей в исследованиях различных школ [5, 2] отличается от терминологии, предложенной А. Е. Янковской [9] и используемой в настоящем исследовании.

Под закономерностями понимаются подмножества признаков с определенными легко интерпретируемыми свойствами, влияющими на различимость объектов из разных образцов, устойчиво наблюдаемыми для объектов из обучающей выборки и проявляющимися на других объектах той же природы, а также весовые коэффициенты признаков, характеризующие их индивидуальный вклад [9] в различимость объектов и информационный

вес, определяемый в отличие от [12] на подмножестве тестов, используемых для принятия итогового решения [9].

К упомянутым подмножествам будем относить константные (принимающие одно и то же значение для всех образов), устойчивые (константные внутри образа, но не являющиеся константными), неинформативные (не различающие ни одной пары объектов), альтернативные (в смысле включения в ДТ), зависимые (в смысле включения подмножеств различимых пар объектов), несущественные (не входящие ни в один безызбыточный ДТ), обязательные (входящие во все безызбыточные ДТ), псевдообязательные (не являющиеся обязательными, но входящие во все ББДТ, участвующие в принятии решений) признаки, подмножества сигнальных признаков первого (различающие исследуемые объекты из двух образов) и второго рода (различающие исследуемые объекты, принадлежащие одному образу, с объектами, принадлежащими другому образу), а также все минимальные и все (либо часть при большом признаковом пространстве) безызбыточные различающие подмножества признаков, являющиеся, по сути, соответственно минимальными и безызбыточными ДТ [13]. Отказоустойчивые ББДТ также входят в число закономерностей [11].

Для выявления различного рода закономерностей при построении ББДТ применяется процедура построения матрицы импликаций (\mathbf{U}) [9], задающей различимость объектов из разных образов.

Матрица \mathbf{U} представляет собой целочисленную матрицу, столбцы которой сопоставлены столбцам матрицы \mathbf{Q} , а строки — всевозможным парам объектов v, l , соответственно из разных образов (классов) a, b ; $v \in \{1, 2, \dots, \sigma(\mathbf{Q}^a)\}$, $l \in \{1, 2, \dots, \sigma(\mathbf{Q}^b)\}$, где $\sigma(\mathbf{Q}^a)$ ($\sigma(\mathbf{Q}^b)$) — количество строк в подматрице \mathbf{Q}^a (\mathbf{Q}^b) матрицы \mathbf{Q} . Строка \mathbf{U}_i матрицы \mathbf{U} представляет собой значение целочисленной вектор-функции различения, j -я ($j \in \{1, 2, \dots, m\}$) компонента u_{ij} которой вычисляется по формуле:

$$u_{i,j} = |q_{v,j}^a - q_{l,j}^b|, \quad (1)$$

где $q_{v,j}^a$ ($q_{l,j}^b$) — значение признака z_j для объекта v (l); а $u_{i,j}$ вычисляется для каждой пары образов,

$$i = \prod_{r=1}^a \sigma(\mathbf{Q}^r) \sum_{s=r+1}^{b-1} \sigma(\mathbf{Q}^s) + \sum_{r=1}^{v-1} \sigma(\mathbf{Q}^b) + vl.$$

Будем говорить, что строка \mathbf{U}_d поглощает строку \mathbf{U}_ρ ($\mathbf{U}_d \succ \mathbf{U}_\rho$), если

$$(\mathbf{U}_d \succ \mathbf{U}_\rho) \leftrightarrow \forall i \in I(u_{di} \geq u_{\rho i}).$$

Безызбыточной матрицей импликаций называется матрица \mathbf{U}' , в которой отсутствуют поглощающие строки.

Далее воспользуемся понятиями константных, устойчивых, альтернативных и зависимых признаков, приведенными в публикации [10].

Если элементы в рассматриваемом столбце матрицы \mathbf{U}' принимают одинаковое значение, то признак, соответствующий этому столбцу, является константным. Иными словами константный признак соответствует столбцу матрицы \mathbf{Q} с одинаковыми значениями элементов.

Устойчивый признак соответствует столбцу матрицы \mathbf{Q} , содержащему одинаковые элементы внутри одного образа, но не являющийся константным.

Альтернативные признаки соответствуют одинаковым столбцам матрицы \mathbf{U}' .

Зависимые признаки соответствуют столбцам матрицы \mathbf{U}' , которые поглощаются другим столбцом (другими столбцами) матрицы \mathbf{U}' .

Одной из наиболее важных проблем при выявлении закономерностей и принятии решений является анализ данных и знаний, включающий выявление наиболее значимых признаков и оценивание величины их значимости [9].

Приведем формулы для вычисления значения весового коэффициента k -значного j -го признака (w_j) и вычисления значения веса i -го теста (W_i), данные в статье [14], и являющиеся модификацией формул, приведенных в публикации [9] для булевых признаков:

w_j — весовой коэффициент целочисленного признака z_j ($j \in \{1, 2, \dots, m\}$);

q_{aj} — значение признака z_j для объекта из образа a ;

δ_{ab}^j — величина j -й компоненты вектор-функции различения между объектами из разных образов a, b ($a \neq b$), вычисляемая по формуле:

$$\delta_{a,b}^j = |q_{a,j} - q_{b,j}|;$$

N_a — число строк в описании a -го образа;

K — число выделенных образов;

S_j — интервал изменения j -го признака, сопоставленного j -му столбцу матрицы \mathbf{Q} ;

p_l — коэффициент повторения l -й строки, задаваемый извне;

$$d_{lj}^- = \begin{cases} 1, & \text{если } q_{lj} = \langle - \rangle \\ 0, & \text{если } q_{lj} \neq \langle - \rangle; \end{cases}$$

σ_a — количество объектов в образе a ($a = 1, \dots, K$), вычисляемое по формуле:

$$\sigma_a = \sum_{l=1}^{N_a} p_l \prod_j (S_j)^{d_{lj}^-};$$

r, s ($r \neq s$) и a, b ($a \neq b$) — номера образов;

L_i — множество признаков, входящих в i -й тест.

Весовой коэффициент w_j целочисленного признака z_j [11] будем определять по его разделяющей способности, аналогично вычислению w_j для двоичного признака z_j [9].

$$w_j = \frac{\sum_{r=1}^{K-1} \sum_{s=r+1}^K \sum_{a=1}^{N_r} \sum_{b=1}^{N_s} \delta_{ab}^j}{S_j \sum_{a=1}^{K-1} \sum_{b=a+1}^K \sigma_a \sigma_b}. \quad (2)$$

Множество обязательных признаков будем называть ядром всех диагностических тестов, поскольку исключение любого признака из ядра нарушает свойство каждого из тестов быть тестом [9].

Каждому i -му тесту соответствует вес теста W_i [9], вычисляемый по формуле:

$$W_i = \sum_{j \in L_i} w_j.$$

Основные этапы интеллектуального анализа данных и знаний

Интеллектуальный анализ данных и знаний сводится к выявлению различного рода закономерностей.

Дано: матрицы \mathbf{Q} и \mathbf{R} (рис. 1) в пространстве k -значных характеристических признаков и диапазон изменения их значений.

Необходимо: построить двоичную матрицу \mathbf{U}'' , найти все ее безызбыточные столбцовые покрытия и соответствующие множества характеристических признаков, вошедших в k -значные ББДТ.

Кратко изложим основные этапы интеллектуального анализа данных и знаний [9, 11, 15].

1. Формирование матрицы \mathbf{R}' по матрице \mathbf{R} .
2. Упорядочивание строк матриц \mathbf{Q} и \mathbf{R}' по принадлежности к образам.
3. Построение по матрицам \mathbf{Q} и \mathbf{R}' матрицы \mathbf{U}' (формула 1) с одновременным удалением поглощающих строк и вычислением весовых коэффициентов признаков по формуле 2.
4. Построение двоичной матрицы \mathbf{U}'' путем замены значений всех элементов, отличных от нуля, в матрице \mathbf{U}' на значение «1».
5. Выявление различного рода закономерностей по матрице \mathbf{U}'' .
6. Построение всех безызбыточных столбцовых покрытий матрицы \mathbf{U}'' либо части при превышении числа безызбыточных столбцовых покрытий в процессе их построения наперед заданной величины g .
7. Построение по найденным безызбыточным столбцовым покрытиям матрицы \mathbf{U}'' всех ББДТ по алгоритму, приведенному в публикации [16].

Описание интеллектуальной системы анализа данных и знаний

Интеллектуальная система анализа данных и знаний предназначена для выявления различного рода закономерностей. Интеллектуальная система разработана в виде динамически подключаемых модулей к интеллектуальному инструментальному средству (ИИС) ИМСЛОГ [17], на базе которого конструируются прикладные интеллектуальные системы.

Архитектура ИИС ИМСЛОГ является открытой и представляет собой иерархическую систему программных модулей, разработанных с использованием средств Builder C++.

Один модуль реализован как резидентный, имеет встроенную систему команд, выполняет функции координирующего центра и называется ядром. Все остальные модули являются динамически подключаемыми, называются плагинами и подразделяются на функциональные модули, модули системных данных и базовый модуль интеллектуального пользовательского интерфейса.

Интеллектуальная система анализа данных и знаний по стентированию коронарных артерий (рис. 3) представляет собой связанный набор функциональных модулей.

Первый сверху модуль предназначен для работы с базой данных и знаний. Выходными данными модуля являются структура базы знаний, объекты базы знаний, название признаков и номера признаков.

Второй модуль предназначен для выбора признаков, включаемых в матрицы \mathbf{Q} и \mathbf{R} .

Третий модуль, получая на вход номера целочисленных характеристических признаков, номера целочисленных классификационных признаков, структуру базы знаний, номера выбранных обучающих объектов и соответствующие им описания объектов из базы знаний, строит матрицы \mathbf{Q} и \mathbf{R} . Также выходными данными этого модуля являются вектор номеров выбранных характеристических признаков, минимальные и максимальные значения этих признаков.

Четвертый сверху модуль предназначен для построения матрицы \mathbf{U}' . В этом модуле используются следующие входные параметры: целочисленные матрицы \mathbf{Q} и \mathbf{R} , вектор номеров характеристических признаков, а также вектора минимальных и максимальных

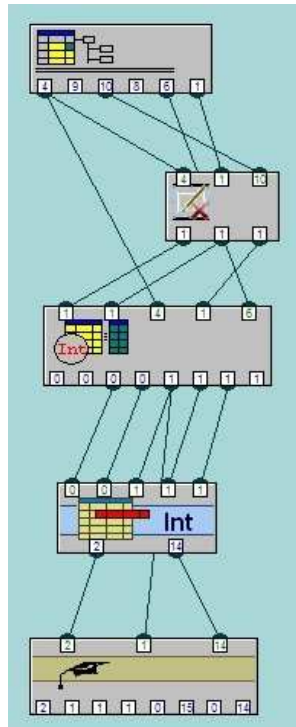


Рис. 3: Интеллектуальная система анализа данных и знаний

значений этих признаков. В результате исполнения модуля формируются выходные значения, содержащие вещественный вектор весовых коэффициентов характеристических признаков и двоичную матрицу U'' , являющуюся входными параметрами для модуля поиска различного рода закономерностей. Выходными параметрами 4-го модуля являются: матрица U'' , неинформативные, обязательные, альтернативные, зависимые характеристические признаки, а также весовые коэффициенты всех характеристических признаков.

Отметим, что в настоящее время не реализована процедура выявления минимальных подмножеств сигнальных признаков.

Исследование интеллектуальной системы

Исследование интеллектуальной системы связано с возможным сокращением признакового пространства. Сокращение признакового пространства позволит сократить число выявляемых значений характеристических признаков для новых исследуемых объектов (обследуемых пациентов), а также сократить временные и стоимостные затраты на исследование и построение диагностических тестов и принятие решений на их основе.

Первые исследования интеллектуальной системы по выявлению различного рода закономерностей по стентированию коронарных артерий описаны в публикации [18]. Исследования проводились по классификационным признакам: наличие окклюзии стента и срок окклюзии стента, наличие рестеноза за все время наблюдения и срок рестеноза, смерть пациента и время смерти.

Система «A multilayer perceptron-based medical decision support system for heart disease diagnosis» [6] включает 40 характеристических признаков и обучающую выборку, состоящую из 352 обучающих объектов, разбитых на 5 классов (используется только один механизм классификации). При этом размер матрицы Q равен 352×40 (14080), а размер матрицы R равен 352×1 . В отличие от вышеупомянутой системы, предлагаемая интеллектуальная система включает обучающую выборку, количество обследуемых в которой

составило 381 человек. Количество характеристических признаков в описании исследуемых объектов (обследуемых) равно 165. Таким образом, размер матрицы Q равен 381×165 (62865). Описание же самого признакового пространства приведено в публикации [18].

В настоящей статье приводятся два исследования по выявлению различного рода закономерностей. В первом исследовании использовались следующие механизмы классификации: 1) наличие окклюзии стента (3 класса); 2) смерть пациента (2 класса). При этом размер матрицы R_1 равен 381×2 (762); Во втором исследовании использовались механизмы классификации: 1) наличие рестеноза за все время наблюдения (3 класса); 2) смерть пациента (2 класса). Размер матрицы R_2 , как и в первом исследовании, равен 381×2 (762). Число образов по каждому из исследований равно 6.

В результате по первому исследованию была построена матрица U'' размером 27352×165 и выявлены следующие закономерности: признак 155 (дата второй повторной операции) является неинформативным; признак 134 (длина второго стента правой коронарной артерии) зависит от признака 139 (длина первого стента диагональной артерии); признак 30 (количество установленных стентов с лекарственным покрытием) зависит от признака 28 (применение баллона со сжатым воздухом); признак 23 (наличие стентирования при прогрессировании атеросклероза) зависит от признака 31 (вид вмешательства). Альтернативных и обязательных признаков в данном исследовании обнаружено не было.

После удаления зависимых и неинформативных признаков признаковое пространство сократилось до 161 признака.

В результате по второму исследованию была построена матрица U'' размером 26149×165 и были выявлены следующие закономерности: признак 155 (дата второй повторной операции) является неинформативным; признак 30 (количество установленных стентов с лекарственным покрытием) зависит от признака 28 (применение баллона со сжатым воздухом); признак 23 (наличие стентирования при прогрессировании атеросклероза) зависит от признака 31 (вид вмешательства); признак 33 (установка стента, покрытого лекарством «сиролimus») зависит от признака 37 (установка стента, покрытого лекарством «эверолимус»). Альтернативных и обязательных признаков в данном исследовании обнаружено не было.

После удаления зависимых и неинформативных признаков, как и в первом исследовании, признаковое пространство сократилось до 161 признака.

Заключение

Дан кратко обзор литературы, посвященный созданным системам по диагностике ишемической болезни сердца. Показано их отличие от предложенной интеллектуальной системы анализа данных и знаний по стентированию коронарных артерий.

Приведены основные этапы интеллектуального анализа данных и знаний по стентированию коронарных артерий, реализованные в интеллектуальной системе.

Описана интеллектуальная система анализа данных и знаний и приведены результаты ее исследования. Результаты исследования показали, что в первом и во втором исследованиях признаковое пространство сократилось на 4 признака, что в ряде случаев существенно снижает временные затраты на построение диагностических тестов и принятие решений на их основе при диагностике рестенозов и окклюзий коронарных артерий.

Применение интеллектуального анализа данных и знаний по стентированию коронарных артерий позволит снизить риски возникновения неблагоприятных сердечнососудистых событий и оптимизировать лечебную тактику при длительном наблюдении за больными, подвергшимся инвазивным вмешательствам на коронарных артериях.

Дальнейшие исследования связаны с реализацией процедуры выявления минимального подмножества сигнальных признаков и алгоритма построения отказоустойчивых тестов [11] в предложенной интеллектуальной системе; с применением подхода к анализу данных и знаний в других проблемных областях, например, для диагностики психических заболеваний [19], а также с созданием гибридной интеллектуальной обучающе-тестирующей системы [20] с применением вышеприведенного подхода к анализу данных и знаний.

Коллектив авторов выражает благодарность профессору, д.м.н., заслуженному деятелю наук РФ, научному руководителю отделения сердечной недостаточности ФГБУ НИИ кардиологии СО РАМН А. Т. Теплякову и к.м.н., старшему научному сотруднику отделения сердечной недостаточности ФГБУ НИИ кардиологии СО РАМН Е. В. Граковой за предоставление данных по стентированию коронарных артерий и консультации по вопросам стентирования.

Литература

- [1] *Загоруйко Н. Г.* Прикладные методы анализа данных и знаний. Новосибирск: ИМ СО РАН, 1999. 270 с.
- [2] *Закревский А. Д.* Логика распознавания. Изд. 2-е, доп. М.: Едиториал УРСС, 2003. 144 с.
- [3] *Финн В. К.* Об интеллектуальном анализе данных // *Новости Искусственного интеллекта*. 2004. № 3. С. 3–18.
- [4] *Янковская А. Е., Гедике А. И.* Интеллектуальный анализ информации на базе инструментального средства ИСМЛОГ // *Новости искусственного интеллекта*. 2005. № 1. С. 36–47.
- [5] *Журавлев Ю. И., Рязанов В. В., Сенько О. В.* Распознавание. Математические методы. Программная система. Практические применения. Москва: Фазис, 2006. 176 с.
- [6] *Hongmei Ya., Yingtao J., Jun Zh., Chenglin P., Qinghui Li.* A multilayer perceptron-based medical decision support system for heart disease diagnosis // *Expert Systems with Applications*. 2005. Vol. 30, no. 2. Pp. 272–281.
- [7] *Kostic P., Vasiljevic Z., Pavlovic S., Milosavljevic I., Milanovic J. G., Blesic S., Milanovic S.* Knowledge management system for clinical decision support — application in cardiology // *Proceedings of 19th Telecommunications Forum (TELFOR)*. 2011. Pp. 1261–1264.
- [8] *Pal D., Mandana K. M., Pal S., Sarkar D., Chakraborty C.* Fuzzy expert system approach for coronary artery disease screening using clinical parameters // *Knowledge-Based Systems*. 2012. Vol. 36. Pp. 162–174.
- [9] *Янковская А. Е.* Логические тесты и средства когнитивной графики. Издательский Дом: LAP LAMBERT Academic Publishing, 2011. 92 с.
- [10] *Янковская А. Е., Петелин А. Е.* Развитие алгоритма многокритериального выбора оптимального подмножества диагностических тестов // *Математические методы распознавания образов (ММРО-14). Сб. докл. 14-й Всеросс. конф.* М.: МАКС Пресс, 2009. С. 212–215.
- [11] *Yankovskaya A. E., Kitler S. V.* Parallel algorithm for constructing k -valued fault-tolerant diagnostic tests in intelligent systems // *Pattern Recognition Image Anal.* 2012. Vol. 22, no. 3. Pp. 473–482.
- [12] *Кренделев Ф. П., Дмитриев А. Н., Журавлев Ю. И.* Сравнение геологического строения зарубежных месторождений докембрийских конгломератов с помощью дискретной математики // *ДАН СССР*. 1967. Т. 173. № 5. С. 1149–1152.
- [13] *Yankovskaya A. E.* New kinds of regularities in knowledge and algorithms of their revealing // *7th Open German/Russian Workshop on Pattern Recognition and Image Understanding*. Ettlingen, Germany. 2007.

- [14] Янковская А. Е., Китлер С. В. Принятие решений на основе параллельных алгоритмов тестового распознавания образов // *Искусственный интеллект*. 2010. № 3. С. 151–159.
- [15] Yankovskaya A. E., Semenov M. E. To the problem about the intelligent extension construction of the geoinformational systems // *Pattern Recognition and Image Understanding (OGRW-8-11). Proceedings of 8th Open German-Russian Workshop*. Nizhny Novgorod: Nizhny Novgorod Lobachevsky State University, 2011. Pp. 349–352.
- [16] Гедике А. И., Янковская А. Е. Построение всех безызбыточных безусловных диагностических тестов в интеллектуальном инструментальном средстве ИМСЛОГ // *Интеллектуальные системы (AIS'05), Интеллектуальные САПР (CAD-2005). Тр. междунар. научно-технических конф.* Москва: Физматлит, 2005. Т. 1. С. 209–214.
- [17] Yankovskaya A. E., Gedike A. I., Ametov R. V., Bleikher A. M. IMSLOG-2002 software tool for supporting information technologies of test pattern recognition // *Pattern Recognition Image Anal.* 2003. Vol. 13. no. 4. Pp. 650–657.
- [18] Янковская А. Е., Китлер С. В., Аметов Р. В., Гракова Е. В. Информационная технология выявления закономерностей в данных и знаниях по стентированию коронарных артерий // *Тр. 5-й междунар. конф. «Системный анализ и информационные технологии» САИТ-2013.* (в печати).
- [19] Янковская А. Е., Китлер С. В., Изюмов А. А., Кривдюк Н. М., Силаева А. В. Основы создания комплекса интеллектуальных систем диагностики и профилактики депрессии // *Приложение к журналу «Вестник Московского университета имени С. Ю. Витте. Серия 1: Экономика и управление»*. 2013. С. 81–84.
- [20] Yankovskaya A. E., Semenov M. E. Application mixed diagnostic tests in blended education and training // *Proceedings of the IASTED International Conference Web-based Education (WBE 2013)*. Innsbruck, Austria. 2013. Pp. 935–939.