УДК 004.9+303.732

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ ПОДДЕРЖКА АВТОМАТИЗИРОВАННОЙ КЛАССИФИКАЦИИ НА ОСНОВЕ ПРЯМЫХ И ОБРАТНЫХ РЕШАЮЩИХ ПРАВИЛ НА ПРИМЕРЕ ПРОГНОЗА ТРОМБОЭМБОЛИИ

Артеменко М.В., Добровольский И.И., Мишустин В.Н.

ФГБОУ ВО «Юго-западный государственный университет», Курск, e-mail: kstu-bmi@yandex.ru, artem1962@mail.ru

Целью работы являлось рассмотрение вопросов разработки структуры информационно-аналитической модели процесса поддержки принятия решений классификационных задач при анализе и управлении состоянием объекта или процесса на основе одновременного применения решающих правил, позволяющих: как по значениям характеристических признаков диагностировать объект исследования, так и по предполагаемой принадлежности объекта к определенному классу определять набор признаков, обладающий наилучшими классификационными возможностями, Рассматривается методика оценки риска принятия классификационных решений в случае применения балльной метрической шкалы задания признаков. Приведен пример набора признаков для прогноза тромбоэмболии и базового интерфейса соответствующей автоматизированной диагностической системы.

Ключевые слова: решение классификационных задач, автоматизированные системы поддержки принятия решений, оценка риска принятия решений, тромбоэмболия

INFORMATION AND ANALYTICAL SUPPORT AUTOMATED CLASSIFICATION BASED ON THE DIRECT AND REVERSE THE DECISION RULE ON THE EXAMPLE OF FORECAST OF THE THROMBOEMBOLISM

Artemenko M.V., Dobrovolsky I.I., Mishustin V.N.

FSEI HT «Southwestern State University», Kursk, e-mail: kstu-bmi@yandex.ru, artem1962@mail.ru

The purpose of work was consideration of questions of development of structure of information and analytical model of process of support of decision-making of classification tasks in the analysis and management of a condition of object or process on the basis of simultaneous application of the decisive rules allowing: as on values of characteristic signs to diagnose object of research, and to determine by estimated belonging of object to a certain class most the feature set possessing the best classification opportunities the technique of an assessment of risk of adoption of classification decisions in case of application of a ball metric scale of a task of signs is considered. The example of a feature set for the forecast of a tromboemboliya and the basic interface of the relevant automated diagnostic system is given.

Keywords: decision classification tasks, automated control systems decision support, risk assessment decision making, tromboemboliya

Автоматизация процесса поддержки принятия решений позволяет лицу, принимающему решение (ЛПР), формировать наиболее оптимальные и приемлемые рекомендации для качественного управления объектом или процессом управления в условиях слабоструктурированной информации, многоальтернативности вариантов решений и многокритериальности при оптимизации. Теоретический базис принятия решений в настоящее время достаточно хорошо проработан – см., например, работы Орлова А.И. (в частности, [8]). Особую важность приобретает применение экспертных систем в составе автоматизированных систем поддержки принятия решений (АСППР) для решения задач, у которых высок риск принятия неадекватного решения, последствия которых могут привести как к уничтожению объекта, так и самой системы управления (кардиология, онкология, хирургия, энергетические объекты и системы, химические объекты и т.п.).

АСППР, как правило, имеет в своем составе базу знаний, позволяющую по определенным решающим правилам классифицировать объект или процесс управления к определенному кластеру состояний, для которого имеется некоторые множество корректирующих или управляющих воздействий, позволяющих управлять объектом как внутри кластера так и переводить его в другой кластер. Поскольку любой объект или процесс для ЛПР характеризуется определенным набором управляемых и управляющих характеризующих его признаков, то решающее правило в конечном итоге сводится к конструкции «ЕСЛИ $F(\{x\})$ соответствует множеству условий $\{c\}$, ТО состояние объекта относится к классу G с уверенностью P и к нему могут быть применены воздействия из множества $\{u\}$» (где x – регистрируемые признаки, c – возможные условия, u – возможные корректирующие и управляющие воздействия, $F(\{x\})$ –

функционалы, позволяющие применить к признакам x операции сравнения). Формирование данной конструкции основывается на методах теории распознавания образов, теории принятия решений, оценки риска, теории нечетких множеств, теории и инструментальных средств искусственного интеллекта. Назовем данную конструкцию «прямым решающим правилом».

В случае наличия большого количества признаков и многоальтернативности решений имеются решающие правила с конструкцией типа: «ЕСЛИ состояние объекта относится к классу G , ТО у него с уверенностью q должны присутствовать признаки $\{x\}$ с определенными значениями (диапазоном значений) $\{RV\}$ и оно управляемо воздействиями из множества $\{u\}$». АСППР с подобными конструкциями нашли меньшее применение. Назовем данную конструкцию как «обратное решающее правило».

Практически все АСППР используют либо прямые, либо обратные решающие правила (например, медицинские АСППР – [3]). Между тем, возникает проблема одновременного применения данных конструкций решающих правил с целью повышения надежности в адекватности рекомендаций СППР ЛПР и снижения риска принятия неправильного решения с катастрофическими последствиями для объекта или процесса управления.

Как отмечалось, к задачам, к которым предъявляются особые требования к качеству управляющих воздействий, относятся связанные в конечном итоге с жизнью человека. Среди медицинских проблем – это снижение смертности от сердечно-сосудистых заболеваний (первое место), травм (второе место), сахарного диабета. Смертность от травм связана с ростом количества венозных тромбоэмболических осложнений. Несмотря на это, в силу множества слабоструктурированных признаков, присутствие которых может говорить о высокой степени риска возникновения тромбов, решение ЛПР (врачом) принимается в настоящее время в основном на основе собственного опыта, опыта коллег и информационных источников без применения АСППР. В работах [9, 10] рассматривается экспертная система на основе прямых решающих правил, построенных с применением нечеткого решающего правила анализа лингвистической переменной, сформированной по определенным признакам.

В связи с этим, приобретает актуальность разработка методов и средств информационно-аналитической поддержки АСППР, позволяющих при формировании рекомендаций ЛПР одновременно применять прямые и обратные решающие правила.

Цель исследования

Разработка модели и метода информационно-аналитического повышения надежности решений, вырабатываемых АСППР при классификации объектов или процессов управления в условиях повышенных требований к адекватности сформированных рекомендаций ЛПР и слабоструктурированных исходных данных.

В процессе исследования решались задачи:

– разработка структуры информационно-аналитической модели принятия решений на основе одновременного применения решающих правил, позволяющих как по значениям признаков идентифицировать принадлежность объекта или процесса управления к определенному классу, так и по гипотезе о принадлежности к определенному классу осуществлять проверку наличия у объекта или процесса информационных характеристических признаков с приемлемыми классификационными возможностями;

– разработка метода оценки риска принятия классификационного решения при применении балльной метрики значений характеризующих признаков;

– формирование набора признаков, позволяющих прогнозировать тромбоэмболию при травмах и эндопротезировании на основе информационно-аналитической модели АСППР;

– формирование предложений по интерфейсу ЛПР и АСППР прогноза тромбоэмболизма.

Для реализации цели и решения исследовательских задач применялись следующие материалы и методы: теория принятия решений, системный анализ, методы синтеза решающих правил для АСППР, прикладной статистический анализ в медицине [6], теория проектирования и исследования систем автономного искусственного интеллекта [4], клинические данные прогноза развития тромбоэмболии.

Результаты исследования и их обсуждение

Для формирования заключения о принадлежности объекта или процесса управления к определенному АСППР имеет в своем составе набор (множество) признаков и базу знаний, состоящую из решающих правил, сформированных в процессе настройки АСППР на обучающей выборке и значения показателей качества (например, диагностические: чувствительность – ДЧ, специфичность – ДС, эффективность – ДЭ и т.п.), полученные на экзаменационной выборке.

Поскольку при решении практических классификационных задач часть признаков в ряде случаев может отсутствовать, то необходимо по возможности базу знаний заполнить решающими правилами всех возможных реализаций множества признаков ($\Omega_k, k = 1,...K$, где K – количество возможных реализаций). Обозначим элементы указанного множества прямых решающих правил как $P_i(\Omega_k)$ и обратных решающих правил как $P_j^{-1}(\omega_l)$ (ω_l – идентификатор класса l, $j = \overline{1,J}, l = \overline{1,L}$, где J – количество обратных решающих правил, L – количество классов состояний объекта). Для $P_i(\Omega_k)$ и $P_j^{-1}(\omega_l)$ в базе знаний формируются соответствующие кортежи показателей качества $Q_i(P_i(\Omega_k))$ и $Q_j(P_j^{-1}(\omega_l))$. В общем случае кортежи представляют собой наборы значений, применяемых в АСППР показателей качества (критериев адекватности и/или приемлемости) рекомендуемых решений согласно применения прямых и обратных решающих правил. Выбор решения по кортежам осуществляется по алгоритмам, реализованным в специализированной АСППР, согласно принципам оценки риска в многокритериальных задачах [5].

Процесс классификации в случае применения прямых и обратных решающих правил, существенно повышающий надежность в адекватности выбора необходимых корректирующих и/или управляющих воздействий, что особенно актуально в задачах с повышенным риском последствий и/или несвоевременности принятия решения, предлагается осуществлять одновременно по алгоритмам, реализующим следующие направления работы АСППР:

1. По зарегистрированным значениям известного для АСППР наблюдаемого набора признаков путем применения прямых решающих правил анализируемый объект или процесс управления соотносится к определенному классу ω_{l1} (в смысле оптимальности согласно кортежам показателей качества $Q_i(P_i(\Omega_k))$), для которого определены управляющие и корректирующие воздействия.

2. Предполагается принадлежность объекта к определенному классу ω_z. Применяется обратное решающее правило, результатом применения которого являются варианты наборов признаков Ω_{k1} с соответствующими показателями качества $Q_j(P_j^{-1}(\omega_z))$. Если признаки регистрируются у объекта или процесса управления, то осуществляется вычисление их значений согласно принятым в АСППР метрикам и правилам и, реализуя прямые решающие правила $P_{i1}(\Omega_{k1})$, осуществляется соотнесение объекта к классу состояний ω_{l2}. Если классы ω_z и ω_{l2} совпадают, то принимается вариант ω_{l2} и процесс классификации заканчивается. Если классы ω_z и ω_{l2} не совпадают, то в качестве ω_z принимается гипотеза о принадлежности объекта к классу ω_{l2} и процесс классификации по обратным решающим правилам продолжается со сменой набора признаков Ω_{k1} до тех пор пока не будут учтены все находящиеся в АСППР реализуемые наборы признаков с приемлемыми показателями качества $Q_j(P_j^{-1}(\omega_z))$ или классы ω_z и ω_{l2} совпадут.

В процессе одновременной реализации рассмотренных направлений АСППР возможно возникновение ситуаций:

1. АСППР не удалось соотнести объект к классу ω_{l1} (прямое решающее правило).

2. АСППР не удалось соотнести к классу ω_{l2} (обратное решающее правило).

3. АСППР соотнесло объект к классу ω_{l1} и ω_{l2}, причем ω_{l1} и ω_{l2} совпадают ($\omega_{l1} = \omega_{l2}$).

4. АСППР соотнесло объект к классу ω_{l1} и ω_{l2}, причем ω_{l1} и ω_{l2} не совпадают ($\omega_{l1} \neq \omega_{l2}$).

Предлагаемые действия ЛПР при различных сочетаниях данных ситуаций представлены в таблице.

Действия ЛПР при различных вариантах классификации объекта АСППР

№ ситуации	ω_{l1}	ω_{l2}	$\omega_{l1} \neq \omega_{l2}$	$\omega_{l1} = \omega_{l2}$	Действия ЛПР
1	+	+	–	+	Согласно принадлежности объекта к классу ω_{l1}
2	+	+	+	–	Дополнительное исследование объекта и/или изменение набора признаков $\Omega_k, k = 1,...K$ и/или изменение прямых и/или обратных решающих правил
3	–	+	–	–	Согласно принадлежности объекта к классу ω_{l2} и/или изменение прямых решающих правил $P_i(\Omega_k)$
4	–	–	–	–	Дополнительное исследование объекта и-или изменение набора признаков $\Omega_k, k = 1,...K$ и/или изменение прямых и/или обратных решающих правил
5	+	–	–	–	Согласно принадлежности объекта к классу ω_{l2} или изменение прямых обратных решающих правил $P_j^{-1}(\omega_l)$

Риск принадлежности объекта или процесса управления к определенному классу ω_l предлагается определять следующим образом.

1. Определяем риск по кортежам Q_{ϖ_l} (формула (1):

$$Q_{\varpi_l}(P_i(\Omega_k), P_j^{-1}(\omega_l)) = Q_{\varpi_l}(P_i(\Omega_k)) +$$

$$+ Q_{\varpi_l}(P_j^{-1}(\omega_l)) - Q_{\varpi_l}(P_i(\Omega_k)) \bullet Q_{\varpi_l}(P_j^{-1}(\omega_l)). \quad (1)$$

2. Определяем риск по «баллам важности» $B(x)$ используемых при классификации признаков (значения $B(x)$ рекомендуется шкалировать в одном диапазоне – например, от 0 до 10 по мере возрастания информационной значимости), которые определяются как экспертным путем, так и статистически (согласно положениям доказательной медицины [6]). Обозначим этот риск как $Risk(B(x))$. Признаки предварительно группируются по семантической нагрузке, образуя G групп, и для каждой группы задаются значения частных (групповых) рисков $Rg_m, m = \overline{1, M_g}$, где M_g – количество выделенных семантических групп. Значения Rg_m определяются как экспертным путем, так и статистически – например, как частота встречаемости у объекта в классе в информативном доверительном интервале. Далее аналогично методике, изложенной в [1], оцениваются значения риска принадлежности к классу ω_l по формулам (2) и (3).

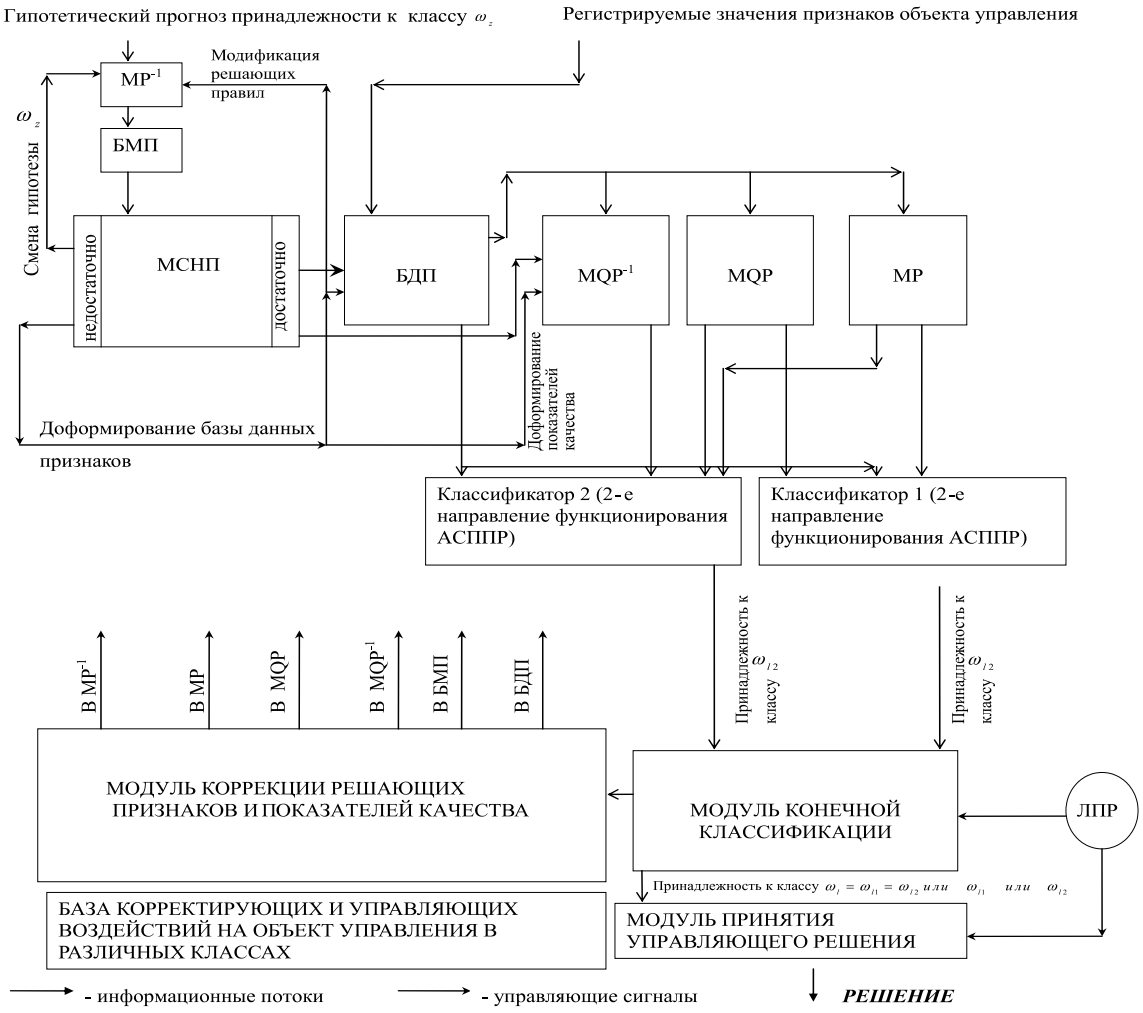

Рис. 1. Структура информационно-аналитической модели принятия решения АСППР по прямым и обратным решающим правилам. Где: МР-1 – модуль обратных решающих правил, МР – модуль прямых решающих правил, МQР-1 – модуль множества кортежей показателей качества для обратных решающих правил, МQР – модуль множества кортежей показателей качества для прямых решающих правил, БМП – база множества признаков, БДП – база данных признаков, МСНП – модуль сравнения наличия признаков, ЛПР – лицо, принимающее решение

$$ORg_m = 1 - \prod_{j1=1}^{J_i} (1 - \frac{B_{m.j1}}{\max(B_{m,j1} / j1 = 1...J_m)} \cdot Rg_m), \quad (2)$$

где J_m – количество учитываемых показателей в m-ой группе, $B_{m,j1}$ – балл, присвоенный j1-му признаку в m-ой группе, ORg_m – итого вое значение риска в группе показателей m (с учетом всех баллов в группе).

$$Risk(B(x)) = 1 - \prod_{m=1}^{M_g} (1 - ORg_m). \quad (3)$$

3. Определяем итоговое значение риска принадлежности объекта к классу ω_l – $Risk_{\varpi_l}$ – по формуле (4):

$$Risk_{\varpi_l} = Risk(B(x)) \bullet Q_{\varpi_l}(P_i(\Omega_k), P_j^{-1}(\omega_l)). \quad (4)$$

Как отмечалось, АСППР при решении классификационных задач в настоящее время находят широкое применение в медицине в диагностическом процессе при наличии большой размерности признакового пространства и высокой ответственности ЛПР (врача) за последствия реализации неправильного решения на основе неадекватной диагностики (классификации состояния человека).

Структура информационно-аналитической модели принятия классификационного решения АСППР согласно предлагаемому методу представлена на рис. 1.

Возможности предложенного подхода к построению АСППР были исследованы для прогноза тромбоэмболии при травмах и эндопротезировании.

В вычислительном эксперименте использовался опыт лечения 400 больных с травмами опорно-двигательной системы, находившихся на стационарном лечении в 1 и 2 травматологическом отделениях городской клинической больницы № 4 г. Курска, травматологическом отделении Курской областной клинической больницы, хирургическом отделении городской клинической больницы № 2 г. Курска, отделении сосудистой хирургии Курской областной клинической больницы, травматологическом отделении Орловской областной клинической больницы, травматологическом отделении городской клинической больницы № 1 г. Белгорода. В исследовании анализировалась информация о состоянии 490 человек с выделением групп [7]:

1 группа – больные с травмой, течение которой осложнилось развитием флеботромбозов различной локализации (200 больных с тромбозом глубоких вен (ТГВ) без тромбоэмболии легочной артерии (ТЭЛА) и 100 больных с ТГВ, осложненной ТЭЛА). Контрольная группа – больные с травмой, у которых возникновение глубокого венозного тромбоза не отмечено.

2 группа – больные с травмой, течение которой осложнилось развитием ТЭЛА (100 больных). Контрольная группа – больные с травмой, течение которой осложнилось возникновением флеботромбоза без развития ТЭЛА.

3 группа – здоровые добровольцы, у которых проводилось измерение давления внутри мышечно-фасциальных футляров голени и измерение объема голени с помощью программного аппаратно-диагностического комплекса [9] (90 человек).

Среди рассмотренных 400 больных: мужчин – 208 (52%), женщин 192 (48%), 308 (77%) обследованных старше 40 лет, умерло 133 пациента (33%).

В ходе анализа статистического материала, литературных источников и собственного опыта выявлено 159 признаков для диагностики возможного прогноза тромбоэмболии (эмболический тромбоз, неэмболический тромбоз, тип ТЭЛА, форма ТЭЛА, исход – смерть, выздоровление), на основании которых предлагается базовый интерфейс ЛПР с АСППР – представлен на рис. 2.

Признаки объединяются в 26 групп: ассоциирующие с больным факторы риска – 14 признаков, тип повреждения – 6 , кость – 13, ушибы сегмента ОДС – 8 , грудная клетка – 8, позвоночник – 7, растяжение связок или разрывов – 2, вывихи – 10, черепно-мозговая травма – 4, травма органов живота – 5, ожоги – 4, ЭКТ – 2, метод фиксации – 9, классификация ТЭЛА – 6, локализация тромба – 17, тип обезболивания – 4, метод репозиции – 5, местные осложнения – 6, анемия – 1, пробы – 6, давление – 3, нарушение венозной гемодинамики – 5, заинтересованный компартмент – 1.

В процессе проведения исследований диагностических возможностей АССПР были синтезированы нечеткие решающие правила прямого типа [7, 10], показавшие следующие значения показателей качества диагностики для:

– неэмбологенного тромбоза: ДЧ = 0,995, ДС = 0,76;

– эмбологенный тромбоз: ДЧ = 0,71, ДС = 0,97;

– ТЭЛА: ДЧ = 0,84, ДС = 0,97;

– летальный исход ТЭЛА: ДЧ = 0,94, ДС = 0,93;

– благоприятный исход ТЭЛА: ДЧ = 0,65, ДС + 0,98.

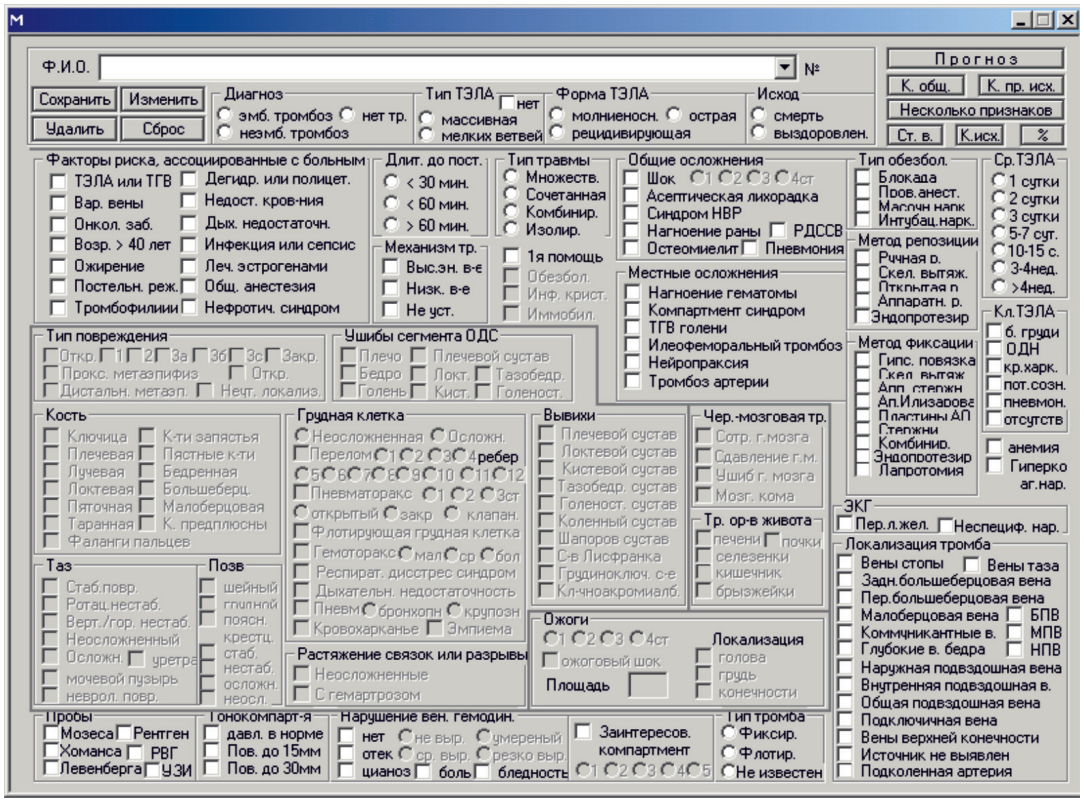

Рис. 2. Базовый интерфейс АС ППР прогноза тромбоэмболии при травмах и эндопротезировании

Примечание: результат по обратным решающим правилам совпал с диагностикой по прямым – см. таблицу (ситуация 1).

Анализ полученной информации показывает, что при посттравматическом тромбозе и ТЭЛА АСППР позволяет ЛПР с приемлемыми значениями диагностических чувствительности и специфичности проводить прогнозирование как ТЭЛА, так и ее исходов у больных с травмой.

Заключение

В процессе проведения исследований:

– разработана структура информационно-аналитической модели функционирования АСППР при работе со слабоструктурированными данными, отличающаяся одновременным применением прямых и обратных классификационных решающих правил, что позволяет повысить надежность рекомендаций ЛПР в принятии определенного решения по управлению объектом или процессом;

– рассмотрены различные варианты действий ЛПР в случае возникновения противоречий при применении прямых и обратных решающих правил, позволяющие оптимизировать его взаимодействие с АСППР и проведение дальнейших исследований;

– предложена методика расчета риска принадлежности объекта к определенному классу, отличающаяся учетом рисков применения прямых и обратных решающих правил и применения определенного набора признаков, представленных в балльной метрике, что позволяет учесть возникающие на различных этапах классификационного процесса риски;

– предложен набор 159 признаков для прогноза тромбоэмболизма с систематизацией по 26 группам, что позволяет оптимизировать осмотр пациента;

– разработано диалоговое окно внешнего интерфейса ЛПР с АСППР;

– исследованы возможности АСППР при различных вариантах тромбоза и ТЭСЛА – получены приемлемые для клинического применения результаты.

Полученные результаты (предлагаемые принципы функционирования АСППР) позволяют повысить надежность принимаемых решений в системе ЛПР-АСППР в условиях слабоструктурированных данных за счет получения непротиворечивой приемлемой классификации объекта и процесса управления путем одновременного применения прямых и обратных классификационных решающих правил и оценки риска формируемых АСППР рекомендаций с учетом

риска применения как самих правил, так и набора признаков объекта, представленных в балльной метрике.

Список литературы

1. Артеменко М.В. Синтез диагностических правил заболевания в условиях массового обследования населения / Медико-экологические информационные технологии – 2015: сб. мат. XVIII Междун.начн.-техн. Конф. (Курск, 20–22 мая 2015). – Курск, 2015. – С. 126–130.

2. Башмаков А.И., Башмаков И.А. Интеллектуальные информационные технологии: учеб. пособие. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2005. – 304 с.

3. Воронцов И.М., Шаповалов В.В., Шерстюк Ю.М. Здоровье. Опыт разработки и обоснование применения автоматизированных систем для мониторинга и скринирующей диагностики нарушений здоровья. – СПб.: ООО «ИПК «Коста» Б, 2006. – 432 с.

4. Жданов А.А. Автономный искусственный интеллект. – М.: БИНОМ. Лаборатория знаний, 2009. – 359 с.

5. Жуковский В.И, Жуковская Л.В. Риск в многокритериальных и конфликтных системах при неопределенности. – М.: Изд-во ЛКИ., 2010. – 272 с.

6. Клюшин Д.А., Петунин Ю.И. Доказательная медицина. Применение статистических методов. – М.: ООО «И.Д. Вильямс», 2008. – 320 с.

7. Мишустин В.Н. Глубокие венозные тромбозы и тромбоэмболия легочной артерии: диагностика, прогнозирование рисков, лечение в раннем периоде травматической болезни: дис…доктора мед. наук. – Курск. 2000. – С. 162–253.

8. Орлов А.И. Теория принятия решений: учебник. – М.: Изд-во «Экзамен», 2006. – 573 с.

9. Распознающий аппаратно-программный диагностирующий комплекс // Приборостроение. – 2005. – № 2. – С. 22–25.

10. Труфанов М.И. Автоматизированное прогнозирование тромбоэмболии легочной артерии на ранней стадии на основе нечеткой логики // Системный анализ и управление в биомедицинских системах. – 2004. – № 3. – С. 204–206.