

РАЗРАБОТКА СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ ДЛЯ ВЫБОРА ТРАЕКТОРИИ ЛЕЧЕНИЯ ДЕТЕЙ С ИЗБЫТОЧНЫМ ВЕСОМ

Марухина Ольга Владимировна

К.т.н., доцент, e-mail: Marukhina@tpu.ru

Берестнева Ольга Григорьевна

Д.т.н., профессор, e-mail: ogb6@yandex.ru

Отделение информационных технологий,

Инженерная школа информационных технологий и робототехники

ФГАОУ ВО НИ Томский политехнический университет, 634050 г. Томск, пр. Ленина 30

Аннотация. Предложен подход к разработке алгоритмической базы системы поддержки принятия врачебных решений при выборе траектории лечения детей с избыточным весом. Проведено исследование структуры пропущенных данных клинико-лабораторных показателей, осуществлено восстановление пропущенных значений методом множественного восстановления и проведено сокращение размерности с использованием пакетов языка R. На основе полученных результатов построена модель принятия решений и определен набор правил для выбора траектории лечения. В результате, сформирован алгоритм обработки данных для принятия решения о выборе группы лечения детей с избыточным весом.

Ключевые слова: многомерные данные, анализ данных, система поддержки принятия врачебных решений, восстановление данных, снижение размерности, логические правила.

Цитирование: Марухина О.В., Берестнева О.Г. Разработка системы поддержки принятия решения для выбора траектории лечения детей с избыточным весом // Информационные и математические технологии в науке и управлении. 2019. № 3 (15). С. 42–52. DOI: 10.25729/2413-0133-2019-3-04

Введение. В настоящее время многие учреждения здравоохранения разного масштаба и профилей во всем мире широко применяют автоматизацию медицинских технологий и соответствующих бизнес-процессов. Для этого используются медицинские информационные системы. Одним из важных направлений является разработка систем поддержки принятия решений (СППР). СППР используют для диагностики и прогнозирования заболеваний. Подобные системы не могут нести ответственность за принятые с ее помощью решения, но способны существенно упростить и облегчить работу врача. Цель СППР заключается в осуществлении кооперации системы и человека в процессе принятия решений.

Анализ научных публикаций по данной тематике показал, что исследования и разработки в области СППР ведутся по всему миру и в различных направлениях. Так, например, авторами [1] представлена система поддержки принятия врачебных решений (СППВР) для использования в области кардиологии. Система предназначена для выбора клинических решений по восьми основным заболеваниям сердечно-сосудистой системы. Позднее на основе использования системы был проведен анализ проблем, возникающих при внедрении технологий персональной телемедицины [6]. Предложены пути преодоления

существующих препятствий и повышения эффективности телемедицинских технологий для управления здоровьем, профилактики и реабилитации сердечно-сосудистых заболеваний. В исследовании на базе Башкирского государственного медицинского университета [2, 3] были проанализированы возможности современных СППР в практике врача-хирурга. Исследования группы авторов [5] направлены на разработку архитектуры базы знаний СППР для инструментальной диагностики стенокардии, в основе которой лежит онтологический подход с использованием графовой базы данных. Ведутся разработки врачебных систем для выбора тактики лечения пациентов с ишемической болезнью сердца, а также в области дерматовенерологии [6, 8-11].

Актуальность разработки систем поддержки принятия решений обусловлена возникновением врачебных ошибок, которые могут возникать по разным причинам. Среди них ограниченность временных ресурсов, недостаточность знаний, отсутствие возможности привлечения компетентных экспертов, неполнота информации о состоянии пациента. Указанные факторы могут привести к серьезным проблемам со здоровьем в будущем. В связи с этим, общей задачей для всех СППР становится принятие «персонализированного» решения о траектории лечения каждого пациента. Новизна исследования заключается в создании нового алгоритма определения индивидуальной траектории лечения пациентов на основе анализа клинико-лабораторных показателей.

1. Постановка задачи. Избыточный вес и ожирение у детей и подростков является одной из самых серьезных проблем общественного здравоохранения в XXI веке. Согласно данным Всемирной организации здравоохранения в 2016 году около 41 миллиона детей в возрасте до 5 лет имели избыточный вес или ожирение. С 1975 по 2016 гг. доля избыточного веса и ожирения среди детей и подростков в возрасте от 5 до 19 лет возросла с 4% до 18% и составило 124 миллионов человек. От последствий избыточного веса и ожирения умирает больше людей, чем от последствий аномально низкой массы тела. Ежегодно в результате избыточного веса или ожирения умирает в среднем 2,6 миллиона человек [15].

В рамках сотрудничества с научно-исследовательским институтом курортологии и физиотерапии г. Томска (ТНИИКиФ) был получен набор данных клинического исследования, проведенного на базе детского отделения. Данные представлены в виде многомерной таблицы и содержат значения групп клинико-лабораторных показателей пациентов. Пациентами являются дети и подростки в возрасте от 7 до 18 лет с избыточным весом. Набор данных содержит 345 объектов (пациентов), 160 показателей (значения до и после лечения) и 5 групп (траекторий лечения). Помимо вышеперечисленных групп показателей, массив данных содержит информацию о принадлежности пациента к группе лечения (1-5).

Разрабатываемая система предназначена для поддержки принятия врачебного решения о назначении пациенту наиболее подходящей для него траектории лечения на основе анализа значений его первичных клинико-лабораторных показателей.

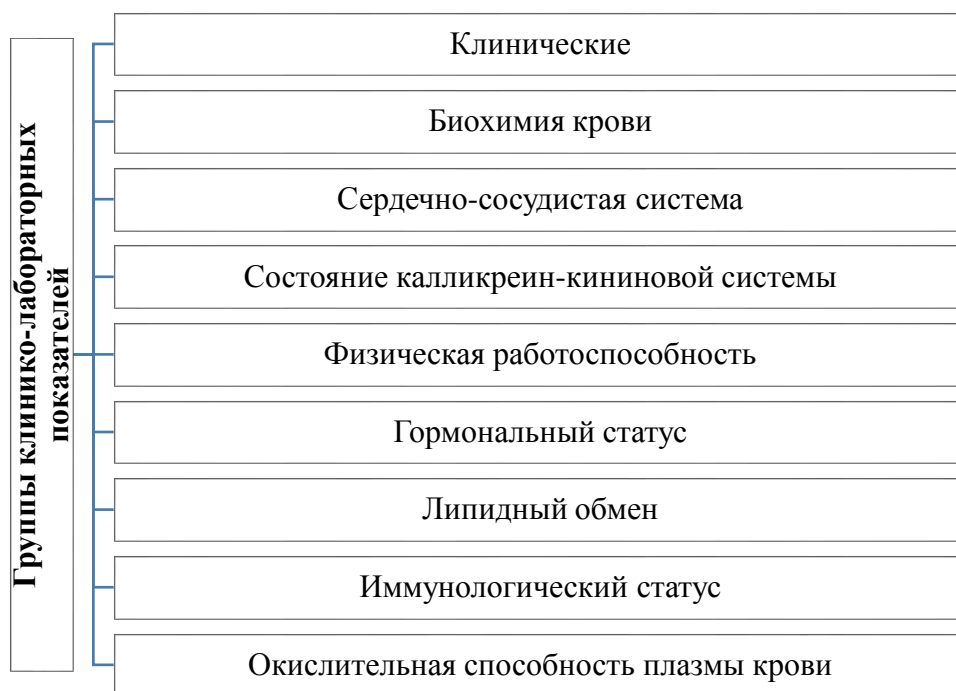


Рис. 1. Группы клиничко-лабораторных показателей

2. Первичная обработка данных. Первичная обработка данных направлена на упорядочивание информации об исследуемом объекте: упорядочивание исходных данных; обнаружение и устранение ошибок, некорректно введенных данных, выбросов, пробелов в сведениях; выявление скрытых закономерностей и связей.

Данные исходного набора очень разрежены, в частности, некоторые «признаки» (столбцы таблицы данных) являются полностью пустыми. Подобные признаки были исключены из рабочей выборки.

Для определения выбросов по каждому показателю были использованы диаграммы размахов (ящики «с усами»), пример такой диаграммы представлен на рис. 2.

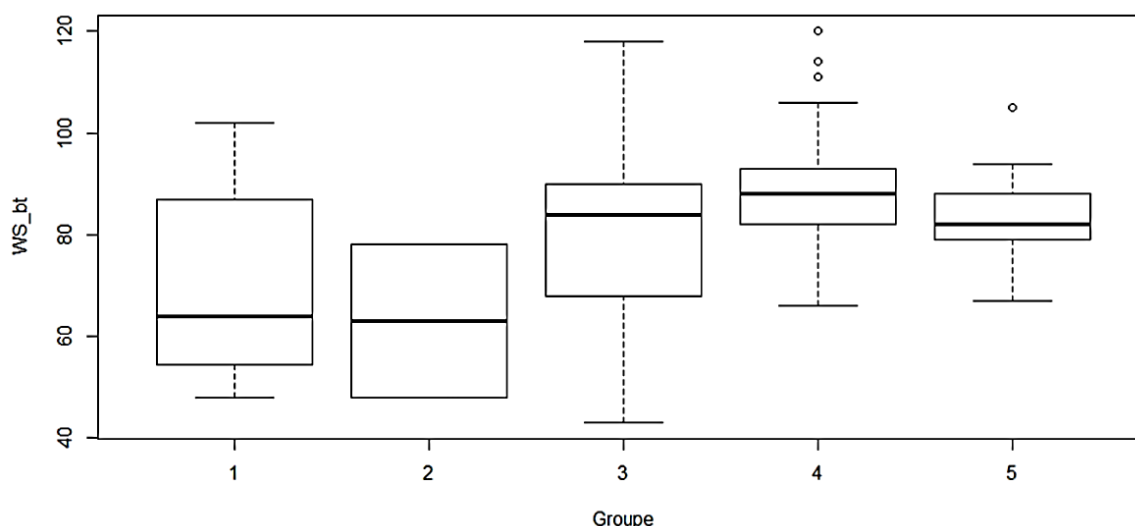


Рис. 2. Диаграмма размаха для показателя *Объем талии до лечения*

Как видно из рисунка, выбросы наблюдаются в четвертой и пятой группах, но существенным выбросом будем считать только максимальное значение 120, остальные

значения выбросов лежат в пределах значений других групп, поэтому могут быть допустимыми. При первичной обработке данных были проанализированы различные случаи наличия и обработки выбросов. Попадались показатели без выбросов, с множеством выбросов, с одним или двумя выбросами. Среди показателей набора данных были и такие, выбросы в которых были удалены полностью, т.к. не попадали в предел допустимых значений в каждой из рассматриваемых групп лечения.

2.1. Исследование и восстановление пропущенных значений. Большое количество статистических методов предполагает, что в ходе наблюдений были получены полностью укомплектованные наборы данных, но на практике пропуски в данных все же являются повсеместным явлением. Прежде чем перейти к анализу данных и построению визуальных образов, необходимо привести обрабатываемый набор данных к "каноническому" виду, т.е. либо удалить фрагменты объектов с недостающими элементами, либо заменить имеющиеся пропуски на некоторые разумные значения. Процедура "борьбы с пропусками" обычно включает следующие шаги:

1. Идентификация недостающих данных.
2. Исследование закономерностей появления отсутствующих значений.
3. Формирование наборов данных, не содержащих пропуски (в результате удаления или замены соответствующих фрагментов).

Идентификация недостающих данных является единственным однозначным шагом. Анализ пропущенных данных зависит от понимания процессов, которые воспроизводят экспериментальную информацию. Решение о способе устранения пропущенных значений также будет зависеть от оценки того, какие процедуры приведут к самым надежным и точным результатам. Идентификация пропущенных значений в пакете языка R обычно производится с использованием функций *is.na()* и *complete.cases()*. В исходном наборе данных было 0 полностью укомплектованных строк и 276 строк, имеющих хотя бы одно пропущенное значение.

Функция *aggr()* графически отображает число наблюдений для каждой отдельной переменной и для каждой комбинации переменных (рис. 3).

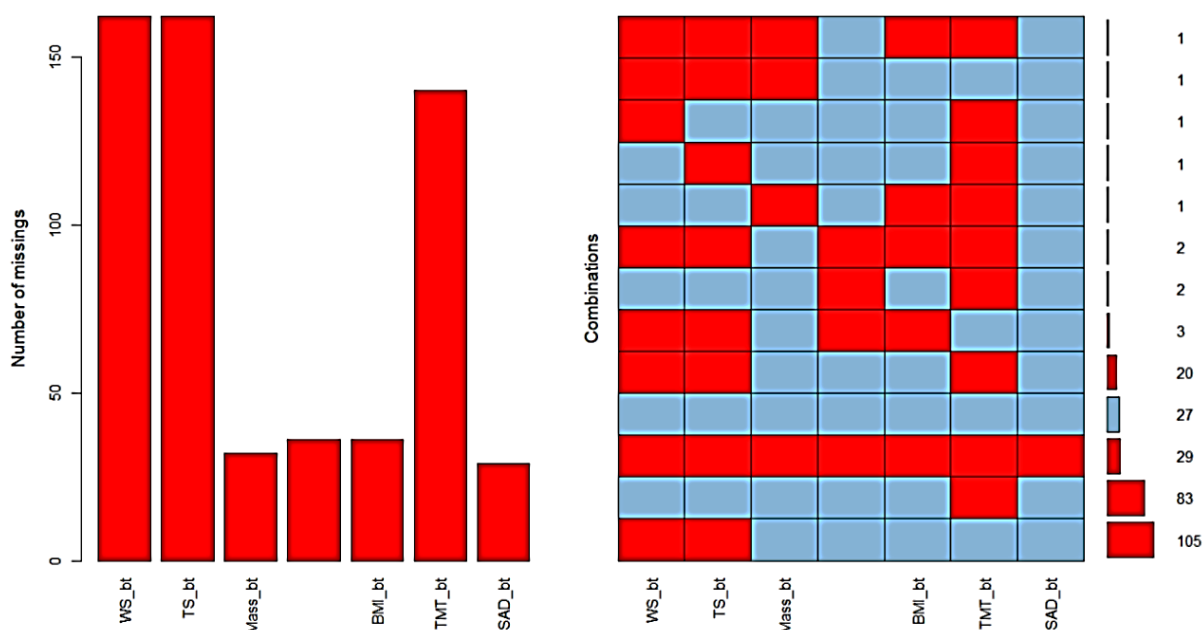


Рис. 3. Визуализация пропущенных значений

На диаграмме слева представлено количество пропущенных значений для каждого показателя до лечения в отдельности (WS – объем талии, TS – объем бедер, $Mass$ – масса тела, BMI – индекс массы тела, TMT – тощая масса тела, CAD – систолическое артериальное давление), а справа – для комбинации показателей (цифры справа обозначают количество пациентов, у которых пропущены значения по комбинациям из указанных показателей).

Для анализа закономерностей появления отсутствующих значений можно заменить данные условными значениями: 1 – обозначает пропущенное значение, 0 – имеющееся. Вычисление корреляций по методу Пирсона между этими преобразованными переменными и между исходными переменными поможет узнать, значения каких переменных имеют тенденцию отсутствовать согласованно, а также выявить связи между отсутствием значений в одной переменной и значениями других переменных. В полученной корреляционной матрице строки являются наблюдаемыми переменными, а столбцы – преобразованными переменными, несущими информацию о наличии пропущенных значений. На рис. 4 приведен фрагмент этой матрицы.

	ОТ_дл	ОТ_пл	ОБ_дл	ОБ_пл	Масса_дл	Масса_пл	избыток_дл	избыток_пл	ИМТ_дл	ИМТ_пл
ОТ_дл	1.0000	0.9606	0.8527	0.8594	0.2691	0.2457	0.2195	0.2088	0.2873	0.2574
ОТ_пл	0.9606	1.0000	0.8742	0.8964	0.2756	0.2756	0.2274	0.2380	0.2943	0.2864
ОБ_дл	0.8527	0.8742	1.0000	0.9775	0.3045	0.2818	0.2622	0.2533	0.3252	0.2983
ОБ_пл	0.8594	0.8964	0.9775	1.0000	0.3023	0.3023	0.2595	0.2709	0.3227	0.3167
Масса_дл	0.2691	0.2756	0.3045	0.3023	1.0000	0.9656	0.8195	0.7962	0.8710	0.8317
Масса_пл	0.2457	0.2756	0.2818	0.3023	0.9656	1.0000	0.8195	0.8270	0.8710	0.8634
избыток_дл	0.2195	0.2274	0.2622	0.2595	0.8195	0.8195	1.0000	0.9723	0.9121	0.8699
избыток_пл	0.2088	0.2380	0.2533	0.2709	0.7962	0.8270	0.9723	1.0000	0.8866	0.8732
ИМТ_дл	0.2873	0.2943	0.3252	0.3227	0.8710	0.8710	0.9121	0.8866	1.0000	0.9557
ИМТ_пл	0.2574	0.2864	0.2983	0.3167	0.8317	0.8634	0.8699	0.8732	0.9557	1.0000

Рис. 4. Фрагмент корреляционной матрицы (наличие связи между признаками)

Как видно из таблицы, появление недостающих значений для признаков $OT_дл$ (объем талии до лечения) и $OT_пл$ (объем талии после лечения) взаимосвязано ($r = 0.9606$). Если коэффициенты корреляции в матрице не слишком велики, можно предположить, что пропуски не сильно отклоняются от полностью случайных величин и могут считаться случайными.

Для восстановления пропущенных значений воспользуемся методом множественного восстановления [4], при котором заполнение пропусков происходит при помощи повторного моделирования. Из существующего набора данных с пропущенными значениями создается несколько полных наборов данных, к каждому из которых применяются стандартные статистические методы для формирования окончательных результатов. Идея множественного восстановления пропущенных данных хорошо реализована такой функцией пакета языка R, как `mice`.

В `mice` многомерное восстановление данных реализуется при помощи связанных уравнений. Пропущенные значения замещаются при помощи выборок Гиббса, который представляет собой способ формирования выборок $(x_1; \dots; x_n)$ из заданных распределений $p(x)$ m -мерных переменных путем многократного извлечения имитируемых значений. По умолчанию, каждая переменная, содержащая недостающие величины, моделируется в зависимости от всех других переменных в наборе данных. Полученные уравнения предсказания используются, чтобы приписать наиболее вероятные значения для пропущенных данных. Процесс повторяется, пока не будет достигнут определенный

критерий сходимости. Для каждой переменной пользователь может выбрать форму модели предсказания, называемую элементарным методом присвоения. Полученные уравнения используются для замещения пропущенных данных подходящими значениями. Этот процесс повторяется, пока значения для пропущенных данных не сойдутся. Процесс заполнения пропусков реализуется в несколько этапов (рис. 4).

Результатом работы функции `mice()` стала база данных с восстановленными значениями. Для оценки точности восстановленных данных были построены графики плотности распределения исходных и заполненных данных, пример представлен на рис. 5. Плотность исходных данных для каждого набора показана темным (синим) цветом, а плотность заполненных показателей – светлым (красным) цветом. Исходя из этого можно сделать вывод о том, что распределения восстановленных данных близки к исходным [12].

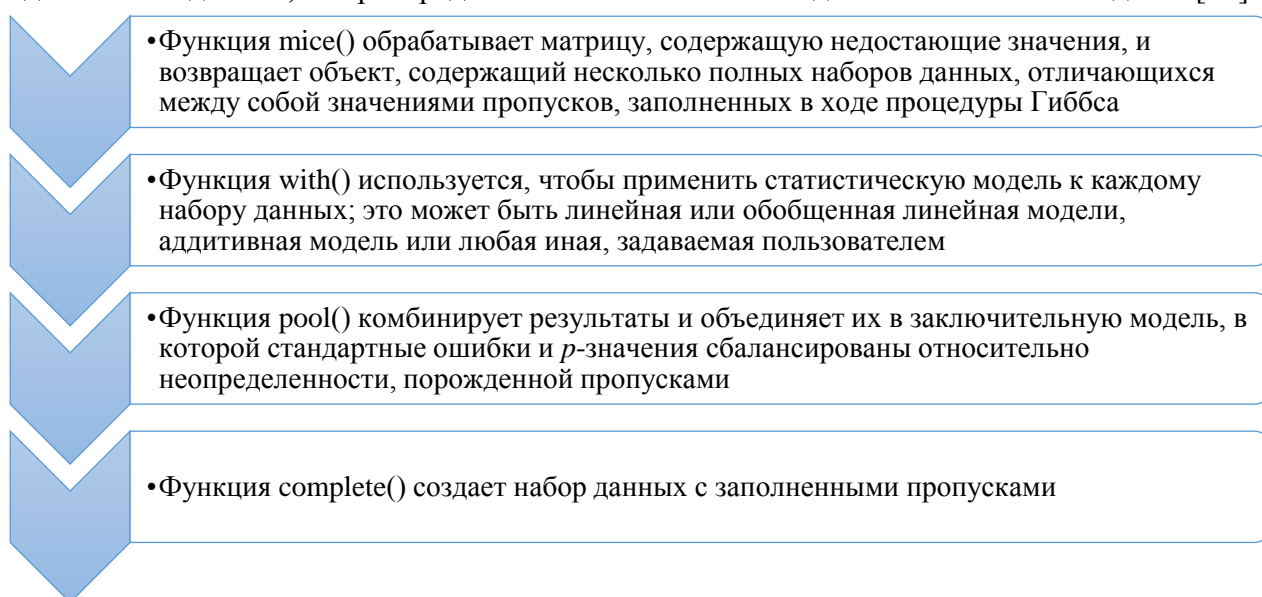


Рис. 4. Этапы заполнения пропусков с использованием метода множественного восстановления

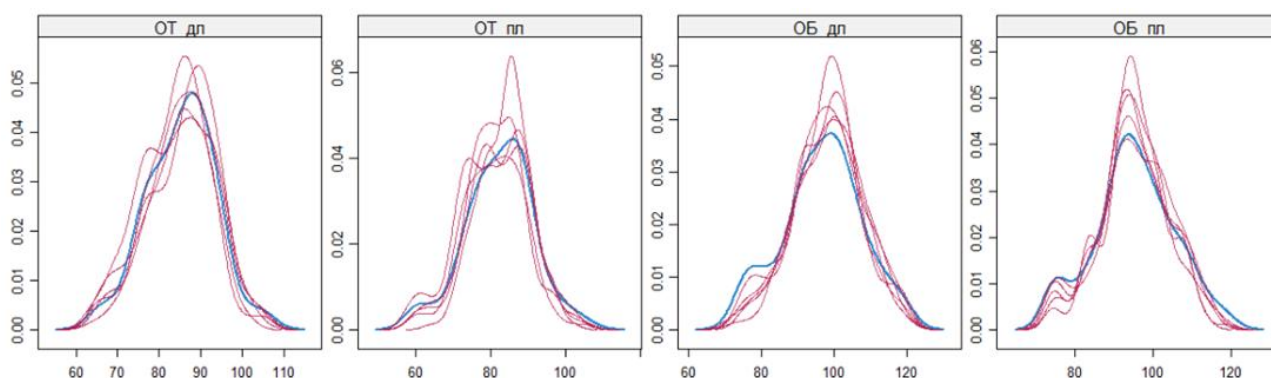


Рис. 5. Графики плотности распределения исходных (синий цвет) и заполненных (красный цвет) данных

3. Сокращение размерности данных. Рассмотрим методы выбора параметров для уменьшения размерности нашего набора данных. Высокая корреляция между двумя переменными означает, что они имеют сходные тенденции и, вероятно, несут схожую информацию. Это может резко снизить производительность некоторых моделей. Возможно рассчитать корреляцию между независимыми числовыми переменными, если коэффициент корреляции пересекает определенное пороговое значение, можно отбросить одну из

переменных. Отбрасывание переменной очень субъективно и всегда должно выполняться с учетом предметной области.

Некоторые методы, такие, как деревья решений, имеют встроенный механизм для сообщения о важности переменной. Для других алгоритмов важность может быть оценена с использованием анализа кривой ROC, проведенного для каждого атрибута. Случайный лес является одним из наиболее широко используемых алгоритмов выбора объектов. В него уже встроена функция важности, поэтому нет необходимости рассчитывать ее отдельно. Случайный лес может помочь выбрать меньшее подмножество функций.

Переменные с высокой важностью являются движущими факторами результата, и их значения оказывают значительное влияние на итоговые значения. Напротив, переменные с низкой важностью могут быть опущены в модели, что упрощает и ускоряет процесс подбора и прогнозирования.

Для выявления информативных признаков были использованы: дисперсионный анализ, классификация показателей по степени важности с использованием пакета Caret, случайный лес (Random Forest). Сравнительный анализ результатов этих методов позволили выделить 35 информативных признаков, сократив, таким образом, пространство исходных признаков: случайный лес дал наиболее полный набор (30 признаков) и еще 5 признаков были добавлены из результатов реализации дисперсионного анализа.

4. Построение прогностической модели. Для решения задачи формирования базы знаний был выбран метод построения дерева решений, как один из наиболее популярных и дающих хороший результат [12-14]. Исходными данными для построения дерева решений является набор данных, полученный в результате сокращения размерности. Данные были поделены на обучающую и тестовую выборки в соотношении 80/20. Для построения дерева решений были использованы пакет RapidMiner и скриптовый язык R. Графическое представление дерева решения (модель для 35 информативных признаков) в R выглядит, как представлено на рис. 7.

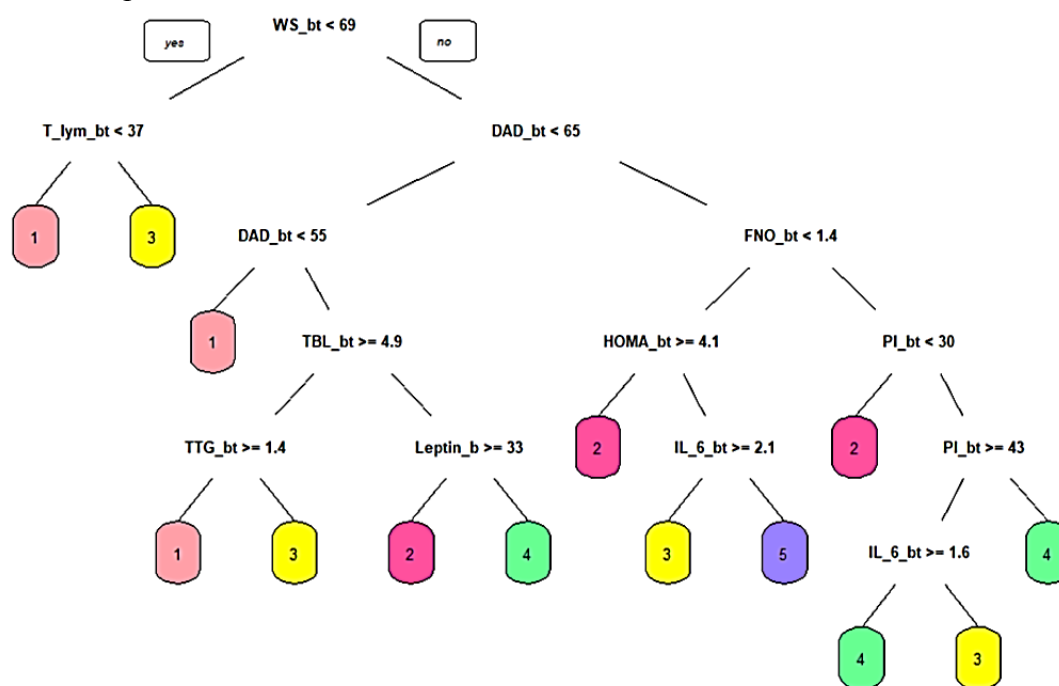


Рис. 7. Дерево решений (прогностическая модель)

Оценка точности построенной модели составляет 0.8284906. Для проверки качества построенной модели было проведено самостоятельное визуальное тестирование модели с исходными данными пациентов. Результат работы модели и априорное отнесение пациента к одной из 5 групп лечения совпали в 95% случаев.

Заключение. В результате проведенного исследования был разработан алгоритм поддержки принятия решения о выборе траектории лечения детей с избыточным весом на основе построения дерева решения. Для этого были реализованы следующие этапы: 1) получены результаты клинического исследования, проведенного на базе детского отделения Томского НИИ курортологии и физиотерапии; 2) проведена первичная обработка данных с помощью методов языка R, был проведен анализ выбросов и восстановление пропущенных значений методом множественного восстановления; 3) сокращена размерность признаков пространства средствами анализа уровня корреляции, линейной зависимости, дисперсии и степени важности признаков; 4) сформирована итоговая выборка для построения и поиска модели дерева решения, которая была разделена на обучающую и тестовую; 5) построено несколько моделей деревьев решений и выбрана модель с наилучшими значениями точности; 6) на основе полученной модели разработан модуль принятия решения в прототипе СППВР.

Исследование выполнено при частичной финансовой поддержке РФФИ в рамках выполнения научного проекта № 18-07-00543.

СПИСОК ЛИТЕРАТУРЫ

1. Атьков О.Ю., Кудряшов Ю.Ю., Прохоров А.А., Касимов О.В. Система поддержки принятия врачебных решений // Врач и информационные технологии. 2013. №6. С. 67–75.
2. Богданова Ю.А., Зарипова Г.Р., др. Современные модели экспертных медицинских систем в прогнозировании операционного риска при наиболее распространенных интраабдоминальных вмешательствах // Медицинский альманах. 2017. С. 9–12.
3. Зарипова Г.Р., Богданова Ю.А., др. Современные модели систем поддержки принятия врачебных решений в хирургической практике. Состояние проблемы // Медицинский вестник Башкортостана. 2016. С. 96–101.
4. Кабаков Р.И. R в действии. Анализ и визуализация данных в программе R / пер. с англ. Полины А. Волковой. М.: ДМК Пресс. 2014. 588 с.
5. Киселев К.В., Ноева Е.А., др. Разработка архитектуры базы знаний системы поддержки принятия врачебных решений, основанной на графовой базе данных // Медицинские технологии. Оценка и выбор. 2018. С.42-48.
6. Кнышов Г.В., Руденко А.В., др. Особенности проектирования медицинской информационной системы поддержки принятия решений, основанной на интеллектуальном анализе данных // Кибернетика и вычислительная техника. 2014. Вып. 177. С. 79–87.
7. Кудряшов Ю.Ю., Атьков О.Ю., Касимов О.В. Телемедицинская профилактика, реабилитация и управление здоровьем: проблемы и решения // Врач и информационные технологии. 2016. №2. С.73–80.

8. Купеева И.А., Разнатовский К.И., др. Оценка эффективности интеллектуальной системы поддержки принятия врачебных решений // Вестник СПбГУ. 2016. сер.10, вып.2. С. 62–68.
9. Мокина Е.Е., Марухина О.В., Шагарова М.Д., Дубинина И.А. Использование методов Data Mining при принятии медицинских диагностических решений. // Фундаментальные исследования. 2016. № 5-2. С. 269–274.
10. Раводин Р.А. Интеллектуальная система поддержки принятия врачебных решений в дерматовенерологии // Проблемы медицинской микологии. 2014. №3. С.59–65.
11. Спирычин А.А., Бурковский В.Л. Интеллектуальная система принятия решений в условиях выбора тактики лечения хронических заболеваний на основе облачных технологий // Вестник Воронежского государственного технического университета. 2014. №5-1.
12. Berestneva, O.G., Marukhina, O.V., Romanchukov, S.V., Berestneva, E.V. Visualization and Cognitive Graphics in Medical Scientific Research // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 11466 LNBI. 2019. Pp. 433–444.
13. Brett Lantz. Machine Learning with R. Packt Publishing, Birmingham – Mumbai. 2013. 396 с.
14. Gerget, O.M., Marukhina, O.V., Cherkashina, Yu.A. System for visualizing and analyzing multivariate data of medico-social research // Key Engineering Materials. 685. 2016. Pp. 957–961.
15. Obesity and overweight [Electronic resource] World Health Organization. Available at: <https://www.who.int/topics/obesity/en/> (accessed 11.07.2019).

UDK 004.822:519.81:613.25-053.4

**DEVELOPMENT THE DECISION SUPPORT SYSTEM FOR SELECTING THE
TREATMENT TRAJECTORY FOR CHILDREN WITH OVERWEIGHT**

Olga V. Marukhina

Ass. Professor, e-mail: Marukhina@tpu.ru

Olga G. Berestneva

Dr., Professor, e-mail: ogb6@yandex.ru

Department of Information Technology, School of Information Technology and Robotics,
Tomsk Polytechnic University, 30, Lenina Str., 634050, Tomsk, Russia

Abstract. An approach to the development of the algorithmic base of the medical decision support system when choosing the trajectory of treatment for children with an excessive weight is proposed. The structure of the missing data of clinical and laboratory indicators was studied, the missing values were restored using the multiple recovery method, and the size was reduced using the R language packages. Based on the results, a decision-making model was constructed and a set of rules for selecting the treatment trajectory was formed. As a result, a data processing algorithm was formed to decide on the choice of treatment group for overweight children.

Keywords: multidimensional data, data analysis, medical decision support system, medical research, data recovery, dimension reduction, logical rules.

Funding: The study was carried out with the partial financial support of the Russian Foundation for Basic Research (RFBR) within scientific project No. 18-07-00543.

References

1. At'kov O.Yu., Kudryashov Yu.Yu., Prohorov A.A., Kasimov O.V. Sistema podderzhki prinyatiya vrachebnyh reshenij [Medical Decision Support System] // Vrach i informacionnye tekhnologii = Physician and Information Technology. 2013. no 6. Pp. 67–75. (in Russian)
2. Bogdanova YU.A., Zaripova G.R. et al. Sovremennye modeli ekspertnyh medicinskih sistem v prognozirovanii operacionnogo riska pri naibolee rasprostranennyh intraabdominal'nyh vmeshatel'stvah [Modern models of expert medical systems in predicting operational risk with the most common intra-abdominal interventions] // Medicinskij al'manah = Medical Almanac. 2017. Pp. 9–12. (in Russian)
3. Zaripova G.R., Bogdanova YU.A. et al. Sovremennye modeli sistem podderzhki prinyatiya vrachebnyh reshenij v hirurgicheskoj praktike [Modern models of medical decision support systems in surgical practice. Problem state] // Medicinskij vestnik Bashkortostana = Medical Bulletin of Bashkortostan. 2016. Pp. 96–101. (in Russian)
4. Kabakov R.I. R v dejstvii. Analiz i vizualizaciya dannyh v programme R [R in action. Analysis and visualization of data in the program R] / translation to Russian by. Polina A. Volkova. Moscow. Izdatel'stvo "DMK Press" = publishing house "DMK Press". 2014. 588 p. (in Russian)
5. Kiselev K.V., Noeva E.A. et al. Razrabotka arhitektury bazy znaniy sistemy podderzhki prinyatiya vrachebnyh reshenij, osnovannoj na grafovoj baze dannyh [Development of the architecture of the knowledge base of the decision support system based on a graph database] // Medicinskie tekhnologii. Ocenka i vybor = Medical technology. Evaluation and selection. 2018. Pp. 42–48. (in Russian)
6. Knyshov G.V., Rudenko A.V. et al. Osobennosti proektirovaniya medicinskoj informacionnoj sistemy podderzhki prinyatiya reshenij, osnovannoj na intellektual'nom analize dannyh [Features of designing a medical information decision-making system based on data mining] // Kibernetika i vychislitel'naya tekhnika = Cybernetics and Computer Engineering. 2014. no 177. Pp.79–87. (in Russian)
7. Kudryashov YU.YU., At'kov O.YU., Kasimov O.V. Telemedicinskaya profilaktika, reabilitaciya i upravlenie zdorov'em: problemy i resheniya [Telemedicine prevention, rehabilitation and health management: problems and solutions] // Vrach i informacionnye tekhnologii = Physician and Information Technology. 2016. no 2. Pp. 73–80. (in Russian)
8. Kupeeveva I.A., Raznatovskij K.I. et al. Ocenka effektivnosti intellektual'noj sistemy podderzhki prinyatiya vrachebnyh reshenij [Evaluating the effectiveness of an intelligent medical decision support system] // Vestnik SPbGU = SPbSU Bulletin. 2016. no 10-2. Pp. 62–68. (in Russian)

9. Mokina E.E., Marukhina O.V., Shagarova M.D., Dubinina I.A. Ispol'zovanie metodov Data Mining pri prinyatii medicinskih diagnosticheskikh reshenij [Using Data Mining methods in making medical diagnostic decisions] // Fundamental'nye issledovaniya = Basic research. 2016. no 5-2. Pp. 269–274. (in Russian)
10. Ravodin R.A. Intel'ktual'naya sistema podderzhki prinyatiya vrachebnyh reshenij v dermatovenerologii [Intellectual system for supporting the adoption of medical decisions in dermatology and venereology] // Problemy medicinskoj mikologii = Medical Mycology Problems. 2014. no 3. Pp. 59–65. (in Russian)
11. Spiryachin A.A., Burkovskij V.L. Intel'ktual'naya sistema prinyatiya reshenij v usloviyah vybora taktiki lecheniya hronicheskikh zabolevanij na osnove oblachnyh tekhnologij [Intellectual decision-making system in terms of the choice of tactics for the treatment of chronic diseases based on cloud technologies] // Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta = Bulletin of the Voronezh State Technical University. 2014. no 5-1. (in Russian)
12. Berestneva, O.G., Marukhina, O.V., Romanchukov, S.V., Berestneva, E.V. Visualization and Cognitive Graphics in Medical Scientific Research // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 11466 LNBI, 2019. Pp. 433–444.
13. Brett Lantz. Machine Learning with R. Packt Publishing, Birmingham - Mumbai, 2013, 396 p.
14. Gerget, O.M., Marukhina, O.V., Cherkashina, Yu.A. System for visualizing and analyzing multivariate data of medico-social research // Key Engineering Materials, 685, 2016. Pp. 957–961.
15. Obesity and overweight [Electronic resource] World Health Organization. Available at: <https://www.who.int/topics/obesity/en/> (accessed 11.07.2019)