

УДК 004.62

## АНАЛИЗ ТРЕБОВАНИЙ К СТРУКТУРЕ АНСАМБЛЯ МОДЕЛЕЙ ДЛЯ ОБРАБОТКИ МЕДИКО-БИОЛОГИЧЕСКИХ ДАННЫХ

**Захарова Алёна Александровна**

Д.т.н., профессор кафедры «Информатика и программное обеспечение»,  
Брянский государственный технический университет,  
241035, Россия, г. Брянск, бул. 50 лет Октября, д. 7, e-mail: [zaa@tu-bryansk.ru](mailto:zaa@tu-bryansk.ru)  
ORCID: 0000-0003-4221-7710

**Подвесовский Александр Георгиевич**

К.т.н., доцент, заведующий кафедрой «Информатика и программное обеспечение»  
Брянский государственный технический университет,  
241035, Россия, г. Брянск, бул. 50 лет Октября, д. 7, e-mail: [apodv@tu-bryansk.ru](mailto:apodv@tu-bryansk.ru)  
ORCID: 0000-0002-1118-3266

**Толстенок Виктория Павловна**

Магистрант, Брянский государственный технический университет,  
241035, г. Брянск, бульвар 50-летия Октября 7, e-mail: [tolstenok21@yandex.ru](mailto:tolstenok21@yandex.ru)

**Аннотация.** В статье рассматриваются особенности медико-биологических данных и их влияние на выбор методов при разработке ансамбля моделей. Все медико-биологические данные условно делятся на пять групп, так как могут иметь различную природу, а для их систематизации необходимо применять шкалирование. Так же в статье приведен ряд особенностей медико-биологических данных, которые оказывают непосредственное влияние на выбор способа формирования ансамбля (в зависимости от требуемых результатов), а также на выбор методов, входящих в ансамбль. В статье приведены следующие методы: корреляционный анализ, ассоциативные правила, карты Кохонена и деревья решений. Таким образом, особенности входных данных при формировании ансамбля необходимо учитывать на начальном этапе, чтобы избежать дальнейших ошибок при работе над ансамблем и интерпретации результатов.

**Ключевые слова:** медико-биологические данные, особенности медико-биологических данных, ансамбль моделей, интеллектуальный анализ данных.

**Цитирование:** Захарова А.А., Подвесовский А.Г., Толстенок В.П. Анализ требований к структуре ансамбля моделей для обработки медико-биологических данных // Информационные и математические технологии в науке и управлении. 2019. № 4 (16). С. 89–98. DOI: 10.25729/2413-0133-2019-4-07

**Введение.** Интеллектуальный анализ стал частью большого числа исследований и технологий в современном мире и включает такие направления как статистика, машинное обучение, искусственный интеллект, распознавание образов, визуализацию и алгоритмизацию. Инструменты Data Mining применяются в различных сферах: от бизнеса до медицины.

Если говорить о применении методов интеллектуального анализа данных в медицине, то данное направление активно развиваться, но главным ограничением являются

особенности медико-биологических данных, для качественного анализа которых одного метода будет недостаточно. При работе с такими данными большинство алгоритмов не позволяет достичь желаемой точности. В подобных случаях целесообразно использовать ансамбли моделей. В этом случае главная задача ансамбля состоит в генерации разнообразных индивидуальных моделей, которые позволят повысить точность и качество результата [2].

Применение инструментов Data Mining в области медицины позволяет получить новые комплексные решения, которые ранее не были доступны.

**1. Особенности медико-биологических данных.** Понятие «медико-биологические данные» включает в себя не только выработанные человеческим организмом различные сигналы, которые свидетельствуют о состоянии человека, а также состояние окружающей среды, которое непосредственно оказывает влияние на жизнедеятельность человека.

Все медико-биологические данные можно условно разбить на пять групп:

а) Количественные данные. Это параметры, характеризующиеся дискретными величинами, такими как рост, вес, показания общего анализа крови или показатель заболеваемости населения.

б) Качественные данные. Качественными данными являются признаки, которым нельзя дать точную оценку, но при этом можно ранжировать, например, систематизировать по ряду условных баллов. К таким данным относятся балльные характеристики уровня боли, качества жизни человека или качества окружающей его среды.

в) Статические картины органов человека или всего тела. Такого рода данные можно получать благодаря средствам медицинской радиологии, которые отображают состояние пациента в целом или отдельных органов. Примером таких данных являются рентгенограмма, томограмма и т.д.

г) Данные, характеризующие динамическую картину органов человека. Данные такого рода можно получить только при постоянном отслеживании движущихся органов человека, таких как сердце или легкие.

д) Данные, характеризующие динамические физиологические функции, такие как электрокардиограмма, электроэнцефалограмма и другие [3].

Медико-биологические данные могут быть отображены в различных форматах (текстовый, числовой, графический и т.д.), а также могут обладать различными признаками, поэтому точной группировки, которая бы отразила все характеристики, особенности и способы регистрации такого рода данных просто не существует.

Особое внимание стоит уделить понятиям «признак» и «параметр». Эти понятия различны, поэтому при работе с медико-биологическими данными их необходимо четко различать, но при этом данное деление весьма условно [4].

Признак – бинарный параметр, принимающий два значения: наличие или отсутствие.

Параметр – это значение, которое можно охарактеризовать в абсолютных или относительных величинах.

Чтобы систематизировать и обработать медико-биологические данные, их необходимо упорядочить. Для этого было предложено использовать ряд шкал измерения:

- Шкала наименований – объединение объектов в ряд непересекающихся классов, при этом принято считать, что объекты, относящиеся к одному классу, являются идентичными.

- Шкала порядка. Данная шкала отражает направление процесса: порядок возрастания или убывания признака.
- Интервальная шкала – шкала, имеющая единицу измерения, что является ее отличительной особенностью.
- Шкала отношений – это разновидность интервальной шкалы, имеющая точку отсчета, то есть точку, в которой исследуемый параметр практически отсутствует [3].

Также медико-биологические данные необходимо привести к единому формализованному виду.

Единый унифицированный вид позволяет сопоставлять результаты различных исследований, анализов и отчетной документации. В Российской Федерации формализация данных регулируется законом «О стандартизации» и общепринятой считается формализация параметрических данных о пациенте на основе международного стандарта СИ. Отдельные стандарты регулируют передачу, обмен данными и архитектуру клинических документов.

Первым шагом при работе с любыми данными, в том числе и с медико-биологическими, необходимо определить их тип. Чаще всего используются два типа данных: качественные и количественные. Качественные данные можно ранжировать, а количественные – квантовать.

Работая с медико-биологическими данными, необходимо учитывать ряд особенностей [1, 8]:

- Нельзя использовать обнуление, то есть в случае пропусков информации пустые значения заменять на ноль нельзя, так как это в большинстве случаев совпадает с кодированием нормы по данному признаку или внесет ложную информацию в дальнейшую работу с данными.
- Не рекомендуется использовать в малых выборках среднее значение по классу ввиду их неоднородности.
- Необходимо четко разделять понятия «переменная» и «признак». Так же необходимо учитывать дихотомичность, так как переменные могут принимать одно из двух значений.
- Необходимо правильно определить значение градации качественного признака. Например, при определении боли по шкале от 1 до 5. Поэтому чтобы уйти от субъективного значения необходимо провести дополнительные исследования.
- Обратит внимание на формат записи дат. Он должен быть записан в единой форме для всех исследуемых данных.
- Обратит внимание на наличие дублирующих значений и ошибки ввода.
- При необходимости структурировать данные для получения адекватных результатов и исключения ошибок.

**2. Разработка ансамбля моделей с учетом особенностей медико-биологических данных.** Ансамбль моделей – это комбинация нескольких методов, которые позволяют создать более точную и эффективную модель, чем модель, в основе которой лежит только один метод. То есть при нахождении решения для одной задачи или доказательства выдвинутой гипотезы будет применена не одна модель, а несколько. При этом будет иметь значение не результат работы одной отдельно взятой модели, а результат работы ансамбля в целом [11].

Обучение ансамбля – это процесс обучения набора базовых методов с последующим объединением результатов их прогнозирования в единый прогноз агрегированного классификатора методов. Таким образом благодаря агрегированному классификатору методов можно получить более точный результат. Под понятием «агрегированный классификатор методов» подразумеваются все методы, входящие в ансамбль моделей.

Цель объединения моделей – улучшить (усилить) решение, которое дает отдельная модель. При этом предполагается, что единственная модель никогда не сможет достичь той эффективности, которую обеспечит ансамбль [9].

Можно выделить три ключевых причины объединения методов (моделей) в ансамбль:

- Уменьшает вероятность влияния разного рода случайностей (например, выбросы) на агрегированную гипотезу.
- Ансамбль моделей ищет глобальным оптимум и избегает локальных оптимумов, так как ведет поиск из разных точек исходного множества гипотез.
- Существует возможность того, что агрегированная гипотеза будет находиться за пределами множества базовых гипотез, в этом случае при построении комбинированной гипотезы можно просто расширить множество возможных гипотез.

Первым шагом для разработки ансамбля является выбор класса ансамбля. Все ансамбли можно разделить на два класса:

- Ансамбли, состоящие из базовых методов только одного типа (только из деревьев решений, нейронных сетей и т. д.).
- Ансамбли, состоящие из базовых методов разного типа (деревья решений, карты Кохонена, различные алгоритмы кластеризации и т. д.).

Для работы с медико-биологическими данными оптимальным будет выбор ансамбля, состоящего из методов различного типа, так как такой класс ансамблей обладает большей гибкостью и точностью.

Большую роль играет правильный выбор метода формирования ансамбля. Во многом выбор зависит от поставленных целей, так как каждый метод направлен на решение определенных задач. Далее рассмотрим два наиболее популярных и часто используемых метода.

Бустинг – алгоритм, относящийся к итерационным и предназначенный для увеличения точности модели. Каждой записи данных на каждой итерации алгоритма присваивается вес. Первый метод обучается на всех примерах с равными весами. На каждой последующей итерации веса расставляются соответственно классифицированным примерам, т. е. веса правильно классифицированных примеров уменьшаются, а неправильно классифицированных – увеличиваются. Следовательно, приоритетными для следующего метода станут неправильно распознанные примеры, обучаясь на которых новый метод будет исправлять ошибки метода на прошлой итерации [5].

Бэггинг направлен на повышение стабильности модели. Данный метод строит модель ансамбля, используя бутстреп-агрегирование, которое генерирует множественные модели для получения более надежных предсказаний. Алгоритм бэггинга подразумевает следующие шаги. Сначала формируется несколько выборок путем случайного отбора из исходного множества данных. Затем на основе каждой выборки строится метод, и выходы всех методов агрегируются с использованием голосования или простого усреднения. Очевидно, что

точность предсказания построенных с помощью беггинга комбинированных методов оказывается значительно выше, чем точность отдельных моделей [7].

Для работы с медико-биологическими данными подойдет любой из методов.

Важным этапом при работе с медико-биологическими данными является предобработка, так как данные могут поступать из нескольких источников и вводиться вручную (поэтому могут присутствовать пустые или дублирующие значения, ошибки ввода и т.д.), также может присутствовать различное шкалирование одних и тех же параметров, поступивших из разных источников.

В большинстве случаев, поступившие данные не пригодны для анализа, так как они слабоструктурированы, в них встречаются пропуски, дублирующие и аномальные значения. Если начать работать с данными в исходном виде, то результат работы ансамбля моделей будет неудовлетворительным. Например, каждому пациенту может быть поставлено до нескольких разных диагнозов, которые занесены в одно поле. Таким образом, для ансамбля моделей входной параметр «Диагноз» вместо 4 различных значений будет иметь только одно, состоящее сразу из нескольких диагнозов.

Пропуски и дублирующие значения также внесут помехи в работу модели, что отрицательно скажется на итоговом результате.

На рис. 1 показаны данные, которые прошли процедуру предобработки.

id_Пациента	Возраст	1-муж;2-жен; Пол	Диагноз (код МКБ)	Название диагноза	Гемоглобин	Отклонение
1	28	1	E78.0	Чистая гиперхолестеринемия	141	0
1	28	1	G90.9	Расстройство вегетативной [автоном]	141	0
1	28	1	H35.0	Фоновая ретинопатия и ретинальнь	141	0
1	28	1	H52.1	Миопия	141	0

Рис. 1. Данные после предобработки

После предобработки появилось новое поле «id\_Пациента», каждый диагноз заносится в отдельное поле, устранены дублирующие и пустые значения. Далее данные можно использовать для разработки ансамбля моделей.

**3. Выбор метода.** Существуют ограничения при выборе методов для обработки медико-биологических данных. Чаще всего используются следующие методы:

а) Корреляционный анализ.

Используется при оценке зависимостей выходных полей данных от входных факторов и при устранении незначущих факторов. Основной принцип заключается в поиске в наименьшей степени коррелированных (взаимосвязанных) значений с выходным результатом. Полученные факторы исключаются из результирующего набора данных почти без потерь важной информации. Для принятия решения об исключении фактора используется порог значимости. Если корреляция (степень взаимозависимости) между входным и выходным факторами меньше порога значимости, то соответствующий фактор отбрасывается как незначущий.

При работе с медико-биологическими данными корреляционный анализ позволит выявить наиболее значимые показатели для каждой группы пациентов.

б) Ассоциативные правила.

Ассоциативные правила [6] позволяют находить закономерности между связанными событиями. Для работы с ассоциативными правилами данные необходимо представить в

виде транзакций. Например, одному пациенту соответствует один и более диагнозов. Пример данных в виде транзакций представлен на рис. 1. На рис. 2 показан пример работы ассоциативных правил.

№	Условие	Следствие	Поддержка		Достоверность	Лифт
			Кол-во	%		
1	Гиперметропия	Нейросенсорная потеря слуха двусторонняя	3	3,00	33,33	11,11
2	Гиперметропия	Хроническая обструктивная легочная болезнь неуточненная	3	3,00	33,33	8,33
3	Хроническая обструктивная легочная болезнь неуточнена	Гиперметропия	3	3,00	75,00	8,33
4	Гиперметропия	Эссенциальная [первичная] гипертензия	3	3,00	33,33	2,38
5	Эссенциальная [первичная] гипертензия	Гиперметропия	3	3,00	21,43	2,38
6	Пресбиопия	Расстройство вегетативной [автономной] нервной системы не	4	4,00	26,67	2,22
7	Расстройство вегетативной [автономной] нервной системы	Пресбиопия	4	4,00	33,33	2,22
8	Пресбиопия	Чистая гиперхолестеринемия	6	6,00	40,00	1,53
9	Чистая гиперхолестеринемия	Пресбиопия	6	6,00	23,08	1,53
10	Расстройство вегетативной [автономной] нервной системы	Чистая гиперхолестеринемия	5	5,00	41,67	1,60
11	Чистая гиперхолестеринемия	Расстройство вегетативной [автономной] нервной системы не	5	5,00	19,23	1,60
12	Чистая гиперхолестеринемия	Эссенциальная [первичная] гипертензия	5	5,00	19,23	1,37
13	Эссенциальная [первичная] гипертензия	Чистая гиперхолестеринемия	5	5,00	35,71	1,37

Рис. 2. Пример работы метода «Ассоциативные правила»

Применение ассоциативных правил позволяет выявить связанные диагнозы и увидеть степень связи и количество случаев.

в) Карты Кохонена.

Самоорганизующиеся карты Кохонена [12] являются удобным графическим инструментом для работы с медико-биологическими данными. С помощью карт можно увидеть значимые различия в данных, подробно описать их состав или выявить закономерности. На рис. 3 показан пример работы метода «Самоорганизующиеся карты Кохонена».

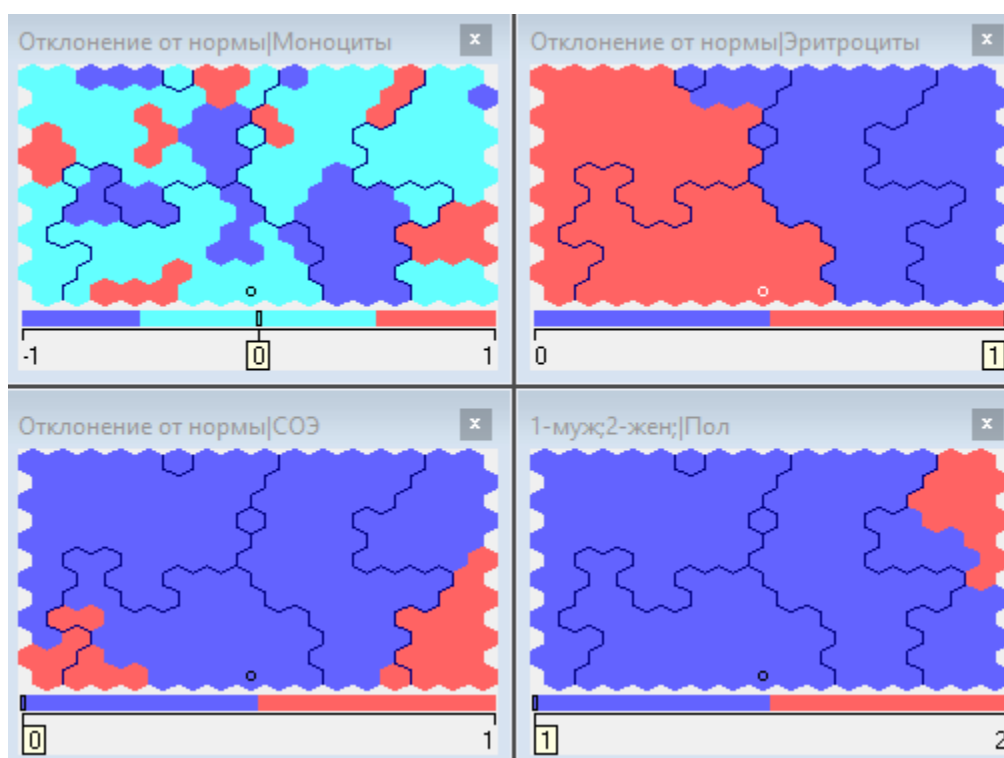


Рис. 3. Пример работы метода «Самоорганизующиеся карты Кохонена»

г) Дерево решений.

Деревья решений [6] – это способ представления правил в иерархической, последовательной структуре, где каждому объекту соответствует единственный узел, дающий решение. Под правилом понимается логическая конструкция, представленная в виде «если ... то ...». Пример работы метода показан на рис. 4.

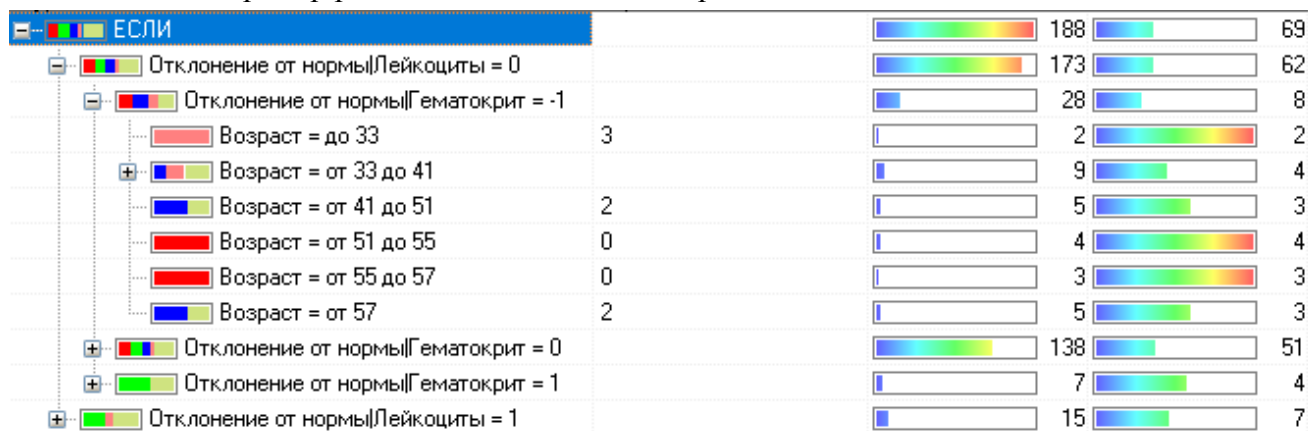


Рис. 4. Пример работы метода «Дерево решений»

Применительно к медико-биологическим данным, деревья решений менее эффективны, чем самоорганизующиеся карты Кохонена. Даже на малых выборках (порядка 100–200 записей) дерево решений получается достаточно большим и трудноинтерпретируемым.

Таким образом, стандартный алгоритм разработки ансамбля моделей предложено скорректировать с учетом особенностей медико-биологических данных и примет следующий вид:

- Шаг 1. Проанализировать исходные данные и при необходимости провести предобработку данных.
- Шаг 2. Определить группу, признак и параметр для медико-биологических данных. Так же необходимо грамотно подобрать единую шкалу для анализа данных.
- Шаг 3. В зависимости от поставленных задач выбрать метод формирования ансамбля.
- Шаг 4. Правильно подобрать методы и последовательность их использования.
- Шаг 5. Интерпретировать полученный результат.

В результате работы ансамбля можно будет увидеть не только скрытые закономерности, но и выдвинуть ряд значимых гипотез, ранее не найденных [10].

**Заключение.** Медико-биологические данные оказывают непосредственное влияние на алгоритм разработки ансамбля моделей. В большинстве случаев исходные данные нуждаются в предобработке: необходимо убрать дублирующие значения, структурировать данные и исправить ошибки ручного ввода. Так же при работе с такими данными необходимо правильно определить их тип (например, количественные или качественные данные), правильно определить признак и параметр, а для упорядочивания использовать единую шкалу.

Медико-биологические данные обладают рядом особенностей, но при работе с ними наиболее значимыми являются пропуски информации (в этом случае нельзя использовать обнуление, так как оно может совпасть с кодированием нормы кого-либо признака), наличие дублирующих значений и ошибок ввода, а также слабая структуризация данных.

Так же предложен алгоритм разработки ансамбля моделей, который учитывает особенности слабоструктурированных выборок медико-биологических данных. Данный алгоритм прошел апробацию на малых выборках и показал свою эффективность.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проект № 19-07-00844.

#### СПИСОК ЛИТЕРАТУРЫ

1. Барьеры и перспективы цифровой трансформации: проблемы управления большими данными в индустрии здравоохранения. Режим доступа: <http://www.medlinks.ru/article.php?sid=83028> (дата обращения 11.11.2019)
2. Кашницкий Ю.С., Игнатов Д.И. Ансамблевый метод машинного обучения, основанный на рекомендации классификаторов // Интеллектуальные системы. Теория и приложения. 2015. Т. 19. № 4. С. 37–55.
3. Королюк И.П. Медицинская информатика. Самара: СамГМУ. 2012. 244 с.
4. Мангалова Е.С. Исследование влияния разнообразия индивидуальных моделей на точность ансамбля // Решетневские чтения. 2015. Т. 2. С. 67–68.
5. Орешков В.И. Методы и модели интеллектуального анализа данных в задачах управления в социальных и экономических системах: автореф. дисс. ... канд. техн. наук. Рязань. 2013. 23 с.
6. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям. СПб.: Питер. 2013. 704 с.
7. Сошенко А.Е., Королева Н.А. Ансамбли моделей для прогнозирования. Режим доступа: [http://arbir.ru/articles/a\\_4053.htm](http://arbir.ru/articles/a_4053.htm) (дата обращения 11.11.2019)
8. Цветкова Л.А., Черченко О.В. Технология Больших Данных в медицине и здравоохранении России и мира // Врач и информационные технологии. 2016. № 3. С. 60–73.
9. Цыганкова И.А. Метод интеллектуальной обработки медико-биологических данных // Программные продукты и системы. 2009. № 3. С. 120–123.
10. Geger E.V., Podvesovskii A.G., Kuzmin S.A., Tolstenok V.P. Methods for the Intelligent Analysis of Biomedical Data // CEUR Workshop Proceedings of the 29th International Conference on Computer Graphics and Vision (GraphiCon 2019). Vol. 2485. Pp. 308–311.
11. Hamed R. Bonab, Fazli Can. A Theoretical Framework on the Ideal Number of Classifiers for Online Ensembles in Data Streams // 25th Conference on Information and Knowledge Management, ACM, New York. 2016. Pp. 2053–2056.
12. Kohonen T. The Self-Organizing Map // Proceeding of the IEEE. 1990. Vol. 78. Pp. 1464–1480.



UDK 004.62

## ANALYSIS OF REQUIREMENTS FOR THE ENSEMBLE OF MODELS STRUCTURE FOR BIOMEDICAL DATA PROCESSING

**Alena A. Zakharova**

Dr. Tech. Sc., Professor of Informatics and Software Engineering Department  
Bryansk State Technical University  
7, 50 let Oktyabrya blvd., 241035, Bryansk, Russia, e-mail: [zaa@tu-bryansk.ru](mailto:zaa@tu-bryansk.ru)  
ORCID: 0000-0003-4221-7710

**Aleksandr G. Podvesovskii**

Cand. Tech. Sc., Associate Professor  
Head of Informatics and Software Engineering Department  
К.т.н., доцент, заведующий кафедрой «Информатика и программное обеспечение»  
Bryansk State Technical University  
7, 50 let Oktyabrya blvd., 241035, Bryansk, Russia, e-mail: [apodv@tu-bryansk.ru](mailto:apodv@tu-bryansk.ru)  
ORCID: 0000-0002-1118-3266

**Viktoria P. Tolstenok**

Graduate Student, Bryansk State Technical University  
7, 50 let Oktyabrya blvd., 241035, Bryansk, Russia, e-mail: [tolstenok21@yandex.ru](mailto:tolstenok21@yandex.ru)

**Abstract.** The article discusses the features of biomedical data and their influence on the choice of classifiers in the development of the ensemble of models. All biomedical data are conventionally divided into five groups, so they can have a different nature, and scaling is necessary for their systematization. The article also presents several features of biomedical data. Biomedical features have a direct impact on the choice of ensemble formation method (depending on the desired results), as well as on the choice of classifiers included in the ensemble. The following classifiers are presented in the article: correlation analysis, associative rules, Kohonen maps and decision trees. Thus, the features of the input data during the formation of the ensemble must be considered at the initial stage in order to avoid further errors when working on the ensemble and interpretation of the results.

**Keywords:** biomedical data, features of biomedical data, ensemble of models, data mining.

### References

1. Bar'ery i perspektivy cifrovoj transformacii: problemy upravlenija bol'shimi dannymi v industrii zdavoohranenija [Barriers and Opportunities of Digital Transformation: Problems of Big Data Control in the Industry of Public Health]. Available at: <http://www.medlinks.ru/article.php?sid=83028> (accessed 11.11.2019) (in Russian)
2. Kashnickij Ju.S., Ignatov D.I. Ansamblevyj metod mashinnogo obuchenija, osnovannyj na rekomendacii klassifikatorov [Ensemble Method of Machine Learning Based on Classifiers Recommendations] // Intellektual'nye sistemy. Teorija i prilozhenija. = Intelligent systems. Theory and applications. 2015. Vol. 19 (4). Pp. 37–55. (in Russian)

3. Koroljuk I.P. Medicinskaja informatika [Medical Infomatics]. Samara: Samara State Medical University. 2012. 244 p. (in Russian)
4. Mangalova E.S. Issledovanie vlijaniya raznoobrazija individual'nyh modelej na tochnost' ansamblja [Research of Influence of Individual Models Diversity on Ensemble Accuracy] // Reshetnevskie chtenija = Reshetnev Readings. 2015. Vol. 2. Pp. 67–68 (in Russian)
5. Oreshkov V.I. Metody i modeli intellektual'nogo analiza dannyh v zadachah upravlenija v social'nyh i jekonomicheskikh sistemah [Methods and Models of Intelligent Data Analysis in Tasks of Management in Social and Economic Systems]: Dissertation abstract. Ryazan. 2013. 23 p. (in Russian)
6. Paklin N.B., Oreshkov V.I. Biznes-analitika: ot dannyh k znanijam [Business Intelligence: from Data to Knowledge]. Saint-Petersburg. Piter. 2013. 704 p. (in Russian)
7. Soshenko A.E., Koroleva N.A. Ansambli modelej dlja prognozirovanija [Ensemble of Models for Forecasting]. Available at: [http://arbir.ru/articles/a\\_4053.htm](http://arbir.ru/articles/a_4053.htm) (accessed 11.11.2019) (in Russian)
8. Tsvetkova L.A., Cherchenko O.V. Tekhnologiya Bol'shikh Danykh v meditsine i zdavookhranenii Rossii i mira [Big Data Technology in medicine and healthcare of Russia and the world] // Vrach i informatsionnyye tekhnologi = Doctor and information technology. 2016. № 3. Pp. 60–73. (in Russian)
9. Tsygankova I.A. Metod intellektual'noj obrabotki mediko-biologicheskikh dannyh [Method of Intellectual Processing of Medical and Biologic Data] // Programmnye produkty i sistemy = Software and Systems. 2009. no. 3. Pp. 120–123 (in Russian)
10. Geger E.V., Podvesovskii A.G., Kuzmin S.A., Tolstenok V.P. Methods for the Intelligent Analysis of Biomedical Data // CEUR Workshop Proceedings of the 29th International Conference on Computer Graphics and Vision (GraphiCon 2019). Vol. 2485. Pp. 308–311.
11. Hamed R. Bonab, Fazli Can. A Theoretical Framework on the Ideal Number of Classifiers for Online Ensembles in Data Streams // 25th Conference on Information and Knowledge Management. ACM, New York. 2016. Pp. 2053–2056.
12. Kohonen T. The Self-Organizing Map // Proceeding of the IEEE. 1990. Vol. 78. Pp. 1464–1480.