Научная статья

УДК 004.021
DOI 10.25205/1818-7900-2025-23-2-53-63

Методика оценки результатов решения задачи отбора признаков

Артем Дмитриевич Черемухин
Андрей Давыдович Рейн

Нижегородский государственный инженерно-экономический университет
Княгинино, Россия

ngieu.cheremuhin@yandex.ru, https//orcid.org/0000-0003-4076-5916
ndr18@yandex.ru, https//orcid.org/0000-0002-1921-483X

Аннотация

Данная статья представляет собой исследование методов оценки эффективности алгоритмов отбора признаков и предлагает новую методику их оценки. Отмечено, что существующие методы и подходы к оценке не всегда способны адекватно отразить действительную эффективность алгоритмов, особенно при применении к реальным задачам. В рамках статьи подробно рассматриваются различные мнения исследователей по проблемам внутренней и внешней валидности существующих методов оценки, оценивается влияние разных параметров, включая объем данных, различия в реализациях алгоритмов и другие факторы. Авторы предлагают новый комплексный подход к оценке эффективности алгоритмов, включающий ряд показателей, среди которых затраты ресурсов, стабильность и качество решения задачи. Одной из особенностей предлагаемого подхода является использование искусственно сгенерированных данных, что позволяет учесть специфические характеристики реальных данных и оценить алгоритмы в контролируемых условиях. Это дает возможность более точно определить их эффективность и надежность. Статья также содержит результаты предварительного тестирования предложенной методики на примере искусственных данных. Анализ этих результатов демонстрирует преимущества нового подхода над традиционными методами оценки. В частности, было установлено, что новая методика позволяет более точно оценивать стабильность и качество работы алгоритмов, что имеет важное значение для принятия решений о выборе подходящего алгоритма для конкретных задач.

В заключение, авторы подчеркивают необходимость дальнейшего развития и расширения предложенной методики, а также ее адаптации для решения различных типов задач. Они также указывают на перспективы интеграции этой методики в специализированные программные пакеты, что сделает ее доступной для широкого круга пользователей и ускорит внедрение инновационных алгоритмов в практику.

Ключевые слова

отбор признаков, методика, стабильность отбора признаков, эффективность отбора признаков, вычислительные ресурсы

Для цитирования

Черемухин А. Д., Рейн А. Д. Методика оценки результатов решения задачи отбора признаков // Вестник НГУ. Серия: Информационные технологии. 2025. Т. 23, № 2. С. xx–xx. DOI 10.25205/1818-7900-2025-23-2-53-63

Methodology for Evaluating the Results of Feature Selection Task Solving

Artem D. Cheremuhin, Andrey D. Rein

Nizhny Novgorod State Engineering and Economic University
Knyaginino, Russian Federation

ngieu.cheremuhin@yandex.ru, https//orcid.org/0000-0003-4076-5916
ndr18@yandex.ru, https//orcid.org/0000-0002-1921-483X

Abstract

Article is a study of methods for evaluating the effectiveness of feature selection algorithms and proposes a new methodology for their evaluation. It is noted that existing methods and approaches to assessment do not always adequately reflect the actual efficiency of algorithms, especially when applied to real problems. The article comprehensively discusses various opinions of researchers on the issues of internal and external validity of existing evaluation methods, assesses the impact of different parameters, including data volume, differences in algorithm implementations, and other factors. The authors propose a new integrated approach to evaluating the effectiveness of algorithms, which includes a set of indicators such as resource costs, stability, and task solution quality. One of the peculiarities of the proposed approach is the use of artificially generated data, which allows considering specific characteristics of real data and evaluating algorithms under controlled conditions. This makes it possible to more accurately determine their efficiency and reliability. The article also contains the results of preliminary testing of the proposed methodology using artificial data. The analysis of these results demonstrates the advantages of the new approach over traditional evaluation methods. In particular, it was found that the new methodology allows more accurately evaluating the stability and quality of algorithm performance, which is crucial for making decisions about choosing an appropriate algorithm for specific tasks. In conclusion, the authors emphasize the need for further development and expansion of the proposed methodology, as well as its adaptation for solving various types of tasks. They also point out the prospects for integrating this methodology into specialized software packages, which will make it accessible to a wide range of users and accelerate the implementation of innovative algorithms in practice.

Keywords

feature selection, methodology, feature selection stability, feature selection efficiency, computational resources

For citation

Cheremuhin A. D., Rein A. D. Methodology for Evaluating the Results of Feature Selection Task Solving. *Vestnik NSU. Series: Information Technologies*, 2025, vol. 23, no. 2, pp. 53–63 (in Russ.) DOI 10.25205/1818-7900-2025-23-2-53-63

Введение

Машинное обучение с учителем – один из наиболее распространенных видов машинного обучения, характеризующийся наличием известных значений зависимой переменной. В практических задачах оно применяется при решении задач регрессии (если зависимая переменная числовая) или классификации (если зависимая переменная факторная); при этом очень часто перед исследователями встает задача определения точного состава признаков, определяющих значение выбранной переменной. Особенно это важно, например, при решении задачи по определению влияющих генов, где необходимо выбрать 1–2 влияющих признака из сотен или тысяч. Эта задача известна как задача отбора признаков (feature selection).

Постановка задачи

В настоящее время разработано и реализовано в программном коде больше 100 различных алгоритмов отбора признаков, что ставит вопрос об определении наиболее эффективных методов. Классическим подходом для доказательства результативности разработанных методов и их пригодности для реальной деятельности является подход с применением бенчмарков:

оценивается эффективность алгоритма (как правило, одним показателем) на основе общедоступного набора датафреймов и сравнивается с результатами прошлых алгоритмов.

Однако практика использования бенчмарков подвергается обоснованной критике – в работе [1] описываются многочисленные проблемы внутренней и внешней валидности, которые возникают при подобном подходе. Во-первых, указывается на то, что лучшие результаты на бенчмарке не гарантируют лучшего результата на задачах «реального мира». Во-вторых, существуют проблемы воспроизводимости алгоритмов – разные реализации одних и тех же алгоритмов зачастую ведут себя как разные алгоритмы. В-третьих, тестируемые алгоритмы в реальных задачах вынуждены работать в более широком диапазоне сценариев (больше/меньше наблюдений, признаков, загрязненные данные), чем реализуемые при тестировании бенчмарками. В-четвертых, когда алгоритм разрабатывается для достижения высоких результатов на каких-то данных, можно уверенно говорить о наличии эффекта переобучения.

Кроме того, как отмечается в работе [2], «даже в самых простых сценариях одного общего показателя недостаточно для точного отражения производительности машинного обучения», соответственно, даже использование показателя величины ошибки (для регрессии) или точности (для классификации) не гарантирует эффективность метода. В работе [3] обращается внимание на наличие разных объектов измерения: точность алгоритма, его устойчивость, вычислительную сложность, объяснимость результатов, ресурсоемкость алгоритмов. При этом обзор литературы показал, что в большинстве исследований, как в [4], игнорируется многообразие объектов оценки, что ставит под сомнения получаемые выводы об эффективности/неэффективности алгоритмов в контексте решения задачи.

Таким образом, можно подтвердить, что в сфере отбора признаков, как и вообще в сфере машинного обучения, сейчас «разработка теории измерений результатов машинного обучения отстает от достижений самого машинного обучения» [1]. Этот факт определил тему исследования, его целью является разработка и апробация методики оценки результатов решения задачи отбора признаков.

Отбор признаков (feature selection) формально является подвидом машинного обучения без учителя и методом сокращения размерности [5], который позволяет из множества переменных отобрать только значимые/влияющие и убрать незначимые/невлияющие. Он применяется в процессе использования как методов обучения с учителем (при решении задач классификации и регрессии), так и методов обучения без учителя (при решении задач кластеризации).

Существуют три основные стратегии решения задачи отбора признаков:

– прямая (на первом шаге множество значимых признаков является пустым, и на каждой итерации происходит добавление в него одного самого важного признака до достижения критерия остановки) [6];

– обратная (на первом шаге множество значимых признаков равно множеству исследуемых признаков, и на каждой итерации происходит удаление самого незначимого признака);

– двусторонняя (на каждом шаге происходит добавление самого значимого из невключенных и удаление самого незначимого из включенных в состав значимых признаков) [7].

Все алгоритмы отбора признаков условно можно разделить на 4 большие группы:

– алгоритмы-фильтры. В них выбирается некоторая метрика «важности» признака (как правило, она является метрикой качества связи между зависимой переменной и переменной-кандидатом на включение в состав влияющих), и решением задачи отбора признаков некоторое является подмножество наиболее важных признаков, выбранных по критерию топ-N или иным подобным признакам [8]. Преимущество данного класса алгоритмов – они быстры и просты в вычислительном отношении и масштабируются для данных любого размера [9], недостаток – рассмотрение признаков «по одному» игнорирует случаи взаимодействия признаков [10], что зачастую приводит к игнорированию важных закономерностей и снижения качества решения задачи;

– «оберточные» алгоритмы. В них выбирается некоторое подмножество признаков, сразу оценивается его качество (в отличие от алгоритмов-фильтров, которые оценивают качество полученного набора признаков только один раз, в конце), и на основе этого принимается решение об изменении подмножества. Примерами таких признаков являются генетические алгоритмы;

– встроенные алгоритмы. К этому семейству алгоритмов относятся те, в которых процесс отбора признаков тем или иным образом встроен в метод определения качества решения, например алгоритм LASSO-регрессии;

– гибридные алгоритмы – алгоритмы, реализующие в себе два или более рассматриваемых подхода.

Методы исследования

Поскольку в общем случае задача отбора признаков является задачей обучения без учителя (потому что правильный набор влияющих признаков неизвестен), а используется он как составляющая часть решения задачи обучения с учителем, то наиболее часто применяемый подход – качеством решения задачи отбора признаков объявить качество решения исходной задачи (например, величину среднеквадратичной ошибки для задачи регрессии и величину точности классификации для задачи классификации). Однако, как показывают современные исследования в области медицины и биоинформатики, важна не только точность метода отбора признаков, но и его стабильность – он должен обеспечивать выбор максимального одинакового подмножества признаков при добавлении или удалении обучающих образцов [11].

Стоит отметить, что в современных исследованиях [3] используется три подхода к оценке стабильности: стабильность по записи (record-stability of a feature selection; оценивает стабильность подмножества при удалении/добавлении наблюдений в датафрейме), стабильность по признакам (feature-stability of a feature selection; оценивает стабильность подмножества при удалении/добавлении признаков в датафрейме) и стабильность по Ляпунову, которая объединяет стабильность и по записи, и по признакам; при этом конкретная мера оценки сходства двух подмножеств может быть любой.

Необходимое для корректного решения практических задач усложнение подхода к оценке результатов алгоритма отбора признаков еще больше актуализирует необходимость разработки понятной методики данного процесса, основой которой должна стать соответствующая система показателей.

Однако есть нюанс, дополнительно усложняющий задачу. Наряду с использованием классических бенчмарков популярным является подход с использованием синтетических и искусственного сгенерированных датафреймов с заранее заданными свойствами [12]. Он обладает рядом преимуществ, главным из которых является знание правильного подмножества признаков; однако при использовании синтетических датафреймов задача отбора признаков превращается в задачу обучения с учителем, что диктует иной подход к оценке ее точности.

По мнению автора, оценка результатов решения задачи отбора признаков должна включать:

– меры затрат ресурсов (количество операций, объем занимаемой памяти, время работы алгоритма) [13];

– меры стабильности (рекомендуется применять все три ранее рассмотренных оценки);

– меры эффективности работы алгоритма.

Дискуссионным является вопрос и о содержании мер стабильности – в работе [5] изучены наиболее часто рассматриваемые расстояния между подмножествами, при этом особо указывается на отсутствие четких теоретических рекомендаций по выбору метрик.

Таким образом, при работе с данными, по которым неизвестно верное подмножество признаков, предлагается использовать следующую систему показателей:

– количество операций, совершенное алгоритмом;

– объем оперативной памяти, затраченной алгоритмом;
– время работы алгоритма;
– индексы Танимото [14] стабильности по записи, признакам и Ляпунову;
– показатели качества решения задачи: среднеквадратичная ошибка, скорректированный коэффициент детерминации в случае решения задачи регрессии; общий показатель точности с учетом дисбаланса классов, F1-мера при решении задачи классификации.

При работе с данными, по которым известно верное подмножество признаков, предлагается дополнить данную систему на основе подхода, представленного автором в [15]:
– количество операций, совершенное алгоритмом;
– объем оперативной памяти, затраченной алгоритмом;
– время работы алгоритма;
– индексы Танимото стабильности по записи, признакам и Ляпунову;
– индекс Танимото расстояния между вычисленным и правильным подмножеством признаков;
– процент правильно определенных влияющих признаков;
– процент правильно определенных невлияющих признаков;
– процент невключенных влияющих признаков;
– процент включенных невлияющих признаков.

Отдельным является вопрос о порядках расчета индекса Танимото для вычисления показателей стабильности – сколько необходимо переменных/наблюдений изменить? По мнению авторов, целесообразно не фокусироваться на одном показателе, а построить график в системе координат «процент сохраненной изначальной выборки – показатель стабильности», что позволит более точно оценивать устойчивость метода отбора признаков.

Таким образом, для определения эффективности некоторого алгоритма отбора признаков предлагается следующая методика:
– генерируется по одним начальным условиям (количество переменных, количество значимых переменных, количество наблюдений, форма и параметры зависимости между переменными, распределение ошибок и переменных) 10 датафреймов;
– в датафрейме выделяется тренировочная часть (50 %), на которой решается задача отбора признаков, идентифицируется выделенное подмножество признаков;
– по выделенному и правильному набору признаков рассчитывается индекс Танимото и 4 рассмотренных выше процентных показателя;
– фиксируются три показателя затрат ресурсов;
– путем увеличения тренировочной базы данных до изначально сгенерированной с шагом в 5 % и уменьшения до 10 % от исходной рассчитывается динамика индекса Танимото стабильности по записи;
– путем добавления в тренировочную базу случайно сгенерированных признаков с шагом в 1 до двукратного увеличения базы данных или снижения числа незначимых признаков до 0 рассчитывается динамика индекса Танимото стабильности по признакам;
– путем одновременного увеличения/уменьшения базы данных и количества незначимых признаков рассчитывается динамика индекса Танимото стабильности по Ляпунову;
– далее показатели усредняются по каждому из 10 сгенерированных датафреймов.

Результаты

Для апробации представленной методики был поставлен следующий эксперимент: в R был сгенерирован искусственный датафрейм (полный код доступен на https://github.com/acheremuhin/FS_methodics_experiment), в котором сгенерированные переменные X и Y, распределенные по нормальному закону, влияют на переменную Z с некоторой нормально распреде-

ленной аддитивной ошибкой, а также присутствуют три нормально распределенные шумовые переменные; и с помощью пакета FSinR [16] решается задача отбора признаков 6 способами: методом муравьиной колонии [17], генетическим алгоритмом [18], методом «Лас-Вегас» [19], методом имитации отжига [20], методом поиска с запретами [21], методом «китовой охоты» [22].

Показатели результативности и затрат ресурсов представлены в табл. 1.

Таблица 1

Показатели ресурсоемкости и результативности решения задачи отбора признаков разными алгоритмами

Table 1

Indicators of resource intensity and effectiveness of solving the problem of feature selection by different algorithms

Метод	Время работы алгоритма, с	Объем потребляемой памяти, Кб	Индекс Танимото	Процент правильно выбранных значимых переменных, %	Процент правильно невыбранных незначимых переменных, %	Процент ошибочно невыбранных значимых переменных, %	Процент ошибочно выбранных незначимых переменных, %
Муравьиная колония	1,59	1097357,6	0,4	100	0	0	100
Генетический алгоритм	3,03	1012	0,4	100	0	0	100
«Лас-Вегас»	0,95	738,4	0,41	100	3	0	97
Имитация отжига	0,9	764	0,5	100	30	0	70
Поиск с запретами	1,94	784	0,4	100	0	0	100
«Китовая охота»	1,07	781,6	0,45	100	17	0	83

Далее была оценена стабильность алгоритмов по базе данных на основании индексов Танимото (табл. 2).

После этого была оценена стабильность алгоритмов по признакам на основании индексов Танимото (табл. 3).

И в целом была оценена стабильность алгоритмов по Ляпунову на основании индексов Танимото (табл. 4).

Обсуждение

На основе анализа таблиц 1–4 можно сделать следующие выводы:

– самым долгим по времени работы является алгоритм генетического алгоритма, наиболее затратным по объему памяти – алгоритм муравьиной колонии;

Таблица 2

Динамика индекса Танимото устойчивости по наблюдениям алгоритма методом отбора признаков

Table 2

Dynamics of the Tanimoto stability index based on observations of the algorithm by feature selection method

Размер датафрейма от естественного, %	Муравьиная колония	Генетический алгоритм	«Лас-Вегас»	Имитация отжига	Поиск с запретами	«Китовая охота»
10	1,00	1,00	0,96	0,75	1,00	0,79
20	1,00	1,00	1,00	0,73	1,00	0,83
30	1,00	1,00	1,00	0,77	1,00	0,86
40	1,00	1,00	0,96	0,72	1,00	0,76
50	1,00	1,00	0.94	0,77	1.00	0,79
60	1,00	1,00	0,98	0,76	1,00	0,76
70	1,00	1,00	1,00	0,65	1,00	0,76
80	1,00	1.00	1,00	0,67	1,00	0,76
90	1,00	1,00	0,98	0,68	1,00	0,81
110	1,00	1,00	0,96	0,73	1,00	0,77
120	1,00	1,00	1,00	0,75	1,00	0,74
130	1,00	1,00	0,96	0,67	1,00	0,76
140	1,00	1,00	1,00	0,68	1,00	0,74
150	1,00	1,00	0,92	0,76	1,00	0,76
160	1,00	1,00	1,00	0,71	1,00	0,78
170	1,00	1,00	1,00	0,62	1,00	0,72
180	1,00	1,00	0,94	0,75	1,00	0,77
190	1,00	1,00	0,98	0,73	1,00	0,82
200	1,00	1,00	0,96	0,67	1,00	0,75

Таблица 3

Динамика индекса Танимото устойчивости по признакам алгоритма методом отбора признаков

Table 3

Dynamics of the Tanimoto index of stability according to the characteristics of the algorithm by the method of feature selection

Количество шумовых переменных в датафрейме по сравнению с исходным	Муравьиная колония	Генетический алгоритм	«Лас-Вегас»	Имитация отжига	Поиск с запретами	«Китовая охота»
–3	0,80	0,80	0,80	0,74	0,80	0,70
–2	0,60	0,60	0,60	0,71	0,60	0,56
–1	0,40	0,40	0,40	0,46	0,40	0,46
+1	0,83	0,83	0,78	0,55	0,83	0,59
+2	0,71	0,71	0,68	0,55	0,71	0,56
+3	0,63	0,63	0,58	0,41	0,63	0,54
+4	0,56	0,56	0,59	0,43	0,56	0,50
+5	0,50	0,50	0,46	0,42	0,50	0,42

Таблица 4

Динамика индекса Танимото устойчивости алгоритма
методом отбора признаков

Table 4

Dynamics of the Tanimoto stability index according to the algorithm
by feature selection method

Шаг изменения датафрейма	Муравьиная колония	Генетический алгоритм	«Лас-Вегас»	Имитация отжига	Поиск с запретами	«Китовая охота»
1	0,83	0,83	0,77	0,64	0,83	0,66
2	0,71	0,71	0,72	0,54	0,71	0,57
3	0,63	0,63	0,63	0,40	0,63	0,53
4	0,56	0,56	0,58	0,47	0,56	0,53
5	0,50	0,50	0,45	0,37	0,50	0,49
6	0,45	0,45	0,44	0,34	0,45	0,39
7	0,42	0,42	0,44	0,29	0,42	0,39
8	0,38	0,38	0,42	0,33	0,38	0,33
9	0,36	0,36	0,39	0,29	0,36	0,33
10	0,33	0,33	0,37	0,29	0,33	0,33

– почти все алгоритмы включают все 5 переменных (две влияющие и три шумовые) в состав факторов, влияющих на зависимую переменную, при этом все алгоритмы правильно включили все действительно влияющие переменные;

– метод «имитации отжига» является самым перспективным, он в некоторых случаях отсекал одну шумовую переменную;

– методы муравьиной колонии, генетического алгоритма, поиска с запретами показали абсолютную стабильность по наблюдениям и не меняли результаты вычислений при сокращении/увеличении выборки;

– все методы показали неустойчивость к включению/удалению шумовых признаков, что объясняется низкой различительной способностью алгоритмов в данных условиях. Этим также объяснены низкие значения стабильности по Ляпунову.

В целом, можно констатировать, что в условиях проведенного эксперимента наиболее эффективным стал метод поиска с запретами – он не тратит больших временных и вычислительных ресурсов, не отстает от остальных по показателям результативности и показывает стабильность по признакам.

Заключение

В данной статье была апробирована методика оценки результатов решения задачи отбора признаков для задачи регрессии. Дальнейшими направлениями исследований будет:

– разработка и апробация аналогичной методики для задач классификации;

– доработка и апробация аналогичной методики на реальных задачах;

– реализация данной методики в виде программного пакета на языке R с оптимизированным кодом;

– продолжение экспериментов и разработка рекомендаций по выбору алгоритмов отбора признаков в разных ситуациях.

Список литературы

1. **Liao T. et al.** Are we learning yet? a meta review of evaluation failures across machine learning // Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2). 2021.
2. **Flach P.** Performance evaluation in machine learning: the good, the bad, the ugly, and the way forward // Proceedings of the AAAI conference on artificial intelligence. 2019. Vol. 33, № 01. P. 9808–9814.
3. **Lazebnik T., Rosenfeld A.** A new definition for feature selection stability analysis // Annals of Mathematics and Artificial Intelligence. 2024. P. 1–18.
4. **Sağbaş E. A.** Performance Evaluation of Feature Selection Methods for Sentiment Classification in Amazon Product Reviews // International Artificial Intelligence And Data Science Congress (ICADA'23). 2023. P. 2.
5. **Mahendran N. et al.** Machine learning based computational gene selection models: a survey, performance evaluation, open issues, and future research directions // Frontiers in genetics. 2020. Vol. 11. P. 603808.
6. **Mohapatra P., Chakravarty S., Dash P. K.** Microarray medical data classification using kernel ridge regression and modified cat swarm optimization based gene selection system // Swarm and Evolutionary Computation. 2016. Vol. 28. P. 144–160.
7. **Abinash M. J., Vasudevan V.** A study on wrapper-based feature selection algorithm for leukemia dataset // Intelligent Engineering Informatics: Proceedings of the 6th International Conference on FICTA. Springer Singapore, 2018. P. 311–321.
8. **Hancer E., Xue B., Zhang M.** Differential evolution for filter feature selection based on information theory and feature ranking // Knowledge-Based Systems. 2018. Vol. 140. P. 103–119.
9. **Ang J. C. et al.** Supervised, unsupervised, and semi-supervised feature selection: a review on gene selection // IEEE/ACM transactions on computational biology and bioinformatics. 2015. Vol. 13. № 5. P. 971–989.
10. **Saeys Y., Inza I., Larranaga P.** A review of feature selection techniques in bioinformatics // Bioinformatics. 2007. Vol. 23, № 19. P. 2507–2517.
11. **Xin B. et al.** Stable feature selection from brain sMRI // Proceedings of the AAAI Conference on Artificial Intelligence. 2015. Vol. 29, № 1.
12. **Bolón-Canedo V., Sánchez-Maroño N., Alonso-Betanzos A.** A review of feature selection methods on synthetic data // Knowledge and information systems. 2013. Vol. 34. P. 483–519.
13. **Sulistiani H. et al.** Performance evaluation of feature selections on some ML approaches for diagnosing the narcissistic personality disorder // Bulletin of Electrical Engineering and Informatics. 2024. Vol. 13, № 2. P. 1383–1391.
14. **Mohammadi M. et al.** Robust and stable gene selection via maximum–minimum correntropy criterion // Genomics. 2016. Vol. 107, № 2–3. P. 83–87.
15. **Черемухин А. Д.** Устойчивость алгоритмов отбора признаков к ошибкам второго рода // Вестник кибернетики. 2021. № 4 (44). С. 78–82. DOI 10.34822/1999-7604-2021-4-78-82. EDN JRYVAF.
16. **Aragón-Royón F. et al.** FSinR: an exhaustive package for feature selection // arXiv preprint arXiv:2002.10330. 2020.
17. **Kashef S., Nezamabadi-pour H.** An advanced ACO algorithm for feature subset selection // Neurocomputing. 2015. Vol. 147. P. 271–279.
18. **Yang J., Honavar V.** Feature subset selection using a genetic algorithm //IEEE Intelligent Systems and their Applications. 1998. Vol. 13, № 2. P. 44–49.

19. **Liu H., Setiono R.** Feature selection and classification–a probabilistic wrapper approach // Industrial and engineering applications or artificial intelligence and expert systems. CRC Press, 2022. P. 419–424.
20. **Posario F., Thangadurai K.** Simulated Annealing Algorithm for Feature Selection // International Journal of Computers & Technology. 2016. Vol. 15, № 2. P. 6471–6479.
21. **Glover F.** Tabu search—part I // ORSA Journal on computing. 1989. Vol. 1, № 3. P. 190–206.
22. **Zamani H., Nadimi-Shahraki M. H.** Feature selection based on whale optimization algorithm for diseases diagnosis // International Journal of Computer Science and Information Security. 2016. Vol. 14, № 9. P. 1243.

References

1. **Liao T. et al.** Are we learning yet? a meta review of evaluation failures across machine learning. *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*. 2021.
2. **Flach P.** Performance evaluation in machine learning: the good, the bad, the ugly, and the way forward. *Proceedings of the AAAI conference on artificial intelligence*, 2019, vol. 33, no. 1, pp. 9808–9814.
3. **Lazebnik T., Rosenfeld A.** A new definition for feature selection stability analysis. *Annals of Mathematics and Artificial Intelligence*, 2024, pp. 1–18.
4. **Sağbaş E. A.** Performance Evaluation of Feature Selection Methods for Sentiment Classification in Amazon Product Reviews. *International Artificial Intelligence And Data Science Congress (ICADA'23)*, 2023, p. 2.
5. **Mahendran N. et al.** Machine learning based computational gene selection models: a survey, performance evaluation, open issues, and future research directions. *Frontiers in genetics*, 2020, vol. 11, p. 603808.
6. **Mohapatra P., Chakravarty S., Dash P. K.** Microarray medical data classification using kernel ridge regression and modified cat swarm optimization based gene selection system. *Swarm and Evolutionary Computation*, 2016, vol. 28, pp. 144–160.
7. **Abinash M. J., Vasudevan V.** A study on wrapper-based feature selection algorithm for leukemia dataset. In: *Intelligent Engineering Informatics: Proceedings of the 6th International Conference on FICTA*. Springer Singapore, 2018, pp. 311–321.
8. **Hancer E., Xue B., Zhang M.** Differential evolution for filter feature selection based on information theory and feature ranking. *Knowledge-Based Systems,* 2018, vol. 140, pp. 103–119.
9. **Ang J. C. et al.** Supervised, unsupervised, and semi-supervised feature selection: a review on gene selection. *IEEE/ACM transactions on computational biology and bioinformatics*, 2015, vol. 13, no. 5, pp. 971–989.
10. **Saeys Y., Inza I., Larranaga P.** A review of feature selection techniques in bioinformatics. *Bioinformatics*, 2007, vol. 23, no. 19, pp. 2507–2517.
11. **Xin B. et al.** Stable feature selection from brain sMRI. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2015, vol. 29, no. 1.
12. **Bolón-Canedo V., Sánchez-Maroño N., Alonso-Betanzos A.** A review of feature selection methods on synthetic data. *Knowledge and information systems,* 2013, vol. 34, pp. 483–519.
13. **Sulistiani H. et al.** Performance evaluation of feature selections on some ML approaches for diagnosing the narcissistic personality disorder. *Bulletin of Electrical Engineering and Informatics*, 2024, vol. 13, no. 2, pp. 1383–1391.
14. **Mohammadi M. et al.** Robust and stable gene selection via maximum–minimum correntropy criterion. *Genomics*, 2016, vol. 107, no. 2–3, pp. 83–87.

15. **Cheremuhin, A. D.** Stability of Algorithms for Feature Selection to Type II Errors. *Cybernetics Bulletin*. 2021, no. 4(44), pp. 78–82. DOI 10.34822/1999-7604-2021-4-78-82. EDN JRYVAF.
16. **Aragón-Royón F. et al.** FSinR: an exhaustive package for feature selection. In: arXiv preprint arXiv:2002.10330. 2020.
17. **Kashef S., Nezamabadi-pour H.** An advanced ACO algorithm for feature subset selection. *Neurocomputing*, 2015, vol. 147, pp. 271–279.
18. **Yang J., Honavar V.** Feature subset selection using a genetic algorithm. *IEEE Intelligent Systems and their Applications*, 1998, vol. 13, no. 2, pp. 44–49.
19. **Liu H., Setiono R.** Feature selection and classification–a probabilistic wrapper approach. In: *Industrial and engineering applications or artificial intelligence and expert systems*. CRC Press, 2022, pp. 419–424.
20. **Posario F., Thangadurai K.** Simulated Annealing Algorithm for Feature Selection. *International Journal of Computers & Technology*, 2016, vol. 15, no. 2, pp. 6471–6479.
21. **Glover F.** Tabu search—part I. *ORSA Journal on computing*, 1989, vol. 1, no. 3, pp. 190–206.
22. **Zamani H., Nadimi-Shahraki M. H.** Feature selection based on whale optimization algorithm for diseases diagnosis. *International Journal of Computer Science and Information Security*, 2016, vol. 14, no. 9, pp. 1243.

Информация об авторе

Черемухин Артем Дмитриевич, кандидат экономических наук, доцент

Рейн Андрей Давыдович, кандидат экономических наук, доцент

Information about the Authors

Artem D. Cheremuhin, Ph.D. in Economics, Associate Professor

Andrey D. Rein, Ph.D. in Economics, Associate Professor

Статья поступила в редакцию 26.02.2025;
одобрена после рецензирования 05.04.2025; принята к публикации 05.04.2025

The article was submitted 26.02.2025;
approved after reviewing 05.04.2025; accepted for publication 05.04.2025