

Систематизация методов применения нейронных сетей для решения задач отбора признаков

А. Д. Черемухин

Нижегородский государственный инженерно-экономический университет

ngie.u.cheremuhin@yandex.ru

Аннотация. В работе представлен обзор современных нейросетевых методов отбора признаков в задачах анализа данных. Рассматриваются подходы, в которых процедура отбора интегрирована в процесс обучения модели и реализуется на основе архитектурных и алгоритмических особенностей нейронных сетей. Установлено, что, несмотря на различия в алгоритмах, все рассмотренные подходы относятся к классу встроенных методов отбора признаков, однако существенно различаются по уровню интеграции в архитектуру модели. Предложенное структурное разбиение нейронных сетей рассматривается как основа для построения таксономии нейросетевых методов отбора признаков и требует дальнейшей верификации на расширенном наборе алгоритмов.

Ключевые слова: отбор признаков; нейронные сети; встроенные методы; архитектура модели; латентное представление; машинное обучение; анализ данных

I. ВВЕДЕНИЕ

Отбор признаков представляет собой фундаментальную подзадачу машинного обучения, направленную на сокращение размерности исходного пространства предикторов при сохранении или повышении качества модели. Формально задача отбора признаков может быть определена как задача выделения подмножества признаков $S \subset \{1, \dots, p\}$, обеспечивающего наилучшее качество аппроксимации зависимости между независимыми переменными и откликом. В вероятностной постановке предполагается существование истинного множества релевантных признаков S_0 , для которого выполняется условие условной независимости отклика от нерелевантных переменных:

$$Y \perp X_{S^c} \mid X_{S_0} \quad (1)$$

что означает отсутствие дополнительной информации о зависимой переменной в признаках, не входящих в S_0 . В функциональной форме задача отбора признаков может быть представлена как задача минимизации функции потерь по подмножеству признаков:

$$\min_{S \subset \{1, \dots, p\}} L(f_\theta(X_S), Y). \quad (2)$$

Актуальность задачи отбора признаков обусловлена как теоретическими, так и практическими причинами. С одной стороны, рост объёмов данных и числа доступных признаков приводит к увеличению размерности пространства признаков, что усиливает эффекты избыточности, мультиколлинеарности и шума, ухудшающие обобщающую способность моделей. С другой стороны, вычислительная сложность полного

перебора всех возможных подмножеств признаков имеет экспоненциальный характер и составляет 2^p , что делает прямое решение задачи невозможным уже при умеренных значениях p . В этих условиях методы отбора признаков выступают как необходимый инструмент построения эффективных моделей, позволяющий одновременно повысить качество прогнозирования, снизить вычислительные затраты и обеспечить интерпретируемость получаемых решений.

На основании проведённых ранее исследований методов отбора признаков и их системного анализа были сформированы краткие идентификационные описания основных классов методов отбора признаков:

- Фильтрационные методы основаны на независимой от модели оценке релевантности каждого признака с последующим формированием подмножества по заданному правилу;
- Обёрточные методы используют модель в качестве элемента оценки качества и реализуют итеративный поиск оптимального подмножества с использованием механизмов обратной связи;
- Встроенные методы интегрируют процедуру отбора непосредственно в процесс обучения модели, объединяя оптимизацию её параметров и структуры;
- Ансамблевые методы формируют итоговое решение путём агрегации результатов нескольких независимых процедур отбора, повышая устойчивость и согласованность результата;
- Гибридные методы представляют собой комбинации указанных подходов, в рамках которых одновременно используются различные механизмы оценки, поиска и агрегации признаков, что отражает современную тенденцию к усложнению и интеграции алгоритмов отбора признаков.

II. ПОСТАНОВКА ЗАДАЧИ

С развитием методов анализа данных и увеличением сложности зависимостей между переменными нейронные сети получили широкое распространение как универсальный инструмент моделирования. Их способность аппроксимировать сложные нелинейные зависимости делает их предпочтительным выбором в задачах регрессии и классификации, особенно в условиях высокой размерности и сложной структуры данных. Однако рост числа признаков и наличие шумовых переменных в таких задачах приводит к

необходимости решения проблемы отбора признаков уже на уровне самой нейросетевой модели.

Одним из первых механизмов, частично решающих данную задачу, стал dropout, изначально предложенный как метод регуляризации. Его работа основана на случайном исключении части входных или скрытых нейронов в процессе обучения, что фактически приводит к обучению модели на случайных подмножествах признаков. В результате нейронная сеть становится менее чувствительной к отдельным входным переменным и начинает опираться на более устойчивые и информативные зависимости. С этой точки зрения dropout можно рассматривать как неявный встроенный механизм отбора признаков, реализующий стохастическое разреживание входного пространства.

Тем не менее, данный подход обладает рядом существенных ограничений. В частности, dropout не формирует явного подмножества значимых признаков, не позволяет интерпретировать их важность и не обеспечивает контроля качества отбора. Более того, его эффективность существенно зависит от параметров модели и структуры данных, а в ряде случаев он не приводит к снижению влияния нерелевантных переменных. Это ограничивает его применение в задачах, где требуется не только повышение качества модели, но и интерпретируемость или строгий контроль структуры входных данных.

В связи с этим в последние годы активно развиваются специализированные методы отбора признаков, встроенные непосредственно в нейросетевые архитектуры. Их ключевая идея заключается в явном моделировании механизма включения и исключения признаков в процессе обучения, что позволяет одновременно оптимизировать параметры модели и структуру используемого подмножества переменных. Целью настоящей работы является краткий обзор таких методов, их систематизация и анализ с точки зрения используемых механизмов отбора признаков и их интеграции в архитектуру нейронных сетей.

III. МАТЕРИАЛЫ И МЕТОДЫ ЭКСПЕРИМЕНТА

A. Методика поиска работ

Поиск научных публикаций осуществлялся с использованием инструмента глубокого анализа (Deep Research), реализованного на базе модели ChatGPT 5.4 Plus. Сформировались целевые поисковые запросы, направленные на выявление работ, опубликованных преимущественно в 2020–2025 гг., в которых задача отбора признаков решается непосредственно с использованием нейронных сетей. Поиск охватывал открытые научные репозитории (в первую очередь arXiv), а также рецензируемые журналы и материалы конференций.

В итоговую выборку включались только те публикации, в которых нейросеть выступает центральным механизмом отбора признаков. Для каждой работы дополнительно проверялись постановка задачи и используемый математический аппарат, что обеспечило формирование релевантного и методологически однородного набора современных методов нейросетевого отбора признаков.

B. Описание современных нейросетевых методов отбора признаков

Ниже дано краткое описание работ с самыми важными и интересными, на взгляд автора, нейросетевыми алгоритмами отбора признаков.

В работе [1] рассматривается задача отбора признаков в глубоких нейронных сетях с акцентом на теоретическую состоятельность процедуры отбора. Авторы вводят класс аналитических нейросетей, в которых отображение задается как композиция аналитических функций; отбор признаков связывается с параметрами первого слоя, так как именно через него осуществляется взаимодействие с входными переменными. Базовая идея заключается в использовании групповой регуляризации, где параметры, соответствующие каждому признаку, объединяются в группы. Классический функционал имеет вид

$$\hat{\alpha}_n = \arg \min_{\alpha} \frac{1}{n} \sum_{i=1}^n (Y_i - f_{\alpha}(X_i))^2 + \lambda_n \sum_k \|u_{[:,k]}\|, \quad (3)$$

где каждая группа $u_{[:,k]}$ соответствует отдельному признаку. Далее вводится адаптивная версия с весами, зависящими от предварительной оценки. Ключевой результат работы — доказательство состоятельности отбора признаков: при росте выборки веса для незначимых признаков стремятся к бесконечности, что приводит к занулению соответствующих параметров, тогда как значимые признаки сохраняются.

В работе [2] предлагается двухэтапный подход к отбору признаков в условиях сверхвысокой размерности, основанный на сочетании нейросетевого представления и статистического скрининга. На первом этапе используется автоэнкодер, который строит низкоразмерное представление данных:

$$L_r(\theta) = \frac{1}{n} \sum_{i=1}^n \|x_i - \Psi(\Phi(x_i))\|_2^2, \quad (4)$$

где Φ – кодировщик, Ψ – декодер.

На втором этапе для каждого исходного признака вычисляется мера зависимости с латентным представлением – многомерная ранговая дистанционная корреляция. Признаки ранжируются по значению этой меры, и выбираются наиболее информативные. Таким образом, нейросеть используется как инструмент извлечения структуры, а сам отбор реализуется через независимую статистическую процедуру.

В работе [3] предлагается встроенный нейросетевой метод отбора признаков для данных одно-клеточной транскриптомики, основанный на итеративном исключении признаков с контролем ложных открытий. Основная идея заключается в оценке важности признака через чувствительность функции потерь к входу:

$$S_j = \frac{1}{n} \sum_{i=1}^n \left| \frac{\partial L(y_i, O_i)}{\partial x_{ij}} \right|. \quad (5)$$

Признаки с наименьшими значениями S_j последовательно удаляются. Для контроля качества отбора вводится оценка доли ложных открытий (FDR), где используются дополнительные «фиктивные» признаки для аппроксимации числа нерелевантных переменных. Алгоритм завершает исключение

признаков, когда оценка FDR становится ниже заданного порога.

В работе [4] предлагается архитектура нейросети с механизмом стохастических «гейтов» и использованием knockoff-подхода для отбора признаков с контролем FDR. Для каждого признака вводится случайная переменная, определяющая его участие:

$$g_j = \max(0, \min(1, \mu_j + \varepsilon_j)), \varepsilon_j \sim \mathcal{N}(0, \sigma^2), \quad (6)$$

что является непрерывной аппроксимацией ℓ_0 -регуляризации. Модель обучается путем минимизации эмпирического риска с учетом регуляризации:

$$\min_{\theta, \mu} \mathbb{E}[L(f_{\theta}(X \odot G), Y) + \lambda \|G\|_0]. \quad (7)$$

Для оценки важности используется разность между оригинальными и knockoff-признаками. Отбор осуществляется по пороговому правилу, обеспечивающему контроль FDR. Таким образом, архитектура объединяет вероятностное включение признаков и сравнительный анализ с контрольными переменными.

IV. РЕЗУЛЬТАТЫ

Нейросеть, применяемую для решения табличных задач машинного обучения с учителем, можно рассматривать как последовательность функциональных блоков: входное пространство признаков, первый слой (где признаки впервые преобразуются в параметры модели), стек скрытых слоев (формирующих представление), латентное пространство, выходной слой и процедура обучения. Такое представление позволяет анализировать методы отбора признаков не с точки зрения алгоритма, а с точки зрения того, в каком месте вычислительного графа происходит вмешательство.

Соответственно, любой нейросетевой метод отбора признаков можно интерпретировать как изменение одного или нескольких из этих блоков: либо меняется вход, либо параметры первого слоя, либо используется латентное представление, либо модифицируется процесс обучения. Применительно к рассмотренным работам получается следующее:

- В [1] изменение происходит на уровне параметризации первого слоя и процедуры обучения. Первый слой переинтерпретируется: веса, связанные с каждым признаком, рассматриваются как отдельные группы; блок входного преобразования не меняется структурно, но изменяется способ оценки и обновления его параметров, в результате чего незначимые признаки автоматически «выключаются» за счёт зануления соответствующих весов.
- В [2] изменение затрагивает блок латентного пространства и добавляется внешний этап анализа после нейросети. Сама архитектура модели не модифицируется, отбор признаков реализуется не внутри сети, а через использование результата работы скрытых слоёв, то есть через интерпретацию сформированного латентного пространства.
- В [3] изменение происходит на уровне процедуры обучения и взаимодействия модели с входными данными. Добавляется механизм оценки важности

признаков на основе чувствительности функции потерь к входу, вводится внешний итеративный цикл: признаки с наименьшей важностью последовательно удаляются, и модель переобучается – в целом, меняется логика использования модели в процессе обучения и отбора признаков.

- В [4] изменение затрагивает сразу несколько блоков: входное пространство, первый слой и процедуру обучения. Входной слой расширяется за счёт добавления искусственных признаков (knockoff), что изменяет исходное пространство признаков. В первый слой вводится механизм стохастических гейтов, управляющих включением признаков в модель, а также реализуется парная структура обработки признаков и их копий. Дополнительно модифицируется процедура обучения (вводится специальная схема обучения и критерий останова). В результате изменяется сама архитектура модели, и отбор признаков становится встроенной частью нейросети.

V. ЗАКЛЮЧЕНИЕ

Рассмотренные в работе подходы к отбору признаков на основе нейронных сетей, несмотря на различие используемых алгоритмических механизмов, объединяются общей принципиальной характеристикой: во всех случаях процедура отбора признаков интегрирована в процесс обучения модели или непосредственно использует её внутренние представления. Это позволяет отнести все проанализированные методы к классу встроенных методов отбора признаков.

При этом проведённый анализ показывает, что даже внутри класса встроенных методов нейросетевые подходы существенно неоднородны. Их различия проявляются не столько в используемых алгоритмах оптимизации или видах регуляризации, сколько в том, на каком уровне архитектуры модели реализуется механизм отбора.

Описанное в работе структурное описание нейросети как композиции функциональных блоков — входного пространства, первого слоя, скрытых преобразований, латентного представления, выходного слоя и процедуры обучения — позволяет формализовать данное различие и задать основу для систематизации методов отбора признаков. В рамках такой схемы любой алгоритм может быть описан через указание блока или набора блоков, в которые вносятся изменения, а также характера этих изменений. Это даёт возможность перейти от описательной классификации методов к более строгой, основанной на архитектурных и функциональных свойствах модели.

В то же время предложенное структурное разбиение носит предварительный характер и требует дальнейшей верификации. Проведённый анализ охватывает ограниченное число современных методов, что не позволяет утверждать полноту или универсальность предложенной схемы. Для её подтверждения необходимо расширение выборки алгоритмов, включая как классические, так и более современные нейросетевые подходы, а также методы, находящиеся на границе между различными классами (например, гибридные и ансамблевые решения). Дополнительного

исследования требует также вопрос о том, в какой степени различные уровни вмешательства в архитектуру модели коррелируют с качеством отбора признаков, устойчивостью результатов и вычислительной эффективностью и насколько они применимы для решения практических задач [5, 6].

СПИСОК ЛИТЕРАТУРЫ

- [1] Dinh V.C., Ho L.S. Consistent feature selection for analytic deep neural networks // *Advances in Neural Information Processing Systems*. 2020. Т. 33. С. 2420-2431.
- [2] Li K. et al. Deep feature screening: Feature selection for ultra high-dimensional data via deep neural networks // *Neurocomputing*. 2023. Т. 538. С. 126186.
- [3] Peng M. et al. scFSNN: a feature selection method based on neural network for single-cell RNA-seq data // *BMC genomics*. 2024. Т. 25. №. 1. С. 264.
- [4] Oh E., Lee H. DeepPIG: deep neural network architecture with pairwise connected layers and stochastic gates using knockoff frameworks for feature selection // *Scientific Reports*. 2024. Т. 14. №. 1. С. 15582.
- [5] Кафиев И.Р. К вопросу автоматизации процесса сортировки белых грибов с использованием нейронных сетей / И.Р. Кафиев, П.С. Романов, И.П. Романова // *Вестник НГИЭИ*. 2024. № 4(155). С. 34-49. – DOI 10.24412/2227-9407-2024-4-34-49. – EDN DENZQG
- [6] Использование нейронных сетей при прогнозировании выхода из строя асинхронного электродвигателя / Е.А. Рожков, Е.В. Белокур, И.А. Израфилов, И.В. Юдаев // *Вестник НГИЭИ*. 2026. № 2(177). С. 46-59. – DOI 10.24412/2227-9407-2026-2-46-59. – EDN TGKHVJ