

Способ диагностики послеоперационных осложнений в хирургии

© В.И. ГОРБАЧЕНКО, А.А. СОЛОМАХА

Пензенский государственный университет, Пенза, Российская Федерация

В России и зарубежных странах научные исследования в хирургии принято выполнять в соответствии с постулатами доказательной медицины. Однако реализация этих принципов требует глубоких знаний хирургии и математического моделирования. Авторы: хирург и программист в области машинного обучения разработали математические модели, участвующие в диагностике послеоперационных осложнений в хирургии. В данной работе исследована глубокая полносвязная нейронная сеть для диагностики послеоперационных осложнений на клиническом примере острого аппендицита. В качестве обучающего множества параметров использовался набор, разработанный авторами на основе реальных клинических данных, имеющий номер государственной регистрации в виде базы данных. Для выбора значимых признаков предложен подход на межквантильном размахе F1-меры и к кодированию составных категориальных признаков, отличающийся компактным представлением. Для предварительной обработки обучающих данных предлагается использовать повышающий автоэнкодер. Автоэнкодер преобразует выбранные признаки в пространство более высокого измерения, что, согласно теореме Ковера, облегчает классификацию признаков. Глубокая нейронная сеть реализована с использованием библиотек Keras и TensorFlow. Для обучения нейросети применен алгоритм Adam с адаптивной скоростью обучения. Для снижения эффекта переобучения использован современный метод регуляризации — dropout. Проведены анализ и выбор метрик качества классификатора. Для оценки характеристик нейросети использована кросс-валидация по k блокам. Обученная нейронная сеть показала высокие диагностические показатели на тестовом наборе данных.

Ключевые слова: медицинская диагностика; послеоперационные осложнения; глубокие нейронные сети; отбор информативных признаков; обучение нейронной сети; переобучение; повышающий автоэнкодер; доказательная медицина

Diagnostic Options for Postoperative Complications in Surgery

© V.I. GORBACHENKO, A.A. SOLOMAKHA

Penza State University, Penza, Russian Federation

Currently, scientific research is usually carried out in accordance with the postulates of evidence-based medicine in Russia and foreign countries. However, the implementation of these principles requires deep knowledge of surgery and mathematical modeling. Authors: a surgeon and a programmer developed mathematical models involved in the diagnosis of postoperative complications in surgery. In this paper, we investigated a deep, fully connected neural network for the diagnosis of postoperative complications on the clinical example of acute appendicitis. As a training set of parameters, we used a set developed by the authors on the basis of real clinical data, which has a state registration number in the form of a database, and includes a knowledge base. The interquartile range of the F1 measure is proposed for the selection of significant features. An approach to coding composite categorical features, characterized by a compact representation, is proposed. For pre-processing of training data, it is proposed to use a step-up autoencoder. The autoencoder converts the selected functions into a higher-dimensional space, which, according to Kover's theorem, facilitates the classification of features. The neural network is implemented using the Keras and TensorFlow libraries. To train the neural network, the Adam algorithm with adaptive learning speed is used. To reduce the effect of overfitting, a modern regularization method — dropout was used. The analysis and selection of the classifier quality metrics are carried out. To evaluate the characteristics of the neural network, k-block cross-validation was used. The trained neural network showed high diagnostic performance on the test data set.

Keywords: medical diagnostics; postoperative complications; deep neural networks; selection of informative features; neural network training; retraining; increasing autoencoder; evidence-based medicine

В данной работе рассматривается нейросетевая диагностика послеоперационных осложнений на клиническом примере острого аппендицита. Проблемы ранней диагностики, лечения, профилактики и прогнозирования осложнений острого аппендицита являются актуальными в клинической хирургии. Гнойно-воспалительные осложнения после аппендэктомии встречаются у 2,7%–39,1% больных [1–2]. До сих пор летальность при остром аппендиците варьируется от 0,1% до 1,6% [1–2]. Решение перечисленных проблем требует разработки методов предсказания послеоперационных осложнений острого аппендицита. Рядом авторов разработаны искусственные нейронные сети, диагностирующие наличие острого аппендицита [3–4].

Проведено исследование факторов риска, связанных с осложнениями острого аппендицита [5], изучены конкретные виды осложнений [6], сведения о нормальном и осложнённом исходах операции представлены нами в разработанной базе данных. Однако математические модели, позволяющие диагностировать риск развития послеоперационных осложнений в хирургии, в соответствии с принципами доказательной медицины, отсутствуют, хотя предсказательная медицина и нейросетевое моделирование применяются в хирургии [7–10]. В настоящее время очень популярны нейронные сети глубокой архитектуры, использующие специфические алгоритмы обучения и функции активации [11–12]. Глубокие нейросети меньше склонны к пере-

обучению и обеспечивают высокую точность по сравнению с обычными нейросетями [12].

Цель

Разработка способа диагностики послеоперационных осложнений в хирургии на основе анализа медицинских данных с помощью нейросетевого моделирования.

2. ОТБОР И ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ПРИЗНАКОВ

Для диагностики послеоперационных осложнений в хирургии будем руководствоваться набором признаков зарегистрированной нами базы данных.

2.1. Метод отбора лучшей комбинации признаков

Для отбора обучающих признаков предложен подход, основанный на методе обёртки [14, 15]. Для отбора значимых признаков предлагается подход, основанный на межквантильном размахе. Формируется таблица размера $N \times M$, где N — количество комбинаций признаков, которые необходимо проверить, а M — количество экспериментов для одной комбинации. Используется M экспериментов для снижения влияния случайного характера обучения сети. Результаты обучения сети в виде F1-меры [16] записываются в таблицу. Использование F1-меры позволяет получить более стабильный результат, чем среднее значение, которое сильно зависит от предельных значений выборки. После формирования таблицы значений, для каждой строки вычисляются нижний (0.25 квантиль, или первый Q1) и верхний (0.75 квантиль, или третий Q3) квантили [16]. Выбирается комбинация признаков (строка таблицы), обеспечивающая наименьший разброс метрики F1-меры — межквантильный размах (разницу между 0.25 и 0.75 квантилями).

2.2. Преобразование признаков

После отбора признаков необходимо произвести преобразование данных: заполнение пропусков в обучающих данных, кодирование и масштабирование данных. Для масштабирования числовых признаков использовали масштабирование по минимуму (min-max scaling) или нормализацию — приведение переменных к одному диапазону [17]. Входные переменные нормализовались так, чтобы признаки изменялись в диапазоне $[0, 1]$. Каждая входная переменная нормализовалась независимо от других переменных. Масштабы входных и выходных переменных не связаны. Использовали линейную нормализацию

$$p_{ij} = \frac{(x_{ij} - x_{j\min})(b - a)}{x_{j\max} - x_{j\min}} + a,$$

где x_{ij} — нормализованное значение j -го признака в i -ом примере, x_j — исходное значение j -го признака в i -ом примере, $[a, b]$ — допустимый диапазон изменения нормализованных признаков $[0, 1]$,

$[x_{j\min}, x_{j\max}]$ — диапазон изменения исходных значений j -го признака.

Для кодирования качественных переменных использовали приём, известный в современном машинном обучении, как One-HotEncoding (унитарное кодирование) [18]. В таком случае, если категориальный признак принимал n значений, то он заменялся вектором из n значений. В строках таблицы, рассматриваемый признак заменялся единственной единицей в позиции, равной номеру значения признака. В остальных позициях располагались нули.

Набор данных для обучения [13] представлен в виде таблицы, имеющей 30 признаков с параметрами анализов крови, до и после операции, и 12 общих признаков, например, возраст, группа крови и так далее. Ниже представлены общие признаки: пол; возраст; койко-день; группа крови и резус-фактор; основной клинический диагноз; сопутствующий диагноз; осложнения; оперативные вмешательства; длительность оперативных вмешательств; электрокардиограмма (ЭКГ); клиническая оценка состояния больного; наличие осложнений.

Признак "пол" представлен в форме строки, имеющей одно из двух значений: "мужской" или "женский". Признак "возраст" — это целое число, являющееся непрерывным значением, большим нуля. Признак "койко-день" является целым числом, определяющим общее количество дней в больнице. Признак "группа крови и резус-фактор" представлен в виде строки. Признак "основной клинический диагноз" описывает в словесной форме диагноз больного. Признак "сопутствующий диагноз" — это дополнительный диагноз больного, если таковой имелся. Признак "осложнения" представляет собой словесное описание вида/видов осложнений у пациента, при наличии таковых. Данный признак будет полезен в случае, когда требуется помимо определения наличия осложнения, дополнительно, определить его вид. Признак "оперативные вмешательства" содержит информацию о виде оперативного вмешательства (например, "аппендэктомия"). Данный признак будет известен в полной мере лишь после проведения операции, поэтому его польза для предсказания осложнений до операции сомнительна. Признак "длительность оперативных вмешательств" представляет собой числовые данные, записанные в форме строки (например, "45 мин."). Признак "ЭКГ" представляет собой результаты электрокардиограммы, записанные в виде строки. Признак "клиническая оценка состояния больного" описывает, в форме строки, мнение врача о состоянии больного до операции. Последний признак, "наличие осложнений" представляет собой данные, представленные в форме строки с двумя вариантами: "да" или "нет". В случае, когда осложнения после операции были (у пациента в данном столбце написано "да"), то в столбце "осложнения" содержится вид/виды осложнений.

Таблица 1. Кодированные значения диагнозов
Table 1. Diagnostic values encoded with the proposed method

Исходное значение / Text value	Кодированное значение / Encoded vector
Острый флегмозный аппендицит / Acute phlegmonous appendicitis	[1, 0, 0, 0, 0, 0, 0, 0]
Острый аппендицит / Acute appendicitis	[0, 1, 0, 0, 0, 0, 0, 0]
Острый катаральный аппендицит / Acute catarrhal appendicitis	[0, 0, 1, 0, 0, 0, 0, 0]
острый катаральный аппендицит, тифлит / Acute catarrhal appendicitis, typhlitis	[0, 0, 1, 1, 0, 0, 0, 0]
Острый гангренозный аппендицит / Acute gangrenous appendicitis	[0, 0, 0, 0, 1, 0, 0, 0]
Острый гангренозный перфоративный аппендицит / Acute gangrenous perforated appendicitis	[0, 0, 0, 0, 0, 1, 0, 0]
Острый катаральный аппендицит, острый панкреатит / Acute catarrhal appendicitis, acute pancreatitis	[0, 0, 1, 0, 0, 0, 1, 0]
Острый деструктивный аппендицит / Acute destructive appendicitis	[0, 0, 0, 0, 0, 0, 0, 1]

Оставшиеся признаки из набора данных — анализы крови до и после операции. Перечисление доступных признаков представлено ниже: гемоглобин; эритроциты; цветовой показатель; лейкоциты; нейтрофилы палочкоядерные; нейтрофилы сегментоядерные; лимфоциты; моноциты; скорость оседания эритроцитов; общий билирубин; общий белок; амилаза; мочевины; креатинин; глюкоза.

Данные параметров анализов крови были представлены в виде действительных чисел. В нашей задаче требовалось диагностировать наличие осложнений.

Для облегчения работы с большим объёмом данных была использована библиотека pandas [19]. Воспользовавшись встроенными в библиотеку средствами, определяли пропуски в медицинских данных. В окончательном наборе данных присутствует 108 записей. Затем были отобраны столбцы без пропусков. Перечислим клинико-лабораторные показатели: пол; возраст; койко-день; основной клинический диагноз; гемоглобин; эритроциты; лейкоциты; нейтрофилы палочкоядерные; лимфоциты; наличие осложнений; пол; возраст; сопутствующий диагноз; гемоглобин; эритроциты; лейкоциты; нейтрофилы палочкоядерные; лимфоциты; наличие осложнений.

2.3. Метод кодирования составных категориальных признаков

Признак "основной клинический диагноз" представлен в форме строки и содержит восемь уникальных значений: острый флегмозный аппендицит; острый аппендицит; острый катаральный аппендицит; острый катаральный аппендицит, тифлит; острый гангренозный аппендицит; острый гангренозный перфоративный аппендицит; острый катаральный аппендицит, мезаденит; острый деструктивный аппендицит.

Заметим, что часть пациентов имеют два диагноза одновременно. В таком случае, необходимо решить, каким образом кодировать такие значения. Традиционное решение заключается в составлении вектора значений, длина которого равна количеству сочетаний диагнозов, и использовать кодирование унитарным кодом. Такой подход приведет к большой избыточности данных. Был предложен отличный от общепринятого

подхода способ кодирования категориальных признаков. Предлагается составить вектор, размерность которого будет равна размерности пространства простых диагнозов (а в общем случае значений). Далее, при наличии конкретного диагноза пациента выставить значение данного элемента вектора в 1, в противном случае в 0. Представим диагнозы больных острым аппендицитом для машинного обучения: острый флегмозный аппендицит; острый аппендицит; острый катаральный аппендицит; тифлит; острый гангренозный аппендицит; острый гангренозный перфоративный аппендицит; острый деструктивный аппендицит; острый панкреатит.

Используя описанный подход, произведём кодирование данных признаков. Результаты представлены в табл. 1.

2.4. Отбор значимых признаков для задачи выявления осложнений

После определения способа кодирования и масштабирования признаков выполнили предобработку. Для облегчения данной работы использовали библиотеку Scikit-learn [17, 20]. Используя предложенный в разделе 2.1 подход, было отобрано 8 информативных признаков:

пол;
 возраст;
 основной клинический диагноз;
 гемоглобин;
 эритроциты;
 лейкоциты;
 нейтрофилы палочкоядерные;
 лимфоциты.

К каждому из выбранных признаков путем перебора добавлялись по одному и другие признаки. С использованием F1-меры оценивалось качество работы нейросетевой модели. По результатам обучения 8 моделей, в качестве начального признака был выбран "основной клинический диагноз". Затем, путём полного перебора комбинаций, были выбраны наиболее значащие для алгоритма комбинации признаков:

основной клинический диагноз, пол, возраст, эритроциты;

основной клинический диагноз, пол, возраст, гемоглобин, лейкоциты, нейтрофилы палочкоядерные;

Полный перебор осуществлялся следующим образом. К признаку "основной клинический диагноз" добавлялись оставшиеся признаки. В результате было составлено 7 новых комбинаций:

основной клинический диагноз, пол;
основной клинический диагноз, возраст;
основной клинический диагноз, гемоглобин;
основной клинический диагноз, эритроциты;
основной клинический диагноз, лейкоциты;
основной клинический диагноз, нейтрофилы палочкоядерные;

основной клинический диагноз, лимфоциты.

Затем добавлялся третий признак:

основной клинический диагноз, пол, возраст;
основной клинический диагноз, пол, гемоглобин;
основной клинический диагноз, пол, эритроци-

ты;

основной клинический диагноз, пол, лейкоциты;
основной клинический диагноз, пол, нейтрофи-

лы палочкоядерные;

основной клинический диагноз, пол, лимфоциты;
основной клинический диагноз, нейтрофилы палочкоядерные, возраст;

основной клинический диагноз, нейтрофилы палочкоядерные, гемоглобин;

основной клинический диагноз, нейтрофилы палочкоядерные, эритроциты;

основной клинический диагноз, нейтрофилы палочкоядерные, лейкоциты;

основной клинический диагноз, нейтрофилы палочкоядерные, лимфоциты.

Этот процесс повторяется до тех пор, пока не будут перебраны все оставшиеся признаки.

Отбор признаков производился с помощью финальной модели машинного обучения, описанной в следующем разделе.

3. ИНСТРУМЕНТАРИЙ, АРХИТЕКТУРА И ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

Разработка и исследование нейронной сети проводились с использованием библиотеки Keras [21].

3.1. Выбор архитектуры нейронной сети

После ряда экспериментов с различными архитектурами, функциями активации, методами обучения и видами нейронов была выбрана нейронная сеть (рис. 1).

В слоях C1, C3, C5, C7 в качестве функции обработки применялся сумматор, вычисляющий взвешенную сумму $\sum w_i x_i$, где w_i и x_i — компоненты вектора весов и входного вектора слоя сети. В скрытых слоях использована функция активации ReLU (RectifiedLinearUnit)[17]

$$f(s) = \max(0, s) = \begin{cases} 0, & s < 0, \\ s, & s \geq 0, \end{cases}$$

где s — взвешенная сумма. В выходном слое используется асимметричная логистическая функция активации

$$f(s) = \frac{1}{1 + \exp(-s)},$$

где s — взвешенная сумма. Данная функция хорошо подходит в качестве выходной функции активации, так как она распределена от 0 до 1, что можно трактовать как вероятность наличия осложнений.

Количество нейронов в слоях классификатора определялось по формуле

где k — коэффициент, n — размерность входного вектора, N — вычисленное количество нейронов. Были экспериментально подобраны коэффициенты k :

для слоя C1 коэффициент $k = 1.0$;

для слоя C3 коэффициент $k = 0.5$;

для слоя C5 коэффициент $k = 0.2$.

В слое C7 имеется один нейрон, определяющий вероятность наличия осложнений.

Классификатор обучался методом Adam [11, 17]. Данный метод обучения имеет адаптивный параметр — скорость сходимости. Благодаря особенности метода, можно избежать попадания в часть локальных минимумов. Для инициализации начальных параметров модели использовался метод Ксавье (Xavierinitialization) [17]. Для преодоления переобучения сети использовалась перекрестная проверка (Cross-Validation) с разделением обучающего набора на три множества: обучающий набор (70%) — для обучения сети, проверочный (15%) — для подсчёта метрик качества на каждой итерации и останова обучения и тестовый (15%) — для оценки качества обученной модели.

Чтобы на каждой итерации обучения работать с лучшей моделью, использован callback-метод [22]. Если на текущей итерации на проверочном наборе точность нейросети не улучшилась, то сохраняются параметры нейросети с предыдущей итерации. Этот подход способствует предотвращению переобучения сети.

3.2. Регуляризация "прореживанием" нейронов

Для предотвращения переобучения нейронной сети использован специфический прием регуляризации — "прореживание" (dropout) нейронов [17]. В этом подходе на каждой итерации обучения для каждого нейрона (кроме нейронов выходного слоя) устанавливалась вероятность, с которой нейрон будет исключён из сети — уровень отключения (dropoutrate). Экспериментальным путём был настроен параметр p — уровень отключения нейронов:

для слоя C2 порог $p = 0.30$;

для слоя C4 порог $p = 0.15$;

для слоя C6 порог $p = 0.05$.

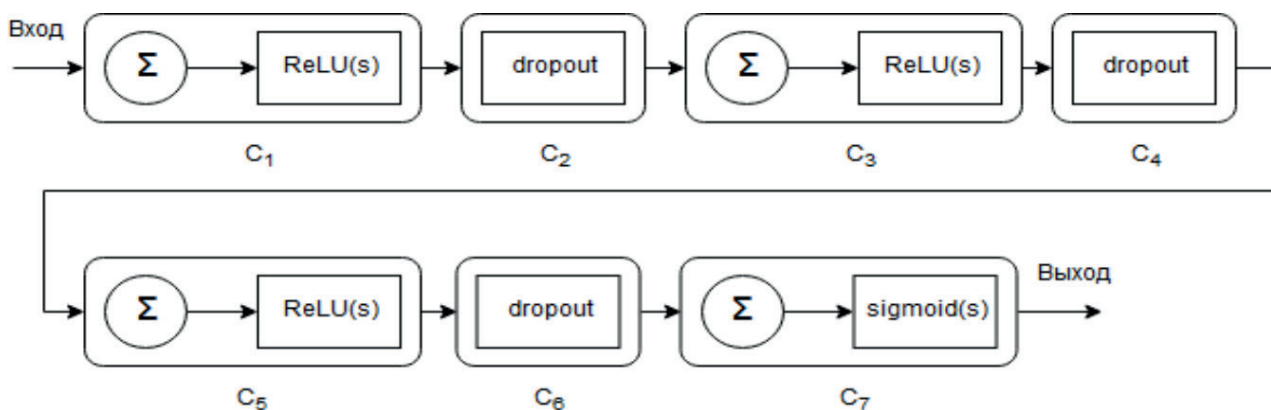


Рис. 1. Архитектура классификатора, построенная с использованием метода "прореживания" нейронов (dropout).
 Fig. 1. Architecture of the classifier built using the method of "dropout" neurons.

Алгоритм обучения строился следующим образом: на каждой новой итерации обучения для каждого нейрона определялась вероятность удаления. Если нейрон удаляется, то его выход полагается равным нулю. Нулевой выход нейрона означает, что нейрон фактически не участвует в вычислениях на данной итерации. При эксплуатации полученной модели участвовали все нейроны, однако, выход каждого нейрона умножался на вероятность сохранения, с которой нейрон оставляли при обучении. Применение обученной с использованием dropout сети эквивалентно усреднению по ансамблю сетей, что лучше, чем применение одной модели.

3.3. Предварительная обработка данных автоэнкодером

Предложено для повышения точности классификации применить повышающий (overcomplete) автоэнкодер. Такой автоэнкодер нелинейно преобразует входные векторы в пространство большей размерности. Согласно теореме Ковера [23], такое преобразование повышает вероятность линейного разделения признаков, следовательно, повышает точность классификации.

Для решения этой проблемы была построена вспомогательная нейронная сеть по типу повышающего автоэнкодера. Перед тем, как выполнять классификацию, данные будут отображены в пространство большей размерности, что теоретически увеличит вероятность успешной классификации. Затем, преобразованные данные будут поданы на вход классификатора, которому выполнить классификацию будет значительно легче.

Автоэнкодер состоит из двух частей: энкодера (encoder) и декодера (decoder). Обученный автоэнкодер на выходе выдаёт вектор значений очень близкий к тому, который был получен на входе. На рис. 2 представлены составные части автоэнкодера, на котором выполнялось разделение признаков.

Энкодер (рис. 2) включает скрытые слои, аналогичные слоям классификатора, описанного ранее. Выходной слой использует линейную функцию активации. На рис. 2 слои энкодера пронумерованы: E1, E2 и E3. Количество нейронов зависит от размерности входных признаков. Коэффициент k задаётся для каждого слоя:

- для слоя E1 коэффициент $k = 1.35$;
- для слоя E2 коэффициент $k = 1.75$;

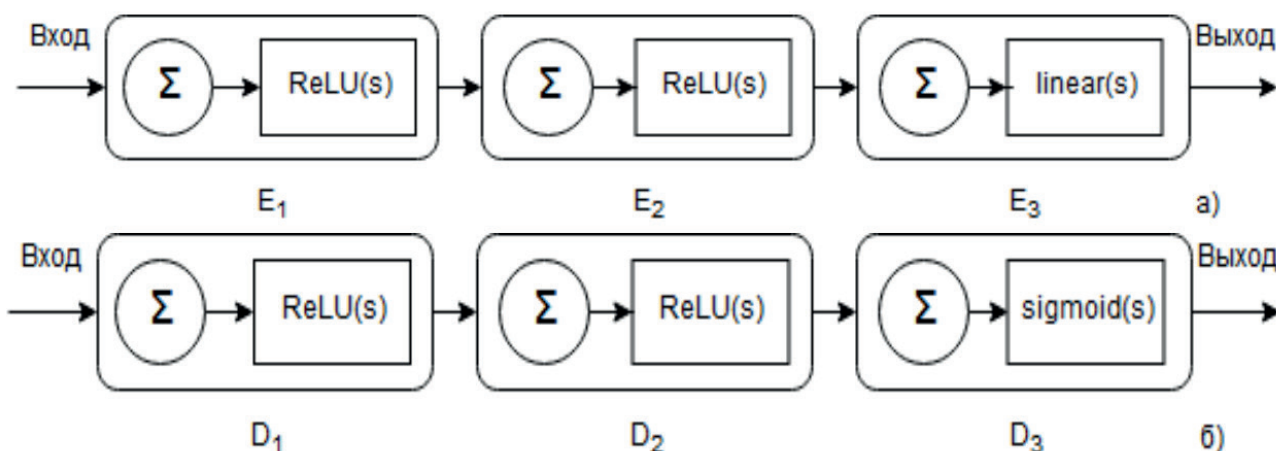


Рис. 2. Архитектура автоэнкодера: а) энкодер, б) декодер.
 Fig. 2. Architecture of autoencoder, (a) encoder (b) decoder.

для слоя E3 коэффициент $k=2.00$.

Данный параметр подбирался экспериментально.

Скрытые слои декодера (рис. 2) аналогичны слоям классификатора. В качестве выходной функции активации использовалась сигмоидальная функция активации, аналогичная выходной функции классификатора.

На рис. 2 слои пронумерованы D1, D2 и D3. Количество нейронов зависит от размерности выходных признаков, как и в энкодере. Коэффициент k был экспериментально подобран для каждого слоя:

для слоя D1 коэффициент $k = 1.75$;

для слоя D2 коэффициент $k = 1.35$;

для слоя D3 коэффициент $k = 1.00$.

Автоэнкодер был обучен методом с адаптивной скоростью обучения *adadelata* [24], показавшим лучшие значения качества обучения, чем метод Adam.

При обучении автоэнкодера, для исключения проблем исчезновения и взрывного роста градиента, для инициализации весов использовался метод Ксавье. Для преодоления эффекта переобучения использовалось обучение с ранним остановом (*earlystopping*) и перекрёстной проверкой. Набор данных для обучения был разделён в пропорциях: обучающая выборка — 90%, проверочная выборка — 10%.

4. АНАЛИЗ РЕЗУЛЬТАТОВ

4.1. Метрики качества нейросетевого классификатора

Задача диагностики послеоперационных осложнений — это задача бинарной классификации, поскольку учитывается наличие или отсутствие осложнения. Чтобы оценить и выбрать лучший классификатор нейронной сети, рассмотрим метрики качества задач бинарной классификации [16–17]. В таких задачах метки классов принадлежат множеству $\{0, 1\}$. Объекты с меткой "1" называются положительными, а с меткой "0" — отрицательными. Нейронная сеть возвращает число с плавающей запятой. В случае сигмоидной функции активации в выходном слое сети — это число от 0 до 1. С помощью точки отсечки — порогового значения, разделяющего классы, ответ классификатора превращается из плавающего в двоичный. Если сигнал нейрона превышает точку отсечки, это положительный результат, в противном случае - отрицательный.

В бинарной классификации каждое индивидуальное предсказание может иметь четыре результата: истинно-положительный (*truepositive* или TP); истинно-отрицательный (*truenegative* или TN); ложно-положительный (*falsepositive* или FP); ложно-отрицательный (*falsenegative* или FN). Истинно положительным исход будет, когда фактический класс данного примера 1 (положительный) и модель на выходе выдаст 1 (положительный). Истинно отрицательным исход будет, когда наблюдаемый класс 0 (отрицательный) и модель выдаст 0 (отрицательный). Ложно-положительное значение имеет место, когда фактический класс 0 (отрицательный), а модель для него сформирует выход 1 (положительный). При ложно-отрицательном выходе целевая переменная принимает значение 1 (положительный), а на выходе модель выдаст 0 (отрицательный). Результаты этих вычислений представляются в виде матрицы ошибок (*confusionmatrix*).

В качестве метрик, оценивающих качество нейросетевой модели, были выбраны метрики, реализованные в Scikit-learn [16–17]:

оценка прецизионности или положительная прогностическая ценность (*positivepredictivevalue*);

оценка чувствительности или доля истинно-положительных срабатываний (*truepositiverate*);

F1-мера;

оценка специфичности (*specificity*) или для истинно-отрицательных срабатываний (*truenegativerate*);

ROC-кривая (*ROC-curve*);

площадь под ROC-кривой (*ROC AUC*);

количество ошибок I рода (количество ложно-положительный – значение метрики FP);

количество ошибок II рода (количество ложно-отрицательный – значение метрики FN);

матрица ошибок.

Оценка прецизионности (*precision*) показывает точность предсказанного положительного исхода (когда у пациента могут быть осложнения)

$$R_{PR} = \frac{N_{TP}}{N_{TP} + N_{FP}}$$

где N_{TN} — число истинно-положительных наблюдений, N_{TN} — число истинно-отрицательных наблюдений, N_{FN} — число ложно-положительных

Таблица 2. Результаты экспериментов для комбинации признаков

Table 2. Experimental results for the first combination of features

Метрика / Metric	Значения метрик / Metric values	
	Среднее / Average	Медиана / Median
Прецизионность / Precision	0.8167	0.8750
Чувствительность / Sensitivity	0.6167	0.7500
F1-мера / F1-score	0.6843	0.7500
Специфичность / Specificity	0.8289	0.8571
ROC AUC	0.7804	0.8036
Вероятность ошибки I рода / Probability of a I type error	0.1833	0.1250
Вероятность ошибки II рода / Probability of a II type error	0.1711	0.1429

наблюдений, N_{FN} — число ложно-отрицательных наблюдений, N_{TP} — общее число наблюдений.

Чувствительность (sensitivity) модели показывает долю положительных исходов, которую алгоритм смог правильно обнаружить

$$S_R = \frac{N_{TP}}{N_{TP} + N_{FN}}$$

Прецизионность и чувствительность модели бывает целесообразно объединить в одну метрику, которая называется F1-мера. Данная метрика является средним гармоническим прецизионности и чувствительности

$$F_1 = 2 \frac{R_{PR} * S_R}{R_{PR} + S_R}$$

Метрика F1-мера будет близка к нулю, если хотя бы один из её аргументов близок к нулю. Причём прецизионность и полнота оказывают одинаковое влияние. Особенность данной метрики в том, что среднее гармоническое придаёт малым значениям больший вес. Поэтому, высокое значение метрики F1-мера можно обеспечить только высокими значениями этих двух метрик.

Специфичность (specificity) модели измеряет способность алгоритма предсказывать отрицательные исходы (долю пациентов без осложнений)

$$F_{PR} = \frac{N_{FP}}{N_{FP} + N_{TN}} = 1 - \frac{N_{TN}}{N_{TN} + N_{FP}}$$

Для оценки качества бинарной классификации применяется ROC-анализ, основанный на построении характеристической кривой обнаружения, или кривой ошибок — ROC-кривой. ROC-кривая показывает соотношение между долей верно классифицированных примеров (True Positive Rate — TPR) т. е. чувствительностью, и долей неверно классифицированных примеров (FalsePositiveRate)

$$F_{PR} = \frac{N_{FP}}{N_{FP} + N_{TN}} = 1 - \frac{N_{TN}}{N_{TN} + N_{FP}}$$

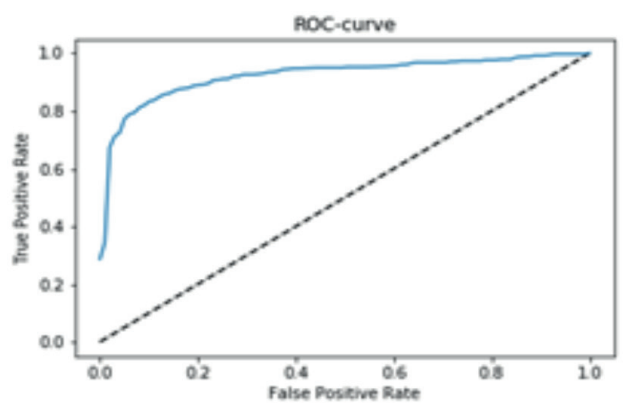


Рис. 3. Усредненная ROC-кривая.
Fig. 3. Averaged ROC curve.

Площадь под ROC-кривой (AUC — AreaUnderROCCurve) говорит о прогностической силе модели, причем AUC=1 соответствует идеальному классификатору. Для реальных классификаторов площадь под ROC-кривой больше 0,7–0,8, соответствует достаточно высокой точности классификации.

Для обучающих данных использовали кросс-валидацию по k блокам (k-foldCV)[25]. В этом методе исходное множество примеров разбивается на k непересекающихся частей. Производится k итераций. На каждой итерации производится модель, которая обучается на $k-1$ частях выборки, тестируется на части обучающей выборки, которая не участвовала в обучении. Каждая из k частей единожды используется для тестирования. Итоговая оценка точности получается усреднением оценок на каждой итерации. В реализации использован класс StratifiedKFold из библиотеки Scikit-learn. Для обучения автоэнкодера и классификатора набор данных был разделен на 10 подмножеств.

4.2. Анализ результатов экспериментов

Эксперименты проводились на компьютере с процессором IntelCore i5 2,4 ГГц с 8 ГБ оперативной памяти DDR3 1600 МГц под управлением операционной системы macOS Mojave версии 10.14.6. Были использованы следующие библиотеки: Scikit-Learn версии 0.22.2.post1, TensorFlow версии 2.1.0, NumPy версии 1.17.0, Keras версии 2.3.1. Время обучения повышающего автоэнкодера составляло 8 секунд. Классификатор обучался за 11 секунд. Представленное время обучения было рассчитано для одной итерации перекрестной проверки.

Описанные показатели качества были рассчитаны для окончательной комбинации значимых атрибутов. В табл. 2 показаны значения показателей качества модели на основе выбранной комбинации функций. Показатели получения с использованием кросс-валидации для k блоков.

На рис. 3 представлена ROC-кривая усредненных значений.

Построенная нейросетевая модель позволяет хирургу объективно оценить риск развития осложнений в хирургии на основе клинического мышления и нейросетевого метода.

Заключение

В статье предложен подход к выбору признаков, который использует межквантильный размах. Такой подход позволил получить высокий и стабильный результат метрики F1-мера. Предлагаемый подход может применяться для решения различных задач классификации табличных данных. Предложен подход к кодированию составных категориальных признаков, отличающийся компактным представлением. Для повышения точности нейросетевой диагностики использован повышающий автоэнкодер, проектирующий признаки в пространство большей размерности.

Применение автоэнкодера позволит повысить качество модели. Разработана и исследована полносвязная нейронная сеть глубокой архитектуры, позволяющая диагностировать послеоперационные осложнения в хирургии. Для снижения эффекта переобучения использован современный метод регуляризации — dropout. Для обучения нейросети применен алгоритм Adam с адаптивной скоростью обучения и кросс-

валидацией по к блокам. Исследование нейросети показало высокие показатели качества диагностики.

Дополнительная информация

Конфликт интересов

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Список литературы

1. Совцов С. А. Острый аппендицит: что изменилось в начале нового века?(с комментарием). *Хирургия. Журнал им. Н. П. Пирогова*. 2013;7:37–42.
2. Simillis C, Symeonides P, Shorthouse AJ, Tekkis PP. A meta-analysis comparing conservative treatment versus acute appendectomy for complicated appendicitis (abscess or phlegmon). *Surgery*. 2010;147:818–829.doi.org/10.1016/s0090-3671(10)79895-5
3. Prabhudesai SG. Artificial neural Networks: Useful Aid in Diagnosing Acute Appendicitis. *World Journal of Surgery*. 2008; 2:305–309.doi.org/10.1007/s00268-007-9299-5
4. Park SY, Kim SM. Acute appendicitis diagnosis using artificial neural networks. *Technology and health care*. 2015;23:559–565.doi.org/10.3233/thc-150994
5. Juliano Y, Rosa OM, Novo NF, Favaro ML. Risk factors associated with complications of acute appendicitis. *Journal of Brazilian College of Surgeons*. 2017;6:560–566.doi.org/10.26226/morressier.58fa1765d462b80290b50f65
6. Bakti N, Hussain A, El-Hassani S. A rare complication of acute appendicitis: Superior mesenteric vein thrombosis. *International Journal of Surgery Case Reports*. 2011; 8: 250–252.doi.org/10.1016/j.ijscr.2011.08.003
7. Brigham KL, Johns MME. Predictive Health: How We Can Reinvent Medicine to Extend Our Best Years. *Basic Books*. 2012; 256.
8. Miner L, Bolding P, Hilbe J. Practical Predictive Analytics and Decisioning Systems for Medicine: Informatics Accuracy and Cost-Effectiveness for Healthcare Administration and Delivery Including Medical Research. *Academic Press*. 2014;1110.
9. Joskowicz L. Computer-aided surgery meets predictive, preventive, and personalized medicine. *EPMA Journal*. 2017;8:1–4. doi.org/10.1007/s13167-017-0084-8
10. Хайкин С. *Нейронные сети: полный курс*. М.: Вильямс. 2006;1104.
11. Гудфеллоу Я., Бенджио И., Курвилль А. *Глубокое обучение*. М.: ДМК Пресс, 2018. 652с.
12. Аггарвал Ч. *Нейронные сети и глубокое обучение*. СПб.: ООО "Диалектика". 2020;752.
13. П. А. Ваченко, А. А. Соломаха, В. И. Горбаченко, А. О. Хазратов. Клинико-лабораторные параметры больных острым аппендицитом. Свидетельство о государственной регистрации базы данных № 2014621431. Дата государственной регистрации в Реестре баз данных 10 октября 2014 г.
14. Zheng A, Casari A. Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists. *O'Reilly Media*. 2018; 218.
15. Chandrashekar G, Sahin F. A survey on feature selection methods. *Computers & Electrical Engineering*. 2014;40: 1: 16–28.doi.org/10.1016/j.compeleceng.2013.11.024
16. Брюс П., Брюс Э. *Практическая статистика для специалистов DataScience*. СПб.: БХВ Петербург. 2018;304.
17. Жерон О. *Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем*. СПб.: ООО "Диалектика". 2020;1040.
18. Brownlee J. Why One-Hot Encode Data in Machine Learning? <https://machinelearningmastery.com/why-one-hot-encode-data-in-machine-learning/>
19. Хейдт М., Груздев А. В. *Изучаем pandas*. М.: ДМК Пресс. 2019;682.
20. Scikit-Learn // <https://scikit-learn.org/stable/>
21. Шолле Ф. *Глубокое обучение на Python*. СПб.: Питер. 2018; 400.
22. Keras: *The Python Deep Learning library*. Usage of callbacks. <https://keras.io/callbacks/>
23. Cover TM. Geometrical and Statistical Properties of Systems of Linear inequalities with Applications in Pattern Recognition. *IEEE Transactions on Electronic Computers*. 1965; EC14: 3: 326–334.doi.org/10.1109/pgec.1965.264137
24. Zeiler MD. ADADELTA: An Adaptive Learning Rate Method. 2012; <https://arxiv.org/abs/1212.5701>
25. Рашка С., Мирджалили В. *Python и машинное обучение: машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow 2*. СПб.: ООО "Диалектика". 2020;848.

References

1. Sovtsov SA. Acute appendicitis: what has changed at the beginning of the new century?(comment). *Khirurgiya. Zhurnal im. N. P. Pirogova*. 2013;7:37-42. (in Russ.)
2. Simillis C, Symeonides P, Shorthouse AJ, Tekkis PP. A meta-analysis comparing conservative treatment versus acute appendectomy for complicated appendicitis (abscess or phlegmon). *Surgery*. 2010;147:818–829.doi.org/10.1016/s0090-3671(10)79895-5
3. Prabhudesai SG. Artificial neural Networks: Useful Aid in Diagnosing Acute Appendicitis. *World Journal of Surgery*. 2008; 2:305–309.doi.org/10.1007/s00268-007-9299-5
4. Park SY, Kim SM. Acute appendicitis diagnosis using artificial neural networks. *Technology and health care*. 2015;23:559–565.doi.org/10.3233/thc-150994
5. Juliano Y, Rosa OM, Novo NF, Favaro ML. Risk factors associated with complications of acute appendicitis. *Journal of Brazilian College of Surgeons*. 2017;6:560–566.doi.org/10.26226/morressier.58fa1765d462b80290b50f65
6. Bakti N, Hussain A, El-Hassani S. A rare complication of acute appendicitis: Superior mesenteric vein thrombosis. *International Journal of Surgery Case Reports*. 2011; 8: 250–252.doi.org/10.1016/j.ijscr.2011.08.003
7. Brigham KL, Johns MME. Predictive Health: How We Can Reinvent Medicine to Extend Our Best Years. *Basic Books*. 2012; 256.
8. Miner L, Bolding P, Hilbe J. Practical Predictive Analytics and Decisioning Systems for Medicine: Informatics Accuracy and Cost-Effectiveness for Healthcare Administration and Delivery Including Medical Research. *Academic Press*. 2014;1110.
9. Joskowicz L. Computer-aided surgery meets predictive, preventive, and personalized medicine. *EPMA Journal*. 2017;8:1–4. doi.org/10.1007/s13167-017-0084-8
10. Khaikin S. *Neironnye seti: polnyi kurs*. M.: Vil'yams. 2006;1104. (in Russ.)
11. Gudfellow Ya, Bendzhio I, Kurvil' A. *Glubokoe obuchenie*. M.: DMK Press, 2018. 652s. (in Russ.)
12. Aggarval Ch. *Neironnye seti i glubokoe obuchenie*. SPb.: OOO "Dialektika". 2020;752. (in Russ.)
13. Vashchenko PA, Solomakha AA, Gorbachenko VI, Khazratov AO. Kliniko-laboratornye parametry bol'nykh ostrym appenditsitom. Svidetel'stvo o gosudarstvennoy registratsii bazy dannykh № 2014621431. Data gosudarstvennoy registratsii v Reestre baz dannykh 10 oktyabrya 2014. (in Russ.)
14. Principles and Techniques for Data Scientists. *O'Reilly Media*. 2018; 218.
15. Chandrashekar G, Sahin F. A survey on feature selection methods. *Computers & Electrical Engineering*. 2014;40: 1: 16–28.doi.org/10.1016/j.compeleceng.2013.11.024
16. Bruce P, Bruce E. *Prakticheskaya statistika dlya spetsialistov DataScience*. SPb.: BKhV Peterburg. 2018; 304. (in Russ.)
17. Geron O. *Prikladnoe mashinnoe obuchenie s pomoshch'yu Scikit-Learn, Keras i TensorFlow: kontseptsii, instrumenty i tekhniki dlya sozdaniya intellektual'nykh sistem*. SPb.: OOO "Dialektika". 2020;1040. (in Russ.)
18. Brownlee J. Why One-Hot Encode Data in Machine Learning? <https://machinelearningmastery.com/why-one-hot-encode-data-in-machine-learning/>
19. Heidt M, Gruzdev AV. *Izuchaem pandas*. M.: DMK Press. 2019;682. (in Russ.)
20. Scikit-Learn // <https://scikit-learn.org/stable/>
21. Scholle F. *Glubokoe obuchenie na Python*. SPb.: Piter. 2018; 400. (in Russ.)
22. Keras: *The Python Deep Learning library*. Usage of callbacks. <https://keras.io/callbacks/>
23. Cover TM. Geometrical and Statistical Properties of Systems of Linear inequalities with Applications in Pattern Recognition. *IEEE Transactions on Electronic Computers*. 1965; EC14: 3: 326–334.doi.org/10.1109/pgec.1965.264137
24. Zeiler MD. ADADELTA: An Adaptive Learning Rate Method. 2012; <https://arxiv.org/abs/1212.5701>
25. Rashka S, Mirjalili V. *Python i mashinnoe obuchenie: mashinnoe i glubokoe obuchenie s ispol'zovaniem Python, scikit-learn i TensorFlow 2*. SPb.: OOO "Dialektika". 2020;848. (in Russ.)

Информация об авторах

1. Горбаченко Владимир Иванович - д.т.н., профессор, заведующий кафедрой «Компьютерные технологии», e-mail: anatoly.solomakha@yandex.ru
2. Соломаха Анатолий Анатольевич - к.м.н., хирург, трансфузиолог, врач высшей категории, доцент, действительный член Российской академии медико-технических наук, e-mail: anatoly.solomakha@yandex.ru

Information about the Authors

1. Vladimir Ivanovich Gorbachenko - Doctor of Technical Sciences, Professor, Head of the Department of Computer Technology, e-mail: anatoly.solomakha@yandex.ru
2. Anatoly Anatolyevich Solomakha - Ph.D., surgeon, transfusiologist, doctor of the highest category, associate professor, full member of the Russian Academy of Medical and Technical Sciences, e-mail: anatoly.solomakha@yandex.ru

Цитировать:

Горбаченко В.И., Соломаха А.А. Способ диагностики послеоперационных осложнений в хирургии. Вестник экспериментальной и клинической хирургии 2022; 15: 4: 314-322. DOI: 10.18499/2070-478X-2022-15-4-314-322.

To cite this article:

Gorbachenko V.I., Solomakha A.A. Diagnostic Options for Postoperative Complications in Surgery. Journal of experimental and clinical surgery 2022; 15: 4: 314-322. DOI: 10.18499/2070-478X-2022-15-4-314-322.