

Федеральное государственное бюджетное учреждение
«Национальный медицинский исследовательский центр онкологии
имени Н.Н. Петрова» Министерства здравоохранения Российской Федерации
(ФГБУ «НМИЦ онкологии им. Н.Н. Петрова» Минздрава России)
Отдел учебно-методической работы

Федеральное государственное бюджетное образовательное учреждение
высшего образования «Северо-Западный государственный
медицинский университет имени И.И. Мечникова»
Министерства здравоохранения Российской Федерации
(ФГБОУ ВО СЗГМУ им. И.И. Мечникова Минздрава России)
Кафедра онкологии

**Беляев А.М., Левченко Е.В.,
Михнин А.Е., Рогачёв М.В.**

Метрики информативности тестов и бинарных алгоритмов

Учебное пособие

Санкт-Петербург
2025

УДК: 61:31:004(07)

ББК: 51.1(2)я7

Беляев А.М., Левченко Е.В., Михнин А.Е., Рогачёв М.В. Метрики информативности тестов и бинарных алгоритмов: учебное пособие для врачей и обучающихся в системе высшего и дополнительного профессионального образования. – Санкт-Петербург: НМИЦ онкологии им. Н.Н. Петрова, 2025. – 28 с.

ISBN 978-5-6051652-1-7

Рецензент: доктор медицинских наук, профессор В.М. Мерабишвили, председатель научно-медицинского Совета по развитию информационных систем онкологической службы СЗФО Российской Федерации, заведующий научной лабораторией онкологической статистики федерального государственного бюджетного учреждения «Национальный медицинский исследовательский центр онкологии имени Н.Н. Петрова» Министерства здравоохранения Российской Федерации, г. Санкт-Петербург.

В учебном пособии представлены принципы построения бинарных классификаторов, базовые метрики информативности, а также рекомендуемые наиболее популярные обобщающие метрики информативности, используемые в медицинских исследованиях.

Учебное пособие предназначено для аспирантов, клинических ординаторов и врачей, начинающих свою научную деятельность в медицине, а также для обучающихся по программам дополнительного профессионального образования.

Утверждено в качестве учебного пособия
Ученым советом ФГБУ «НМИЦ онкологии
им. Н.Н. Петрова» Минздрава России
протокол № 9 от 28 октября 2025 г.

© Беляев А.М. Коллектив авторов, 2025

ISBN 978-5-6051652-1-7

Содержание

Введение	4
Глава 1. Матрица ошибок бинарного классификатора	5
Глава 2. Базовые метрики информативности бинарных классификаторов	8
Глава 3. Дополнительные метрики информативности бинарных классификаторов	10
Глава 4. Обобщающие метрики информативности бинарных классификаторов	12
Глава 5. Выбор обобщающей метрики для сравнения медицинских классификаторов	18
Приложения	20
Приложение 1. Геометрические обоснования сходства метрик BA и AUC	20
Приложение 2. Стандартные ошибки и доверительные интервалы некоторых метрик	21
Контрольные вопросы	23
Тестовые задания	25
Список литературы	28

Введение

В современных клинических исследованиях широко используются диагностические и прогностические модели и алгоритмы.

Их разработку, обучение и практическое использование, как правило, осуществляют специалисты по информатике, имеющие дополнительную подготовку в области биологии и медицины.

Тем не менее, для правильной постановки задач и корректного использования таких моделей врачу необходимо понимание заложенных в основу математических методов, их возможностей и ограничений.

В исследованиях нередко возникает необходимость оценки эффективности бинарных классификаторов: моделей, алгоритмов и лабораторных тестов, разделяющих объекты (пациентов) на две категории, например, «норма – патология» или «прогноз благоприятный – прогноз неблагоприятный».

Математическое содержание такой задачи заключается в сравнении эффективности используемого классификатора с эталонным, которым может являться референсный тест или более информативный метод исследования.

Оценка состоит в анализе матрицы ошибок (confusion matrix) классификатора с вычислением метрик информативности.

При сравнении классификаторов обычно требуется указывать уровень статистической значимости различий либо доверительные интервалы, что для некоторых метрик достаточно сложно.

Глава 1.

Матрица ошибок бинарного классификатора

Бинарный классификатор – это физическая (например, лабораторный тест) или виртуальная (компьютерная) модель, позволяющая разделить объекты (лабораторные пробы, данные инструментальной диагностики и т.д.) на два класса: положительный и отрицательный.

В бинарных классификаторах могут быть реализованы как классические статистические методы, например, логистическая регрессия, дискриминантный анализ, так и не всегда прозрачные технологии, в частности, обучение нейронных сетей.

Независимо от этого результат работы бинарного классификатора может иметь только два значения: группирующий признак обнаружен (например, заболевание) и группирующий признак не выявлен.

Матрица ошибок (confusion matrix, матрица неточностей) – представляет собой таблицу несоответствия между предсказанными и фактическими значениями для рассматриваемого классификатора.

Она позволяет визуализировать эффективность алгоритма, выявляя не только количество правильных прогнозов, но и типы ошибок, которые допускает модель.

Для не бинарной (многоклассовой) классификации матрица ошибок расширяется до размера $N \times N$, где N – количество классов. В такой матрице диагональные элементы представляют правильные предсказания, а все остальные – ошибки.

Таблица сопряженности результатов используемого бинарного классификатора с референсным классификатором – матрица ошибок представлена в таблице 1.

Таблица 1

Матрица ошибок бинарного классификатора
[оригинальная таблица]

Результат классификации*	Фактический положительный класс, D(+)	Фактически отрицательный класс, D(-)
Предсказанный положительный класс	TP	FP
Предсказанный отрицательный класс	FN	TN

*True positive – TP; True negative – TN; False positive – FP;
False negative – FN; $D(+)$ = TP+FN; $D(-)$ = FP+TN

Каждая ячейка матрицы содержит важную конкретную информацию:

True Positive (TP): Модель правильно предсказала положительный класс. Например, тест определил наличие болезни у действительно больного.

True Negative (TN): Модель правильно предсказала отрицательный класс. Например, тест оказался отрицательным у действительно здорового.

False Positive (FP): Модель ошибочно предсказала положительный класс (ошибка I рода, или «ложная тревога»). Например, тест определил наличие болезни у действительно здорового.

False Negative (FN): Модель ошибочно предсказала отрицательный класс (ошибка II рода, или «пропуск»). Например, тест

определил отсутствие болезни у действительно больного.

Матрица ошибок даёт возможность анализировать поведение модели:

Если модель демонстрирует высокий уровень FP, она допускает много ошибок I рода и склонна к «паранойе» – часто видит положительный класс там, где его нет.

Если модель показывает высокий уровень FN, т.е. допускает много ошибок II рода, она склонна к «халатности» – пропускает важные случаи положительного класса.

Если модель имеет высокие значения как FP, так и FN, она либо плохо обучена, либо задача сформулирована некорректно.

Интерпретация матрицы ошибок напрямую связана с контекстом задачи.

В медицинской диагностике критически важно минимизировать FN (пропуск больных пациентов).

Для системы безопасности аэропорта необходимо снижать FN (пропуск опасных предметов), одновременно поддерживая разумный уровень FP, чтобы не создавать чрезмерных неудобств для пассажиров.

Выбор балансировки между FN и FP в значительной мере зависит от реального соотношения объектов положительного и отрицательного классов.

Например, при разработке скринингового теста необходимо обязательно учитывать распространённость заболевания в тестируемой популяции.

Глава 2.

Базовые метрики информативности бинарных классификаторов

Показатели, характеризующие матрицу ошибок классификатора, в информатике называются метриками информативности. Несмотря на то, что матрица ошибок бинарного классификатора имеет формат 2×2 и состоит всего из четырех элементов, известно более 70 метрик, являющихся комбинациями этих элементов.

Базовыми являются метрики: Recall (Sensitivity) и Specificity.

Recall (Полнота) = Sensitivity, Se (Чувствительность) – доля положительных объектов, правильно идентифицированных моделью:

$$Recall = Se = \frac{TP}{TP+FN} \quad (1)$$

Метрика Recall показывает, насколько хорошо модель определяет все положительные случаи. Эта метрика приобретает решающее значение, когда цена пропуска положительного случая высока.

Для лабораторных тестов метрика **Recall** чаще называется **Sensitivity** (чувствительность) и означает способность теста выявлять больных среди действительно больных:

$$Se = \frac{TP}{D(+)} \quad (2)$$

Specificity, Sp (Специфичность) – доля объектов, идентифицированных как отрицательные среди действительно отрицательных:

$$Sp = \frac{TN}{TN+FP} \quad (3)$$

Специфичность лабораторного теста – это его способность идентифицировать здоровых среди действительно здоровых.

Precision (Точность) – доля объектов, идентифицированных как положительные, которые действительно являются положительными:

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

Метрика Precision показывает уровень доверия положительным предсказаниям модели. Эта метрика особенно важна, когда положительный класс редок и ценен, а цена ложной тревоги высока.

В бизнес-моделях компромисс между Precision и Recall известен как Precision-Recall trade off. Управление этим компромиссом для вероятностных моделей осуществляется варьированием порога отсечения (cut-off threshold). Оптимизация медицинских вероятностных моделей в поисках компромисса между Sensitivity и Specificity проводится аналогичным образом.

Глава 3.
Дополнительные метрики
информативности бинарных классификаторов

Positive predictive value, PPV (Предсказательная ценность положительного результата) – вероятность того, что заболевание присутствует при положительном тесте:

$$PPV = \frac{TP}{TP+FP} \quad (5)$$

Negative predictive value, NPV (Предсказательная ценность отрицательного результата) – вероятность того, что заболевание отсутствует при отрицательном тесте:

$$NPV = \frac{TN}{TN+FN} \quad (6)$$

Positive likelihood ratio, PLR (Отношение правдоподобия положительного результата теста) – отношение вероятности положительного результата теста при наличии заболевания к вероятности положительного результата теста при отсутствии заболевания:

$$PLR = \frac{Se}{1-Sp} \quad (7)$$

Negative likelihood ratio, NLR (Отношение правдоподобия отрицательного результата теста) – отношение вероятности

отрицательного результата теста при наличии заболевания к вероятности отрицательного результата теста при отсутствии заболевания:

$$NLR = \frac{1-Se}{Sp} \quad (8)$$

Apparent prevalence, $Aprev$ (Кажущаяся распространённость болезни среди пациентов, участвующих в исследовании) – вероятность того, что результат теста будет положительным относительно всех проведённых тестов:

$$Aprev = \frac{TP+FP}{TP+TN+FP+FN} \quad (9)$$

True prevalence, $Tprev$ (Истинная распространённость болезни среди пациентов, участвующих в исследовании) – вероятность того, что результат теста будет положительным относительно всех проведённых тестов:

$$Tprev = \frac{TP+FN}{TP+TN+FP+FN} \quad (10)$$

Метрики (1-4) являются базовыми взаимозависимыми характеристиками бинарного классификатора, например, чувствительность и специфичность и позволяют в паре оценить качество классификатора.

Тем не менее, разработан целый ряд метрик, позволяющих представить оценку эффективности классификатора в виде единственного параметра (11-15).

Глава 4.

Обобщающие метрики информативности бинарных классификаторов

Ассурасу, Асс (Правильность) – доля правильных предсказаний среди всех предсказаний модели:

$$Асс = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

Для лабораторного теста – это доля пациентов с правильно определенным статусом среди всех пациентов, прошедших тест.

Эта метрика находится в сильной зависимости от соотношения положительного и отрицательного классов (т.е. пропорции действительно положительных и действительно отрицательных объектов). Метрика Ассурасу может применяться лишь в хорошо сбалансированных выборках, в которых количество объектов в классах различается менее, чем на один порядок.

Balanced Accuracy, BA (Сбалансированная правильность) – среднее чувствительности и специфичности: специальный аналог правильности для несбалансированных данных:

$$BA = \frac{1}{2} \left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP} \right) = \frac{1}{2} (Se + Sp) \quad (12)$$

Метрика Balanced Accuracy гораздо менее чувствительна к дисбалансу и может быть рекомендована при дисбалансе классов, оставляющем несколько порядков. Заметим, что эта мет-

рика соответствует метрике AUC (15), вычисленной для единственной точки отсечения. Более подробно взаимосвязь метрик ВА и AUC рассматривается в приложении.

F1-score – среднее гармоническое метрик Precision и Recall, позволяющее сбалансировать их значения:

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} = \frac{2TP}{2TP+FP+FN} \quad (13)$$

Geometric mean, GM (Геометрическое среднее точности и чувствительности):

$$GM = \sqrt{Precision * Recall} \quad (14)$$

Метрики F1 (13) и GM (14) часто используются в машинном обучении при дисбалансе классов, составляющем два-три порядка, однако их информативность снижается при большем дисбалансе.

Area under ROC-curve, AUC (Площадь под ROC-кривой) – совокупный показатель производительности теста по всем возможным диагностическим порогам (cut off threshold), который не имеет аналитического выражения.

AUC может быть измерена в прикладных статистических пакетах, которые одновременно позволяют определить уровень значимости различий при сравнении двух тестов.

Если порог отсечения теста является фиксированным, AUC

может быть рассчитана для этой единственной точки ROC-кривой по формуле:

$$AUC = \frac{1}{2}(Se + Sp)* \quad (15)$$

*Простое геометрическое обоснование формулы (14) см. в приложении.

Расчет AUC по одной точке не вполне корректен, однако часто встречается в медицинских публикациях.

Метрика (15) в точности соответствует метрике (12) – сбалансированная правильность (*BA*). Именно эта метрика может быть рекомендована для оценки сравнения большинства бинарных классификаторов, применяемых в медицинских исследованиях.

Кроме метрик, основанных на вероятности (доля, частота), существуют показатели информативности, основанные на шансах.

Шансы (Odds) – это отношение вероятности того, что некое событие произойдет (p) к вероятности того, что оно не произойдет:

$$Odds = \frac{p}{1 - p}$$

Например, если вероятность события равна 5% ($p=0,05$), то шансы этого события составляют 1/19. При вероятности 10% ($p=0,1$) шансы равны 1/9.

DOR (Diagnostic Odds Ratio) – диагностическое отношение шансов: отношение шансов получить положительный результат теста, если пациент болен (TP/FN), к шансам получить положительный результат теста, если он здоров (FP/TN).

$$DOR = \frac{TP/FN}{FP/TN} = \frac{TP*TN}{FN*FP} \quad (16)$$

Значения этой метрики максимальны при FN или FP, стремящихся к нулю, т.е. при отсутствии ошибок классификации объектов одного из двух классов. Например, если пациента всегда относить к классу здоровых, метрика DOR будет иметь максимальное значение.

Диагностическое отношение шансов, выраженное через показатели Sensitivity (Se) и Specificity (Sp), выглядит следующим образом:

$$DOR = \frac{Se*Sp}{(1-Se)(1-Sp)} \quad (17)$$

Диагностическое отношение шансов также может быть выражено через метрики правдоподобия:

$$DOR = \frac{PLR}{NLR} \quad (18)$$

DOR имеет важное преимущество: возможность построения доверительного интервала на основе матрицы ошибок. Стандартную ошибку (SE) логарифма DOR вычисляют по формуле:

$$SE(\ln DOR) = \sqrt{\frac{1}{TP} + \frac{1}{TN} + \frac{1}{FP} + \frac{1}{FN}} \quad (19)$$

Границы доверительного интервала (95%CI) получают путем антилогарифмирования уравнения:

$$95\%CI = \ln DOR \pm 1,96 * SE(\ln DOR) \quad (20)$$

Известны также некоторые другие, не всегда интуитивно понятные, но популярные обобщенные метрики информативности:

Youden's Index, J (Индекс Юдена) – Разница вероятностей истинно положительных и ложноположительных результатов теста:

$$J = Se + Sp - 1 \quad (21)$$

Индекс Юдена является линейной комбинацией чувствительности и специфичности и связан с метрикой AUC:

$$J = 2AUC - 1 \quad (22)$$

Predictive Summary Index, PSI (Сводный прогнозный индекс) - Общий выигрыш в уверенности в наличии заболевания от выполнения диагностического теста:

$$PSI = PPV + NPV - 1 \quad (23)$$

Cohen's kappa, k (Каппа Коэна) – показатель уровень согласия между двумя оценщиками, классифицирующими элементы по взаимоисключающим категориям. В отличие от простого процентного согласия, каппа Коэна учитывает возможность случайного совпадения.

$$k = \frac{(TP+FP)(TP+FN)+(TN+FN)(TN+FP)}{n^2} \quad (24)$$

Matthews correlation coefficient, MCC (Коэффициент корреляции Мэттьюса) – дискретный случай коэффициента корреляции Пирсона, является произведением средних геометрических точности и полноты:

$$MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (25)$$

Метрики (24-25) используются в бизнес-моделях, однако редко применяются в медицинских классификаторах в силу малой интуитивной понятности и сложности вычисления уровней статистической значимости различий.

Глава 5.

Выбор обобщающей метрики для сравнения медицинских классификаторов

Для бизнес-моделей разработаны общие рекомендации по выбору обобщающих метрик информативности классификатора и оптимальной настройке порога отсечения. С определенными ограничениями эти рекомендации могут быть использованы и в медицинских исследованиях, в особенности тогда, когда решение основывается на экономических последствиях каждого типа ошибок, например, при оценке «цена-качество» скрининговых программ.

Тем не менее, в медицине стандартные решения, основанные на учете степени дисбаланса положительного и отрицательного классов, являются вполне применимыми.

Accuracy остаётся золотым стандартом при анализе сбалансированных данных, когда положительный и отрицательный классы представлены примерно одинаково, и ошибки I и II рода имеют примерно равноценные последствия.

Balanced accuracy следует использовать при дисбалансе данных в пределах нескольких порядков, что на практике встречается гораздо чаще. Следует понимать, что эта метрика фактически сближается с AUC.

AUC является наиболее универсальной и наглядной метрикой, реализованной во многих статистических пакетах, в которых могут быть вычислены как доверительные интервалы, так и статистическая значимость различий качества классификаторов.

По этой причине именно эта метрика может быть рекомендована для сравнения медицинских бинарных классификаторов в условиях дисбаланса классов.

Важно понимать, что ни одна метрика не является идеальной. Профессиональный подход нередко требует использования комбинации метрик с учетом их соотношения с предъявляемыми требованиями. Правильная метрика – это та, которая лучше всего отражает реальную полезность модели в конкретном применении.

Приложения

Приложение 1

Геометрические обоснования сходства метрик BA и AUC

Метрики **Balanced accuracy** и **AUC** для бинарного классификатора достаточно близки и совпадают в случае построения AUC по одной фиксированной точке отсечения.

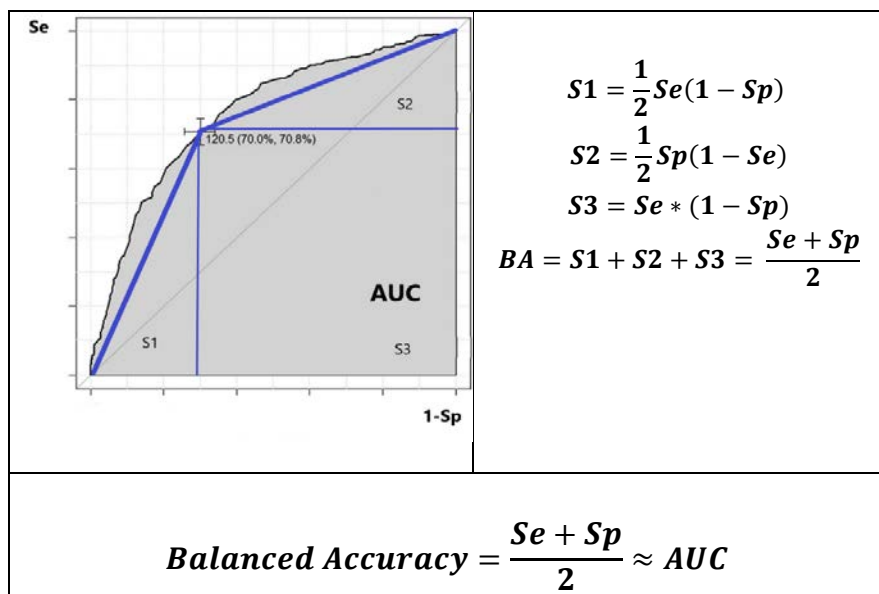


Рис. 1. ROC-кривая бинарного классификатора и метрика сбалансированной точности (*BA*) при фиксированной точке отсечения с координатами (*Se*; *1-Sp*) [оригинальный рисунок].

Как следует из рис. 1, *AUC* (Area under ROC-curve) представляет собой закрашенную серым площадь под

многоступенчатой кривой, построенной по всем возможным порогам отсечения.

В случае единственной точки отсечения эта площадь состоит из суммы площадей прямоугольных треугольников S_1 , S_2 и прямоугольника S_3 , отмеченных синими границами, которая легко вычисляется по формулам в правой части рис. 1.

Суммарная площадь в точности соответствует метрике **Balanced accuracy** и незначительно меньше **AUC**.

**Стандартные ошибки
и доверительные интервалы некоторых метрик**

Большинство метрик представляют собой относительные величины, выраженные в долях единицы или процентах.

Ошибка относительной величины (m) определяется по известной формуле:

$$m = \pm \sqrt{\frac{p(1-p)}{n}}$$

где p – доля, n – число наблюдений (при числе наблюдений менее 30 используется $n-1$).

95% доверительный интервал рассчитывается по формуле:

$$95\%CI = p \pm 1,96m$$

Контрольные вопросы

1. Бинарный классификатор, определение, общая характеристика.
2. Значения бинарного классификатора.
3. Матрица ошибок, определение, общая характеристика.
4. Сопряженность результатов используемого бинарного классификатора с референсным классификатором.
5. Модель True Positive (TP).
6. Модель True Negative (TN).
7. Модель False Positive (FP).
8. Модель False Negative (FN).
9. Возможности матрицы ошибок.
10. Интерпретация матрицы ошибок в медицинской диагностике.
11. Показатели, характеризующие матрицу ошибок классификатора.
12. Базовые метрики информативности.
13. Метрика Recall (Полнота).
14. Метрика Specificity (Специфичность).
15. Метрика Precision (Точность).
16. Предсказательная ценность положительного результата.
17. Предсказательная ценность отрицательного результата.
18. Отношение правдоподобия положительного результата теста.
19. Отношение правдоподобия отрицательного результата теста.

20. Кажущаяся распространённость болезни среди пациентов.
21. Истинная распространённость болезни среди пациентов.
22. Метрика Accuracy, Acc (Правильность).
23. Метрика Balanced Accuracy, BA (Сбалансированная правильность).
24. Метрика Geometric mean, GM (Геометрическое среднее точности и чувствительности).
25. Метрика Area under ROC-curve, AUC (Площадь под ROC-кривой).
26. Метрика DOR (Diagnostic Odds Ratio) – диагностическое отношение шансов.
27. Метрика Youden's Index, J (Индекс Юдена) – Разница вероятностей истинно положительных и ложноположительных результатов теста.
28. Метрика Predictive Summary Index, PSI (Сводный прогнозный индекс) – Общий выигрыш в уверенности в наличии заболевания от выполнения диагностического теста.
29. Метрика Cohen's kappa, k (Каппа Коэна) – показатель уровень согласия между двумя оценщиками.
30. Метрика Matthews correlation coefficient, MCC (Коэффициент корреляции Мэттьюса) – дискретный случай коэффициента корреляции Пирсона.

Тестовые задания

Инструкция: выберите один или несколько правильных ответов.

1. Объекты положительного класса в бинарных классификаторах - это:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	объекты, отнесенные классификатором к положительному классу	
б	объекты, отнесенные классификатором к положительному классу плюс объекты, ошибочно отнесенные классификатором к отрицательному классу	+
в	объекты, отнесенные референтным классификатором к положительному классу	+
г	объекты, ошибочно отнесенные классификатором к отрицательному классу	

2. Сбалансированность классов бинарной классификации - это:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	равенство числа объектов положительного класса числу объектов отрицательного класса	
б	разница числа объектов положительного класса и числа объектов отрицательного класса менее 10	
в	разница числа объектов отрицательного класса и числа объектов положительного класса менее 100	
г	отношение числа объектов положительного класса к числу объектов отрицательного класса в пределах одного порядка	+

3. Метрики информативности классификатора – это:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	показатели, характеризующие матрицу ошибок классификатора	+
б	показатели, характеризующие матрицу ошибок в положительном классе	
в	показатели, позволяющие представить оценку эффективности классификатора в виде единственного параметра	

4. Матрица ошибок классификатора – это:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	таблица расхождения результатов изучаемого классификатора по отношению к референсному	
б	четырёхпольная таблица сопряжённости результатов классификации	+
в	таблица правильных и неправильных результатов классификации объектов положительного и отрицательного классов	+
г	таблица неправильных результатов классификации	

5. Точка отсечения бинарного классификатора – это:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	значения метрик Recall и Selectivity, при которых классификатор обеспечивает требуемую эффективность	+
б	значения метрик Sensitivity и Selectivity, при которых классификатор обеспечивает наибольшую эффективность	
в	значения метрик Accuracy и Sensitivity ROC-кривой, на которые настроен порог отсечения классификатора	
г	значения метрик Accuracy и Selectivity, при которых классификатор обеспечивает наибольшую эффективность	

6. К базовым метрикам информативности относятся:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	DOR	
б	Recall	+
в	Specificity	+
г	Accuracy	

7. К дополнительным метрикам информативности относятся:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	NPV	+
б	PLR	+
в	NLR	+
г	Selectivity	

8. К обобщающим метрикам информативности относятся:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	DOR	+
б	PLR	+
в	NLR	
г	AUC	

9. Метрики информативности, наименее чувствительные к дисбалансу классов:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	Accuracy	
б	Sensitivity	
в	DOR	+
г	AUC	+

10. Метрики информативности, высоко чувствительные к дисбалансу:

Поле для выбора ответа	Варианты ответов	Поле для отметки правильного ответа (+)
а	Recall	+
б	Accuracy	+
в	Balanced accuracy	
г	F1	

Список литературы

1. Корнеенков А.А. Меры информативности диагностических медицинских технологий в оториноларингологии: вычисление и интерпретация / А.А. Корнеенков, С.В. Рязанцев, Е.Э. Вяземская, М.А. Будковая // Российская оториноларингология. – 2020. – Т. 19, № 1. – С. 46-55. <https://doi.org/10.18692/1810-4800-2020-1-46-55>.

2. Старовойтов В.В. Сравнительный анализ оценок качества бинарной классификации / В.В. Старовойтов, Ю.И. Голуб // Информатика. – 2020. – Т. 17, № 1. – С. 87-101. <https://doi.org/10.37661/18160301-2020-17-1-87-101>.

ISBN 978-5-6051652-1-7



Отпечатано в ООО «АРТЕК»,
СПб, 6-я линия В.О., д. 3/10
E-mail: artek-1@mail.ru, т. +7(911) 239-25-32
Подписано в печать 26.11.25
Формат 60x90/16. Тираж 50 экз.