Задачи на совокупность значений показателей качества классификации

А. И. Майсурадзе

# Основные определения и обозначения

Рассмотрим матрицу ошибок (confusion matrix, часто употребляют более общий термин contingency table) для задачи классификации с двумя классами (binary classification task), где один из классов назван положительным (positive, P), а другой – отрицательным (negative, N). Матрица ошибок получается как результат применения классификатора к размеченной выборке. В её ячейках стоят целые неотрицательные числа – частоты ответов. Получается 4 рода ответов. При статистической проверке гипотез отрицательный класс соответствует нулевой гипотезе (нет тревоги), а положительный – альтернативе (бьём тревогу).

Таблица . Матрица ошибок для задачи классификации с двумя классами.

|  |  |  |
| --- | --- | --- |
| PredictedclassActualclass | Результат распознавания P (выбрана альтернатива) | Результат распознавания N (выбрана нулевая гипотеза) |
| Истинный класс P (верна альтернатива) | TP – количество истинноположительных ответов (лечим больного) | FN – количество ложноотрицательных ответов, ошибки II рода, пропуск цели (упустили больного) |
| Истинный класс N (верна нулевая гипотеза) | FP – количество ложноположительных ответов, ошибки I рода, ложная тревога (лечим здорового) | TN – количество истинноотрицательных ответов (отпустили здорового) |

Если не оговаривается иного, то предполагается, что матрица ошибок построена по конечной выборке, содержащей объекты обоих классов, т. е. $1\leq TP+FN$ и $1\leq FP+TN$. Назовём это стандартным случаем.

Частоты (frequencies), стоящие в ячейках матрицы ошибок, являются исходными значениями для различных показателей качества классификации. Поскольку хочется, чтобы значения показателей качества были сопоставимы для выборок разного размера, принято от частот переходить к долям (rates, ratios).

Показатель качества классификации «общая точность» (accuracy, ACC) характеризует матрицу ошибок в целом и определяется формулой $ac=\frac{TP+TN}{TP+FN+FP+TN}$. Случай деления на ноль при положительном числителе невозможен. Неопределённость 0/0 возникает тогда и только тогда, когда выборка пустая, соответственно, эту неопределённость не раскрывают. При непустой выборке и в стандартном случае неопределённость 0/0 невозможна. Случаи $ac<0$ или $ac>1$ невозможны. Любое рациональное значение из [0,1] возможно даже в стандартном случае.

Случай $ac=1$ означает классификатор, который правильно ответил на всех объектах выборки. Случай $ac=0$ означает классификатор, который ошибается на всех объектах выборки. В случае бинарной классификации инверсия такого классификатора с $ac=0$ даст классификатор с $ac=1$. Статистически худший бинарный классификатор – подкидывание монеты.

Показатель качества классификации «полнота» (recall, sensitivity, true positive rate, TPR) характеризует распознавание положительного класса и определяется формулой $rc=\frac{TP}{TP+FN}$. Случай деления на ноль при положительном числителе невозможен. Неопределённость 0/0 возникает тогда и только тогда, когда выборка не содержит положительных объектов, соответственно, эту неопределённость не раскрывают. В стандартном случае неопределённость 0/0 невозможна. Случаи $rc<0$ или $rc>1$ невозможны. Любое рациональное значение из [0,1] возможно даже в стандартном случае.

Показатель качества классификации «точность» (precision, positive predictive value, PPV) характеризует распознавание положительного класса и определяется формулой $pr=\frac{TP}{TP+FP}$. Случай деления на ноль при положительном числителе невозможен. Неопределённость 0/0 возникает тогда и только тогда, когда выборка пуста или классификатор отвергает (объявляет отрицательными) все объекты выборки. Неопределённость 0/0 возможна даже в стандартном случае. Если допустим классификатор, отвергающий все объекты, то неопределённость 0/0 возможна всегда. Случаи $pc<0$ или $pc>1$ невозможны. Любое рациональное значение из [0,1] возможно даже в стандартном случае.

Если матрица ошибок построена по конечной выборке, то значения всех этих показателей являются рациональными числами.

# Общий подход к задачам

Не любые сочетания значений $ac$, $rc$ и $pc$ возможны. Соответственно, возникают задачи на их допустимое сочетание.

Традиционно предполагается, что тестирование классификатора проходило на конечной выборке, возникла матрица ошибок, которая содержит целые неотрицательные числа. Основной подход к решению таких задач состоит в том, что проверяется существование подходящих неотрицательных целых частот TP, TN, FP, FN. Возникает система из 4 неравенств неотрицательности и некоторых равенств по исходным данным. Логически полное решение должно упомянуть все 4 неравенства. Дополнительного внимания заслуживают крайние ситуации с нулевыми частотами.

# Задача на возможные значения точности

Рассматривается задача классификации на два класса: положительный и отрицательный. В ходе тестирования классификатора получены следующие результаты. Полнота равна $rc$, общая точность равна $ac$. Какие значения может принимать точность $pc$?

Поскольку задано определённое значение $rc$, то $1\leq TP+FN$, выборка содержит положительные объекты.

Сразу отметим, что при $ac<0$, или $ac>1$, или $rc<0$, или $rc>1$ исходные данные следует считать неверными. Ответ: $pc\in ∅$.

Если $rc=0$, то $TP=0$ и $1\leq FN$. Следовательно, $ac<1$.

Если $rc=0$ и $ac=1$, то ответ: $pc\in ∅$.

Если $rc=0$ и $0<ac<1$, то ответ: $pc\in \{0, 0/0\}$. Даже в стандартном случае.

Если $rc=0$ и $0=ac$, то ответ: $pc\in \{0, 0/0\}$. В стандартном случае только $pc=0$.

Если $ac=0$, то $TN=0$, $TP=0$, $1\leq FN$. Следовательно, $rc=0$.

Если $ac=0$ и $rc\ne 0$, то ответ: $pc\in ∅$.

Случай $ac=0$ и $rc=0$ есть выше.

Ниже $ac>0$ и $rc>0$.

Если $ac=1$, то $FN=0$, $FP=0$, $1\leq TP$. Следовательно, $rc=1$.

Если $rc=1$ и $ac=1$, то ответ: $pc=1$.

Если $rc\ne 1$ и $ac=1$, то ответ: $pc\in ∅$.

Ниже $1>ac>0$ и $rc>0$. Тогда $1\leq TP$, $pc>0$. Выразим остальные частоты как доли положительного TP.

$\frac{TN}{TP}=x$, надо обеспечить неотрицательность;

$\frac{FN}{TP}=\frac{1}{rc}-1$, всегда неотрицательно;

$\frac{FP}{TP}=\frac{1}{ac}-\frac{1}{rc}+x(\frac{1}{ac}-1)$, надо обеспечить неотрицательность.

$pc=\frac{1}{1+\frac{1}{ac}-\frac{1}{rc}+x(\frac{1}{ac}-1)}$. Интересно отметить, что с ростом TN точность $pc$ падает.

Неотрицательность FP соответствует условию $x\left(\frac{1}{ac}-1\right)\geq \frac{1}{rc}-\frac{1}{ac}$. Случай $ac=1$ есть выше. Сейчас множитель при $x$ положителен. Получили систему

$$\left\{\begin{matrix}x\geq 0\\x\geq \left(\frac{1}{rc}-\frac{1}{ac}\right)/\left(\frac{1}{ac}-1\right) \end{matrix}\right.$$

Если $rc<ac$, то $\frac{FP}{TP}\geq 0$, иначе $\frac{FP}{TP}\geq \frac{1}{ac}-\frac{1}{rc}$.

Если $0<rc\leq ac<1$, то ответ: $pc\in (0,1]$, рациональное. Примечание: нужно большое TN, можно любое FP.

Если $0<ac<rc\leq 1$, то ответ: $pc\in (0,\frac{1}{1+\frac{1}{ac}-\frac{1}{rc}} ]$, рациональное. Примечание: можно любое TN, нужно большое FP.