

Научная статья
УДК 51-7
DOI 10.26297/2312–9409.2025.6.10

Инструменты нахождения вероятностей, характеризующих достоверность бинарной и многоклассовой классификации

**Николай Андреевич Белоусов^{1✉}, Анастасия Валерьевна Бирбасова¹,
Михаил Юрьевич Захаров²**

¹*Краснодарское высшее военное орденов Жукова и Октябрьской Революции Краснознаменное училище имени генерала армии С.М. Штеменко, Краснодар, Россия,*

²*Кубанский государственный университет, Краснодар, Россия*
loki911112@mail.ru[✉]

Аннотация. Исследуются способы нахождения вероятностей, системно характеризующих достоверность бинарной и многоклассовой классификации. Проанализированы соответствующие инструменты, основанные на формуле полной вероятности и формуле Байеса.

Ключевые слова: системный подход, системный анализ, статистические гипотезы, бинарная и многоклассовая классификация, ошибки классификации, случайное событие, вероятность, условная вероятность, формула полной вероятности, формула Байеса.

Tools for finding probabilities characterizing the validity of binary and multi-class classification

Nikolay A. Belousov^{1✉}, Anastasia V. Birbasova¹, Michail Y. Zaharov²

¹*Krasnodar Higher Military Orders of Zhukov and the October Revolution Red Banner School named after General of the Army S.M.Shtemenko, Krasnodar, Russia,*

²*Kuban State University, Krasnodar, Russia*
lloki911112@mail.ru[✉]

Abstract. This paper investigates methods for finding probabilities that systemically characterize the validity of binary and multiclass classification. The corresponding tools, based on the law of total probability and Bayes' theorem, are analyzed.

Keywords: systems approach, systems analysis, statistical hypotheses, binary and multiclass classification, classification errors, random event, probability, conditional probability, law of total probability, Bayes' theorem.

Очевидно, что очень часто реальные явления, процессы, устройства весьма сложны и «многомерны». Но во многих случаях для целей системного анализа, моделирования и классификации указанные объекты исследования удается разбить

на условные «подобъекты», либо «классы», соответствующие различным альтернативным «условиям» функционирования или «свойствам» исходных объектов. Формализовать такие альтернативные «условия» («свойства») можно, в частности, с помощью аппарата вероятностных событий-гипотез (используемых в формуле полной вероятности и в формуле Байеса). Направления исследований объектов также нередко определяются указанными гипотезами. Естественно, что при таких направлениях исследований и соответствующие «решения» по объектам будут иметь форму данных гипотез. То есть указанные «решения» будут созвучны гипотезам, соответствовать, либо противоречить им и, как и гипотезы, могут быть правильными, либо ошибочными. В последнем случае, такие «решения» будут реализовывать ошибки 1-го и 2-го рода при бинарной классификации или, в общем случае, ошибки при многоклассовой классификации.

Общая теория и сведения по ошибкам 1-го и 2-го рода изложены, в частности, в [1]–[3]. Одному из методов многоклассовой классификации посвящена, например, работа [4]. В [5], [6] исследовались методы нахождения вероятностей ошибок 1-го и 2-го рода при бинарной классификации, а также проводился анализ вероятностей, связанных с ошибками 1-го и 2-го рода при бинарной классификации.

В данной работе делается попытка определенной систематизации и обобщения способов нахождения вероятностей, характеризующих достоверность бинарной и многоклассовой классификации, основанных на формуле полной вероятности и формуле Байеса ([3]).

Актуальность данной работы заключается в том, что применение полученных в ней результатов позволит проще выполнять корректное моделирование и оптимизацию вероятностей ошибок указанных видов классификации, что позволит улучшать достоверность системного анализа, моделирования или классификации реальных явлений, процессов, устройств.

Рассмотрим (априорные) вероятности гипотез при бинарной классификации: основной (нулевой) гипотезы H_0 и альтернативной гипотезы H_1 . Введем также соответствующие события ([6]):

$$H_0 = \{ \text{гипотеза } H_0 \text{ верна} \}, \quad (1)$$

$$H_1 = \{ \text{гипотеза } H_1 \text{ верна} \}. \quad (2)$$

В силу рассмотрения процесса бинарной классификации, будем считать события (1), (2) противоположными друг другу:

$$H_1 = \overline{H_0}. \quad (3)$$

В силу соотношения (3) выполняется:

$$P(H|0) + P(H|1) = 1, \quad (4)$$

причем в (4) достаточно знать хотя бы одну вероятность в левой части равенства.

Аналогично, рассмотрим (априорные) вероятности гипотез при многоклассовой классификации: альтернативных гипотез H_i , $i = 1, 2, \dots, n$. Соответствующие события будут такими:

$$H_i = \{ \text{гипотеза } H_i \text{ верна} \}, i = 1, 2, \dots, n. \quad (5)$$

В силу рассмотрения процесса многоклассовой классификации, будем считать события (5) полной группой несовместных событий:

$$H_i \cdot H_j = \emptyset, i \neq j, \sum_{i=1}^n H_i = \Omega, \quad (6)$$

где \emptyset и Ω – соответственно, невозможное и достоверное события, $i, j = 1, 2, \dots, n$.

В силу соотношений (6) выполняется:

$$\sum_{i=1}^n P(H|i) = 1. \quad (7)$$

Рассмотрим далее вероятности ошибок и правильных «решений» соответствующего рода (по различным гипотезам). Это будут условные вероятности следующих видов:

$$P(H_{1\text{реш}}|H_0) = \alpha, P(H_{0\text{реш}}|H_1) = \beta, \quad (8)$$

– вероятности ошибок 1-го и 2-го рода и

$$P(H_{0\text{реш}}|H_0) = 1 - \alpha, P(H_{1\text{реш}}|H_1) = 1 - \beta \text{ (мощность критерия)} \quad (9)$$

– вероятности правильных решений при бинарной классификации;

$$P(H_{i\text{реш}}|H_j), i, j = 1, 2, \dots, n, i \neq j, \quad (10)$$

– вероятности ошибок « i -го и j -го рода» и

$$P(H_{i\text{реш}}|H_i), i = 1, 2, \dots, n, \quad (11)$$

– вероятности правильных решений при многоклассовой классификации.

В формулах (8)–(11) события вида $H_{i\text{реш}}$, $i = 0, 1$, либо $1, 2, \dots, n$, означают, что соответствующие принятые «решения» по объектам имеют вид гипотез H_i . Эти решения могут быть как правильными ((9), (11)), так и ошибочными ((8), (10)). Соответственно, события вида H_i , $i = 0, 1$, либо $1, 2, \dots, n$, означают истинные

(неизвестные для исследователя) гипотезы, при которых принимаются «решения» $H_{i_{реш}}$.

Правильные и ошибочные «решения» $H_{i_{реш}}$, описываемые формулами (8)–(11), очевидно, могут иметь место только при истинности соответствующих гипотез-условий H_i . Поэтому для системности исследования необходим также учет соответствующих вероятностей $P(H|i)$, $i=0,1$, либо $1,2,\dots,n$. Условно назовем формируемые указанным образом вероятности (12) – (15) «полными» вероятностями соответствующих ошибок, либо правильных решений при бинарной и многоклассовой классификации:

$$P(H|0)P(H_{1_{реш}}|H_0), P(H|1)P(H_{0_{реш}}|H_1), \quad (12)$$

—«полные» вероятности ошибок 1-го и 2-го рода и

$$P(H|0)P(H_{0_{реш}}|H_0), P(H|1)P(H_{1_{реш}}|H_1), \quad (13)$$

—«полные» вероятности правильных решений при бинарной классификации;

$$P(H|j)P(H_{i_{реш}}|H_j), i, j=1,2,\dots,n, i \neq j, \quad (14)$$

—«полные» вероятности ошибок « i -го и j -го рода» и

$$P(H|i)P(H_{i_{реш}}|H_i), i=1,2,\dots,n, \quad (15)$$

—«полные» вероятности правильных решений при многоклассовой классификации.

Любое решение $H_{i_{реш}}$, $i=0,1$, либо $1,2,\dots,n$, может быть правильным, либо ошибочным и быть принято при любой из гипотез-условий H_i . Следовательно, вероятность $H_{i_{реш}}$, $i=0,1$, либо $1,2,\dots,n$, необходимо определять по формуле полной вероятности (16):

$$P(H|i_{\text{реш}}) = \sum_{j=0(1)}^{1(n)} P(H|j) P(H_{i_{\text{реш}}}|H_j), \quad (16)$$

где $i=0,1$, либо $1,2,\dots,n$; суммирование от 0 до 1 соответствует бинарной классификации, а от 1 до n – многоклассовой классификации.

Перейдем к построению вероятностей общего правильного решения (Π) и общей ошибки (O) при 2 и при n альтернативных гипотезах. Эти вероятности будут определяться по следующим формулам соответственно:

$$P(\Pi) = \sum_{i=0(1)}^{1(n)} P(H|i) P(H_{i_{\text{реш}}}|H_i), \quad (17)$$

$$P(O) = \sum_{j=0(1)}^{1(n)} P(H|j) P\left(\sum_{\substack{i=0(1) \\ i \neq j}}^{1(n)} H_{i_{\text{реш}}}|H_j\right). \quad (18)$$

Можно показать, что выражения (17) и (18) также являются формулами полной вероятности. Действительно, событие «общее правильное решение», при условии верности гипотезы H_i (обозначим: $\Pi|H_i$), представляет собой событие $H_{i_{\text{реш}}}$ и формула (17) становится формулой полной вероятности для события Π . Аналогично, событием «общая ошибка», при условии верности гипотезы H_j (обозначим: $O|H_j$), будет являться событие $\sum_{\substack{i=0(1) \\ i \neq j}}^{1(n)} H_{i_{\text{реш}}}$ и формула (18) становится формулой полной вероятности для события O .

Построим теперь условные вероятности всех гипотез по всем решениям – также для случая бинарной и многоклассовой классификации. Пусть требуется

найти условную вероятность гипотезы H_j при условии события $H_{i\text{реш}}$, $i, j=0,1$, либо $1,2,\dots,n$. Воспользуемся формулой Байеса:

$$P(H_j|H_{i\text{реш}}) = \frac{P(H|j)P(H_{i\text{реш}}|H_j)}{P(H|i\text{реш})}, \quad (19)$$

где $P(H|i\text{реш})$ может быть вычислена по формуле (16), $i, j=0,1$, либо $1,2,\dots,n$.

Видно, что формула (19) подходит как для ситуации, когда истинная гипотеза и решение совпадают ($i = j$), так и для ситуации ошибочного решения ($i \neq j$). Формула (19) является в определенном смысле обратной к формулам (8)– (11), что делает ее полезной в определенных прикладных задачах.

В статье выполнена определенная систематизация и обобщение способов нахождения вероятностей, характеризующих достоверность бинарной и многоклассовой классификации, основанных на формуле полной вероятности и формуле Байеса. Применение полученных в статье результатов позволит проще выполнять корректное моделирование и оптимизацию вероятностей ошибок и правильных решений при указанных видах классификации, что позволит улучшать достоверность системного анализа, моделирования или классификации реальных явлений, процессов, устройств.

Список источников

1. ГОСТ Р 8.731-2010. СИСТЕМЫ ДОПУСКОВОГО КОНТРОЛЯ. Основные положения (2010) // Государственная система обеспечения единства измерений. Москва: Стандартинформ.
2. Корн, Г. и Корн, Т. Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – 2-е изд. – Москва : Издательство «Наука», 1973. – 832 с. – Текст : непосредственный.

3. Гмурман, В.Е. Теория вероятностей и математическая статистика: учебник для вузов / В.Е. Гмурман. — 12-е изд. — Москва: Издательство Юрайт, 2023. — 479 с. — (Высшее образование). — ISBN978-5-534-00211-9. — Текст : непосредственный.
4. Бубнов, А.А., Гусев, М.А., Майков, К.А., Пылькин, А.Н. Интерпретируемый метод многоклассовой классификации / Журнал «Вестник РГРТУ». 2021. № 76. С. 74-80.
5. Белоусов, Н.А., Махов, Д.С., Захаров, М.Ю. О методах нахождения вероятностей ошибок 1-го и 2-го рода при бинарной классификации / Электронный сетевой политематический журнал «Научные труды КубГТУ». 2024. № 3. С. 74-79.
6. Белоусов, Н.А., Махов, Д.С., Махмутов, Р.Д., Захаров, М.Ю. Анализ вероятностей, связанных с ошибками 1-го и 2-го рода при бинарной классификации / Электронный сетевой политематический журнал «Научные труды КубГТУ». 2024. № 5. С. 111-118.

References

1. GOST R 8.731-2010. SISTEMY DOPUSKOVOGO KONTROLYA. Osnovnyepolozheniya (2010) // Gosudarstvennayasistemaobespecheniyaedinstvaizmerenii. Moskva: Standartinform.
2. Korn, G. i Korn, T. Spravochnikpomatematikedlyanauchnykhrabotnikov i inzhenerov / G. Korn, T. Korn. — 2-e izd. — Moskva :Izdatel'stvo «Nauka», 1973. — 832 s. — Текст :neposredstvennyi.
3. Gmurman, V. E. Teoriyaveroyatnostei i matematicheskayastatistika :uchebnikdlyavuzov / V. E. Gmurman. — 12-e izd. —

Moskva :Izdatel'stvoYurait, 2023. — 479 s. — (Vyssheebrazovanie). — ISBN 978-5-534-00211-9. — Tekst :neposredstvennyi.

4. Bubnov, A.A., Gusev, M.A., Maikov, K.A., Pyl'kin, A.N. Interpretiruemyimetodmnogoklassovoiklassifikatsii / Zhurnal «Vestnik RGRTU». 2021. № 76. S. 74-80.
5. Belousov, N.A., Makhov, D.S., Zakharov, M.Yu. O metodakhnazhdeniyaveroyatnostei oshibok 1-go i 2-go rodapribinarnoiklassifikatsii / Elektronnyisetevoipolitematicheskii zhurnal «NauchnyetrudyKubGTU». 2024. № 3. S. 74-79.
6. Belousov, N.A., Makhov, D.S., Makhmutov, R.D., Zakharov, M.Yu. Analiz veroyatnostei, svyazannykh s oshibkami 1-go i 2-go rodapribinarnoiklassifikatsii / Elektronnyisetevoipolitematicheskii zhurnal «NauchnyetrudyKubGTU». 2024. № 5. S. 111-118.

Дата поступления рукописи в редакцию: 14.10.2025

Дата принятия рукописи в печать: 14.11.2025