УДК 519.7

А.А. Галкин

ГЛУБИННЫЙ МЕТОД КЛАССИФИКАЦИИ НА ОСНОВЕ МАСШТАБИРУЕМОГО РАССТОЯНИЯ МАХАЛАНОБИСА ДЛЯ МНОЖЕСТВ С НЕРАВНЫМИ АПРИОРНЫМИ ВЕРОЯТНОСТЯМИ

Введение

Основываясь на результатах научных исследований в области распознавания образов, можно утверждать, что непараметрические глубинные методы классификации могут использоваться в качестве эффективного и независимого от распределения инструмента классификации в случае равных априорных вероятностей множеств данных и при соответствии модели смещения расположения [1–3]. Однако на практике различные множества данных могут иметь неравные априорные вероятности, а также не принадлежать к общему семейству эллиптических распределений. Данная работа посвящена исследованию данной проблемы, в результате предлагается глубинный классификатор, который позволяет получить достаточно хорошие результаты при решении задач классификации, которые включают случай неравных априорных вероятностей.

Постановка задачи

Рассмотрим двухклассовую задачу, где необходимо классифицировать элемент данных z_0 в один из двух классов, используя элементы $z_{l1}, z_{l2}, ..., z_{lm_l}$ с l-го множества данных, где $l = 1{,}2$. Сначала вычисляются эмпирические функции полупространственной глубины $E_{m_l}(l, z_0)$ от элемента данных z_0 относительно данных из двух множеств, где $l = 1{,}2.$ Далее элементы данных проектируются от плотности c_l в некотором фиксированном направлении φ $(\|\varphi\| = 1)$. После этого находим две точки: x_1 и x_2, которые расположенны на противоположной стороне от центра и имеют функцию эмпирической глубины $E_{m_l}(l, z_0)$. Заметим, что оценка масштабируемого расстояния Махаланобиса $\beta_l\{E(l, z_0)\}\sqrt{\varphi' \Xi_l \varphi}$ является половиной расстояния между точками x_1 и x_2, т.е. $|x_1 - x_2|/2$.

Лемма 1. Оптимальный байесовский классификатор может быть задан как

$$\Im_{\text{opt}}(z) = \arg\max_{l} p_l \overline{\lambda}_l \{E(l, z)\}, \qquad (1)$$

если распределения множеств данных эллиптически-симметричны, а для функций глубины Махаланобиса, полупространственной, симплициальной, мажоритарной, симплициальной объемной и проекционной глубины существуют некоторые функции $\overline{\lambda}_l(\cdot)$ множественной глубины $E(l, z)$, зависящие от типа функции глубины.

Доказательство. Определим $C_l = \{(Z_l - \varepsilon_l)' \, \Xi_l^{-1} (Z_l - \varepsilon_l)\}^{1/2}$, где ε_l — параметр расположения, Ξ_l — матрица разброса l-го множества данных, которое имеет функцию плотности c_l, а $Z_l \approx c_l$.

Расстояние Махаланобиса от элемента z с параметром расположения ε_l определяется как

$$d_l = \{(z - \varepsilon_l)' \Xi_l^{-1} (z - \varepsilon_l)\}^{1/2}. \tag{2}$$

Поэтому распределения C_l задаются следующим образом:

$$\psi_l(d_l) = \frac{p^{r/2}}{\mathrm{I}r/2} \left| \Xi_l \right|^{1/2} d_l^{r-1} c_l(z), \;\; 0 < d_l < \infty, \tag{3}$$

когда функция плотности c_l эллиптически симметрична [4].

Стоит отметить, что расстояние Махаланобиса d_l является функцией от множественной глубины $E(l, z)$ в случае эллиптических множеств данных. Кроме того, $p_l c_l(z) > p_i c_i(z)$ тогда и только тогда, когда $p_l \left| \Xi_l \right|^{-1/2} \psi_l(d_l) / d_l^{r-1} >$ $> p_i \left| \Xi_i \right|^{-1/2} \psi_i(d_i) / d_i^{r-1}$.

В результате, поскольку $d_l = \beta_l \{E(l, z)\}$, оптимальный байесовский классификатор может быть задан как

$$\Im_{\mathrm{opt}}(z) = \arg\max_l p_l \bar{\lambda}_l \{E(l, z)\}, \tag{4}$$

где $\bar{\lambda}_l(\kappa) = \left| \Xi_l \right|^{-1/2} \psi_l \{\beta_l(\kappa)\} / \{\beta_l(\kappa)\}^{r-1}$. Лемма доказана.

Лемма 2. Определим $\rho_{\pi, \varphi, m}$ как π-й эмпирический квантиль $\varphi' Z$ для некоторого φ с $\| \varphi \| = 1$. Предположим, что существует множество данных с m независимыми и одинаково распределенными случайными величинами, в отношении которого величина a_m является эмпирической глубиной элемента z. Величины Z и a_m имеют эллиптически-симметричную плотность распределения с параметром расположения ε и матрицей разброса Ξ. Величина $(\rho_{1 - a_m, \varphi, m} - \rho_{a_m, \varphi, m}) / 2$ асимптотически сходится к $\{(z - \varepsilon)' \Xi^{-1} (z - \varepsilon)\}^{1/2} \sqrt{\varphi' \Xi \varphi}$ при $m \to \infty$.

Доказательство. Если величина $\rho_{\pi, \varphi}$ является π-м квантилем от $\varphi' Z$, $\rho_{a_m, \varphi, m}$ асимптотически сходится к $\rho_{a, \varphi}$, а $\rho_{1 - a_m, \varphi, m}$ асимптотически сходится к $\rho_{1 - a, \varphi}$, что следует из теоремы 2.

Таким образом, $(\rho_{1 - a_m, \varphi, m} - \rho_{a_m, \varphi, m}) / 2$ асимптотически сходится к $(\rho_{1 - a, \varphi} - \rho_{a, \varphi}) / 2$, где

$$(\rho_{1 - a, \varphi} - \rho_{a, \varphi}) / 2 = \{(z - \varepsilon)' \Xi^{-1} (z - \varepsilon)\}^{1/2} \sqrt{\varphi' \Xi \varphi}.$$

Полученный результат следует из теоремы 1.

Лемма доказана.

Функции $\bar{\lambda}_l$ являются монотонными и одинаковыми для всех множеств данных, когда функции плотности уменьшаются с расстоянием Махаланобиса от центра симметрии, а распределения множеств данных удовлетворяют модели

ISSN 0572-2691

смещения расположения. Поэтому при приведенных выше условиях байесовский классификатор (4) является классификатором максимальной глубины в случае равных априорных вероятностей [5]. В результате следующим шагом является нахождение соответствующих выборочных модификаций для $\overline{\lambda}_l\{E(l, z)\}$, что позволяет построить правило классификации на основе элементов данных выборки. Однако, поскольку $\overline{\lambda}_l\{E(l, z)\}$ — сложная функция от $E(l, z)$, для большинства функций глубины нахождение ее последовательной оценки на основе выборки данных — достаточно сложная задача.

Теорема 1. Пусть М — сферически распределенная случайная величина, ε — параметр расположения, а Ξ — параметр масштабирования. Кроме того, величина Z имеет эллиптическое распределение R такой формы:

$$Z = \varepsilon + \Xi^{1/2}\text{M}.$$

Тогда при $\rho_{\pi,\varphi}$, что является π-м квантилем от $\varphi' Z$, и $\rho = F_{hsd}(R, z)$, что является полупространственной глубиной элемента z относительно распределения R, имеет место неравенство

$$[\rho_{1-a,\varphi} - \rho_{a,\varphi}]/2 = \{(z-\varepsilon)'\Xi^{-1}(z-\varepsilon)\}^{1/2}\sqrt{\varphi'\Xi\varphi} \tag{5}$$

для любого z и заданного направления $\varphi(\|\varphi\| = 1)$.

Доказательство. Очевидно, что величина $\dfrac{\varphi'(Z-\varepsilon)}{\sqrt{\varphi'\Xi\varphi}}$ распределена как $j'\text{M}$ с $\|j\| = 1$, где функция полупространственной глубины элемента z относительно распределения R может быть выражена следующим образом:

$$F_{hsd}(R, z) = 1 - \sup_{\varphi} P\{\varphi'(Z - z) < 0\} = 1 - \sup_{\varphi} P\left\{\frac{\varphi'(Z-\varepsilon)}{\sqrt{\varphi'\Xi\varphi}} < \frac{\varphi'(z-\varepsilon)}{\sqrt{\varphi'\Xi\varphi}}\right\} \tag{6}$$

Поэтому

$$a = F_{hsd}(R, z) = 1 - \sup_{\varphi}\Phi\left[\frac{\varphi'(z-\varepsilon)}{\sqrt{\varphi'\Xi\varphi}}\right] = 1 - \Phi\left[\sup_{\varphi}\frac{\varphi'(z-\varepsilon)}{\sqrt{\varphi'\Xi\varphi}}\right] = 1 - \Phi[\{(z-\varepsilon)'\Xi^{-1}(z-\varepsilon)\}^{1/2}], \tag{7}$$

где Φ — функция распределения от $j'\text{M}$ для каждого j с $\|j\| = 1$.

Итак,

$$P\{\varphi' Z < \rho_{\pi,\varphi}\} = \Phi\left(\frac{\rho_{\pi,\varphi} - \varphi'\varepsilon}{\sqrt{\varphi'\Xi\varphi}}\right) = \pi, \tag{8}$$

что следует из определения $\rho_{\pi,\varphi}$.

Далее получаем, что

$$\Phi\left(\frac{\rho_{a,\varphi} - \varphi'\varepsilon}{\sqrt{\varphi'\Xi\varphi}}\right) = a = 1 - \Phi[\{(z-\varepsilon)'\Xi^{-1}(z-\varepsilon)\}^{1/2}] = \Phi[-\{(z-\varepsilon)'\Xi^{-1}(z-\varepsilon)\}^{1/2}] \tag{9}$$

и

$$\Phi\left(\frac{\rho_{1-a,\varphi} - \varphi'\varepsilon}{\sqrt{\varphi'\Xi\varphi}}\right) = 1 - a = \Phi[\{(z-\varepsilon)'\Xi^{-1}(z-\varepsilon)\}^{1/2}], \tag{10}$$

где $\pi = a$ и $\pi = 1 - a$.

В результате, учитывая, что Φ — строго монотонная функция,

$$\frac{\rho_{1-a,\,\varphi}-\varphi'\varepsilon}{\sqrt{\varphi'\Xi\varphi}}=-\frac{\rho_{a,\varphi}-\varphi'\varepsilon}{\sqrt{\varphi'\Xi\varphi}}=\{(z-\varepsilon)'\Xi^{-1}(z-\varepsilon)\}^{1/2}, \tag{11}$$

откуда следует, что

$$[\rho_{1-a,\,\varphi}-\rho_{a,\,\varphi}]/2=\{(z-\varepsilon)'\Xi^{-1}(z-\varepsilon)\}^{1/2}\sqrt{\varphi'\Xi\varphi}. \tag{12}$$

Теорема доказана.

Теорема 2. Пусть a и a_m — функции полупростанственной глубины элемента данных z относительно Φ и Φ_m соответственно. Кроме того, пусть θ_π — уникальное решение уравнения $\Phi(z)=\pi$, а $\theta_{\pi,\,m}$ — ее эмпирическая модификация на основе эмпирической функции распределения Φ_m, где $(0<\pi<1)$. Тогда $\left|\theta_{a_m,\,m}-\theta_a\right|$ асимптотически стремится к 0 при $m\to\infty$**.**

Доказательство. Для каждого $\phi>0$ существует такое $\overline{\omega}>0$, что $\left|a_m-a\right|<\overline{\omega}$. Отсюда следует, что $\left|\theta_{a_m}-\theta_a\right|<\phi/2$, поскольку θ_π — непрерывная функция от π. В результате

$$P\{\left|\theta_{a_m}-\theta_a\right|>\phi/2\}<P\{\left|a_m-a\right|>\overline{\omega}<2m^r e^{-2m\overline{\omega}^2}. \tag{13}$$

Поскольку $h_m=\min\{\Phi(\theta_{a_m}+\phi/2)-a_m,\ a_m-\Phi(\theta_{a_m}-\phi/2)\}$ для каждого a_m $(0<a_m<1)$, имеет место неравенство

$$P\{\left|\theta_{a_m,\,m}-\theta_{a_m}\right|>\phi/2\}<2e^{-2mh_m^2}, \tag{14}$$

полученное с использованием теоремы Серфлинга. Поэтому

$$P\{\left|\theta_{a_m,\,m}-\theta_a\right|>\phi\}<2m^r e^{-2m\overline{\omega}^2}+2e^{-2mh_m^2}. \tag{15}$$

Можно утверждать, что всегда существует такое целое число m_0 и такой интервал $\Lambda=[a-b,\ a+b]$, что $a_m\in\Lambda$ для всех $m>m_0$ и $(0<a-b<a+b<1)$. Данный вывод получен на основе результатов сходимости функции эмпирической полупростанственной глубины, где a_m асимптотически сходится к a при $m\to\infty$.

Поскольку $\min_m h_m>\inf_{\kappa\in\Lambda}[\min\{\Phi(\theta_\kappa+\phi/2)-\kappa,\ \kappa-\Phi(\theta_\kappa-\phi/2)\}]=n>0$, а также, используя лемму Бореля–Кантелли, получаем, что

$$P\{\left|\theta_{a_m,\,m}-\theta_a\right|>\phi\}<2m^r e^{-2m\overline{\omega}^2}+2e^{-2mn^2}, \tag{16}$$

для всех $m>m_0$.

Теорема доказана.

Как уже было показано в доказательстве теоремы 1, существует определенная связь между расстоянием Махаланобиса и функцией полупространственной глубины с параметром расположения ε_l и параметром разброса Ξ_l. Поэтому, используя функцию полупространственной глубины, величину $\overline{\lambda}_l\{E(l,z)\}$ можно выразить в простой форме, а именно:

$$\overline{\lambda}_l\{E(l,z)\}=\left|\Xi_l\right|^{-1/2}\psi_l(\beta_l\{E(l,z)\})/(\beta_l\{E(l,z)\})^{r-1}, \tag{17}$$

где $\psi_l(\cdot)$ — функция плотности от $\beta_l\{E(l,z)\}$. В данном случае выражение

$$E(l,z) = 1 - \Phi_l(\beta_l\{E(l,z)\}) \tag{18}$$

указывает на связь между расстоянием Махаланобиса $\beta_l\{E(l,z)\}$ и функцией глубины $E(l,z)$, а $\Phi_l(\cdot)$ — функция распределения от $\varphi'\Xi_l^{-1}(Z_l - \varepsilon_l)$ для любого φ с $\|\varphi\| = 1$.

Для построения правила решения, которое обеспечивает получение коэффициентов ошибочной классификации на уровне, близком к оптимальному байесовскому риску, применяются последовательные оценки Ξ_l, $\beta_l\{E(l,z)\}$ и ее функции плотности $\psi_l(\cdot)$. Функция полупространственной глубины используется для оценки β_l, а для построения модифицированного глубинного классификатора и для оценки ψ_l — ядерные оценки плотности [6]. Применение данного подхода исключает применение любых оценок на основе момента ε_l и Ξ_l, однако требует последовательной оценки для $|\Xi_l|$, т.е. включение соответствующей оценки для Ξ_l обеспечивает устойчивость алгоритма классификации к выбросам и экстремальным значениям.

Кроме того, имея постоянные величины $V_1, V_2, ..., V_L$, данную проблему можно решить, записав классификатор в форме

$$\Im(z) = \arg\max_l V_l \psi_l(\beta_l\{E(l,z)\}) / (\beta_l\{E(l,z)\})^{r-1}, \tag{19}$$

частота ошибок которого зависит от этих постоянных величин. Поэтому получение оптимального байесовского правила требует проведения минимизации частоты ошибок относительно $V_1, V_2, ..., V_L$. Используя выборку данных, можно минимизировать частоту ошибок результирующего классификатора для построения окончательного правила классификации. Это возможно только после нахождения последовательных оценок для ψ_l и β_l. Также можно взять $V_1 = 1$ и минимизировать частоту ошибок по другим $L-1$ независимым параметрам.

Оценочный алгоритм и его свойства сходимости

Используя различные направления $\varphi_1, \varphi_2, ..., \varphi_W$, происходит итеративное повторение оценочного алгоритма. Поэтому, принимая

$$\alpha_l = \sum_{w=1}^{W} \sqrt{\varphi'_w \sum_l \varphi_w} / W, \tag{20}$$

среднее значение оценок определяется конечной оценкой $\bar{b}_0^{(l)}$ для масштабируемого расстояния Махаланобиса $b_0^{(l)} = \alpha_l \beta_l\{E(l, z_0)\}$. Заметим, что при таком масштабном преобразовании были изменены только выражения V_1, V_2 на V_1^*, V_2^*, где $V_2^* / V_1^* = (\alpha_1/\alpha_2)^r V_2 / V_1$, а форма модифицированного глубинного классификатора осталась неизменной.

Оставляя соответствующую точку данных, оцениваем масштабируемое расстояние Махаланобиса в каждой такой точке с помощью метода остаточного прохода, т.е. характерно, что данная процедура применяется не только для точки z_0. Итак, масштабируемые расстояния Махаланобиса от z_{li} формируют центры первого и второго множества данных. Поскольку $\overline{b}_{li} = (\overline{b}_{li}^{(1)}, \overline{b}_{li}^{(2)})$ определяет оценку $(b_{il}^{(1)}, b_{li}^{(2)})$, получено определенное количество двумерных элементов данных $\overline{b}_{11}, \overline{b}_{12}, ..., \overline{b}_{1m_1}$ и $\overline{b}_{21}, \overline{b}_{22}, ..., \overline{b}_{2m_2}$. Учитывая, что $\mathrm{A}_l(b) = \alpha_l \psi_l(\alpha_l b)$, функция плотности A_l от масштабируемого расстояния Махаланобиса оценивается с помощью ядерного метода с использованием $\overline{b}_{li}^{(l)}$, которые являются элементами данных с l-го множества $(i = 1, 2, ..., m_l,\ \ l = 1,2)$.

Заметим, что для решения задач классификации более эффективно использование самой большой полосы пропускания, которая минимизирует коэффициент ошибочной классификации, используя метод перекрестной проверки. Кроме того, такая стратегия обеспечивает более высокие результаты, нежели минимизация с использованием полос пропускания для оценки среднеквадратичной интегрированной ошибки ядерных оценок плотности распределения данных [7]. Итак, поскольку необходимо оценить значение $V^* = V_2^* / V_1^*$, используем остаточный метод перекрестной проверки для одновременной оценки V^* и полос пропускания, что позволяет найти частоты ошибок на основе метода перекрестной проверки [8].

Предположим, что $\overline{\mathrm{A}}_{1a_{1m_1}}^*$ и $\overline{\mathrm{A}}_{2a_{2m_2}}^*$ — ядерные оценки плотностей масштабируемых расстояний Махаланобиса для двух множеств данных, а a_{1m_1} и a_{2m_2} — оцененные полосы пропускания. Когда величина $\overline{V}^*$ является оценкой величины V^*, полученной с помощью метода перекрестной проверки, классификация элемента данных z_0 в первое множество данных может быть осуществлена только в случае, если $\overline{\mathrm{A}}_{1a_{1m_1}}^* (\overline{b}_0^{(1)}) / \{\overline{b}_0^{(1)}\}^{r-1} > \overline{V}^* \overline{\mathrm{A}}_{2a_{2m_2}}^* (\overline{b}_0^{(2)}) / \{\overline{b}_0^{(2)}\}^{r-1}$.

Отметим, что величина $\overline{b}_0^{(l)}$ сходится почти наверное к $b_0^{(l)}$ при $l = 1,2$, что следует из леммы 2. Кроме того, величина $\overline{\mathrm{A}}_{la_{lm_l}}^* (\overline{b}_0^{(l)})$ сходится к $\mathrm{A}_l(b_0)$ при соответствующих условиях регулярности, которые приведенны в теореме 3. В результате оценки V^* обеспечат получение коэффициентов ошибочной классификации, близких к оптимальному байесовскому риску [9].

Для решения задач с более чем двумя множествами данных был применен аналогичный подход для нахождения $\overline{b}_0^{(l)}$ и $\overline{\mathrm{A}}_{la_{lm_l}}^* (\overline{b}_0^{(l)})$, где $l = 1, 2, ..., L.$ В результате на основе предложенного подхода можно построить модифицированный глубинный классификатор такой формы:

$$\Im_{E^*}(z_0) = \arg\max_l V_l \overline{\mathrm{A}}_{la_{lm_l}}^* (\overline{b}_0^{(l)}) / \{\overline{b}_0^{(l)}\}^{r-1}. \tag{21}$$

Заметим, что $L - 1$ независимых параметров $V_2/V_1, V_3/V_1, ..., V_L/V_1$ имеют непосредственное влияние на частоту ошибок классификатора (21). Учитывая данный факт, результирующий классификатор можно получить путем минимизации частоты ошибок по указанным параметрам $V_2/V_1, V_3/V_1, ..., V_L/V_1$.

Теорема 3. Пусть величина $\rho_{\pi,\varphi}$ является π-м квантилем $\varphi' Z$, а $\rho_{\pi,\varphi,m}$ — π-м эмпирическим квантилем $\varphi' Z$ для некоторого φ с $\|\varphi\| = 1$. Также пусть Z имеет эллиптическое распределение, формы $Z = \varepsilon + \Xi^{1/2}\mathrm{M}$, где $z_1, z_2, ..., z_m$ — элементы данных с R. Определим b_0 и $\overline{b}_0$ для нового элемента данных z_0, а также $b_i = \{(z_i - \varepsilon)'\Xi^{-1}(z_i - \varepsilon)\}^{1/2}\sqrt{\varphi'\Xi\varphi}$ и $\overline{b}_i = (\rho_{1-a_{i_m},\varphi,m} - \rho_{a_{i_m},\varphi,m})/2$ для $i = 1, 2, ..., m$, где a_{i_m} — эмпирическая полупространственная глубина z_i для некоторого заданного $\varphi(\|\varphi\| = 1)$. Предположим, что b_i имеют функцию плотности А. Для полосы пропускания $\widetilde{a}_m > 0$ и некоторой ядерной функции Θ определим ее ядерную оценку плотности $\overline{\mathrm{A}}^*_{\widetilde{a}_m}(b) = \dfrac{1}{m\widetilde{a}_m}\sum_{i=1}^{m}\Theta\{(b - \overline{b}_i)/\widetilde{a}_m\}$. Предположим также, что функция А имеет ограниченную третью производную, а функция Θ является симметричной, имеет ограниченную первую производную, а также выполняется неравенство $\int|\kappa|^3\Theta^2(\kappa)\,d\kappa < \infty$. Кроме того, пусть $\widetilde{a}_m \to 0$ и $m\widetilde{a}_m \to \infty$ при $m \to \infty$. Тогда, $\overline{\mathrm{A}}^*_{\widetilde{a}_m}(\overline{b}_0)$ вероятностно сходится к $\mathrm{A}(b_0)$ при $m \to \infty$.

Доказательство. Для каждого $\phi > 0$

$$P\{|\overline{\mathrm{A}}_{\widetilde{a}_m}(b_0) - \overline{\mathrm{A}}^*_{\widetilde{a}_m}(\overline{b}_0)| > \phi\} < mP\{|\Theta\{(b_0 - b_1)/\widetilde{a}_m\} - \Theta\{(\overline{b}_0 - \overline{b}_1)/\widetilde{a}_m| > \widetilde{a}_m\phi\}, \quad (22)$$

что следует из определения $\overline{\mathrm{A}}_{\widetilde{a}_m}(.)$ и $\overline{\mathrm{A}}^*_{\widetilde{a}_m}(.)$, а также, принимая тот факт, что

$$\overline{\mathrm{A}}_{\widetilde{a}_m}(b) = \frac{1}{m\widetilde{a}_m}\sum_{i=1}^{m}\Theta\{(b - b_i)/\widetilde{a}_m\}.$$

Для некоторого b, что находится между значениями $(b_0 - b_1)$ и $(\overline{b}_0 - \overline{b}_1)$, имеет место такое неравенство:

$$\Theta\{(b_0 - b_1)/\widetilde{a}_m\} - \Theta\{(\overline{b}_0 - \overline{b}_1)/\widetilde{a}_m\} = \frac{1}{\widetilde{a}_m}\{(b_0 - b_1) - (\overline{b}_0 - \overline{b}_1)\}\,\Theta'(b/\widetilde{a}_m). \quad (23)$$

Итак,

$$P\{|\overline{\mathrm{A}}_{\widetilde{a}_m}(b_0) - \overline{\mathrm{A}}^*_{\widetilde{a}_m}(\overline{b}_0)| > \phi\} < mP\{|(b_0 - b_1) - (\overline{b}_0 - \overline{b}_1)| >$$

$$> \widetilde{a}_m N\phi\} < 2mP\{|b_0 - \overline{b}_0| > \widetilde{a}_m N\phi/2\}, \quad (24)$$

когда первая производная функции $\Theta(.)$ ограничена N.

Используя теорему 1, можно увидеть, что

$$|b_0 - \overline{b}_0| \le \frac{1}{2}\{|\rho_{a,\varphi} - \rho_{a_m,\varphi,m}| + |\rho_{1-a,\varphi} - \rho_{1-a_m,\varphi,m}|\}. \quad (25)$$

Итак, для $\overline{\omega} > 0$ и $n > 0$

$$P\{|b_0 - \overline{b}_0| > \widetilde{a}_m N\phi/2\} < mP\{|\rho_{a,\varphi} - \rho_{a_m,\varphi,m}| >$$

$$> \widetilde{a}_m N\phi/2\} < 4(m^r e^{-m\widetilde{a}_m^2\overline{\omega}^2/2} + e^{-m\widetilde{a}_m^2\overline{\omega}^2/2}), \quad (26)$$

откуда следует, что

$$P\{|\overline{\mathrm{A}}_{\widetilde{a}_m}(b_0) - \overline{\mathrm{A}}^*_{\widetilde{a}_m}(b_{0_m})| > \phi\} < 8(m^{r+1}e^{-m\widetilde{a}_m^2\overline{\omega}^2/2} + e^{-m\widetilde{a}_m^2 n^2/2}). \quad (27)$$

Из неравенства (27) можно сделать вывод, что $\left|\overline{\mathrm{A}}^{*}_{\tilde{a}_m}(\bar{b}_0) - \overline{\mathrm{A}}_{\tilde{a}_m}(b)\right|$ вероятностно сходится к 0 при $m \to \infty$. Поэтому $\left|\overline{\mathrm{A}}_{\tilde{a}_m}(b) - \mathrm{A}(b)\right|$ вероятностно сходится к 0 при $m \to \infty$, учитывая соответственно форму математического ожидания и дисперсии $\overline{\mathrm{A}}_{\tilde{a}_m}(b)$, а именно:

$$\Omega\{\overline{\mathrm{A}}_{\tilde{a}_m}(b)\} = \mathrm{A}(b) + O(\tilde{a}_m^2), \tag{28}$$

$$D\{\overline{\mathrm{A}}_{\tilde{a}_m}(b)\} = O(m^{-1}\tilde{a}_m^{-1}). \tag{29}$$

Заметим, что полученная вероятностная сходимость $\left|\overline{\mathrm{A}}_{\tilde{a}_m}(b) - \mathrm{A}(b)\right|$ имеет место только при вышеприведенных условиях.

В результате $\overline{\mathrm{A}}^{*}_{\tilde{a}_m}(\bar{b}_0)$ вероятностно сходится к $\mathrm{A}(b)$.

Теорема доказана.

Заключение

В данной работе предложен непараметрический глубинный метод решения задач распознавания, когда различные множества данных имеют неравные априорные вероятности, а также когда они могут не принадлежать к общему семейству эллиптических распределений. Разработан универсальный глубинный классификатор, производительность которого не снижается в результате отклонения от модели, смещения расположения или нарушения монотонного характера функций плотности. Учитывая вычислительную сложность одновременной минимизации частоты ошибок относительно независимых параметров $V_1, V_2, ..., V_L$ и полос пропускания $a_1, a_2, ..., a_L$, мультиклассовая задача была разделена на ряд двухклассовых задач, принимая одновременно каждую пару классов и используя алгоритм для случая равных априорных вероятностей. Окончательные результаты всех парных сравнений были объединены с использованием метода голосования для получения конечного решения задачи классификации.

О.А. Галкін

ГЛИБИННИЙ МЕТОД КЛАСИФІКАЦІЇ НА ОСНОВІ МАСШТАБОВАНОЇ ВІДСТАНІ МАХАЛАНОБІСА ДЛЯ МНОЖИН З НЕРІВНИМИ АПРІОРНИМИ ЙМОВІРНОСТЯМИ

Запропоновано непараметричний глибинний метод класифікації для випадку, коли множини даних мають нерівні апріорні ймовірності та не належать до спільного сімейства еліптичних розподілів. Розроблено універсальний глибинний класифікатор, що не залежить від відхилення в моделі зсуву розташування або порушення монотонного характеру функцій щільності. Масштабовану відстань Махаланобіса оцінено в кожній точці з використанням методу залишкового проходу.

A.A. Galkin

DEPTH BASED CLASSIFICATION METHOD ON THE BASIS OF A SCALABLE MAHALANOBIS DISTANCE FOR SETS WITH UNEQUAL PRIOR PROBABILITIES

A nonparametric depth based method of classification is proposed for the case when data sets have unequal prior probabilities and do not belong to a common family of elliptical distributions. The multipurpose depth based classifier is developed that is independent of deviations in the location, shift model or violation of monotonous character of density functions. Scalable Mahalanobis distance is estimated at each point using the residual passage method.

1. *Pollard D.* Convergence of stochastic processes. — New York: Springer-Verlag, 1984. — P. 1–10.
2. *Zuo Y., Serfling R.* Structural properties and convergence results for contours of sample statistical depth functions // The Annals of Statistics. — 2000. — **28**. — P. 484–497.
3. *Holmes C.C., Adams N.M.* A probabilistic nearest neighbor method for statistical pattern recognition // Journal of the Royal Statistical Society. — 2002. — **64**. — P. 295–306.
4. *Godtliebsen F., Marron J.S., Chaudhuri P.* Significance in scale space for bivariate density estimation // Journal of Computational and Graphical Statistics. — 2002. — **11.** — P. 1–22
5. *Hall P.* Large sample optimality of least squares cross validations in density estimation // The Annals of Statistics. — 1983. — **11.** — P. 1156–1174.
6. *Mosler K.* Multivariate dispersions, central regions and depth. — New York: Springer-Verlag, 2002. — P. 271–292..
7. *Silverman B.W.* Density estimation for statistics and data analysis. — London: Chapman and Hall, 1986. — P. 1–22.
8. *Lachenbruch P., Mickey M.* Estimation of error rates in discriminant analysis // Technometrics. — 1968. — **10**. — P. 1–11.
9. *Hoeffding W.* Probability inequalitites for sums of bounded random variables // Journal of the American Statistical Association. — 1963. — **58**. — P. 14–27.

Получено 13.07.2015

Статья представлена к публикации чл.-корр. НАН Украины А.В. Анисимовым.