УДК: 004.93 + 519.7

© 2018 г. **А.В. Лапко** [1,2], д-р. техн. наук,
В.А. Лапко [1,2], д-р. техн. наук,
В.В. Молоков [3], канд. техн. наук
[1](Институт вычислительного моделирования СО РАН, Красноярск),
[2](Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, Красноярск),
[3](Сибирский юридический институт МВД РФ, Красноярск)

НЕПАРАМЕТРИЧЕСКИЕ АЛГОРИТМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ И МОДЕЛИ СТОХАСТИЧЕСКИХ ЗАВИСИМОСТЕЙ В УСЛОВИЯХ МАЛЫХ ВЫБОРОК

Рассматривается методика построения линейных и нелинейных непараметрических коллективов решающих функций в задачах распознавания образов и восстановления стохастических зависимостей. Предлагаемые системы обеспечивают эффективную обработку информации большой размерности. Приводятся результаты вычислительных экспериментов.

Ключевые слова: непараметрическая статистика, распознавание образов, непараметрические модели, плотность вероятности, малые выборки.

DOI: 10.22250/isu.2018.58.112-121

Введение

Традиционные непараметрические алгоритмы принятия решений ориентированы на представительные обучающие выборки [1 – 6]. Однако при решении прикладных задач часто располагают ограниченным объемом наблюдений – малой выборкой, что обусловливается нестационарностью объекта исследования, высокой стоимостью и сложностью получения необходимой информации. Получаемые на их основе решающие функции и правила не всегда обеспечивают приемлемые результаты, так как информация малых обучающих выборок недостаточна для оценивания вероятностных характеристик изучаемых закономерностей. Подобные условия встречаются при исследовании медико-биологических, экологических и социально-экономических объектов, когда отношение «объем/размерность» (n/k) обучающих выборок мало.

Для «обхода» проблем малых выборок широкое распространение получили методы группового учета аргументов (МГУА) [7]. Данный подход реализует последовательную процедуру усложнения решающего правила путем целенаправленного отбора (селекции) пар, состоящих из первичных и промежуточных признаков. Каждый этап синтеза алгоритма характеризуется отношением $n/k = n/2$. Алгоритмы МГУА отличаются друг от друга критерием селекции, количеством промежуточных моделей и их сложностью.

В работе предлагаются методика синтеза непараметрических алгоритмов распознавания образов и моделей стохастических зависимостей, основанная на декомпозиции обучающей выборки по ее размерности с целью увеличения отношения n/k и принципах коллективного оценивания.

Идея предлагаемого подхода при синтезе непараметрических систем принятия решений состоит в построении семейства частных решающих функций на основе результатов декомпозиции обучающей выборки по ее размерности с целью увеличения отношения n/k.

Частные решающие функции формируются на основе частей обучающей выборки, которые удовлетворяют одному или нескольким требованиям, – например, наличие однотипных признаков, возможность декомпозиции исходных признаков на группы в соответствии со спецификой решаемой задачи. Это порождает широкий круг постановок задач синтеза непараметрических решающих правил. При интеграции частных решающих функций используется методика построения их линейных и нелинейных коллективов.

Синтез непараметрических систем распознавания образов

Пусть $V = \left(x^i, \sigma\left(x^i\right), i = \overline{1, n}\right)$ – обучающая выборка объема n, составленная из значений признаков $x^i = \left(x_1^i, x_2^i, \ldots, x_k^i\right)$ классифицируемых объектов и соответствующих «указаний учителя» об их принадлежности к одному из двух классов

$$\sigma\left(x^i\right) = \begin{cases} -1, & \text{если } x^i \in \Omega_1 \\ 1, & \text{если } x^i \in \Omega_2 . \end{cases}$$

Условные плотности вероятности $p_j(x)$, $j = \overline{1, 2}$ распределения значений признаков x_v, $v = \overline{1, k}$ в области определения классов неизвестны. Причем отношение «объем выборки / размерность» (n/k) недостаточно для построения эффективных статистик оценивания $p_j(x)$, $j = \overline{1, 2}$.

Идея предлагаемого подхода к решению задачи распознавания образов в данных условиях состоит в выполнении следующих действий.

В соответствии с особенностями задачи классификации сформировать наборы признаков $\left(x(t),\ t=\overline{1,T}\right)$ и на этой основе осуществить декомпозицию исходной выборки $V=\left(x^i, \sigma\left(x^i\right), i=\overline{1,n}\right)$ на части $V(t)=\left(x^i(t), \sigma\left(x^i\right), i=\overline{1,n}\right),\ t=\overline{1,T}$. Здесь приняты обозначения $x(t)=\left(x_v, v\in I_t\right)$, где I_t – множество номеров признаков, составляющих их набор $x(t)$.

По полученным данным построить решающие правила [8]:

$$\overline{m}_t(x(t)):\begin{cases} x\in\Omega_1\,,\ \text{если}\ \bar{f}_{12}(x(t))\le 0 \\ x\in\Omega_2\,,\ \text{если}\ \bar{f}_{12}(x(t))>0\,,\ t=\overline{1,T}\,. \end{cases} \tag{1}$$

В качестве оценок $\bar{f}_{12}(x(t))$ частных решающих функций $f_{12}(x(t))$ между классами в пространстве признаков $x_v,\ v\in I_t\subset I=\left(v=\overline{1,k}\right)$ используются непараметрические статистики

$$\bar{f}_{12}(x(t))=\left[n\prod_{v\in I_t}c_v\right]^{-1}\sum_{i=1}^{n}\sigma\left(x^i\right)\prod_{v\in I_t}\Phi\left(\frac{x_v-x_v^i}{c_v}\right),\ t=\overline{1,T}\,, \tag{2}$$

где $\Phi(u_v)$ – ядерные функции, удовлетворяющие условиям [9, 10]:

$$\Phi(u_v)=\Phi(-u_v),\ 0\le\Phi(u_v)<\infty,\ \int_{-\infty}^{\infty}\Phi(u_v)du_v=1,\ \int_{-\infty}^{\infty}u_v^2\,\Phi(u_v)du_v=1,$$

$$\int_{-\infty}^{\infty}u_v^m\,\Phi(u_v)du_v<\infty\,,\ 0\le m<\infty\,,\ v=\overline{1,k}\,,$$

а $c_v=c_v(n)$ – последовательность положительных чисел (коэффициенты размытости ядерных функций) такая, что $c_v\to 0$, $n c_v\to\infty$, при $n\to\infty$.

Оптимизация частных решающих правил (1) по коэффициентам размытости ядерных функций $c_v,\ v\in I_t$ осуществляется в режиме «скользящего экзамена» из условия минимума статистической оценки вероятности ошибки распознавания образов

$$\overline{\rho}(t)=\frac{1}{n}\sum_{j=1}^{n}1\left(\sigma\left(x^j\right),\overline{\sigma}\left(x^j(t)\right)\right),\ t=\overline{1,T}\,, \tag{3}$$

$$1\left(\sigma(j),\overline{\sigma}(j)\right)=\begin{cases}0,\ \text{если}\ \ \sigma\left(x^j\right)=\overline{\sigma}\left(x^j(t)\right)\\ 1\,,\ \text{если}\ \ \sigma\left(x^j\right)\ne\overline{\sigma}\left(x^j(t)\right),\end{cases}$$

где $\overline{\sigma}\left(x^j(t)\right)$ – решение алгоритма (1) о принадлежности ситуации $x^j(t)$ к одному из двух классов. При формировании решения $\overline{\sigma}\left(x^j(t)\right)$ ситуация $x^j(t)$ исключается из процесса обучения в непараметрической статистике (2).

Линейный коллектив решающих функций. На основе частных непараметрических уравнений разделяющих поверхностей (2) определим обобщенную решающую функций в виде

$$\bar{f}_{12}(x) = \sum_{t=1}^{T} \alpha_t \, \bar{f}_{12}(x(t)), \ \sum_{t=1}^{T} \alpha_t = 1 . \qquad (4)$$

Очевидно, что меньшей оценке вероятности ошибки распознавания образов $\overline{\rho}(t)$ (3) соответствует больший весовой коэффициент α_t в коллективе (4). Поэтому значения α_t будем рассчитывать по формуле

$$\alpha_t = \frac{\overline{\rho}^{-1}(t)}{\sum_{j=1}^{T} \overline{\rho}^{-1}(j)}, \ t = \overline{1, T} .$$

Тогда обобщенное решающее правило распознавания образов запишется в виде

$$\overline{m}(x): \begin{cases} x \in \Omega_1 , \ \textit{если} \ \bar{f}_{12}(x) \le 0 \\ x \in \Omega_2 , \ \textit{если} \ \bar{f}_{12}(x) > 0 . \end{cases}$$

<u>Нелинейный коллектив решающих функций</u>. Будем полагать, что в соответствии с вышеприведенными рекомендациями получены частные непараметрические уравнения разделяющей поверхности $\bar{f}_{12}(x(t)), \ t = \overline{1, T}$. Необходимо построить обобщенную решающую функцию в виде нелинейного преобразования $\bar{f}_{12}(x) = \Psi\left(\bar{f}_{12}(x(t)), t = \overline{1, T}\right)$. Для этого будем использовать непараметрические статистики ядерного типа.

Используя непараметрических оценок решающих функций (2), сформируем массив данных

$$\left(\bar{f}_{12}\left(x^i(1)\right), \bar{f}_{12}\left(x^i(2)\right), \ldots, \bar{f}_{12}\left(x^i(T)\right), \sigma\left(x^i\right), \ i = \overline{1, n}\right).$$

На этой основе построим решающее правило в пространстве значений $\bar{f}_{12}(x) = \left(\bar{f}_{12}(x(t)), t = \overline{1, T}\right)$:

$$\overline{m}\left(\bar{f}_{12}(x)\right): \begin{cases} x \in \Omega_1 , \ \textit{если} \ \overline{F}_{12}\left(\bar{f}_{12}(x)\right) \le 0 \\ x \in \Omega_2 , \ \textit{если} \ \overline{F}_{12}\left(\bar{f}_{12}(x)\right) > 0 , \end{cases} \qquad (5)$$

где непараметрическая оценка обобщенной решающей функции между классами имеет вид

$$\overline{F}_{12}\left(\bar{f}_{12}(x)\right) = \left[n \prod_{t=1}^{T} c_t \right]^{-1} \sum_{i=1}^{n} \sigma\left(x^i\right) \prod_{t=1}^{T} \Phi\left(\frac{\bar{f}_{12}(x(t)) - \bar{f}_{12}\left(x^i(t)\right)}{c_t} \right).$$

На первом уровне структуры системы классифицируемая ситуация $x = \left(x(t), \ t = \overline{1, T}\right)$ преобразуется в значения непараметрических оценок $\bar{f}_{12}(x(t))$, $t = \overline{1, T}$, в пространстве которых принимается решение $\overline{\sigma}(x)$ правилом (5) о принадлежности ситуации x к тому или иному классу.

Предлагаемая система классификации обеспечивает не только эффективное решение задач распознавания образов в условиях малых выборок, но и позволяет учитывать априорные сведения о виде частных решающих функций.

Ближайшими аналогами данного алгоритма являются метод комитетов Вл.Д. Мазурова [11] и алгебраический подход Ю.И. Журавлева [12].

Сравним вычислительную эффективность предложенной двухуровневой системы классификации и традиционного непараметрического алгоритма [6]. Будем считать, что время расчета одной ядерной функции составляет τ, размерность наборов признаков $x(t)$, $t=\overline{1,T}$ одинакова и равна $k1=k/T$.

Тогда максимальное время, необходимое для принятия решения традиционным непараметрическим алгоритмом и предлагаемой системой классификации при использовании технологии параллельных вычислений, составляют значения $t_{mp} \approx n\tau k$, $t_n \approx nk1\tau + nT\tau$. Определим отношение $t_{mp}/t_n = k/(k1+T)$, т.е. разработанная система классификации обладает большей вычислительной эффективностью, если $k1+T<k$.

Синтез коллектива непараметрических регрессий

Разовьем методику синтеза коллектива непараметрических решающих функций на задачу моделирования стохастических зависимостей в условиях малых выборок. Имеются статистические данные $V=\left(x^i, y^i, i=\overline{1,n}\right)$, составленные из значений аргументов $x=\left(x_\nu, \nu=\overline{1,k}\right)$ и функции y восстанавливаемой однозначной стохастической зависимости $y=\varphi(x)$. Объем n выборки V недостаточен для эффективного восстановления искомой функции.

<u>Линейный коллектив частных моделей.</u> В соответствии со спецификой решаемой задачи сформируем наборы признаков $\left(x(j), j=\overline{1,T}\right)$ из исходных $x=\left(x_\nu, \nu=\overline{1,k}\right)$ и построим семейство частных моделей $\varphi_j(x(j))$, $j=\overline{1,T}$ с использованием обучающих выборок $V_j=\left(x^i(j), y^i, i=\overline{1,n}\right)$, $j=\overline{1,T}$. Например, если в качестве частных моделей используется непараметрическая регрессия [13, 14], то

$$\bar{y}_j=\bar{\varphi}_j(x(j))=\frac{\sum\limits_{i=1}^{n} y^i \prod\limits_{\nu\in I_j} \Phi\left(\dfrac{x_\nu - x_\nu^i}{c_\nu}\right)}{\sum\limits_{i=1}^{n} \prod\limits_{\nu\in I_j} \Phi\left(\dfrac{x_\nu - x_\nu^i}{c_\nu}\right)},\ j=\overline{1,T}, \qquad (6)$$

где I_j – множество номеров аргументов из исходных $x=\left(x_\nu, \nu=\overline{1,k}\right)$, образующих их набор $x(j)$.

Интеграция частных моделей может быть осуществлена в линейном их коллективе

$$\bar{y} = \sum_{j=1}^{T} \alpha_j \, \bar{\varphi}_j(x(t)), \; \alpha_j = W_j^{-1} \Big/ \sum_{t=1}^{T} W_t^{-1} .$$

Здесь оценка среднеквадратической ошибки восстановления искомой зависимости с использованием модели $\bar{\varphi}_j(x(j))$ (6) определяется выражением

$$W_j = \frac{1}{n} \sum_{i=1}^{n} \left(y^i - \bar{\varphi}_j\left(x^i(j)\right)\right)^2, \; j = \overline{1,T} . \tag{7}$$

Параметры частных моделей (6) находятся из условия минимума критерия (7).

Нелинейный коллектив частных моделей. Обобщение частных моделей $\bar{\varphi}_j(x(j))$, $j = \overline{1,T}$ в единой решающей функции будем осуществлять с помощью непараметрической статистики

$$\bar{\bar{y}} = \bar{\bar{\varphi}}(x) = \frac{\sum_{i=1}^{n} y^i \prod_{j=1}^{T} \Phi\left(\frac{\bar{\varphi}_j(x(j)) - \bar{\varphi}_j\left(x^i(j)\right)}{c_j} \right)}{\sum_{i=1}^{n} \prod_{j=1}^{T} \Phi\left(\frac{\bar{\varphi}_j(x(j)) - \bar{\varphi}_j\left(x^i(j)\right)}{c_j} \right)}, \tag{8}$$

формируемой по выборке $\bar{V} = \left(\bar{\varphi}_j\left(x^i(j)\right), \; y^i, \; j = \overline{1,T}, \; i = \overline{1,n}\right)$.

Оптимизация непараметрического коллектива (8) по коэффициентам размытости ядерных функций c_j, $j = \overline{1,T}$ производится в режиме «скользящего экзамена» из условия минимума эмпирического критерия

$$W(c) = \frac{1}{n} \sum_{t=1}^{n} \left[y^t - \frac{\sum_{\substack{i=1 \\ i \neq t}}^{n} y^i \prod_{j=1}^{T} \Phi\left(\frac{\bar{\varphi}_j\left(x^t(j)\right) - \bar{\varphi}_j\left(x^i(j)\right)}{c_j} \right)}{\sum_{\substack{i=1 \\ i \neq t}}^{n} \prod_{j=1}^{T} \Phi\left(\frac{\bar{\varphi}_j\left(x^t(j)\right) - \bar{\varphi}_j\left(x^i(j)\right)}{c_j} \right)} \right]^2 .$$

Преимущества предлагаемой процедуры по сравнению с традиционной непараметрической регрессией состоят в возможности учета частичных априорных сведений о виде взаимосвязи между переменными исследуемой зависимости и «обходе» проблем малых выборок за счет снижения размерности задачи.

Анализ результатов вычислительных экспериментов

Проводится сравнение предлагаемой непараметрической системы классификации (5) с хорошо зарекомендовавшим себя на практике традиционным непараметрическим алгоритмом распознавания образов, который основывается на решающей функции типа (2) в пространстве признаков $x = \left(x_v\,,\, v = \overline{1,k}\right)$.

Исследования осуществлялись при решении двухальтернативной задачи распознавания образов с нелинейной разделяющей поверхностью. Законы распределения признаков в области первого класса формировались в соответствии с датчиками случайных чисел

$$x_v = a + \varepsilon\,(b - a),$$

$$x_{v+1} = (x_v)^2 - 6x_v + 10 + \sigma_1 \left(\sum_{i=1}^{p_1} \varepsilon^i - 0{,}5 p_1 \right) \frac{6}{\sqrt{3p_1}}\,,\ v \in I_н\,,$$

где параметры распределений $a = 1{,}5$; $b = 4{,}5$; $p_1 = 5$; среднеквадратическое отклонение $\sigma_1 = 0{,}7$; $\varepsilon \in [0;1]$ – случайная величина с равномерным законом распределения; $I_н = (1\,,3\,,5\,...)$ – множество нечетных чисел меньших k.

Признаки второго класса генерировались с нормальным законом

$$x_v = m + \sigma_2 \left(\sum_{i=1}^{p_2} \varepsilon^i - 0{,}5 p_2 \right) \frac{6}{\sqrt{3p_2}}\,,\ v = \overline{1,k}\,,$$

при $p_2 = 5$, $\sigma_2 = 0{,}7$, $m = 3$.

Априорные вероятности классов принимались равными. Вычислительный эксперимент при фиксированных условиях исследований осуществлялся N раз с использованием контрольной выборки объема 2000, полученной с помощью приведенных выше датчиков случайных чисел. Достоверность различия эмпирических оценок вероятности ошибки распознавания образов сравниваемых методов рассчитывалась в соответствии с критерием Смирнова.

Анализ результатов вычислительного эксперимента показал, что законы распределения оценок вероятностей ошибок классификации сравниваемых алгоритмов распознавания образов являются симметричными. Поэтому средние значения оценок исследуемых показателей эффективности близки к их наиболее вероятным значениям.

Установлено достоверное преимущество предлагаемого алгоритма (5) над традиционным непараметрическим классификатором. Данная закономерность сохраняется для малых объемов обучающих выборок. Обнаружен экстремальный характер зависимости оценки вероятности ошибки непараметрической классификации с использованием системы (5) от количества T частных решающих правил (рис. 1). Причем с ростом размерности k признаков классифицируемых объектов его преимущество при оптимальных значениях T над традиционным непараметрическим алгоритмом возрастает. Отношение средних значений их оценок вероятности ошибки достигает трех на контрольных выборках, что особенно проявляется при малых объемах экспериментальных данных.

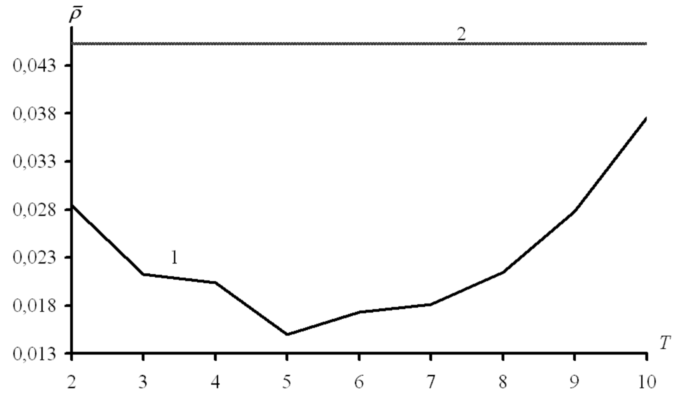

Рис. 1. Зависимость оценки вероятности ошибки распознавания образов от количества частных решающих правил T при n =100, N =50, k =10. Кривые 1, 2 соответствуют нелинейному непараметрическому коллективу и традиционному непараметрическому классификатору.

Методом статистического моделирования исследовались свойства непараметрического коллектива типа (8) в зависимости от размерности k признаков x, объема обучающей выборки n и уровня помех r, накладываемых на значения восстанавливаемой зависимости. Полученные результаты сравнивались со свойствами традиционной непараметрической регрессии [13, 14].

В качестве искомой зависимости $\varphi(x)$ использовались полиномы третьей степени при $k \in [3;12]$. При формировании исходной выборки $V = \left(x^i, y^i, i = \overline{1, n}\right)$ на значения функции накладывалась аддитивная помеха $y^i = \varphi\left(x^i\right)\left(1 + 2r\left(\xi^i - 0.5\right)\right)$, где ξ – случайная величина с равномерным законом распределения на интервале [0; 1].

Эффективность исследуемых непараметрических моделей оценивалась среднеквадратической ошибкой аппроксимации на основе контрольной выборки объема n_k=1000. При конкретных условиях исследования вычислительный эксперимент осуществлялся 10 раз, а полученные оценки ошибки аппроксимации усреднялись.

Анализ результатов вычислительных экспериментов показывает, что предлагаемый непараметрический коллектив моделей (8) обладает более высокими аппроксимационными свойствами в условиях малых выборок по сравнению с традиционной непараметрической регрессией, когда отношение «объем / размерность» обучающей выборки меньше 10 (рис. 2).

С ростом объема обучающих выборок показатели эффективности исследуемой модели (8) и непараметрической регрессии сопоставимы, что согласуется с результатами аналитических исследований. Установлена большая устойчивость непараметрического коллектива (8) при уровне помех $r > 0{,}3$.

Предлагаемая непараметрическая модель, связанная с представлением сложной стохастической зависимости в виде коллектива более простых аппроксимаций, важна в задачах обработки информации значительной размерности.

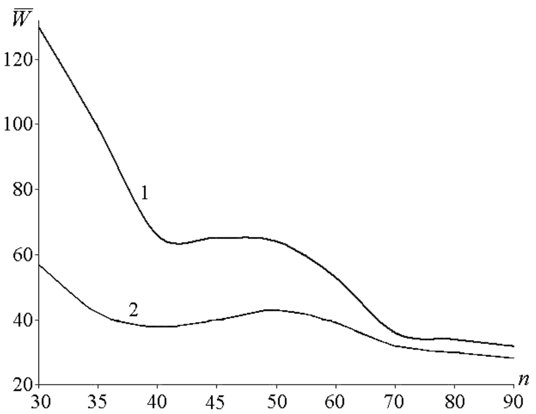

Рис. 2. Зависимость показателей эффективности моделей от объема выборки при $k = 7$ и $r = 0{,}2$. Кривая 1 соответствует многомерной непараметрической регрессии; кривая 2 – нелинейному непараметрическому коллективу (8).

Их применение обеспечивает значительное снижение ошибки аппроксимации по сравнению с традиционной непараметрической регрессией при малых объемах статистических данных.

Заключение

Использование принципов декомпозиции обучающих выборок по их размерности и коллективного оценивания позволяет повысить эффективность решения задач распознавания образов и восстановления стохастических зависимостей в условиях малых выборок. Реализация предлагаемых непараметрических систем открывает возможность использования технологии параллельных вычислений. Обобщение частных решающих функций, синтез которых осуществляется по результатам декомпозиции исходных статистических данных, может быть реализовано на основе их линейных и нелинейных преобразований. В первом случае весовые коэффициенты частных решающих функций определяются путем анализа их показателей эффективности. Во втором случае задача обнаружения скрытых закономерностей переносится в пространство значений частных решающих функций.

Установлено достоверное преимущество нелинейной непараметрической системы классификации над традиционным алгоритмом распознавания образов, что соблюдается при малых объемах обучающих выборок. Обнаружен экстремальный характер зависимости оценки вероятности ошибки распознавания образов от количества частных решающих правил. По сравнению с традиционным непараметрическим классификатором парзеновского типа применение предлагаемой непараметрической системы позволяет (в 1,5– 3 раза) снизить ошибку распознавания образов в условиях малых выборок.

Непараметрический нелинейный коллектив моделей стохастических зависимостей в условиях малых выборок при значениях отношения $n/k < 10$ имеет более высокие аппроксимационные свойства по сравнению с традиционной непараметрической регрессией. С ростом объема обучающих выборок показатели эффективности сравниваемых моделей приближаются к сопоставимым.

Актуальность данного направления определяется необходимостью обработки данных испытаний объектов, которые требуют больших расходов ресурсов при их получении. Дальнейшее развитие предложенного подхода связано с учетом в исследуемых моделях значений показателей эффективности частных решающих функций, которые определяются в процессе синтеза нелинейных непараметрических коллективов.

ЛИТЕРАТУРА

1. *Лапко А.В., Лапко В.А.* Непараметрическая оценка уравнения разделяющей поверхности в условиях больших выборок и ее свойства // Системы управления и информационные технологии. – 2010. – 1.2 (39). – С. 300-304.
2. *Лапко В.А., Капустин А.Н.* Синтез нелинейных непараметрических коллективов решающих правил в задачах распознавания образов // Автометрия. – 2006. – Т.42, №6. – С.26-33.
3. *Лапко А.В., Лапко В.А.* Непараметрические алгоритмы распознавания образов в задаче проверки статистической гипотезы о тождественности двух законов распределения случайных величин // Автометрия. – 2010. – № 6. – С. 47-53.
4. *Лапко А.В., Лапко В.А.* Сравнение эмпирической и теоретической функций распределения случайной величины на основе непараметрического классификатора // Автометрия. – 2012. – Т.48, № 1. – С. 45-49.
5. *Лапко А.В., Лапко В.А., Ярославцев С.Г.* Разработка и исследование гибридных алгоритмов в задачах распознавания образов // Автометрия. – 2006. – Т.42, №1. – С. 32-39.
6. *Лапко А.В., Лапко В.А.* Асимптотические свойства многомерной непараметрической оценки уравнения разделяющей поверхности в двуальтернативной задаче распознавания образов // Системы управления и информационные технологии. – 2010. – 1 (39). – С. 16-19.
7. *Ивахненко А.Г.* Непараметрический комбинаторный алгоритм МГУА на операторах поиска аналогов // Автоматика. – 1990. – № 5. – С. 14-27.
8. *Лапко А.В., Лапко В.А., Соколов М.И., Ченцов С.В.* Непараметрические системы классификации. – Новосибирск: Наука, 2000.
9. *Parzen E.* On estimation of a probability density function and mode // Ann. Math. Statistic. – 1962. – Vol.33. – P. 1065-1076.
10. *Епанечников В.А.* Непараметрическая оценка многомерной плотности вероятности // Теория вероятности и ее применения. – 1969. – Т.14, №1. – С. 156-161.
11. *Мазуров Вл.Д.* Метод комитетов в задачах оптимизации и классификации. – М.: Наука, 1990.
12. *Журавлев Ю.И.* Об алгебраическом подходе к решению задач распознавания образов или классификации // Проблемы кибернетики. – 1978. – №33. – С. 5-68.
13. *Хардле В.* Прикладная непараметрическая регрессия. – М.: Мир, 1993.
14. *Лапко А.В., Лапко В.А.* Анализ асимптотических свойств многомерной непараметрической регрессии // Вестник СибГАУ. – 2012. – 2 (42). – С. 41-44.

E-mail:
Лапко Александр Васильевич – lapko@icm.krasn.ru;
Лапко Василий Александрович – lapko@icm.krasn.ru;
Молоков Вячеслав Витальевич – vvmolokov@mail.ru.