

IV

Навчання та самонавчання розпізнаванню образів

Training and Selftraining Problems in Pattern Recognition

Matrix of Between Clusters Distances Symmetry Criterion in the Algorithm of Binary Perceptron Construction

OLEKSIJ IVAKHNENKO, HRYHORIJ IVAKHNENKO

NAS Institute of Cybernetics

*Ukraine, 252022 Kyiv,
prosp. Akademika Hlushkova 40
Tel.: (044) 267-6743
E-mail: gai@insight.kiev.ua*

Олексій Івахненко, Григорій Івахненко. Критерій симетрії матриці міжкластерних відстаней в алгоритмі побудови перцептрону.

Перший рядок зв'язків перцептрону S-A має відповідати координатам центрів кластерів зображень, що задані у вибірці даних. Щоб одержати кластери та знайти їх центри, застосовується ієрархічний аналіз за допомогою побудови дерев кластеризації. Оптимальними на цих деревах будуть кластеризації, які відповідають мінімуму критерія симетрії матриці міжкластерних відстаней, яка розглядається відносно другої (побічної) діагоналі. Справедливість критерія симетрії доведено методом Монте-Карло на множині задач кластеризації, де оптимальність було визначено за двома критеріями. Оптимальність кластеризації за критерієм симетрії збігається з оптимальністю, визначеній за критерієм балансу змінних [4]. Побудова перцептрону за критерієм балансу дає впевненість у тому, що він залишиться оптимальним не тільки на заданій вибірці, але й на наступних даних.

Другий ряд зв'язків перцептрону S-R здійснює логічне об'єднання центрів оптимальних кластерів зображень, які подано на той чи інший вихід перцептрону. Як приклад, показано побудову перцептрону з використанням критерія симетрії за даними роботи [5].

Introduction. Theory of perceptrons developed by F. Rosenblatt stated that the optimal number of associating (hidden) units A is equal to the number of clusters in clusterisation of input data sample. The optimal weights of first layer links S-A are equal to the coordinates of each cluster [1,2]. Because of that, the first step of the

algorithm considered is to construct the tree of possible clusterisations to find optimal one [3,4]. New symmetry criterion [4]:

$$CR_{\text{symm}} = \sum_1^{\frac{N^2}{2}} (l_{\text{above}} - l_{\text{under}}),$$

where: CR_{symm} - criterion;

N - number of lines in input data sample;

l_{above} - the value of distance between two given lines of data sample, situated under second diagonal of matrix of between lines distances;

l_{under} - the value of distance in symmetrical situated element of the matrix.

It can be used to find the optimal clusterisation among the outputs of clusterisation tree.

In process of tree construction, the number of clusters sequentially decreased. The between lines distances matrix can be calculated for each clusterisation. The optimal result corresponds to the clusterisation, for which the symmetry criterion reaches minimum (zero) value.

For example, if we have received following consequence of distances matrices of clusterisations:

1			
0	2	2	1
1	0	1	1
1	1	0	0
1	1	1	0

2		
0	1	1
1	0	1
1	1	0

3	
0	1
1	0

Symmetry on second diagonal is reached in the second clusterisation, which is optimal one.

Optimal clusterisation, according to its definition, corresponds to physical model of the object of observation. It means, that the balance of variables criterion, calculated for optimal clusterisation, is equal to zero. To calculate the balance criterion, the data sample is to be divided to two equal parts. Two clusterisations trees are constructed for both parts and difference of clusterisations is measured by balance of variables criterion. For optimal clusterisation this criterion is equal to zero $BL=0$.

The optimality of clusterisation choice by symmetry criterion is proved by Monte-Carlo method: a set of clusterisation problems, where solved and for each of them, the optimal clusterisation chosen by symmetry criterion coincides with the optimal clusterisation, chosen by balance of variables criterion [4].

Algorithm of binary perceptron construction, using the symmetry criterion. Input data sample contents $(M+L) \times N$ elements (M - number of binary features, L - number of outputs, showing the classes, N - number of images (lines)). Every element of sample is binary, equal to 0 or 1.

A. Preliminary proceeding of information.

Step 1. Evaluation of features by criterion $F(x_i)$ equal to percentage of coincidence of elements of given column of the sample and output variables. If $F(x_i) < 50\%$, the feature should be changed to its supplement $x_i = -x_i$.

Step 2. Extension of data sample, by calculation of the covariances of features $F(x_{ij}) = x_i \cdot x_j$. If $F(x_{ij}) < 0.5$, the covariance should be changed to its supplement $F(x_{ij}) = 1 - x_i \cdot x_j$.

Step 3. The choice of the about 40% of features, having the bigger values of $F(x_i)$ criterion.

In the case of two classes classification, we get two samples, containing $(M+1) \times N$ elements.

B. First layer of perceptron links S-A construction.

Step 4. Constructing the clusterisation tree for each pattern and calculation of symmetry criterion of each clusterisation. The choice of optimal clusterisation by "left corner rule": optimal clusterisation is obtained, when the symmetry criterion reaches its minimum value. We obtain the definite number of clusters for each pattern.

Step 5. Calculation of middle (center) points of each cluster coordinates r_1, r_2, \dots, r_M . These coordinates are used for outputs of perceptron first layer calculation:

$$u_1 = r_1 x_1 + r_2 x_2 + \dots + r_M x_M.$$

The values of outputs u_1, u_2, \dots, u_S (S - number of clusters) are calculated for each image, presented in the input data sample ($N_1 + N_2 = N$).

	u_1	u_2	...	u_S	
1					1
≡					
N_1					
1					2
≡					
N_2					

C. Second layer of perceptron links A-R construction.

Number of association (hidden) elements is equal to number of clusters. The choice of thresholds, switched after each association unit allowed us to include every cluster to every pattern, which we like. There are two

outputs in the case, when two patterns are distinguished:

$$\varepsilon_1 = w'_1 u_1 + w'_2 u_2 + \dots + w'_S u_S$$

$$\varepsilon_2 = w''_1 u_1 + w''_2 u_2 + \dots + w''_S u_S$$

By choice of weights, we can realize any linear and pieces-linear classification in binary perceptron [2].

In example, the algorithm was shown on the data sample, taken from [5]. The accuracy of clusterisation is higher, than in usual polynomial pattern recognition system, described there.

References

1. Rosenblatt F. Principles of Neurodynamics, Spartan Books, 1962, p. 450.
2. Ивахненко А.Г. Самообучающиеся системы распознавания и автоматического управления, Киев: Техника, 1969, с. 390.
3. Жамбю М. Иерархический кластер-анализ и соответствия, М: Финансы и статистика, 1988, с. 343.
4. Ivakhnenko A.G., Ivakhnenko G.A., Müller J.A. Selforganization of optimal physical data sampling clusterisation for wicked description and undefined objects forecasting. *Pattern recognition and image analysis*, 1993, Vol. 3, No. 3.
5. Круг Г.К., Круг О.О. Математический метод классификации древней керамики, М:Наука, Труды Института археологии, 1965, с.318-325.



Автоматичне навчання розпізнавати образи методом побудови бази багаторівневих алгоритмічних квантів знань (БАКЗ-метод)

ІГОР СІРОДЖА

Авіаційний інститут

Україна, 310070 Харків

вул.Чкалова 17

Тел.: (0572) 44-2734 Факс.: (0572) 44-0046

Ігорь Сироджа. Автоматическое обучение распознаванию образов методом построения базы многоуровневых алгоритмических квантов знаний (МАКЗ-метод).

Предлагаемый МАКЗ-метод автоматического распознавания образов обеспечивает логический вывод требуемого результата в тех случаях, когда существующие способы не эффективны. Это достигается за счет усовершенствования способа извлечения знаний при обучении компьютера построению решающего правила распознавания путем манипулирования специальными алгоритмическими структурами данных, названными k -знаниями (квантами знаний), и построения базы k -знаний (БкЗ) как системы импликативных закономерностей природы исследуемых классов объектов. Импликативной закономерностью r -го ранга при фиксированном числе n признаков объектов называется связь между r признаками ($r < n$), представленная запретом хотя бы одной комбинации их значений. Знания о таких связях неизбежно имеют характер гипотез. Поэтому для их поиска среди n признаков в процессе обучения по выборочным знаниям объема m предметной области доказана теорема об оценке достоверности гипотез о существовании импликативных закономерностей ранга r в зависимости от величины параметров m, n и r . Опираясь на этот результат, обоснован и синтезирован алгоритм автоматического обучения распознаванию образов по выборке k -знаний и различных источников (от экспертов, измерителей, датчиков, из научных данных и т.п.), составляющий основу оператора индуктивного вывода k -знаний и построения БкЗ.

Принципи комп'ютерного навчання розпізнавати образи БАКЗ-методом.
Запропонований БАКЗ-метод автоматичного навчання розпізнавати образи, на відміну від існуючих, ґрунтується на трьох принципах.

1. Уявлення різнотипових даних як багаторівневих алгоритмічних структур - квантів знань (k -знань), що допускають алгебраїчні та логічні перетворення у векторно-матричній і аналітичній формах.
2. Індуктивний принцип пошуку узагальнених k -знань (закономірностей) як бази знань (БкЗ) через аналіз часткових k -знань (фактів).
3. Дедуктивний принцип логічного виводу часткових k -знань (рішень, що приймаються), опираючись на БкЗ.

В роботах [1,2] викладені теоретичні основи БАКЗ-методу, формалізація k -знань та алгоритмічні засоби маніпулювання ними. Реальні об'єкти розпізнавання (ОР) будемо уявляти точками (доменизованими векторами) багатовимірного простору БАКЗ-моделей Ω , яким відповідають елементні кванти знань 1-го рівня $k1Ye$:

$$\begin{aligned} k1Ye &= [\alpha(1,1)\dots\alpha(1,p_1):\alpha(2,1)\dots\alpha(2,p_2)::\dots:\alpha(n,1)\dots\alpha(n,p_n)] = \\ &= [0\dots 1\dots 0:0\dots 1\dots 0::\dots:0\dots 1\dots 0], \end{aligned} \quad (1)$$

де кількість доменів, що виділені двома крапками, дорівнює числу n ознак об'єкта, які приймають значення $\alpha(j,p_j)$, $1 < j < n$, $2 < p_j < p_n$ із скінченних множин з потужностями p_1, p_2, \dots, p_n ; $k1$ -символи, що кодують зміст (семантику) k -знань 1-го рівня; Ye - ім'я кванта, бінарні компоненти доменів якого містять лише по одній

"1" на i -му місці ($1 < i < p_j$), що вказує на єдине зафіксоване значення відповідної ознаки $x(j)$.

Наприклад, деякі свідчення описуються елементарним квантом

$$k1Ye = [010:0001:10:00100] , \quad (2)$$

що має семантику : "досліджуваний ОР характеризується ознаками $x(1)$, $x(2)$, $x(3)$, $x(4)$, що приймають значення із множин відповідних потужностей $||X(1)||=p_1=3$, $||X(2)||=p_2=4$, $||X(3)||=p_3=2$, $||X(4)||=p_4=5$; в даний момент зафіксовано $x(1)=\alpha(1,2)$, $x(2)=\alpha(2,4)$, $x(3)=\alpha(3,1)$, $x(4)=\alpha(4,3)$ як координати точки L -вимірного простору БАКЗ-моделей ($L=p_1 \times p_2 \times p_3 \times p_4 = 3 \times 4 \times 2 \times 5 = 120$)". Аналогічними засобами можна описувати імплікативні закономірності.

Визначення 1. Імплікативною закономірністю (заборонною) рангу r при фіксованому числі n ознак ОР називається такий зв'язок між r ознаками ($r < n$), коли є хоча б одна недопустима (заборонена) комбінація їх значень. Наприклад, структурою (2) тільки заборонним квантом з новим ім'ям $k1_Y$ і змістом описується імплікативна закономірність рангу $r=4$ з семантикою: "не існує об'єкту, у якого ознаки $x(1)$, $x(2)$, $x(3)$, $x(4)$ одночасно визначаються значеннями відповідно (1,2), (2,4), (3,1), (4,3)". В аналітичному вигляді цей зміст можна передати еквівалентними заборонами 4-го рангу у формі предикатних рівнянь :

$$x(1) \text{Щ} x(2) \text{Щ} x(3) \text{Щ} x(4) = 0, \quad x(1) \text{Щ} x(2) \rightarrow \neg x(3) \vee \neg x(4) = 0,$$

$$x(2) \text{Щ} x(4) \rightarrow \neg x(1) \vee \neg x(3) = 0, \quad x(2) \text{Щ} x(3) \text{Щ} x(4) \rightarrow \neg x(1) = 0 .$$

Пояснимо індуктивний принцип щодо навчання розпізнавати образи.

Нехай в результаті експеримента або спостережень, чи свідчень експертів з досліджуваної предметної галузі, об'єктивно охарактеризована лише деяка частина ΣO підмножини ΣD допустимих об'єктів множини $\Sigma \Omega$ всіх можливих об'єктів класу, що вивчається.

Назвемо ΣO множиною навчальних об'єктів (навчальною вибіркою) і приймемо до уваги співвідношення потужностей визначених множин :

$$|\Sigma O| << |\Sigma D| << |\Sigma \Omega| . \quad (4)$$

Очевидно, таке співвідношення має місце і відносно відповідних k -знань 2-го рівня (матричних квантів) $k2\Sigma O$, $k2\Sigma D$, $k2\Sigma \Omega$, якими можна описати вказані множини. Тому загальні імплікативні закономірності у формі бази k -знань (БкЗ) треба шукати на основі надто малочисельних вибірових k -знань $k2\Sigma O$. При цьому, якщо деякі ознаки ОР пов'язані між собою імплікативною залежністю, то існує принаймні одна заборонена комбінація їх значень. Це еквівалентно тому, що в інтервал простору Ω , який відповідає даній комбінації властивостей, не попадає жоден елемент множини ΣD , а отже, жоден з елементів вибірових знань ΣO , бо $\Sigma O \subset \Sigma D$. Тобто, інтервальний заборонний квант є порожньою множиною відносно ΣO . Навпаки, якщо деякий заборонний квант виявився пустим, то це ще не означає, що існує задана ним імплікативна закономірність. Тому висувається гіпотеза про існування імплікативної закономірності, яка відповідає виявленому кванту, що не перетинається з вибіровими знаннями $k2\Sigma O$. Далі встає необхідність оцінки ступеня вірогідності гіпотези величиною ймовірності того, що вона помилкова.

Постановка задачі навчання. Взагалі задача полягає у створенні алгоритму індуктивного виводу системи імплікативних закономірностей як бази k -знань про узагальнену модель досліджуваного середовища, посиляючись на вибіркові k -знання $k\Sigma O$.

Нехай задано: стандартний вибірковий квант $k\Sigma O$, що містить m елементних квантів типу $k1Y_e(1)$ з рівномірним законом розподілу відносно n ознак об'єктів розпізнавання, максимально допустимий ранг R_m імплікативних закономірностей і структуру заборонних k -знань вигляду $k1Y(3)$. Необхідно синтезувати алгоритм індуктивного виводу $Bk3$ у вигляді кванту $k\Sigma_{\Sigma m}$, який складається із заздалегідь невідомого числа qz заборон виду $k1Y$ рангу не вище R_m , що задовольняє умовам заданого порогу вірогідності гіпотез і співвідношенням

$$k1Y \cap k\Sigma O = \emptyset, \quad k1Y \cap k\Sigma D = \emptyset, \quad k\Sigma O \subseteq k\Sigma D. \quad (5)$$

Ім'я $\Sigma_{\Sigma m}$ кванту $Bk3$ символізує її мінімізацію з метою отримання простих заборон у розумінні імплікант, що не виводяться одна з іншої.

Алгоритм навчання і його обґрунтування. Для синтезу алгоритму індуктивного виводу $Bk3$ в процесі навчання оцінимо вірогідність гіпотези про існування імплікативної залежності між деякими ознаками, що відповідає наявності порожнього інтервалу в підпросторі вибіркових k -знань $k\Sigma_{\Sigma m} \subseteq k\Sigma D$. Вважаємо гіпотезу помилковою у тому випадку, коли в заборонному інтервалі, описаному квантом $k1Y$ рангу r , з'являться деякі елементи допустимих знань $k\Sigma D$, які випадково не попали в навчальну вибірку $k1\Sigma O$. Тоді потрібну вірогідність гіпотези обчислимо як ймовірність P_s події $S[n, m, r]$: "довільний заборонний квант $k1Y$ r -го рангу не перетинається із заданим навчальним квантом $k\Sigma O$ розмірами $m \times n$, бо в дійсності не існує зв'язку між зафіксованими ознаками".

Теорема. Ймовірність $P_s(m, n, r)$ події S , яка стосується існування заборонного кванта $k1Y$ r -го рангу відносно навчального кванта $k\Sigma O$ розмірами $m \times n$, оцінюється величиною математичного сподівання $M_s\{m, n, r\}$ числа заборонних квантів r -го ранга, що не перетинаються з випадковою вибіркою $k\Sigma O$, за формулою

$$P_s(m, n, r) \leq M_s\{m, n, r\} = \sum_{\alpha \in A} a_{\alpha} \left(\frac{a_{\alpha} - 1}{a_{\alpha}} \right)^m \quad (6)$$

де α - індекс номера комбінації значень ознак OP з множини індексів A потужністю $|A|$:

$$|A| = C_n^r = \frac{n!}{r!(n-r)!}; \quad a_{\alpha} = \prod_{i=1}^r \rho_i(\alpha);$$

$\rho_i(\alpha)$ - кількість значень i -ої ознаки у відповідному домені кванта.

Згідно цієї теореми гіпотеза про існування імплікативної закономірності $k1Y$ в ΣD , судячи по ΣO , приймається, якщо оцінка вірогідності

$$M_s\{m, n, r\} = \sum_{\alpha \in A} a_{\alpha} \left(\frac{a_{\alpha} - 1}{a_{\alpha}} \right)^m \leq M_s^* \quad (7)$$

при достатньо малому допустимому значенні M_s^* . Вибір порогу M_s^* не критичний через значну залежність M_s від рангу r , про що свідчать розрахунки

в таблицях [1]. Таким чином, при заданих m, n і M_s^* за формулою (7) обчислюється максимально допустиме значення рангу R_m імплікативних закономірностей, що дозволяє знайти їх кількість qz для бінарних ознак за формулою

$$qz = \sum_{r=1}^{R_m} \binom{n}{r} 2^n. \quad (8)$$

Наведені вище викладки зумовлюють наступний алгоритм розв'язку поставленої задачі комп'ютерного навчання [1]:

1. Визначення максимально допустимого значення R_m рангу і кількості qz -шуканих імплікативних закономірностей за відомими m, n, M_s^* , використовуючи формули (6)-(8).
2. Виявлення відсутніх заборонних квантів типу $k_1 \uparrow Y$ різних рангів не вище R_m і формування матричного кванта заборон 2-го рівня $k_2 \uparrow \Sigma_v$ в об'ємі qz .
3. Перетворення кванта $k_2 \uparrow \Sigma_v$ засобами мінімізації [1] в Bk_3 , що складається з мінімальної системи простих заборонних квантів $k_2 \uparrow \Sigma_{vm}$.

Наведений алгоритм складає основу ефективного оператора індуктивного виводу k -знань, який успішно пройшов випробування і працює в діючій знанняорієнтованій системі прийняття рішень.

Література

1. Сироджа И.Б. Математическое и программне обеспечение интеллектуальных компьютерных систем. - Харьков: ХАИ, 1993, 100 с.
2. Сироджа І.Б., Прудников Г.В., Резниченко О.В. Нова інформаційна технологія розпізнавання образів і прийняття рішень, що ґрунтується на використанні багаторівневих алгоритмічних квантів знань (БАКЗ-метод). - Київ, Доповіді Всеукраїнської міжнародної конференції з обробки сигналів і зображень та розпізнавання образів "УКРОБРАЗ-94" 1994.



Application of the Reduction Theory to Restore Membership Functions in Fuzzy Sets

VOLODYMYR VASYL'JEV, VALERIJ SUSHKO

Інститут кібернетики НАН

*Україна, 252022 Київ-22,
проспект Академіка Глушкова 40
Тел.: (044) 266-2204*

When solving pattern recognition problems the most important part of research is to find numerical estimate and to measure some qualitative attributes of objects. This is especially true where objects could not be described by measurements directly.

In the basic work on fuzzy set theory [1] it is proposed to describe objects by use of the so-called linguistic variables (LV), that is ones having as values not quantities, but words of a natural language. The possibility to use words, but not quantities, attracts our attention because the linguistic description is less specific than numerical one, and so words „a young man“ are less specific than words „a man of 25“. Here the linguistic value of variable the „age“ plays the same part as numerical value 25,

but is less precise, more blurred and so less informative. The LV the „appearance“ provides still more blurred notion. This variable can not be measured at all, and can not be presented as a function of one or another physical value. Of course, each man can be assigned a grade of membership (GM) into the class of „attractive ones“, but this characteristic will be subjective. And here the GM is defined not through a set of strictly determined objects, but through a set of impressions being represented by some symbols. In other words, certain calculations are performed „behind the scenes“ [1], and thereafter a linguistic approximation is used to map quantities into words. So the quantities still exist, but they are hidden from the observer, and the problem is to find inverse mapping - from words into quantities. The idea of membership differs radically from that of probability. So if a GM of the quantity „25 years“ in the notion the „age“ equals to 0.4, it does not mean at all that the value of the LV the „age“ equals to 25 with probability 0.4.

Most commonly scales, on ground of which the membership functions (MF) are based, are alternative by their nature. So if LV the „age“ is plotted along the time axes, the value of the MF of LF „young“ decreases with time, whereas the value of the „old“ increases. The same is true for pairs „attractive“-„ugly“, „clever“-„stupid“ etc. In each case there is a „neutral“ zone, equivalent to the intersection zone (for continuous spaces) in so-called-method algorithm of learning [2]. The algorithm is advanced in the context of limiting reduction method (LMR). Thus if direct and alternative MFs are anyway got, it would be possible to use fuzzy variables in the context of LRM to search such subspaces of features, wherein there exist solutions that insure quality and reliability of pattern recognition being not worse than preset values.

In the report a new method of MF restoration is proposed, based on search of such metric features, in the space of which interrelations, having been determined during previous learning, between objects remain invariant. Here the pattern recognition learning (PRL) falls into four stages:

1. The trainer or the expert sets an order ratio between learning sequence objects by each LV corresponding to a certain notion;
2. Using a procedure to LRM-learning, such metric spaces are formed, in which presetted order ratios between objects by each LV hold;
3. In the chosen metric space distances are calculated of all objects from the reference one of this notion (having got a maximal estimate during the first stage of learning); the distance determines a GM of each object in the notion;
4. The obtained MF are used in the sequel at the PRL stage, the patterns may be specified blurredly.

The proposed procedure, like all algorithms based on the reduction theory and the LRM, insures preset quality and reliability of pattern recognition on new objects.

References

1. L.A.Zadeh. The Concept of Linguistic Variable and Its Application to Approximate Reasoning. Am. Elsevier PC, NY, 1973.
2. V.I. Vasil'yev, F.P. Ovsyannikova, V.I. Sushko, D.Ye. Skvaletskiy: Synthesis of Continuous Spaces That Guarantee the Separation of Patterns with a Specified Reliability. Soviet journal of Automation and Information Sciences, v.23, N6, 1990, pp. 15-20.

Визначення оптимальної множини ознак в дискримінантному аналізі в умовах нерівних коваріаційних матриць (схема з поділом спостережень на навчальні та перевірні підвибірки)

ОЛЕКСАНДР САРИЧЕВ

Інститут кібернетики НАН

*Україна, 252022 Київ
проспект Академіка Глушкова, 40
Тел.: (044) 267-6743*

Oleksandr Sarychev. The Optimal Set Features Determination in Discriminant Analysis under the Conditions of Nonequal Covariance Matrices (Scheme with Dividing of Observations into Training and Checking Subsamples).

In article, it is shown analytically that the scheme with dividing of observations into the training and checking subsamples allow to solve the problem of discriminant analysis in the broad sense (with unknown set of features). It is shown that there exists an optimal set of features corresponding to the maximum mathematical expectation of some generalized distance between the observations from two general sets. It is shown analytically that parameters of the general sets and samples sizes influence on a complexity of optimal discriminant function, and conditions under which the optimal discriminant function is simplified, are exhibited. These conditions depend on parameters of the general sets and samples sizes.

Постановка задачі дискримінантного аналізу в широкому розумінні. Нехай $(m * n_k)$ — матриця X_k є вибіркою n_k незалежних спостережень m -вимірного випадкового вектора η_k , підпорядкованого m -вимірному нормальному розподілу з невідомим математичним сподіванням χ_k та невідомою коваріаційною матрицею Σ_{kX} , де $k=I, II$ — номери двох незалежних генеральних сукупностей P_k , причому $\chi_I \neq \chi_{II}$ та $\Sigma_{IX} \neq \Sigma_{IIX}$. Нехай X — множина всіх компонент векторів η_k , над якими проведено спостереження. Якщо апріорно невідомо, які саме компоненти з множини X треба включати в дискримінантну функцію, то кажуть про задачу дискримінантного аналізу в широкому розумінні.

Схема порівняння дискримінантних функцій, що ґрунтується на поділі спостережень на навчальні та перевірні підвибірки. Нехай V — поточна множина компонент, що аналізується. Нехай множині V відповідають: V_I і V_{II} — матриці спостережень з сукупностей P_I і P_{II} , v_I і v_{II} — вектори математичних сподівань та Σ_{IV} і Σ_{IIV} — коваріаційні матриці. Нехай поділ на навчальні (A) та перевірні (B) підвибірки деяким способом проведено. Розглядаємо відстань:

$$\hat{D}_{AB}^2(V) = \frac{\hat{d}_A^T (\tilde{v}_{IB} - \tilde{v}_{IIB}) (\tilde{v}_{IB} - \tilde{v}_{IIB}) \hat{d}_A}{\hat{d}_A^T S_{IB} \hat{d}_A + \hat{d}_A^T S_{IIB} \hat{d}_A} \quad (1)$$

В (1) вектор \hat{d}_A — обчислена на підвибірці A оцінка коефіцієнтів дискримінантної функції: $\hat{d}_A = S_A^{-1} * (\tilde{v}_{IA} - \tilde{v}_{IIA})$, де вектори \tilde{v}_{IA} і \tilde{v}_{IIA} — оцінки v_I і v_{II} ; $S_A = (S_{IA} + S_{IIA}) / 2$, де S_{kA} — незсунена оцінка Σ_{kV} . В (1) вектори

$\tilde{v}_{IВ}$ і $\tilde{v}_{IIВ}$ — оцінки v_I і v_{II} , а матриця $S_{kВ}$ — оцінка Σ_{kV} , обчислені на підвибірці B .

Означення 1. Оптимальною множиною компонент (множиною ознак) будемо називати таку множину компонент $V_{opt} \subseteq X$, для якої виконується:

$$V_{opt} = \arg \max_{V \subseteq X} E \{ D_{AB}^2(V) \},$$

де $E\{\bullet\}$ — знак математичного сподівання.

Означення 2. Оптимальною за кількістю та складом компонент будемо називати фішерівську дискримінантну функцію, побудовану на множині компонент V_{opt} .

В доповіді, по-перше, доведено існування оптимальної множини ознак, що відповідає максимуму математичного сподівання згадуваної статистики; по-друге, сформульовано умови, за яких дискримінантна функція, оптимальна за кількістю та складом ознак, що включаються до неї, спрощується в залежності: а) від параметрів генеральних сукупностей χ_k і Σ_{kX} ; б) від обсягу вибірок спостережень n_k .



Розпізнавання образів у задачах структурної самоорганізації інформаційної бази навчальних систем

АНАТОЛІЙ ТИМЧЕНКО, ЮРІЙ ТЕСЛЯ, ЛЮБОВ ОКСАМИТНА

Інженерно-технологічний інститут

*Україна, 257006 Черкаси,
вул.Шевченка,460
Тел.: (047) 435-628*

Анатолий Тимченко, Юрий Тесля, Любовь Оксамитная. Распознавание образов в задачах структурной самоорганизации обучающих систем.

Доклад посвящен актуальным вопросам построения обучающих систем на основе структурной самоорганизации моделей предметных областей. В основе такой самоорганизации лежит распознавание существующих между элементами обучающей среды отношений и формирование в информационной базе обучающей системы множества связей адекватных структуре обучаемой дисциплины. Предложен метод использования коротких выборок в процессе распознавания системы отношений предметной области.

В основі сучасних підходів до розв'язування задач розпізнавання образів лежить побудова інформаційних баз розпізнавальних систем, що адекватно відображають предметну область задач як в декларативному, так і в процедурному наповненні. Найбільш ефективним способом побудови інформаційної бази є навчання розпізнаванню образів. Формулюється задача структурної самоорганізації інформаційної бази навчальних систем на основі розпізнавання системи відношень в заданій предметній області /1/.

В безпосередньо навчальних системах той, хто навчається, одержує знання, навички та вміння в деякій предметній області під управлінням навчальної програми; при цьому програма приймає на себе всі функції викладача по організації представлення навчального матеріалу, контролю його засвоєння та виявленню помилок того, хто навчається. Для подолання недоліків, властивих аналогічним системам, пропонується аткий підхід. В основі створюваної системи знаходиться інформаційна база предмету навчання (контролю) та алгоритм, який реалізує процес самоорганізації структурних зв'язків інформаційної бази. Процес самоорганізації здійснюється на основі відображення моделі опитування на інформаційну базу системи. Самоорганізація являє собою процес побудови системи зв'язків інформаційної бази в залежності від спрямованості навчання. Таким чином, структура зв'язків інформаційної бази попередньо не готується, а створюється в процесі розпізнавання структурних зв'язків в предметній області дисципліни. Це дозволяє значно зменшити трудомісткість побудови структури інформаційної бази по кожній із дисциплін, при необхідності легко пристосувати її до нових умов функціонування. Структура інформаційної бази системи, в цьому випадку, повинна відображати як відношення між елементами предметної області (каузальні, входження та інші), так і забезпечувати формування методики навчання на основі відображення рівня знань того, хто навчається, на структурні зв'язки бази навчання.

В режимі навчання автоматизованої системи відбувається автоматична корекція ваги зв'язків G інформаційної бази, обґрунтованих на апріорній інформації a і на одержаній імовірнісній характеристиці взаємозв'язку елементів інформаційної бази $P/2/$: $G=f(a,P)$, де a - апріорна інформація про зв'язки між елементами інформаційної бази; P - імовірнісна характеристика зв'язків.

Імовірнісні характеристики взаємозв'язку елементів визначаються на основі моделювання порядку опитування контролюючих в режимі навчання навчальної системи. Основною для визначення ваги зв'язку служить відхилення величини G від абсолютної імовірності $P_j=P(V_j)$. Одержана структура зв'язків інформаційної бази є відображенням реальної взаємодії запитань і тем по дисципліні, що дозволяє на її основі створити модель опитування або навчання студента. Розв'язок проблеми, пов'язаної з інформаційною неповнотою коротких виборок при обмеженому часі навчання і значній кількості зв'язків, інформаційної бази системи розв'язується шляхом підбору характеристичних значень нормального розподілу імовірностей відхилення ваги зв'язків від нульового значення.

Запропонований підхід до розробки навчальних систем дозволяє створювати на одній і тій же програмній основі інформаційні бази різноматних дисциплін, гнучко змінювати їх структуру в залежності від поставлених задач. Поведінка таких систем при виборі контрольного набору запитань максимально наближується до поведінки викладача. Широке впровадження таких систем дозволить організувати процеси автоматизованого контролю знань та навчання з мінімізацією затрат на побудову інформаційної бази навчальних систем.

Література

1. В.И.Васильев. Распознающие системы. Справочник. К.:Наукова думка, 1983.-421 с.
2. Ю.Н. Тесля. Самоорганизация структурных связей в системах контроля знаний и обучения. - Программир. обучение, 1984, вып.21, с.91-94.

