Determination of Informative Features Based on the Criteria of Positive Type

Narzillo Mamatov, Abdurashid Samijonov, Nilufar Niyozmatova, Erkin Rahmonov, Bakhromjon Abdurakhmonov

Abstract*: Currently, pattern recognition systems are widely used to solve practical problems. The choice of criterion in pattern recognition problems is fundamental. The choice of informative features depends on the criterion of informativeness. A single criterion has not been developed to select a set of informative features. Therefore, the criterion is selected from the formulation of a practical problem. For each selected criterion, it is necessary to develop special methods and algorithms. For homogeneous criteria, there is no single method.A new method and algorithm for determining sets of informative features based on generalized criteria of a positive order is proposed in this work.

***Keywords*: feature, positive-types criteria, informative feature, vector, functional, gradient, operator.**

I. INTRODUCTION

Currently, pattern recognition systems are widely used to solve practical problems. The choice of criterion in pattern recognition problems is fundamental. The choice of informative features depends on the criterion of informativeness. A single criterion has not been developed to select a set of informative features. Therefore, the criterion is selected from the formulation of a practical problem. For each selected criterion, it is necessary to develop special methods and algorithms. For homogeneous criteria, there is no single method. The main problem in pattern recognition problems is the choice of informative features and the effectiveness of its solution largely depends on the criterion used for the content. A single criterion has not been developed to select a set of informative features. Therefore, the criterion is selected from the formulation of a practical problem. Currently, more than 30 homogeneous criteria with zero order have been developed, but in many practical tasks the task of selection informative features requires the use of homogeneous criteria with a positive degree.

The article suggests a "Gradient" method of choosing a set of informative features using a uniform criterion with a positive degree.

Using the concepts and notation introduced in [1,2], we choose a criterion in the following form:

$$\Re(\eta)=\frac{\prod_{j=1}^{t}\left(a^{(j)},\eta\right)}{(b,\eta)} \qquad (1)$$

where, a,b and η - are n dimensional vectors, $\eta_i\in\{0,1\}, i=\overline{1,n}, \; t\in N(t>1)$,

Functional (1) expresses a generalized form of a homogeneous functional with a positive degree of the Fisher type.

II. STATEMENT AND SOLUTION OF THE PROBLEM

Consider the following task of finding informative subsystems of signs:

$$\begin{cases}\Re(\eta)\to\max\\ \eta\in\Lambda^{1},\end{cases} \qquad (2)$$

where Λ^{l} is the set of all l informative features.

To simplify the calculation, we introduce the following notation:

$$A^{(j)}=\left(a^{(j)},\eta\right),\; B=(b,\eta),\; j=\overline{1,t},$$

$$A_1^{(j)}=\left(a^{(j)},\mu\right),\; B_1=(b,\mu).$$

To solve problem (2), we introduce the vector $D=(d_1,d_2,...,d_N)$ directing to the optimal solution. This vector is called the gradient of functional (1) and its components are defined as:

$$d_i=\frac{\sum_{s=1}^{t}a_i^{(s)}}{w^{t-1}}-\Re_\eta b_i,\; i=\overline{1,N}$$

where $w=\frac{N}{\min\limits_{a_j^{(i)}\neq 0}a_j^{(i)}}$, $\Re_\eta=\Re(\eta)=\frac{\prod_{i=1}^{t}A^{(i)}}{B}$;

Theorem 1. If η and μ are l informative vectors and $b_s>0,\; s=\overline{1,N}$, then $\Re(\eta)<\Re(\mu)$ if and only if $(D,\mu)>0$.

Proof.

Let be $(D,\mu)>0$.

Revised Manuscript Received on January 30, 2020.
* Correspondence Author

Mamatov Narzillo*, Tashkent University Information Technologies named after Al-Kharezmi, Tashkent, Uzbekistan.

Niyozmatova Nilufar, Tashkent University Information Technologies named after Al-Kharezmi, Tashkent, Uzbekistan.

Samijonov Abdurashid, Bauman Moscow State Technical University, Moscow, Russia Federation.

Rahmonov Erkin, Tashkent Pharmaceutical Institute, Tashkent, Uzbekistan.

Abdurakhmonov Bakhromjon, Tashkent Pharmaceutical Institute, Tashkent, Uzbekistan.

Then $(\mathrm{D},\mu)=\dfrac{\sum_{i=1}^{t}\left(a^{(i)},\mu\right)}{w^{t-1}}-\Re_{\eta}(b,\mu)>0$ or

$$\frac{\sum_{i=1}^{t}A_{1}^{(i)}}{w^{r-1}}>\Re_{\eta}B_{1} \qquad (3)$$

The left side of inequality (3) is expressed as follows:

$$\frac{\sum_{j=1}^{t}A_{1}^{(j)}}{w^{t-1}}=\frac{\sum_{j=1}^{t}wA_{1}^{(j)}}{w^{t}} \qquad (4)$$

Given $\sum_{j=1}^{t}wA_{1}^{(j)}\geq N$, we have

$$\sum_{j=1}^{t}wA_{1}^{(j)}\leq w^{t}\prod_{j=1}^{t}A_{1}^{(j)} \qquad (5)$$

From (5) and (3), obtained the following (6)

$$\prod_{j=1}^{t}A_{1}^{(j)}\geq\frac{\sum_{j=1}^{t}wA_{1}^{(j)}}{w^{t}}>\Re_{\eta}B_{1} \qquad (6)$$

We reduce inequality (6) to the following form:

$$\prod_{j=1}^{t}A_{1}^{(j)}>\Re_{\eta}\cdot B_{1}\Rightarrow\frac{\prod_{j=1}^{t}A_{1}^{(j)}}{B_{1}}>\Re_{\eta}=\frac{\prod_{j=1}^{t}A^{(j)}}{B}\Rightarrow\Re(\mu)>\Re(\eta)$$

.

Thus, the theorem is proved.

We introduce the operator (of succession) $\mu:\Lambda^{1}\rightarrow\Lambda^{1}$, such that

$$\left(\mathrm{D},\mu(\eta)\right)=\max_{\nu\in\Lambda^{1}}(D,\nu).$$

The operator μ can be written in the following constructive form. By ordering the components of the vector D , a set of pairwise different indices $j_{1},j_{2},\dots,j_{N}$ is determined, such that $d^{j_{1}}\geq d^{j_{2}}\geq\dots\geq d^{j_{N}}$, then the components of the vector $\mu(\eta)$ will be defined as

$$\mu^{j_{1}}(\eta)=1,\mu^{j_{2}}(\eta)=1,\dots,\mu^{j_{l}}(\eta)=1,$$
$$\mu^{j_{l+1}}(\eta)=0,\mu^{j_{l+2}}(\eta)=0,\dots,\mu^{j_{N}}(\eta)=0.$$

Moreover, the components of the vector $\mu(\eta)$, the initial l are the maximum components of the vector D are equal to one, the rest are zero.

Based on the stated Theorems 1, the proposed method for maximizing functional (1) is based, which is implemented as an iterative procedure. Moreover, at the first step, an arbitrary l is choosed - an informative vector η , for example,

$$\eta=\begin{pmatrix}6\ 7\ 8\\ 1,1,\dots,1,0,0,\dots,0\end{pmatrix}.$$

Then, in each iteration, the new vector η is calculated from the previous one using the operator $\mu(\eta)$, that is, assigning $\eta=\mu(\eta)$.

The iteration continues until the $\Re(\eta)$ grows. In case growth stops, i.e. $\Re(\eta)=\Re(\mu(\eta))$, then η is the optimal solution.

III. RESULT AND DISCUSSION

The recognition task is carried out on the basis of a pre-formed sample. The sample consists of 3 classes, each class contains 20 objects and the size of the attribute space is 15. The task of classifying objects is carried out using informative signs based on the k-nearest neighbor method. To select informative features, the criteria of "-1" and "-2" order were used. The tables show the corresponding results of the selection of informative features according to the criteria.

Table 1

N	l	Number of objects			Informative features (1)															$\max I(\lambda)$	k	Stability of objects
		class 1	class 2	class 3																		
15	1	20	20	20	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	15,68	7	81
15	2	20	20	20	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	14,92	35	83
15	3	20	20	20	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	13,72	27	89
15	4	20	20	20	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	11,86	3	91
15	5	20	20	20	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	10,55	3	91
15	6	20	20	20	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	9,11	3	91
15	7	20	20	20	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	7,95	3	91
15	8	20	20	20	1	1	1	1	1	1	0	0	0	1	0	0	0	0	1	7,04	3	92
15	9	20	20	20	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	6,29	31	91
15	10	20	20	20	1	1	1	1	1	1	0	0	1	1	1	0	0	0	1	5,69	37	100
15	11	20	20	20	1	1	1	1	1	1	0	1	1	1	1	0	0	0	1	5,14	37	100
15	12	20	20	20	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	4,69	37	100
15	13	20	20	20	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	4,30	37	100
15	14	20	20	20	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	3,96	37	100
15	15	20	20	20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3,67	37	100

Table 2

N	l	Number of objects			Informative features (1)															$\max I(\lambda)$	k	Stability of objects
		class 1	class 2	class 3																		
15	1	20	20	20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	5,3E-6	21	85
15	2	20	20	20	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1,1E-6	13	88
15	3	20	20	20	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	4,5E-7	25	100
15	4	20	20	20	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	2,3E-7	3	92
15	5	20	20	20	1	0	1	0	1	1	0	0	0	0	0	0	0	0	1	1,2E-7	7	94
15	6	20	20	20	1	1	1	0	1	1	0	0	0	0	0	0	0	0	1	7,7E-8	15	95
15	7	20	20	20	1	1	1	0	1	1	0	0	0	1	0	0	0	0	1	4,9E-8	3	91
15	8	20	20	20	1	1	1	0	1	1	0	0	1	1	0	0	0	0	1	3,3E-8	3	93
15	9	20	20	20	1	1	1	0	1	1	0	0	1	1	1	0	0	0	1	2,3E-8	27	92
15	10	20	20	20	1	1	1	1	1	1	0	0	1	1	1	0	0	0	1	1,7E-8	37	100
15	11	20	20	20	1	1	1	1	1	1	0	1	1	1	1	0	0	0	1	1,2E-8	37	100
15	12	20	20	20	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	9,5E-9	37	100
15	13	20	20	20	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	7,3E-9	37	100
15	14	20	20	20	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	5,7E-9	37	100
15	15	20	20	20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4,5E-9	37	100

Experimental studies of the proposed algorithm for the choosing of informative features using the example of solving the model problem showed that the results of the selection of informative features obtained using this algorithm coincide with similar results obtained using the exhaustive search method.

IV. CONCLUSION

In the work, methods for the formation of informative features for homogeneous criteria by the Fisher type positive order are developed. A software package designed to solve the problems of determining a set of informative features and pattern recognition has also been developed.

Retrieval Number: B6369129219/2020©BEIESP
DOI: 10.35940/ijitee.B6369.019320
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

It should be noted that the use of heterogeneous criteria for the formation of informative features in pattern recognition is considered for the first time.

REFERENCES

1. Mamatov N, Samijonov A, and Yuldashev Z, "Selection of features based on relationships," J. Phys. Conf. Ser., vol. 1260, no. 10, p. 102008, 2019.
2. Fazilov Sh and Mamatov N, "Formation an informative description of recognizable objects," in Journal of Physics: Conference Series, 2019.
3. Mamatov N.S. Heuristic criteria for the informativeness of signs. Materials of the XVII International Scientific and Methodical Conference "Informatics: Problems, Methodology, Technologies". -Voronezh. 9-10 February 2017. Volume 3, P.114-120.
4. Mamatov N, "Selection of informative features using heuristic criteria of fisher type," BEST Int. J. Manag. Inf. Technol. Eng. (BEST IJMITE), 2016.
5. Ayvazyan S.A., Buchstaber V.M., Enyukov I.S., Meshalkin L.D. Applied statistics: Classification and reduction of dimension. - M .: Finance and Statistics, 1989. - 607 p.
6. Gorelik A.L., Skripkin V.A. Building recognition systems. - M .: Sov.radio, 1974. - 224 p.
7. Zhuravlev Yu.I., Kamilov M.M. Tulyaganov Sh.E. Algorithms for calculating estimates and their application. - Tashkent: Fan, 1974. - 124 p.
8. Kutin G.I. Methods of ranking complexes of features. Review // Foreign Radioelectronics. -1981, -№9. -WITH. 54-70.
9. Fazilov Sh and Mamatov N 2019 IOP Conf. Series: Journal of Physics: Conf. Series 1210 (2019) 012043 doi:10.1088/1742-6596/1210/1/012043
10. Methods, criteria and algorithms used in the conversion, selection and selection of features in data analysis / K.A.Cheponis, D.A. articles. - Vilnius, 1988. - 150 p.