Pattern Recognition Based on Classes Distinctive Features

Viktor Krasnoproshin
Faculty of Applied Mathematics and Computer Science
Belarusian State University
Minsk, Belarus
krasnoproshin@bsu.by

Vadim Rodchenko
Faculty of Mathematics and Informatics, Yanka Kupala State University of Grodno
Grodno, Belarus
rovar@grsu.by

Anna Karkanitsa
Faculty of Mathematics and Informatics, Yanka Kupala State University of Grodno
Grodno, Belarus
a.karkanica@grsu.by

Abstract. **In pattern recognition, the approach where Supervised Learning is reduced to the construction of decision rules is considered to be classical. These rules should ensure an extremum of some criterion. The paper proposes an alternative solution based on the search for combinations of features that ensure classes separation. The results of a numerical experiment on model data confirm the effectiveness of the proposed approach.**

Keywords: **pattern recognition, classification, instance-based learning, discovering patterns**

I. Introduction

Pattern recognition is one of the most important fundamental problems in computer science. The progress in developing artificial intelligence technologies largely depends on theoretical achievements in this area [1–4].

According to classical definition, pattern recognition is the assignment of initial data to a certain class. This is met by the selection of significant distinctive features from the general data set [5].

The growing computing power, progress in big data sets availability and the accumulated practical experience have led to a new perspective on the pattern recognition problem. In particular, C.M. Bishop writes in his monograph: “the field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions such as classifying the data into different categories” [6].

We propose to refrain from the traditional approach when training is focused on building a classification algorithm. An alternative approach is proposed in the paper. The combinations of features that ensure classes distinction are identified in the learning process.

The effectiveness of the approach is confirmed by the results of a numerical experiment on model data.

II. About Statement of the Recognition Problem

The classical description of the recognition (classification) problem is as follows:

Let the set of objects is given. The set is divided into subsets called classes. Partial information about classes, descriptions of the entire set and an object are given. Does object belong to a certain class is unknown. It is required to establish object’s belonging to one of the classes according to the given information [7].

Currently, this problem is traditionally being solved in the following statement:

Let the objects’ descriptions X and the acceptable answers Y for objects classification are given. It is assumed that there is an unknown target dependency $y^* : X \rightarrow Y$, *which values* $X^m=\{(x_1, y_1),\dots, (x_m, y_m)\}$ *are known only for a finite set of objects (training set objects).*

It is required to construct an algorithm $a : X \rightarrow Y$, *that would approximate this target dependency on the entire set X* [8].

To solve such a problem, a family of algorithms is firstly set (up to parameters). Then, in the learning process (Supervised Learning), such parameters values are found that provide the extremum of the selected criterion.

Practical experience of usage the described scenario has revealed a number of serious shortcomings:

1) Selecting a family (model) of algorithms is a nontrivial problem. So, we are dealing not so much with a science as with the art of building algorithms;

2) Learning can be implemented only in an automated mode. The resulting algorithm a: $X \rightarrow Y$ is a «black box» approximating the unknown target dependency. The obtained dependency can hardly if any be interpreted in terms of the subject domain;

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

3) Finding a solution (based on the training set X^m) is carried out only in the initial space of object description. The question of the existence of subspaces in which the problem is solved more efficiently remains open [9].

To overcome the mentioned disadvantages an alternative approach is proposed. It is based on the hypothesis of compactness: «in the objects space, classes form compactly localized subsets» [10, 11].

In this case, the mathematical statement of recognition problem can be formulated as follows:

Let the objects' descriptions X and the acceptable answers of objects classification Y are given. There is an unknown target dependency $y^*: X \to Y$, which values $X^m=\{(x_1, y_1),\dots,(x_m, y_m)\}$ are only known for the training set objects.

It is required to find feature subspaces where class patterns do not intersect. Based on these patterns it is necessary to construct an algorithm $a: X \to Y$ that would approximate the target dependency on the entire set X.

To solve the recognition problem in such statement, it is firstly proposed to find feature spaces in which the classes do not intersect. After that, the construction of the classification algorithm becomes a trivial procedure.

III. Supervised Learning Algorithm

Let the training set $X^m=\{(x_1, y_1),\dots,(x_m, y_m)\}$ is formed on the basis of an a priory dictionary of features $F=\{f_1, \dots, f_n\}$.

Let $V=\{v_1, \dots, v_q\}$ denote the set of all possible combinations of features from F. Then V contains $q=2^n-1$ subsets.

The searching algorithm for combinations of features on the set $V=\{v_1, \dots, v_q\}$ for which the class patterns do not intersect is as follows.

Algorithm

Step 1. Select from V a subset $V^+=\{v^+_1, \dots, v^+_n\}$, where v^+_i contains only one feature.

Step 2. For each v^+_i build the class patterns and compare their mutual placement.

Step 3. If the class patterns do not intersect, then the feature v^+_i is included in the set $V^*=\{v^*_1, \dots, v^*_k\}$.

Step 4. Exclude the subset $V^+=\{v^+_1, \dots, v^+_n\}$ from the set $V=\{v_1, \dots, v_q\}$ and obtain $V^\wedge=\{v^\wedge_1, \dots, v^\wedge_p\}$.

Step 5. Exclude from $V^\wedge$ all combinations $v^\wedge_i$that contain any combination from $V^*=\{v^*_1, \dots, v^*_k\}$.

Step 6. Pick the next combination $v^\wedge_i$ from $V^\wedge$ and, on its basis, construct the feature subspace.

Step 7. In this feature subspace, build class patterns and compare their mutual placement.

Step 8. If class patterns do not intersect then the combination of features $v^\wedge_i$ is included in the set V^*, and from $V^\wedge$ we exclude all combinations that contain $v^\wedge_i$.

Step 9. The process is repeated until $V^\wedge$ becomes empty.

As a result of the described algorithm, the set $V^*= =\{v^*_1, \dots, v^*_t\}$ where $0 \le t \le q$, will be constructed.

Based on the combinations $v^*_i \in V^*$, we formulate the previously hidden and empirically revealed pattern: «in the feature space of the subset v^*_i the classes do not intersect».

It should be noted that within a specific applied problem the revealed combinations of features v^*_i can be interpreted. This, in turn, means the possibility to interpret all the revealed patterns.

The combinations of features $v^*_i \in V^*$ define decision spaces where class patterns do not intersect. For class patterns inside such spaces, the condition of the compactness hypothesis is hold. Therefore, the construction of classification algorithms does not cause any difficulties.

IV. Results of the Numerical Experiment

Let's show the results of solving the recognition problem on the model data example.

Example. Let the given:

– two classes of objects: **M5 (multiples of 5)** and **NM5 (not multiples of 5)**;

– a priori dictionary of features F = {units, tens, hundreds, thousands, tens of thousands, hundreds of thousands, millions};

– a training set consisting of 20000 seven-bit integers, including 10000 multiples and 10000 not multiples of 5.

Table I shows the research results of class patterns intersection based on the features *units* and *tens*, where $NM5_i$ = Number of $M5_i$ for the i-th digit; $NNM5_i$ = Number of $NM5_i$ for the i-th digit;

TABLE I. Results for features units, tens

Digit	Units		Tens	
	M5	***NM5***	***M5***	***NM5***
0	5043	0	1004	1043
1	0	1253	1003	1049
2	0	1252	966	939
3	0	1238	1020	942
4	0	1235	1034	1010
5	4957	0	1095	992
6	0	1218	933	1018
7	0	1226	992	991
8	0	1266	984	1003
9	0	1312	969	1013

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

$$a_i = \begin{cases} NM5_i + NNM5_i, NM5_i = 0 \lor NNM5_i = 0 \\ 0, NM5_i > 0 \land NNM5_i > 0 \end{cases}$$

$$Intersection = \frac{20000 - \sum_{i=0}^{9} a_i}{20000} * 100\%$$

Table II shows the study results of class patterns intersection based on the features *hundreds* and *thousands*.

TABLE II. RESULTS FOR FEATURES HUNDREDS, THOUSANDS

Digit	Hundreds		Thousands	
	M5	NM5	M5	NM5
0	988	989	1064	1005
1	987	986	974	1020
2	999	1006	1048	968
3	1034	989	956	991
4	980	972	994	979
5	1043	1045	978	1002
6	935	996	1006	999
7	994	980	1042	1036
8	1020	981	966	1025
9	1020	1056	972	975

Table III shows the study results of class patterns intersection based on the features *tens of thousands*, *hundreds of thousands* and *millions*.

TABLE III. RESULTS FOR FEATURES TENS OF THOUSANDS, HUNDREDS OF THOUSANDS, MILLIONS

Digit	Tens of thousands		Hundreds of thousands		Millions	
	M5	*NM5*	*M5*	*NM5*	*M5*	*NM5*
0	1042	1070	998	1020	1072	1012
1	1013	978	975	1018	1014	1026
2	970	990	997	994	948	1029
3	1009	972	957	987	950	1003
4	1026	961	1009	1027	999	994
5	1008	1028	1005	943	1039	989
6	990	973	942	992	979	978
7	988	1017	1045	964	961	980
8	953	1009	1025	1032	1022	1001
9	1001	1002	1047	1023	1016	988

Table IV shows the analysis results for all features from the a priori dictionary.

TABLE IV. EXPERIMENT RESULT FOR ALL FEATURES

Feature name	Intersection (%)
units	0.0
tens	100.0
hundreds	100.0
thousands	100.0
tens of thousands	100.0
hundreds of thousands	100.0
millions	100.0

Tables 1 and 4 show that:

- multiples of 5 have no digits 1, 2, 3, 4, 6, 7, 8, 9 at the units' place; not multiples of 5 – have no digits 0, 5;
- the feature *units*provides an absolute separation of classes M5 and NM5, since *Intersection*=0.0%.

In conclusion, let's note that the algorithm running time is 0.35 seconds.

V. CONCLUSION

An alternative option of the recognition problem statement and solution are considered in the paper.

In the present case, in the learning process we determine combinations of features of a priori dictionary that ensure classes distinction. Then, on the base of revealed feature combinations a classification algorithm is constructed.

The results of solving the recognition problem are shown on the model data example.

REFERENCES

[1] N. G. Zagoruiko, Applied methods of data analysis and knowledge. Izdatel'stvo Instituta matematiki SO RAN. Novosibirsk, 1999, pp. 268 (in Russian).

[2] P. A. Flach, Machine learning. The art and science of algorithms that make sense of data. Cambridge University Press, 2012, pp. 396.

[3] V. I. Vasilev, The problem of training for pattern recognition. Vishashkola. Golovnoe izdatel'stvo. Kiev, 1989, pp. 64 (in Russian).

[4] V. V. Krasnoproshin and V. G. Rodchenko, "Learning by Precedents Based on the Analysis of the Features Properties", Doklady BGUIR, №6 (2017), p. 35–41 (in Russian).

[5] (2021, June) Pattern recognition – Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Pattern_recognition.

[6] C. M. Bishop, Pattern Recognition and Machine Learning. Springer-Verlag, New York, 2006, pp. 738.

[7] Yu. I. Zhuravlev, "On an algebraic approach to solving problems of recognition and classification", Problems of Cybernetics, № 33 (1978), pp. 5–68.

[8] (2021, June) Supervised learning – Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Supervised_learning.

[9] V. V. Krasnoproshin, V. G. Rodchenko, "Classification Based on Decision Spaces", Doklady BGUIR, №6 (2019), pp. 20–25 (in Russian).

[10] M. M. Bongard, The problem of recognition. Nauka, Moscow, 1967, pp. 320 (in Russian).

[11] V. V. Krasnoproshin, V. G. Rodchenko, "Cluster Structures and Their Applications in Data Mining", Informatics, № 2, 2016, pp. 71–77 (in Russian).

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.