Supervised Learning: from the allocation algorithm to the identification of hidden interpretable patterns

Vadim G. Rodchenko
Yanka Kupala State University of Grodno
Grodno, Republic of Belarus
rovar@grsu.by

***Abstract*—In Supervised Learning, traditionally, training is aimed at constructing an algorithm that in the future should carry out the correct classification of objects. The initial training sample is formed on the basis of classes alphabet and an a priori features dictionary, and then, in the process of training, separating surface classes are constructed. A practically useful classifier is constructed in result, but nothing can be learned about the properties of classes. An alternative approach to Supervised Learning is proposed, which is aimed at studying of the of classes properties at revealing the hidden interpreted patterns.**

***Keywords*—machine learning, data mining, classification, supervised learning, instance-based learning, training dataset**

I. Introduction

The development and implementation of new methods and algorithms of machine learning in order to increase the efficiency of processing the accumulated data arrays is one of the most important tasks of computer science [1].

Today, the progress in practical use of artificial intelligence technologies is completely depends on an increase in the effectiveness of Machine Learning (ML) methods and tools [2]. The most important and determining stage in ML is training, which is implemented on the basis of the data training samples and is aimed at identifying empirical patterns [3].

As part of a traditional approach, the result of Supervised Learning is the classification algorithm, which is actually a practically useful "black box", but which can hardly be interpreted.

The article proposes an alternative approach to Supervised Learning, which is based on study of the class properties and identification of the informative features combinations that provide class distinction. The results of the practical application of the approach based on a real dataset are described.

II. Implementing Machine Learning Based on Supervised Learning

Formally, the Machine Learning process based on Supervised Learning can be represented as the following chain of transformations:

$$S \xrightarrow{F_1} C \xrightarrow{F_2} A \xrightarrow{F_3} T \xrightarrow{F_4} I \xrightarrow{F_5} P \xrightarrow{F_6} R$$

where S is the alphabet of classes; C is a set of observed characteristics; A – a priori dictionary of features; T – training samples; I – a refined dictionary of informative features combinations to construct the decision spaces; P – classifier built on the basis of class patterns; R is the set of solutions; F1 – algorithm for obtaining the observed characteristics; F2 – algorithm to construct an a priori features dictionary; F3 – training samples generation algorithm; F4 – an algorithm for identifying the informative attributes combinations of an a priori dictionary to construct the decision spaces; F5 – an algorithm for constructing a classifier based on class patterns in decision spaces; F6 – P-based classification.

The alphabet of classes S, the dictionary C of observable characteristics, and the a priori dictionary of attributes A are developed by experts. The training samples T is formed on the basis of the states of objects observations taking into account S, C and A as a result of algorithms F1 and F2 execution.

Based on the analysis of the training data samples, a study of features combinations from the a priori dictionary is made and a refined dictionary I is formed. Combinations of features that turned out to be uninformative in terms of the class patterns separation are not included in the refined dictionary I.

We exclude from the training set the lines corresponding to the characteristics that did not fall into the updated dictionary I. Using the new training data set, we first construct the class patterns in the form of cluster structures and then use them to build a classifier.

The decision to assign a recognized object to a certain class is based on the study of the object belonging to the class pattern.

III. An Alternative Approach to Supervised Learning

In Machine Learning, based on Supervised Learning, actually two tasks are solved sequentially: 1) training and 2) classification. As a result of solving the first problem, a classifier constructed, which then is used to solve the second problem.

The results of existing methods and tools in the area of Supervised Learning analysis suggest two approaches to the implementation of the learning process.

At present, such an approach in which training is reduced to constructing decision rules that provide an extremum for a certain criterion is traditionally accepted and is universally used. The class of decision rules is a priori specified up to parameters, and training involves finding such parameter values that provide an extremum for the selected criterion.

On use of a priori dictionary of features applied to construct the training set attention does not payed. It is believed that the space to describe objects is given, and it is only necessary to build a separating surface in this space within the criterion framework.

An alternative approach to learning is based on the idea of finding such features combinations from the a priori dictionary that are most informative from the point of distinguishing classes. The identification of these features combinations occurs as a result of the feature spaces construction in which class patterns do not intersect [4]. After this, the procedure for constructing a classifier becomes trivial.

In the framework of the traditional approach to Supervised Learning, it is believed that a space for describing objects is given and it is necessary to build a separating surface in this space.

The practical application of the traditional approach has effectively solved a large number of different applied problems.

Especially exciting results were obtained in the area where the artificial neural network technologies are used. However, although neural network technologies provide for virtually automatic training, the problem of interpreting the revealed patterns has not yet been solved. Then, it turns out that the useful result obtained from resource-intensive process of preparing the training samples (up to 80% of all costs) is only the classification algorithm, which is actually a “black box”.

Note that at present, in machine learning, the focus is done on all methods to construct the classification algorithms is the “Achilles heel”, since they allow us to separate class patterns, but do not find out anything about the properties of classes.

In order to study and identify the properties of classes, an alternative approach to learning can be proposed, which is based on the assumption of the compactness hypothesis that compact sets in the attribute space correspond to class patterns.

Obviously, in the a priori dictionary there may be features that are not informative from the separating class patterns point of view, and then objects of one class are either placed non-compactly in the attribute space or are scattered among objects of another class.

Based on the analysis of data from the training samples, it is proposed to identify such combinations of features that provide separation of class patterns. In this case, the training will actually be aimed at identifying feature spaces in which the compactness hypothesis is confirmed [5].

Note that, firstly, found features combinations can be interpreted within the framework of the subject area, and, secondly, they can be used to construct the classification algorithms [6].

IV. Supervised Learning Based on Analysis of Class Pattern Properties

The classical formulation of the Supervised Learning problem assumes that there are many descriptions of objects X and many acceptable answers for their classification Y. There is an unknown target dependence $y^*: X \to Y$, values $X^m=\{(x_1,y_1),\dots,(x_m,y_m)\}$ which are known only for objects of the training set. It is necessary to construct an algorithm $a: X \to Y$ that would approximate this target dependence, not only on the objects of the finite sample, but also on the whole set X.

To solve the problem, a certain class of algorithms is preliminarily specified up to parameters, and training is reduced to finding the values of the parameters providing an extremum for the selected criterion.

When solving the tusk, a number of problem points arise:

1) The choice of the model $A = \{a : X \to Y\}$ is a non-trivial task and requires the participation of a qualified specialist, which ultimately allows to implement only an automated, but not automatic learning mode.
2) The class-separating surface is constructed on the basis of the data of the training sample X^m, and the question of the information content of the used features from the a priori dictionary remains open.
3) The constructed algorithm $a : X \to Y$ approximates the unknown target dependence, but is actually a “black box” that cannot be interpreted.

If an alternative approach to Supervised Learning is used, one can avoid these disadvantages. The statement of the learning problem in this case is as follows: let there be a lot of descriptions of objects X and a set of valid answers for their classification Y. There is an unknown target dependence $y^*: X \to Y$, values $X^m = \{(x_1, y_1), \dots, (x_m, y_m)\}$ which is known only for the objects of the training set. Then it is required to find the feature subspaces in which class patterns do not intersect.

Suppose that training samples $X^m = \{(x_1, y_1),\dots, (x_m, y_m)\}$ is formed on the basis of the dictionary features $F = \{f_1,\dots, f_n\}$. Denote by $V = \{v_1,\dots, v_q\}$ the set of all possible combinations of features from F. Then V contains q=$\sum_{i=1}^{n} C_n^i$=2^n-1 subsets.

The search algorithm of feature subspaces in which class patterns do not intersect on data set V = {v_1,..., v_q} consists of the following nine steps:

STEP 1. Choose from V a subset of V^+={v^+_1,...,v^+_n}, where v^+_i contains only one attribute.

STEP 2. For each v^+_i build the class patterns and evaluate their relative placement.

STEP 3. Include v^+_i in result set V^*={v^*_1,...,v^*_k} if the class patterns do not intersect.

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

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

STEP 6. Take the next combination $v^\wedge_i$ from $V^\wedge$ and based on it construct a feature subspace.

STEP 7. In this feature subspace, we build class patterns and evaluate their relative positioning.

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

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

As a result of the analysis of all elements V={v_1,..., v_q}, the set V^*={v^*_1,...,v^*_t} will be constructed, where $0 \le t \le q$. Based on the combinations $v^*_i \in V^*$, we formulate previously hidden and empirically revealed patterns: "in the feature space of the subset v^*_i the classes do not intersect".

Since the combinations of features v^*_i can be interpreted within the framework of a specific applied problem, then all the revealed patterns can be interpreted.

The combination of features $v^*_i \in V^*$ defines the solutions spaces in which class patterns do not intersect. For class patterns inside such spaces, the compactness hypothesis is confirmed, and therefore the construction of classification algorithms is straightforward.

V. Results of dataset analysis

We demonstrate the results of applying the proposed algorithm to detect patterns based on cluster structures. We have used the Mushroom dataset, which is hosted in the UCI Machine Learning Repository [7].

The Mushroom dataset contains data on 8124 instances, 3916 of which belong to the Poisoned class, and 4208 to the Eatable class. 22 attributes were used to describe instances: *odor*, *gill-color*, *ring-type*, *stalk-color-below-ring*, *stalk-color-above-ring*, *spore-print-color*, *ring-number*, *veil-color*, *cap-surface*, *cap-shape*, *cap-color*, *gill-attachment*, *gill-spacing*, *stalk-shape*, *bruises*, *stalk-root*, *gill-size*, *veil-type*, *stalk-surface-above-ring*, *stalk-surface-below-ring*, *population*, *habitat*.

Table 1 shows the results of a study of the intersection of class patterns based on a single attribute. The table shows that the odor attribute provides a good separation of the Poisoned and Eatable classes.

Table I
Results of the first numerical experiment

Attribute	Intersection (%)
odor	3.06
gill-color	55.26
ring-type	65.99
stalk-color-below-ring	87.44
stalk-color-above-ring	87.85
spore-print-color	98.16
ring-number	99.08
veil-color	99.80
cap-surface	99.90
cap-shape	99.90
cap-color	100
gill-attachment	100
gill-spacing	100
stalk-shape	100
bruises	100
stalk-root	100
gill-size	100
veil-type	100
stalk-surface-above-ring	100
stalk-surface-below-ring	100
population	100
habitat	100

Table 2 presents the results of two attributes combinations analysis. Table shows data on the 10 most suitable for separating the **Poisoned** and **Eatable** classes from the 230 possible combinations.

Table II
Results of the second numerical experiment

Combination attributes	Intersection (%)
odor, spore-print-color	1.23
odor, habitat	2.15
odor, cap-color	2.45
odor, gill-color	2.45
odor, stalk-color-below-ring	2.45
odor, stalk-root	2.86
odor, stalk-color-above-ring	2.86
odor, veil-color	2.86
odor, cap-surface	2.96
odor, cap-shape	2.96

It turns out that patterns of the **Poisoned** and **Eatable** classes stopped to intersect only in feature spaces formed by combinations of 4 features.

Table 3 shows the results of a numerical experiment based on combinations of four attributes. Only 13 features from 7315 possible combinations were identified in which the intersection of the **Poisoned** and **Eatable** classes patterns have not intersected.

Table III
RESULTS OF THE THIRD NUMERICAL EXPERIMENT

Combination attributes	Intersection (%)
odor, habitat, population, stalk-color-below-ring	0.0
odor, habitat, population,spore-print-color	0.0
bruises, gill-spacing, spore-print-color, stalk-root	0.0
bruises, cap-color, spore-print-color, stalk-root	0.0
bruises, gill-size, spore-print-color, stalk-root	0.0
bruises, ring-number, spore-print-color, stalk-root	0.0
bruises, population, spore-print-color, stalk-root	0.0
odor, bruises, habitat, stalk-surface-above-ring	0.0
odor, habitat, ring-type, stalk-root	0.0
odor, population, spore-print-color, stalk-root	0.0
bruises, habitat, spore-print-color, stalk-root	0.0
odor, cap-color, habitat, stalk-root	0.0
odor, habitat, stalk-color-below-ring, stalk-root	0.0

CONCLUSION

This paper presents an alternative approach for executing the Supervised Learning procedure, which is based on an analysis of the attributes properties of an a priori dictionary.

The purpose of training is to search for feature subspaces in which class patterns do not intersect. Class patterns are represented as cluster structures.

The learning algorithm allows to automatically analyze the training data samples and identify the most informative features in terms of class separation.

The results of the learning algorithm application to solve the classification problem are presented. The Mushroom dataset from the UCI Machine Learning Repository has been used.

ACKNOWLEDGMENT

The author thanks Boris Assanovich for his assistance in preparing the paper final version.

REFERENCES

[1] J. T. Tou and R. C. Gonzales, “Printsipy raspoznavaniya obrazov [Pattern recognition priceples],” Moscow: Mir, 1978, 412 p.

[2] P. Flach, “Mashinnoe obuchenie. Nauka i iskusstvo postroeniya algoritmov, kotorye izvlekayut znaniya iz dannykh [Machine learning. Science and art of constructing algorithms that extract knowledge from data],” Moscow: DMK Press, 2015, 400 p.

[3] N. G. Zagoruiko, “Prikladnye metody analiza dannykh i znanii [Applied methods of data analysis and knowledge],” Novosibirsk: IMSO RAN, 1999, 270 p.

[4] V. V. Krasnoproshin and V. G. Rodchenko, “[Cluster structures and their applications in Data Mining],” Informatika [Informatics], 2016, vol. 50, no 2, pp. 43-57.

[5] MachineLearning. Available at: http://www.machinelearning.ru/wiki/index.php?title (accessed 2019, Dec).

[6] Data mining. Available at: https://en.wikipedia.org/wiki/Data_mining (accessed 2019, Dec).

[7] UCI Machine Learning Repository. Mashroom Data Set. Available at: https://archive.ics.uci.edu/ml/datasets/mushroom (accessed 2019, Dec).

Обучение с учителем: от построения алгоритма классификации к выявлению скрытых интерпретируемых закономерностей

Родченко В.Г.

Обучение с учителем направлено на построение алгоритма, который в дальнейшем должен осуществлять правильную классификацию объектов. Исходная обучающая выборка формируется на основе алфавита классов и априорного словаря признаков, а затем в процессе обучения строятся разделяющие классы поверхности. В результате получается практически полезный классификатор, но о свойствах классов ничего узнать не удается. Предложен альтернативный подход к обучению с учителем, который направлен на исследование свойств классов и на выявление скрытых интерпретируемых закономерностей. Целью обучения является поиск признаковых подпространств, в которых паттерны классов не пересекаются. Паттерны классов представляются в виде кластерных структур. Алгоритм обучения позволяет автоматически провести анализ данных обучающей выборки и выявить наиболее информативные признаки с точки зрения разделения классов. Представлены результаты применения алгоритма обучения для решения задачи классификации.

Received 27.01.2020