BULLETIN OF THE POLISH
ACADEMY OF SCIENCES
TECHNICAL SCIENCES
Vol. 35, No. 5-6, 1987

COMPUTER AND INFORMATION SCIENCES

Learning from Examples – the Case of an Imperfect Teacher

by

Zdzisław PAWLAK

Presented by Z. PAWLAK on September 11, 1986

Summary. In this article we consider a machine learning mechanism with an imperfect teacher. We show that in some cases the teacher's lack of knowledge (ignorance) does not affect the learning process – and cannot be discovered by a "student". To show this we employ the rough set approach.

1. Introduction. In this article we analyse a machine learning process with an imperfect teacher, and we show that in some cases the lack of teacher's knowledge is not relevant and cannot be detected by a "student".

Our approach is based on the concept of a rough set (see [1, 2]) which seems to be a very suitable mathematical tool to deal with this kind of problems.

2. Rough sets. In this section we introduce, after [1], basic notions concerning rough sets being necessary in this paper. A learning system is a 4-tuple $S = (E, A, V, f)$, where

E – is a finite set of examples,

$A = B - \{e\}$ – is a finite of attributes; B – is the set of "student" attributes and e is a teacher (expert) attribute,

$V = \bigcup_{a \in A} V_a$, V_a – is a finite set of values of attribute $a \in A$ $f: E \times A \to V$ – is a training function (total) such that $f(x, a) \in V_a$ for every $a \in A$, and $x \in E$.

If $C \subseteq A$, then by $\tilde{C}$ we mean a binary relation defined thus

$$\tilde{C} = \{(x, y) \in E^2 : f(x, a) = f(y, a) \quad \text{for every } a \in C\}.$$

Obviously $\tilde{C}$ is an equivalence relation for any $C \subseteq A$. The family of all equivalence classes will be denoted by C^*. Any union of equivalence classes of $\tilde{C}$ is called C-discernible set in S; sets which are not C-discernible are called C-indiscernible in S, or rough sets with respect to C in S.

The C-lower approximation of $X \subseteq E$ is defined as:

$$\underline{C}X = \{x \in E: [x]_{\tilde{C}} \subseteq X\},$$

and C-upper approximation of $X \subseteq E$ is defined thus:

$$\bar{C}X = \{x \in E: [x]_{\tilde{C}} \cap X \neq \emptyset\},$$

where $[x]_{\tilde{C}}$ denotes an equivalence class of the relation $\tilde{C}$, containing the element x.

The set

$$BN_C(X) = \bar{C}X - \underline{C}X,$$

will be called a C-boundary of X in S.

The number

$$\alpha_C(X) = \frac{\text{card } C\underline{X}}{\text{card } C\bar{X}},$$

is referred as an accuracy of the approximation of X by C in S.

Obviously $X \subseteq E$ is C-discernible ($C \subseteq A$) iff $\underline{C}X = \bar{C}X$; otherwise set X is rough with respect to C in S. We shall distinguish the following classes of rough sets:

a) if $\underline{C}X \neq \emptyset$ and $\bar{C}X \neq E$, set X will be called roughly C-discernible,
b) if $\underline{C}X \neq \emptyset$ and $\bar{C}X = E$, set X will be called externally C-indiscernible,
c) if $\underline{C}X = \emptyset$ and $\bar{C}X \neq E$, set X will be called internally C-indiscernible,
d) if $\underline{C}X = \emptyset$ and $\bar{C}X = E$, set X will be called totally C-indiscernible.

Let us observe that if X is discernible (roughly discernible or totally indiscernible) so is $-X$; if X is externally indiscernible (internally indiscernible), then $-X$ is internally indiscernible (externally indiscernible) – for any $C \subseteq A$.

3. Learning from examples. The application of rough sets to learning from examples can be found in [2]. Here we define only some notions needed in the rest of the paper.

Any subset $X \subseteq E$ will be called a concept (in S). Elements of X are called examples of X and elements of $-X$ are called counterexamples of X.

We assume that the classification of examples of E into X and $-X$ is provided by a teacher (expert, enviornment etc.) and a "student", who is able to observe the values of attributes from B, is supposed to describe (i.e. to learn) the expert classification. From the previous considerations it follows that the concept X (and consequently $-X$) can be learned if $\underline{B}X = \bar{B}X$; otherwise, i.e. if $\underline{B}X \neq \bar{B}X$, the concept X (and $-X$) cannot be learned using the set of attributes B. This is to mean that only B-discernible sets can be learn, employing set of attributes B, and set rough with respect to B cannot be learned. Using, however, the classification of rough set we can say something more about sets which cannot be learned in the learning system S, namely:

a) if X is roughly B-discernible, the concept X and $-X$ can be learned with certain approximation (expressed by the accuracy coefficient $\alpha_B(X)$, i.e. the student can learn only some examples and counterexamples of the concept X – but not the whole concept X,

b) if X is externally B-indiscernible, the student is able to learn only some examples of the concept X, but is unable to learn any counter-examples of X,

c) if X is internally B-indiscernible, the student is unable to learn any examples of X but is able to learn some counterexamples of X,

d) if X is totally B-indiscernible, the student is unable to learn neither examples nor counterexamples of X.

Thus the case (d) is the worst one, because the student is unable to learn anything from the teacher; in cases (a), (b) and (c) he is able to learn some example and/or counterexample of the concept considered.

4. The case of an imperfect teacher. In what follows we assume that the teacher is imperfect, i.e. he is unable to classify all examples from E, that is to mean he does not know whether some examples belong to X of $-X$.

The problem arises how the teacher's lack of knowledge, i.e. ignorance in classifying examples, affects learning of concept X by the student.

This question can be very easily answered, employing notions introduced in the previous section of this paper.

Let $V_e = \{+, 0, -\}$ and

$$X_+ = \{x \in E : f_x(e) = +\},$$
$$X_0 = \{x \in E : f_x(e) = 0\},$$
$$X_- = \{x \in E : f_x(e) = -\},$$

where

X_+ – is the set of teacher's examples of X $(X_+ \subseteq X)$,

X_- – is the set of teacher's counterexamples of X $(X_- \subseteq -X)$,

X_0 – is the set of examples which cannot be classified by the teacher (he says: I do not know).

The set $X^* = X_+ \cup X_-$ will be called the teacher's knowledge, and the set X_0 we will call the teacher's ignorance. The teacher's knowledge and ignorance can be measured by the coefficient $\alpha_B(X^*)$ and $\alpha_B(X_0)$, respectively; $\alpha_B(X_0)$ will be referred as an ignorance coefficient.

Our problem, whether the student can detect the teacher's ignorance X_0 employing the set of attributes B, is reduced now to the question whether

X_0 is B-discernible or not. From the consideration in the previous section it follows that if X_0 is B-discernible the teacher's ignorance can be fully detected by the student (the ignorance coefficient $\alpha_B(X_0) = 1$); if X_0 is B-indiscernible (rough with respect to B) we have the following four possibilities:

a) X_0 is roughly B-discernible; the teacher's ignorance can be recognized only partially $(0 < \alpha_B(X_0) < 1)$,

b) X_0 is internally B-indiscernible; the teacher's ignorance cannot be recognized $(\alpha_B(X_0) = 0)$,

c) X_0 is externally B-indiscernible; the teacher's ignorance can be recognized only partially $(0 < \alpha_B(X_0) < 1)$,

d) X_0 is totally B-indiscernible; neither the teacher's ignorance nor knowledge can be recognized by the student $(\alpha_B(X_0) = 0)$.

Thus, in cases (a) and (c) the teacher's ignorance can be discovered by the student and in cases (b) and (d) cannot. The case (d) is not interesting from our point of view because neither the ignorance can be discovered nor the knowledge can be learned. Thus let us discuss the case (b) in a little more detail.

Let X_0 be internally B-indiscernible and suppose that $\underline{B}X_+ \neq \emptyset$ and $\underline{B}X_- \neq 0$. It is easy to check that in this case $X_0 \subseteq BN_B(X^*)$, that is to say that the teacher's ignorance is "hidden" in the boundary region of the teacher's knowledge. In other words, it does not matter whether the teacher is unable to classify some examples or not because the student can learn in both cases exactly the same concepts $\underline{B}X_+$ and $\underline{B}X_-$. Thus the theacher's ignorance does not influence the learning ability of the student, in this case.

5. Examples

Example 1. Let us consider the learning system shown in Table 1.

E	a	b	c
1	0	2	+
2	0	1	+
3	1	0	+
4	0	1	0
5	1	0	0
6	1	1	−
7	2	1	−
8	0	1	−
9	1	0	−

Table 1

where $\{a, b\} = B$ are student attributes and c is teacher attribute.

In this system we have:

$$X_0 = \{4, 5\}$$
$$X_+ = \{1, 2, 3\}$$
$$X_- = \{6, 7, 8, 9\}$$

and

$$\underline{B}X_+ = \{1\} \qquad \bar{B}X_+ = \{1, 2, 3, 4, 5, 8, 9\}$$
$$\underline{B}X_0 = \emptyset \qquad \bar{B}X_0 = \{2, 3, 4, 5, 8, 9\}$$
$$\underline{B}X_- = \{6, 7\} \qquad \bar{B}X_- = \{2, 3, 4, 5, 6, 7, 8, 9\}$$
$$BN_B(X^*) = \{2, 3, 4, 5, 8, 9\}.$$

Obviously $X_0 \subset BN_B(X^*)$, and X_0 is internally B-indiscernible. Thus the student, using the set of attributes $B = \{a, b\}$ is unable to discover that the teacher failed in classifying examples 4 and 5. The ignorance coefficient $\alpha_B(4, 5) = 0$.

Example 2. In the learning system as shown in Table 2

E	a	b	c
1	0	2	+
2	0	1	+
3	1	0	+
4	0	1	+
5	1	0	0
6	1	1	0
7	2	1	–
8	0	1	–
9	1	0	–

Table 2

attributes a and b are also student attributes and c is the teacher attribute. In this table we have

$$X_0 = \{5, 6\}$$
$$X_+ = \{1, 2, 3, 4\}$$
$$X_- = \{7, 8, 9\}$$

and

$$\underline{B}X_+ = \{1\} \qquad \bar{B}X_+ = \{1, 2, 3, 4, 5, 8, 9\}$$
$$\underline{B}X_0 = \{6\} \qquad \bar{B}X_0 = \{3, 5, 6, 9\}$$
$$\underline{B}X_- = \{7\} \qquad \bar{B}X_- = \{2, 3, 4, 5, 8, 9\}$$
$$BN_B(X^*) = \{2, 3, 4, 5, 8, 9\} \text{ and } X_0 \not\subseteq BN_B(X^*).$$

In this example the student can discover that the teacher failed in classifying example 6 (but not 5), so he is able discover partially the teacher's ignorance. The ignorance coefficient in this case is $\alpha_B(5, 6) = 0.25$.

6. Conclusion. In general case the theacher is supposed to classify examples into n classes, where $n > 2$. In this case the teacher's ignorance is manifested by pointing out some classes (instead of one class) to which an example may belong. This case will be discussed in detail elsewhere.

DEPARTMENT OF COMPLEX CONTROL SYSTEMS, POLISH ACADEMY OF SCIENCES, BAŁTYCKA 10, 44-100 GLIWICE
(ZAKŁAD SYSTEMÓW AUTOMATYKI KOMPLEKSOWEJ PAN)
DEPARTMENT OF COMPUTER SCIENCE, NORTH CAROLINA UNIVERSITY, NC 28223, USA

REFERENCES

[1] Z. Pawlak, *Rough sets*, Int. J. Inform. Comp. Sci., **2** (1982), 341–356.
[2] Z. Pawlak, *On learning – a rough set approach*, in: Lecture Notes in Computer Sciences, No. 208 ed. A. Skowron, Springer-Verlag, Berlin, (1986), 197–227.

З. Павляк, **Обучение по примерам — случай несовершенного учителя**

В настоящей статье рассматривается механизм компьютерного обучения с несовершенным учителем. Доказывается, что в некоторых случаях пробелы в знаниях учителя не воздействуют на процесс обучения и не могут быть обнаруженными „студентом".