machine learning, knowledge acquisition,
medical diagnosis support systems

Michał WOŹNIAK*

PROPOSITION OF SUBJECTIVE QUALITY MEASURE FOR EXERT RULES

Paper deals with the knowledge acquisition process. Different experts formulate the rules for decision support systems. We assume they have different knowledge about the problem and therefore obtained rules have different qualities. We will formulate the proposition of the confidence measure and its application to the decision process. We will propose how calculate the value of measure under consideration for typical statistical learning process. On the base on the proposed measure of the knowledge quality we propose the procedure of the contradictions elimination for the set of logical rules.

1. INTRODUCTION

Machine learning is the attractive approach for building decision support systems[6]. For this type of software, the quality of the knowledge base plays the key-role. In many cases we can meet those problems was partly described for the induction learning [2,3] and for the concept description [1]. The following paper concerns on subjective quality measure of rule.

The content of the work is as follow: Section 2 defines statistical confidence measure of the knowledge. In this section we also present the interpretation of proposed measure for the estimation process based on the typical statistical model. Section 3 proposes the procedure of contradictions elimination for the set of rules obtained from different sources. The last section concluded the paper.

2. PROPOSITION OF KNOWLEDGE CONFIDENCE MEASURE

2.1. DEFINITION

Let consider the first order rule (conclusion of rule points at the final, separate class number) and each of them is obtained under the following assumption:
– learning set is noise free (or expert tell us always true),
– target concept contained in the set of class number $\{1, ..., M\}$.

We consider decision under the first assumption given by the following formulae:

* Chair of Systems and Computer Networks, Faculty of Electronics, Wroclaw University of Technology. Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland, e-mail: Michal.Wozniak@pwr.wroc.pl

$$P(\text{If A then B}) = 1 .$$

In the expert system designing process the rules are obtained from different sources and the sources have the different confidence. For the knowledge given by experts we can not assume that expert tell us true or/and if the rule set is generated by the machine learning algorithms we can not assume the learning set is noise free.

Therefore we postulate that we have not to trust all information we got or believing on it only with the γ factor, proposed as the confidence (quality) measure. It can be formulated as [8]

$$P(\text{If A then B }) = \gamma \leq 1 .$$

Let $\gamma_i^{(k)}$ denotes the value of the confidence measure of rule $r_i^{(k)}$ (k-th rule points at the class *i*).

2.2. CONFIDENCE MEASURE FOR THE STATISTICAL ESTIMATION

The central problem of our proposition is how to calculate the confidence measure. For human experts the values for their rules is fixed arbitrary according to the quality of creator. The presented problem we can also find in the typical statistical estimation of unknown parameter β , where we assume the significant level[7]. The significant level can be interpreted as the confidence measure. Each rule gives the index of the class. If the feature vector value belongs to the decision area given by the rule, the decision depends on the previous state and on the applied therapy. While constructing the artificial rule set, we have to define somehow the decision areas for the new rule set. For example we can want to obtain for each rule *posterior* probability estimator, which is not less than a fixed value or in the practice we can use the one of very well known machine learning algorithms like AQ,CN2 [6].

For each of the given intervals *k* we have to obtain the estimator of the *posterior* probability. We use the following statistical model [7]:

– the learning set is selected randomly from a population and there exists two class of points: marked (point at the class $i \in \{1, ..., M\}$) and unmarked (point at the class *l*, where $l \in \{1, ..., M\}$ and $l \neq i$),
– the expected value for the population is *p*,
– the best estimator of *p* is

$$\hat{p} = \frac{m}{n} , \tag{1}$$

where *n* means the sample size and *m* - the number of the marked elements.
Let us concentrate on two cases
Small sample ($n \leq 100$)

$$t = \frac{\hat{p} - p}{s/\sqrt{n}} \tag{2}$$

has the Student's t-distribution. We want to estimate one parameter - the expected value, therefore we use t-distribution with n-1 degree of freedom.

For the fixed significance level α we get

$$P(|t| < \mu_\alpha) = 1 - \alpha \tag{3}$$

using the short-cut formula

$$s^2 = \frac{\sum (x_i - \hat{p})^2}{(n-1)} \approx \sum x_i^{\,2} + \frac{\left(\sum x_i\right)^2}{n} \tag{4}$$

we obtain

$$P\left(\frac{m}{n} - \mu_\alpha \sqrt{\frac{\frac{m}{n}\left(1 - \frac{m}{n}\right)}{n}} < p < \frac{m}{n} + \mu_\alpha \sqrt{\frac{\frac{m}{n}\left(1 - \frac{m}{n}\right)}{n}}\right) \leq 1 - \alpha \tag{5}$$

The μ_α is the value of the t-distribution on n-1 degrees of freedom and for the significance level α. In this case we get rule $r_i^{(k)}$, for which confidence measure of rule $\gamma_i^{(k)}$ is given by the following equation

$$\gamma_i^{(k)} = 1 - \alpha , \tag{6}$$

$$\underline{\beta}_i^{(k)} = \frac{m}{n} - \mu_\alpha \sqrt{\frac{\frac{m}{n}\left(1 - \frac{m}{n}\right)}{n}} \quad and \quad \overline{\beta}_i^{(k)} = \frac{m}{n} + \mu_\alpha \sqrt{\frac{\frac{m}{n}\left(1 - \frac{m}{n}\right)}{n}} \tag{7}$$

Big sample (n> 100)

For the big sample, the distribution is similar to the normal distribution. We thus have the equation of the same for as (15), but in this case μ_α is the value of normal standardized N(0, 1) distribution for the significance level α.

3. CONTRADICTIONS ELIMINATION ALGORITHM

As we have mentioned the proposed method of the quality management can be applied to the logical knowledge representation (where "if-then" means logical implication). E.g. for the unordered set of first-order logical rules acquisition process we can attribute the value of confidence to each of rule. It could be used in the case if the contradiction in the set of rules would be detected.

First we note the set of rule R consists of the M subsets

$$R = R_1 \cup R_2 \cup ... \cup R_M \ , \qquad (8)$$

where R_i denote subset of rule pointed at the *i*-th class.

For this form of rule the two of them contradict each other if

$$\exists\, x \in X \wedge \exists\, k, l \in \{1, 2, ..., M\} \wedge \exists\, i, j \qquad x \in D_i^{(k)} \wedge x \in D_j^{(l)} \qquad (9)$$

Where *i, j* denote the number of rule.

The equation (9) means that we can find observation, which belongs to the decision area of the rule pointed at class *i* ($R_i^{(k)}$) and decision area of the rule pointed at different class *j* ($R_j^{(l)}$) –for the first order rules set this two rules contradict each other.

Let us propose the idea of the contradictions elimination algorithm.

```
//for each class number
for i:= 1 to M:
//for each rule in R_i
    for k:=1 to |R_i|
        //for each class number bigger than i
        for j:= i to M:
            //for each rule in R_j
            for l:= 1 to |R_j| :
                //if r_i^(k) and r_j^(l) contradict each other
                if D_i^(k) ∩ D_j^(l) ≠ ∅
                    then
                        //if confidence of r_i^(k) is higher then
                        //confidence of r_j^(l)
                        if γ_i^(k) ≥ γ_j^(l)
                            then
                                //remove the common part from decision
                                //area of rule r_j^(l)
                                D_j^(l) := D_j^(l) \ (D_j^(l) ∩ D_i^(k))
                            else
                                //remove the common part from decision
                                //area of rule r_j^(l)
                                D_i^(k) := D_i^(k) \ (D_j^(l) ∩ D_i^(k))
                            fi
                fi
            endfor
        endfor
    endfor
endfor
```

4. CONCLUSION

The paper concerned the proposition of the quality measure for that formulated decision problems. We proposed the idea of the contradiction detection and elimination method for logical representation of experts' knowledge (first order rule), but it could be used for any other form of rule (we have to define contradiction conditions for this type of knowledge representation). Presented methods need the analytical and simulation researches, but we hope this idea of confidence management can be helpful for other problems whose can be met during the knowledge acquisition process from different sources (for example to modify `PERFORMANCE` function in the algorithms based on *sequential covering* procedure). Now we are working on the modification of C4.5 and AQ algorithms and their applications to the real medical decision problem (the diagnosis of acute abdominal pain). The results will be presented on the MIT conference and in the final version of this paper.

BIBLIOGRAPHY

[1] BERGADANO F., MATWIN S., MICHALSKI R.S., Zhang J., Measuring of Quality Concept Descriptions, *Proc. Of the 3rd European Working Session on Learning*, Aberdeen, Scotland, 1988.

[2] BRUHA I., KOCKOVA S., Quality of decision rules: Empirical and statistical approaches, *Informatica*, no 17.

[3] DEAN P., FAMILI A., Comparative Performance of Rule Quality Measures in an Inductive Systems, *Applied Intelligence*, no 7, 1997.

[4] GIAKOUMAKIS E., PAPAKONSTANTIOU G., SKORDALAKIS E., Rule-based systems and pattern recognition, *Pattern Recognition Letters*, No 5, 1987.

[5] GUR-ALI O., WALLANCE W.A., Induction of rules subject to a quality constraint: probabilistic inductive learning, *IEEE Transaction on Knowledge and Data Engeineering*, vol. 5, no 3, 1993.

[6] MITCHELL T., *Machine Learning*, McGraw Hill, 1997.

[7] SACHS L., *Applied Statistic. A Handbook of Techniques*, Springer-Verlag, New York Berlin Heideberg Tokyo, 1984.

[8] WOZNIAK M., Idea of Knowledge Acquisition for the Probabilistic Expert Systems, *Proceedings of the International Conference on Computational Intelligence for Modeling, Control and Automation*, Vienna (Austria), 12-14 February 2003.