Cost-Sensitive Classification for Medical Diagnosis

Gerald Schaefer
Aston University, UK

Tomoharu Nakashima
Osaka Prefecture University, Japan

Yasuyuki Yokota
Osaka Prefecture University, Japan

INTRODUCTION

Medical diagnosis is often regarded as a pattern classification problem: based on a certain input, the task is to assign it to one of a set of classes. Often the number of classes is two: for example, malignant vs. benign, disease vs. no disease. The task of a pattern classification system is to assign as many input samples as possible to the correct class whereas the behaviour of the classifier is often optimised through the learning of some ground truth data. Conventional classifiers treat each sample of this learning set equally, yet in medical diagnosis this is often not desirable as different classes are associated with different costs. While the misdiagnosis of a malignant case as being benign can be very costly (e.g., when the time for effective treatment has passed), mistaking a benign case as malignant (though of course it should be avoided) will involve relatively lower costs (such as some further tests).

In this article, we present a cost-sensitive approach to medical diagnosis based on fuzzy rule-based classification (Schaefer, Nakashima, Yokota, & Ishibuchi, 2007). While fuzzy rule-based systems have been mainly employed for control problems (Lee, 1990) more recently they have also been applied to pattern classification problems (Ishibuchi & Nakashima, 1999; Nozaki, Ishibuchi, & Tanaka, 1996). We modify a fuzzy rule-based classifier to incorporate the concept of weight which can be considered as the cost of an input pattern being misclassified. The pattern classification problem is thus reformulated as a cost minimisation problem. Based on experimental results on the Wisconsin breast cancer dataset, we demonstrate the efficacy of our approach. We also show that the application of a learning algorithm can further improve the classification performance of our classifier.

FUZZY RULE-BASED CLASSIFICATION

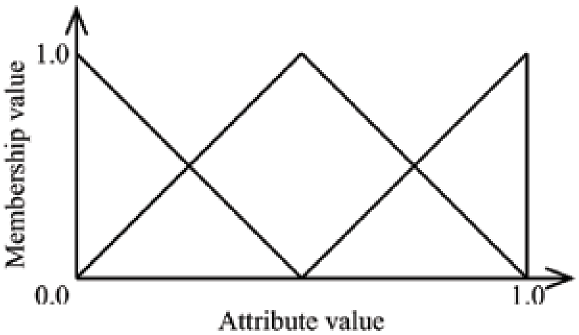
Let us assume that our pattern classification problem is an n-dimensional problem with M classes (in medical diagnosis M is typically 2) and m given training patterns $\mathbf{x}_p = (x_{p1}, x_{p2},...,x_{pn})$, $p = 1,2,...,m$. Without loss of generality, we assume each attribute of the given training patterns to be normalised into the unit interval $[0,1]$; that is, the pattern space is an n-dimensional unit hypercube $[0,1]^n$. In this study we use fuzzy if-then rules of the following type as a base of our fuzzy rule-based classification systems:

$$\begin{aligned}&\text{Rule } R_j:\ \text{If } x_1 \text{ is } A_{j1} \text{ and } \dots \text{ and } x_n \text{ is } A_{jn}\\&\text{then Class } C_j \text{ with } CF_j \qquad j = 1,2,\dots,N\end{aligned} \qquad (1)$$

where R_j is the label of the j-th fuzzy if-then rule, $A_{j1},...,A_{jn}$ are antecedent fuzzy sets on the unit interval $[0,1]$, C_j is the consequent class (i.e., one of the M given classes), and CF_j is the grade of certainty of the fuzzy if-then rule R_j. As antecedent fuzzy sets we use triangular fuzzy sets as in Figure 1 where we show a partition of the unit interval into a number of fuzzy sets.

Our fuzzy classification system consists of N fuzzy if-then rules each of which has a form as in Equation (1). There are two steps in the generation of the rules: specification of antecedent part, and determination of consequent class C_j and grade of certainty CF_j. Once the antecedent part of a rule is specified the consequent part (i.e., consequent class and the grade of certainty) is determined from a set of given training patterns (Ishibuchi, Nozaki, & Tanaka, 1992). In Ishibuchi and Nakashima (2001), it is shown that the use of the grade of certainty in fuzzy if-then rules allows us to generate comprehensible fuzzy rule-based classification systems with high classification performance.

Figure 1. Triangular membership function



Fuzzy Rule Generation

Let us assume that m training patterns $\mathbf{x}_p = (x_{p1},...,x_{pn})$, $p = 1,...,m$ are given for an n-dimensional C-class pattern classification problem. The consequent class C_j and the grade of certainty CF_j of the if-then rule are determined in the following manner:

1. Calculate $\beta_{Class\ h}(j)$ for Class h as:

$$\beta_{Classh\ h}(j) = \sum_{\mathbf{x}_p \in Class\ h} \mu_j(\mathbf{x}_p) \tag{2}$$

where

$$\mu_j(\mathbf{x}_p) = \mu_{j1}(x_{p1}) \cdot ... \mu_{jn}(x_{pn}) \tag{3}$$

and $\mu_{jn}(\cdot)$ is the membership function of the fuzzy set A_{jn}. In this article, we use triangular fuzzy sets as in Figure 1.

2. Find Class $\hat{h}$ that has the maximum value of $\beta_{Class\ h}(j)$:

$$\beta_{Class\ \hat{h}}(j) = \max_{1 \le k \le C}\{\beta_{Class\ k}(j)\} \tag{4}$$

If two or more classes take the maximum value, the consequent class C_j of the rule R_j cannot be determined uniquely. In this case, we specify C_j as $C_j = \{\}$. If a single class $\hat{h}$ takes the maximum value, let C_j be Class $\hat{h}$.

The grade of certainty CF_j is determined as:

$$CF_j = \frac{\beta_{Class\ \hat{h}}(j) - \bar{\beta}}{\sum_h \beta_{Class\ h}(j)} \tag{5}$$

with

$$\bar{\beta} = \frac{\sum_{h \neq \hat{h}} \beta_{Classh}(j)}{c-1} \tag{6}$$

Fuzzy Reasoning

Using the rule generation procedure outlined above we can generate N fuzzy if-then rules as in Equation (1). After both the consequent class C_j and the grade of certainty CF_j are determined for all N rules, a new pattern $\mathbf{x} = (x_1,...x_n)$ can be classified by the following procedure:

1. Calculate $\alpha_{Class\ h}(\mathbf{x})$ for Class h, j=1,...,C, as:

$$\alpha_{Class\ h}(\mathbf{x}) = \max\{\mu_j(\mathbf{x}) \times CF_j \mid C_j = h\} \tag{7}$$

2. Find Class h' that has the maximum value of $\alpha_{Class\ h}(\mathbf{x})$:

$$\alpha_{Class\ h'}(\mathbf{x}) = \max_{1 \le k \le C}\{\alpha_{Class\ k}(\mathbf{x})\} \tag{8}$$

If two or more classes take the maximum value, then the classification of **x** is rejected (i.e., **x** is left as an unclassifiable pattern), otherwise we assign **x** to Class h'.

COST-SENSITIVE FUZZY CLASSIFICATION

The standard fuzzy rule-based classifier as detailed above treats each class and hence each sample equally. In medical diagnosis however this is often not desirable. Misdiagnosing a malignant case as benign should be penalised more than diagnosing healthy patients as having a certain disease. While in the first case, the result might be that of late treatment in the best and missing of the treatable time in the worst scenario, the latter case will usually involve some further tests which should then identify the misdiagnosis. Clearly, the costs involved in the first case will exceed those of the latter.

Related Content

REDAlert+: Medical/Fire Emergency and Warning System using Android Devices
Arjun Shakdherand Kavita Pandey (2017). *International Journal of E-Health and Medical Communications (pp. 37-51).*
www.irma-international.org/article/redalert/171161

ICT Use and Multidisciplinary Healthcare Teams in the Age of e-Health
Bolanle A. Olaniran (2016). *International Journal of Reliable and Quality E-Healthcare (pp. 18-31).*
www.irma-international.org/article/ict-use-and-multidisciplinary-healthcare-teams-in-the-age-of-e-health/145760

A Prototype Agent Based Model and Machine Learning Hybrid System for Healthcare Decision Support
Marek Laskowski (2013). *Digital Advances in Medicine, E-Health, and Communication Technologies (pp. 230-253).*
www.irma-international.org/chapter/prototype-agent-based-model-machine/72980

ICT Supplier Evaluation Criteria in Large University Hospitals: An Empirical Examination in Turkey
Canan Gamze Baland Serkan Ada (2013). *Handbook of Research on ICTs and Management Systems for Improving Efficiency in Healthcare and Social Care (pp. 1208-1223).*
www.irma-international.org/chapter/ict-supplier-evaluation-criteria-large/78077

An Exploration of Demographic Inconsistencies in Healthcare Information Environments
Larbi Esmahi, James W. Edwardsand Elarbi Badidi (2008). *Encyclopedia of Healthcare Information Systems (pp. 561-572).*
www.irma-international.org/chapter/exploration-demographic-inconsistencies-healthcare-information/12985