A method of classifier design for fuzzy data

Witold Pedrycz
Department of Automatic Control
and Computer Sciences
Silesian Technical University
Gliwice,Poland

We discuss classifier construction based on fuzzy data that are available while dealing with object classification described in linguistic fashion.Fuzzy relation equations will be studied as a flexible framework for classifier design.An approach toward feature extraction being of a great importance in pattern recognition techniques is shown as well.

1. Introductory remarks

In practice of pattern recognition we are faced with objects described in linguistic fashion that are usually classified by any human being.Due to well-known properties of a process of human cognition the object classified may be characterized by specification names of linguistic labels attached to it.Also the classification realized is not crisp but a fuzzy one.

A general problem setting is the following.Let be $X=\{X_1,X_2,\ldots,X_n\}$a collection of objects classified.X_i denotes a feature vector of the i-th object.Further on we assume X_i is represented by a fuzzy relation defined in cartesian product of coordinates of the feature space $\mathcal{X}$,viz.

$$X_i:\underbrace{\mathcal{X}_1\times\mathcal{X}_2\times\ldots\times\mathcal{X}_N}_{X}\rightarrow[0,1] \qquad /1/$$

Each element of X is assigned to exactly one class/category/ among a fixed list of classes $\Omega=\{\omega_1,\omega_2,\ldots,\omega_c\}$.Nevertheless,we may relax this requirement discussing the case the object is classified in a fuzzy manner,viz.more than one class is involved with the object.The task of the classifier design is to construct a mapping"f",

$$f:\ X\in\mathcal{X}\longrightarrow\Omega\in\boldsymbol{\Omega} \qquad /2/$$

that has some"plausible" properties.Especially"f" should minimize a classification error/error of misrecognition/.Here we shall speak about a fuzzy relation R defined in $\mathcal{X}\times\Omega$.

2. Fuzzy classifier design

Being equipped with the above preliminary statements,let us formulate the method of construction of the classifier. Remember we have at our disposal a list of data/training set of patterns/,

$$\{X_i, \Omega_i\}, i=1,2,\ldots,n, \qquad /3/$$

where X_i is the fuzzy relation standing for the feature representation of the i-th object. Ω_i is a fuzzy set of classes $\Omega_i=[\Omega_i(\omega_1)\Omega_i(\omega_2)\ldots\Omega_i(\omega_c)]$ The j-th coordinate of the membership function of Ω_i, viz $\Omega_i(\omega_j)\in[0,1]$ expresses a strength of belongingness of the i-th object to the j-th class; and as usually $\Omega_i(\omega_j)=1$ indicates highest degree of membership, while $\Omega_i(\omega_j)=0$ excludes it. Searching a form of the fuzzy classifier, notice that without any additional information we can speak the objects represented by their feature vector are r e l a t e d with the classes. Therefore one may pay attention to fuzzy relation equations as appropriate apparatus for discovering ties between features and classes. Generally we put down,

$$X \, R \, \Omega \qquad /4/$$

with $X: \mathfrak{X}\to[0,1]$, $\Omega:\mathfrak{N}\to[0,1]$ $R:\mathfrak{X}\times\mathfrak{N}\to[0,1]$. For a broad discussion on fuzzy relation equations the reader is referred e.g. to[4][5][7]. Here we focus our attention on the following form of the equation[3]

$$\Omega = X \circledphi R \qquad /5/$$

viz.

$$\Omega(\omega) = \inf_{x\in\mathfrak{X}} [X(x)\varphi R(x,\omega)] \qquad , \qquad /6/$$

"φ" stands for pseudocomplement associated with t-norm, see[4][6],

$$a\varphi b = \sup\{ c\in[0,1] \mid atc \le b\} \qquad /6/$$

It could be seen "φ" is also defined as two-placed function $\varphi:[0,1]\times[0,1]\to[0,1]$ such that $a\varphi\max(b,c)\geq\max(a\varphi b, a\varphi c)$, $a\varphi(a\varphi b)\leq b$, $a\varphi(atb)\geq b$, $a,b,c\in[0,1]$. Some results reported in[3] or[5] that will be of significant interest may be summarized in the following propositions. Denote by $\mathfrak{R}$, $\mathfrak{X}$ families of fuzzy relations,

$$\mathfrak{R}=\{ R : \mathfrak{X}\times\mathfrak{N}\to[0,1] \mid X\circledphi R=\Omega\}, \quad \mathfrak{X}=\{ X : \mathfrak{X}\to[0,1] \mid X\circledphi R = \Omega\} \qquad /7/$$

Proposition 1. If $\mathfrak{R}\neq\emptyset$ then $\hat{R}= X \circledt \Omega$, $\hat{R}(x,\omega)=X(x)\,t\,\Omega(\omega)$ is the least element of $\mathfrak{R}$, $\hat{R}=\min\mathfrak{R}$.

Proposition 2. If $\mathfrak{X}\neq\emptyset$ then $\check{X}=\Omega\circledphi R$, $\check{X}(x) = \inf_{\omega\in\mathfrak{N}}[\Omega(\omega)\varphi R(x,\omega)]$ is the greatest element of $\mathfrak{X}$, $\check{X}=\max\mathfrak{X}$.

Proposition 3. If $\bigcap_{i=1}^{n}\mathfrak{R}_i\neq\emptyset$ then $\hat{R}=\bigcup_{i=1}^{n} X_i\circledt\Omega_i$, $\hat{R}(x,\omega)=\max_i[X_i(x)\,t\,\Omega_i(\omega)]$ is the least element of $\bigcap_{i=1}^{n}\mathfrak{R}_i$, where $\mathfrak{R}_i=\{ R_i: \mathfrak{X}\times\Omega\to[0,1] \mid X_i\circledphi R=\Omega_i\}$..

Bearing in mind the result of the Proposition 3, the fuzzy relation $\hat{R}$ of the classifier is computed as union of t-composition of X_i and Ω_i for $i=1,2,\ldots,n$.

3. On evaluation of discrimination power of features

One of central problems of pattern recognition is concentrated on feature extraction [1] that leads to dimension diminishing of the classification problem. Consider the classifier equation /5/ with the learning set /3/. The discrimination power of the j-th feature /viz. the j-th coordinate of the fuzzy relation X/ is measured by its influence on the vector Ω resulting from $\circledast$ composition of X and R. Let the j-th feature is "unknown" modelled by a fuzzy set with unit membership function $X_j(x_j)=0$ for all $x_j \in \mathcal{X}_j$, where the fuzzy relation X is equal to,

$$X(x_1,x_2,\dots,x_{j-1},x_j,x_{j+1},\dots,x_N)=X_1(x_1)\,t\,X_2(x_2)\,t\dots t\,X_{j-1}(x_{j-1})\,t\,X_j(x_j)\,t\,X_{j+1}(x_{j+1})\,t\dots t\,X_N(x_N).$$

The j-th feature, j=1,2,...,N, is evaluated with regard to a sum of distances resulting by treatment the j-th coordinate of the fuzzy relation as "unknown". For concise notation put down, X_i^j as the fuzzy relation X_i modified as stated above, $\Omega_i^j = X_i^j \circledast \hat{R}$ /$\hat{R}$ calculated with the aid of Proposition 3/ The sum of distances "d" between fuzzy sets Ω_i^j and Ω_i, i=1,2,...,n

$$\frac{1}{c}\sum_{i=1}^{n} d(\Omega_i^j, \Omega_i) \qquad /8/$$

creates an index reflecting discrimination power of the feature. The higher the value of /8/ the more significant the j-th feature is. We shall call /8/ a classification error with respect to the j-th feature.

4. Illustrative example

A numerical example makes use of data contained in [2] and concerns the analysis of characteristics of the working situations in a lumber factory. The features describing each worker /pattern/ are contained in 14-dimensional vector, $X=[x_1,x_2,\dots,x_{14}]$, (e.g. x_1-professional training, x_{11}-supervision of tasks). Each entry of X lies in [0,1]-interval and its value reflects degree of satisfaction of the property listed. The patterns have been classified into four nonfuzzy categories of workers /for instance the workers having relatively high academic standards and intellectual activity/. The training set consists of 18 patterns. The classification error calculated as,

$$D= 1/4 \sum_{k=1}^{18} d(\Omega_k, \hat{\Omega}_k), \quad \hat{\Omega}_k = X_k \circledast \hat{R} \qquad /9/$$

where "d" stands for the Hamming distance, is equal to 2.21.
Determining discriminating power of the features we have realized the most significant are: physical effort, supervision of tasks, incidence on

the production process,and adverse environmental conditions.

5.R e f e r e n c e s

1.R.O.Duda,P.E.Hart,Pattern Classification and Scene Analysis,J.Wiley, N.York,1973.

2.R.Lopez de Mantaras,J.Aguilar Martin,Classification and linguistic characterization of non-deterministic data,Pattern Recognition Letters, 2,1983,33-41.

3.M.Miyakoshi,M.Shimbo,Solutions of composite fuzzy relational equations with triangular norms,Fuzzy Sets and Systems/to appear/.

4.W.Pedrycz,Fuzzy relational equations with triangular norms,BUSEFAL, 11,1982,24-32.

5.W.Pedrycz,Application of fuzzy relational equations for methods of reasoning in presence of fuzzy data,Fuzzy Sets and Systems/to appear/.

6.S.Weber,A general concept of fuzzy connectives,negations and implications based on t-norms,and t-conorms,Fuzzy Sets and Systems,11,1983, 115-134.